Savings Behavior

of

Private Households

in

Germany

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

Lothar D. Essig

vorgelegt im Wintersemester 2004/2005

Referent:	Prof. Axel Börsch-Supan, Ph.D.
Korreferent:	Prof. Dr. Ulrich Schlieper
Dekan:	Prof. Konrad Stahl, Ph.D.
Prodekan:	Prof. Dr. Paul Gans
Tag der mündlichen Prüfung:	22. Dezember 2004

Preface

This study is based on data which were collected under the authority of the *Mannheim Research Institute of the Economics of Aging (MEA)*. It was an extremely heavy financial task to make the *SAVE* survey possible. Therefore, I am deeply indebted to the *Deutsche Forschungsgemeinschaft (DFG)* who provided us with the necessary financial means to run this panel and with a leap of faith in our work.

While preparing this thesis I was a doctoral student at the *Gradudiertenkolleg Allokation auf Finanz- und Gütermärkten* and a research fellow at *MEA* at the University of Mannheim. I am deeply indebted to my advisors, Professors Axel Börsch-Supan and Ulrich Schlieper at Mannheim and Joachim Winter at Munich, for their encouragement and for many helpful discussions and comments. I also benefited a lot from the great support - through encouragement and advice - from my colleagues at the *Graduiertenkolleg* and the *MEA*.

The lively seminars at the *MEA* have proved to be a great opportunity to structure own ideas as well as receive many comments pointing the way ahead.

Many other people have backed me during the work of this thesis. Naming some of them would simultaneously neglect others. Therefore, my special thanks go to all my colleagues at *MEA* not only for their academic skills, but also for having been very dear friends making the time for this thesis not only an appealingly intellectual, but also a very funny one. All the common activities like hiking, canoing, and especially skiing made my stay at Mannheim an unforgettable time.

I am deeply indebted to my parents, providing me with great support during these many years. I am also grateful for the patience and encouragement I received from Tanja.

Lothar Essig

Contents

Chapter 1

Introduction

The purpose of this dissertation is to contribute to the understanding of households' saving behavior. Savings and savings behavior was subject to large and comprehensive research, see Deaton (1992), Browning and Lusardi (1996) and Attanasio (1999), and still, the complexity of households' savings behavior is not completely understood. This is astonishing since the allocation of available income on spending and saving is one of the most important economic decisions made by a household. The intertemporal aspect of saving is fundamental for the understanding of how a household plans for the long term. Saving behavior encompasses not only the sober economic thinking of perfectly informed planners but also (often only seemingly) unstructured reactions deeply rooted in human psychology and socio-cultural norms. Actual behavior may deviate (e.g. Thaler and Shefrin (1981), Laibson (1997), O'Donoghue and Rabin (1999)) from the models economists are used to work with (e.g. Kotlikoff (1989); Hurd (1990); Jappelli and Modigliani (1998)). To understand saving, it therefore helps to be open for economic as well as psychological and sociological explanations.

Germany, in particular, is an interesting country to study saving, especially among older households. Even though the pension and health insurance system in Germany is one of the most generous systems in the world, private savings are high until old age. Börsch-Supan *et al.* (2001) refer to that phenomenon as the "German savings puzzle".

Until recently, there has been a lack of data that records detailed information on savings in Germany in conjunction with sociological and psychological characteristics. The German Socio-Economic Panel (*GSOEP*) only records rough indicators such as "Did you spend all of your income last year or was there anything left over?" and "Do you have a savings book?", etc. but does not cover the quantitative composition and any changes in the amount of wealth. The situation is similar for the 'Soll und Haben' (Debit and Credit) survey. While this contains very detailed data on the composition of various forms of investment it does not quantify these in greater detail.

The income and expenditure survey (*EVS*) conducted every five years by the Federal Statistical Office with its detailed information on the amount and composition of income, expenditure and wealth is the main source of data on the savings behavior of households in Germany. Unfortunately, several variables that are important for savings behavior are now missing. More importantly, extensive sociological or psychological factors are completely absent in the income and consumption surveys, because these very expensive surveys are primarily intended for the work of the Federal Statistical Office.

Weaknesses of existing data material can only be rectified by new surveys. It is important to record variables which can also describe psychologically determined behavioral phenomena for a better

understanding of actual savings behavior. This insight in the need for new data material was the motivation to create an additional data base for the analysis of savings behavior.

In 2001, the first wave of the *SAVE* survey was conducted. It was initiated and is administered by members of the Mannheim Research Institute for the Economics of Aging (MEA). Considering the experiences of researchers and external experts with other surveys, a questionnaire was designed approaching the subject of savings from different angles, taking into consideration economical and sociological as well as psychological perspectives. The main financial support for this survey stems form the German Science Foundation (Deutsche Forschungsgemeinschaft (DFG)). So far, data from three years and five different subsamples are available. The next wave will be guided in 2005.

This dissertation is based on the *SAVE* survey. Since *SAVE* has in many respects an experimental nature, methodological aspects are an integral part of *SAVE*. These will be analyzed and discussed in this work before discussing substantial questions on savings and savings behavior.

Structure of this work

I summarize the seven chapters concisely at this point. Every chapter, though, has a more thorough separate introduction highlighting the research interest and methodology.

Chapter 2 presents a basic methodological review of the *SAVE* dataset. It explains the general construction of the questionnaire and the sampling schemes, and it shows descriptive findings of similarities and differences between the four different *SAVE* subsamples available until spring 2004. Furthermore, the representativeness of the data is discussed. It also investigates whether weights constructions can help to eliminate potential differences if they were significant between subsamples and between representative population values.

Chapter 3 is joint work with Joachim Winter. We analyze nonresponse to questions on financial items such as income and asset holdings in household surveys using data from a controlled field experiment. As part of *SAVE*, questions on household income and financial assets were administered using different modes (personal interview vs. drop-off questionnaire). The data also allow to investigate the influence of interviewer characteristics on nonresponse. Our results are in line with predictions derived from models of survey response behavior that have been developed in survey research and social psychology.

In Chapter 4, I discuss first panel results for saving behavior, the effect of changes between 2001 and 2003 for pension information level and pension expectations, and present results for risk variables and their effect on financial behavior.

The literature on precautionary saving gives very contrary results for the importance and size of precautionary saving. The *SAVE* data offer the possibility to generate some of the frequently used instruments for the precautionary motive known from the literature. Chapter 5 compares the influence of these instruments on long-run and short-run saving measures. Additionally, the *SAVE* questionnaire contains information on saving motives. I compare the explanatory power of these saving motives to other subjective instruments.

General purpose surveys typically refrain from using an exhaustive list of consumption items since the trade-off between gaining more precise data on consumption and losing a lot of both, interview

time and respondent effort, generally forbids this procedure. An alternative is to ask respondents a non-exhaustive list of sub-items and use those to impute total expenditures by the use of an external data source. Beginning with the 2003 wave, the *SAVE* questionnaire was enriched by a short section of expenditure questions in a way enabling this imputation method. This is done in Chapter 6.

The results from Chapter 6 are preparatory for the discussion of savings measures in Chapter 7. In *SAVE*, savings are measured using a one-shot question for total annual savings. This recall question might cause significant problems concerning the precise measure for actual savings. Unlike expenditures, savings is a far more complicated concept which most respondents might not be fully aware of. This chapter shows potential flaws of this kind of questioning as well as potential remedies to squeeze the most of reliable measures given the information at hand.

The demographic change presents major financing problems for the pay-as-you-go pension system. For this reason, the 2001 pension reform entailed a reduction in the level of statutory pensions and created a significantly strengthened framework for the funded second and third pillars of old-age pension provision. The population at large is only dimly aware of the gap in provision created by the 2001 pension reform and the introduction in 2003 of the sustainability factor. Chapter 8, which is joint work with Axel Börsch-Supan, examines the extent to which households are in a position to close this gap with their personal assets without changing their savings and asset accumulation behavior. Four critical factors are relevant to this issue. 1. Anticipated life expectancy: In the 2001 and 2003 *SAVE* samples, only qualitative information was surveyed. To calculate the specific subjective life expectancy, the *SAVE* 2004 sample was modified to raise this information. 2. The level of personal assets on retirement, 3. the age of retirement itself, and 4. anticipated interest rates. The results show that, unless they were to change their savings behavior, more than half of households would not be able to bridge the pension gap from their future financial assets. The results for that analysis shed light on two points. First, households, even though having a good idea of the *average* unconditional[1] life expectancy, systematically underestimate their own, even relative to the time-specific life tables which do not include probable medical progress for each cohort, and even more, if compared to actual research results on demographics. Second, one third of the households will not be able, if continuing their financial behavior, to fill the pension gap. The situation becomes even more unfavorable when assuming more realistic individual life expectancies.

[1] I.e., the life expectancy of newborns.

Chapter 2

Methodological aspects of the *SAVE* data set

2.1 Introduction

Savings and savings behavior is still not fully understood. In fact, it is partially even hard to tell what economic agents have in mind when thinking about savings. The introduction of this dissertation listed some of the many contributions to this topic. The recent research development suggests to enlarge the economist's view to other research fields, especially to psychology and sociology.

Additionally, the data situation for the savings analysis is limited in Germany. Weaknesses of existing data material can only be rectified by new surveys. It is important to record variables which can also describe psychologically determined behavioral phenomena for a better understanding of actual savings behavior. Taking as a basis the examples of the Dutch CentER Panels, the US Health and Retirement Surveys, and the Bank of Italy's Survey on Household Income and Wealth (SHIW), the Mannheim Research Institute on the Economics of Aging (MEA) has cooperated with the Mannheim Center for Surveys, Methods and Analyses (ZUMA), NFO Infratest (Munich), and Psychonomics (Cologne) to produce a questionnaire consisting of six sections. The questionnaire has been designed in such a way that the interview should not exceed 45 minutes. It is downloadable in English and German as PDF-file.[2]

Surveys are an important source of data for the empirical analysis of household behavior. Unfortunately, data problems such as unit nonresponse (sample selection), item nonresponse, and measurement error are the rule rather than the exception in survey data. Well-designed studies using household survey data carefully proceed to detect outliers, to impute missing values, and to correct for selection caused by missing observations.

The *SAVE* panel attempts to collect a large set of variables shedding light on many household characteristics. The *SAVE* data were collected in 2001 and 2003. In the year 2001, one of the tense aspects of the survey was to check whether a major survey can be established in Germany which directly asks so called 'hard' financial, and, therefore, most private questions. The 2001 wave consisted of two parts. The first one was a computer-assisted personal interview (CAPI) quota sample which was itself divided into four different interview modes. For an analysis of potential interview mode effects, see Chapter 3. The second part was a paper & pencil (P&P) interview which drew households from a standing German access panel. In 2003, the survey again consisted of two parts. The first one

[2] Visit www.mea.uni-mannheim.de, select the language, go to 'research/Forschung', 'household savings behavior / Sparverhalten der Haushalte', then 'SAVE-Project', where you can find the questionnaire.

assembled the recontacted households from the 2001 CAPI samples, while the second one was a new 'refreshment' sample constructed as a random ('Random Route') sample.

The plan of this chapter is as follows. In Section 2.2 I briefly review the general design of the *SAVE* survey and the sampling differences between the four embedded subsamples. In Section 2.3, I discuss problems and opportunities of the sampling design considering the income question as an example. Section 2.4 discusses the representativeness of the data; probit regressions with nonresponse dummies for income and two key assets as dependent variables show potential subsample differences. Section 2.4 also shows the weights constructions to rectify potential deviations of representative population values. Section 2.5 summarizes the results and discusses implications for the use of the *SAVE* data material in estimation procedures.

2.2 SAVE

This section describes the general design of the *SAVE* survey: the design of the questionnaire, interviewer and interviewee motivation, and sampling differences between the two subsamples conducted in 2001 (Section 2.2.2) and 2003 (Section 2.2.3). Contributions in Gabler *et al.* (1997) discuss different sampling procedures and their experiences for German data.

2.2.1 General design of the *SAVE* survey

The *SAVE* survey seeks to achieve several goals. The most important one is to shed more light on households' saving behavior. This substantive goal can certainly only be accomplished if severe threats to the data validity are excluded or reduced as far as possible.[3] Research perspectives from six different groups are worth to be taken into account when designing surveys and evaluating survey data: statisticians, psychologists, sociologists, anthropologists, political scientists, and economists.

Groves (1989) classifies three major languages of error which are applied to survey data, associated with three different disciplines: (i) statistics (mostly, sampling theory) (ii) psychology (psychometric test and measurement theory) and (iii) economics (mostly, econometrics). The other three disciplines mentioned above employ in Grove's view similar languages to these three. Andersen *et al.* (1979) depict a conceptual structure of error sources in surveys, accumulating in the total mean square error. Variance and bias, the two components of the mean square error criterion, are split up into errors of *nonobservation* and *observational errors*. Errors of nonobservations are due to three sources, *coverage*, *nonresponse* (if not located or refusals), and *sampling error* (depending on the subset of the population). Observational errors can be due to *interviewer errors* (wrong [manipulative or ignorable] guidance through the interview process), *instrument errors* (stemming from the wording of the question, a large field in social psychology; see, e.g., Schwarz (1999)), *respondent errors* (arising from different cognitive abilities or motivation to answer questions), and the *mode of data collection* (different effects of CAPI vs. P&P or CATI interview modes).

[3] Statistical and econometric models, e.g., try to minimize sampling errors, but are generally not tailored for nonsampling errors

In addition to the potential errors leading to errors in survey data, it is possible that errors would be made after receiving answers from the respondent: interviewers could enter wrong values, variables can be wrongfully matched to respondents, skip patterns might be erroneous; in general, other procedures proceeding and following the data collection phase.

Apart from these more or less 'trivial' technical errors, the questionnaire might be designed in a way not suited or incomplete for the topic of interest. For example, if one is interested in studying saving behavior, wealth variables are a necessary list of variables which are even theoretically hard to assess and disentangle.

The Mannheim Research Institute for the Economics of Aging (MEA) has cooperated with the Mannheim Center for Surveys, Methods and Analyses (ZUMA), TNS Infratest (Munich), Psychonomics (Cologne) and members of the Sonderforschungsbereich 504 at the University of Mannheim to design a questionnaire which reduces the extend of instrument and respondent errors. In addition, experiences with other surveys, especially with the *HRS* and the Bank of Italy Survey of Household Income and Wealth (*SHIW*) data sets inspired certain wordings of questions and their associated answering scale.

The task to reduce interviewer errors was undergone by the survey agency, TNS Infratest, by intensive interviewer training and motivation for the subject.[4]

To check the influence of interview modes on nonobservations (unit and item nonresponse) and on respondent errors, the first *SAVE* wave additionally included an experimental component. The CAPI part was divided into four subsamples, differing in interview mode and questionnaire design in the central part, see below. Dillman (2000) discusses extensively issues on questionnaire construction, survey implementation and mixed-mode surveys. Many issues implemented in the *SAVE* design are discussed in that survey.

So far, the arguments for data quality and error minimization neglected a non-trivial component: survey costs.[5] Surveys are very expensive; and some interview modes are much more expensive than others, e.g., CAPI interviews are more expensive than CATI[6] or P&P interviews. Obviously, there are trade-offs between the modes' results; if not, the cheapest interview mode would be the only one available at the market. The question is whether survey results justify the cost differences. Given budget constraints, the first *SAVE* wave included P&P interviews from a standing access panel. This opens the opportunity to check for which variables these much cheaper data work and where they don't.

The questionnaire has been designed in such a way that the interview should not exceed 45 minutes. Table 2.1 provides an overview of the *SAVE* questionnaire.

The survey's sensitive topic requires careful convincing by the interviewer. A letter which was handed to the interviewees explaining the scientific and political concern about the topic was thought to raise the willingness for participation, see also Dillman (2000).

[4] For the Survey on Health, Retirement and Ageing in Europe (*SHARE*), the principals of the survey personally encountered the interviewers in addition to the survey agency's effort to motivate the topic's importance.

[5] Ignoring legal problems (e.g., holding a gun to uncooperative respondents' heads).

[6] Computer-assisted telephone interview.

The brief first section of the questionnaire explains the purpose of the questionnaire and describes the precautions that have been taken with respect to data protection. This introduction was considered appropriate because the survey particularly deals with the personal affairs of those surveyed. The interviewer then asks to speak to a member of the household who knows about household income and assets. If this person is not at home, the interviewer must make a return visit.

Part 2 lasts about 15 minutes and is the standard initial interview in which questions are asked about the composition and socio-economic structure of the household, including age, education and participation in the labor force of the person surveyed and his or her partner.

The interviewer deals with the key issues in Part 3 of the questionnaire. This part contains qualitative and simple quantitative questions on saving behavior and how households deal with income and assets, such as the type of investment selected for one-off injections of cash, the importance of a series of savings motives, whether there is actually anything left over to save, how regularly savings are made, etc. Questions are also asked about decision processes and possible rules of thumb[7], past patterns of behavior as well as their parents and attitude to money.

Part 4 is the critical part of the questionnaire because this is where a complete "financial review" is made of the household. A detailed survey is made of income according to the types of income, changes in income, the level of assets according to the various kinds of wealth and changes in the types of wealth over the last year. Apart from financial assets, the questions also cover private and company pensions, ownership of property and business assets. Questions are also asked about debt. Part 4 is kept separate from the other parts, see Section 2.2.2.

Part 5 contains questions about psychological and social factors. It includes the social environment, expectations about income, the economic situation, health, life expectancy and general attitudes to life.

Part 6, the final part, ends the interview with the standard questions about the interview situation and leaves both the person surveyed and the interviewer considerable scope for their own comments. Typically, comments about confidentiality, the length and accuracy of the questionnaire are expected. Questions are also asked about Internet access and the possibility of conducting a further survey.

The survey's topic demands careful convincing by the interviewer and, in order to motivate interviewers, by the principal. We did not reward participants by financial incentives,[8] even though there is a huge amount of literature describing possible advantages of monetary incentives, thereby possibly reducing reducing unit nonresponse. See Brennan *et al.* (1991), Singer (2002), Porst (1996), and Klein and Porst (2000) for surveys of incentives.

[7] See Baumol and Quandt (1964) for a theoretical foundation on the use rules of thumb under uncertainty and Rodepeter and Winter (1999) for the use of rules of thumb in life-cycle savings models

[8] There were mainly two reasons for not paying incentives. The first is that for CAPI interviews, the amount needed to raise interview participation is unclear. The cited literature mainly addresses P&P mail surveys. Second, there were concerns by the survey agency for harming firm policy regulations regarding the treatment of TPI members (by destroying 'market prices').

2.2.2 *SAVE* 2001

The surveys took place in early summer 2001 and 2003. In 2001, the fieldwork for the personal interviews took place between May 29 and June 26, 2001, whereas the fieldwork for the Access Panel took place between June 29 and July 24, 2001.

Experimental design of the *SAVE* 2001 survey

The first four versions were computer aided personal interviews (CAPI); they were carried out by NFO Infratest, Munich. In contrast, the fifth version was a conventional paper questionnaire ("paper and pencil", P&P). The CAPI interviews were carried out using quota samples whereas conventional P&P questionnaires were given to a so-called Access Panel operated by the company TPI (Test Panel Institute, Wetzlar).[9]

The only difference in the four versions of the CAPI interview is in the critical part 4 of the questionnaire. In versions 1 and 2, all questions were administered by CAPI in the presence of the interviewer. The difference between these versions is that the questions on asset holdings were presented using an open-ended format with follow-up brackets (range cards) in version 1 and with 'forced' brackets in version 2.[10]

Because many of these questions relate to intensely personal matters of income and wealth, there is another modification in versions 3 and 4. In these two versions, part 4 was not part of the personal CAPI interview, but left as a paper-and-pencil questionnaire by the interviewer (this mode is termed "P&P drop off" in the sequel). In version 3, the interviewer came back personally to collect the drop-off questionnaire; in version 4, the questionnaire had to be returned by mail using a pre-paid envelope. If this was not done within a specified number of days, the respondent was reminded by telephone several times. This helped increase response rates for the drop-off questionnaire, but nevertheless, they were significantly lower in version 4 than in version 3 (90.5% vs. 98.0%).

Both the CAPI (quota sample) and the P&P (TPI Access Panel) segments were targeted at households with head of the household aged between 18 and 69 years. For the CAPI versions, the quota performance targets were related to the dimension gender (male respondent ratio of 75 percent) and age (a distribution in age classes under 25, 25-34, 35-50 and 50-70 years) according to the current official population statistics (and, in particular, the 2000 micro census).

For the TPI interviewees, the quota targets were also based on the 2000 micro census and either related to the dimensions gender (male respondent ratio of 75 percent) and age (a distribution in age classes 18-29: 13%; 30-39: 24%; 40-49: 22%; 50-59: 21%; 60-69: 20%), and, additionally, whether the respondent is a wage earner or a salaried employee, and the size of the household.

Table 2.2 shows the sample sizes for the five survey versions. In total, 1,829 households were surveyed.

[9] In other words, a standing panel of households surveyed at regular intervals.

[10] This experimental manipulation of question format is not investigated in the present chapter; this is part of Chapter 3.

Quota sampled surveys are heavily debated concerning their representativeness and arising statistical problems. King (1983) lists four principal sources of bias possibly induced by quota sampling: *Differences in respondent availability, insufficient control strata, interviewer selection bias and incorrect information on stratum sizes.* Even though these arguments are well known and taken into account, there are still arguments in favor of quota sampling. A survey of this is kind is new to Germany, and caution with regard to the survey design therefore was a driving force. In a quota sample, interviewers try to contact easily reachable persons which typically are acquainted households. The presumption was that unit and item nonresponse would be significantly lower than in random samples. Or, talking economics, we were seeking output maximization under given budget constraints.[11]

2.2.3 *SAVE* 2003

The *SAVE* 2003 wave consisted of two major samples. The first one consisted of the households which already participated in the *SAVE* 2001 CAPI sample. The second one was a newly added "refreshment"[12] random sample. Interview modes for the two subsamples were identical. They were CAPI interviews except for part 4 (drop-off with mail-back / collection by the interviewer).

Panel CAPI sample

One of the major interests of the *SAVE* study is to analyze behavioral and financial changes over time. Therefore, we tried to re-contact the interviewees from the 2001 personal interviews (N=1169) again in 2003.

The German data protection act prohibits to keep interviewees' addresses when they denied a future follow-up corporation. This has to be checked at the end of an interview. While there is no precise law article, there exists an agreement between the *ADM* (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., where Infratest is also a member) and the official data protection agency.[13] As a result of the denials in 2001, only 72% (= 840 households) were available as gross sample in 2003. After different stages of losses (moved away/died, refused, no time, not available) and rejecting some incomplete interviews, only 483 completed interviews were available.[14]

The fieldwork for the 2001 CAPI sample in 2003 took place between June 2 and July 18, 2003.

Random Route sample

The most favorable argument for the quota sample in 2001 was the expectedly lower unit and item nonresponse rates. Since item nonresponse rates were in line with comparable surveys in other coun-

[11] As will be shown, item and unit nonresponse rates in the quota samples are below those from the Random Route sample.

[12] The quotation marks indicate that this sample size is actually much larger than the original panel. See section 2.2.3.

[13] The agreement itself is sometimes called "Schweinoch"-agreement since Mr. Schweinoch conducted negotiations on behalf of the official site.

[14] Chapter 3 analyzes both the probability of refusals and the probability of interviewing households another time given that they agreed to in the first place. While in the former case the interview mode in part 4 of the questionnaire played a significant role (see Section 2.2.2), the latter was also influenced by income (pos. influence) and age.

tries, and also descriptive statistics compared to other German data sources, the decision was made that the design of the *SAVE* 2003 refreshment sample was to be a Random Route sample.

Sample design The data universe for the *SAVE* 2003 random sample were all German speaking households in Germany with the households' head being eighteen years and older. Interviewees were selected from a multiply stratified multistage random sample. All communities were segmented into stratifications by regional criteria. Stratification criteria were states (Bundesländer), districts and community types. For further sampling details, see Heien and Kortmann (2003).

Unit response rates Random Route sampling requires more careful planning than quota sampling. In contrast to quota sampling schemes where the interviewer is actually in control of sampling the interviewees as long as they fulfill the quota targets and where no information is available on unit nonresponse, this information is available for the Random Route sampling.[15] The contract with the field agency *Infratest Sozialforschung* aimed at a net sample of 2,200 households. It turned out that a gross sample of 4,772 addresses was needed to get a net sample of 2,184 interviews. The most important reason for losses was, as expected, refusal (directly indicated or indirectly as "no time") which accounted for 36.7% of the losses.

The fieldwork for the Random Route sample began on May 26 and ended on July 14.

2.3 Reported income in the *SAVE* survey

This section explains problems and opportunities which arise in P&P interviews / interview parts (part 4 of the *SAVE* questionniare). In the first part of this section, I will explain how income was asked in the questionnaire, what problems arose, and how they can be dealt with.

2.3.1 Income questions in *SAVE*

Income was asked in a three-step process. Interviewees were first given a list of 20 types of income from which composes monthly household income. Afterwards, an open-ended question for the amount of monthly net household income followed. In case of nonresponse, a brackets list was presented including 14 income classes.[16] The brackets list was asked as a range card. See, e.g., Juster and Smith (1997) or Hurd *et al.* (2003) for more advanced unfolding brackets methods.

2.3.2 Imputation of income values

Table 2.4 shows differences between the different *SAVE* subsamples. An unintended effect of the questionnaire design and interview mode will be used to correct for income outliers. As the fifth line

[15] There is an ongoing discussion about the required minimum unit response rate in surveys. Numbers between 50% and 80% were proposed, see Porst (1996) for a review. The assumption that missing values due to unit nonresponse are missing at random might be misleading. See Little and Rubin (1987).

[16] Income brackets range from <500, 500-1000, 1000-1500, 1500-2000, 2000-2500, 2500-3000, 3000-3500, 3500-4000, 4000-4500, 4500-5000, 5000-7500, 7500-10000, 10000-15000 and ≥15000 €.

in Table 2.4 shows, we observe income values for both the open-ended question as well as for the range-card follow-up brackets question in 1,263 cases. This results from the fact that respondents overlooked the filter instructions to skip the follow-up question in case they answered the open-ended question.[17] Further inquiries at the survey agency support the fact that respondents typically have problems following filtering instructions in a P&P questionnaire, even though these instructions were very clearly pronounced. If respondents fully understand the questions and the values being addressed, responses in the open values and in the brackets question should lie in the same brackets class. For a comparison, Table 2.6 shows the class distances when subtracting actual given classes from class analog values imputed from the open value question. Household income from those data was imputed assuming class means.

Table 2.6 shows that about 90% of both given income values lie in the same or in an adjacent income class (marked as bold). This shows that for an overwhelming majority of responses, income can be believed as a reliable measure.

Answers for brackets questions[18] were used when no answer was given in the previous question. This was done in 881 cases of *SAVE* 2001 and 2003. One is tempted to claim that large class differences in Table 2.6 may be due to a misperception of yearly and monthly income. A different possibility might be that errors are simple input errors when the P&P data were electronically transferred. We had this double-checked by the survey agency.

In a second step, we propose the hypothesis that respondents are less likely to mix up monthly with yearly income because brackets induce a readaptation due to a cognitive process: relatively more lower income brackets are linking obviously not to yearly but to monthly data.[19] The correction procedure uses the following ideas:

1. If both values available: compare brackets values to open values. If open values between 7 and 17 times the brackets means values: divide open values by 12. This leads to 42 changes.

2. Use panel information: when data differ more than by factor five between two years → supposedly yearly income → divide by 12. 13 cases reimputed for the 2003 CAPI sample, 11 for 2001.

These two at least partially hypotheses-driven correction procedures still leave us with 79 observations where the monthly net household income is still at least 10,000 €. Even though one might be tempted to divide these remaining large income values by 12, I refrain from this procedure for two reasons. First, this would completely exclude any 'true' measure of high income, which, even though unlikely, are still possible, even in small samples. Second, this is no hypothesis driven procedure. One might, of course, look at different indicators implicitly excluding such high values. But which to pick is rather vague and a matter of ongoing discussions.

[17] Table 2.4 also shows that this did happen significantly less frequently for the 2001 TPI subsample. This most probably stems from the fact that the TPI respondents have some questionnaire experience.

[18] Class mean values.

[19] See, e.g., Winter (2002a) for an experimental study on bracketing effects in survey questions.

2.4 Representativeness

This section discusses the quality and representativeness of the *SAVE* data. Figure 2.1 shows the number of observations for each subsample, the refusal rate for future interview participation and the actual loss of observations from the CAPI 2001 subsample to 2003. Panel attrition rates will also decrease over time, which can be seen from the drop of the CAPI 2001 refusal rate of 28.1% to 12.0%, since reluctant respondents already disappeared in the second wave.

Household surveys underly two major stages. The first one is the design of the study (random route, quota sample etc.), while the second one is the field work itself (systematic and idiosyncratic observation losses). The inclusion probability of a "target person / household" might or might not be equal to its relative population frequency counterpart. The two mentioned stages might influence and bias this inclusion probability; resulting data might therefore be "weighted" relative to its population frequency. So called "weighting procedures", or, correctly spoken, "unweighting[20] procedures", try to reduce or, in best case eliminate these effects.[21] See also Von der Heyde (1994).

Table 2.13 shows item nonresponse to income, and conditional item nonresponse to savings accounts and stocks for the four different samples. Like the regression results presented in tables 2.14 - 2.16, item nonresponse is depending on the sampling method. See the following sections for a discussion.

2.4.1 Subsample differences: Regression results

This section presents estimation results from probit regressions on income and assets (saving accounts and stocks) with dummies for item nonresponse of each of the three variables as dependent variable and a set of household (and interviewer) characteristics as well as subsample dummies as independent variables as dummies to check whether sampling procedures (access panel, quota, random route) influence response behavior.

Regression results: income

Table 2.14 shows conditional probit estimates for nonresponse for open-ended question of monthly net household income of the *SAVE* subsamples (four / three). For better comparability and, in order to eliminate mode effects, observations for the non-P&P modes for the *SAVE* 2001 CAPI subsample were discarted. The second two columns show estimates with interviewer variables, ignoring the *SAVE* 2001 TPI subsample[22]. The relative influence of the sample dummies remained nearly completely constant. Table 2.14 shows that a change from quota samples to a random sample significantly reduces the willingness to reveal sensitive data (raises nonresponse). Thus, the response rates achieved in 2001 with the quota samples could be attained; this supports the hypothesis from Section 2.2.3 that quota samples promise higher response rates. But another effect is also astonishing. One might wonder whether respondents in a quota sample would react to an interviewer change. This is not supported

[20] Assuming the total population as being unweighted, a sample not being representative due to different sorts of sample selection is then weighted in that sense.

[21] Indeed, the procedure rather tries to correct presumed survey's biases.

[22] Remember, this was a pure P&P sample

by the data. Even in a probit regression keeping only households when observed in 2001 and 2003, a dummy variable for interviewer change is not significant. The problem here is that I ignore[23] the effect that an interviewer change could already have affected unit nonresponse which eliminates the item nonresponse effect. Interviewees of the quota sample typically are more likely to collaborate with an interviewer they know and trust. If there was an interviewer change between 2001 and 2003, they might refuse to participate in the 2003 survey if the known interviewer would be replaced by someone unknown to them. Thus, the interviewer change might well lead to unit nonresponse, and does not translate into different item response behavior.

Regression results: assets

Tables 2.15 and 2.16 show regression results from probit estimates of conditional[24] item nonresponse to financial variables on a set of respondent characteristics, interviewer characteristics and dummies[25] for each subsample.[26] The results show a strong influence of the sampling design on item nonresponse. Interview 'professionals' like the sampled respondents in the TPI sample prove to have the highest response probability. This result is as expected since they actually have agreed to collaborate with the survey agency on a regular basis.[27] Quota sampled respondents in the 2001 CAPI sample have the second highest response probability. On the other hand, one result is puzzling: respondents in the panel sample 2003 seem to be more reluctant to answer to financial questions. Two hypotheses were tested. First, regressions were run to test for the influence that the willingness to further participation influences the answering probability in the 2001 CAPI sample. Second, it was tested if there is a time effect when only including respondents into the regression when observed in both subsamples.[28] Interestingly, neither dummy variable controlling for each of the two effects is significant. The dummy variable for the 2003 random route sample is soundly significant in any specification and has the expected sign: as hypothesized earlier, respondents in a random route sample typically have lower response rates.

This brings back the trade-off between costs and errors. Even if item nonresponse is unsystematic, so that values are missing at random and thus ignorable, a larger net sample is needed to produce the same amount of responses than the quota sample.

[23] Since I cannot control for it.

[24] Conditional on the fact that people claimed to own assets of this type (in tables 2.15 and 2.16) but gave no actual value to the follow-up questions.

[25] The basic sample is the 2001 TPI, 1 stands for the CAPI 2001, 2 for the panel 2003, and 3 for the new random route 'refreshment' sample.

[26] Observations were excluded from the regressions when the interview mode differed in the corresponding part 4, see Table 2.2. Moreover, the last two columns in tables 2.14 - 2.16 only refer to samples 1-4 since the 2001 TPI sample was a full P&P interview with no interviewers involved.

[27] Still, they have the right to refuse the participation in unpleasant interview topics.

[28] While the results are not shown here, they are available by the author on request.

2.4.2 Weights constructions

Tables 2.7 and 2.8 show how representative the *SAVE* sample is in comparison with the German micro-census of 2000 and 2002, respectively. The figures in this table compare the proportion of households in an age and income class with the comparable proportion of the same type of households in the micro-census. A figure of 1.2 means that the micro-census covers 20% more households of this type than are present in our random sample. If we take the micro-census as the benchmark, a figure of less than 1 indicates underrepresented household types, and figures over 1 indicate overrepresented household types. Tables 2.7 and 2.8 were stratified for each subsample and for two variables: income / age and income / household size. The reason for using these two different methods lies in the fact that on the one hand, it is common to use income and age as classical spanning variables, but on the other hand, age itself was used as a quota target variable for the *SAVE* 2001 CAPI subsample. See Gabler *et al*. (1994) for a discussion of weighting criteria. Differentiation by more variables imposes the problem of too small cell sizes. In comparison to the micro-census, the random sample contains considerably more middle-aged households but fewer older households. This applies to both sample groups (CAPI variants and Access Panel). Young households are represented approximately correctly. With regard to income, we can see really pronounced shifts towards richer households. This is particularly pronounced in the Access Panel: here the micro-census indicates four times as many households with a monthly net income of less than DM 2,500 / 1,300 Euros than in our sample group but only half as many households with an income of over DM 5,000 / 2,600 Euros.

Tables 2.7 and 2.8 show the importance of using the variable 'subsample type'. This weighting criterion variable was used implicitly by constructing the weight factors separately for each subsample in each year. While especially the 2003 random sample proves to fit the 2002 German micro-census data extremely well (especially regarding the age / household size part of Table 2.8 where values orbit around 1), we see large deviations in the distribution when comparing the 2001 Access Panel sample to the 2000 micro-census (Table 2.7); the *SAVE* sample contains considerably more middle-aged households but fewer older households. This applies to both sample groups (CAPI variants and Access Panel). Young households are represented approximately correctly. With regard to income, we can see a really pronounced shift towards richer households. This is particularly salient in the Access Panel: here the micro-census indicates four times as many households with a monthly net income of less than DM 2,500 / 1,300 Euros than in our sample group but only half as many households with an income of over DM 5,000 / 2,600 Euros.

While the following paragraph will show the influence of the weights use on the distribution of certain key variables, the weights used in the following chapters of this thesis refer to the dimensions subsample type, age, and income. The reason for not using the the dimension household size instead of age is a continuity reason, since Börsch-Supan and Essig (2002) used these weights in the first examination of the *SAVE* data.

2.4.3 Weighting effects

The results of tabes 2.7 and 2.8 demand a further investigation of the influence of weighting procedures on key variables in the *SAVE* data set. Therefore, income, savings and wealth will be displayed by each subsample with and without the usage of weights. Results are presented in tables 2.9 and 2.10. The use of weights shifts the distributions of all presented measures to the left; theses effects are translations from the results of tables 2.7 and 2.8: weighing variables have the strongest effect when distributions of income and age (or income and household size) deviate the most from the German microcensus.[29] The higher means of income in the *SAVE* 2003 RR sample are due to remaining high outliers: 52 households (or 2.6%)[30] in this subsample claim to earn more than 10,000 € net each month. Not considering values higher than 15,000 € in this subsample reduces the mean net monthly household income to about 2,100 €.

Similar effects are observed for the *GSOEP* 2000 to 2003 (Table 2.11) and the *EVS* 1998 and 1998 (Table 2.12) income measures: in both surveys, richer households seem to be oversampled in comparison to the microcensus (Table 2.11).

2.5 Conclusions

This chapter briefly surveys the objective and structure of the questionnaire and the sampling methods of the 2001 and 2003 *SAVE* study. Unit and item nonresponse, a measure of acceptance of a survey of this kind, are absolutely in line with surveys in other countries.[31] I also show how representative the data are in comparison to the German microcensus and other comparable data sources. It proves to be the case that the *SAVE* data actually show similar effects as, for example, the *GSOEP* data. The sampled persons are slightly richer (or, biased towards middle classed households; the strength of this bias depends on the sampling criteria for each subsamle). Using sample weights tailored individually for each subsample, values are obtained that fit the microcensus population means almost perfectly, exemplified using the income measure. Contributions in Gabler *et al.* (1994) discuss the use of weights for different data sources. While weighting might be a probate method for descriptive analysis, it is unclear whether weights should be used for estimation procedures. There is still ongoing research on this topic; see Wooldridge (2001b) and Wooldridge (2001a) for a discussion of the use of weights.

Clearly, data quality could be enhanced by more sophisticated survey methodology in future waves.[32] This, on the other hand, comes at the cost of inconsistencies across time. In such cases, one is tempted to renounce to improved survey methodology to avoid those inconsistencies and simply freeze survey methodology over time, thereby eliminating any quality enhancement. However, as a long run strategy, this is clearly a bad idea - robust empirical findings cannot be obtained from poor

[29] And this is the reason why the use of weights for the RR 2003 subsample does merely affect the means and medians of the presented variables.

[30] Only about 1.5% in the 2000, 2001 and 2002 *GSOEP*, and about 0.5% in the 1998 *EVS*.

[31] Compare the figures for unit and item nonresponse for eight European surveys in Alkemade *et al.* (2003).

[32] See, for example, Van Soest and Hurd (2004) for a review.

data. Juster *et al.* (2002) develop methods of recovering time series consistency in the face of data enhancements. These ideas are beyond the scope of this chapter.

2.A Figures

Figure 2.1: Sample scheme of the *SAVE* data set

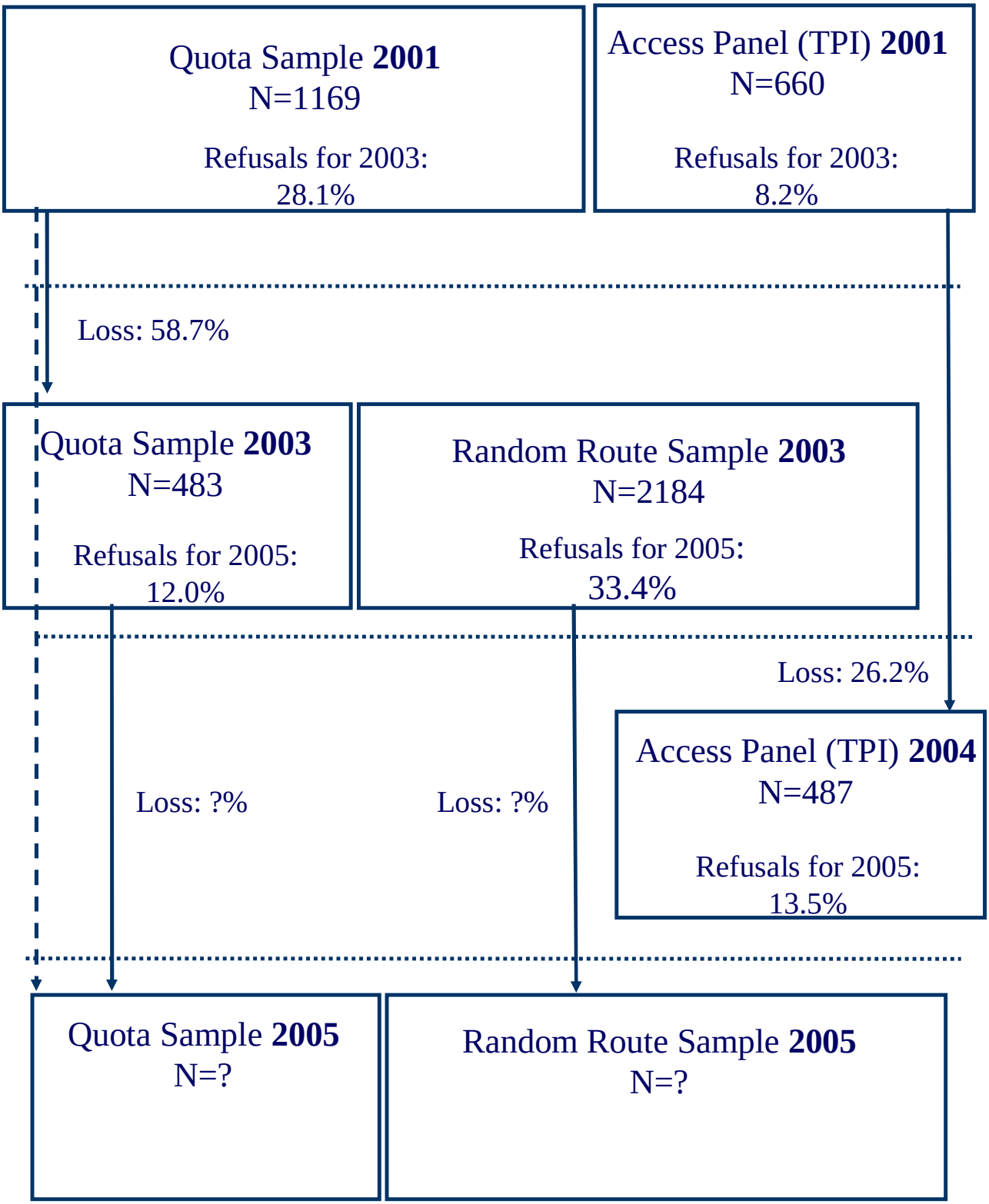
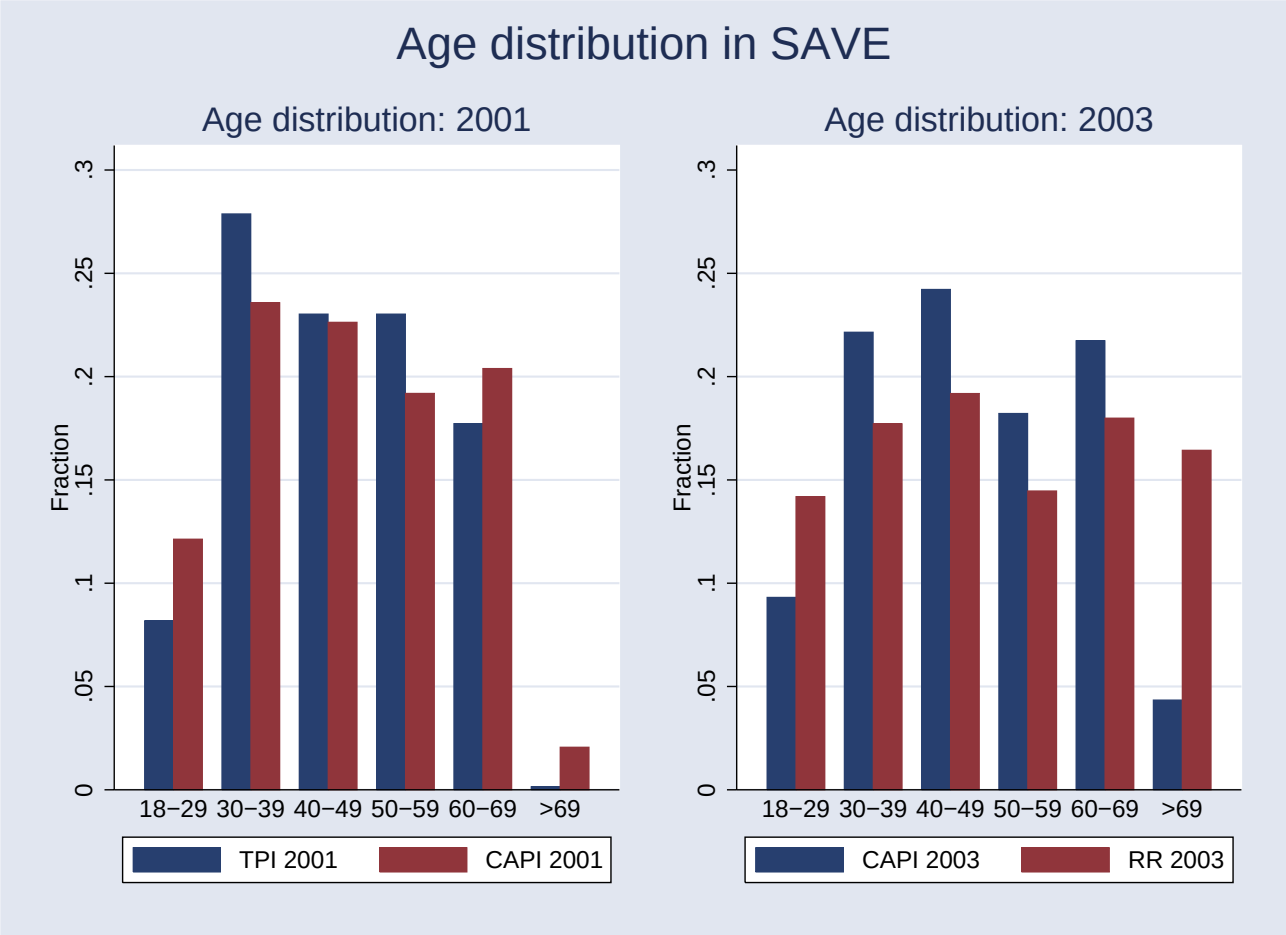


Figure 2.2: Age distribution in the *SAVE* 2001 and 2003 data set



Note: Unweighted values

2.B Tables

Table 2.1: Structure of the *SAVE* questionnaire

Part 1:	Introduction, determining which person will be surveyed in the respective household
Part 2:	Basic socio-economical data of the household
Part 3:	Qualitative questions concerning saving behavior, income and wealth
Part 4:	Budget balance: Quantitative questions concerning income and wealth
Part 5:	Psychological and social determinants of saving behavior
Part 6:	Conclusion: Interview-situation

Table 2.2: Experimental design of the *SAVE* 2001 data set

	Version 1	Version 2	Version 3	Version 4	Version 5
Sampling scheme	Quota	Quota	Quota	Quota	Access panel
Mode: Parts 1, 2, 3, 5	CAPI	CAPI	CAPI	CAPI	P&P
Mode: Part 4 (sensitive items)	CAPI	CAPI	P&P (pick-up)	P&P (mail-back)	P&P (mail-back)
Response rate P&P			98.0%	90.5%	
Question format: income	open-end	open-end	open-end	open-end	open-end
Question format: assets	open-end	brackets	open-end	open-end	open-end
Number of households	295	304	294	276	660

Source: Essig and Winter (2003), Chapter 3

Table 2.3: Design of the *SAVE* 2003 data set

	Panel sample	Refreshment Sample
Sampling scheme	Quota	Random Route
Mode: Parts 1, 2, 3, 5	CAPI	CAPI
Mode: Part 4 (sensitive items)	P&P (pick-up)	P&P (mail-back)
Response rate P&P	98.1%	97.0%
Question format: income	open-end	open-end
Number of households	483	2184

Source: *SAVE* 2003

Table 2.4: Income values: single and double measures

	2001 TPI	2001 CAPI[a]	2001 CAPI[b]	2003 old	2003 new
N	660	599	570	483	2,184
No part D	0	0	32	9	65
Open values	88.2% (582)	88.15% (528)	79.82% (455)	72.88% (352)	65.29% (1426)
Bracket values	23.9% (158)	3.0% (18)	53.86% (307)	63.56% (307)	62% (1354)
Both (open+brackets)	12.9% (85)	0	40.53% (231)	40.79% (197)	34.34% (750)
at least 5' in open field	3.5% (23)	4.67% (28)	2.11% (12)	7.45% (36)	5.95% (130)
at least 10' in open field	0	3.33% (20)	0.18% (1)	4.97% (24)	3.53% (77)
Mean (open values)	2520.11	2922.90	2191.53	4264.2	3385.79
Median (open values)	2351.94	2045.17	2045.17	2200	1800

[a] Only Mode 1 and 2 (full CAPI interview, see section 2.2.2

[b] Mode 3 and 4 (CAPI with dropoff)

Source: *SAVE* 2001 / 2003

Table 2.5: Age distribution in *SAVE* 2001 and 2003

	2001				2003			
	TPI		CAPI		CAPI		Random Route	
Age class	N	Percent	N	Percent	N	Percent	N	Percent
18 - 29	54	8.2	141	12.1	45	9.3	310	14.2
30 - 39	184	27.9	274	23.4	107	22.2	387	17.7
40 - 49	152	23.0	263	22.5	117	24.2	419	19.2
50 - 59	152	23.0	223	19.1	88	18.2	316	14.5
60 - 69	117	17.7	237	20.3	105	21.7	393	18.0
> 69	1	0.2	31	2.7	21	4.4	359	16.4

Source: *SAVE* 2001 / 2003

Table 2.6: Income values class differences

	Number of double answers			
Class differences	2001 TPI	2001 (only CAPI dropoff)	2003 old	2003 new
-11		2		
-10	2	5		2
-9		4		
-8		3		1
-7				
-6	2			
-5				
-4				
-3			1	3
-2	1			1
-1	**1**	**3**		**12**
0		**173**	**137**	**540**
1	**68**	**39**	**39**	**148**
2	5	1	2	5
3	4		1	2
4	2		2	2
5				
6				2
7			1	3
8		1		2
9			3	7
10			6	12
11			4	4
12			1	3
13				1
Total	85	231	197	750

Source: *SAVE* 2001 / 2003

Table 2.7: Representativeness and weights of the *SAVE* 2001 samples

	Low income <2500		Medium income 2500-5000		High income >=5000		All income classes	
Age	CAPI	TPI	CAPI	TPI	CAPI	TPI	CAPI	TPI
up to 35 years	1.18	3.43	0.81	0.74	0.58	0.57	0.88	1.06
	81	17	116	77	52	32	249	126
from 35 up to 55 years	1.18	3.33	0.77	0.71	0.68	0.44	0.79	0.67
	65	14	225	148	201	190	491	352
55 years and older	3.34	6.45	1.12	1.37	0.81	0.69	1.40	1.62
	57	18	177	88	100	71	334	177
All age classes	1.79	4.51	0.90	0.90	0.70	0.52		
	203	49	518	313	353	293		
Household size								
Single	1.86	8.82	0.69	2.88	0.59	2.47	1.18	5.22
	142	18	160	23	28	4	330	45
Two	2.40	2.89	0.60	1.10	0.28	0.52	0.53	0.96
	30	15	329	108	314	103	673	226
3 and more	0.87	1.20	5.26	0.53	11.89	0.46	4.44	0.52
	32	14	30	179	12	185	74	378
All HH size classes	1.78	4.66	0.90	0.90	0.70	0.51		
	204	47	519	310	354	292		

Source: *SAVE* 2001 and German micro-census 2000

Table 2.8: Representativeness and weights of the *SAVE* 2003 samples

	Low income <1300		Medium income 1300-2600		High income >=2600		All income classes	
Age	Panel	RR new	Panel	RR new	Panel	RR new	Panel	RR new
up to 35 years	1.42	0.95	0.97	0.91	0.86	0.75	1.10	0.89
	27	177	38	179	16	81	81	437
from 35 up to 55 years	1.01	0.93	0.74	1.02	0.73	0.91	0.78	0.96
	33	158	99	317	91	319	223	794
55 years and older	2.36	1.17	1.16	1.04	0.74	1.10	1.27	1.10
	32	283	75	366	51	150	158	799
All age classes	1.60	1.05	0.93	1.01	0.74	0.94		
	92	618	212	862	158	550		
Household size								
Single	2.62	1.56	1.12	1.47	1.00	0.99	1.74	1.49
	41	302	48	161	9	40	98	503
Two	0.90	0.69	0.86	0.98	0.56	0.83	0.75	0.87
	32	184	96	371	79	231	207	786
3 and more	0.57	0.36	0.89	0.81	0.92	1.02	0.87	0.81
	19	131	68	330	70	279	157	740
All HH size classes	1.60	1.05	0.93	1.01	0.74	0.94		
	92	617	212	862	158	550		

Source: *SAVE* 2003 and German micro-census 2002

Table 2.9: Effect of weights usage: 2001

	TPI 2001			CAPI 2001		
Weights	None	Inc./Age	Inc./HHSize	None	Inc./Age	Inc./HHSize
Net Income						
Mean	2577.34	1962.76	1933.83	2300.81	2060.59	1941.71
Median	2300.81	1789.52	1789.52	2045.17	1738.39	1636.13
Gross savings						
Mean	5928.24	5903.74	5086.12	4246.96	3586.98	4163.52
Median	2556.46	2812.11	2556.46	2556.46	2045.17	2045.17
Financial Wealth						
Mean	35248.00	25765.22	24293.87	28043.36	22610.99	25842.35
Median	15364.32	8691.96	8078.41	8947.61	5777.60	5112.92
Total Wealth						
Mean	159472.10	152342.60	119679.90	125759.70	104399.10	110759.20
Median	92901.73	51020.79	27090.80	26127.02	15717.11	19684.74

Source: *SAVE* 2001 / 2003

Notes: When no information on weights construction variables (income/age/household size) was available, weights were set to 1.

Table 2.10: Effect of weights usage: 2003

	Panel 2003			RR 2003		
Weights	None	Inc./Age	Inc./HHSize	None	Inc./Age	Inc./HHSize
Net Income						
Mean	2397.00	2091.79	2108.63	2732.43	2635.16	2641.33
Median	2100	1800	1800	1800	1750	1750
Gross savings						
Mean	5160.68	4745.93	4759.52	4333.62	4267.64	4193.16
Median	3000	2500	2500	2400	2400	2400
Financial Wealth						
Mean	29239.61	23393.43	22650.22	21312.56	21062.73	20629.94
Median	7530	4500	3500	2190	2300	2330
Total Wealth						
Mean	140537.40	116894.90	109512.60	139554.90	140014.50	133325.50
Median	38198	21990	18928	8600	9000	8000

Source: *SAVE* 2001 / 2003

Notes: When no information on weights construction variables (income/age/household size) was available, weights were set to 1.

Table 2.11: Income measures: German microcensus and the German Socio-Economic Panel *GSOEP*

	German micro-census[a]		GSOEP			
			Not weighted		Weighted values	
Year	Mean	Median	Mean	Median	Mean	Median
2000	1973.04	./.	2075.99	1891.78	1967.57	1738.39
2001	2015.40	./.	2127.49	1942.91	2000.77	1789.52
2002	2103.78	./.	2525.07	2096	2077.30	1800

[a] Income classes changed from 2001 to 2003. For the lowest class, 400 €were assumed, for the highest, 7800 €.
Source: *SAVE* 2001 / 2003

Table 2.12: Income measures: Income and expenditure survey *EVS*

	Not weighted		Weighted values	
Year	Mean	Median	Mean	Median
1998	2844.30	2510.94	2301.91	1947.56
2003[a]	2612.29	2450	2120.59	1850.00

[a] EVS 2003 income values are self-classified measures for January income. Class means were assumed for the imputation.
Source: *SAVE* 2001 / 2003

Table 2.13: Item nonresponse: descriptive results

	TPI 2001		CAPI 2001[a]		Panel 2003		RR 2003	
	N	Percent	N	Percent	N	Percent	N	Percent
Income								
Nonresponse	78	11.82	83	15.43	122	25.74	693	32.7
Savings accounts								
Ownership	513	78.08	407	76.36	303	65.58	1,153	58.44
Value nonresponse	99	19.3	100	24.57	77	25.41	331	28.71
Stocks								
Ownership	304	46.27	147	27.58	105	22.73	304	15.41
Value nonresponse	57	18.75	33	22.45	38	36.19	134	44.08

[a] Only modes 3 and 4, see table 2.2
Source: *SAVE* 2001 / 2003

Table 2.14: Nonresponse regressions: household net income

Income	All samples		CAPI only	
	Coef.	$P > z$	Coef.	$P > z$
Respondent				
Age	-0.002	0.817	-0.007	0.464
Age squared	0.000	0.555	0.000	0.473
Secondary school (D)	**0.179**	0.002	**0.162**	0.019
Graduation diploma (D)	0.068	0.419	0.008	0.942
University degree (D)	-0.016	0.831	0.012	0.905
Partner (D)	**0.177**	0.001	**0.200**	0.000
East Germany (D)	0.009	0.889	0.073	0.265
Female (D)	0.047	0.389	0.034	0.555
Worker (D)	-0.019	0.827	-0.019	0.849
Civil Servant (D)	0.169	0.155	0.136	0.307
Freelancer (D)	**0.570**	0.002	**0.615**	0.002
Self-employed (D)	**0.233**	0.047	**0.268**	0.037
Part-time working (D)	0.022	0.832	0.013	0.910
Little working (D)	0.118	0.278	0.138	0.239
Not working (D)	0.022	0.813	0.025	0.803
Retired (D)	0.070	0.520	0.071	0.542
Unemployed (D)	-0.004	0.971	-0.004	0.971
Small Community (D)	0.004	0.962	0.040	0.650
Version				
Sample: CAPI 2001 (D)	**0.212**	0.028		
Sample: CAPI 2003 (D)	**0.588**	0.000	**0.358**	0.000
Sample: RR 2003 (D)	**0.792**	0.000	**0.590**	0.000
Interviewer				
Interviewer changed in 2003			0.100	0.569
Experienced > 4 years (D)			-0.023	0.655
Female (D)			0.123	0.017
Older than resp. (D)			-0.156	0.027
Higher schooling (D)			0.009	0.913
Lower schooling (D)			-0.057	0.460
Constant	**-1.563**	0.000	-1.110	0.000
Number of obs	3684		3066	
LR	210.74		126.12	
Prob larger chi2	0.0000		0.0000	
Pseudo R2	0.05030		0.03440	
Log likelihood	-1987.9579		-1768.4779	

Source: *SAVE* 2001 / 2003

Note: Interview versions dropped when part 4 was not P&P

Table 2.15: Nonresponse regressions: savings accounts

Saving accounts	All samples				CAPI only	
	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent						
HH income			0.000	0.737	0.000	0.811
HH income squared			0.000	0.992	0.000	0.928
Age	0.012	0.313	0.016	0.203	0.016	0.229
Age squared	0.000	0.542	0.000	0.350	0.000	0.373
Secondary school (D)	0.032	0.669	0.015	0.838	0.048	0.609
Graduation diploma (D)	0.005	0.963	0.009	0.929	0.068	0.630
University degree (D)	0.006	0.949	-0.018	0.855	0.047	0.731
Partner (D)	**0.191**	0.007	**0.206**	0.005	**0.217**	0.006
East Germany (D)	**-0.162**	0.047	**-0.149**	0.075	-0.137	0.142
Female (D)	0.099	0.156	0.096	0.176	0.125	0.104
Worker (D)	-0.057	0.580	-0.060	0.568	-0.100	0.420
Civil Servant (D)	-0.074	0.584	-0.074	0.584	0.002	0.988
Freelancer (D)	0.168	0.438	0.209	0.339	-0.030	0.909
Self-employed (D)	-0.084	0.578	-0.087	0.577	-0.130	0.464
Part-time working (D)	0.114	0.382	0.124	0.347	-0.007	0.964
Little working (D)	-0.106	0.450	-0.123	0.388	-0.170	0.293
Not working (D)	-0.058	0.639	-0.033	0.797	-0.023	0.869
Retired (D)	-0.108	0.444	-0.143	0.320	-0.150	0.343
Unemployed (D)	0.111	0.510	0.079	0.646	0.072	0.696
Small Community (D)	0.013	0.903	-0.036	0.739	-0.009	0.938
Version						
Sample: CAPI 2001 (D)	**0.189**	0.097	**0.176**	0.072		
Sample: CAPI 2003 (D)	0.210	0.104	**0.191**	0.080	-0.002	0.986
Sample: RR 2003 (D)	**0.310**	0.082	**0.288**	0.001	0.103	0.251
Interviewer						
Interviewer changed in 2003					0.051	0.817
Experienced > 4 years (D)					-0.062	0.358
Female (D)					**0.187**	0.006
Older than resp. (D)					0.036	0.681
Higher schooling (D)					0.079	0.342
Lower schooling (D)					0.142	0.166
Constant	**-1.398**	0.000	**-1.452**	0.000	**-1.451**	0.000
Number of obs	2320		2284		1802	
LR	40.55		40.11		40.39	
Prob larger chi2	0.0064		0.0149		0.0611	
Pseudo R2	0.0154		0.0157		0.0195	
Log likelihood	-1298.4317		-1259.9054		-1017.1128	

Source: *SAVE* 2001 / 2003

Note: Interview versions dropped when part 4 was not P&P

Table 2.16: Nonresponse regressions: stocks

Stocks	All samples				CAPI only	
	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent						
HH income			0.000	0.346	-6.81E-05	0.263
HH income squared			0.000	0.121	3.88E-09	0.156
Age	-0.001	0.971	0.002	0.942	0.005	0.860
Age squared	0.000	0.876	0.000	0.972	0.000	0.983
Secondary school (D)	-0.019	0.885	-0.027	0.843	0.253	0.172
Graduation diploma (D)	-0.008	0.961	-0.014	0.934	0.160	0.515
University degree (D)	-0.136	0.360	-0.158	0.317	0.254	0.301
Partner (D)	0.070	0.579	0.103	0.448	0.107	0.491
East Germany (D)	0.147	0.271	0.158	0.254	0.139	0.415
Female (D)	0.153	0.193	0.179	0.133	0.122	0.389
Worker (D)	-0.100	0.590	-0.088	0.636	-0.083	0.735
Civil Servant (D)	-0.043	0.818	-0.004	0.984	0.129	0.570
Freelancer (D)	0.163	0.543	0.020	0.945	-0.032	0.929
Self-employed (D)	-0.145	0.452	-0.130	0.510	-0.343	0.170
Part-time working (D)	0.161	0.412	0.152	0.444	0.254	0.280
Little working (D)	0.262	0.264	0.282	0.232	0.157	0.612
Not working (D)	0.282	0.154	0.260	0.196	0.279	0.253
Retired (D)	**-0.548**	0.019	**-0.560**	0.018	**-0.687**	0.016
Unemployed (D)	-0.349	0.221	-0.344	0.229	-0.378	0.252
Small Community (D)	-0.072	0.682	-0.046	0.793	-0.033	0.873
Version						
Sample: CAPI 2001 (D)	0.151	0.317	0.143	0.347		
Sample: CAPI 2003 (D)	**0.586**	0.000	**0.541**	0.002	**0.425**	0.040
Sample: RR 2003 (D)	**0.722**	0.000	**0.662**	0.000	**0.509**	0.004
Interviewer						
Interviewer changed in 2003					-0.161	0.681
Experienced > 4 years (D)					0.120	0.332
Female (D)					0.177	0.148
Older than resp. (D)					0.105	0.502
Higher schooling (D)					0.210	0.190
Lower schooling (D)					-0.090	0.573
Constant	**-1.013**	0.089	**-0.987**	0.104	**-1.326**	0.079
Number of obs	840		828		538	
LR	71.97		78.2		52.15	
Prob larger chi2	0.0000		0.0000		0.0037	
Pseudo R2	0.0700		0.0778		0.0743	
Log likelihood	-477.98488		-463.78174		-325.07301	

Source: *SAVE* 2001 / 2003

Note: Interview versions dropped when part 4 was not P&P

Chapter 3

Item nonresponse to financial questions in household surveys: An experimental study of interviewer and mode effects*

3.1 Introduction

Surveys are an important source of data for the empirical analysis of household behavior. Unfortunately, data problems such as unit nonresponse (sample selection), item nonresponse, and measurement error are the rule rather than the exception in survey data. Well-designed studies using household survey data are careful to detect outliers, to impute missing values, and to correct for selection caused by missing observations.

Economists and econometricians have traditionally addressed such data problems using *ex post* approaches such as various imputation schemes or sample selection models. These methods have reached a high level of sophistication, as summarized for instance in the monograph by Wansbeek and Meijer (2000) and in the chapter on "Measurement error in survey data" by Bound *et al.* (2001) in the most recent volume of the *Handbook of Econometrics*. An important drawback of such approaches is that they either require imposing untestable assumptions about the data generating process to ensure point identification of parameters of interest or allow only for much weaker conclusions if weaker assumptions are imposed; see Horowitz and Manski (1995, 1998) for an extensive discussion.

Complementary to correcting data problems *ex post*, researchers have recently increased their efforts to improve survey administration and the design of survey questionnaires so that problems such as item nonresponse can be avoided or at least mitigated *ex ante*. In particular, economists who design survey questions are beginning to use knowledge about the sources of data problems that has been accumulated in other disciplines. For instance, Bound *et al.* (2001) devote a section of their handbook chapter to results from survey research and social psychology that apply to the measurement of quantities that are of economic interest. However, this approach has not been widely used yet. In this Chapter, we show how economists can use knowledge about survey response behavior accumulated in psychology and survey research not only in their analysis of existing data, but also in the design of future household surveys.

We concentrate on a specific aspect of response behavior that is of interest in the empirical analysis households' saving and asset allocation decisions: item nonresponse to questions on financial items in household surveys.[33] Nonresponse in household surveys has been analyzed by various authors,

* This chapter is joint work with Joachim Winter.

[33] In the remainder of this chapter, the term "nonresponse" refers to item nonresponse. We do not address issue of unit nonresponse.

beginning with the work by Ferber (1966); see Schnell (1997) and Beatty and Herrmann (2002) for reviews.[34] However, empirical evidence on response behavior in surveys that focus on financial variables such as income, saving, and asset choice, is still sparse. Recent examples for Germany are Biewen (2001), Riphahn and Serfling (2002), and Schräpler (2003) who work with data from the German Socio-Economic Panel (GSOEP); Nicoletti and Peracchi (2001) investigate nonresponse in the European Household Panel (EHCP). In contrast to these papers, we use data from a controlled experiment that was conducted as part of a representative household survey specifically to analyze the effects of interview mode and question format.

As part of the *SAVE* study, questions on household net income and on six key financial assets were administered using different modes (computer-assisted personal interview vs. self-administered drop-off questionnaire) that were assigned randomly to sample households. We show that nonresponse rates to these sensitive questions are lower in the drop-off questionnaires than in the personal interviews. These results are in line with predictions from models of survey response behavior developed in social psychology that stress, *inter alia*, the role of privacy in answering sensitive questions. Our analysis also confirms earlier findings on the influence of characteristics of the interviewer on response rates in personal interviews.

The plan of the Chapter is as follows. In Section 3.2 we briefly review models of survey response behavior from social psychology that motivate our analysis. The design of the 2001 *SAVE* survey and the embedded experiments on mode effects are described in Section 3.3. In Section 3.4, we present our results, primarily a series of probit regressions with nonresponse dummies for income and six key assets as dependent variables. Section 3.5 summarizes our results and discusses implications for the design of survey questions on financial variables.

3.2 Item nonresponse in household surveys

Why should survey mode and question format influence responses? If respondents are perfectly certain about the quantity in question, they should be able to give the correct answer. However, respondents are rarely certain about quantities they are asked to report in household surveys. Therefore, the formation of answers to survey questions is a complicated process. As a starting point for thinking about item nonresponse and other data problems *ex ante*, or to correct for resulting bias in survey data *ex post*, it is useful to review existing research by psychologists and survey methodologists in some detail.

Since the early 1980's, psychologists and survey methodologists have worked together, trying to understand the cognitive and communicative processes that govern survey response behavior. One of the first systematic attempts to analyze survey response behavior as an interaction between the interviewer and the respondent, and to uncover the cognitive processes involved in answering survey questions, is Tourangeau (1984). There is now an extensive literature in survey research on cognitive processes that generate survey responses and on pitfalls that should be avoided in survey design; Sudman *et al.* (1996) and Tourangeau *et al.* (2000) provide overviews of the literature on survey response

[34] The edited volumes by Groves and Couper (1998) and Groves *et al.* (2002) are devoted entirely to survey nonresponse.

behavior and question design in cognitive and social psychology. Cognitive issues in households' reports of financial variables, in particular with respect to reports of household income, are discussed by Moore *et al.* (1999).

An important insight from survey research is that the process of forming the response to a survey question consists of several steps, each of which might contribute to the fact that answers often do not provide reliable measures of the quantity in question. Survey respondents first have to understand the question and determine which quantity they are to report on. To do so, they draw on a wide range of contextual information in ways that researchers are often unaware of. Second, respondents have to recall information on the quantity from memory. In many instances, respondents will have imperfect recall and need to apply various inference and estimation strategies to arrive at an answer; this is the third step of the response process. Fourth, once respondents have arrived at an answer, they need to map it onto the response alternatives provided by the researcher (unless the question format is open-ended). Finally, respondents may edit their answer because of social desirability and self-representation concerns (i. e., even though they might be aware of the "true" value of some quantity, they on purpose or unconsciously report a higher or lower value).

In order to derive hypotheses about factors that influence item nonresponse on financial questions in household surveys, we use a conceptual model by Tourangeau and Smith (1996). For our purpose, the advantage of such a model is that it makes cognitive processes and social interaction between the interviewer and the respondents explicit. This conceptual model links interview modes, psychological variables, and data quality. Specifically, dimensions of data quality, such as accuracy, reliability, and item nonresponse are influenced by three psychological variables: privacy, legitimacy, and cognitive burden. The signs of the relationships form the basis for hypotheses about survey response behavior. For instance, privacy reduces the problem of item nonresponse on sensitive questions while increased cognitive burden reduces response accuracy. The key variables privacy, legitimacy, and cognitive burden are influenced, in turn, by the mode of data collection (face-to-face interviews, self-administrated surveys, computerized surveys such as internet surveys, or telephone interviews with auditory presentation).

In our analysis, we specifically concentrate on the trade-off between using a computer-assisted personal interview (CAPI) and a self-administered paper-and-pencil questionnaire (P&P) for collecting data on sensitive financial variables. The prediction from models of survey response behavior such as the one outlined above is that relative to CAPI, the self-administered P&P interview should result in higher perceived levels of privacy, which in turn increases responses accuracy and decreases the rate of item nonresponse.

A second hypothesis we test is related to the social interaction between interviewer and respondent. It seems plausible that in personal interviews, characteristics of the interviewer may influence response behavior; interviewer effects on survey response have been analyzed as far back as Rice (1929). For instance, there is evidence that interviewer attributes such as age and gender affect response rates in surveys. Interestingly, the effects of interviewer experience on response behavior seem to be stronger than those of personal characteristics; see Groves and Couper (1998), Chapter 7, and Hox and de Leeuw

(2002) for reviews. Riphahn and Serfling (2002) provide empirical evidence on interviewer effects in the German Socio-Economic Panel (GSOEP).

Finally, response behavior may depend on the respondent's motivation and on incentives for providing accurate answers. These aspects are analyzed in the rational-choice approach to survey response behavior; see Philipson (1997), Philipson (2001), Philipson and Malani (1999), Singer (2002), and Stocké (2003). In the analysis presented in this Chapter, we do not address these issues – our field experiment does not contain reliable indicators of respondents' motivation, and no incentives were used to increase response accuracy.

3.3 The field experiment embedded in *SAVE* 2001

In this section, we discuss the embedded survey experiments. For a short overview of the *SAVE* 2001 study, please see Chapter 2.2.1.

The embedded experimental design of the *SAVE* 2001 study is summarized in Chapter 2, Table 2.2. The first four versions were computer aided personal interviews (CAPI); they were carried out by NFO Infratest, Munich. In contrast, the fifth version was a conventional paper questionnaire ("paper and pencil", P&P). The CAPI interviews were carried out using quota samples whereas conventional P&P questionnaires were given to a so-called Access Panel operated by the company TPI (Test Panel Institute, Wetzlar), in other words a standing panel of households surveyed at regular intervals.

The only difference in the four versions of the CAPI interview is in the critical part 4 of the questionnaire. In versions 1 and 2, all questions were administered by CAPI in the presence of the interviewer. The difference between these versions is that the questions on asset holdings were presented using an open-ended format with follow-up brackets (range cards) in version 1 and with "forced" brackets in version 2. The experimental manipulation of the question format with respect to follow-up vs. forced range-card questions is not investigated in the present Chapter. For a discussion of how follow-up questions alleviate the problem of item nonresponse, see Juster and Smith (1997). Hurd *et al.* (1998) and Winter (2002a) investigate response biases such as anchoring that arise in follow-up questions that use unfolding brackets or range cards, respectively.

Because many of these questions relate to intensely personal matters of income and wealth, we went one step further in versions 3 and 4. In these versions, part 4 of the questionnaire was not part of the personal CAPI interview, but left as a paper-and-pencil questionnaire by the interviewer (this mode is termed "P&P drop off" in the sequel). In version 3, the interviewer came back personally to collect the drop-off questionnaire; in version 4, the questionnaire had to be returned by mail using a pre-paid envelope. If this was not done within a specified number of days, the respondent was reminded by telephone several times. Nevertheless, response rates for the drop-off questionnaire were significantly lower in version 4 than in version 3 (90.5% vs. 98.0%).

Summarizing, in order to test our hypothesis that there is an anonymity/privacy effect on nonresponse to sensitive financial questions, we could compare response behavior in versions 1 and 2 to that in versions 3, 4, and 5. In this Chapter, we use data from versions 1, 3, and 4 only, for reasons detailed below.

The survey took place in early summer 2001. The fieldwork for the personal interviews took place between May 29 and June 26, 2001, whereas the fieldwork for the Access Panel took place between June 29 and July 24, 2001. Both the CAPI (quota sample) and the P&P (TPI Access Panel) segments were targeted at households with head of the household aged between 18 and 69 years. For the CAPI versions, the quota performance targets were related to the dimension gender (male respondent ratio of 74 percent) and age (a distribution in age classes under 25, 25-34, 35-50 and 50-70 years) according to the current official population statistics (and, in particular, the 2000 micro census). For the TPI interviewees, the quota targets were also based on the 2000 micro census and related to whether the respondent is a wage earner or a salaried employee, and the size of the household. Table 2.2 shows the sample sizes for the five survey versions. In total, 1,829 households were surveyed.

3.4 Results

In this section, we present our findings on response behavior in the 2001 *SAVE* survey and relate these findings to the hypotheses presented in Section 3.2. We present summary statistics for the dependent and independent variables in Sections 3.4.2 and 3.4.2, respectively. We then turn to the main part of the analysis, a series of probit regressions that allow us to test for factors that influence item nonresponse in Section 3.4.3. Other aspects of response behavior that might be related to interview mode and interviewer characteristics are briefly discussed in section 3.4.4.

In the following, we restrict the analysis to a comparison of version 1 (CAPI) with versions 3 (P&P drop-off, pick-up) and 4 (P&P drop-off, mail-back). As noted above, version 2 differs in the format of asset questions ("forced" brackets rather than open-ended questions with follow-up brackets). An analysis of the effects of question format on nonresponse is a separate issue from the mode effects we are interested in here, so we leave that to future work. Version 5 of the *SAVE* 2001 study used a different sample (drawn from a standing Access Panel) that exhibits a significant middle-class bias.[35] We decided not to use data from version 5 in most of the subsequent analysis because of these differences in sample composition. While differences in observable characteristics could potentially be resolved using matching techniques, the problem is actually deeper: It is very likely that households in the Access Panel differ not only in observable characteristics from a random or quota sample, but also in unobservable characteristics that are relevant for our substantive analysis. For instance, members of an Access Panel typically have some survey experience, and their response behavior might therefore differ from that in a representative sample.[36]

[35] While the average household size in all four CAPI versions is about 1.9, the average household in the version 5 has about 3 persons. Specifically, the number of single households is much lower in the sample used for version 5. Furthermore, in version 5, less households were interviewed in Eastern Germany. Also, there are significant differences concerning education and the proportion of workers and employees between the four CAPI versions and version 5.

[36] Such unobserved differences would violate the "ignorability" assumption that is required for the application of matching techniques.

3.4.1 Dependent variables

In the main part of our analysis, we use probit regressions with indicators for item nonresponse on the household net income and on six asset questions as the dependent variables. In addition, we use the incidence of focal (rounded) values in responses to the income question and the response to the question of whether respondents would be willing to participate in a future wave of the *SAVE* survey as additional dependent variables. Descriptive statistics for all dependent variables are reported in Table 3.1.

3.4.2 Independent variables

The independent variables we use in the subsequent analysis fall into four categories: (i) characteristics of the respondent, (ii) characteristics of the interviewer, (iii) self-reported feedback by the respondents, and (iv) interview mode (i. e., the different versions of the survey assigned by our experimental design). Summary statistics for the first three sets of independent variables are reported in Table 3.2.

The set of demographic and economic characteristics of the respondent contains age and, in some regressions, household net income (and the squared values of these variables to allow for nonlinear effects). Note that there is some nonresponse to the open-ended income question itself, so we imputed these missing values using the information from the follow-up bracket question. To check whether these imputations affect the regression results, we include a dummy for households with imputed values in those regressions that contain income as an independent variable. Other respondent characteristics we use are three dummy variables for level of education (the reference category is primary school); a dummy for households in East Germany; a set of dummy variables for occupation and labor market status (the reference category is white-collar employees)[37]; and a dummy variable for households in small communities with a population of less than 5000.

As can be seen in Table 3.2, there are no striking differences in respondent and household characteristics between versions 1, 3, and 4 which were randomly assigned to households within the quota sample. One exception is income which will be used as an independent variable in the nonresponse regressions for assets, but this variable was administered differently across versions[38], so differences in responses are not surprising.

A second set of variables contains characteristics of the interviewer. In total, 267 interviewers administered the 1169 CAPI interviews, with a maximum of eight interviews by interviewer and a minimum of one. Considering this rather small average number of interviews by interviewer, we refrain from using interviewer dummies as explanatory variables. From the survey agency that administered the survey, we obtained data on gender, age, the level of schooling of the interviewer, and his interview experience (as measured by the number of years she has been working for the survey agency). In the regressions, we use a dummy variable for experienced interviewers, defined as having more than median experience (4 years); a dummy variable for female interviewers; a dummy variable for interviewers who are older than the respondent; and two dummy variables for interviewers with lower and higher

[37] The dummy variable for farmers is dropped from the regressions since we have only two farmers in the sample.

[38] The income questions were asked in the drop-off part of the questionnaire.

education level than the respondent (with the categories defined as described above). We experimented with more complicated specifications for the age relation between interviewer and respondent, but we did not find results that were qualitatively different for those based on just a dummy for older interviewers, so we report only those results below.

The third set of explanatory variables is based on respondents self reports to an open-ended feedback question that was administered at the end of the interview. We have classified the responses to this question using a set of keywords which resulted in four indicator variables for whether the respondent mentioned specific aspects in a negative or positive way. A negative statement is coded as -1, a positive statement as 1, no statement as 0. The four aspects are: (1) overall reaction to the interview; (2) concerns about privacy; (3) length of the interview; (4) questions easy to answer. We also constructed a dummy variable that indicates whether at least one of these four aspects was mentioned in a positive way.

The final set of variables controls for mode effects. We are interested in the effects of a CAPI interview vs. a P&P drop-off questionnaire on nonresponse rates. In addition, we would like to distinguish between drop-off questionnaires that are picked up by the interviewer (version 3) and mailed back directly to the survey agency (version 4). We therefore include two dummy variables for versions 3 and 4, respectively, in our regressions. The frequencies of the interview mode indicators are reported in Table 2.2, Chapter 2.

3.4.3 Nonresponse regressions

Tables 3.3 through 3.9 contain probit estimates for nonresponse rates for absolute values of monthly net household income and the balances held in six asset categories. We should note that in the case of the asset regressions, the dependent variable is always nonresponse conditional on holding that asset. For each of those assets, households were first asked whether they hold it, and they were asked for the amount only if they do. Since ownership rates for the six assets vary, so do the numbers of observations used in the asset nonresponse regressions. For all but savings accounts, they are actually quite small, and we therefore discuss only the results of the regression for nonresponse on the income and savings account questions in detail.

For the income nonresponse regression, we report two specifications, one with interviewer characteristics and one without (Table 3.3). Few of the respondent and household characteristics are significant – nonresponse rates are higher in East Germany, lower for blue-collar workers, and higher in small communities. Interestingly, the willingness to report income is not affected by the interview mode – the coefficients of the dummy variables for the P&P drop-off versions are not significant. These results also hold when interviewer characteristics are included. The only interviewer variable that is significant is the dummy for older interviewers; they seem to have a positive effect on willingness to report income. Overall, nonresponse on the household net income question appears to be very heterogenous and hard to explain with respondent and household characteristics, interview mode, and interviewer characteristics. However, self-reported feedback measures that characterize how the respondents have perceived the interview situation have explanatory power.

The situation is different for assets. In the regression results for nonresponse to the question on the balance held in saving account (Table 3.4), we see that respondent and household characteristics still have few significant characteristics. Nonresponse is higher for the unemployed and lower in small communities. The latter is an interesting sign change compared with the income regression. Most importantly, we see strongly negative coefficients of the dummy variables for the P&P drop-off versions. Respondents are more willing to report their saving account balance in the private P&P interview mode. The dummy variable for imputed income (i. e., nonresponse to the income question) is significantly positive which indicates that there is some consistency in nonresponse across questions. Finally, there are effects of interviewer characteristics, and there is again some evidence that self-reported feedback is related to item nonresponse.

We do not comment on the nonresponse regressions for the other five assets in detail because the number of households who hold these assets is smaller. However, we should note that the negative coefficients of the dummy variables for the P&P drop-off modes can be found in most of these regressions. This is strong evidence for a mode effect in nonresponse to questions on asset holdings – respondents are much more willing to answer if such questions are self-administered and private. In some but not all cases, the coefficient of the version 4 dummy is larger in magnitude than that of the version 3 dummy. This finding is consistent with even lower rates of nonresponse when the drop-off questionnaire is mailed back rather than being picked up by the interviewer. However, as reported in Table 2.2 in Chapter 2.B, response rates for the drop-off questionnaires are lower for version 4 than for version 3 in the first place.

3.4.4 Other dimensions of survey response behavior

The incidence of focal points ("round" values) in the responses to open-ended questions is a direct measure of data quality. There are two primary reasons why survey respondents give focal-point responses to open-ended questions (see Tourangeau *et al.* (2000), section 8.1). First, rounding could reflect uncertainty about the exact value of the quantity asked. Second, even if a respondent knows the exact value of a quantity, she could round this value because of privacy concerns and other aspects of the interviewer-respondent interaction. Moreover, there could be dependence between these two effects. For instance, in the presence of an interviewer, a respondent who is uncertain about the quantity in question might report a guess, reflected in a focal-point response (so as not to disappoint the interviewer by a complete refusal). In contrast, in a self-administered mode, the respondent might be more willing to admit that she does not have an exact answer by giving an explicit "Don't know" response. Finally, other aspects of the interviewing process such as time pressure (which is more intense in personal interviews than in self-administered surveys such as a paper-and-pencil questionnaire) could also induce the respondent to report a focal-point guess rather than thinking a little longer – that is, evoking a more elaborate cognitive process – to come up with an exact answer.

The present study was not designed to disentangle these potential explanations of focal responses. However, our data allow to assess whether the factors that influence item nonresponse also affect the incidence of focal-point (rounded) rounded answers from those who respond. For this purpose, we focus on the income question for which nonresponse regressions were reported above in Table

3.3. Tables 3.10, 3.11, and 3.12 contain probit regressions in which the dependent variable is an indicator of whether the response is a multiple of 100, 500, or 1000, respectively. The most interesting observation in these regressions is that Version 3 (drop-off questionnaire picked up by the interviewer) has generated the smallest fraction of focal responses. This is evidence against the hypothesis that rounding is caused by privacy concerns or other aspects of the interviewer-respondent interaction. Rather, rounding seems to be related to response uncertainty and the opportunity and incentives to look up correct values. While these results are interesting, more research is needed to substantiate these claims. In particular, having a direct measure of respondents' uncertainty about the quantity in question seems important; see Winter (2002b) for a recent attempt in that direction.

Finally, we analyze responses to the question of whether the respondent would be willing to participate in a future wave of the *SAVE* survey. The fraction of "yes" responses varies between about 60% and 70% in versions 1 through 4 which were administered with quota samples (see Table 3.1).[39] To check whether respondent and interviewer characteristics affect the willingness to participate in future surveys, we run probit regressions similar to those for nonresponse reported in Section 3.4.3 above. Table 3.13 reports the results. Few of the independent variables are significant. However, a consistent pattern emerges: Respondents in version 3 (which included a drop-off questionnaire on financial items collected by the interviewer) are less willing to participate in future surveys. In the specifications that also control for income, dummies for version 3 also has a significant negative coefficients. From these results, it follows that having to return the drop-off questionnaire to the interviewer at a future point reduces stated willingness to participate in future surveys.

The present design does not allow to disentangle all possible explanations for this observation. One explanation might be that the drop-off questionnaire is perceived as a burden – respondents know that they have to do additional work after the interviewer has left. The absolute value of the coefficient of the version 3 dummy is larger than that of the version 4 dummy; this is consistent with the hypothesis that in version 4 (drop-off questionnaire mailed in rather than being picked up), respondents feel less obliged to do the extra work of filling in the drop-off questionnaire. This corresponds the lower return rate in version 4, 91% compared with 98% in version 3, see Table 3.1.

A weakness of this explanation is that the proxy variable for overall satisfaction (willingness to participate in future surveys) was obtained before the drop-off questionnaire has been answered; it might be possible that respondents are more willing to participate in future surveys after having answered (because they realized that the burden was less than they expected) so that the mode effect disappears. A reliable analysis of alternative explanations for the finding that the drop-off modes reduce willingness to participate in future surveys would therefore require a more complex experimental design; this is left for future research.

In any case, the survey protocol reflects that German law which requires that only those respondents who agree explicitly to participate in future waves can be contacted for a re-interview. The results on the reported willingness to participate in future surveys indicate that the higher response rates on sensitive financial items achieved in interviews with drop-off questionnaires might come at a

[39] Not surprisingly, it is significantly higher (about 90%) in version 5 that used a standing access panel of persons who had already agreed to answer household surveys in the past.

price. Drop-off questionnaires appear to increase the perceived burden of the interview and to reduce respondents' overall satisfaction.

The sample of the 2003 *SAVE* survey contains re-interviews of 2001 sample members who said they are willing to participate in a future wave (augmented by a refreshment sample). This opens the door to judge whether stated willingness to participate in future survey translates in actual behavior, but this is an issue that deservers attention in future research.

3.5 Conclusions

The present Chapter investigated the effects of interview mode on nonresponse to sensitive questions on items such as income and asset holdings using data from a field experiment. The main hypothesis we tested was that a self-administered interview mode results in lower rates of nonresponse than a personal interview, as suggested by models of survey response behavior developed in social psychology and survey research. We found that in comparison to the CAPI mode, rates of nonresponse are lower in a paper-and-pencil drop-off questionnaire that could be answered in private and independently of the rest of the survey interview. This effect is very strong for all six asset categories we analyzed while it is not significant for the question on household net income.

Respondent and household characteristics as well as interviewer characteristics do not appear to have strong and consistent effects on nonresponse to sensitive financial questions. This raises the question of whether correcting for item nonresponse using complex designs that require an explicit model of the nonresponse process offer much gain over straightforward imputation schemes that invoke a "missing at random" assumption. This aspect is certainly worth further investigation; our work illustrates the usefulness of controlled survey experiments for such an analysis.

Another finding is that data quality for those households who actually answer also seems to be better in the P&P drop-off modes, as judged by the lower frequency of focal-point responses (which suggests that these responses are more accurate). This observation could be explained by the fact that respondents have more time to answer a drop-off questionnaire. We know from survey research that respondents are more likely to invoke more elaborate cognitive processes when they have more time to answer questions. These more elaborate processes should result in more accurate responses. Alternatively, respondents may be more likely to look up exact quantities when they fill in a drop-off questionnaire than in a personal interview situation. We cannot distinguish between these explanations with the experimental data obtained from the present study. This issue should be explored in future research.

Our results have a number of practical implications. Data quality seems to be better if sensitive questions on financial items are administered in a private interview mode, and drop-off questionnaires seem to be a practical way to implement private data collection modes within a random or quota sampling scheme (such as random route) that requires personal interviewer contact. However, it should be noted that response rates for the drop-off questionnaire might cause a problem – while the 98% achieved in *SAVE* 2001 when the interviewer came back to pick up the drop-off questionnaire is

acceptable, a 90.5% response rate for mail-back questionnaires might be too low. This aspect should be analyzed in future work.

Finally, we have seen that stated willingness to participate in future surveys is lower if the survey interview consists not only of the CAPI component but also has a drop-off questionnaire. While it is unclear whether respondents would state that they are more willing to participate in a future survey *after* they have actually filled in the drop-off questionnaire, the interview protocol that is typically followed requires that the question on future surveys be asked at the end of the CAPI interview. A piece of practical advice would be to move that question to the end of the drop-off questionnaire. In any case, the effect that the interview mode chosen for sensitive financial questions has on participation in a follow-up survey will be investigated with data from the upcoming *SAVE* 2003 wave.[40]

[40] This is work in progress.

3.A Tables

Table 3.1: Dependent variables: Response rates and willingness to participate in future surveys

	Version 1 (CAPI)	Version 3 (P&P pickup)	Version 4 (P&P mail-in)
Income			
HH refuse values	56	58	57
in % of HH	19.0%	19.7%	20.7%
Saving Accounts			
HH owning	215	215	192
in % of all HH	72.9%	73.1%	69.6%
HH refuse values	100	50	31
in % of HH owning	46.5%	23.3%	16.2%
Building Society Contracts			
HH owning	85	90	83
in % of all HH	28.8%	30.6%	30.1%
HH refuse values	38	23	22
in % of HH owning	44.7%	25.6%	26.5%
Life Insurances			
HH owning	128	131	123
in % of all HH	43.4%	44.6%	44.6%
HH refuse values	71	46	36
in % of HH owning	55.5%	35.1%	29.3%
Retirment Savings			
HH owning	43	45	34
in % of all HH	14.6%	15.3%	12.3%
HH refuse values	34	24	14
in % of HH owning	79.1%	53.3%	41.2%
Bonds			
HH owning	38	38	49
in % of all HH	12.9%	12.9%	17.8%
HH refuse values	18	17	15
in % of HH owning	47.4%	44.7%	30.6%
Stocks, Funds			
HH owning	97	92	55
in % of all HH	32.88	31.29	19.93
HH refuse values	50	23	11
in % of HH owning	51.6%	25.0%	20.0%
Would participate another time			
yes	66.4%	58.8%	60.1%
no	26.4%	31.3%	30.8%
no answer	7.1%	9.9%	9.1%
Number of households	295	294	276

Table 3.2: Independent variables: respondent and interviewer characteristics

	Version 1		Version 3		Version 4	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Respondent						
HH income	5731.8	8870.3	4499.9	3931.4	4396.3	2345.7
HH income imputed (D)	4.7%	21.3%	16.0%	36.7%	10.5%	30.7%
Age	45.0	13.4	46.6	13.3	46.6	13.4
Secondary school (D)	36.3%	48.2%	36.1%	48.1%	40.2%	49.1%
Graduation diploma (D)	13.2%	33.9%	11.9%	32.4%	10.1%	30.2%
University degree (D)	14.9%	35.7%	16.7%	37.3%	11.2%	31.6%
East Germany (D)	22.4%	41.7%	19.4%	39.6%	19.6%	39.7%
Worker (D)	19.7%	39.9%	14.1%	34.9%	19.6%	39.7%
Civil Servant (D)	5.8%	23.4%	5.2%	22.1%	5.4%	22.7%
Farmer (D)	0.3%	5.8%	0.3%	5.9%	0.0%	0.0%
Self-employed (D)	6.8%	25.2%	5.8%	23.5%	7.6%	26.6%
Retired (D)	17.6%	38.2%	21.4%	41.1%	17.8%	38.3%
Unemployed (D)	5.8%	23.3%	7.2%	26.0%	3.6%	18.8%
Small Community (D)	13.6%	34.3%	15.0%	35.7%	10.9%	31.2%
Interviewer						
Experienced $>$ 4 years (D)	43.8%	49.7%	61.4%	48.8%	55.1%	49.8%
Female (D)	43.7%	49.7%	59.5%	49.2%	29.7%	45.8%
Older than resp. (D)	51.2%	50.1%	58.5%	49.4%	55.4%	49.8%
Higher schooling (D)	38.3%	48.7%	37.1%	48.4%	42.0%	49.5%
Lower schooling (D)	25.1%	43.4%	26.9%	44.4%	19.9%	40.0%
Feedback						
Positive opinion / Interesting subject	6.4%	52.1%	10.5%	57.8%	1.1%	56.9%
Privacy	-15.6%	37.3%	-18.4%	39.7%	-10.5%	33.0%
Interview not too long	-0.7%	16.5%	-0.7%	16.5%	1.4%	12.0%
Easy to answer	-1.4%	16.4%	0.3%	10.1%	-0.4%	10.4%
at least one of the 4 latter positive	-9.8%	67.5%	-4.8%	73.3%	-7.2%	66.8%

Note: Dummy variables are marked (D).

Table 3.3: Nonresponse regressions: income

Income	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
Age	-0.001	0.964	-0.007	0.813	-0.008	0.784	-0.010	0.735
Age squared	0.000	0.962	0.000	0.882	0.000	0.885	0.000	0.928
Secondary school (D)	0.053	0.695	0.090	0.551	0.048	0.754	0.049	0.747
Graduation diploma (D)	-0.151	0.444	-0.070	0.751	-0.152	0.499	-0.138	0.537
University degree (D)	**-0.409**	0.030	-0.355	0.111	**-0.451**	0.047	**-0.448**	0.048
East Germany (D)	**0.313**	0.019	**0.345**	0.013	**0.383**	0.006	**0.392**	0.005
Worker (D)	**-0.409**	0.012	**-0.413**	0.013	**-0.440**	0.009	**-0.435**	0.010
Civil Servant (D)	-0.442	0.118	-0.453	0.115	-0.453	0.122	-0.476	0.104
Self-employed (D)	0.262	0.175	0.301	0.126	0.319	0.111	0.334	0.094
Retired (D)	0.056	0.776	0.055	0.787	0.048	0.812	0.054	0.790
Unemployed (D)	-0.264	0.279	-0.223	0.375	-0.207	0.414	-0.230	0.365
Small Community (D)	**0.350**	0.017	**0.394**	0.010	**0.394**	0.010	**0.399**	0.009
Version								
Interview version 3 (D)	0.050	0.689	0.115	0.377	0.117	0.375	0.130	0.323
Interview version 4 (D)	0.093	0.460	0.157	0.226	0.181	0.169	0.169	0.195
Interviewer								
Experienced > 4 years (D)			-0.041	0.703	-0.008	0.940	-0.020	0.855
Female (D)			0.089	0.418	0.080	0.471	0.075	0.496
Older than resp. (D)			**-0.410**	0.003	-0.431	0.002	**-0.429**	0.002
Higher schooling (D)			-0.126	0.323	-0.147	0.254	-0.160	0.215
Lower schooling (D)			-0.215	0.156	-0.188	0.220	-0.211	0.168
Feedback								
Positive / Interesting subject					**-0.182**	0.064		
Privacy					**-0.293**	0.029		
Interview not too long					-0.503	0.209		
Easy to answer					0.053	0.899		
at least one of the 4 latter pos.							**-0.242**	0.002
Constant	-0.820	0.204	-0.272	0.698	-0.243	0.731	-0.172	0.808
Number of obs	847		836		836		836	
LR	30.75		46.5		56.26		56.35	
Prob larger chi2	0.0060		0.0004		0.0001		0.0000	
Pseudo R2	0.03710		0.05700		0.06890		0.06900	
Log likelihood	-399.43856		-384.90723		-380.03051		-379.98598	

Table 3.4: Nonresponse regressions: savings accounts

Saving Accounts	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
HH income	-1.04E-06	0.978	1.97E-06	0.960	3.85E-06	0.922	3.97E-06	0.919
HH income squared	1.10E-10	0.880	1.19E-10	0.876	7.85E-11	0.918	7.20E-11	0.925
HH income imputed (D)	**0.483**	0.006	**0.487**	0.008	**0.477**	0.010	**0.470**	0.011
Age	-0.020	0.559	-0.008	0.825	-0.007	0.842	-0.007	0.839
Age squared	0.000	0.568	0.000	0.686	0.000	0.705	0.000	0.703
Secondary school (D)	0.206	0.191	0.182	0.310	0.162	0.369	0.158	0.381
Graduation diploma (D)	0.039	0.859	-0.124	0.635	-0.161	0.541	-0.156	0.551
University degree (D)	-0.182	0.429	-0.336	0.226	-0.408	0.147	-0.410	0.144
East Germany (D)	-0.201	0.263	-0.181	0.328	-0.152	0.414	-0.152	0.414
Worker (D)	-0.048	0.784	-0.126	0.489	-0.128	0.482	-0.119	0.512
Civil Servant (D)	-0.204	0.416	-0.299	0.246	-0.303	0.245	-0.321	0.217
Self-employed (D)	-0.414	0.126	**-0.579**	0.041	-0.578	0.045	-0.569	0.048
Retired (D)	-0.159	0.508	-0.225	0.358	-0.213	0.384	-0.207	0.399
Unemployed (D)	**0.554**	0.067	**0.513**	0.098	0.542	0.081	0.528	0.089
Small Community (D)	**-0.442**	0.035	**-0.494**	0.032	-0.498	0.031	-0.488	0.035
Version								
Interview version 3 (D)	**-0.575**	0.000	**-0.529**	0.000	**-0.552**	0.000	**-0.532**	0.000
Interview version 4 (D)	**-0.871**	0.000	**-0.785**	0.000	**-0.793**	0.000	**-0.789**	0.000
Interviewer								
Experienced > 4 years (D)			**-0.381**	0.002	**-0.352**	0.006	**-0.362**	0.004
Female (D)			0.190	0.140	0.181	0.164	0.179	0.165
Older than resp. (D)			0.210	0.181	0.221	0.163	0.212	0.178
Higher schooling (D)			0.008	0.960	-0.017	0.913	-0.024	0.875
Lower schooling (D)			**0.342**	0.063	**0.368**	0.048	**0.360**	0.052
Feedback								
Positive / Interesting subject					-0.113	0.291		
Privacy					-0.261	0.121		
Interview not too long					-0.128	0.751		
Easy to answer					-0.024	0.953		
at least one of the 4 latter pos.							**-0.168**	0.054
Constant	0.233	0.754	-0.266	0.742	-0.308	0.704	-0.282	0.728
Number of obs	586		579		579		579	
LR	58.38		74.6		78.19		78.32	
Prob larger chi2	0.0000		0.0000		0.0000		0.0000	
Pseudo R2	0.0868		0.113		0.1184		0.1186	
Log likelihood	-307.28906		-292.87995		-291.08286		-291.01689	

Table 3.5: Nonresponse regressions: buildings society savings

Build. Soc. Contr.	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
HH income	**0.000**	0.059	**0.000**	0.043	**0.000**	0.035	**0.000**	0.043
HH income squared	**1.83E-09**	0.093	**2.07E-09**	0.080	**2.11E-09**	0.070	**2.07E-09**	0.080
HH income imputed (D)	**0.641**	0.067	**0.6281**	0.089	**0.709**	0.060	**0.6327**	0.091
Age	0.026	0.610	0.0190	0.710	0.015	0.775	0.0188	0.712
Age squared	0.000	0.868	0.0001	0.802	0.000	0.701	0.0001	0.799
Secondary school (D)	0.028	0.906	0.1107	0.676	0.158	0.557	0.1122	0.673
Graduation diploma (D)	-0.173	0.597	-0.0694	0.863	0.017	0.967	-0.0684	0.865
University degree (D)	-0.156	0.646	-0.0949	0.814	-0.001	0.998	-0.0903	0.825
East Germany (D)	-0.272	0.278	-0.3151	0.241	-0.324	0.235	-0.3182	0.242
Worker (D)	0.011	0.964	-0.0356	0.889	-0.039	0.879	-0.0365	0.886
Civil Servant (D)	-0.221	0.471	-0.1658	0.599	-0.230	0.476	-0.1646	0.602
Self-employed (D)	-0.385	0.456	-0.4252	0.409	-0.474	0.369	-0.4282	0.407
Retired (D)	**-0.717**	0.093	**-0.8189**	0.056	-0.897	0.041	-0.8214	0.056
Unemployed (D)	0.193	0.673	0.1207	0.796	0.071	0.880	0.1213	0.795
Small Community (D)	**-0.598**	0.078	**-0.6907**	0.067	**-0.697**	0.067	**-0.6903**	0.068
Version								
Interview version 3 (D)	**-0.491**	0.029	**-0.4658**	0.057	**-0.449**	0.075	-0.4675	0.058
Interview version 4 (D)	**-0.473**	0.034	**-0.4587**	0.047	**-0.469**	0.050	-0.4607	0.048
Interviewer								
Experienced > 4 years (D)			-0.1575	0.434	-0.220	0.293	-0.1592	0.432
Female (D)			-0.0342	0.866	-0.038	0.854	-0.0325	0.873
Older than resp. (D)			**0.5641**	0.032	**0.572**	0.031	**0.5649**	0.032
Higher schooling (D)			0.0590	0.799	0.079	0.738	0.0618	0.792
Lower schooling (D)			0.0010	0.997	-0.039	0.890	0.0015	0.996
Feedback								
Positive / Interesting subject					-0.034	0.833		
Privacy					0.337	0.216		
Interview not too long					0.060	0.917		
Easy to answer					0.191	0.787		
at least one of the 4 latter pos.							0.0106	0.937
Constant	-0.440	0.683	-0.9369	0.416	-0.844	0.466	-0.9371	0.416
Number of obs	243		241		241		241	
LR	26		29.98		29.98		29.98	
Prob larger chi2	0.0745		0.1190		0.1190		0.1499	
Pseudo R2	0.0866		0.1014		0.1014		0.1014	
Log likelihood	-137.1763		-132.81767		-132.81767		-132.81456	

Table 3.6: Nonresponse regressions: life insurance contracts

Life Insurances	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
HH income	**-0.0001**	0.057	**-0.0001**	0.063	**0.000**	0.054	**0.000**	0.062
HH income squared	**2.90E-09**	0.053	**2.96E-09**	0.053	**3.03E-09**	0.045	**2.96E-09**	0.053
HH income imputed (D)	0.238	0.303	0.272	0.248	0.285	0.228	0.276	0.244
Age	-0.040	0.366	-0.028	0.531	-0.027	0.551	-0.028	0.530
Age squared	0.000	0.365	0.000	0.426	0.000	0.432	0.000	0.425
Secondary school (D)	-0.034	0.862	-0.138	0.529	-0.113	0.610	-0.135	0.540
Graduation diploma (D)	0.010	0.971	-0.181	0.566	-0.120	0.709	-0.176	0.578
University degree (D)	-0.304	0.250	-0.518	0.112	-0.464	0.160	-0.511	0.120
East Germany (D)	-0.025	0.903	0.067	0.752	0.069	0.746	0.064	0.761
Worker (D)	0.074	0.734	0.006	0.979	0.023	0.919	0.006	0.978
Civil Servant (D)	-0.161	0.543	-0.201	0.458	-0.236	0.393	-0.199	0.465
Self-employed (D)	-0.069	0.781	-0.103	0.684	-0.097	0.703	-0.106	0.676
Retired (D)	-0.221	0.435	-0.292	0.307	-0.300	0.295	-0.293	0.305
Unemployed (D)	**-0.669**	0.067	**-0.810**	0.030	-0.837	0.026	-0.811	0.030
Small Community (D)	**-0.612**	0.013	**-0.651**	0.013	-0.654	0.013	-0.651	0.013
Version								
Interview version 3 (D)	**-0.400**	0.019	**-0.309**	0.087	**-0.304**	0.094	**-0.310**	0.086
Interview version 4 (D)	**-0.669**	0.000	**-0.679**	0.000	**-0.702**	0.000	**-0.679**	0.000
Interviewer								
Experienced > 4 years (D)			-0.172	0.255	-0.211	0.171	-0.173	0.252
Female (D)			-0.230	0.144	-0.225	0.159	-0.229	0.146
Older than resp. (D)			0.240	0.198	0.239	0.201	0.241	0.196
Higher schooling (D)			0.018	0.923	0.029	0.877	0.021	0.910
Lower schooling (D)			**0.357**	0.080	0.341	0.096	0.357	0.080
Feedback								
Positive / Interesting subject					-0.0340454	0.785		
Privacy					0.2400224	0.203		
Interview not too long					0.1096004	0.798		
Easy to answer					0.0428822	0.954		
at least one of the 4 latter pos.							0.0144056	0.883
Constant	1.489	0.129	1.067	0.305	1.073081	0.306	1.064985	0.307
Number of obs	368		364		364		364	
LR	37.24		42.77		44.55		42.79	
Prob larger chi2	0.0031		0.0050		0.0132		0.0073	
Pseudo R2	0.076		0.0883		0.092		0.0884	
Log likelihood	-226.31737		-220.66524		-219.77686		-220.6544	

Table 3.7: Nonresponse regressions: retirement savings contracts

Retirement Sav.	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
HH income	0.000	0.393	0.000	0.284	-0.0001953	0.209	-0.0001593	0.293
HH income squared	0.000	0.398	0.000	0.307	8.49E-09	2.57E-01	7.76E-09	0.301
HH income imputed (D)	0.801	0.106	0.782	0.121	0.660	0.223	0.650	0.221
Age	-0.012	0.888	-0.015	0.859	0.007	0.939	-0.025	0.771
Age squared	0.000	0.940	0.000	0.789	0.000	0.977	0.000	0.711
Secondary school (D)	0.526	0.191	0.400	0.390	0.409	0.390	0.386	0.406
Graduation diploma (D)	0.298	0.556	0.132	0.847	0.119	0.867	0.195	0.777
University degree (D)	0.518	0.312	0.369	0.576	0.337	0.615	0.357	0.589
East Germany (D)	**-1.052**	0.006	**-1.244**	0.005	**-1.397**	0.003	**-1.188**	0.008
Worker (D)	0.143	0.742	0.069	0.879	0.056	0.903	0.084	0.852
Civil Servant (D)	0.164	0.784	0.172	0.790	0.146	0.825	0.163	0.801
Self-employed (D)	0.009	0.981	-0.046	0.902	-0.102	0.801	-0.008	0.982
Retired (D)	0.356	0.598	0.235	0.732	0.169	0.811	0.325	0.645
Unemployed (D)	-0.377	0.538	-0.393	0.539	-0.477	0.477	-0.375	0.555
Small Community (D)	-0.405	0.317	-0.447	0.303	-0.576	0.198	-0.437	0.316
Version								
Interview version 3 (D)	**-0.865**	0.010	**-0.889**	0.018	**-0.844**	0.031	**-0.850**	0.025
Interview version 4 (D)	**-1.055**	0.002	**-1.132**	0.002	**-1.102**	0.004	**-1.149**	0.002
Interviewer								
Experienced > 4 years (D)	1.504	0.427	-0.200	0.505	-0.276	0.396	-0.204	0.500
Female (D)			0.076	0.812	0.087	0.796	0.057	0.860
Older than resp. (D)			0.518	0.144	0.499	0.187	0.504	0.158
Higher schooling (D)			-0.431	0.270	-0.567	0.164	-0.445	0.256
Lower schooling (D)			-0.141	0.746	-0.186	0.674	-0.155	0.722
Feedback								
Positive / Interesting subject					-0.007	0.980		
Privacy					0.122	0.739		
Interview not too long					dropped			
Easy to answer					dropped			
at least one of the 4 latter pos.							-0.160	0.430
Constant	1.504	0.427	1.514	0.462	1.387	0.512	1.720	0.406
Number of obs	115		114		110		114	
LR	24.16		27.73		27.96		28.35	
Prob larger chi2	0.1151		0.1850		0.2619		0.2028	
Pseudo R2	0.154		0.178		0.2619		0.182	
Log likelihood	-66.370471		-64.029633		-60.788319		-63.717127	

Table 3.8: Nonresponse regressions: bonds

Bonds	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
HH income	0.000	0.102	0.000	0.127	0.000	0.100	0.000	0.113
HH income squared	0.000	0.110	0.000	0.109	**2.48E-09**	0.081	2.33E-09	0.096
HH income imputed (D)	0.357	0.405	0.320	0.491	0.465	0.330	0.323	0.487
Age	-0.059	0.512	-0.072	0.447	-0.092	0.352	-0.076	0.427
Age squared	0.000	0.665	0.000	0.633	0.001	0.524	0.001	0.595
Secondary school (D)	-0.614	0.141	-0.721	0.127	-0.472	0.344	-0.647	0.178
Graduation diploma (D)	**-1.321**	0.019	**-1.921**	0.007	**-2.057**	0.006	**-1.892**	0.008
University degree (D)	-0.660	0.162	**-1.173**	0.094	-0.902	0.216	-1.087	0.127
East Germany (D)	-0.003	0.994	0.061	0.880	-0.096	0.826	0.006	0.988
Worker (D)	-0.592	0.206	-0.667	0.204	-0.629	0.248	-0.636	0.230
Civil Servant (D)	-0.282	0.701	-0.591	0.472	-0.288	0.729	-0.602	0.463
Self-employed (D)	-0.653	0.242	-0.928	0.135	**-1.645**	0.054	-1.030	0.106
Retired (D)	-0.413	0.369	-0.534	0.270	-0.579	0.240	-0.568	0.245
Unemployed (D)	0.200	0.774	0.224	0.782	0.309	0.720	0.319	0.700
Small Community (D)	**-1.713**	0.002	**-1.961**	0.002	**-1.893**	0.003	**-1.968**	0.002
Version								
Interview version 3 (D)	0.498	0.233	**0.832**	0.082	0.830	0.103	0.880	0.069
Interview version 4 (D)	-0.283	0.458	-0.151	0.720	-0.245	0.584	-0.155	0.715
Interviewer								
Experienced > 4 years (D)			-0.425	0.247	-0.416	0.279	-0.434	0.237
Female (D)			-0.013	0.971	-0.012	0.976	-0.003	0.993
Older than resp. (D)			-0.257	0.518	-0.217	0.606	-0.261	0.508
Higher schooling (D)			0.055	0.885	0.112	0.778	0.087	0.820
Lower schooling (D)			**0.911**	0.066	0.767	0.137	0.886	0.076
Feedback								
Positive / Interesting subject					0.374	0.190		
Privacy					-0.367	0.377		
Interview not too long					dropped			
Easy to answer					-0.275	0.814		
at least one of the 4 latter pos.							0.196	0.357
Constant	2.917	0.173	3.648	0.133	4.067	0.109	3.721	0.128
Number of obs	116		115		112		115	
LR	30.32		38.29		42.3300		39.14	
Prob larger chi2	0.0241		0.0170		0.0166		0.0192	
Pseudo R2	0.19		0.2518		0.2856		0.2574	
Log likelihood	-61.322773		-56.872127		-52.930077		-56.445915	

Table 3.9: Nonresponse regressions: stocks

Stocks	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
HH income	0.000	0.105	0.000	0.134	0.000	0.146	0.000	0.136
HH income squared	0.000	0.107	0.000	0.118	1.29E-09	0.126	1.30E-09	0.121
HH income imputed (D)	0.216	0.479	0.164	0.602	0.222	0.495	0.161	0.614
Age	-0.024	0.706	-0.030	0.652	-0.036	0.590	-0.030	0.651
Age squared	0.000	0.574	0.001	0.461	0.001	0.403	0.001	0.460
Secondary school (D)	0.054	0.858	0.125	0.697	0.153	0.637	0.125	0.698
Graduation diploma (D)	0.020	0.958	0.267	0.533	0.333	0.444	0.267	0.532
University degree (D)	-0.053	0.880	0.247	0.561	0.293	0.495	0.245	0.565
East Germany (D)	0.021	0.938	-0.093	0.749	-0.085	0.771	-0.092	0.753
Worker (D)	-0.347	0.288	-0.313	0.358	-0.316	0.355	-0.313	0.358
Civil Servant (D)	-0.125	0.686	-0.059	0.850	-0.056	0.858	-0.060	0.849
Self-employed (D)	-0.099	0.778	-0.090	0.801	-0.098	0.788	-0.089	0.804
Retired (D)	-0.661	0.116	-0.652	0.129	-0.671	0.121	-0.654	0.130
Unemployed (D)	-0.655	0.319	-0.513	0.433	-0.536	0.419	-0.514	0.433
Small Community (D)	-0.280	0.321	-0.239	0.423	-0.196	0.515	-0.240	0.423
Version								
Interview version 3 (D)	**-0.570**	0.011	**-0.496**	0.042	**-0.516**	0.036	**-0.496**	0.042
Interview version 4 (D)	**-0.673**	0.008	**-0.593**	0.025	**-0.588**	0.028	**-0.594**	0.025
Interviewer								
Experienced > 4 years (D)			-0.017	0.935	0.004	0.983	-0.016	0.941
Female (D)			0.074	0.710	0.072	0.723	0.074	0.715
Older than resp. (D)			0.378	0.126	0.403	0.106	0.379	0.126
Higher schooling (D)			0.330	0.197	0.364	0.160	0.330	0.198
Lower schooling (D)			-0.159	0.563	-0.191	0.493	-0.158	0.565
Feedback								
Positive / Interesting subject					0.010	0.959		
Privacy					-0.221	0.402		
Interview not too long					0.465	0.316		
Easy to answer					0.328	0.626		
at least one of the 4 latter pos.							-0.007	0.957
Constant	0.643	0.647	0.073	0.960	0.067	0.964	0.072	0.961
Number of obs	223		219		219		219	
LR	19.88		22.91		24.97		22.91	
Prob larger chi2	0.2806		0.4068		0.5205		0.4658	
Pseudo R2	0.072		0.0849		0.0926		0.085	
Log likelihood	-128.01658		-123.40718		-122.37628		-123.40573	

Table 3.10: Focal points in responses to the income question (multiples of 100)

Focal Points 100	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
Age	0.005	0.894	0.014	0.731	0.021	0.605	0.014	0.730
Age squared	0.000	0.796	0.000	0.906	0.000	0.815	0.000	0.904
Secondary school (D)	**0.423**	0.023	0.350	0.102	0.330	0.127	0.347	0.105
Graduation diploma (D)	0.410	0.142	0.241	0.456	0.255	0.441	0.236	0.467
University degree (D)	**0.439**	0.076	0.321	0.312	0.303	0.352	0.309	0.333
East Germany (D)	**-0.583**	0.001	**-0.534**	0.004	**-0.547**	0.004	**-0.527**	0.005
Worker (D)	-0.069	0.731	-0.069	0.734	-0.002	0.992	-0.067	0.741
Civil Servant (D)	-0.103	0.753	-0.140	0.673	-0.212	0.526	-0.138	0.678
Self-employed (D)	0.459	0.317	0.422	0.351	0.405	0.380	0.423	0.348
Retired (D)	**-0.492**	0.070	-0.431	0.115	-0.475	0.087	-0.433	0.114
Unemployed (D)	0.069	0.823	0.085	0.785	0.068	0.828	0.084	0.787
Small Community (D)	0.319	0.237	0.292	0.300	0.333	0.247	0.288	0.306
Version								
Interview version 3 (D)	**-0.312**	0.074	**-0.336**	0.067	**-0.319**	0.088	**-0.334**	0.070
Interview version 4 (D)	**-0.296**	0.094	-0.266	0.138	-0.254	0.167	-0.267	0.136
Interviewer								
Experienced > 4 years (D)			-0.047	0.754	-0.066	0.670	-0.040	0.792
Female (D)			0.155	0.313	0.153	0.329	0.152	0.323
Older than resp. (D)			-0.263	0.186	-0.261	0.195	-0.262	0.186
Higher schooling (D)			-0.088	0.630	-0.111	0.556	-0.096	0.602
Lower schooling (D)			0.166	0.488	0.203	0.412	0.163	0.495
Feedback								
Positive / Interesting subject					-0.027	0.824		
Privacy					0.177	0.361		
Interview not too long					-0.800	0.077		
Easy to answer					-1.435	0.039		
at least one of the 4 latter pos.							-0.042	0.669
Constant	0.984	0.255	1.152	0.207	1.022	0.265	1.153	0.207
Number of obs	684		676		676		676	
LR	28.31		32.1		42.83		32.28	
Prob larger chi2	0.0129		0.0304		0.0073		0.0404	
Pseudo R2	0.06590		0.07500		0.10010		0.07540	
Log likelihood	-200.63494		-197.93678		-192.57297		-197.84542	

Table 3.11: Focal points in responses to the income question (multiples of 500)

Focal Points 500	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
Age	0.031	0.286	0.038	0.189	0.038	0.186	0.038	0.191
Age squared	0.000	0.374	0.000	0.216	0.000	0.215	0.000	0.217
Secondary school (D)	0.205	0.117	0.157	0.300	0.151	0.321	0.155	0.308
Graduation diploma (D)	**0.443**	0.022	0.344	0.125	0.328	0.147	0.341	0.128
University degree (D)	**0.383**	0.029	0.281	0.205	0.263	0.238	0.273	0.220
East Germany (D)	**-0.701**	0.000	**-0.679**	0.000	**-0.677**	0.000	**-0.675**	0.000
Worker (D)	-0.083	0.570	-0.058	0.698	-0.044	0.769	-0.058	0.698
Civil Servant (D)	0.295	0.202	0.268	0.249	0.257	0.267	0.267	0.250
Self-employed (D)	**0.681**	0.007	**0.767**	0.003	**0.768**	0.003	**0.773**	0.003
Retired (D)	-0.190	0.321	-0.158	0.414	-0.165	0.394	-0.158	0.416
Unemployed (D)	**-0.405**	0.077	-0.322	0.170	-0.323	0.168	-0.321	0.171
Small Community (D)	-0.020	0.903	0.036	0.834	0.045	0.796	0.034	0.842
Version								
Interview version 3 (D)	**-0.228**	0.062	**-0.272**	0.035	**-0.260**	0.045	**-0.270**	0.036
Interview version 4 (D)	-0.188	0.128	-0.166	0.183	-0.156	0.212	-0.167	0.182
Interviewer								
Experienced > 4 years (D)			0.041	0.700	0.048	0.656	0.045	0.676
Female (D)			0.203	0.060	0.198	0.067	0.200	0.063
Older than resp. (D)			-0.142	0.300	-0.139	0.312	-0.141	0.301
Higher schooling (D)			-0.092	0.486	-0.100	0.450	-0.096	0.466
Lower schooling (D)			0.066	0.675	0.073	0.642	0.064	0.685
Feedback								
Positive / Interesting subject					-0.006	0.947		
Privacy					-0.027	0.851		
Interview not too long					-0.140	0.672		
Easy to answer					-0.445	0.294		
at least one of the 4 latter pos.							-0.032	0.657
Constant	-0.460	0.470	-0.511	0.453	-0.535	0.433	-0.507	0.457
Number of obs	684		676		676		676	
LR	68.93		76.28		77.81		76.48	
Prob larger chi2	0.0000		0.0000		0.0000		0.0000	
Pseudo R2	0.07370		0.08260		0.08420		0.08280	
Log likelihood	-433.16659		-423.8715		-423.10341		-423.77277	

Table 3.12: Focal points in responses to the income question (multiples of 1000)

Focal Points 1000	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
Age	0.001	0.985	0.009	0.759	0.010	0.747	0.009	0.760
Age squared	0.000	0.955	0.000	0.731	0.000	0.727	0.000	0.732
Secondary school (D)	0.188	0.160	0.074	0.635	0.074	0.638	0.073	0.637
Graduation diploma (D)	0.141	0.461	-0.052	0.815	-0.053	0.811	-0.052	0.813
University degree (D)	**0.396**	0.025	0.173	0.439	0.175	0.436	0.172	0.443
East Germany (D)	**-0.637**	0.000	**-0.568**	0.000	**-0.568**	0.000	**-0.567**	0.000
Worker (D)	-0.123	0.417	-0.112	0.466	-0.096	0.535	-0.112	0.466
Civil Servant (D)	-0.160	0.456	-0.196	0.362	-0.210	0.330	-0.196	0.362
Self-employed (D)	0.296	0.171	0.324	0.141	0.326	0.142	0.326	0.141
Retired (D)	-0.317	0.113	-0.271	0.180	-0.270	0.182	-0.271	0.180
Unemployed (D)	-0.314	0.205	-0.257	0.309	-0.275	0.279	-0.257	0.310
Small Community (D)	-0.145	0.371	-0.097	0.569	-0.082	0.631	-0.097	0.569
Version								
Interview version 3 (D)	-0.201	0.105	**-0.254**	0.051	**-0.236**	0.073	**-0.254**	0.051
Interview version 4 (D)	-0.096	0.438	-0.076	0.546	-0.067	0.593	-0.076	0.545
Interviewer								
Experienced > 4 years (D)			0.101	0.350	0.094	0.389	0.102	0.349
Female (D)			0.162	0.134	0.161	0.139	0.162	0.135
Older than resp. (D)			-0.167	0.219	-0.167	0.223	-0.167	0.219
Higher schooling (D)			-0.145	0.286	-0.155	0.257	-0.145	0.286
Lower schooling (D)			0.139	0.364	0.136	0.377	0.139	0.366
Feedback								
Positive / Interesting subject					0.012	0.891		
Privacy					0.077	0.594		
Interview not too long					0.136	0.678		
Easy to answer					-0.506	0.241		
at least one of the 4 latter pos.							-0.005	0.948
Constant	-0.309	0.635	-0.324	0.642	-0.340	0.627	-0.323	0.644
Number of obs	684		676		676		676	
LR	38.12		44.49		44.49		44.5	
Prob larger chi2	0.0005		0.0008		0.0027		0.0013	
Pseudo R2	0.04330		0.05110		0.05320		0.05110	
Log likelihood	-421.00946		-413.15682		-412.24823		-413.15466	

Table 3.13: Willingness to participate in future surveys

Would participate another time	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Respondent								
HH income	0.000	0.267	0.000	0.266	0.000	0.352	0.000	0.352
HH income squared	0.000	0.167	0.000	0.160	-6.21E-10	0.209	-6.11E-10	0.216
HH income imputed (D)	-0.202	0.181	-0.228	0.140	-0.166	0.287	-0.168	0.280
Age	0.000	0.996	0.005	0.863	0.007	0.817	0.009	0.759
Age squared	0.000	0.813	0.000	0.692	0.000	0.679	0.000	0.628
Secondary school (D)	0.064	0.619	-0.010	0.946	0.035	0.811	0.024	0.868
Graduation diploma (D)	0.144	0.430	0.003	0.989	0.092	0.667	0.063	0.765
University degree (D)	-0.134	0.433	-0.287	0.175	-0.172	0.424	-0.191	0.373
East Germany (D)	**0.246**	0.070	**0.285**	0.044	**0.240**	0.095	**0.241**	0.093
Worker (D)	0.174	0.228	0.132	0.369	0.148	0.320	0.136	0.356
Civil Servant (D)	0.113	0.593	0.082	0.703	0.040	0.853	0.074	0.731
Farmer (D)	-0.755	0.426	-0.924	0.341	-1.121	0.266	-1.076	0.288
Self-employed (D)	-0.124	0.529	-0.174	0.387	-0.216	0.292	-0.223	0.273
Retired (D)	**0.302**	0.097	0.273	0.139	0.274	0.139	0.271	0.143
Unemployed (D)	0.147	0.502	0.260	0.255	0.201	0.383	0.245	0.287
Small Community (D)	**0.493**	0.002	**0.559**	0.001	**0.601**	0.000	**0.577**	0.001
Version								
Interview version 3 (D)	**-0.407**	0.001	**-0.352**	0.004	**-0.344**	0.006	**-0.375**	0.003
Interview version 4 (D)	**-0.367**	0.002	**-0.350**	0.005	**-0.364**	0.004	**-0.356**	0.004
Interviewer								
Experienced > 4 years (D)			**-0.199**	0.051	**-0.260**	0.013	**-0.233**	0.024
Female (D)			-0.104	0.302	-0.090	0.380	-0.093	0.361
Older than resp. (D)			-0.056	0.656	-0.058	0.646	-0.051	0.685
Higher schooling (D)			-0.018	0.887	0.007	0.956	0.023	0.852
Lower schooling (D)			**0.259**	0.069	0.235	0.106	0.271	0.060
Feedback								
Positive / Interesting subject					**0.249**	0.004		
Privacy					**0.471**	0.000		
Interview not too long					**0.591**	0.080		
Easy to answer					-0.206	0.594		
at least one of the 4 latter pos.							**0.283**	0.000
Constant	0.490	0.419	0.553	0.394	0.556	0.396	0.463	0.478
Number of obs	773		764		764		764	
LR	38.65		49.69		73.85		66.4	
Prob larger chi2	0.0032		0.0010		0.0000		0.0000	
Pseudo R2	0.0381		0.0495		0.0736		0.0662	
Log likelihood	-487.86982		-476.60255		-464.51875		-468.24662	

Chapter 4

Extensions to *Saving in Germany*

4.1 Introduction

This chapter represents a seamless extension to the paper of Börsch-Supan and Essig (2003), "*Saving in Germany*". They exploited the 2001 *SAVE* study to shed more light on the the Germans' savings behavior.

The main findings of Börsch-Supan and Essig (2003), briefly, are: Overall, they find extraordinarily stable savings patterns. More than 40% of German households save regularly a fixed amount. About 25% of German households plan their savings and have a clearly defined savings target in mind. Most of German household saving is in the form of contractual saving, such as saving plans, whole life insurance and building society contracts; see Walliser and Winter (1999) for review. This makes the flow of saving rather unresponsive to economic fluctuations, such as income shocks. Most households prefer to cut consumption if ends do not meet. In particular the elderly do not like to use credit cards, and they eschew debt.

This chapter will add new insight to the results from Börsch-Supan and Essig (2003) with the new available *SAVE* data from the year 2003.[41] These new data are composed by two different subsamples. The first one is a panel sample from households being asked in both available waves, while the second one is an additional 'refreshment sample'. Methodological aspects of the *SAVE* data will not be discussed in this chapter; they are summarized in Chapter 2. In order to distinguish year and sample effects, all results will be presented pooled and separately for each subsample if possible.[42]

At the time field work for this survey has been done (June 2001 and June 2003) several political incidents had strong influence

The gap between the June 2001 and June 2003 - when the field work has been done - was a time of destructive political events worldwide, beginning with the attacks against the WTC, followed by the war against Afghanistan and ending with the third, still ongoing gulf war. These events brought up a time of insecurity all over the world.

In Germany, two additional events took place in that time span. First, the Euro was introduced as hard cash in Germany on January 1, 2002[43], and ever since there was a discussion of an additional inflation induced by that introduction. Even the German Minister of Finance, Hans Eichel,

[41] This chapter was written excluding the 2004 TPI subsample, cf. Table 2.1, since this subsample was not available until the end of August 2004. An exception is Section 4.4 where results for the new included risk variables are presented.

[42] Panel analysis is, by construction, only possible for the CAPI 2001 and 2003 subsamples.

[43] while it was already introduced as the official currency of the European Monetary Union on January 1, 1999

publicly called to boycott presumed inflationary industries using the Euro to 'raise prices and profits' in May 2002. While some researches demonstrate that this inflationary feeling is due to risen prices of only some products in the basket[44], see Ifo Institut (2002), others refer to psychological phenomena like biased assimilation and attitude polarization, see Frey *et al.* (2002). Second, even though the demographic changes of the aging population in Germany are well known for years, there has not been a large public debate on the pension and social security system in Germany. This definitely changed with the installation of the so-called 'Ruerup-Commission'[45] in 2003; recommendations of the commission were also published before the field work of the second wave. The main message of the commission's report are, very briefly, that private old-age provision and health-care will become more and more important. Even if the majority of the population might have implicitly doubted the long-run stability of the social security systems, the problem is now much more obvious.

So, indeed, the situation in 2003 is different to the one in 2001 in many respects. And for that reason reading the *SAVE* data again might be an instructive exercise.

The plan of this chapter is as follows. Section 4.2 discusses the present, the past and the future of the Germans' judgement of their well-being in the light of the possible worsening of the population's confidence. Section 4.3 reports findings for respondents' information level on and expectations for the public pension systems. Section 4.4 describes results for the recently included[46] risk questions. It tests the reliability of the experimental question design and links self-assessed risk to respondents' financial behavior. Section 4.5 summarizes the findings on these three topics.

Due to sampling issues of the *SAVE* dataset (see Chapter 2), three major questions will be discussed in this chapter. First, will there be difference in the strict panel structure of the data, i.e., what are, if any, the changes between 2001 and 2003 of households being exclusively observed in both years? This clearly offers the most interesting research fundament since values can compare households one by one. Second, do households differ in the two 2003 subsamples? If not, this would allow pooling of the two 2003 subsamples, as was the case in 2001, see Börsch-Supan and Essig (2003). Third, pooling subsamples in years, are there significant changes between 2001 and 2003? The abundance of tables and figures requires showing only significant ones. All other not shown but mentioned results will be available by the author on request.

All presented values underlying every table and figure in this chapter will be weighted across subsamples. See Chapter 2 for a discussion of the data's representativeness and the construction of used weights. The chosen weights from Chapter 2.4.2 refer to the dimensions subsamples, net household income, and age.

4.2 Well-being and living situation

This section shows values for the respondents' evaluation of the present, the past and the future. Börsch-Supan and Essig (2003) found that the majority of those surveyed consider that their situation

[44] which were often due to higher taxes, e.g. on alcohol and tobacco

[45] Cf. the report of the Kommission zur Nachhaltigkeit der Finanzierung der sozialen Sicherungssysteme (2003)

[46] In the *SAVE* 2004 wave, surveyed in June/July 2004.

is really positive, what may be a surprise because of the Germans' rather pessimistic reputation. The question asked here is whether these findings can be affirmed two years later.

4.2.1 The present

What variables would be influenced by the time change between 2001 and 2003? This is the topic of this subsection. As argued in the introduction, the exogenous shocks in 2001, 2002 and 2003 might have caused or induced a change of behavior. But has it also changed personal contentment? This will be analyzed looking at five variables: health, work, housing, income, and the general living situation.[47] It might well be the case that in general, respondents feel worse in general 2001 than in 2003 (living situation), but when directly asked to certain specific categories, one does not observe significant changes at all.

Panel results

Figure 4.1 compares households observed in both years,[48] while Figure 4.2 shows the distribution of the differences between the two years. Taking a scale from 0 to 10, which ranges from "completely discontent" to "very content", the majority of those surveyed classed themselves in the middle (the rating 5) or to the right of this (values 6 to 10). One can see no tremendous change in the pictures of the satisfaction items. The influence of deteriorated exogenous factors on the individual well-being is thus not directly reflected in the contentment variables. Running *Ordered probit* regressions[49] separate for all of these five variables on a constant and a dummy for the second wave one can see the results suggested by Figure 4.1 proved: only in the regression of the general living contentment variable shows a significant deterioration in 2003 can be observed. This seems to back the hypothesis from the previous paragraph that respondents, when asked specifically, will not translate these changed environmental conditions into an adaption of evaluating present contentment.[50]

Other results

Figures 4.3 and 4.4 show the analog results for the comparison of the two *SAVE* 2003 subsamples and between the pooled 2001 and 2003 subsamples. The figures leave us with two insights. First, there

[47] The exact wording was: "You will see a list of things that carry a certain significance in life. Please state according to a scale of 1 to 10 in how far you are content with the respective aspect. 0 means very discontented and 10 means very content. How content are you in regard to: (a) your health? (b) Your job? (c) Your housing? (d) Your income? (e) Your standard of living in general?"

[48] A so-called 'flag' filter (a variable controlling for respondent identity within each household) was introduced to eliminate observations when a respondent change in 2003 was suspected (e.g., the partner of the year 2001 is the respondent in 2003); this was done when the gender of the interviewee differed. One thereby loses 4 observations (2 in each wave since only a fully balanced panel is considered.). A different flag filter based on the interviewee's birth year proved to be problematic since it changed sometimes only by one year, leaving the possibility that respondents rather fib about their age (32 cases) than actually observing a real respondent change. This is much less less likely for the gender variable, reducing the type-I – error (incorrectly dropping observations) practically to zero.

[49] Not shown here, but available on request.

[50] Calculating t tests for the means of these variables, we see in every case lower means in 2003 than in 2001, only one of them (for housing) being significant. The problem of this procedure is that values of these variables cannot be interpreted as cardinal values.

are no profound differences between the two *SAVE* 2003 subsamples,[51] and second, the satisfaction level did not change in a large scope between 2001 and 2003. A figure comparable to Figure 4.2 is of course not feasible since this can only be done when we observe households more than once.

4.2.2 The past

Panel results

These assessments are supported by past experience. Table 4.1 focuses the attention on the income situation. Those surveyed were asked to compare their current income situation with the situation five years ago.[52] The table shows that more households are in a significantly better situation than are in a significantly worse situation. The improvement in income in the middle category is particularly noticeable: whereas 25.1% of households verify that their income situation is 'slightly better' than it was five years ago, only 13.8% ascertain that they are in a slightly worse position. Just under a third of those households surveyed remain at the same level.[53]

Table 4.1: Change of the income situation

	2001				
2003	Significantly better	Slightly better	About the same	Slightly worse	Sign. worse
Significantly better	**28**	16	9	3	0
	31.5%	11.4%	6.5%	5.9%	0.0%
Slightly better	28	**54**	26	7	7
	31.5%	**38.6%**	18.8%	13.7%	18.0%
About the same	17	46	**64**	15	8
	19.1%	32.9%	**46.4%**	29.4%	20.5%
Slightly worse	9	14	25	**14**	8
	10.1%	10.0%	18.1%	**27.5%**	20.5%
Significantly worse	7	10	14	12	**16**
	7.9%	7.1%	10.1%	23.5%	**41.0%**

Source: *SAVE* 2001 / 2003
Note: Row values add to 100% per column.

The direct comparison of the evaluation of the income situation is shown in Figure 4.5. Compared to the evaluation in 2001, the evaluation in 2003 is more evenly distributed: the share of respondents claiming that their income situation worsened in the last five years almost doubled in 2003 compared to 2001.

[51] Remember, all tables and figures present weighted values.

[52] The exact wording was: "Is your income situation, compared with five years ago (a) significantly better (b) slightly better (c) about the same (d) slightly worse (e) significantly worse?" This question was followed by: "During the last five years, did your income (a) Fluctuate significantly (b) Fluctuate slightly (c) Not fluctuate at all?"

[53] Zaller (1992) investigates variability over time for attitude changes within a four months period. The five item answers (where one of them was a "don't know"-item) were only equivalent in 48% of the cases; in Table 4.1, we find equivalent answers after two years in 38.5% of the cases (which are marked along the cross-diagonal in the table). On the one hand, keeping the Zaller's results in mind, one has to be careful interpreting the observed changes as real time effects; on the other hand, we observe a skewness towards a more pessimistic assessment, analog to the results in Section 4.2.1.

Table 4.2: Income fluctuations during the last 5 years

	2001		
2003	Significant fluctuation	Slight fluctuation	No fluctuation at all
Significant fluctuation	**35**	41	16
	40.7%	18.7%	11.3%
Slight fluctuation	37	**121**	62
	43.0%	**55.3%**	43.7%
No fluctuation at all	14	57	**64**
	16.3%	26.0%	**45.1%**

In households where any change at all occurred, it appears that these tended to be minimal, as can be seen in Table 4.2. Only around a quarter of those surveyed indicated significant fluctuations in income.

Other results

Figure 4.6 expands the results from the past section to all *SAVE* 2001 and 2003 subsamples. The results show that differences between the two years, if pooling the subsections by year, is not pronounced.

4.2.3 The future

From this view of the past it might be interesting to compare the results to the view of the future. Figure 4.9 shows the future expectations of the households surveyed. First, I look at the situation of the individual. The perspectives for their own health, the health of their partners and their own financial situation are largely seen in a positive light. This is all the more astonishing because the overall economic situation at the time the survey was conducted in spring 2001 and 2003 already showed considerable signs of a downturn - a fact that was very clear in the survey (chart on the left "Economic trends in Germany"). The objection that the households in *SAVE* are considerably wealthier than in the average population if no weighting is applied (see Tables 2.9 and 2.10) does not explain this observation, because the weighted income does not show any statistically significant differences over and above the 2000 and 2002 micro-census.

The insight gained at the end of this section is that, in contrast to current voices of foreboding in respect of the downturn in the economy (which tends to be cyclical and therefore perhaps not so very surprising for many people) but also the threat to economic growth resulting from the general pessimism following the terrorist attacks, it can be seen that people give their responses in a context of a healthy assessment of their own situation and an economic basis of this kind.

Panel results

In contrast to the results for the *present* contentment variables presented in Section 4.2.1, we see a rather pronounced shift for the *expectations* of the economic situation. Figures 4.7 and 4.8 show that especially for the economic development of Germany, households became more pessimistic in 2003;

the lowest value, zero, indicating a pure pessimistic expectation, was chosen five times as often in 2003 than in 2001 (from 3 to 15% of all respondents). In contrast, the own health situation as well as the health situation of the partner has not changed significantly. This supports the hypothesis from the introduction that indeed households adapted to the changed environmental conditions; Figures 4.7 and 4.8 also prove that there is no general gloomy prophecy but households rather know to differ between dimensions indeed affected by environmental changes between 2001 and 2003 of the kind we discussed before, and dimensions like health, which are rather exogenous due to these changes. Another interesting result which is also known from other studies and from psychological literature can be seen in Figures 4.7 and 4.8: households tend to think that they might not be beaten as hard as the average when conditions worsen.

Effects on financial behavior Given the reduced optimism concerning Germany's future economic development, one might interested whether this translates into changed financial behavior. I will examine this question by two topics. First, will the precautionary saving motive become more important, and second, do households hold less stocks.[54]

Table 4.3: Differences in stock ownership rates by expectations for Germany's economic development

	2001	2003	Difference
Non-Pessimists	29.7%	23.1%	6.6%
Pessimists	32.5%	22.4%	10.2%

Notes: Unweighted values. The *difference-in-difference* value of 3.6% is significant at the 1%-level

Table 4.3 shows differences for stock ownership rates stratified by expectations for Germany's economic development. Pessimists were defined as respondents whose difference between the expectations of 2001 and 2003 was below the median value of -2. The significant larger drop of ownership rates of the pessimists seems to support the findings of Börsch-Supan and Essig (2003) that intentions, or in this case, assessments, are supported by households' actions.

Saving motives The *SAVE* questionnaire asks for judging the importance of nine savings motives on a scale from 0 to 10, where 0 means 'completely unimportant' and 10 'very important', respectively.[55] Figure 4.12 shows the panel results for nine savings motives as the share of respondents judging the motives being important (defined as judgements ranging from 8 to 10).

The analogue presentation to the previous paragraph can thus be done only implicitly since the answers to the savings motives are measured on an ordinal scale (cf. Footnote 50), and so a *difference-in-difference* approach is presented implicitly in Figure 4.12. We see slightly positive differences for the pessimistic households.

To identify the effect of deteriorated confidence in Germany's economic development, I apply the following methodology. Again, pessimists were defined as respondents whose difference between the

[54] The small sample size for this panel analysis is a serious issue and limits some analysis to bivariate correlation tables. This leaves room for more detailed analysis at a later stage of the *SAVE* panel when more observations will be available.

[55] Chapter 5 will analyze the importance of each of the nine saving motives in detail. In this chapter, the point of interest is whether we can observe changes in judging the savings motives *over time*.

expectations of 2001 and 2003 was below the median value of -2. This adaption of expectations will be called the 'treatment', i.e. whether respondents reacted to the possibly changed economic environment. On the other hand, the 'control group' are respondents who did not adjust their expectations in that manner. Since the CAPI subsample was observed in 2001 and 2003, it is possible to control for effects which were experienced by both groups ('time effect'). When a control group is present, the design in psychology has been called the untreated control group design with pretest and posttest; see Meyer (1995). This can be identified with the *difference-in-difference* technique.

Table 4.4 shows results from *difference-in-difference* ordered probit regressions of savings motives on dummy variables for the year 2003 and for pessimistic households[56].

Again, the outcome, $savingmotiveX_i$, was modeled by the following equation

$$savingmotiveX_i = \alpha + \beta Time_i + \gamma Pessimist_i + \delta(Time_i \cdot Pessimist_i) + \varepsilon_i \tag{4.1}$$

where

α = constant term
β = time trend common to control and treatment group
γ = treatment group specific effect (to account for average permanent differences between treatment and control)
δ = true effect of treatment
ε_i = random, unobserved 'error' term which contains all determinants of $savingmotiveX_i$ which the model omits

Of course, this is not some *difference-in-difference* model in the classical sense, since in this model here, respondents self-selected themselves into the treatment group. In other words, I take account for the fact that the time trend is not common to all respondents, dividing them into two different groups. The objections to such an approach are summarized in Meyer (1995) who gives an excellent review for the validity threats to models of this kind.

Table 4.4 reveals three very interesting results. First, respondents becoming more pessimistic significantly rate the precautionary savings motive more importantly (the highly significant positive *treatment* dummy takes account for that fact). Second, savings for non-durable consumption like travelling is negatively affected by pessimistic evaluations. Third, the old-age provision savings motive does not seem to be affected by more pessimistic economic evaluations (all parameters being insignificant in that regression), not even a pure time effect seems to be relevant. Coming back to the reasoning in Section 4.1, one can state the following: events having taken place between 2001 and 2003 affect the evaluation of Germany's economic development in a negative way. Respondents whose drop in this rating is more negative than the median also rate the precautionary savings motive as well as supporting their offspring more important. The ongoing discussion about the German pension system

[56] Pessimism, again, is defined as the difference between the expectations of 2001 and 2003 being below the median value of -2.

Table 4.4: Ordered probit estimates: savings motives by time and pessimism

	Purchase of real estate		Precautions for unexpected events		Paying off debts		Old-age provision		Travelling	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Time effect	-0.099	0.227	**-0.202**	0.009	-0.045	0.572	-0.010	0.902	-0.090	0.234
Group effect	-0.397	0.212	**-0.485**	0.092	0.071	0.812	-0.402	0.168	**0.530**	0.071
Treatment effect	0.304	0.365	**0.762**	0.012	-0.100	0.750	0.368	0.232	**-0.621**	0.045
	Major purchases		Education / support of children /grandchildren		Bequests for children / grandchildren		Taking advantage of state subsidies			
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$		
Time effect	**-0.284**	0.000	-0.035	0.653	-0.079	0.326	-0.023	0.769		
Group effect	-0.051	0.859	**-0.661**	0.037	-0.358	0.268	-0.381	0.214		
Treatment effect	0.143	0.638	**0.546**	0.100	0.253	0.455	0.251	0.437		

Notes: Cut-off parameters are not reported. Significant parameters are highlighted.

and the rising importance of private contributions for old-age did not seem to affect respondents' evaluation of this savings motive.

4.3 Pensions

The German private pension system ("public retirement insurance" / 'Gesetzliche Rentenversicherung' *[GRV]*) which covers about 85% of the German workforce[57] has undergone continuous reforms concerning retirement *payments*[58] and *pathways* into retirement[59]. Börsch-Supan and Wilke (2003) provide a description of the recent history of institutional changes of the German pension system.

This section will not go into details of the pension system's institutional background. Rather, it checks whether the broad discussion on the future of the German pension system which was intensified by the 2001 *"Riester"*-Reform and reinforced by proposals of the 2003 *Rürup*-commission induced changes of expectations concerning the generosity of the pension system with respect to the replacement rate (direct financial effect) and the pension entry age (indirect financial effect).

Equation 4.2 presents a simplified stylized model for the effects of the two mentioned parameters:

$$PW_A = \sum_{t=A}^{T} \frac{1}{1+r_t} \alpha_t Y_{A-1} \tag{4.2}$$

where

α_t = pension replacement rate in year t
Y_{A-1} = income in last year before pension entry
T = time of death
A = pension entry age

[57] Cf. Berkel and Börsch-Supan (2003)

[58] Cf. Börsch-Supan *et al.* (2004c)

[59] Cf. Berkel and Börsch-Supan (2003)

PW = discounted pension wealth
r_t = discount factor in year t

While T and r_t are exogenous for the insured persons, this is only true for the pension entry age A and the replacement rate α, when there exists no compulsory retirement age. In an actuarial fair pension system, α and A are directly connected: the higher the pension entry age, the higher will be the replacement rate to compensate for (1) the shorter time span of pension receipts until T and (2) for higher accrued claims during the working life (since the expected pension wealth should equal the accrued pension contributions at time A). Actuarial fairness would then determine the replacement rate α_t to ascertain this equality. Departures from actuarially fair replacement rates generate incentive effects, mostly to early retirement due to more 'generous' pension reforms. Many studies have looked at these incentives effects by microeconometric modelling (Börsch-Supan (1992), Schmidt (1995), Siddiqui (1997), Börsch-Supan (2000), Börsch-Supan (2001), and Börsch-Supan *et al.* (2004c). See also Börsch-Supan *et al.* (2004a) for a review on the Germans' knowledge on the pension system and willingness to accept pension reforms.

This section is structured as follows. First, I will analyze whether the mentioned discussion about the future of the German security systems induced changes on respondents' information level by exploring refusals and 'don't know' answers to pension parameter questions. As a second step, I examine whether respondents changed and their expectations concerning parameters of Equation 4.2, namely, the pension entry age, A, and the replacement rate, α. Third, I will inspect the importance of private pensions before and after the pension reform discussion.

4.3.1 Information level on pension parameters

Table 4.14 shows probit regression results for refusals for the expected pension entry age. Several results are worth noting. First, as expected, refusal is a function of age. But interestingly, the age polynomial's *minimum* is about 37 - 38 years.[60] One would rather suspect that the closer one is away from the event of pension entry, the higher the probability that the actual year is known.[61] Interpreting refusals for this question as insecureness for knowledge, this seems surprising. Second, the year dummy for 2003 is highly negative significant in the respondents' regression. This can be interpreted that respondents in the year 2003 cared more about retirement than in 2001. Third, there seems to be evidence that refusals and therefore knowledge about own and partner's pension entry is correlated: respondents giving answers for themselves have a higher probability of answering also to the partner's question.

[60] This means that after that age, nonresponse will rise with age

[61] Or, that the variance of knowledge is a positive function of time to retirement: the closer the retirement age, the better the knowledge and the smaller the variance.

4.3.2 Pension age

In the *SAVE* questionnaire, several pension related questions were asked. As mentioned in the introduction, one of the interesting research topics of savings behavior is the relevance of households' information and expectations about their pension level. As mentioned earlier in Section 4.2, one would expect some reaction from households to changed economic and political environments. Apart from relatively soft questions like expectations, we are interested whether there is also some significant reaction concerning more specific questions. One major debate of the year 2003 was the adaption of Germany's pension system to the challenge of demographic changes. By that time, one of the most discussed questions was the increase of the normal pension entry age from 65 to 67 years. Did respondents react to that discussion, and if so, how large was the reaction? Table 4.5 shows average pension entry ages, observed ones (from the *GSOEP* and the *VDR*) and expected values from *SAVE* .

Table 4.5: Estimated and expected pension entry age

	SOEP 1999[a]		*VDR* 2001[b]	*SAVE* 2001[c]		*SAVE* 2003[c]	
	Men	Women	Men and Women	Men	Women	Men	Women
Mean	59.7	60.7	60.2	62.9	61.9	63.5	63.0
Median	/	/	/	64	62	65	65
Standard Deviation	/	/	/	3.22	2.89	4.01	3.72
Standard Error	/	/	/	0.10	0.16	0.14	0.13
N	/	/	/	941	340	852	830

[a] *Source*: Berkel and Börsch-Supan (2003)

[b] *Source*: Own calculations / data were supported by Christina B. Wilke.

[c] *Notes*: Weighted values, pooled over subsamples. T-Tests for differences between 2001 and 2003 *SAVE* values are significant.

The expected pension entry age is higher in *SAVE* than average observed values in the *GSOEP* and in data we received from the *Verband der Rentenversicherungsträger (VDR)*. This divergence can have several sources. One might be that respondents are more confident concerning their future health or employment situation, or they are more pessimistic with regards to future pension regulations.

Table 4.13 shows regression results from the expected pension entry age on a set of household and respondent characteristics. The strong positive significant effect of the time gap from 2001 to 2003 seems to support the formerly stated hypothesis that expectations have differed between 2001 and 2003. Also, while in the second specification dummies for both 2003 subsamples are significantly positive, they are not significantly different from each other. The time effect is strong. A very good and likely explanation is the exacerbated pension system discussion in 2002 and 2003.

Given these findings, a compelling question arises: Do changed expectations concerning the future pension entry age alter the savings motive for old-age provision? The reason for that might be that respondents whose expected pension entry age difference between 2003 and 2001 is above the median might be more pessimistic concerning the future generosity of the pension system; their loss of faith in the pension system might therefore correspond to a higher emphasize for private old-age provisions.[62]

[62] Raising the pension entry age actually *stabilizes* the pension system and should induce *more* faith in the public pension system. The argument here works the other way around: if there is a necessity to cut benefits (in any form,

The results for this procedure are insignificant. In that sense, the evaluation of the importance of savings motives is not affected by the changes in expectations concerning pension entry age. Adding first differences in expectations for the future development of Germany's economic situation does not bring qualitative changes.

4.3.3 Pension levels

How well are people informed about pension levels? This is as much a serious question as the estimated pension entry age.

Table 4.15 lists probit estimation results for 'don't knows' and refusals for the replacement rates of pension payments in relation to the last income before retirement, while Table 4.16 lists results for the estimation for partner's values.[63]

Coefficients for age are highly significant[64] as one would expect them to be: the closer a person comes to retirement, the higher the probability that he will inform himself about future pension payments. Astonishingly, the coefficients for females are insignificant when looking at the information level for the partner; the female dummy in the refusal regression is even significantly negative. Since typically males are the financial officers[65] in a household and typically have more complicated working histories which in turn make replacement rate calculation more complicated[66], one would suspect their knowledge on their partner's replacement rates to be higher. This is not supported by the data.

Pension level and pension age

As Equation 4.2 implies, pension entry age and the replacement rate are positively interrelated, since the replacement, if adapted actuarially fair, is the degree of freedom parameter to equal retirement contributions and pension payments *when the pension entry age is chosen.*

4.3.4 Private pensions

This Section compares the ownership rates for private pensions between 2001 and 2003.

Table 4.8 lists descriptive findings of private pension ownership rates for 2001 and 2003. At first glance, the different rates for item nonresponse seem disturbing, but since the questioning was slightly more complex for 2003, this can easily explain the different rates. The ownership rate presented for 2003 therefore can be viewed as a lower bound for ownership, since the rate was calculated *assuming the missing values to be zero.* Just calculating the ratio for the values present would raise it from

like reducing the pension payment period by altering the entry age), one might begin to wonder how weak the system already is.

[63] The exact wording of the question was: "What percentage of your last wage/salary do you estimate your pension and your partner's pension will be?"

[64] with a negative influence on the 'don't know' and a positive influence on the 'refusal' probability

[65] E.g. in about 75% of respondents claiming to be the household's financial officer are male in the *GSOEP*

[66] Opposingly, one could argue that women's pensions levels are more difficult to calculate (effect of chid-raising allowance, incomplete work histories, low wages etc.), but the author favors the previous argument

Table 4.6: Expectations for pension entry age and replacement rates by age

	Men		Women	
Age	Pension entry age	Replacement rate	Pension entry age	Replacement rate
under 30	63.6	50.9%	63.7	48.0%
30-39	64.0	57.6%	62.7	52.2%
40-49	63.0	59.4%	62.8	59.4%
50-59	62.8	62.5%	62.9	61.5%
60 and older	64.1	59.4%	63.2	61.6%

Notes: Means for weighted values. Random sample 2003 only. Self-employed respondents were excluded.

Table 4.7: Expectations for pension entry age and replacement rates by age

	2001				2003			
	Men		Women		Men		Women	
Age	Age	Repl. rate	Age	Repl. rate	Age	Repl. rate	Age	Repl. rate
under 30	62.7	59.7%	61.8	52.9%	63.7	52.4%	63.7	48.4%
30-39	63.2	61.3%	61.6	60.9%	64.0	58.8%	62.8	52.7%
40-49	62.5	61.8%	61.8	62.0%	63.1	60.7%	62.7	59.6%
50-59	62.5	65.9%	61.4	61.1%	63.0	63.8%	62.8	61.4%
60 and older	64.2	63.9%	63.2	61.1%	64.1	59.8%	63.1	61.6%

Notes: Means for weighted values. Self-employed respondents were excluded.

20.4% to 22.7%, and assuming all missing values were indeed respondents *not willing to admit or not knowing* the ownership, the ratio would rise to 30.6% as an upper bound.

Table 4.17 shows probit regression results for private pension ownership. Again, the time effect is strongly significant: the ownership rates rose between 2001 and 2003.

Table 4.18 depicts probit estimation results for the three second pillar pension types the *SAVE* questionnaire asks for. Income and age are significant for occupational pension schemes and other private old-age provisions, while income is insignificant for 'Riester'-type pensions - an interesting

Table 4.8: Private pension ownership rates

	2001[a]		2003	
	N	Percent	N	Percent
Private Pensions	297	16.53	529	20.4
Occupational pensions		./.	285	10.99
'Riester' pensions		./.	122	4.7
Other private pensions		./.	219	8.45
Refusals	15	0.83	265	10.22
Number of observations	1797[b]		2593[b]	

Note: Unweighted values.

[a] Difference in question design: in 2003, private pensions were asked for each category separately, whereas in 2001, the question directly addressed to all forms of private pensions.

[b] Observations differ from full sample (1829 obs. in 2001 and 2667 in 2003) because of not returned drop-off parts.

insight. Further, as already was pointed out by Börsch-Supan and Essig (2003), there is evidence for a relationship between households' self-assessment and actual behavior: the old-age provision motive is highly significant in any of the three regressions - the higher the rated importance of savings for old-age, the higher the probability to own any of the second pillar or two third pillar pension types. For the 'Riester'-type pensions, the motive to take advantage of state/tax subsidies is also highly significant.

Private pensions: panel results

Table 4.9 shows private pension ownership rates for households asked in 2001 and 2003.

Table 4.9: Private pension ownership rates: panel results

	2001			
2003	No	Yes	Refusals	Total
No	**279**	21	2	302
	71.4%	31.8%	50.0%	65.5%
Yes	91	**39**	1	131
	23.3%	**59.1%**	25.0%	28.4%
Refusals	21	6	**1**	28
	5.4%	9.1%	**25.0%**	6.1%
Total	391	66	4	461

Private pension ownership more than doubled. This result has to be taken with caution, as already mentioned in Section 4.3.4, since the questionnaire has changed between both periods of time.[67]

A further analysis determining the contribution rates to private pensions can be found in Chapter 7.3.2.

4.4 Risk behavior

I now turn to topic remote to ones discussed in the two previous sections. In this section, I will present first empirical results for risk measures, which I will link to financial behavior.

The *SAVE* survey contains risk variables on five different domains: health, job, financial, leisure and sports, and when driving. Additionally, the *SAVE* 2005 sample will include questions for engaging on four different activities, followed by a judgement on how risky these activities are evaluated. In very recently received data[68], these questions were already included to pretest their feasibility in the field. Moreover, the new risk engagement and evaluation questions allow a consistency test for the existing five risk domain questions.

[67] And since in 2003, the survey design was more detailed than in 2001, it is possible that this leads to overestimation, or, in 2001, to underestimation of private pensions rates, and thereby, pension wealth.

[68] The 2001 TPI subsample was surveyed again in June/July 2004; data were received by August 20, 2004. This panel panel sample contains 487 observations, which corresponds to a relatively high net percentage of 74% of the original 660 observations from 2001. I will *not* refer to any other survey results except the risk questions in this chapter.

Table 4.10: Risk assessment question

"To what extent do the following statements apply to you. Please answer on a scale of 0 to 10, where 0 means 'does not apply at all' and 10 means 'applies very well'. I do not mind taking risks with respect to :
1. My own health
2. My career
3. In money matters
4. With respect to leisure time and sports
5. When driving"

This section proceeds in the following way. First, I briefly summarize the sets of risk questions contained in *SAVE* and show descriptive and regression results for the newly added questions. Second, I describe the influence of risk measures on the financial behavior, mainly, on the portfolio composition.

4.4.1 Risk variables in *SAVE*

Risk assessment questions

In order to allow to link savings behavior to risk, the *SAVE* survey contains risk questions common to all subsamples. The exact format of that question is shown in Table 4.10.

Weber *et al.* (2002) distinguish between five content domains:[69] Financial, health/safety, recreational, ethical and social decisions. Applying this scheme to the risk questions in *SAVE* , we see that the five direct risk questions distinguish between 3 content domains: (a) Health /Safety (1. + 5.), (b) financial (3.) and (c) recreational (4). Considering the results from Weber *et al.* (2002), the correlation coefficients between the risk domains would be positive, but significantly lower than 1.[70] Table 4.11 shows that the highest correlation coefficient is 0.61 - well below 1.

Table 4.11: Correlations between risk domains

	Health	Career	Money matters	Leisure and sports
Health	1			
Career	0.4852	1		
Money matters	0.4277	0.5211	1	
Leisure and sports	0.5061	0.5264	0.5419	1
When driving	0.4418	0.4583	0.5001	0.6072

Notes: N = 4476. Weighted values by income and age.
Source: All *SAVE* subsamples

The distributions of responses for the risk variables are shown in Figure 4.14, separately for each subsample. Two things are worth noting. First, the value 0 is used particularly often in respect of

[69] This bases on the work of Slovic and Lichtenstein (1986).

[70] Simply spoken, this accounts for the fact that the preference for one risk domain is *relatively* offset by a less pronounced risk preference in a different domain.

Table 4.12: Risky activities

"Four different activities are described below. What is the probability that you would do one of the following activities:
1. Walking around alone in an unknown neighborhood
2. Investing 5% of your yearly income in one stock
3. Putting a day's income on a bet
4. Go climbing"

health, driving and investments, whereas 5 is fairly frequent for career and leisure/sport. Zero values are also given more frequently by the respondents in the Random Route subsample. Second, changes over time within subsamples, e.g. between the two times the TPI sample was surveyed, are very low, thus indicating a high stability of the answers.

In Tables 4.19 and 4.20, the results of Figure 4.14 are analyzed multi-dimensionally. In the 2001 and 2004 TPI samples, respondents have a higher probability associating higher values to the risk-assessment than the respondents in the 2001 CAPI set for all risk categories. Freelancers and self-employed respondents describe themselves as showing more risky behavior in the fields of career and money matters, which coincidences with the widespread outlook for these two employment categories. Income variables are jointly significant for the risk categories money matters, leisure and sports, and when driving, but not for career and health categories. A higher schooling degree also raises the probability for higher risk values in four of the five categories. Still, one puzzling result is that civil servants do not describe themselves as being less risky in career matters. Women have an overall lower probability of associating high values to risk assessment. This is also in line with other findings from sociopsychological literature, cf. Slovic (1997). Age is jointly significant for all risk domains; age decreases the probability for high risk values. This means that older respondents, though not having to care for children, are more risk averse.

Risky activities

It has been found in the psychological sciences that risk-taking is influenced by both, the characteristics of the decision-maker and by the situation (cf. Bromiley and Curley (1992)). The *SAVE* TPI 2004 subsample thus was enriched by additional questions: engaging on four different activities. The exact wording is shown in Table 4.12. The following two subsections present the results for these variables using the data from the *SAVE* TPI 2004 subsample; a more thorough analysis will be possible after the reception of the 2005 Random Route subsample data.

This question is followed by the request to judge the riskiness of each of the listed activities on a scale from 0 to 10. The descriptive results are shown in Figure 4.15. It can be well seen that on average, risk evaluation and the probability for exercising the activity are negatively correlated, thus indicating that the questions were very well understood. Furthermore, and quite amazing, response rates are close to 100%.

Interaction of both sets of risk variables

The next step of analysis is to analyze the link between the two blocks of risk questions. This is done by regressing the probabilities of exercising each of the four activities on a set of respondent and household characteristics and the risk judgement of the corresponding activity, as well as the five abstract risk variables using an Ordered Probit regression model. The results, which are shown in Table 4.21, prove two things. First, the findings of Figure 4.15 that risk judgement and exercise are positively correlated also holds in a more differentiated framework: for each regression, the risk judgement associated with the risky activity is strongly negative significant. Second, for each activity, the risk assessment question corresponding to each risky activity is positively significant. A high self-assessed risk for leisure and sports raises the probability of walking alone through an unknown neighborhood, which is also true for money matters risk and investing 5 % of household income in one stock and for betting a day's income; risking more concerning own's health raises the probability of climbing. These results show that obviously the abstractness of the risk questions does not prevent respondents giving thoughtful and coherent answers. The next section will thus analyze the connection of risk assessment and financial behavior.

Lottery questions

Before leaving this field, it might be interesting to take a look at a third set of risk attitude questions. In 2004 and 2005, *SAVE* contains so-called lottery questions. Lottery, or gamble questions, can be used to elicit individual value functions[71], see Farquhar (1984), Abdellaoui (2000), Pennings and Smidts (2003), or Wakker and Deneffe (1996). While lottery questions are a relatively common tool in the economic literature to assess risk parameters, it is less common in the psychological literature in which risk is assessed by directly asking for certain situations/events. The risk associated with these situations then has to be evaluated by the respondent; see Weber *et al.* (2002).

In the *SAVE* lottery questions, respondents are given a set of hypothetical choices where each choice is between a certain cash value and the toss of a coin when head pays a different value from tail. In the first three lottery questions, the value of a head draw increases from 1,700 to 2,000 and 2,300 €, where tail pays nothing and the certain cash value is 1,000 €. In the second set of lottery questions, the certain cash value is 0, head brings a loss of 100 €, and the value of a tail draw changes from -150 to -200 and to -250 €.[72] Figure 4.16 show the share of respondents choosing the toss of a coin when the expected value of the lottery is increased. Again, as with the risk activity questions, we see that values are consistent, since the fraction of respondents choosing the risky alternative rises when the incentives to do so are increased. The risky alternative choice is also more frequent for the second set of lotteries when payments are on a lower level, even if the the risky choice of the last question of the first set offers the highest gain over the certainty equivalence. Overall, the choice probabilities for the

[71] Value functions assign subjective values to a stated (objective) value, cf. Schunk and Betsch (2004).

[72] This means that in the first set, a large number of repetitions would present the choice between 1,000 and 850, then 1000 and 1000, then 1000 and 1150 €, while in the second set, the values are 0 and 25, 0 and 50, and 0 and 75 €. Risk neutral agents would simply compare these values and choose the alternative with the higher payments, while risk avers agents need to be reimbursed for the taken risks, and risk seeking agents are willing to pay for the chance of a higher gain.

coin toss are relatively low given that, at least in the second set and partly in the first set, the choice of the uncertain alternative offers the higher expected value.

Table 4.22 shows Probit and Ordered Probit regression results for the sets lottery questions. For the Ordered Probit case, the depend variable was constructed in the following way. It takes on the value 0 if no lottery was chosen, 1 if only the lottery with the highest payoff, 2 if additionally the lottery with the second highest and 3 if the lottery with the lowest expected payoff was chosen. In most cases, respondents choosing lower expected value payoff lotteries also chose higher payoff lotteries. This can be seen by comparing the results from the Ordered Probit regression results with the Binary Probit regressions[73], which are very similar. A very interesting effect, again, is that the shown risk and lottery questions are reasonably connected. In Table 4.22, I included for all four regression types three risk variables dealing with monetary risk. This is the direct self-assessed general monetary risk questions, the probabilities of investing 5% of net income in a stock and for betting a day's income. The risky activity most directly connected to lotteries is betting. It is positively significant in all four regressions which means that respondents more willing to place money on a bet are also more likely to accept the lottery instead of the security equivalent. In the second set of lotteries, monetary risk taking also rises the probability of choosing the lottery. Nearly every other personal and household characteristic does not provide additional explanatory power.

Of course, one might tend to argue that TPI sampled households are highly selective since they are used to those kinds of questions. This is not the entire truth. TPI comprises topics from market research, consumption and living habits, to political and personal attitudes to the knowledge about newest technologies. Additionally, in the case of the lottery and risk questions, the two set of questions were rather remote in the questionnaire; between the two blocks lie about 25 minutes of interview time,[74] where the most complex interview part is in between. Even more surprising is the strong consistency of the data.

Summarizing the results, the risk data are a promising base to work with. Skeptics might check the additional data when risk questions will be incorporated into the Random Route and CAPI panel subsamples in 2005.

In the following subsection, I will investigate the interrelation between risk questions and portfolio compositions, controlling for personal and household characteristics.

4.4.2 Risk and portfolio composition

This section checks whether self-assessed risk affects financial behavior; I investigate this by using portfolio shares of financial wealth asset categories.

[73] The dependent variable takes on the value 1 if any of the lotteries was chosen instead of the risk-free payoff and 0, otherweise.

[74] Compared with the interview time from CAPI interviews; TPI interview mode is P&P thus offering no time tracking.

Portfolios in *SAVE*

In this paragraph, I will very quickly review the portfolio choice for *SAVE* 2001 and 2003. This work has been done prior to the arrival of the *SAVE* 2004 subsample.

Figure 4.13 shows ownership rates for the six financial wealth categories. Ownership rates are comparable between subsample but for one item: stock ownership is more than twice as high for the TPI 2001 subsample compared to the RR 2003 sample, considering weighted values. The values are altogether higher than compared to the *EVS* for stock ownership.

Figure 4.17 shows the portfolio shares of each financial category to total financial wealth.[75]. Again, subsample effects are strong: The relative value of saving accounts is almost twice as high in the RR 2003 sample as in the TPI 2001 sample. This cannot be attributed to a low number of cases: the number of observations for *conditional* stock asset shares in the TPI sample is 160, as much as for the CAPI 2001 sample, and more than for the RR 2003 sample. Dividing Figure 4.17 into three age classes adds the insight that especially younger households in the TPI sample are more prone for stock holding, cf. Figure 4.18.

Regression results

I proceed in the following way. Since the additional risk and lottery questions are not available but for the TPI 2004 subsample, I limit the analysis to the set of risk questions present in all subsamples. This means that there is only one question connected to monetary behavior, which will be included in the analysis: self-assessed risk for money matters. Respondent were also given a question to what extent seven statements apply to them. I picked one of them, optimism[76], which in several other analyses seemed to characterize households very well.

Since not every household holds all financial wealth categories, an OLS approach would not account for this selectivity. I therefore apply the Heckman selection model. Finding the correct exclusion restrictions is not an easy task since most variables are very likely to affect both, asset ownership and the amount invested. Specification problems quickly lead to non-convergence of the ML-estimation of the selection model. I therefore apply a very slim specification for the regression stage which lead to convergence for all six asset categories.[77]

The regression results are presented in Tables 4.23 - 4.25.[78] For the two most extreme asset categories[79], saving accounts and stocks/fonds, the effects are very pronounced. Risk averse households have a higher probability for saving account ownership, and their portfolio share rises with risk aversion, as the negative significant coefficient for monetary risk indicates. For stocks and fonds, the

[75] As numbers might be difficult to retrieve, results are also depicted in Figure 4.26

[76] One of the seven questions was also whether respondents were pessimistic. This statement is basically redundant, but was added as a plausibility check. The consistency of the response to the two contrary statements optimism and pessimism is very high. The response scale ranges from 0 to 10 for the statements, and, hence, the sum of the evaluations for optimism and pessimism should equal 10. 79% of respondents had this sum ranging from 8 to 12.

[77] This was a second order polynomial for net income and age, and the risk variable.

[78] I included the TPI 2004 panel for these regressions.

[79] in the sense of return risk

opposite applies. The probability of ownership rises with risk preference, and also the relative amount of financial wealth invested in stocks.

For the other four wealth categories, the effects are a little less clear. Bond ownership rises with risk preference, but not the portfolio share, which also applies to private old-age provisions.[80] As for whole life insurances, risk preferences increases the ownership probability, but affects the portfolio share negatively.

4.5 Conclusions

This chapter adds three branches to the investigation of Börsch-Supan and Essig (2003). The first one is whether a deterioration in households' contentment and confidence can be observed on a general level. While the judgement of the present and the past is comparable between 2001 and 2003, the evaluations for future expectations became more pessimistic. This can be linked to a re-evaluation of the savings motive for old-age, which became more important.

The second checks whether this deterioration translates into one specific aspect for households' saving behavior: the old-age provision, namely the public pension system. I observe two things: the nonresponse and the 'don't know' level have fallen between 2001 and 2003. The political discussion thus might have been the catalyst to induce more individual concern for the respective pension situation. Second, a slight rise of the expected pension entry age and a drop in the expected pension level can be observed. This does not translate into an adaption of the evaluation of savings motives.

Risk assessment theoretically seems a promising procedure when determining household financial behavior. The empirical implementation in *SAVE* was done including 5 risk questions covering 3 risk domains. Using the 'monetay risk' variable, I find significantly higher probabilities of stock ownership and portfolio shares. The opposite is true when looking at saving accounts.

In addition to these results, I analyze the experimental risk questions included in *SAVE* 2004. In contrast to possibly sceptical objections to this kind of risk questions, these results prove two things. First, nonresponse is neglectably low. Considering the unusual character of these questions even for interview-skilled TPI respondents, this is a very positive result. Second, responses are sensible for both sets of risk questions, lotteries and risk assessments / risky activities. Risk assessment can be linked to financial behavior, as the regression results indicate. This is a promising base for further analysis concerning risk and portfolio choice. The *SAVE* 2005 samples will show whether the response quality level can be maintained for less interview skilled respondents.

[80] In 2001, this category was asked less differentiated than in 2003 and 2004; I therefore summarize the three questions for private pensions from 2003 and 2004 in one category.

4.A Tables

Table 4.13: Expected pension entry age

	Coef.	$P > t$	Coef.	$P > t$
Net HH income / 10000	0.395	0.207	0.396	0.172
(Net HH income) squared / 10000	0.000	0.363	0.000	0.309
Age	**-0.074**	0.054	**-0.074**	0.054
Age squared	**0.001**	0.047	**0.001**	0.046
Secondary school (D)	**0.474**	0.006	**0.487**	0.005
Graduation diploma (D)	**0.707**	0.002	**0.710**	0.002
University degree (D)	**1.253**	0.000	**1.256**	0.000
Partner (D)	**-0.365**	0.021	**-0.394**	0.014
East Germany (D)	0.027	0.881	0.026	0.887
Female (D)	**-0.724**	0.000	**-0.726**	0.000
Job: blue collar (D)	0.330	0.115	0.324	0.122
Job: civil servant (D)	**-1.190**	0.000	**-1.188**	0.000
Job: freelancer (D)	**1.271**	0.006	**1.285**	0.006
Job: self employed (D)	0.171	0.564	0.171	0.563
Work parttime (D)	0.084	0.758	0.068	0.801
Work little (D)	**0.797**	0.009	**0.773**	0.011
Work not (D)	0.250	0.322	0.233	0.357
Unemployed (D)	-0.350	0.223	-0.345	0.229
Village (D)	-0.140	0.545	-0.119	0.608
Year 2003 (D)	**0.814**	0.000		
Sample: CAPI 2001 (D)			-0.276	0.201
Sample: Panel 2003 (D)			**0.579**	0.031
Sample: RR 2003 (D)			**0.647**	0.003
Constant	64.064	0.000	64.276	0.000
Number of obs	2799		2799	
Prob > F	0.000		0.000	
R-squared	0.048		0.048	
Adj. R-squared	0.041		0.041	
Root MSE	3.496		3.496	

Table 4.14: Probit regression results for nonresponse to the pension entry age

	Respondent				Partner					
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net HH income/10'	-0.163	0.415	-0.101	0.593	-0.294	0.177	-0.292	0.180	**-0.620**	0.016
Net HH income/10' sq.	0.053	0.160	0.041	0.240	0.064	0.125	0.064	0.124	**0.112**	0.020
Age/10	**-0.443**	0.010	**-0.469**	0.006	**-0.820**	0.000	**-0.827**	0.000	**-0.912**	0.000
Age/10 squared	**0.060**	0.001	**0.063**	0.001	**0.107**	0.000	**0.108**	0.000	**0.121**	0.000
Secondary school (D)	-0.148	0.098	**-0.178**	0.044	-0.050	0.568	-0.053	0.545	-0.014	0.889
Graduation diploma (D)	**-0.265**	0.033	**-0.262**	0.033	**-0.275**	0.036	**-0.280**	0.032	**-0.276**	0.073
University degree (D)	-0.093	0.441	-0.092	0.440	-0.041	0.720	-0.039	0.737	0.006	0.961
Partner (D)	0.116	0.194	**0.166**	0.056						
East Germany (D)	-0.104	0.308	-0.110	0.278	**-0.259**	0.015	-0.262	0.014	**-0.294**	0.019
Female (D)	0.107	0.227	0.129	0.141	**-0.359**	0.000	**-0.350**	0.000	**-0.779**	0.000
Job: blue collar (D)	0.038	0.732	0.058	0.598	-0.110	0.336	-0.107	0.349	-0.170	0.205
Job: civil servant (D)	-0.117	0.478	-0.121	0.459	-0.207	0.193	-0.208	0.192	-0.222	0.237
Job: freelancer (D)	-0.031	0.902	-0.059	0.815	**-0.676**	0.072	**-0.678**	0.071	-0.694	0.101
Job: self-employed (D)					-0.260	0.132	-0.256	0.137	**-0.533**	0.017
Work parttime (D)	-0.059	0.691	-0.034	0.817	**-0.358**	0.065	**-0.351**	0.069	-0.218	0.362
Work little (D)	0.160	0.291	0.198	0.187	-0.022	0.890	-0.016	0.921	0.152	0.398
Work not (D)	**0.473**	0.000	**0.503**	0.000	**0.317**	0.004	**0.321**	0.003	**0.398**	0.001
Unemployed (D)	**-0.362**	0.011	-0.362	0.011	**-0.301**	0.053	**-0.292**	0.059	**-0.400**	0.024
Village (D)	-0.110	0.410	-0.152	0.249	0.036	0.761	0.033	0.785	0.030	0.826
Sample: CAPI 2001 (D)	**-0.290**	0.004			-0.004	0.966	-0.072	0.379		
Sample: Panel 2003 (D)	**-1.047**	0.000			-0.145	0.292	0.625	0.153		
Sample: RR 2003 (D)	**-0.524**	0.000			-0.053	0.606				
Year 2003 (D)			**-0.424**	0.000	0.613	0.167			0.096	0.312
Refusal respondent(D)									**2.434**	0.000
Constant	-0.327	0.403	-0.494	0.193	0.613	0.167	0.625	0.153	0.477	0.348
Number of obs	2879		2879		2250		2250		2250	
LR chi2	129.1		110.73		274.81		274.24		704.71	
$Prob > F$	0.000		0.000		0.000		0.000		0.000	
Pseudo-R squared	0.076		0.066		0.141		0.141		0.362	
Log likelihood	-780.487		-789.672		-835.104		-835.387		-620.155	

Notes: Observations were excluded if partner was retired or self-employed.

Table 4.15: Probit regression results for 'don't know' and refusals for own pension replacement rates

	Don't Know				Refusal				Don't Know or Refusal			
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net HH income/10'	**-0.207**	0.038	**-0.213**	0.032	0.008	0.968	0.126	0.538	**-0.210**	0.033	**-0.203**	0.040
Net HH income/10' sq.	**0.020**	0.157	**0.021**	0.138	-0.015	0.734	-0.040	0.406	0.017	0.227	0.016	0.259
Age/10	**0.247**	0.102	**0.245**	0.103	**-0.575**	0.000	**-0.563**	0.000	**-0.244**	0.079	**-0.241**	0.082
Age/10 squared	**-0.049**	0.006	**-0.048**	0.007	**0.080**	0.000	**0.078**	0.000	0.018	0.273	**0.017**	0.285
Secondary school (D)	-0.093	0.136	-0.088	0.156	-0.096	0.246	-0.119	0.147	**-0.137**	0.029	**-0.142**	0.023
Graduation diploma (D)	**-0.168**	0.048	**-0.167**	0.049	-0.134	0.234	-0.121	0.275	**-0.225**	0.008	**-0.224**	0.008
University degree (D)	-0.135	0.118	-0.136	0.115	**-0.204**	0.084	**-0.211**	0.073	**-0.220**	0.010	**-0.219**	0.011
Partner (D)	0.030	0.599	0.015	0.790	**-0.228**	0.004	**-0.148**	0.053	-0.060	0.296	-0.042	0.463
East Germany (D)	0.028	0.675	0.029	0.668	0.059	0.511	0.057	0.524	0.046	0.487	0.045	0.497
Female (D)	**0.134**	0.031	**0.131**	0.033	**0.135**	0.097	**0.144**	0.076	0.194	0.002	**0.196**	0.001
Job: blue collar (D)	-0.119	0.110	-0.122	0.100	0.085	0.410	0.103	0.313	-0.084	0.256	-0.080	0.275
Job: civil servant (D)	**-0.600**	0.000	**-0.599**	0.000	-0.050	0.746	-0.060	0.694	**-0.543**	0.000	-0.544	0.000
Job: freelancer (D)	0.193	0.242	0.197	0.232	0.110	0.627	0.099	0.655	0.255	0.115	0.250	0.121
Job: self employed (D)												
Work parttime (D)	**0.164**	0.092	**0.160**	0.099	0.011	0.936	0.020	0.886	**0.163**	0.093	**0.167**	0.086
Work little (D)	0.166	0.117	0.158	0.136	**0.326**	0.016	**0.356**	0.008	**0.326**	0.003	**0.334**	0.002
Work not (D)	**0.252**	0.005	**0.248**	0.006	**0.495**	0.000	**0.507**	0.000	**0.538**	0.000	**0.543**	0.000
Unemployed (D)	0.020	0.845	0.020	0.843	**-0.248**	0.051	**-0.241**	0.057	-0.155	0.143	-0.155	0.143
Village (D)	-0.057	0.511	-0.047	0.584	-0.066	0.586	-0.116	0.334	-0.081	0.346	-0.093	0.276
Sample: CAPI 2001 (D)	0.113	0.149			**-0.524**	0.000			**-0.136**	0.077		
Sample: Panel 2003 (D)	0.034	0.730			**-0.617**	0.000			**-0.244**	0.012		
Sample: RR 2003 (D)	-0.003	0.964			**-0.417**	0.000			**-0.209**	0.006		
Year 2003 (D)			-0.068	0.210			**-0.133**	0.076			**-0.130**	0.016
Constant	-0.335	0.291	-0.249	0.422	0.079	0.828	-0.326	0.354	**0.978**	0.001	**0.872**	0.003
Number of obs	2835		2835		2835		2835		2835		2835	
LR chi2	181.16		178.88		126.78		97.84		242.76		239.47	
$Prob > F$	0.000		0.000		0.000		0.000		0.000		0.000	
Pseudo-R squared	0.047		0.046		0.065		0.050		0.062		0.061	
Log likelihood	-1835.521		-1836.659		-917.202		-931.671		-1839.564		-1841.207	

Notes: Observations were excluded if respondent was retired or self-employed.

Table 4.16: Probit regression results for 'don't know' and refusals for partner's pension replacement rates

	Don't Know				Refusal				Don't Know or Refusal			
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net HH income/10'	**-0.573**	0.002	**-0.569**	0.002	-0.118	0.609	-0.054	0.813	**-0.625**	0.000	**-0.597**	0.001
Net HH income/10' sq.	**0.086**	0.018	**0.085**	0.019	0.031	0.484	0.018	0.692	**0.100**	0.004	**0.094**	0.007
Age/10	-0.057	0.753	-0.054	0.767	**-0.472**	0.024	**-0.451**	0.031	**-0.394**	0.034	**-0.385**	0.038
Age/10 sq.	0.000	0.981	-0.001	0.966	0.066	0.005	**0.063**	0.007	**0.045**	0.032	**0.044**	0.036
Secondary school (D)	-0.031	0.680	-0.032	0.668	-0.095	0.321	-0.109	0.251	-0.087	0.254	-0.098	0.198
Graduation diploma (D)	-0.106	0.310	-0.105	0.314	-0.101	0.447	-0.095	0.472	-0.164	0.120	-0.165	0.117
University degree (D)	0.023	0.820	0.023	0.822	**-0.361**	0.009	**-0.354**	0.010	-0.159	0.121	-0.156	0.128
East Germany (D)	-0.016	0.848	-0.016	0.848	-0.070	0.530	-0.072	0.518	-0.048	0.569	-0.051	0.545
Female (D)	0.009	0.912	0.008	0.915	-0.097	0.354	-0.078	0.449	-0.041	0.612	-0.033	0.677
Job: blue collar (D)	-0.110	0.209	-0.109	0.213	0.025	0.830	0.038	0.740	-0.107	0.228	-0.098	0.268
Job: civil servant (D)	-0.130	0.259	-0.129	0.261	0.220	0.128	0.211	0.143	-0.029	0.801	-0.031	0.789
Job: freelancer (D)	-0.258	0.225	-0.260	0.220	-0.031	0.915	-0.088	0.761	-0.255	0.219	-0.275	0.184
Work parttime (D)	0.103	0.381	0.103	0.378	-0.090	0.587	-0.073	0.655	0.065	0.587	0.078	0.514
Work little (D)	-0.132	0.311	-0.130	0.320	0.053	0.755	0.088	0.604	-0.114	0.390	-0.091	0.492
Work not (D)	-0.095	0.388	-0.095	0.384	0.230	0.104	**0.245**	0.081	0.030	0.791	0.040	0.718
Unemployed (D)	0.104	0.426	0.103	0.429	0.010	0.949	0.013	0.934	0.119	0.379	0.120	0.376
Village (D)	**-0.263**	0.011	**-0.267**	0.009	0.131	0.310	0.089	0.486	**-0.187**	0.069	-0.211	0.038
Sample: CAPI 2001 (D)	-0.044	0.612			**-0.384**	0.001			**-0.240**	0.006		
Sample: Panel 2003 (D)	-0.106	0.358			**-0.350**	0.021			**-0.286**	0.015		
Sample: RR 2003 (D)	-0.119	0.174			**-0.218**	0.044			**-0.233**	0.008		
Year 2003 (D)			-0.092	0.188			-0.053	0.552			-0.113	0.106
Constant	0.599	0.136	0.568	0.154	-0.091	0.851	-0.348	0.465	**1.664**	0.000	**1.503**	0.000
Number of obs	1922		1922		1922		1922		1922		1922	
LR chi2	38.67		38.39		51.69		39.02		53.49		45.75	
Prob > F	0.007		0.003		0.000		0.003		0.000		0.000	
Pseudo-R squard	0.0145		0.014		0.036		0.027		0.021		0.018	
Log likelihood =	-1312.436		-1312.573		-695.871		-702.207		-1255.600		-1259.466	

Notes: Observations were excluded if partner was retired or self-employed.

Table 4.17: Probit regression results for private pension ownership

	Coeff.	$P > z$	Coeff.	$P > z$
Net HH income/10'	**0.310**	0.050	**0.288**	0.069
Net HH income/10' sq.	**-0.082**	0.033	**-0.075**	0.050
Age/10	**0.533**	0.000	**0.515**	0.000
Age/10 squared	**-0.066**	0.000	**-0.064**	0.000
Secondary school (D)	**0.123**	0.051	**0.126**	0.046
Graduation diploma (D)	**0.285**	0.001	**0.277**	0.001
University degree (D)	**0.252**	0.002	**0.256**	0.002
Partner (D)	**0.161**	0.007	**0.142**	0.018
East Germany (D)	-0.016	0.804	-0.015	0.820
Female (D)	-0.030	0.606	-0.020	0.729
Job: blue collar (D)	-0.114	0.162	-0.115	0.157
Job: civil servant (D)	**-0.389**	0.001	**-0.386**	0.001
Job: freelancer (D)	**0.293**	0.080	**0.293**	0.081
Job: self-employed (D)	**0.503**	0.000	**0.513**	0.000
Work parttime (D)	-0.039	0.706	-0.038	0.709
Work little (D)	**-0.415**	0.000	**-0.419**	0.000
Work not (D)	**-0.446**	0.000	**-0.443**	0.000
Unemployed (D)	0.010	0.926	0.012	0.916
Village (D)	**0.200**	0.013	**0.213**	0.008
Sample: CAPI 2001 (D)			**-0.231**	0.003
Sample: Panel 2003 (D)			**0.426**	0.000
Sample: RR 2003 (D)			**0.324**	0.000
Year 2003 (D)	**0.492**	0.000		
Constant	**-2.203**	0.000	**-2.008**	0.000
Number of obs	3894		3894	
LR chi2	398.84		409.23	
$Prob > F$	0.000		0.000	
Pseudo-R squared	0.102		0.105	
Log likelihood	-1748.039		-1742.840	

Notes: A t-test showed that the hypothesis of equality of the two 2003 subsample dummies could not be rejected.

Table 4.18: Probit regression results for ownership of different private pension types

	Occ. pensions[a]		'Riester' pensions		Other old-age provisions	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net HH income/10'	**0.540**	0.023	0.235	0.765	**1.092**	0.083
Net HH income/10' sq.	**-0.115**	0.043	-0.399	0.505	**-0.852**	0.070
Age/10	**0.549**	0.001	**0.763**	0.008	**0.437**	0.032
Age/10 squared	**-0.050**	0.003	**-0.099**	0.002	**-0.062**	0.005
Secondary school (D)	0.121	0.224	**0.254**	0.049	-0.092	0.398
Graduation diploma (D)	**0.394**	0.005	0.282	0.102	0.115	0.413
University degree (D)	**0.312**	0.016	0.192	0.285	**0.252**	0.060
Partner (D)	0.135	0.147	**0.235**	0.073	-0.057	0.577
East Germany (D)	**-0.558**	0.000	**0.241**	0.057	0.109	0.309
Female (D)	-0.020	0.825	0.013	0.914	-0.085	0.366
Job: blue collar (D)	-0.161	0.221	0.228	0.189	-0.176	0.235
Job: civil servant (D)			0.057	0.813	-0.086	0.649
Job: freelancer (D)					**0.574**	0.031
Job: self-employed (D)			0.165	0.468	**0.747**	0.000
Work parttime (D)	**-0.285**	0.050	0.196	0.304	-0.031	0.846
Work little (D)	**-0.557**	0.001	-0.054	0.815	**-0.372**	0.064
Work not (D)	**-0.694**	0.000	0.059	0.729	**-0.266**	0.051
Unemployed (D)	**-0.526**	0.023	0.237	0.206	-0.107	0.551
Village (D)	**0.267**	0.047	0.089	0.598	0.079	0.576
Saving Reason: State subsidies			**0.053**	0.000		
Saving Reason: Old-age provision	**0.027**	0.038	**0.046**	0.019	**0.072**	0.000
Constant	**-2.563**	0.000	**-3.982**	0.000	**-2.473**	0.000
Number of obs	2002		2158		2209	
LR chi2	190.65		116.21		209.74	
$Prob > F$	0.000		0.000		0.000	
Pseudo-R squared	0.129		0.131		0.152	
Log likelihood	-646.715		-385.181		-582.202	

[a] Self-employed, civil servants and freelancers were excluded from the regression.

Table 4.19: Ordered probit regressions: self-assessed risk: health, career, money matters

	Health		Career		Money matters	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net income / 10,000	0.106	0.730	0.124	0.687	**0.736**	0.016
(Net inc./ 10,000) sq.	-0.105	0.730	-0.067	0.823	-0.365	0.215
Age / 10	**0.726**	0.000	**0.704**	0.000	**0.583**	0.000
Age / 10 sq.	**-0.053**	0.000	**-0.055**	0.000	**-0.075**	0.000
Secondary school (D)	0.004	0.919	0.061	0.147	0.067	0.107
Graduation diploma (D)	0.018	0.757	**0.153**	0.009	**0.273**	0.000
University degree (D)	-0.040	0.449	**0.143**	0.009	**0.214**	0.000
Kids (D)	-0.008	0.892	0.005	0.930	**-0.145**	0.013
Kids living in same house (D)	0.050	0.285	0.042	0.388	0.061	0.204
Job: blue collar (D)	0.046	0.408	-0.087	0.121	**-0.127**	0.026
Job: civil servant (D)	0.072	0.357	-0.018	0.813	0.113	0.145
Job: freelancer (D)	0.158	0.215	**0.552**	0.000	**0.273**	0.031
Job: self-employed (D)	0.101	0.204	**0.374**	0.000	**0.287**	0.000
Retired (D)	-0.023	0.779	**-0.368**	0.000	-0.054	0.516
Work parttime (D)	0.014	0.853	0.116	0.110	-0.030	0.688
Work little (D)	0.120	0.115	0.116	0.137	0.010	0.900
Work not (D)	-0.068	0.319	-0.264	0.000	**-0.180**	0.010
Unemployed (D)	-0.014	0.866	0.213	0.012	-0.074	0.382
Unemp.> 1 month (D)	**0.146**	0.002	**0.128**	0.007	0.014	0.772
Unemp.> 6 months (D)	-0.047	0.385	-0.035	0.536	0.033	0.544
Partner (D)	**-0.024**	0.661	0.010	0.865	0.055	0.321
Separated or divorced (D)	**0.217**	0.001	0.140	0.035	0.162	0.013
Widowed (D)	**0.158**	0.019	0.117	0.091	**0.145**	0.036
Female (D)	**-0.258**	0.000	**-0.304**	0.000	**-0.407**	0.000
Sample: TPI 2001 (D)	**0.230**	0.000	**0.337**	0.000	**0.190**	0.001
Sample: CAPI 2003 (D)	0.009	0.891	-0.046	0.496	0.030	0.651
Sample: RR 2003 (D)	0.055	0.264	**0.149**	0.004	0.027	0.592
Sample: TPI 2004 (D)	**0.266**	0.000	**0.361**	0.000	**0.119**	0.066
Number of obs	4516		4176		4483	
LR chi2(29)	290.51		727.77		636.18	
Prob > chi2	0.000		0.000		0.000	
Pseudo R2	0.0161		0.0416		0.0373	
Log likelihood	-8896.788		-8379.676		-8214.139	
F-Test income variables	0.940		0.863		0.003	
F-Test age variables	0.001		0.000		0.001	

Notes: Self-assessed risk is coded from 0 to 10, where 0 means 'does not apply at all' and 10 'does fully apply'. The wording of the questions is shown in Table 4.10

Table 4.20: Ordered probit regressions: self-assessed risk: leisure and sports, when driving

	Leisure and sports		When driving	
	Coeff.	$P > z$	Coeff.	$P > z$
Net income / 10,000	**0.552**	0.071	**0.637**	0.039
(Net inc./ 10,000) sq.	-0.274	0.363	-0.420	0.160
Age / 10	**-0.201**	0.012	**-0.171**	0.046
Age / 10 sq.	0.001	0.932	0.003	0.721
Secondary school (D)	0.030	0.461	**0.121**	0.004
Graduation diploma (D)	**0.175**	0.002	**0.187**	0.001
University degree (D)	**0.138**	0.008	**0.140**	0.010
Kids (D)	0.001	0.984	-0.079	0.185
Kids living in same house (D)	-0.012	0.797	0.022	0.653
Job: blue collar (D)	0.016	0.776	-0.017	0.762
Job: civil servant (D)	0.060	0.435	-0.105	0.179
Job: freelancer (D)	**0.201**	0.112	0.054	0.674
Job: self-employed (D)	0.084	0.287	0.109	0.169
Retired (D)	0.013	0.870	0.036	0.671
Work parttime (D)	0.061	0.398	-0.059	0.432
Work little (D)	0.124	0.100	-0.102	0.198
Work not (D)	**-0.156**	0.021	**-0.274**	0.000
Unemployed (D)	-0.002	0.982	0.007	0.932
Unemp.> 1 month (D)	**0.096**	0.037	0.066	0.164
Unemp.> 6 months (D)	**-0.120**	0.025	-0.045	0.419
Partner (D)	**-0.023**	0.677	0.010	0.866
Separated or divorced (D)	0.093	0.150	0.191	0.004
Widowed (D)	**0.130**	0.054	0.004	0.949
Female (D)	**-0.385**	0.000	**-0.446**	0.000
Sample: TPI 2001 (D)	**0.318**	0.000	**0.094**	0.098
Sample: CAPI 2003 (D)	0.062	0.335	-0.069	0.299
Sample: RR 2003 (D)	**0.113**	0.021	0.042	0.401
Sample: TPI 2004 (D)	**0.357**	0.000	**0.235**	0.000
Number of obs	4477		4364	
LR chi2(29)	786.13		603.74	
Prob > chi2	0.000		0.000	
Pseudo R2	0.0416		0.0359	
Log likelihood	-9062.239		-8110.990	
F-Test income variables	0.030		0.056	
F-Test age variables	0.000		0.000	

Notes: Self-assessed risk is coded from 0 to 10, where 0 means 'does not apply at all' and 10 'does fully apply'. The wording of the questions is shown in Table 4.10

Table 4.21: Ordered probit regressions: exercising risky activities

	Walk alone at night		5% in stocks		Bet a day's inc.		Climbing	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net income / 10,000	-0.852	0.269	**1.893**	0.015	**2.668**	0.002	**2.786**	0.002
(Net inc./ 10,000) sq.	0.323	0.402	**-0.726**	0.063	**-1.117**	0.016	**-1.206**	0.006
Age / 10	0.790	0.101	0.072	0.884	-0.218	0.679	0.296	0.606
Age / 10 sq.	**-0.090**	0.077	-0.023	0.660	0.007	0.907	-0.069	0.266
Secondary school (D)	**0.276**	0.037	0.292	0.031	**0.269**	0.068	**0.287**	0.066
Graduation diploma (D)	**0.303**	0.091	0.343	0.062	0.148	0.453	0.021	0.919
University degree (D)	0.142	0.409	0.259	0.141	-0.125	0.515	-0.048	0.817
Kids (D)	**0.368**	0.076	0.140	0.515	0.244	0.288	0.144	0.554
Kids living in same house (D)	-0.050	0.741	0.134	0.391	-0.101	0.549	-0.027	0.882
Job: blue collar (D)	0.000	0.999	0.023	0.900	-0.049	0.802	0.302	0.124
Job: civil servant (D)	0.067	0.771	-0.060	0.801	0.035	0.887	0.136	0.613
Job: freelancer (D)	**0.894**	0.091	-0.350	0.504	-0.617	0.282	-0.641	0.295
Job: self-employed (D)	-0.222	0.346	0.118	0.613	**0.567**	0.017	-0.195	0.471
Retired (D)	-0.102	0.733	-0.222	0.451	0.280	0.378	0.070	0.843
Work parttime (D)	0.058	0.812	-0.182	0.445	**0.485**	0.054	0.244	0.362
Work little (D)	-0.140	0.603	0.316	0.217	0.071	0.801	-0.143	0.637
Work not (D)	0.074	0.784	0.404	0.118	-0.001	0.996	0.278	0.374
Unemployed (D)	-0.016	0.958	-0.504	0.102	0.095	0.782	-0.235	0.531
Unemp.> 1 month (D)	-0.007	0.962	0.111	0.448	0.036	0.819	-0.035	0.835
Unemp.> 6 months (D)	-0.027	0.884	0.026	0.892	-0.146	0.480	0.023	0.920
Partner (D)	0.077	0.775	**0.635**	0.025	-0.205	0.487	0.361	0.268
Separated or divorced (D)	-0.009	0.973	0.446	0.102	-0.260	0.382	-0.314	0.306
Widowed (D)	0.243	0.392	0.568	0.050	0.348	0.242	0.159	0.619
Female (D)	**-0.861**	0.000	-0.180	0.248	-0.140	0.418	-0.105	0.573
Risk judgement of act.	**-0.247**	0.000	**-0.171**	0.000	**-0.179**	0.000	**-0.225**	0.000
Risk: health	0.013	0.596	-0.033	0.179	0.003	0.922	0.020	0.476
Risk: career	0.016	0.512	0.008	0.759	0.017	0.541	0.016	0.567
Risk: money matters	-0.014	0.610	**0.147**	0.000	**0.115**	0.000	**-0.075**	0.022
Risk: leisure and sports	**0.048**	0.064	0.011	0.686	0.022	0.459	**0.176**	0.000
Risk: driving	0.016	0.520	-0.013	0.608	**0.054**	0.052	0.031	0.308
Number of obs	434		432		433		433	
LR chi2(29)	300.79		178.08		191.32		229.23	
Prob > chi2	0.000		0.000		0.000		0.000	
Pseudo R2	0.1524		0.102		0.1334		0.1649	
Log likelihood	-836.62743		-783.98228		-621.18513		-580.62361	

Notes: Self-assessed risk is coded from 0 to 10, where 0 means 'does not apply at all' and 10 'does fully apply'. The wording of the questions is shown in Table 4.12

Table 4.22: Ordered probit and probit regressions: lottery questions

	Lottery set 1				Lottery set 2			
	Ordered probit reg.		Probit reg.		Ordered probit reg.		Probit reg.	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net income / 10,000	0.056	0.981	0.065	0.980	0.078	0.971	0.110	0.960
(Net inc./ 10,000) sq.	0.504	0.855	0.197	0.950	2.161	0.441	1.790	0.547
Age / 10	0.751	0.435	0.666	0.509	-0.405	0.504	-0.365	0.555
Age / 10 sq.	-0.101	0.345	-0.092	0.409	0.052	0.419	0.051	0.435
Secondary school (D)	0.159	0.503	0.239	0.339	0.052	0.766	0.101	0.562
Graduation diploma (D)	-0.104	0.741	0.032	0.921	**-0.441**	0.070	-0.390	0.115
University degree (D)	0.156	0.615	0.351	0.278	0.045	0.834	0.116	0.604
Kids (D)	-0.185	0.577	-0.348	0.322	-0.355	0.188	-0.370	0.177
Kids living in same house (D)	-0.199	0.475	-0.099	0.738	0.010	0.959	-0.051	0.794
Job: blue collar (D)	0.184	0.507	0.304	0.305	**0.446**	0.049	**0.445**	0.059
Job: civil servant (D)	-0.506	0.297	-0.697	0.197	-0.251	0.446	-0.286	0.406
Job: freelancer (D)	0.227	0.782	0.439	0.614	-0.210	0.788	-0.114	0.870
Job: self-employed (D)	0.083	0.802	0.115	0.744	0.187	0.508	0.169	0.573
Retired (D)	-0.384	0.509	-0.311	0.605	-0.482	0.199	-0.618	0.107
Work parttime (D)	0.119	0.768	0.012	0.976	-0.427	0.221	-0.368	0.281
Work little (D)	0.093	0.844	0.129	0.788	**0.647**	0.041	**0.821**	0.013
Work not (D)	0.255	0.584	0.257	0.590	0.358	0.261	0.451	0.174
Unemployed (D)	-0.976	0.127	-1.145	0.090	-0.311	0.417	-0.423	0.284
Unemp.> 1 month (D)	-0.177	0.460	-0.087	0.728	-0.198	0.294	-0.190	0.334
Unemp.> 6 months (D)	0.379	0.238	0.428	0.200	0.344	0.154	0.364	0.144
Partner (D)	-0.081	0.846	-0.088	0.846	-0.131	0.710	-0.239	0.502
Separated or divorced (D)	0.105	0.813	0.063	0.896	0.470	0.167	0.520	0.136
Widowed (D)	0.109	0.781	0.180	0.673	-0.177	0.643	-0.100	0.793
Female (D)	-0.312	0.275	-0.242	0.406	0.180	0.359	0.269	0.181
Risk: money matters	0.002	0.960	-0.008	0.868	**0.093**	0.003	**0.106**	0.002
Risky act.: invest 5% in stocks	0.054	0.107	**0.075**	0.036	-0.029	0.272	-0.019	0.491
Risky act.: bet a day's inc.	**0.169**	0.000	**0.172**	0.000	**0.080**	0.007	**0.092**	0.004
Constant			-2.637	0.253			-0.286	0.849
Number of obs	434		434		427		435	
LR chi2(29)	82.25		84.32		54.14		64.11	
Prob > chi2	0.000		0.000		0.002		0.000	
Pseudo R2	0.1945		0.265		0.078		0.1246	
Log likelihood	-170.301		-116.925		-319.951		-225.121	

Table 4.23: Portfolio shares: saving accounts and building society contracts. Heckman selection regression

	Saving accounts				Buiding society contracts			
	Regression stage		Selection stage		Regression stage		Selection stage	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net income / 10,000	**-0.375**	0.011	**1.883**	0.000	**1.229**	0.000	**2.493**	0.000
(Net inc./ 10,000) sq.	0.159	0.246	**-1.310**	0.001	**-1.100**	0.000	**-2.140**	0.000
Age / 10	**-0.199**	0.000	-0.055	0.635	**0.154**	0.038	**0.500**	0.000
Age / 10 sq.	**0.026**	0.000	0.018	0.114	**-0.023**	0.002	**-0.063**	0.000
Secondary school (D)			**0.119**	0.046			**0.056**	0.036
Graduation diploma (D)			**0.192**	0.033			**0.099**	0.025
University degree (D)			0.035	0.664			**0.113**	0.002
Kids (D)			-0.005	0.951			0.032	0.383
Kids living in same house (D)			**-0.145**	0.038			-0.003	0.917
Job: blue collar (D)			-0.057	0.496			**0.070**	0.055
Job: civil servant (D)			0.188	0.137			0.051	0.264
Job: freelancer (D)			-0.161	0.412			-0.063	0.389
Job: self-employed (D)			**-0.347**	0.008			-0.008	0.884
Retired (D)			0.180	0.127			0.006	0.910
Work parttime (D)			**-0.338**	0.003			-0.003	0.949
Work little (D)			**-0.328**	0.004			0.000	0.999
Work not (D)			**-0.352**	0.001			0.040	0.469
Unemployed (D)			-0.171	0.157			-0.054	0.456
Unemp.> 1 month (D)			**0.129**	0.071			0.032	0.301
Unemp.> 6 months (D)			**-0.190**	0.018			-0.039	0.260
Partner (D)			**0.167**	0.031			0.036	0.382
Separated or divorced (D)			**-0.190**	0.041			-0.024	0.650
Widowed (D)			-0.042	0.684			**0.093**	0.064
Female (D)			0.075	0.199			0.010	0.765
Exp.: Germany's ec. developm.			-0.006	0.647			0.004	0.511
Exp.: own ec. developm.			**0.085**	0.000			**0.019**	0.004
Sample: TPI 2001 (D)			0.195	0.029			**0.136**	0.001
Sample: CAPI 2003 (D)			-0.096	0.316			**0.096**	0.045
Sample: RR 2003 (D)			**-0.292**	0.000			**0.063**	0.070
Sample: TPI 2004 (D)			0.011	0.914			**0.179**	0.000
Self-assess: optimist			-0.008	0.451			-0.005	0.378
Risk: money matters	**-0.027**	0.000	**-0.021**	0.059	-0.003	0.624	-0.001	0.960
Constant	0.879	0.000	-0.500	0.121	-0.682	0.000	-2.248	0.000
Number of obs	2755				2755			
Uncensored obs	1641				701			
Prob > chi2	0.000				0.000			
Log likelihood	-2168.966				-1474.937			
Rho	0.628				0.997			
Sigma	0.379				0.566			
Lambda	0.238				0.564			

Table 4.24: Portfolio shares: whole life insurances and private old age provisions. Heckman selection regression

	Whole life insurances				Private old age provisions			
	Regression stage		Selection stage		Regression stage		Selection stage	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net income / 10,000	0.031	0.859	**1.777**	0.000	**-0.870**	0.000	**1.420**	0.028
(Net inc./ 10,000) sq.	-0.078	0.554	**-0.897**	0.042	**0.511**	0.013	-0.947	0.104
Age / 10	**0.336**	0.000	**0.952**	0.000	**0.354**	0.004	**0.876**	0.000
Age / 10 sq.	**-0.035**	0.000	**-0.107**	0.000	**-0.037**	0.008	**-0.106**	0.000
Secondary school (D)			0.190	0.005			0.060	0.500
Graduation diploma (D)			0.104	0.293			**0.207**	0.090
University degree (D)			0.108	0.232			**0.253**	0.027
Kids (D)			-0.004	0.965			-0.118	0.364
Kids living in same house (D)			0.033	0.687			0.042	0.693
Job: blue collar (D)			-0.066	0.473			0.041	0.713
Job: civil servant (D)			-0.025	0.841			**-0.344**	0.035
Job: freelancer (D)			0.182	0.383			0.255	0.285
Job: self-employed (D)			0.083	0.542			**0.588**	0.000
Retired (D)			0.115	0.409			-0.164	0.412
Work parttime (D)			-0.096	0.443			0.001	0.992
Work little (D)			-0.112	0.416			-0.235	0.192
Work not (D)			**-0.227**	0.060			-0.200	0.199
Unemployed (D)			-0.184	0.228			0.226	0.210
Unemp.> 1 month (D)			**0.142**	0.076			0.037	0.712
Unemp.> 6 months (D)			**-0.163**	0.076			0.044	0.710
Partner (D)			0.155	0.108			-0.112	0.391
Separated or divorced (D)			-0.073	0.503			**-0.295**	0.056
Widowed (D)			**-0.246**	0.041			-0.078	0.605
Female (D)			-0.055	0.434			-0.068	0.455
Exp.: Germany's ec. developm.			0.005	0.745			0.010	0.600
Exp.: own ec. developm.			**0.075**	0.000			0.020	0.301
Sample: TPI 2001 (D)			**0.237**	0.008			0.011	0.437
Sample: CAPI 2003 (D)			0.016	0.878			0.000	0.982
Sample: RR 2003 (D)			**-0.162**	0.055			**0.244**	0.040
Sample: TPI 2004 (D)			-0.074	0.520			**0.581**	0.000
Self-assess: optimist			-0.003	0.804			**0.256**	0.023
Risk: money matters	**-0.013**	0.002	**0.016**	0.174	0.006	0.375	**0.761**	0.000
Constant	-0.263	0.245	-3.345	0.000	-0.419	0.186	-3.474	0.000
Number of obs	2755				2755			
Uncensored obs	790				289			
Prob > chi2	0.000				0.000			
Log likelihood	-1477.699				Twostep			
Rho	0.3687				0.2984			
Sigma	0.2924				0.2769			
Lambda	0.1078				0.0826			

Table 4.25: Portfolio shares: bonds and stocks, fonds. Heckman selection regression

	Bonds				Fonds, Stocks			
	Regression stage		Selection stage		Regression stage		Selection stage	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net income / 10,000	-0.200	0.360	1.046	0.103	0.098	0.632	**2.181**	0.000
(Net inc./ 10,000) sq.	0.043	0.739	0.005	0.993	-0.111	0.438	**-0.918**	0.064
Age / 10	-0.110	0.150	**0.366**	0.058	-0.106	0.124	**0.385**	0.019
Age / 10 sq.	0.013	0.071	-0.023	0.235	**0.013**	0.072	**-0.029**	0.081
Secondary school (D)			**0.202**	0.042			**0.322**	0.000
Graduation diploma (D)			0.159	0.277			**0.397**	0.000
University degree (D)			**0.313**	0.007			**0.442**	0.000
Kids (D)			-0.063	0.656			-0.151	0.179
Kids living in same house (D)			-0.108	0.349			-0.016	0.859
Job: blue collar (D)			-0.115	0.421			-0.099	0.362
Job: civil servant (D)			-0.160	0.361			-0.120	0.370
Job: freelancer (D)			-0.057	0.845			0.227	0.308
Job: self-employed (D)			0.039	0.830			0.011	0.942
Retired (D)			-0.017	0.930			-0.109	0.489
Work parttime (D)			-0.194	0.325			0.139	0.330
Work little (D)			-0.219	0.289			-0.034	0.829
Work not (D)			0.021	0.905			0.013	0.923
Unemployed (D)			-0.158	0.512			-0.030	0.869
Unemp.> 1 month (D)			**-0.223**	0.071			0.084	0.341
Unemp.> 6 months (D)			0.224	0.119			**-0.252**	0.020
Partner (D)			0.099	0.436			0.168	0.127
Separated or divorced (D)			-0.252	0.130			-0.058	0.664
Widowed (D)			0.082	0.627			0.096	0.502
Female (D)			0.083	0.388			0.004	0.959
Exp.: Germany's ec. developm.			-0.008	0.665			0.025	0.123
Exp.: own ec. developm.			**0.083**	0.000			**0.084**	0.000
Sample: TPI 2001 (D)			0.068	0.599			**0.562**	0.000
Sample: CAPI 2003 (D)			0.111	0.432			0.066	0.593
Sample: RR 2003 (D)			-0.176	0.129			-0.027	0.782
Sample: TPI 2004 (D)			-0.199	0.224			**0.473**	0.000
Self-assess: optimist			0.011	0.550			-0.015	0.324
Risk: money matters	-0.013	0.189	**0.067**	0.000	**0.020**	0.010	**0.127**	0.000
Constant	0.665	0.082	-3.653	0.000	0.426	0.064	-3.649	0.000
Number of obs	2755				2755			
Uncensored obs	232				548			
Prob > chi2	0.007				0.002			
Log likelihood	-679.936				-1170.428			
Rho	-0.024				0.176			
Sigma	0.234				0.302			
Lambda	-0.006				0.053			

Table 4.26: Financial wealth allocations by subsamples and income

	TPI 2001			CAPI 2001			CAPI 2003			RR 2003		
	Own.[a]	inc.[b]	PF share	Own.	inc.	PF sh.	Own.	inc.	PF sh.	Own.	inc.	PF sh.
N[c]	657		324	1125		576	462		208	1973		773
Saving	71.20%	2128.4	33.4%	70.8%	2247.0	48.7%	62.1%	2357.8	48.9%	58.9%	2754.2	57.4%
accounts	467		264	797		481	287		170	1162		614
Build.soc.	45.0%	2143.8	17.2%	25.6%	2268.2	11.3%	25.6%	2339.6	9.2%	23.9%	3156.1	13.2%
contracts	296		164	288		175	118		66	471		221
Whole	50.0%	2330.5	22.9%	39.2%	2460.7	23.1%	32.4%	2299.7	21.3%	26.3%	3210.2	15.8%
life ins.	328		184	441		256	150		85	520		212
Private old-	19.7%	1896.0	6.3%	11.7%	2577.5	2.7%	26.6%	2512.5	7.0%	20.2%	3189.8	4.2%
age provisions	129		52	131		58	119		44	380		93
Bonds	16.7%	2412.4	4.2%	13.3%	3402.3	4.8%	11.6%	2703.2	5.5%	7.8%	2908.1	3.3%
	109		54	150		79	54		29	153		64
Stocks/	38.6%	2309.6	16.1%	23.8%	3091.4	9.4%	20.10%	2810.6	8.2%	5.1%	3513.1	6.1%
Fonds	253		160	268		160	93		47	298		123
None of these	12.8%	1117.4		16.5%	1242.1		23.9%	1242.1		25.2%	1978.6	
items	84			185			185			497		
Refusals	0.8%	1022.6	43.5%	1.2%	1638.9	38.7%	6.9%	2327.7	24.9%	12.6%	3413.9	47.6%
	5		249	14		364	33		69	267		703

Notes: Weighted values. Percentages and absolute observations presented.

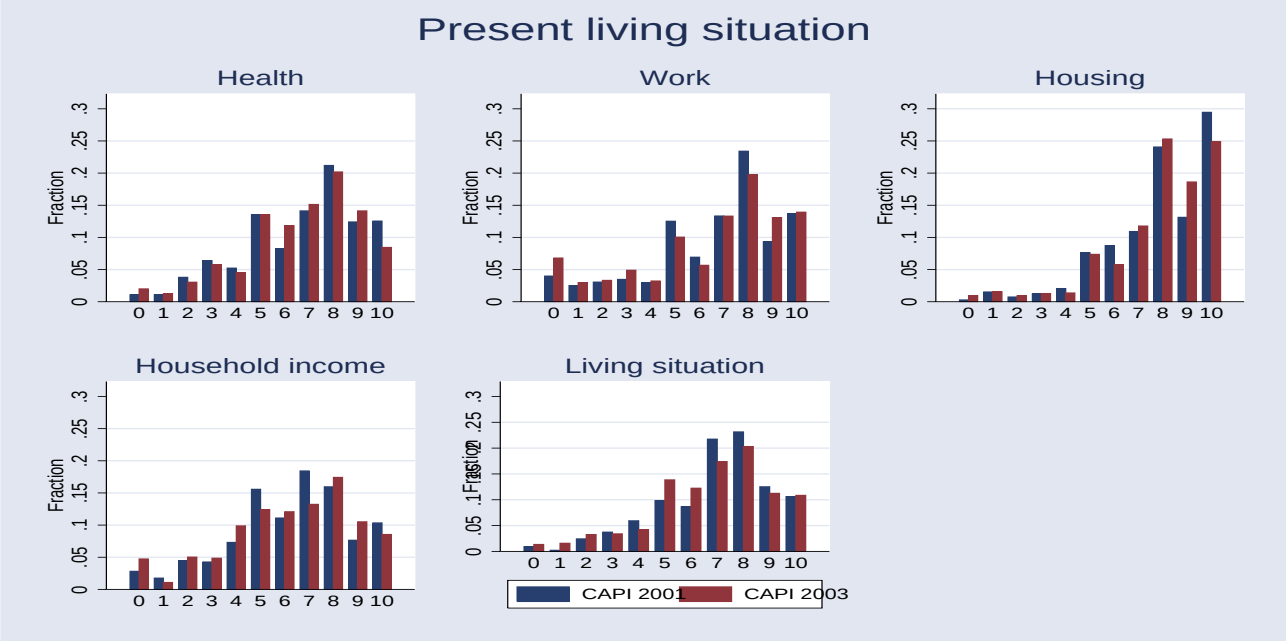
[a] Ownership of the asset category.

[b] Conditional income means on item ownership.

[c] Nonmissing observations on ownership question.

4.B Figures

Figure 4.1: Contentment with health, work, housing, income, and the general living situation

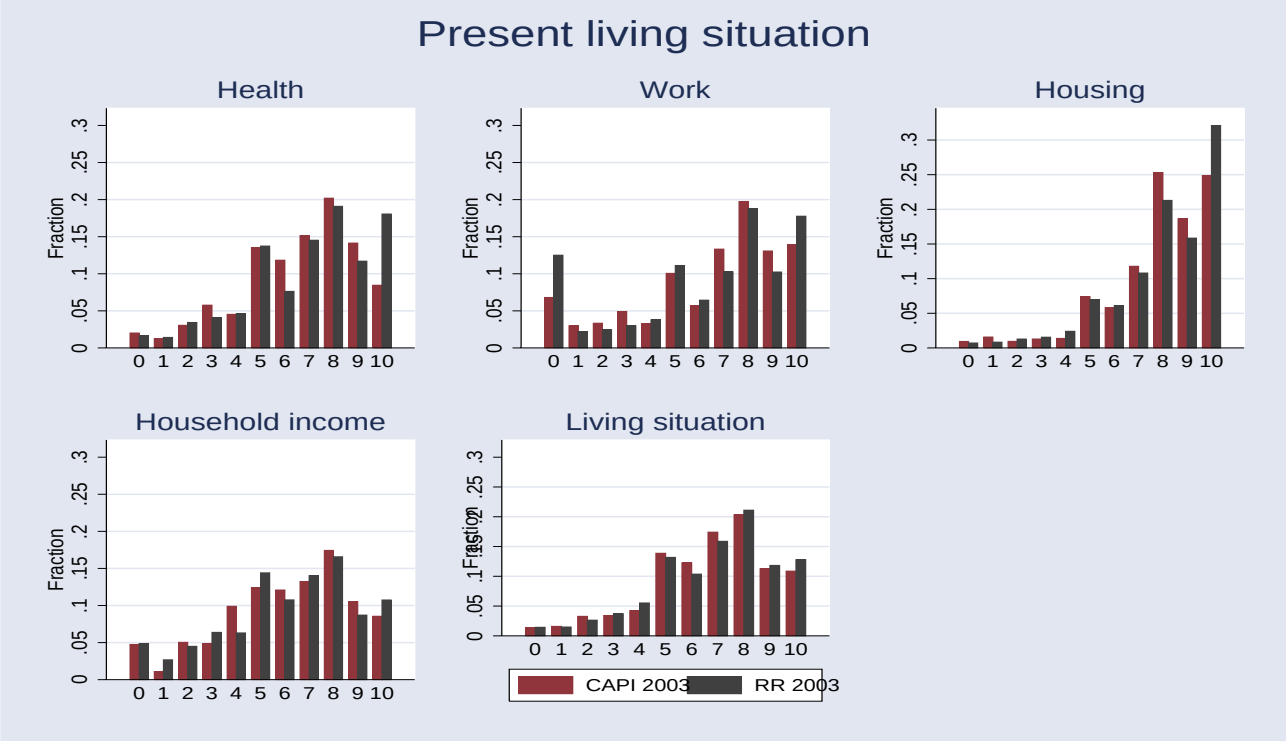


Notes: Weighted values by subsample, hh income, and age. Panel results: observations if observed in 2001 and 2003 CAPI samples. Answers are measured on a scale from 0 to 10, where 0 means 'completely dissatisfied' and 10 'totally satisfied'.

Figure 4.2: Differences in contentment with health, work, housing, income, and the general living situation

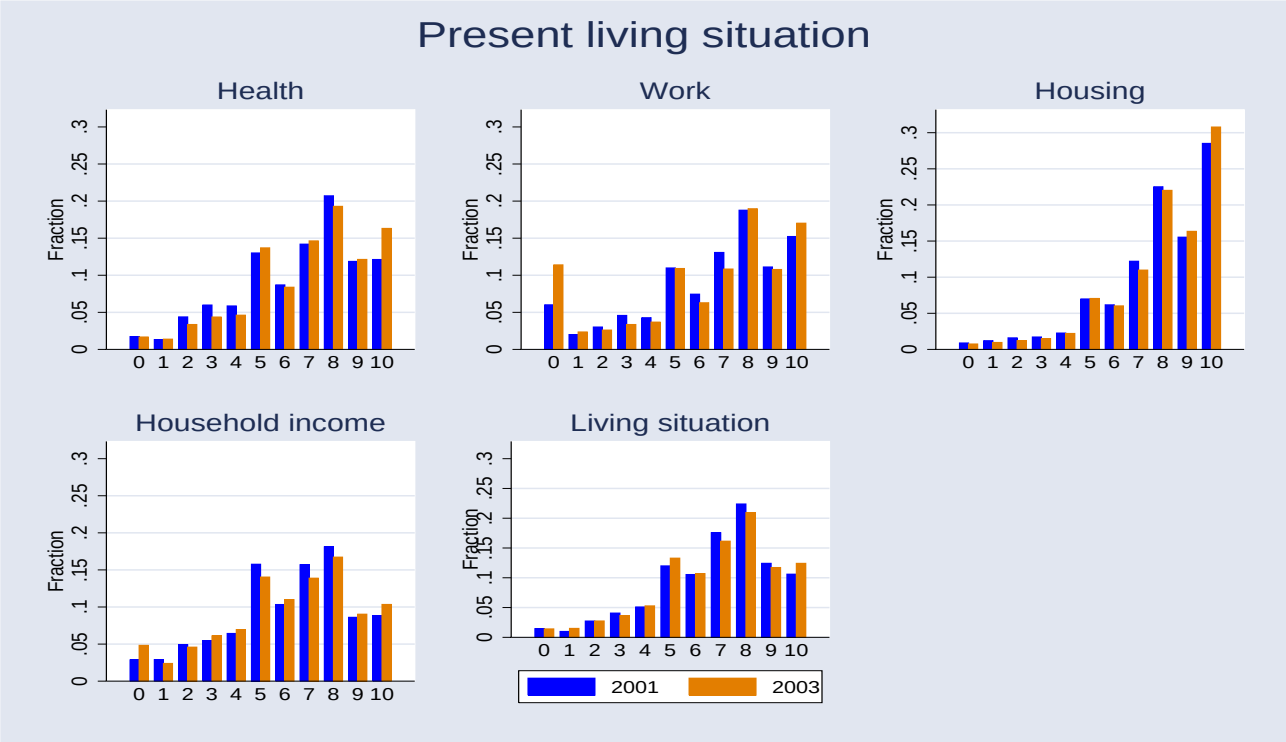
Notes: Weighted values by subsample, hh income, and age. Panel results: observations if observed in 2001 and 2003 CAPI samples.

Figure 4.3: Contentment with health, work, housing, income, and the general living situation



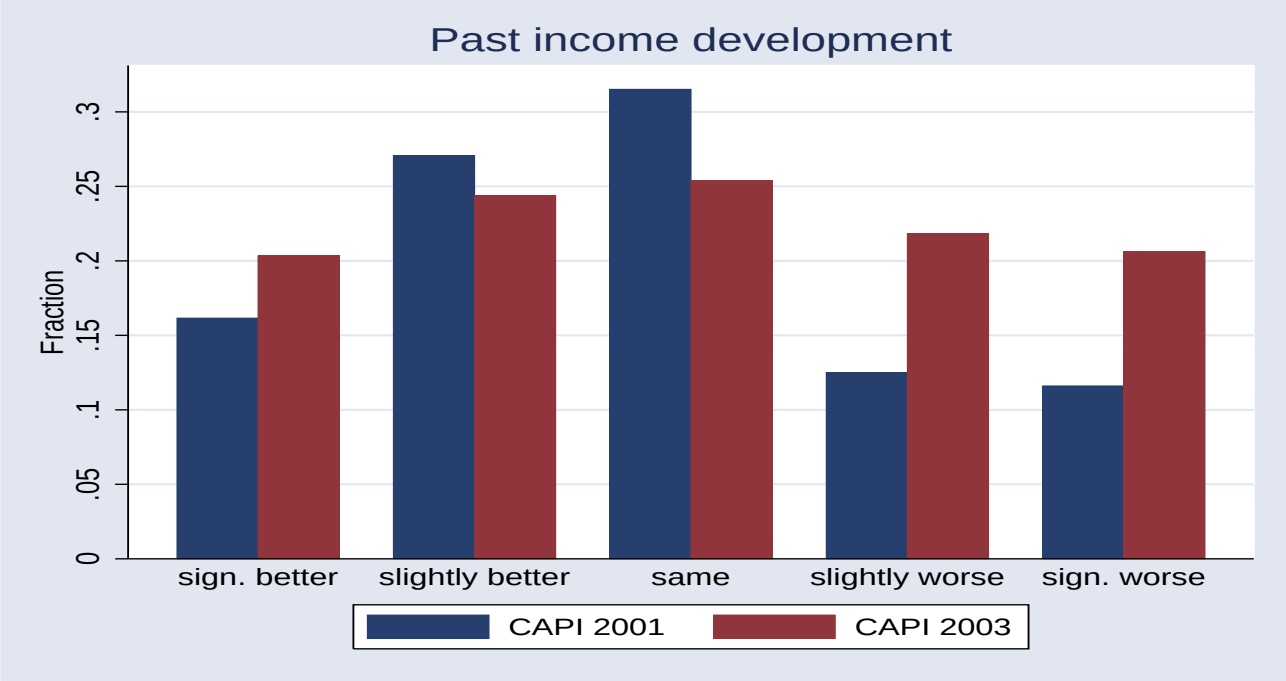
Note: Weighted values by subsample, hh income, and age. Values for both 2003 subsamples.

Figure 4.4: Contentment with health, work, housing, income, and the general living situation



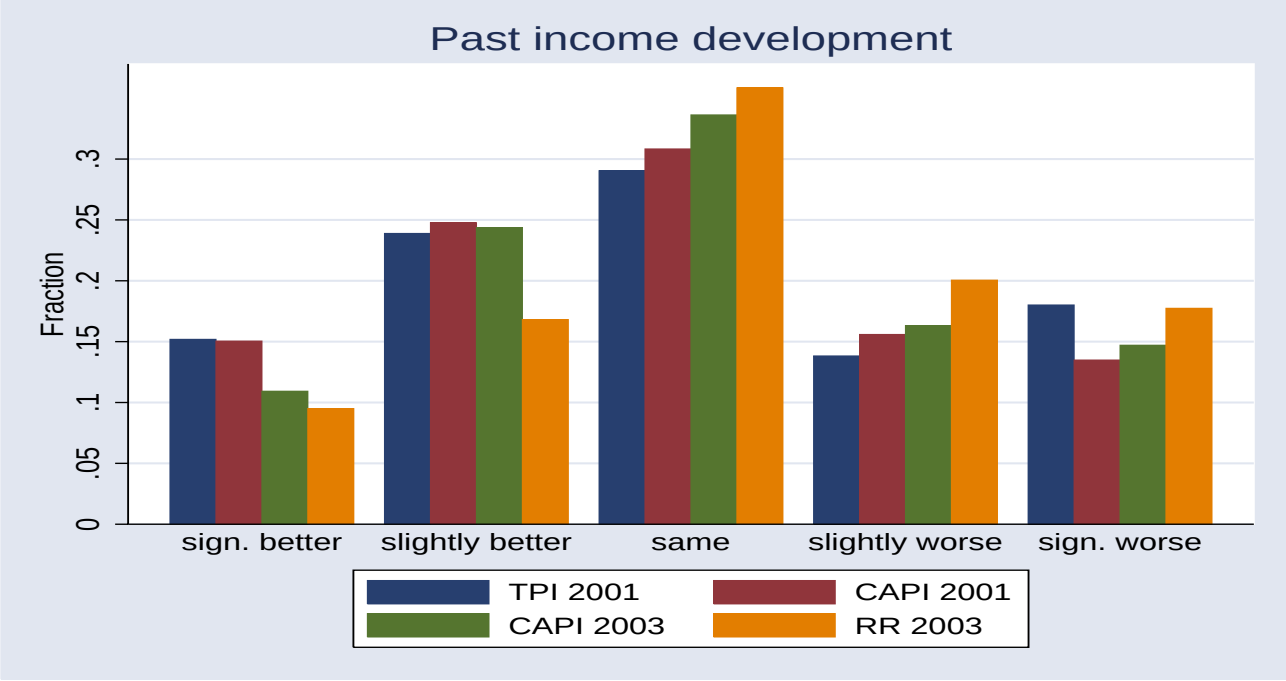
Note: Weighted values by subsample, hh income, and age. Pooled samples in 2001 and 2003.

Figure 4.5: Income development: panel comparison



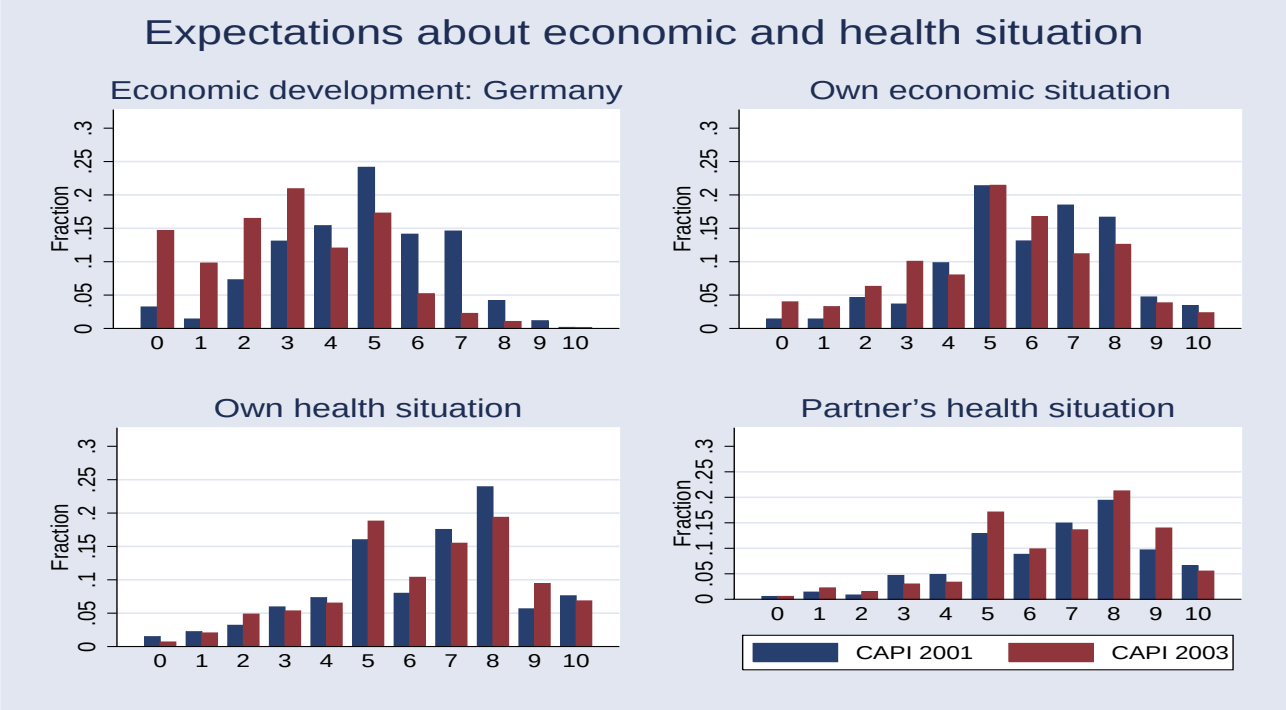
Note: Weighted values by subsample, hh income, and age. Panel results: observations if observed in 2001 and 2003 CAPI samples.

Figure 4.6: Income development: all *SAVE* subsamples



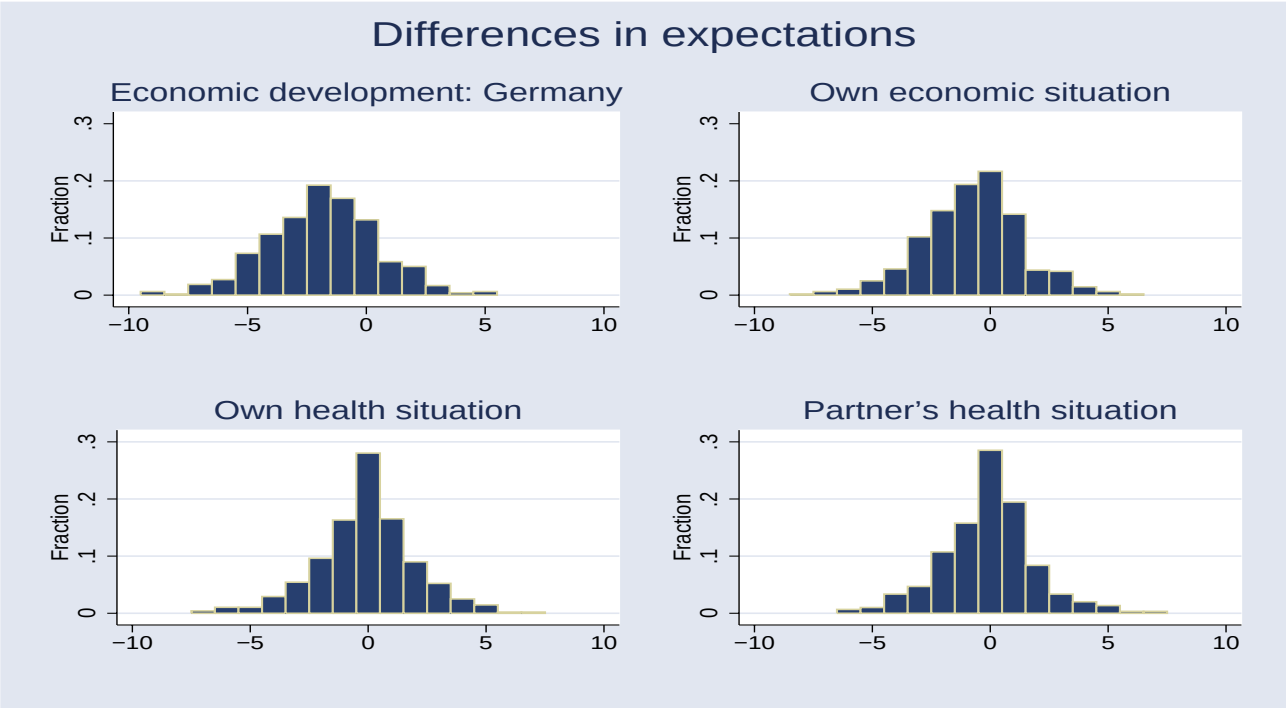
Note: Weighted values by subsample, hh income, and age.

Figure 4.7: Expectations for health and financial situation



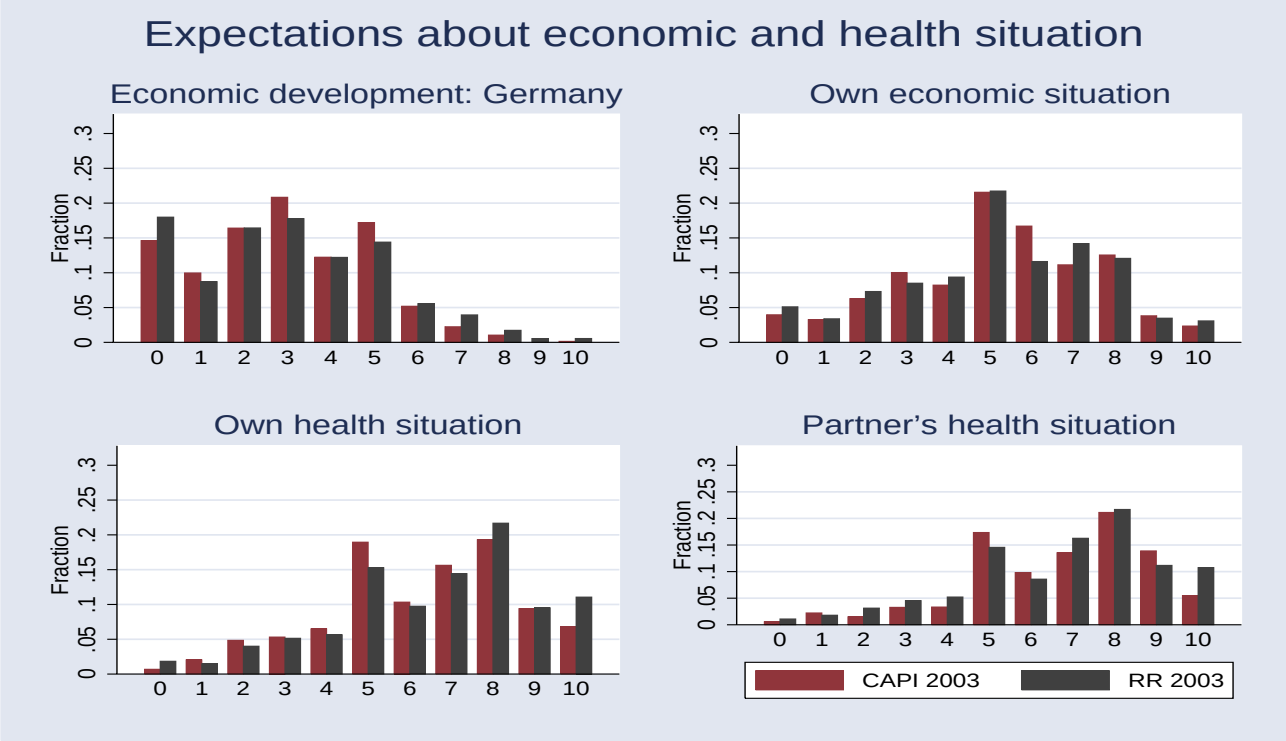
Notes: Weighted values by subsample, hh income, and age. Panel results: observations if observed in 2001 and 2003 CAPI samples.

Figure 4.8: Differences in expectations for health and financial situation



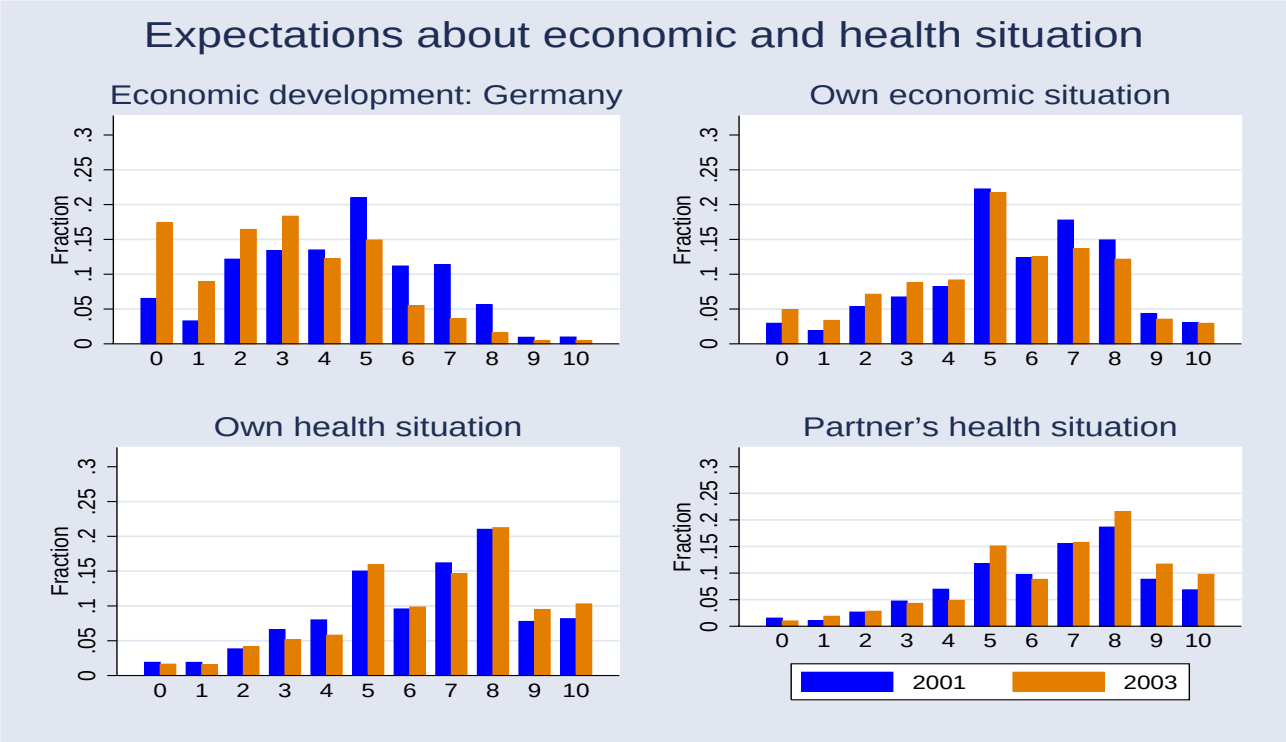
Notes: Weighted values by subsample, hh income, and age. Panel results: observations if observed in 2001 and 2003 CAPI samples.

Figure 4.9: Expectations for health and financial situation



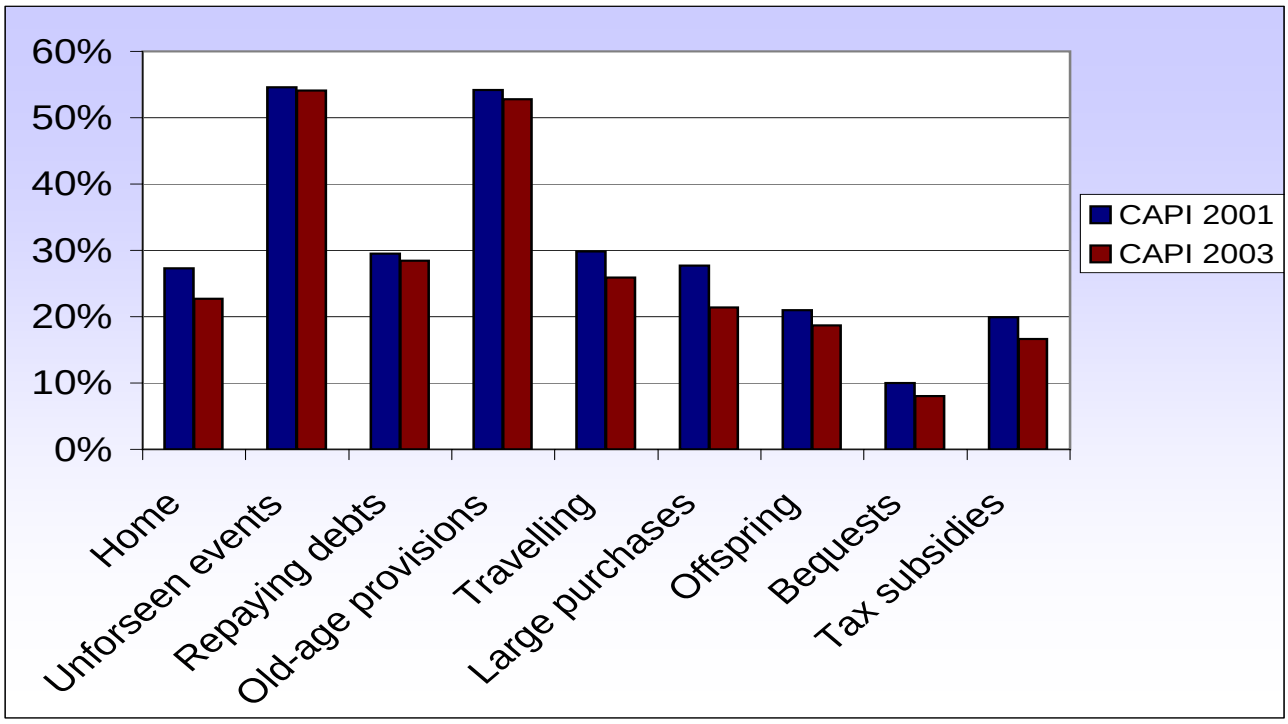
Note: Weighted values by subsample, hh income, and age.

Figure 4.10: Expectations for health and financial situation



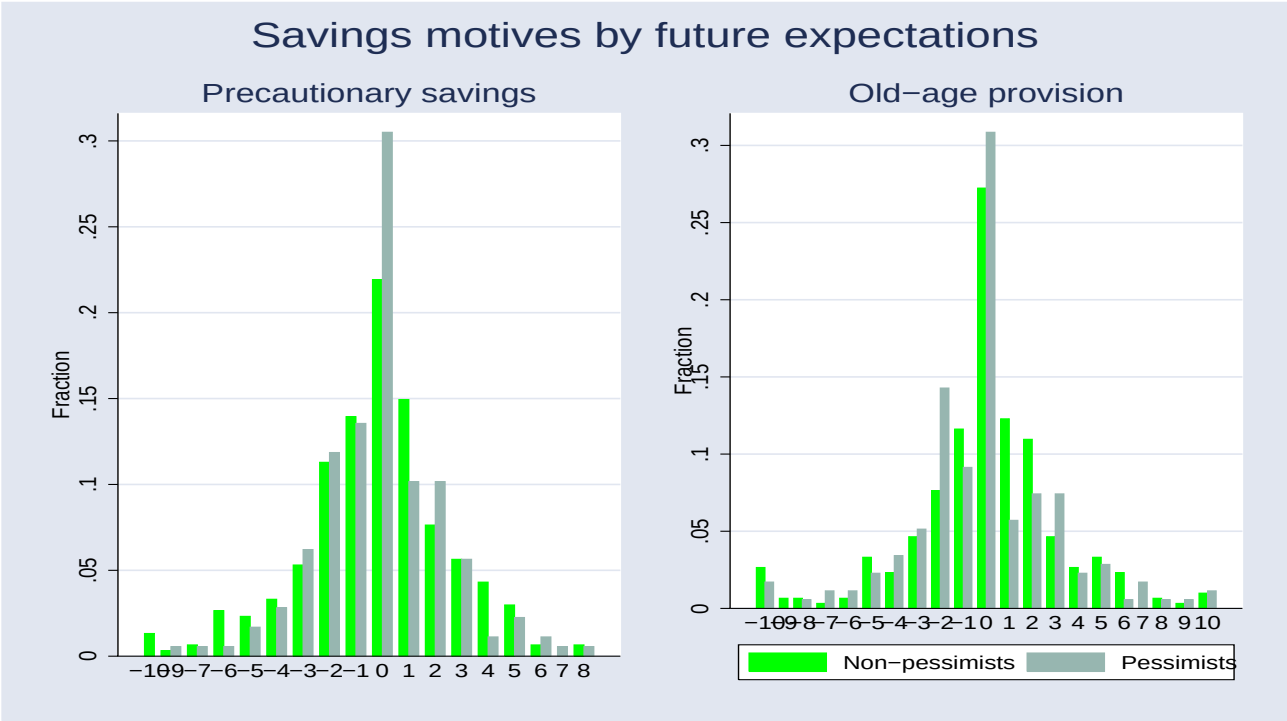
Notes: Weighted values by subsample, hh income, and age. Pooled samples in 2001 and 2003.

Figure 4.11: Importance of different saving motives



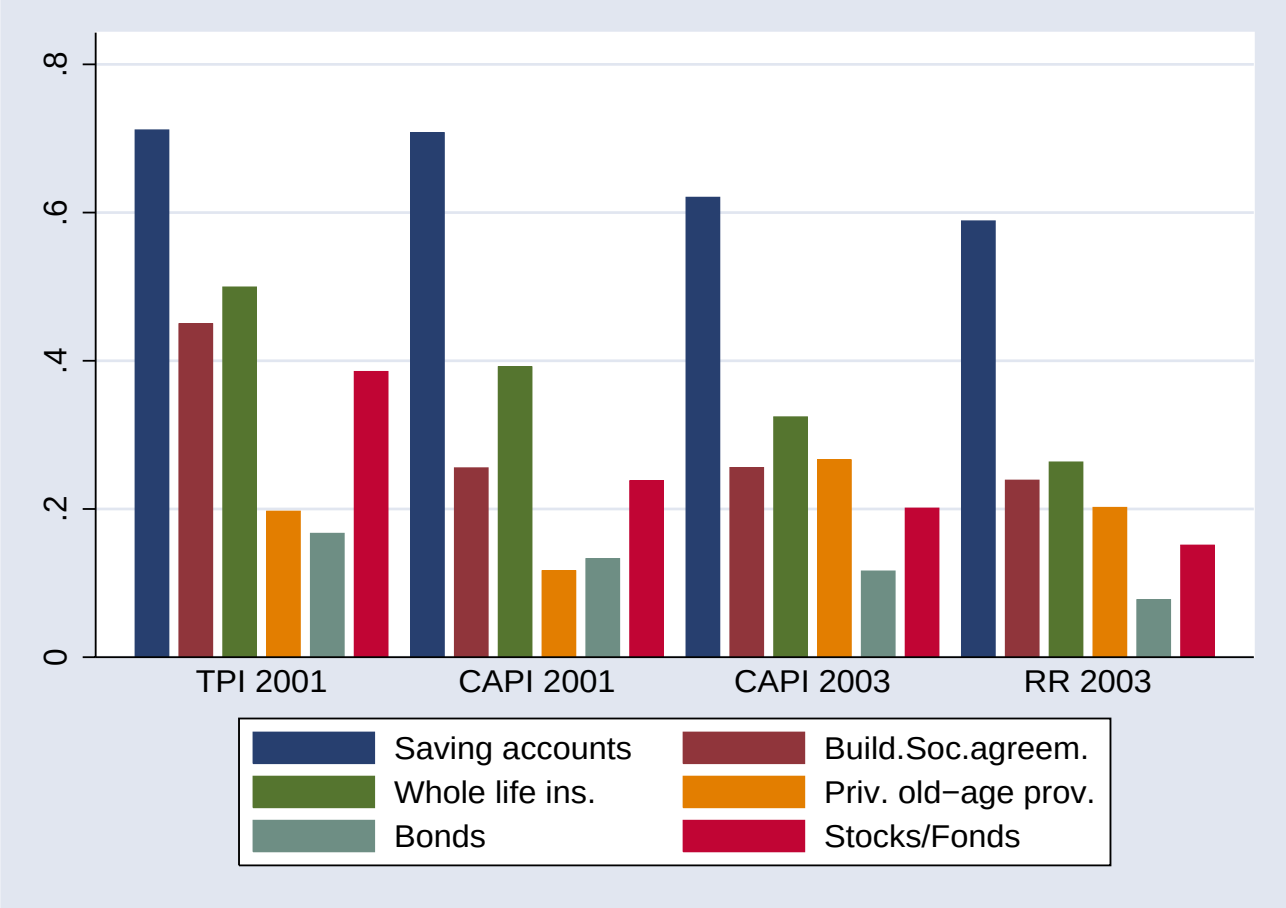
Notes: Weighted values.

Figure 4.12: Differences in savings motives by expectations for Germany's economic development



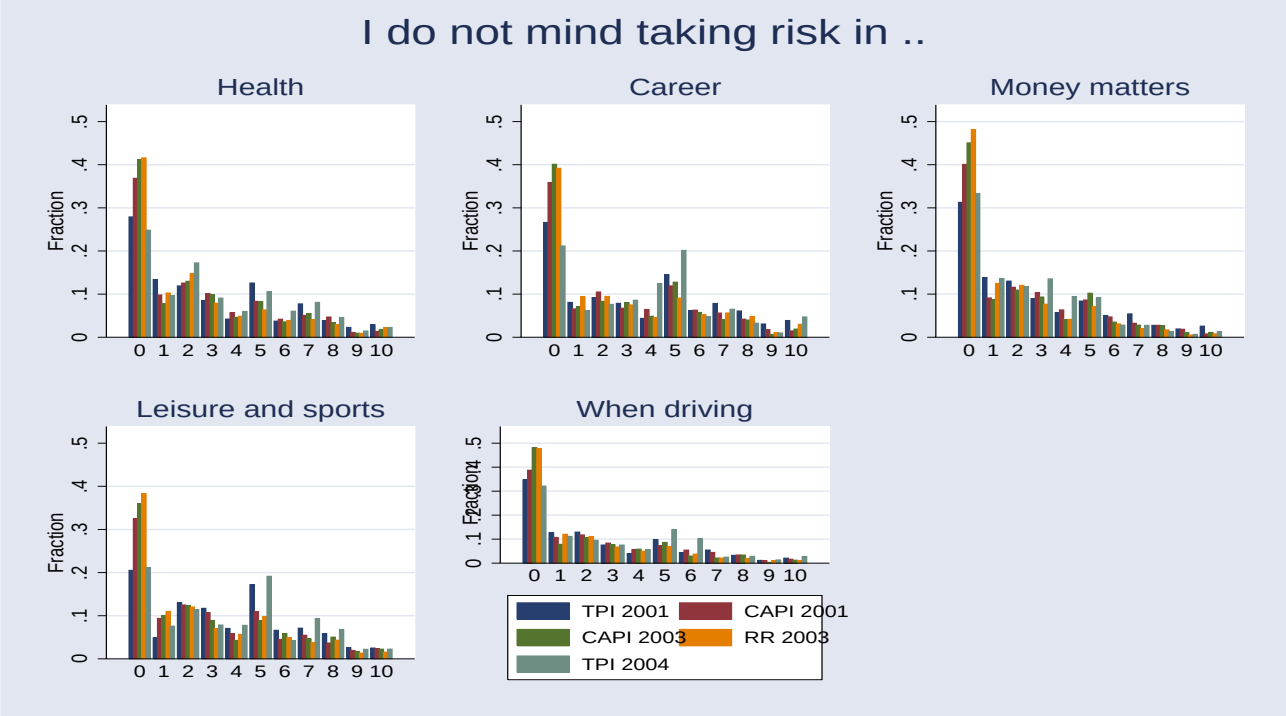
Notes: Unweighted values. Pessimists are defined as respondents whose expectation for Germany's economic development deteriorated by more than two points on a scale from 0 to 10.

Figure 4.13: Investment allocation of financial assets (ownership rates)



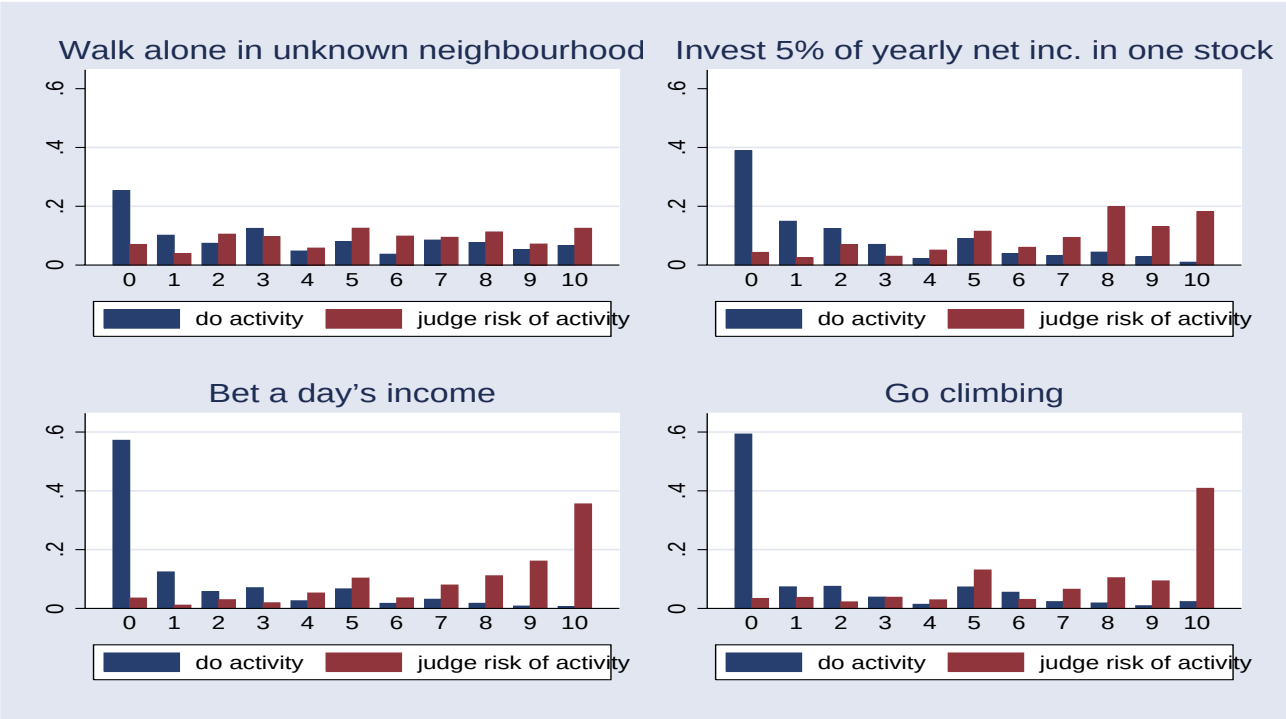
Notes: Weighted values. Private old-age provision question changed between 2001 and 2003.

Figure 4.14: Risk taking



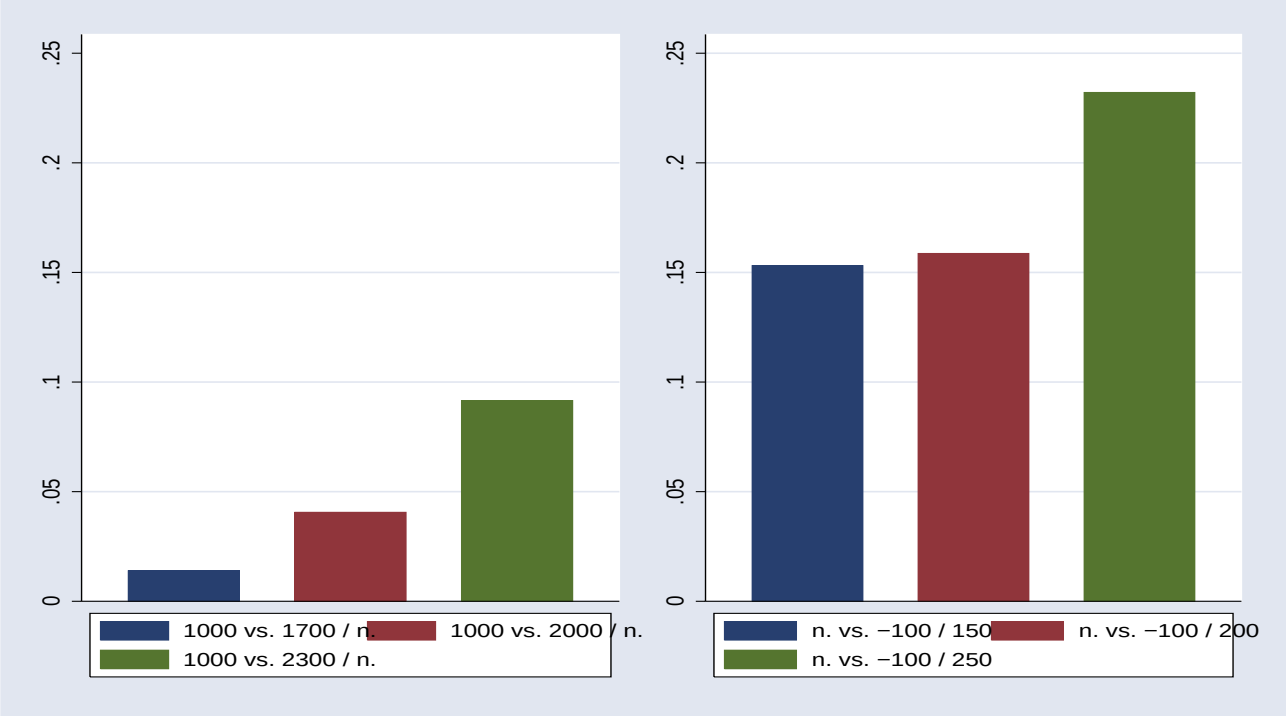
Note: Weighted values by subsample, hh income, and age.

Figure 4.15: Risk judgement and involvement probabilities



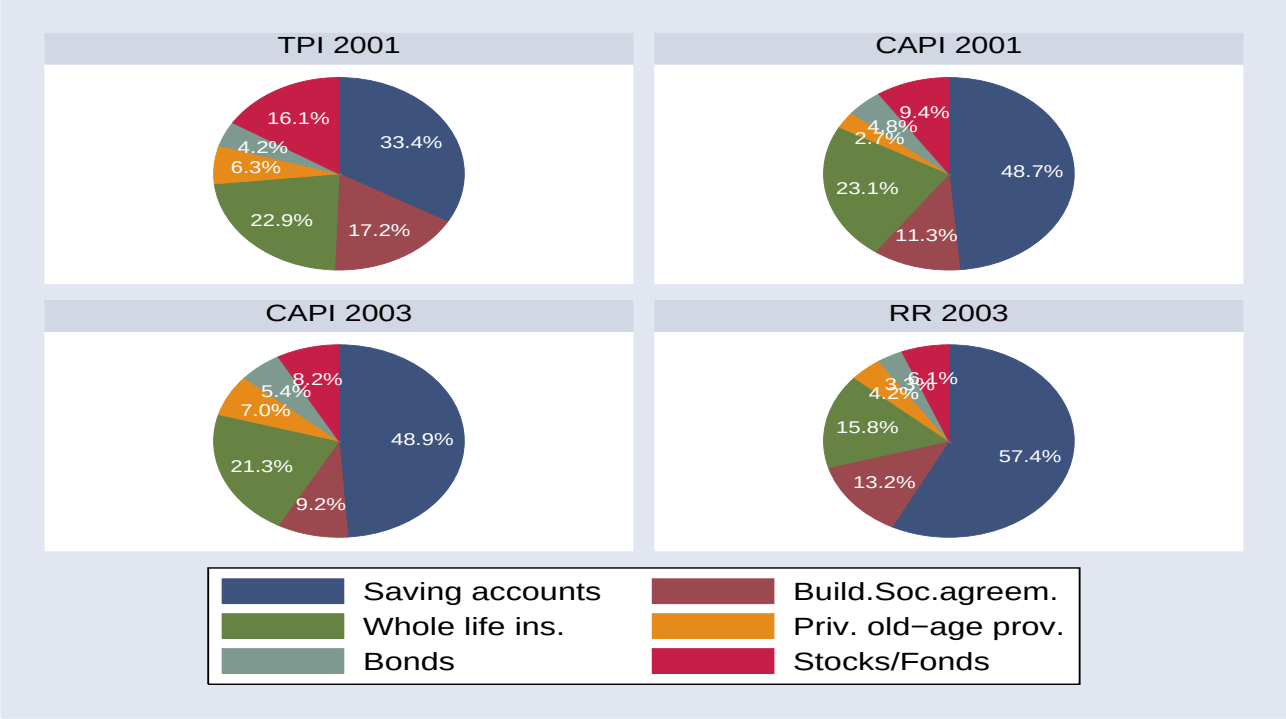
Notes: Weighted values by subsample, hh income, and age. Values for the 2004 TPI subsample only. Weighted values. 0 means very unprobable to engage in activity / not judging risky at all.

Figure 4.16: Likelihood of choosing lotteries with different expected values



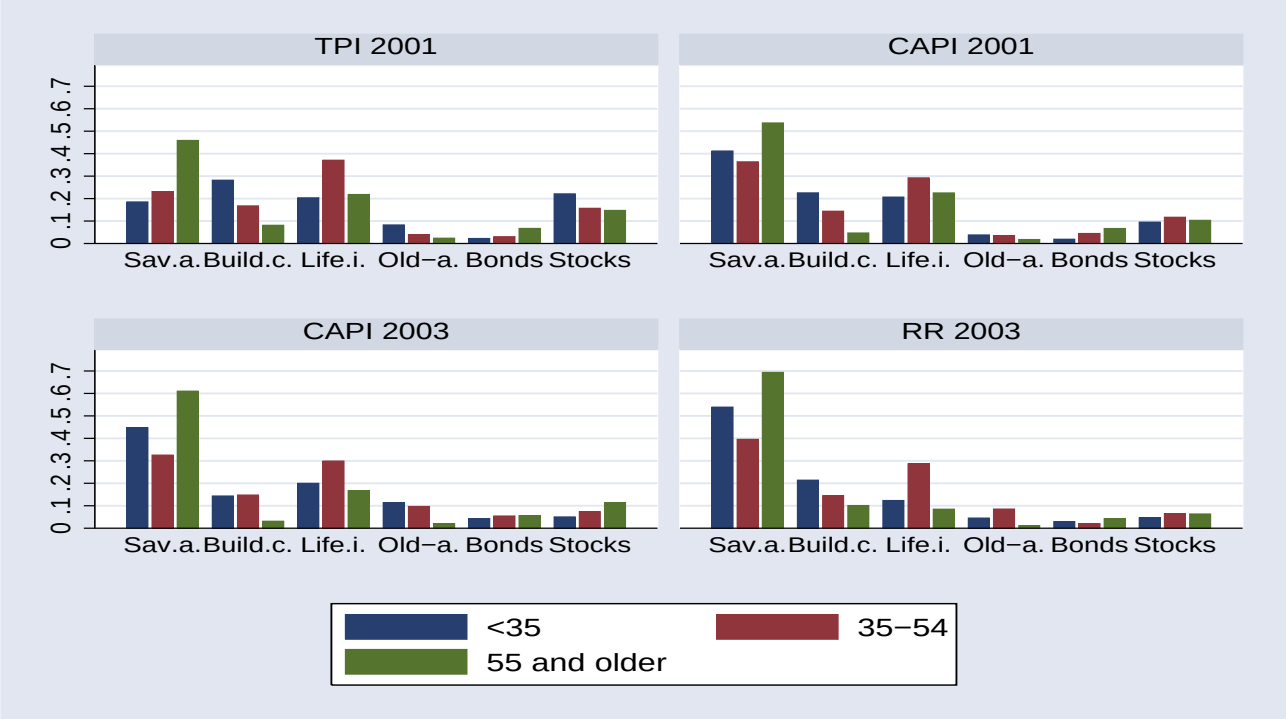
Notes: Weighted values by subsample, hh income, and age. Values for the 2004 TPI subsample only.

Figure 4.17: Portfolio compositions by subsamples



Notes: Weighted values. Private old-age provision question changed between 2001 and 2003.

Figure 4.18: Portfolio compositions by age classes



Notes: Weighted values. Private old-age provision question changed between 2001 and 2003.

Chapter 5

Precautionary savings and old-age provisions: Do subjective savings motive measures work?

5.1 Introduction

The theory of precautionary saving has challenged and enriched the literature on consumers' behavior. Precautionary saving leads to consumption cut-backs and the accumulation of wealth to insure against several sorts of uncertainty or risk, the income risk being the most frequently stated (see, e.g., Deaton (1992), Carroll and Samwick (1998)). Lusardi (1998) emphasizes that, within the life cycle / permanent income model, saving and wealth are not only related to the first moment of income, but also to higher moments, especially to the second one (variance of income).

The empirical approaches to precautionary savings have to deal with a couple of major challenges all of which make it hard to disentangle and identify its quantitative effects. As briefly reviewed in Section 5.2, economic theory provides a good deal of foundations and predictions for household behavior.

In this chapter, I will first take advantage of already used different empirical procedures to identify and quantify the precautionary savings motive. The *SAVE* data provide a variety of subjective measures for income uncertainty which have been used in the existing literature. As a next step, I adopt a new approach to map the importance of precautionary savings. I use short-run and long-run savings motives to describe differences in savings, saving rates and wealth accumulation. Even though these measures are also subjective and not quantitative like the ones used in Kennickell and Lusardi (2004), they can provide additional information explaining the heterogeneity in households' saving behavior.

This chapter is organized as follows. In Section 5.2, I briefly reconsider the importance of precautionary savings and summarize different groups of the most important results along with problems to identify precautionary savings. In Section 5.3, I examine the two main variables of the *SAVE* data set concerning precautionary savings at hand: (a) the measure of subjective earnings variance and (ab) the savings motives for precautionary savings and old-age provision (as an extended precautionary savings motive). Section 5.4 shows results for the two measures on wealth accumulation, while Section 5.5 leaves with some concluding remarks.

I will use the *SAVE* 2003 data random route subsample, since thereby I can circumvent possible sample selection problems; see Chapter 2 for a review of the different *SAVE* sample characteristics.

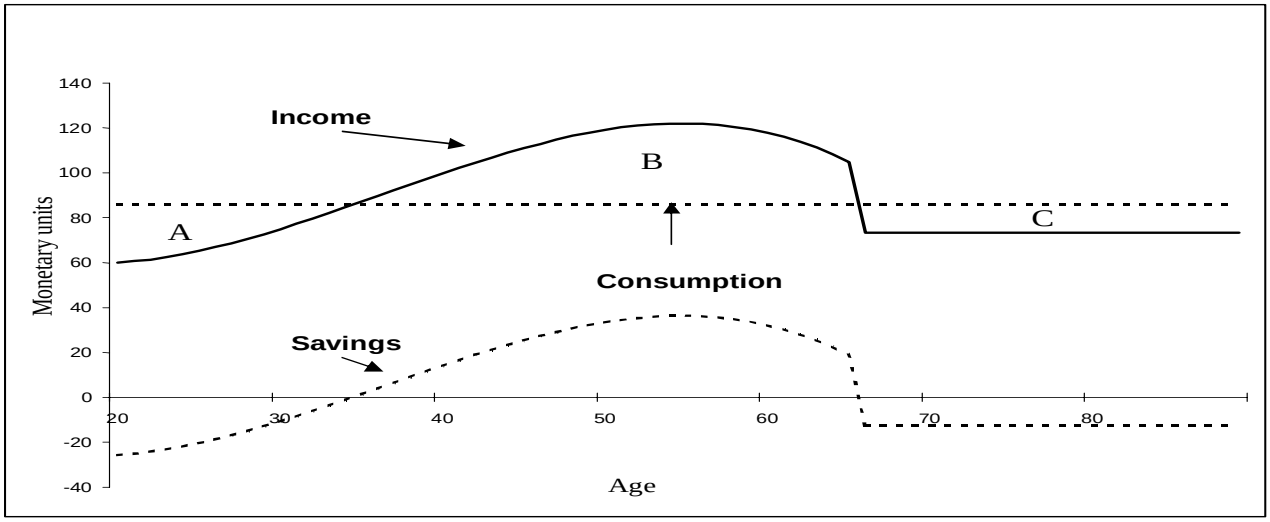
Like in the Health and Retirement Survey (*HRS*), it is only the individual deemed most knowledgeable about the family's assets, debts, and retirement planning, who is asked questions on demographics, savings, housing, net worth and income of the family.

5.2 Precautionary savings: theory and empirical findings

The life cycle [Modigliani and Brumberg (1954)]-permanent income [Friedman (1957)] model [LCPI] has become the basic theoretical framework for analyses on saving. The theory has opened the door to many refinements over the years; and its importance has thus repeatedly been acknowledged over decades, see Meghir (2004) for an actual recognition. The fundamental insight of this model describes household consumption smoothed over the life cycle, which in turn implies that individuals spend more in earlier stages of the life and build up wealth in the middle part of the life cycle. The "underlying idea of the life-cycle hypothesis - that people save for their old age - is of course not new; nor is it Modigliani's own. His achievement lies primarily in the rationalization of the idea into a formal model which he has developed in different directions and integrated within a well-defined and established economic theory, and secondly in the drawing of macroeconomic implications from that model and in performing a number of empirical tests of these implications"[81]. The LCPI builds the basis for many empirical investigations; e.g. it has proved an ideal tool for analyses of the effects of different pension systems and the discussion whether an introduction of a general pension system leads to a decline in private saving. See Barro (1974 and 1978) and Feldstein (1974 and 1978) for a controversial discussion on this topic.

Figure 5.1 depicts the simplest form of the life cycle-permanent income model.

Figure 5.1: Income, consumer behavior, and savings during the lifespan

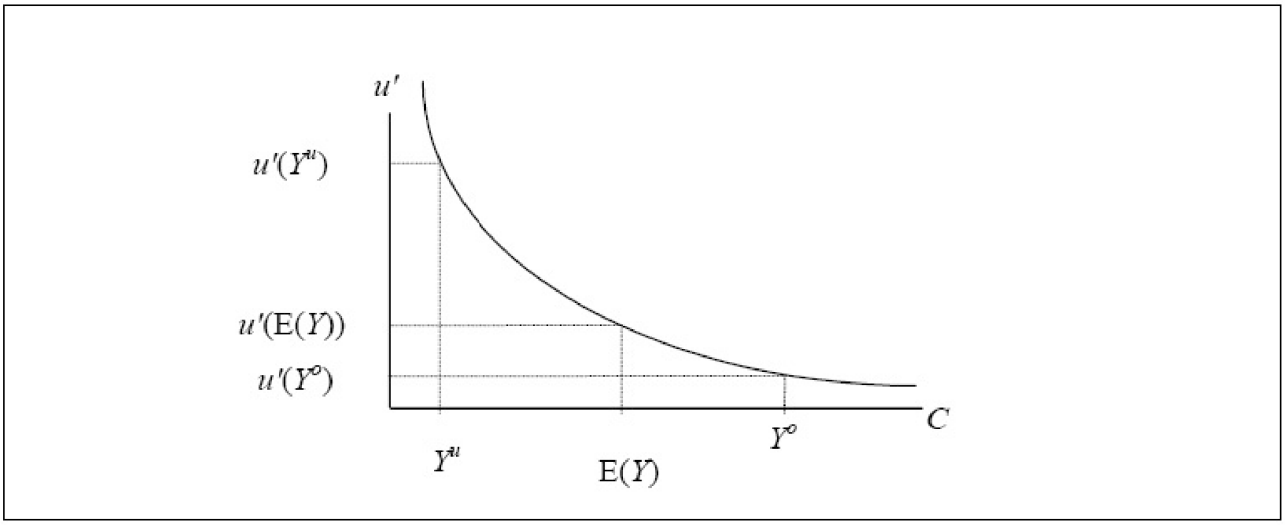


Source: Börsch-Supan and Essig (2002)

[81] Press Release: The Sveriges Riksbank (Bank of Sweden) Prize in Economic Sciences in Memory of Alfred Nobel for 1985.

One of the extensions to the LCPI is the theory of precautionary savings. It says that savings are not only functions as an income reallocation over the life cycle, but also as an insurance against income shocks. This theory implicitly presupposes some classes of utility functions. Leland (1968), Sandmo (1970) and Kimball (1990) showed that degree of prudence depends on the third derivative of the utility function (and so a quadratic utility function cannot represent the precautionary savings motive since the third derivative is zero).

Figure 5.2: Example for a two-period-model with a certain income Y_1 and uncertain income Y_2, which can take on the two values Y_u and Y_o with a probability of 0.5



Source: Rodepeter (1999)

Figure 5.2 shows that savings depends on the the range of the Y_u and Y_o. The degree of prudence depends of the third derivative of the utility function. If the third derivative is zero, the first derivative of the utility function is linear, and individuals will face no utility loss through income uncertainty, since $\mathrm{E}(u'(Y)) = u'(\mathrm{E}(Y))$.

The inability of quadratic utility functions to model the precautionary motive leads to the more realistic modelling by the family of *constant relative risk aversion* (CRRA) functions.[82] The problem in assuming CRRA utility functions, which means skipping the problematic but convenient[83] assumption of quadratic utility is that closed form solutions can no longer be derived; see Zeldes (1989).

The basic trigger for the precautionary savings motive is that insurance markets are not existent or imperfect. The theory of precautionary saving from literature on consumer's behavior predicts that in this case, risk depresses consumption and increases the accumulation of wealth. Alternatively, households could try to hedge themselves by a social network like family / friends. Wealth accumulation thus can be less important if this social network is large enough and can easily be accessed. This would mean that one needs to distinguish who has access and who doesn't. Social Security represents another

[82] An isoelastic utility function $u_t = \frac{1}{1-\gamma} c_t^{1-\gamma}$ has the property of constant relative risk aversion $\left(-\frac{u''(c_t)c_t}{u'(c_t)}\right)$ and constant relative risk prudence $\left(-\frac{u'''(c_t)c_t}{u''(c_t)}\right)$ (cf. Kimball (1990)).

[83] in the sense of computability

source of insurances. Two incentive problems for privately insuring against shocks by accumulation wealth accompany the presence of these insurances. Firstly, social security insurances tend to maintain a *relative* living standard (since the minimum level depends on the average working income); this means that the absolute insurance level rises with the productivity progress, allowing a higher living standard in a worst-case scenario. Secondly, if social security insurances are means-tested, incentives for wealth holding are lowered, see Hubbard *et al.* (1995).

Drèze and Modigliani (1972) have shown that consumption and portfolio decisions are not separable. But so far, many of the saving and portfolio choice models have been estimated separately. Heaton and Lucas (2000) find that business owners (taking a high income risk) have a lower probability to invest in stocks. In *SAVE*, business owners have a highly significant positive probability of owning stocks (Probit results show marginal effects of about 14 to 15%, depending on specification).

5.2.1 Literature review

The literature on precautionary savings leaves us with quite mixed results. E.g. Skinner (1988) calculates precautionary savings up to 54% of total life cycle saving and that precautionary savings are higher when consumers are more risk averse and when borrowing constraints are more immediate, in accordance with Zeldes (1989). Other simulation results (Caballero (1991) and Gourinchas and Parker (2002)) studies do support these theoretical findings. Cagetti (2003) finds in his simulations that wealth accumulation is driven mostly by precautionary motives at the beginning of the life cycle, whereas savings for retirement purposes become significant only closer to retirement.

Browning and Lusardi (1996) and Kennickell and Lusardi (2004) list the main contributions to empirical evidence for precautionary saving; the results are rather mixed. More precisely, these papers can be grouped to have shown four different ranges of results:

1. Skinner (1988) and Dynan (1993) find little or no evidence for precautionary savings.

2. Guiso *et al.* (1992), Lusardi (1997, 1998 and 2000) and Arrondel (2002) report modest values for the accumulation of precautionary wealth using subjective risk measures.

3. Dardanoni (1991), Hubbard *et al.* (1995), Kazarosian (1997), Carroll and Samwick (1997 and 1998), Engen and Gruber (2001) and Carroll (1997), in contrast, find that precautionary savings produce a considerable share of wealth. E.g. Dardanoni (1991), using data on British households, found the average consumption across occupation and industry groups to be significantly lower when income variance is greater; he estimates that more than 60 percent of saving is due to precautionary motives.

4. Murata (2003) and Kennickell and Lusardi (2004) provide mixed results for different types of households each of which associated with different risk exposure.

5.2.2 Problems associated with the empirical assessment of the precautionary savings motive

The meanwhile more or less basic procedure for identifying the existence and degree of the precautionary motive in the empirical literature is to identify the relation between household wealth W_h, permanent income Y_h^P, a set of covariate control variables X_h typically including all sets of household variables (socio-economic variables like age, job variables and other characteristics) and some risk measure R_h:

$$f(W_h) = g(Y_h^P, X_h, R_h) \tag{5.1}$$

As mentioned in Kennickell and Lusardi (2004), the large range of estimates are due to differences in the data and the methodologies used. The two most important variables to deal with are wealth and risk, both very difficult to assess. In the following, I quickly review the problems mentioned in Kennickell and Lusardi (2004) associated with the assessment of these variables.

Wealth

The basic question associated with wealth is: which wealth measure should be used in estimations? Wealth consists of different components, which differ in terms of accessibility and liquidity. E.g. the total wealth measure in *SAVE* consists of 13 different wealth items (8 of them being financial wealth items with different liquidity). Typically, the largest share is housing wealth, as Table 5.1 shows. Unconditional values confirm the typical positively skewed distribution: few large values, many small values, especially zeros.

Table 5.1: Shares of different wealth items

	Owner occ. housing	Other hous.	Business	Financial	Credits[a]	Other real wealth
Unconditional						
Mean	42.89%	3.61%	1.65%	50.28%	41.85%	1.57%
Median	26.53%	0	0	30.00%	0	0
Std. Error	1.57%	0.47%	0.34%	1.61%	13.07%	0.32%
Obs.	787	787	787	787	787	787
Conditional[b]						
Mean	83.56%	36.00%	32.42%	57.02%	171.53%	16.88%
Median	91.35%	30.95%	23.70%	77.21%	26.59%	7.53%
Std. Error	0.99%	2.60%	4.60%	1.67%	52.59%	2.87%
Obs.	404	79	40	694	192	73

[a] Total wealth was calculated gross of credits (which makes look shares larger than they are since the denominator is thus larger). Therefore, shares, neglecting credits, sum to 1.

[b] Conditional on having positive values of that share; this means that medians are larger than 0.

Source: *SAVE* 2003 Random sample.

As emphasized in Kennickell and Lusardi (2004), business owners might cause problems for assessing the precautionary motive since their behavior differs largely from the rest of the sample. In *SAVE* 2003, they hold 23.0% of total wealth, though representing only 4.7% of the sample (problem: 5.3 % nonresponse to question of business wealth ownership.) There are 40 observations for wealth if

respondent is a business owner, 1109 altogether. Numbers are not as high as in Gentry and Hubbard (2000) where this group accounts for 42.1%; but in their sample, 11.4% are business owners.[84]

The quintessential point here is that it might be highly misleading to neglect certain wealth categories and simply concentrate on liquefiable financial assets like saving accounts, since that procedure might neglect a much higher stock of more long-run precautionary wealth. This, again, recurs the question of the time horizon of precautionary savings, or, more principally, what exactly is precautionary savings? Against which risk should it protect / insure households?

Risk measurement

As mentioned in the introduction, precautionary savings are supposed to be some sort of replacement for incomplete or even non-existing insurance markets. Much of the typical long-term risks are normally covered by compulsory insurance plans, like the insurance against the longevity risk (public and private pension systems / occupational pension plans, and also the public long term care insurance), and insurances against health risks (public and private health insurances). Also many of the more short-run risks are typically insured by compulsory public insurances, like unemployment insurance. See Börsch-Supan (2004b) for a review of the history, negative incentives and possible threats to the German social security system.

So the question remains, which risks does the household need to insure against and build up a wealth stock to rely on. Long-run and mostly unforseen shocks are those affecting the life time income path (cf. Figure 5.1) due to job loss or wanted, undesired changes to a less-paid job, wealth shocks, premature death of he bread-earner, or the political risk of pensions, only to name some. This would change the curvature of life time income thus lowering the possible permanent consumption path. A second risk would be the mentioned transitory short-run income risk proposed by Friedman (1957).

These reasons lead to the usage of income risk most common in the empirical literature on precautionary savings.

The empirical problem, though, remains. What are good measures for income risk? This is typically proxied by the variance of total income. Still, there are two objections to this approach. (1) Caballero (1991) and Browning and Lusardi (1996) point out that the calculated income variation could be well-known by the respondent and hence already been insured against privately. (2) Measurement errors could possibly wrongly be identified as transitory income in panel analysis.

To circumvent the mentioned problems, one can use subjective measures for income and risk. Of course, as with all subjective data, the door is open to all sorts of cognitive problems. Do respondents process the information given in the questionnaire well? There is now an extensive literature in survey research on cognitive processes that generate survey responses and on pitfalls that should be avoided in survey design; Sudman *et al.* (1996) and Tourangeau *et al.* (2000) provide overviews of the literature on survey response behavior and question design in cognitive and social psychology. Cognitive issues in households' reports of financial variables, in particular with respect to reports of household income, are discussed by Moore *et al.* (1999).

[84] Same measurement: when respondents answered to owning business assets, they were classified as business owners even if the business ownership value is zero or below zero.

The longevity and health risk are risk factors less frequently used in empirical studies, mostly for the reason of a lack of available data, and for the existence of the above mentioned social security systems. I will use two subjective variables rudimentarily covering these two risk factors, see Section 5.3.1.

Permanent income

Another possible challenge is the determination of the third variable entering Equation 5.1: how can the household's permanent income Y_h^P be identified? In panel studies, this issue raises possible identification problems already mentioned (to differ between measurement error and transitory income). But in cross sectional analysis, data are not available but for one single income observation which requires a rather different approach. King and Dicks-Mireaux (1982) propose a measure of permanent income which can be calculated from cross-sectional data which was also used by Starr-McCluer (1996) and Kazarosian (1997)[85] to measure the effects of health insurances on precautionary savings. Since I will use the *SAVE* RR 2003 subsample only leaving me with one cross-sectional data base I will follow this approach and therefore quickly review the basic ideas of this measure.

Permanent income is modelled as function of $\mathbf{Z_i}\gamma$, a vector of observable characteristics with γ, the associated parameter vector, and s_i being an unobservable variable measuring characteristics (skills, luckiness, power), and $c(A_i)$ controlling for technical progress and, therefore, sets younger individuals better off (cohort effect).

$$\ln Y_i^{perm.} = \mathbf{Z_i}\gamma + s_i - c(A_i) \tag{5.2}$$

Permanent and current income differ for two reasons: (i) the existence of an age-earnings profile and (ii) the transitory earnings component u_{it} such that

$$\ln Y_i^{current} = lnY_i^{perm.} + h(A_i - \overline{A}) + u_i, \tag{5.3}$$

where h represents the age-earnings profile, restricted to be constant across the population. Inserting 5.2 into 5.3 gives the estimation equation

$$\ln Y_i^{current} = \mathbf{Z_i}\gamma - c(A_i) + h(A_i - \overline{A}) + s_i + u_i. \tag{5.4}$$

Since the earnings profile and the cohort effect cannot be separately identified in this equation, King and Dicks-Mireaux (1982) suggest to use data from outside the sample. I will use the wage index development[86] in Germany separately for respondents and for their partners in the regression, assuming that every cohort enters the labor market at age 20. With the parameter estimates $\hat{\gamma}$ and $\hat{c}$ it would be possible to impute $\ln Y_i^{perm.}$ if it was possible to disentangle the error term $s_i + u_i$ to receive the individual-specific effect s_i. King and Dicks-Mireaux propose the share of s_i of the total error term to be 0.5 after considering longitudinal studies on earnings. Therefore, the measure of of permanent income I will use here is the predicted value $\hat{Y}_i$ from the earnings regression plus half of the difference between observed and imputed income.

[85] This work expands this model using panel data

[86] Alternatively, Kapteyn *et al.* (2004) use GDP per capita as a cohort productivity measure.

I use the permanent income imputation only for households which have at least one member not yet retired. The reason for this is that pension income, in contrast to permanent *earnings*, is determined once a person has entered retirement.[87] One could have, though, used all households since pensions represent claims which were earned during the working life thus reflecting a fraction (typically about 70% of net earnings, see Braun *et al.* (2000)) of working life permanent earnings.

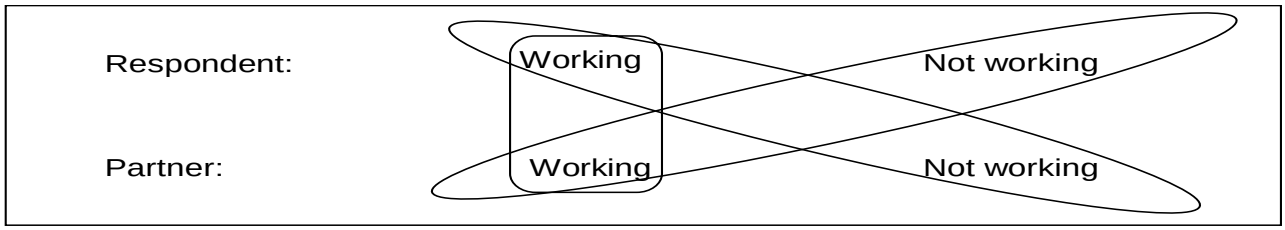
The *SAVE* 2003 RR sample has a rather high fraction of female respondents. This might cause a problem if the wage earner is the husband, and thus regressing reported household income on a set of regressors to use the predictions from this regression as a proxy for permanent income should not only include the respondents' characteristics, but also the partners'.

An alternative would be to consult an external data source, as proposed by Browning *et al.* (2003) for improving the precision of consumption question by asking a non-exhaustive list of consumption questions in a survey and imputing total consumption using the country's official consumption and expenditure survey; in this case, the German Income and Expenditure Survey (*EVS*). For expenditures, this is done in Chapter 6. For income, this is due to further research.

At the time being, I will content myself with the within-survey imputation of permanent income using the prediction of household earnings from a regression as proxy. As regressors I choose a set of respondent and, since I cannot separate incomes for each earner, of partner characteristics. This is especially important since the *SAVE* questionnaire does not seek explicitly the household head as respondent which typically is the main income earner, see last paragraph in Section 5.1.

I include all households if either the respondent, the partner or both are fully employed. This is depicted in Figure 5.3. All variables are therefore interacted accordingly to this selection, e.g. partner's variables are set to zero if there is no partner.

Figure 5.3: Household selection for permanent income regression



The chance of losing observations due to missing values in one of the variables is rather high because of the large number of regressors. This can lead to more missing observations of predicted permanent income than observed earnings. Two correction steps follow: (1) I replace those missing predicted missings of $Y_{HH}^{perm.}$ with the observed values of $Y_{HH}^{obs.}$. (2) Also, since nothing restricts predicted income values being positive, I replace all negative predicted values by the observed household earnings. Regression results are listed in Table 5.7. I used two specifications, one where absolute marginal effects are modelled constantly (absolute income as dependent variable), and another where percentage marginal effects are modelled constantly (income in logarithm), which is the more common estimation

[87] The main sources of variation of pension income is policy interventions like, e.g. the shift from gross to net earnings indexation, or changes in household composition (divorce or death of the partner).

procedure in the literature. The single elements of the fourth-order age polynomials are more or less insignificant. Still, they are jointly highly significant, tested separately for the respondent's and the partner's age.

Further problems to deal with

There are other sources of problems associated with the measurement of precautionary savings I will only briefly discuss. All the following examples are thoroughly highlighted in Kennickell and Lusardi (2004), so I refer to that source for further reading. The arguments are mentioned here to show possible problems in the following estimation procedure.

Liquidity constraints can affect individuals differently, which may lead to different wealth accumulation other things being equal since households could borrow in emergency situations.

Restrictions on the functional form can lead to difficulties if it systematically excludes certain groups of households. E.g. if one restricts the function $f(W)$ in Equation 5.1 being logarithmic, this automatically excludes zero and negative wealth (indebted households). For transformation: see Burbidge *et al.* (1988) who revisited the inverse hyperbolic sine function proposed by Johnson (1949); MacKinnon and Magee (1990) who developed them further and Carroll *et al.* (2003) for an application of wealth data.

Unfortunately, there exist no embedded ado-files containing the Inverse Hyperbolic sine function transformation ML-implementation. There seems to be a group of people working on it and this implementation is "in the pipeline"[88], so it can and should be used in further research. The transformation parameter θ from the inverse hyperbolic sine transformation of y is

$$\begin{aligned} g(y,\theta) &= \frac{\ln[\theta\, y + (\theta^2\, y^2 + 1)^{\frac{1}{2}}]}{\theta} \\ &= \frac{\sinh^{-1}(\theta\, y)}{\theta} \end{aligned} \tag{5.5}$$

Carroll *et al.* (2003) estimated the parameter θ to be 3.87 in a regression using Equation (5.5) for $\frac{W}{y_p}$ for SCF 1983 data. I will compare three values for θ: 1, 2 and 3.[89] Figure 5.4 shows the transformed wealth values for θ: 1, 2 and 3 and also for $\theta = 1.05$ and $\ln(wealth)$. The transformations for the $\theta = 1.05$ and the ln function nearly coincide (for strictly positive wealth values);[90] since the ln function is the preferred transformation in the literature, I will use 1.05 for θ in all transformations. I stress here that normally, functional form restrictions are typically set without further mentioning. The most common values for θ, for the Box-Cox transformation $(y^\theta - 1)/\theta$ are $\theta = 0$ and $\theta = 1$, giving the log and the linear form, respectively. Assuming $\theta = 1$, further restrictions imposed in wealth regressions are: $\ln W$ using only $W >$ \$4000 [Diamond and Hausman (1984)], $\ln W/Y^P$ using only $W >$ \$2500 [King and Dicks-Mireaux (1982)], $\ln[W - \min(W, 0) + 1]$ [Starr-McCluer (1996), Carroll and Samwick

[88] 10th UK Stata Users Group meetings: Abstracts Monday, 28 June 2004. "A comment on infrequency of purchase models in Stata" by Julian A. Fennema, j.a.fennema@hw.ac.uk, Centre for Economic Reform and Transformation, Heriot-Watt University.

[89] Using l'Hopital's rule it is easy to show that $\lim_{\theta \to 0} g = y$.

[90] For smaller wealth values, the difference is slightly larger; the higher the values, the smaller the gap.

(1997 and 1998)]; in the first two cases, the authors confronted the selection problem by the Heckman correction method.

While there has been an extensive discussion of the importance of macro shocks in the estimation of Euler equations,9 this topic has been largely ignored in the estimation of precautionary saving. However, this problem is important in this context as well. It is not possible to estimate the extent of precautionary accumulation using a single cross-section of wealth data. The problem may be best understood by using a simple example. Suppose that, because of a national housing market bust, the wealth of home-owners was substantially reduced. Suppose further, as it is not unreasonable, that home-owners are less likely to face high earnings risk. Simple regressions of wealth on income risk lead to biased estimates of the extent of precautionary accumulation.

In cross sectional analyses, macro shocks may cause problems due to biases. E.g., if the stock market went up, and if stock owners are more likely to have risky earnings, then estimating the extent of precautionary wealth accumulation possibly leads to biased estimates. This argument also holds for portfolio choice models.

Apart from the precautionary motive, other reasons may be and presumably are present which also account for a great deal of the wealth accumulation. The two most prominent long run motives certainly are the wealth accumulation for old age (which I argued before can also be viewed as a long run precautionary motive), and the bequest motive. But though a substantial amount of empirical evidence contradicts the predictions of the life-cycle model, this does not necessarily support the latter motive necessarily. Still, especially in Germany, the life-cycle profile of discretionary household saving is rather flat, much flatter than, e.g., in the US. The question emerges why saving remains positive in old age, even for most low income households. This is particularly puzzling given the generous pensions and health insurance in Germany, and still, German households do not seem to draw it down but even accumulate real and financial wealth. This is why Börsch-Supan *et al.* (2001) refer to that observation as the "German savings puzzle."

Another important theoretical insight concerning factors determining savings and wealth accumulation are time and risk preferences[91]. Even if households face the the same source and hight of risk, they might completely differently deal with it. Risk preferences can be inferred domain-specifically, see Weber *et al.* (2002). They also find that situational characteristics as well as person-centered characteristics jointly influence risk-taking. For precisely assessing preferences, a relatively large number of domain-specific risk questions is needed which normally prohibits the implementation of these questions in general purpose surveys. E.g. the *HRS* experimentally contained risk and preferences questions; the estimates of the coefficient of risk aversion, varies substantially, see Barsky *et al.* (1997).

Another threat to the validity of estimates is the possible self-selection into safe jobs. Kimball (1990) or Lusardi (1997) refer to this as the prudence motive[92]; this causes an endogeneity problem since people choose occupations on the basis of their degree of risk aversion.

[91] Or risk attitudes.

[92] Kimball (1990), p.54, gives 'the name "prudence" to the sensitivity of the optimal choice of a decision variable to risk..[] The term is meant to suggest the propensity to prepare and forearm oneself in the face of uncertainty, in contrast to "risk aversion", which is how much one dislikes uncertainty and would turn away from uncertainty if possible.' And, on the same page, in Footnote 4: 'In different contexts, "prudence" will have different meanings. In

5.3 Household savings and measures for precautionary savings

This section describes the measures used to determine the importance of precautionary savings. In brief, theory tells us that precautionary savings is mainly driven by income risk. I will compare the effects of three different sets of measures for identifying precautionary savings.

The first set are respondents' expectations concerning the development of three domains (Germany's economic situation, the own economic situation, and the own and partner's health situation); more specifically, included in *SAVE* are also questions concerning the near future (job, labor income, inheritances). Additionally, I use information about the income development over the past five years (level and fluctuations, measured on a scale from 1 to 5). These questions of the *SAVE* questionnaire were influenced by Kotlikoff (1989) who states that new surveys are needed covering two issues to empirically assess precautionary savings: (1) implicit family insurance agreements and (2) the extent of subjective uncertainty.[93]

The second set contains the job variables available for the household and construct risk classes. Lusardi (1997 and 1998) argues that jobs might be selected by risk preferences since risk averse households would have a higher probability choosing a safer job. In that case, the estimated coefficient would be biased downwards. A crude instrumental procedure would be to use regional information about unemployment rates, assuming that households do not choose the living region by its unemployment rate. While this assumption might be plausible for some countries, it is definitely problematic for Germany with its extreme east-west slants of employment and the still ongoing migration of younger households caused by that. Nevertheless, I will use this approach for comparability reasons to other papers.

A third set of measures are the direct questions for savings motives implemented in the *SAVE* questionnaire. I will analyze what household characteristics influence these motives, and in a second step, I will test whether these motives are a proper way to map savings and wealth accumulation.

5.3.1 Subjective measures

A different approach to disentangle the influence of savings motives are subjective measures capturing the individual assessment of the different motives. Since empirical work is in nearly no domain free from any problems, these measures also entail two types of potential problems, see Jürges (2001). The first one is misreporting of the motives, consciously or unconsciously. There might be errors in different stages of the cognitive process involved in answering to survey questions. Apart from having problems with allocating probabilities or importance weights to questions, there might also be the problem of privacy effects and social desirability, see Tourangeau *et al.* (2000) for an overview or

the paradigmatic example of the consumption-saving decision under uncertainty, "prudence" represents the intensity of the precautionary saving motive.'

[93] One of the first surveys covering a subjective probability question of earnings was the 1989 Survey of Household Income and Wealth (*SHIW*), run every two years by the Bank of Italy. It was established in 1965. It is a series of independent cross sections, including a small panel component. See Guiso *et al.* (1992). Another data set which implemented subjective probability questions was the Health and Retirement Survey, (*HRS*). These data were used in Lusardi (1998).

Stocké (2001) for a study from social psychology. The second problem might be the endogeneity of the wealth formation processes and the savings motives.

Expectations for the future which are held less concrete than income questions were also included in the *SAVE* questionnaire. They cover three domains, Germany's economic development (macro level income), own economic development, own and partner's health development. Figure 5.6 shows the histograms for each of the four variables. Partner's and own health development are nearly congruent; if health is age-dependent and the age differential between couples is not too large, this is also rational. Since the partner does not file for that question himself, a more simple explanation is that respondents simply assign the same value for both if the health differential is not too large. A little bit more striking is the pessimism concerning Germany's development compared to the own one, or putting it differently, the overconfidence for the own situation rating macro and micro risks differently on a larger scale.

As mentioned in Section 5.1, the *SAVE* data contain, additionally to objective measures, expectation measures for different domains. Concerning earnings expectations, two questions are included. The first one asks for the the self-reported probability[94] of a net income raise in comparison to the previous year for the respondent and his/her partner. The second question asks for expectations concerning the employment situation (exactly, about how probable it is that respondent and/or partner will become unemployed in next year). Additionally, as a third variable, I use the self-assessed variable measuring the probability that respondents receive an inheritance in the following two years.

Since the *SAVE* questionnaire measures income and savings on a household basis, the single probabilities for respondents and partners have to be combined to the joint probability that at least one of self/partner is affected:

$$\begin{aligned} prob(X)_{household} &= prob(X)_{respondent} \vee prob(X)_{partner} \\ &= 1 - \{[(1 - prob(X)_{respondent})] \times [(1 - prob(X)_{partner})]\} \end{aligned} \tag{5.6}$$

Table 5.2: Subjective probabilities for job loss, income raise, and inheritances

	Job loss[c]		Income raise		Inheritance within 2 years	
	Respondent	Partner	Respondent	Partner	Respondent	Partner
Zeros	61.0% (594)	62.0% (490)	65.8% (1438)	66.5% (936)	89.6% (1956)	89.8% (1281)
Refusals	0.5% (5)	1.0% (8)	0.87% (19)	1.3% (19)	0.6% (12)	0.6% (8)
Mean perc.	14.6	12.9	13.0	12.5	3.5	3.6
N[a]	970[b]	782[b]	2165	1408	2172	1419

[a] Number of reported nonmissing values.

[b] If respondent or partner is not at least part-time employed and retired or unemployed, zero values will be imputed for further use.

[c] This question is identical to the one asked in the *HRS* and applies only if respondent or partner is at least partially employed.

The variance of net income can directly be computed by the household measure (see Equation 5.6) using the corresponding values from the first two columns of Table 5.2 and the income variance expression $p(1-p)(1-\alpha)^2Y^2$ where α is the replacement rate in case of job loss, which is 67% if

[94] Probabilities were given as 10%-steps on a scale from 0 to 100%.

person has at least one child and 60% without children corresponding to the definition of children after the tax law (§32 EStG).[95]

The other two variables from Table 5.2 will be used directly as a crude proxy for some positive income risk, since it cannot be linked directly to current household's income. They do not contain information on the amount of additional permanent income (income raise) or wealth (inheritances), just on the probabilities.

Table 5.3 contains the corresponding values from the mentioned variables measuring respondents' expectations for three domains.[96]

Table 5.3: Expectations for economic and health situation development

	Economic Development		Health Development	
	Germany	Own	Respondent	Partner
Zeros	17.7% (387)	5.0% (110)	1.8 % (39)	1.1% (15)
Refusals	0.6% (13)	0.7% (15)	0.6% (14)	1.0% (14)
Median value[a]	3	5	7	7
N	2171	2169	2170	1413

[a] Median instead of mean values are reported since values are measured on an ordinal scale.

Table 5.4 lists values for the measures on past income development.[97] These two questions have been asked within the drop-off part of the questionnaire.

Table 5.4: Assessment of past income development

	N	Percent		N	Percent
significantly better	215	10.27	Fluctuate significantly	535	25.8
slightly better	363	17.34	Fluctuate slightly	868	41.85
about the same	731	34.91	Not fluctuate at all	671	32.35
slightly worse	415	19.82			
significantly worse	370	17.67			
Refusals	25			45	

[95] Of course, this is only an approximation since these replacement rates can only be claimed for a certain time horizon (6-32 months depending on years of contribution payment). Replacement rates drop to 57% / 53% of a generalized net income after that time period *Arbeitslosenhilfe*. This might cause biased estimates of the motive since the variance would be measured too low.
For simplification reasons and since the *SAVE* data set does not contain information on the age of the children, I define having children in the sense of the law if at least one child still lives in the same household. Also, since there are no information on the duration of unemployment in the data set, I settle for the *Arbeitslosengeld* (67% / 60%) replacement rates.

[96] The exact wording of that question was: "We would now like to know a little about your views on future developments. Please indicate, according to a scale of 1 to 10. 0 means very negative 10 means very positive. (a) The economic development of Germany (b) Your own financial situation (c) Your own health situation (d) The health situation of your partner"

[97] The exact wording was: "Is your income situation, compared with five years ago (a) significantly better (b) slightly better (c) about the same (d) slightly worse (d) significantly worse?" This question was followed by: "During the last five years, did your income (a) Fluctuate significantly (b) Fluctuate slightly (c) Not fluctuate at all?"

In Table 5.5 Ordered Probit results for the unemployment probabilities are shown. I included respondent's characteristics for partner's probabilities to check whether they might be connected or even dominate partner's characteristics in their explanatory power. This is obviously not the case. Past unemployment is highly significant, especially in the regression for partner's probabilities.

There is one major problem with the reported subjective probabilities of unemployment, income change and inheritances, and that is the time horizon. In the first two cases (unemployment and income change), the time horizon is only about 6 months (since data were collected in June 2003, and the questions ask for changes until the end of the current year), while for inheritances, the time horizon is 2 years. First of all, respondents might have problems of adapting exactly to the time horizon given. Secondly, the short time horizon might make it difficult to extend the measure of uncertainty to human wealth uncertainty (see Lusardi (1997)). Thirdly, zero values were often given as answers to that question which might be due to the fact that most labor contracts are already determined for the given time period and respondents don't face any uncertainty for the given period. Guiso *et al.* (1992) propose two assumptions to estimate the effect of uncertainty on consumption and wealth accumulation: (1) the degree of persistence in the income generating process is identical for all households and (2) the probability distribution from which earnings are drawn is time-invariant. The second argument, though, might interact with the objection mentioned above that labor contracts might already be determined for the concerning time period, but undergo further negotiations in different points of time.

5.3.2 Environment measures

Apart from the subjective measures in Section 5.3.1, it is possible to construct a risk variable constructed by the job risk information contained in the job variables available in *SAVE*. Lusardi (1997) shows that saving rates are more or less independent from occupations using the 1989 *SHIW* data. Table 5.6 shows Tobit and OLS regression results for gross savings and saving rates[98]. Interestingly, having a riskier occupation like freelancers or a safer like civil servants do not show different savings behavior from the basic occupation category, employees. This is not the case for self-employed respondents, whose saving rates are significantly higher. This might be due to the fact of a higher polynomial degree of the income function which is not mapped here. I therefore also included a dummy variable for net household income larger than 10,000 € which turned out to be negatively significant, but it does not take away explaining power of the self-employed dummy. I also controlled for limited job contracts (not reported in Table 5.6), but the coefficient is totally insignificant.

The findings of Table 5.6 are generally in line with results from Lusardi (1997), Skinner (1988), and Jappelli and Pagano (1994).

These results suggest that a construction of risk index by occupation characteristics is little promising. I instead follow the procedure in Lusardi (1997) which in turn was motivated by Carroll (1992) and Engen and Gruber (2001) to use regional information on unemployment since unemployment is

[98] Gross savings were the direct savings measure given by the respondents and not corrected for net credit uptake / downpayments for the reason of otherwise higher data loss.

one of the major sources for income variation. Table 5.8 lists the available unemployment data for the 16 German states. There are two minor differences to the *SAVE* data states: (1) Berlin is, in contrast to official data, still separated in Berlin-West and Berlin-East in *SAVE*. (2) Rheinland-Pfalz and Saarland are pooled in *SAVE*, so I used the average of the unemployment rates weighted by unemployed persons for these two states (which then is 8.1%). The income variance will be constructed as in Section 5.3.1 ($p(1-p)(1-\alpha)^2Y^2$) where p is the states' unemployment rate. p will be set to zero for respondents who are either civil servants, retired or otherwise not working (unemployed, housewife, student etc.).[99]

The implied restriction in the use of these index unemployment rates is that individuals are equally affected by this unemployment risk.

5.3.3 Savings motives

Another set of possible variables to explaining savings and thereby wealth accumulation might be preferences, not only with regard to risk exposure, but also by the curvature of the utility function itself. Of course, determining risk preferences and the utility function is quite a hard task; individual heterogeneity causes the curvature of utility functions to vary between people. Schunk and Betsch (2004) suggest that there exists a relationship between the mode in which a person usually makes a decision and the curvature of the individual utility function. Their results suggest that individually stable traits might help explain observed economic behavior, such as portfolio choice and stock market decisions. Contributions in economic and psychological literature have investigated the question how people resolve decision problems under risk and uncertainty for quite a long time. Starmer (2000) gives a comprehensive review of the evolvement of the different approaches of both fields and their raising approach; see, e.g., the emergence of the whole field of behavioral finance, especially the prospect theory (Kahnemann/Tversky, 1979 and 1992), and their importance highlighted again by Laibson and Zeckhauser (1998).

As this discussion shows, the determination of individual utility functions is not trivial, though important. As a proxy for the utility function, the *SAVE* questionnaire also includes direct questions concerning the importance of a list of nine different savings motives.

Figure 5.5 shows histograms for the nine different savings motives. The bimodality is less pronounced for the saving motives precautionary savings ("unforseen events") and old-age provision (especially compared to "home"). The high fraction of zeros for repaying debts accounts for non-indebted households which in turn rate the motive for repaying debts low. Saving for buying a home is very bimodal; either households are interested in buying a house and therefore rate this motive high or mutatis mutandis very low.

How do respondents evaluate the savings motives and how do they compare to other findings? This is shown in Tables 5.9 - 5.11, where results are listed from ordered probit regressions for the nine listed savings motives. I do not comment each single regression or parameter estimate, since I will focus in the following on the two motives of interest in this context: the short run and long run precautionary

[99] The results from the *subjective* probabilities for p will be discussed in Section 5.4.

savings (unforseen events and old-age provision). Only one comment to the regression for leaving bequests: it it obvious that the the parameter estimate for children is highly significant reflection the accentuated bimodality of this saving motive which certainly inheres in this motive's nature.

Business owners or older households were found significant when identifying the precautionary motive in Kennickell and Lusardi (2004), are *not* likely to allocate higher values to the precautionary savings motive (cf. Table 5.9). In fact, the dummy for business owners is not significant but for the regression of the bequest motive.

Kennickell and Lusardi (2004) use the 1995 and 1998 Survey of Consumer Finances. Included in this survey is also a question for savings motives; respondents are given a list of 12 savings motives (and additional items "no reason" and "have no money to save"), but in difference to *SAVE*, respondents should only name the most important savings reason. These were the "emergency", or precautionary motive (36.2%), and the old-age provision motive (32.4 %). *SAVE*: even though a direct comparison is not possible, since *SAVE* only provides an ordinal measurement of the savings motives, the results are very similar: the precautionary motive has, on average, the highest importance (on a scale from 0 to 10), followed by the old-age provision motive.

As described in Section 5.3.1, I will interact the savings motives with the financial decision maker dummy.

5.4 Estimation results for precautionary savings measures

This section presents results for each of the mentioned three groups of variables. Dependent variables are saving rates, savings and a relative wealth measure. In addition to disentangle the more 'abstract' or long-rung saving goals from short ones for consumption reasons, I add another constructed variable to each of the regressions to control for this savings motive. In the *SAVE* questionnaire there is small set of questions included whether households seek to reach a definite savings goal.[100] It was then asked how large the desired saved amount is and by when it should be reached. The control variable I constructed from these three variables is whether a household has a savings goal, whether it should be reached within the next two years and whether its value is below € 20,000 to catch all planned larger expenditures (including holidays, cars etc.).

Concerning the use of the scaled variables discussed in Section 5.3, I circumvent the problem of ordinal measurement of the relevant variables by building three classes, the low one ranging from 0 to 2, the middle one from 3 to 7 and the high one from 8 to 10.[101] This procedure was not chosen for the probability questions in which case I interpret the percentage classes cardinally to calculate the household probabilities.

I interacted the household probability for receiving inheritances within the next two years with the follow-up question whether this inheritance would at least slightly improve the income situation.

[100] The exact wording of that question was: "Do you or your partner currently have a fixed objective in mind for which you are saving at least 500 €? If yes, what it is your objective?" Nonresponse is very small (1.1%).

[101] This classification was chosen such that at the low and high class comprise the same amount of values. See Börsch-Supan and Essig (2003) who used the same classification.

This means that I will have 43 positive probabilities in the sample. In all specifications for wealth or financial wealth accumulation and saving rates, the expectation of significant inheritances is not significant. This can be linked to the discussion between Barro and Thaler (1990) whether agents rationalize and privately offset different wealth sources. In this case, the expectance of a windfall income like inheritances or bestowals should ceteris paribus reduce the wealth accumulation, which is not the case with the data at hand.

5.4.1 Estimation procedure

Dependent variables

For each of the three set of explanatory variables, I check their influence on saving rates, which is a rather short-run measure since it is only measured at one point in time and represents no accumulation over time; the second dependent variable is financial wealth relative to permanent income, and the third one is total wealth over permanent income, both latter dependent variables being transformed by the inverse hyperbolic sine function with $\theta = 1.05$.[102] For saving rates and relative financial wealth, I choose the Tobit regression model since only positive values are observed. Total wealth can be negative if the household is indebted.

Independent variables common for all regressions

In Tables 5.12 - 5.23, the first set of variables is unchanged for all regressions. Two things are worth noting. First, there is a pronounced age and income pattern. This is in line with the theoretical literature;[103] second, past unemployment, as a further risk proxy variable, is negative significant in all specifications. Third, the control variable whether the household aims achieving a savings target also has explanatory power in all regressions. Fourth, as proxy variable for risk preference or tolerance, whether or not the household has a private occupational disability insurance, is significantly positive in most of the regressions. Fifth, occupation variables are only partly significant for self-employed. Civil servants, who face a significant lower occupation and income risk, do not show different financial behavior to employees.

5.4.2 Subjective measures

Development variables

Tables 5.12 - 5.23 show that the macro level expectations (Germany's economic development) provide insignificant results for all dependent variables regressions but for saving rates where pessimistic beliefs lead to *smaller* saving rates. In contrast, the own development beliefs are linked to financial behavior. Optimistic respondents show higher saving rates and higher financial wealth accumulation, while the opposite is true for pessimistic respondents. This is clearly in contrast to the classical vision of any precautionary savings argument from Section 5.2 as long as the causality is correctly captured

[102] Cf. Section 5.2.2.

[103] Keeping in mind, however, that the findings are based on cross-sectional data thereby ignoring cohort effects.

here. If, in contrast, respondents belief that the own economic situation will be better in the future *since their financial background due to wealth accumulation* is better, then estimates are biased. In some specifications, the same is true for respondents' and partners' health situation expectations. If these questions are judged more optimistically, this gives raise to higher saving rates and higher wealth accumulation. This finding would support the hypothesis by Börsch-Supan and Stahl (1991) that households might be consumption restricted due to age-related health problems. If in contrast, households do not feel that these health restrictions might take place (or not *that* badly as primarily expected), wealth accumulation should be higher to provide funds for the unrestricted consumption. On the other hand, a high wealth accumulation could open possibilities for better health care measures leading causality in the opposite direction. Given the (still) generous German health care system, I discard the latter hypothesis.

Income uncertainty

The effect of income variation and development variables and their effect on the set of dependent variables are listed in Tables 5.15 - 5.17. These explanatory variables are partially only significant in the saving rates regression. Again, results are counter-intuitive. If income development was positive in the past five years, saving rates are *higher*, and also, if income was highly volatile and therefore more risky, saving rates are *lower*. If a household would face significant changes to its economic situation if, e.g., one or both members would finish schooling and start their working life, this would be comprehensible. In nearly all other cases, these findings would contradict the above mentioned theoretical findings.

The variance of net income, which proved to be significant in Lusardi (1998), is insignificant for all three tested dependent variables. An explanation for the different findings might be that in Germany, the shock on the permanent income might be lower in the Germany than in the U.S., since the replacement rate, α, is higher in Germany. The so-called *Hartz IV - reform* can in this case been seen as a natural experiment. This also points the way ahead to reconsider this variable with the then available *SAVE* 2005 data.

Job risk

The proxy variable for job risk measure, the state unemployment rate, is significant in the Tobit regression for saving rates (see Table 5.18), but has only a very small effect. In contrast, it is insignificantly negative for relative financial and total wealth (Tables 5.19 and 5.20).

Savings motives

Tables 5.21 - 5.23 show Tobit and OLS estimates for saving rates, relative financial and total wealth. The precautionary savings motive is, in all three regressions, strongly negative significant if the motive is rated low; when it is highly rated, respondents do not behave statistically different to the median rated motive group. The old-age provision motive is insignificant in all regressions. Interestingly,

the self-estimation of living longer than some self-estimated average does not alter financial behavior significantly.[104].

All variables

Table 5.24 includes all discussed sets of subjective variables. The identical block of other variables used in previous regressions was also applied here, but is not shown in the table for obvious reasons. The basic patterns prove to be stable when including the additional sets of variables. Still, the savings motive for unforseen events is, when ranked high, now significant in the saving rates regression.

5.5 Conclusions

The capture of a short-run or long-run precautionary savings motive is empirically a hard task. I approached this challenge by three different sets of variables.

1. Expectations for the future: A negative evaluation of the own or the economy's economic future situation or health situation might be the reason for households' increased need to insure themselves - by higher savings or capital formation - to mitigate a negative development. The results contradict this hypothesis. Households with low expectations for their own economic development have significantly lower saving rates, and relative financial and total wealth.
   Expectations capturing the job risk (by building a variance variable which measures the household's risk of unemployment) show no significant effects, while income development variables support the findings for economic development expectations: a *positive* development increases saving rates, while a *higher* volatility of past income reduces saving rates.

2. Environmental situation: Substituting the expectations variable for job risk by local unemployment rates (which probably are not disaggregated enough) shows a positively significant effect for saving rates, controlling for East Germany. As for occupation variables, which are included in any regression specification, self-employed have a higher saving rate. Attribution this to job risk seems difficult for two reasons. First, the dummy for self-employed might only produce a spline for the influence of income since typically self-employed have a monthly income above the median. Second, a dummy for retired households is also positive - the income risk for retirees rather lies in the development of the Social Security system. One could argue, though, that retirees have higher saving rates for health risks, but as seen in the specifications including expectations for health development, the health expectations dummies are insignificant while the dummy for retirees maintains positive significant.

3. Savings motives: The short-run precautionary savings motive for unforseen events shows the expected negative coefficient if the motive is ranked unimportant, but is insignificant when ranked

[104] The relative life expectancy is the following. 65.1% belief to live as long as the average, 22.4% belief to live longer and 12.5% to live shorter than the average. These numbers only slightly compare to surveys about self-judgement for driving capabilities or being a good professor (better than the average) where these numbers typically reach about 90%

important[105]. The long-run precautionary motive, old-age provisions, is positive significant when ranked important in the saving rates regression, but insignificant in the other two regressions (relative financial and total wealth). Expectations concerning the relative life expectancy have unexpected coefficients or are insignificant.

By using three different dependent variables, one can see that the evaluation of the precautionary saving motive is not homogeneous. Within one set of independent variables, the coefficients change when applying each set in the estimation of the three variables saving rate, relative financial and relative total wealth.

[105] relative to the median group

5.A Tables

Table 5.5: Ordered probit results for future unemployment probabilities

	Prob. job loss respondent		Prob. job loss partner	
	Coef.	$P > z$	Coef.	$P > z$
Respondent				
Permanent income / 10000	**-3.436**	0.006	**-3.551**	0.052
Permanent income / 10000 sq.	2.475	0.142	3.266	0.146
Age / 10	-2.162	0.479	-6.273	0.141
Age / 10 sq.	0.780	0.456	2.134	0.132
(Age / 10) cub.	-0.121	0.423	-0.308	0.123
$(Age/10)^4$	0.006	0.412	0.016	0.122
# inc. sources	-0.043	0.314	0.017	0.744
Secondary school (D)	-0.095	0.362	-0.056	0.667
Graduation diploma (D)	-0.208	0.166	-0.228	0.238
University degree (D)	-0.076	0.608	0.004	0.984
Kids (D)	**0.409**	0.005	0.117	0.522
Kids living in same house (D)	-0.211	0.102	-0.035	0.813
East Germany (D)	**0.228**	0.045	0.114	0.409
Job: blue collar (D)	0.081	0.496	0.068	0.688
Job: civil servant (D)	**-0.960**	0.000	0.008	0.973
Job: freelancer (D)	0.254	0.408	0.550	0.190
Job: self-employed (D)	-0.462	0.015	-0.050	0.822
Work parttime (D)	-0.111	0.379	0.050	0.725
Work little (D)	**-0.479**	0.010	-0.105	0.566
Female (D)	0.080	0.434	4.142	0.593
Past unemployment 1-6 months	**0.287**	0.009	0.171	0.189
Past unemp.> 6 months	0.047	0.705	-0.117	0.422
Partner				
# inc. sources			**-0.140**	0.090
Age / 10			-3.729	0.597
Age / 10 sq.			1.189	0.604
(Age / 10) cub.			-0.153	0.627
$(Age/10)^4$			0.007	0.654
Secondary school (D)			0.047	0.773
Graduation diploma (D)			0.010	0.972
University degree (D)			0.044	0.845
Job: blue collar (D)			-0.142	0.364
Job: civil servant (D)			**-1.339**	0.007
Job: freelancer (D)			-0.685	0.295
Job: self-employed (D)			-0.082	0.707
Work parttime (D)			-0.192	0.776
Work little (D)			-7.053	1.000
Past unemployment 1-6 months			**0.412**	0.003
Past unemp.> 6 months			**0.335**	0.028
Number of obs	887		666	
LR chi2(23)	139.4		111.12	
$Prob > chi2$	0		0	
Pseudo R2	0.0549		0.0585	
Log Likelihood	-1199.9729		-894.89974	

Note: The probability questions were only asked for respondents and their partners if they were fully, partly or little employed.

Table 5.6: Regression results for saving rates and savings on age, income and job variables

	Tobit estimates				OLS estimates			
	Saving rates		Savings		Saving rates		Savings	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net HH income/10'	-0.021	0.372	**2756.702**	0.001	**-0.064**	0.007	**3337.686**	0.000
Net HH income/10' sq.	0.001	0.778	**-290.777**	0.004	**0.005**	0.081	**-326.019**	0.001
Age/10	0.039	0.108	**1580.671**	0.065	0.018	0.485	789.069	0.403
Age/10 squared	-0.004	0.122	**-145.323**	0.079	-0.002	0.356	-81.014	0.368
Secondary school (D)	**0.028**	0.078	**1573.795**	0.006	-0.001	0.946	955.604	0.117
Graduation diploma (D)	**0.086**	0.000	**2803.431**	0.001	**0.045**	0.057	1418.290	0.106
University degree (D)	**0.090**	0.000	**4659.829**	0.000	**0.053**	0.007	**3908.034**	0.000
Partner (D)	**0.064**	0.000	**2833.860**	0.000	0.011	0.480	**1521.987**	0.009
Kids (D)	0.010	0.634	581.270	0.432	0.018	0.379	867.396	0.264
Kids living in same house (D)	-0.044	0.015	**-1335.333**	0.040	-0.028	0.123	-702.688	0.304
East Germany (D)	-0.018	0.259	**-1180.589**	0.043	0.003	0.845	-714.526	0.256
Female (D)	-0.001	0.930	-166.434	0.744	-0.010	0.487	-418.612	0.437
Job: blue collar (D)	0.004	0.879	109.641	0.904	0.004	0.876	464.205	0.617
Job: civil servant (D)	0.051	0.136	**2720.648**	0.023	0.006	0.847	1128.355	0.314
Job: freelancer (D)	0.049	0.471	**4336.068**	0.069	0.080	0.214	**6806.315**	0.005
Job: self-employed (D)	**0.142**	0.000	**6135.093**	0.000	**0.172**	0.000	**7371.621**	0.000
Retired(D)	**0.068**	0.025	**2415.502**	0.025	0.022	0.490	1008.425	0.395
Work parttime (D)	-0.030	0.295	-1059.731	0.290	0.031	0.273	1143.156	0.284
Work little (D)	**-0.071**	0.021	-1677.571	0.121	0.019	0.558	1599.189	0.184
Work not (D)	**-0.074**	0.004	**-2564.480**	0.004	0.010	0.709	-87.758	0.929
Unemployed (D)	**-0.067**	0.022	**-2168.852**	0.037	0.008	0.809	138.344	0.912
Constant	**-0.123**	0.038	**-7419.495**	0.000	0.092	0.142	-1492.492	0.523
Number of obs	1751		1751		1005		1005	
left-censored	746		746					
uncensored	1005		1005					
LR chi2	159.740		258.730					
$Prob > F$	0		0.000		0.000		0.000	
Pseudo-R sq. / R sq.	0.150		0.012		0.050		0.124	
Log likelihood	-451.995		-10940.126					
F(21, 983)					2.480		6.600	
Adj. R sq.					0.030		0.105	
Root MSE					0.195		7326.500	

Table 5.7: Regression results for permanent income imputation

	Net income		ln(Net income)	
	Coef.	$P > t$	Coef.	$P > t$
Respondent				
Salary index partner	4.202	0.254	-0.001	0.748
Household size	**179.115**	0.000	**0.102**	0.000
Age / 10	-77.808	0.966	1.155	0.146
Age / 10 sq.	418.941	0.366	-0.266	0.187
(Age / 10) cub.	-76.241	0.200	0.024	0.355
$(Age/10)^4$	3.855	0.167	-0.001	0.528
# inc. sources	**77.205**	0.010	0.018	0.169
Secondary school (D)	**339.235**	0.000	**0.185**	0.000
Graduation diploma (D)	**264.785**	0.014	**0.109**	0.020
University degree (D)	829.677	0.000	**0.368**	0.000
Kids (D)	**189.060**	0.054	0.068	0.111
Kids living in same house (D)	**-517.452**	0.000	**-0.205**	0.000
East Germany (D)	**-364.048**	0.000	**-0.210**	0.000
Job: blue collar (D)	**-221.743**	0.048	-0.084	0.087
Job: civil servant (D)	638.495	0.000	**0.214**	0.002
Job: freelancer (D)	212.344	0.472	0.013	0.922
Job: self-employed (D)	**530.715**	0.001	**0.131**	0.062
Retired(D)	130.125	0.382	**0.150**	0.021
Work parttime (D)	**-405.935**	0.001	**-0.229**	0.000
Work little (D)	**-558.564**	0.000	**-0.373**	0.000
Work not (D)	**-602.099**	0.000	**-0.395**	0.000
Unemployed (D)	-198.885	0.103	**-0.145**	0.007
Widowed (D)	**-461.373**	0.000	**-0.269**	0.000
Separated or divorced (D)	**-465.760**	0.000	**-0.250**	0.000
Partner				
Partner (D)	-6457.766	0.518	-3.648	0.402
Salary index partner	1.414	0.849	0.000	0.982
# inc. sources	144.317	0.029	0.039	0.172
Age / 10	5755.009	0.335	3.805	0.143
Age / 10 sq.	-1978.272	0.262	**-1.316**	0.087
(Age / 10) cub.	292.717	0.221	**0.189**	0.070
$(Age/10)^4$	-15.188	0.197	**-0.009**	0.065
Secondary school (D)	141.893	0.301	0.053	0.371
Graduation diploma (D)	169.040	0.474	0.093	0.367
University degree (D)	**1027.127**	0.000	**0.263**	0.001
Job: blue collar (D)	**-258.862**	0.074	-0.091	0.147
Job: civil servant (D)	238.089	0.371	0.067	0.563
Job: freelancer (D)	**-926.080**	0.059	-0.283	0.187
Job: self-employed (D)	**671.694**	0.001	0.085	0.327
Retired(D)	521.327	0.276	0.212	0.310
Work parttime (D)	-379.482	0.650	-0.043	0.906
Work little (D)	-493.868	0.378	**-0.521**	0.033
Work not (D)	**-1278.834**	0.002	**-0.554**	0.002
Unemployed (D)	663.911	0.128	0.217	0.253
Constant	-1606.059	0.699	**5.867**	0.001
Number of obs	1694		1694	
F(33, 1100 / F(20, 661)	22.450		31.390	
Prob > F	0.000		0.000	
R squared	0.3691		0.45	
Adj. R sq.	0.3526		0.4356	
Root MSE	1155.4		0.50335	

Notes: Conditional regression that at least one household member is not yet retired. In 18 cases, monthly income was most probably mixed up with yearly income by the respondent when respondent's occupation was blue collar worker with low schooling and was thus divided by 12.

Table 5.8: Unemployment information for German Bundesländer (states); average values for 2003

State	# Unemployed	Unemployement rate	Open jobs	Short-time workers
Baden-Wrttemberg	336,540	6.1	49,022	34,623
Bayern	447,349	6.9	56,863	26,991
Berlin	306,462	18.1	9,291	4,485
Brandenburg	253,028	18.8	9,125	5,675
Bremen	42,366	13.2	3,411	1,653
Hamburg	86,388	9.9	7,633	2,032
Hessen	242,059	7.9	25,989	17,651
Mecklenburg-Vorpommern	181,710	20.1	7,484	2,939
Niedersachsen	379,811	9.6	34,444	13,936
Nordrhein-Westfalen	880,053	10.0	65,394	47,205
Rheinland-Pfalz	154,610	7.7	27,308	9,170
Saarland	47,718	9.5	4,953	2,495
Sachsen	403,529	17.9	17,063	10,641
Sachsen-Anhalt	268,293	20.5	9,795	4,613
Schleswig-Holstein	136,159	9.7	10,771	4,740
Thringen	210,693	16.7	10,115	6,524
Germany	4,376,767	10.5	354,762	195,371

Source: Bundesagentur für Arbeit

Table 5.9: Ordered probit regression results for saving motives: part 1

	Buying a home		Unforseen events		Paying off debts	
	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Perm. income / 10,000	**3.380**	0.001	**2.648**	0.002	**1.687**	0.079
(Perm. income / 10,000) sq.	**-4.835**	0.001	**-3.063**	0.008	**-2.206**	0.100
Two income earners (D)	**0.320**	0.000	0.027	0.704	**0.221**	0.004
Age/10	**-0.375**	0.000	0.075	0.426	0.055	0.598
Age/10 squared	**0.028**	0.006	-0.007	0.454	-0.013	0.215
Secondary school (D)	**0.103**	0.122	**0.164**	0.007	0.050	0.446
Graduation diploma (D)	**0.206**	0.030	0.104	0.243	-0.098	0.305
University degree (D)	**0.171**	0.057	0.062	0.451	0.103	0.245
Kids (D)	-0.091	0.309	0.027	0.745	0.063	0.480
Kids living in same house (D)	0.108	0.164	-0.070	0.323	0.078	0.300
East Germany (D)	**-0.457**	0.000	**-0.145**	0.021	**-0.191**	0.005
Job: blue collar (D)	0.044	0.681	0.019	0.855	0.012	0.911
Job: civil servant (D)	-0.113	0.457	-0.048	0.739	**-0.270**	0.075
Job: freelancer (D)	0.089	0.740	-0.076	0.762	-0.148	0.576
Job: self-employed (D)	0.055	0.725	0.007	0.965	-0.180	0.243
Work parttime (D)	**-0.208**	0.069	-0.035	0.747	0.029	0.798
Work little (D)	-0.067	0.582	-0.054	0.638	**0.249**	0.039
Work not (D)	0.197	0.063	0.034	0.732	0.074	0.486
Unemployed (D)	**-0.299**	0.008	-0.127	0.224	0.107	0.327
Female (D)	-0.043	0.480	0.067	0.222	-0.004	0.943
Partner	0.112	0.201	0.015	0.853	-0.050	0.567
Widowed (D)	0.039	0.718	-0.157	0.118	**-0.234**	0.029
Separated or divorced (D)	-0.032	0.748	**-0.296**	0.001	-0.092	0.345
Retired(D)	**-0.386**	0.002	-0.027	0.813	**-0.352**	0.004
Business owner (D)	-0.072	0.565	0.041	0.731	0.175	0.159
Number of obs	1957		1966		1935	
LR chi2(23)	357.94		98.41		222.1	
Prob > chi2	0.000		0.000		0.000	
Pseudo R2	0.0496		0.0118		0.0299	
Log Likelihood	-3429.6856		-4121.3349		-3603.2229	

Table 5.10: Ordered probit regression results for saving motives: part 2

	Old-age provisions		Traveling		Major purchases	
	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Perm. income / 10,000	0.703	0.424	**4.228**	0.000	**2.846**	0.001
(Perm. income / 10,000) sq.	-0.273	0.820	**-5.310**	0.000	**-3.780**	0.002
Two income earners (D)	0.059	0.415	**0.142**	0.047	-0.029	0.688
Age/10	**0.295**	0.002	-0.083	0.388	**-0.193**	0.043
Age/10 squared	**-0.029**	0.002	-0.003	0.780	0.008	0.358
Secondary school (D)	**0.185**	0.003	**0.185**	0.002	**0.186**	0.002
Graduation diploma (D)	0.051	0.569	**0.365**	0.000	**0.217**	0.014
University degree (D)	**0.142**	0.092	**0.317**	0.000	0.121	0.139
Kids (D)	0.057	0.501	0.073	0.378	-0.022	0.786
Kids living in same house (D)	**-0.227**	0.002	**-0.216**	0.002	-0.069	0.325
East Germany (D)	**-0.172**	0.007	**-0.153**	0.016	**-0.339**	0.000
Job: blue collar (D)	-0.071	0.488	**0.195**	0.051	0.060	0.547
Job: civil servant (D)	**-0.452**	0.002	**0.249**	0.081	-0.052	0.710
Job: freelancer (D)	-0.056	0.826	-0.079	0.751	0.058	0.812
Job: self-employed (D)	0.117	0.447	-0.121	0.416	-0.074	0.617
Work parttime (D)	-0.008	0.945	0.135	0.204	0.023	0.827
Work little (D)	-0.093	0.418	-0.144	0.202	0.007	0.948
Work not (D)	-0.093	0.356	0.066	0.503	-0.131	0.184
Unemployed (D)	-0.108	0.302	**-0.354**	0.001	**-0.235**	0.024
Female (D)	-0.028	0.615	0.006	0.912	-0.029	0.599
Partner	0.076	0.342	-0.138	0.081	**0.179**	0.024
Widowed (D)	**-0.177**	0.081	0.091	0.364	0.142	0.157
Separated or divorced (D)	**-0.243**	0.008	-0.138	0.130	-0.042	0.648
Retired(D)	**-0.450**	0.000	0.065	0.562	-0.024	0.828
Business owner (D)	0.070	0.567	-0.192	0.104	-0.041	0.726
Number of obs	1953		1967		1966	
LR chi2(23)	202.4		246.36		273.58	
Prob > chi2	0		0		0	
Pseudo R2	0.0241		0.0278		0.0308	
Log Likelihood	-4094.5818		-4304.2849		-4303.7851	

Table 5.11: Ordered probit regression results for saving motives: part 3

	Subsidizing offspring		Leaving bequests		To receive tax subsidies	
	Coef.	$P > z$	Coef.	$P > z$	Coef.	$P > z$
Perm. income / 10,000	**2.096**	0.019	0.428	0.632	**3.638**	0.000
(Perm. income / 10,000) sq.	**-1.934**	0.108	0.317	0.792	**-4.939**	0.000
Two income earners (D)	**0.160**	0.029	0.090	0.228	**0.181**	0.015
Age/10	**-0.610**	0.000	**-0.570**	0.000	**-0.235**	0.026
Age/10 squared	**0.049**	0.000	**0.052**	0.000	0.005	0.596
Secondary school (D)	-0.053	0.398	-0.073	0.255	0.023	0.723
Graduation diploma (D)	**0.164**	0.076	**-0.191**	0.045	-0.015	0.871
University degree (D)	**0.190**	0.027	**-0.149**	0.089	0.017	0.846
Kids (D)	**1.069**	0.000	**0.988**	0.000	**0.180**	0.045
Kids living in same house (D)	**0.220**	0.002	**-0.226**	0.002	-0.068	0.368
East Germany (D)	**-0.129**	0.047	**-0.255**	0.000	**-0.297**	0.000
Job: blue collar (D)	-0.074	0.482	-0.076	0.482	0.123	0.236
Job: civil servant (D)	-0.065	0.666	-0.102	0.511	-0.182	0.226
Job: freelancer (D)	-0.049	0.855	-0.044	0.872	-0.402	0.143
Job: self-employed (D)	-0.133	0.388	0.000	0.999	0.041	0.787
Work parttime (D)	**0.225**	0.041	0.006	0.959	-0.067	0.544
Work little (D)	0.083	0.482	0.002	0.986	-0.008	0.946
Work not (D)	0.161	0.117	0.091	0.386	0.006	0.955
Unemployed (D)	-0.005	0.962	-0.084	0.448	-0.083	0.448
Female (D)	-0.018	0.757	-0.034	0.562	-0.079	0.176
Partner	**-0.216**	0.008	-0.106	0.198	-0.024	0.780
Widowed (D)	**-0.332**	0.002	**-0.304**	0.006	**-0.181**	0.094
Separated or divorced (D)	**-0.310**	0.001	**-0.309**	0.001	-0.025	0.800
Retired(D)	**-0.199**	0.086	**-0.253**	0.031	**-0.319**	0.008
Business owner (D)	0.182	0.130	**0.276**	0.021	-0.037	0.760
Number of obs	1943	1943	1938	1938	1942	
LR chi2(23)	496.77		286.4		361.09	
Prob > chi2	0		0		0	
Pseudo R2	0.0586		0.0365		0.0466	
Log Likelihood	-3988.6158		-3781.9001		-3690.9245	

Table 5.12: Regression results: development expectations and **saving rates**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	0.306	0.166		
(Perm. Inc./ 10,000) sq.	-0.242	0.408		
Age / 10	**0.050**	0.071	-0.004	0.912
Age / 10 sq.	**-0.005**	0.071	0.001	0.890
Secondary school (D)	0.004	0.814	0.023	0.237
Graduation diploma (D)	0.037	0.103	0.032	0.300
University degree (D)	0.032	0.142	**0.051**	0.057
Kids (D)	0.023	0.251		
Kids living in same house (D)	**-0.056**	0.002		
Separated or divorced (D)	**-0.040**	0.095		
Widowed (D)	-0.012	0.642		
East Germany (D)	-0.001	0.966		
Female (D)	-0.010	0.514		
Job: blue collar (D)	0.025	0.294	0.015	0.575
Job: civil servant (D)	0.016	0.623	-0.034	0.489
Job: freelancer (D)	0.033	0.589	-0.017	0.833
Job: self-employed (D)	**0.114**	0.002	**-0.102**	0.012
Retired (D)	**0.062**	0.036	-0.016	0.580
Work parttime (D)	0.009	0.735	-0.017	0.581
Work little (D)	-0.023	0.427	-0.039	0.304
Work not (D)	-0.040	0.103	-0.018	0.514
Unemployed (D)	0.017	0.563	-0.013	0.729
Unemp.> 1 month (D)	-0.001	0.948	-0.010	0.678
Unemp.> 6 months (D)	**-0.054**	0.006	0.013	0.623
Village (D)	-0.008	0.742		
Partner (D)	0.026	0.790		
Savings goal ahead (D)	**0.055**	0.021		
Business owner (D)	0.013	0.688		
Prob(inheritance)	0.025	0.456		
Occ. disab. insur. (D)	**0.061**	0.000		
Live shorter than av. (D)	**-0.009**	0.640		
Live longer than av. (D)	0.005	0.733		
Expectations (D)				
Low: Germany's ec. situation	**-0.021**	0.096		
High: Germany's ec. situation	-0.007	0.822		
Low: Own economic situation	**-0.113**	0.000		
High: Own economic situation	**0.045**	0.004		
Low: Own health situation	0.028	0.280		
High: Own health situation	0.021	0.160		
Low: Partner's health situation	-0.021	0.555		
High: Partner's health situation	-0.020	0.261		
Constant	-0.163	0.038		
Number of obs	1573			
uncensored obs	909			
LR chi2(68)	317.46			
$Prob > chi2$	0			
Pseudo R2	0.4074			
Log likelihood	-230.92823			

Table 5.13: Regression results: development expectations and **financial wealth/permanent income**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**11.826**	0.000		
(Perm. Inc./ 10,000) sq.	**-11.645**	0.000		
Age / 10	**0.560**	0.039	0.590	0.127
Age / 10 sq.	-0.035	0.164	-0.051	0.185
Secondary school (D)	**0.326**	0.049	**-0.352**	0.096
Graduation diploma (D)	0.255	0.290	-0.112	0.738
University degree (D)	0.151	0.501	-0.048	0.871
Kids (D)	0.140	0.494		
Kids living in same house (D)	**-0.602**	0.001		
Separated or divorced (D)	-0.340	0.144		
Widowed (D)	-0.245	0.358		
East Germany (D)	**-0.319**	0.047		
Female (D)	-0.233	0.114		
Job: blue collar (D)	0.020	0.937	-0.168	0.572
Job: civil servant (D)	-0.425	0.246	-0.156	0.758
Job: freelancer (D)	-0.562	0.383	-0.760	0.360
Job: self-employed (D)	-0.589	0.133	-0.605	0.144
Retired (D)	0.163	0.565	-0.016	0.955
Work parttime (D)	-0.326	0.300	0.021	0.951
Work little (D)	-0.393	0.193	0.099	0.804
Work not (D)	-0.444	0.083	-0.041	0.887
Unemployed (D)	-0.195	0.498	0.284	0.444
Unemp.> 1 month (D)	0.188	0.311	-0.058	0.828
Unemp.> 6 months (D)	-0.581	0.005	-0.082	0.780
Village (D)	0.044	0.865		
Partner (D)	**-1.637**	0.096		
Savings goal ahead (D)	**0.611**	0.016		
Business owner (D)	**0.862**	0.008		
Prob(inheritance)	**0.705**	0.053		
Occ. disab. insur. (D)	**0.376**	0.043		
Live shorter than av. (D)	-0.054	0.778		
Live longer than av. (D)	-0.037	0.806		
Expectations (D)				
Low: Germany's ec. situation	0.112	0.389		
High: Germany's ec. situation	-0.217	0.583		
Low: Own economic situation	**-0.839**	0.000		
High: Own economic situation	0.238	0.142		
Low: Own health situation	**-0.528**	0.034		
High: Own health situation	-0.194	0.212		
Low: Partner's health situation	-0.172	0.609		
High: Partner's health situation	0.173	0.362		
Constant	-1.721	0.029		
Number of obs	1140			
uncensored obs	431			
LR chi2(68)	402.4			
$Prob > chi2$	0			
Pseudo R2	0.1034			
Log likelihood	-1744.6798			

Table 5.14: Regression results: development expectations and **total wealth/permanent income**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**12.520**	0.000		
(Perm. Inc./ 10,000) sq.	**-12.950**	0.000		
Age / 10	**0.872**	0.003	**1.025**	0.015
Age / 10 sq.	**-0.050**	0.065	-0.068	0.108
Secondary school (D)	**0.573**	0.002	-0.129	0.592
Graduation diploma (D)	**0.472**	0.071	-0.472	0.215
University degree (D)	**0.571**	0.029	-0.075	0.826
Kids (D)	**-0.574**	0.010		
Kids living in same house (D)	0.053	0.791		
Separated or divorced (D)	**-0.424**	0.100		
Widowed (D)	-0.313	0.275		
East Germany (D)	**-0.428**	0.015		
Female (D)	**-0.340**	0.037		
Job: blue collar (D)	0.313	0.269	-0.449	0.179
Job: civil servant (D)	0.026	0.952	0.316	0.596
Job: freelancer (D)	-0.993	0.156	-0.353	0.712
Job: self-employed (D)	-0.185	0.707	0.391	0.451
Retired (D)	-0.345	0.270	-0.382	0.252
Work parttime (D)	0.297	0.394	0.302	0.448
Work little (D)	**0.679**	0.041	-0.245	0.589
Work not (D)	**0.555**	0.048	0.512	0.122
Unemployed (D)	**-0.813**	0.006	-0.123	0.763
Unemp.> 1 month (D)	-0.048	0.817	-0.324	0.290
Unemp.> 6 months (D)	**-0.386**	0.089	0.350	0.287
Village (D)	**1.157**	0.000		
Partner (D)	**-3.496**	0.001		
Savings goal ahead (D)	**0.713**	0.011		
Business owner (D)	**1.959**	0.000		
Prob(inheritance)	0.472	0.256		
Occ. disab. insur. (D)	0.179	0.399		
Live shorter than av. (D)	-0.215	0.307		
Live longer than av. (D)	**-0.367**	0.028		
Expectations (D)				
Low: Germany's ec. situation	0.216	0.134		
High: Germany's ec. situation	-0.693	0.112		
Low: Own economic situation	**-0.611**	0.002		
High: Own economic situation	**0.427**	0.019		
Low: Own health situation	**-0.422**	0.109		
High: Own health situation	0.128	0.458		
Low: Partner's health situation	-0.488	0.183		
High: Partner's health situation	0.100	0.644		
Constant	-1.754	0.037		
Number of obs	1016			
F(56, 959)	10.36			
Prob ¿ F	0			
R-squared	0.3769			
Adj R-squared	0.3405			
Root MSE	2.0339			

Table 5.15: Regression results: probability of job loss expectations and **saving rates**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**0.476**	0.033		
(Perm. Inc./ 10,000) sq.	-0.402	0.174		
Age / 10	0.043	0.128	0.000	0.993
Age / 10 sq.	-0.004	0.120	0.000	0.974
Secondary school (D)	0.012	0.448	0.022	0.252
Graduation diploma (D)	0.048	0.034	0.029	0.356
University degree (D)	0.042	0.049	0.052	0.052
Kids (D)	0.021	0.305		
Kids living in same house (D)	**-0.055**	0.002		
Separated or divorced (D)	**-0.045**	0.057		
Widowed (D)	-0.015	0.578		
East Germany (D)	-0.013	0.422		
Female (D)	-0.011	0.454		
Job: blue collar (D)	**0.040**	0.096	0.014	0.589
Job: civil servant (D)	0.003	0.936	-0.043	0.389
Job: freelancer (D)	0.049	0.425	-0.063	0.463
Job: self-employed (D)	**0.125**	0.001	**-0.107**	0.010
Retired (D)	**0.069**	0.018	-0.024	0.408
Work parttime (D)	0.000	0.994	-0.022	0.478
Work little (D)	-0.011	0.712	-0.039	0.308
Work not (D)	-0.028	0.253	-0.009	0.746
Unemployed (D)	0.003	0.915	-0.034	0.358
Unemp.> 1 month (D)	-0.002	0.895	-0.004	0.851
Unemp.> 6 months (D)	**-0.054**	0.006	0.014	0.605
Village (D)	-0.002	0.938		
Partner (D)	0.002	0.987		
Savings goal ahead (D)	**0.055**	0.024		
Business owner (D)	0.018	0.575		
Prob(inheritance)	0.035	0.296		
Occ. disab. insur. (D)	**0.063**	0.000		
Live shorter than av. (D)	-0.009	0.605		
Live longer than av. (D)	0.008	0.586		
Income variation				
Income development: pos. (D)	**0.039**	0.006		
Inc. dev.: highly volatile (D)	**-0.039**	0.025		
Inc. dev.: slightly volatile (D)	-0.009	0.529		
Variance of net income	0.000	0.271		
Constant	-0.179	0.023		
Number of obs	1566			
uncensored obs	906			
LR chi2(68)	267.7			
$Prob > chi2$	0			
Pseudo R2	0.3462			
Log likelihood	-252.73504			

Table 5.16: Regression results: probability of job loss expectations and **financial wealth/permanent income**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**12.975**	0.000		
(Perm. Inc./ 10,000) sq.	**-12.601**	0.000		
Age / 10	**0.613**	0.025	0.632	0.101
Age / 10 sq.	-0.041	0.107	-0.058	0.125
Secondary school (D)	**0.406**	0.015	-0.318	0.136
Graduation diploma (D)	0.272	0.265	-0.225	0.514
University degree (D)	0.148	0.517	0.042	0.886
Kids (D)	0.159	0.440		
Kids living in same house (D)	**-0.707**	0.000		
Separated or divorced (D)	-0.274	0.243		
Widowed (D)	-0.243	0.369		
East Germany (D)	**-0.352**	0.030		
Female (D)	**-0.284**	0.058		
Job: blue collar (D)	0.072	0.779	-0.183	0.545
Job: civil servant (D)	-0.430	0.245	-0.253	0.622
Job: freelancer (D)	-0.485	0.458	-0.880	0.293
Job: self-employed (D)	-0.551	0.168	-0.649	0.129
Retired (D)	0.214	0.453	-0.113	0.696
Work parttime (D)	-0.293	0.355	0.041	0.906
Work little (D)	-0.317	0.299	-0.040	0.921
Work not (D)	-0.355	0.175	-0.007	0.981
Unemployed (D)	-0.337	0.250	0.132	0.725
Unemp.> 1 month (D)	0.153	0.416	-0.102	0.707
Unemp.> 6 months (D)	**-0.584**	0.005	-0.054	0.855
Village (D)	0.072	0.784		
Partner (D)	-1.544	0.110		
Savings goal ahead (D)	**0.581**	0.025		
Business owner (D)	**0.899**	0.007		
Prob(inheritance)	**0.756**	0.040		
Occ. disab. insur. (D)	**0.375**	0.047		
Live shorter than av. (D)	-0.121	0.524		
Live longer than av. (D)	-0.054	0.722		
Income variation				
Income development: pos. (D)	0.151	0.329		
Inc. dev.: highly volatile (D)	-0.194	0.276		
Inc. dev.: slightly volatile (D)	-0.014	0.923		
Variance of net income	-0.001	0.311		
Constant	-2.158	0.007		
Number of obs	1130			
uncensored obs	706			
LR chi2(68)	360.65			
$Prob > chi2$	0			
Pseudo R2	0.0934			
Log likelihood	-1749.5148			

Table 5.17: Regression results: probability of job loss expectations and **total wealth/permanent income**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**15.123**	0.000		
(Perm. Inc./ 10,000) sq.	**-15.517**	0.000		
Age / 10	**0.877**	0.003	**1.169**	0.005
Age / 10 sq.	**-0.053**	0.053	**-0.084**	0.044
Secondary school (D)	**0.657**	0.000	-0.146	0.544
Graduation diploma (D)	**0.641**	0.014	**-0.847**	0.030
University degree (D)	**0.645**	0.014	-0.120	0.726
Kids (D)	**-0.566**	0.011		
Kids living in same house (D)	-0.046	0.819		
Separated or divorced (D)	-0.513	0.047		
Widowed (D)	-0.430	0.139		
East Germany (D)	**-0.444**	0.012		
Female (D)	**-0.451**	0.006		
Job: blue collar (D)	0.311	0.279	-0.541	0.109
Job: civil servant (D)	-0.033	0.938	0.331	0.581
Job: freelancer (D)	-1.033	0.142	-0.404	0.674
Job: self-employed (D)	-0.340	0.493	0.637	0.232
Retired (D)	-0.458	0.144	-0.425	0.203
Work parttime (D)	0.352	0.315	0.249	0.529
Work little (D)	**0.772**	0.022	-0.438	0.347
Work not (D)	**0.649**	0.022	**0.550**	0.098
Unemployed (D)	**-1.117**	0.000	-0.233	0.571
Unemp.> 1 month (D)	-0.049	0.818	-0.500	0.109
Unemp.> 6 months (D)	**-0.376**	0.100	0.516	0.120
Village (D)	**0.975**	0.001		
Partner (D)	**-3.853**	0.000		
Savings goal ahead (D)	**0.716**	0.012		
Business owner (D)	**1.972**	0.000		
Prob(inheritance)	0.616	0.137		
Occ. disab. insur. (D)	0.177	0.406		
Live shorter than av. (D)	**-0.389**	0.063		
Live longer than av. (D)	**-0.401**	0.017		
Income variation				
Income development: pos. (D)	-0.059	0.730		
Inc. dev.: highly volatile (D)	-0.068	0.721		
Inc. dev.: slightly volatile (D)	-0.144	0.359		
Variance of net income	-0.001	0.241		
Constant	-1.686	0.045		
Number of obs	1010			
F(33, 1100 / F(20, 661)	10.53			
$Prob > F$	0			
R squared	0.364			
Adj. R sq.	0.3294			
Root MSE	2.0449			

Table 5.18: Regression results: local unemployment probability and **saving rates**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**0.560**	0.011		
(Perm. Inc./ 10,000) sq.	**-0.483**	0.099		
Age / 10	0.041	0.141	-0.006	0.868
Age / 10 sq.	-0.004	0.131	0.001	0.860
Secondary school (D)	0.012	0.444	0.017	0.374
Graduation diploma (D)	**0.051**	0.022	0.029	0.353
University degree (D)	**0.042**	0.052	**0.059**	0.028
Kids (D)	0.019	0.342		
Kids living in same house (D)	**-0.055**	0.002		
Separated or divorced (D)	**-0.048**	0.041		
Widowed (D)	-0.017	0.513		
East Germany (D)	-0.006	0.690		
Female (D)	-0.012	0.423		
Job: blue collar (D)	0.029	0.220	0.014	0.601
Job: civil servant (D)	0.011	0.720	-0.031	0.531
Job: freelancer (D)	0.030	0.621	-0.023	0.769
Job: self-employed (D)	**0.109**	0.003	**-0.117**	0.004
Retired (D)	**0.064**	0.029	-0.020	0.497
Work parttime (D)	0.002	0.927	-0.014	0.635
Work little (D)	-0.021	0.473	-0.033	0.379
Work not (D)	-0.036	0.147	-0.006	0.832
Unemployed (D)	-0.010	0.736	-0.045	0.221
Unemp.> 1 month (D)	-0.006	0.732	-0.003	0.882
Unemp.> 6 months (D)	**-0.056**	0.005	0.010	0.712
Village (D)	-0.006	0.806		
Partner (D)	0.011	0.904		
Savings goal ahead (D)	**0.056**	0.020		
Business owner (D)	0.018	0.566		
Prob(inheritance)	0.046	0.166		
Occ. disab. insur. (D)	**0.062**	0.000		
Live shorter than av. (D)	-0.014	0.448		
Live longer than av. (D)	0.013	0.355		
Income risk (local unemp.rate)	**0.000**	0.024		
Constant	**-0.172**	0.026		
Number of obs	1589			
uncensored obs	918			
LR chi2(68)	263.550			
$Prob > chi2$	0.000			
Pseudo R2	0.336			
Log likelihood	-260.606			

Table 5.19: Regression results: local unemployment probability and **financial wealth/permanent income**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**13.671**	0.000		
(Perm. Inc./ 10,000) sq.	**-13.397**	0.000		
Age / 10	**0.516**	0.059	**0.726**	0.060
Age / 10 sq.	-0.032	0.204	**-0.065**	0.086
Secondary school (D)	0.399	0.017	-0.303	0.156
Graduation diploma (D)	0.297	0.224	-0.078	0.820
University degree (D)	0.191	0.403	0.094	0.753
Kids (D)	0.156	0.448		
Kids living in same house (D)	**-0.735**	0.000		
Separated or divorced (D)	-0.332	0.155		
Widowed (D)	-0.250	0.357		
East Germany (D)	**-0.325**	0.044		
Female (D)	**-0.275**	0.067		
Job: blue collar (D)	0.087	0.735	-0.175	0.563
Job: civil servant (D)	-0.379	0.306	-0.311	0.559
Job: freelancer (D)	-0.463	0.494	-0.884	0.293
Job: self-employed (D)	-0.633	0.116	**-0.735**	0.081
Retired (D)	0.154	0.592	-0.170	0.558
Work parttime (D)	-0.286	0.369	0.013	0.969
Work little (D)	-0.346	0.259	0.137	0.733
Work not (D)	-0.331	0.205	0.045	0.878
Unemployed (D)	-0.442	0.127	-0.044	0.905
Unemp.> 1 month (D)	0.154	0.414	-0.075	0.781
Unemp.> 6 months (D)	**-0.588**	0.005	-0.013	0.964
Village (D)	0.106	0.687		
Partner (D)	-1.898	0.050		
Savings goal ahead (D)	**0.614**	0.017		
Business owner (D)	**0.933**	0.005		
Prob(inheritance)	**0.779**	0.034		
Occ. disab. insur. (D)	**0.353**	0.062		
Live shorter than av. (D)	-0.147	0.440		
Live longer than av. (D)	-0.047	0.754		
Income risk (local unemp.rate)	-0.002	0.141		
Constant	-1.976	0.012		
Number of obs	1140			
uncensored obs	713			
LR chi2(68)	361.06			
$Prob > chi2$	0			
Pseudo R2	0.0925			
Log likelihood	-1771.2515			

Table 5.20: Regression results: local unemployment probability and **total wealth/permanent income**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**14.821**	0.000		
(Perm. Inc./ 10,000) sq.	**-15.231**	0.000		
Age / 10	**0.872**	0.003	**1.217**	0.004
Age / 10 sq.	**-0.053**	0.055	**-0.087**	0.036
Secondary school (D)	**0.640**	0.000	-0.119	0.620
Graduation diploma (D)	**0.605**	0.020	**-0.661**	0.087
University degree (D)	**0.656**	0.012	**-0.050**	0.884
Kids (D)	**-0.566**	0.011		
Kids living in same house (D)	-0.065	0.749		
Separated or divorced (D)	**-0.520**	0.043		
Widowed (D)	-0.450	0.121		
East Germany (D)	**-0.453**	0.010		
Female (D)	**-0.466**	0.005		
Job: blue collar (D)	0.314	0.274	-0.531	0.115
Job: civil servant (D)	-0.064	0.880	0.204	0.741
Job: freelancer (D)	-1.055	0.153	-0.476	0.620
Job: self-employed (D)	-0.398	0.427	0.442	0.395
Retired (D)	-0.453	0.150	-0.470	0.157
Work parttime (D)	0.343	0.327	0.234	0.554
Work little (D)	**0.750**	0.025	-0.287	0.533
Work not (D)	**0.660**	0.019	0.566	0.088
Unemployed (D)	**-1.114**	0.000	-0.296	0.467
Unemp.> 1 month (D)	-0.049	0.817	-0.463	0.134
Unemp.> 6 months (D)	-0.381	0.094	0.505	0.125
Village (D)	**1.008**	0.001		
Partner (D)	**-4.025**	0.000		
Savings goal ahead (D)	**0.735**	0.009		
Business owner (D)	**2.027**	0.000		
Prob(inheritance)	0.617	0.135		
Occ. disab. insur. (D)	0.152	0.475		
Live shorter than av. (D)	**-0.390**	0.061		
Live longer than av. (D)	**-0.377**	0.024		
Income risk (local unemp.rate)	-0.002	0.144		
Constant	-1.700	0.039		
Number of obs	1016			
F(33, 1100 / F(20, 661)	11.22			
Prob > F	0			
R squared	0.3626			
Adj. R sq.	0.3303			
Root MSE	2.0467			

Table 5.21: Regression results: Savings motives and **saving rates**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**0.490**	0.028		
(Perm. Inc./ 10,000) sq.	-0.436	0.137		
Age / 10	0.048	0.103	-0.017	0.665
Age / 10 sq.	-0.004	0.128	0.002	0.691
Secondary school (D)	0.009	0.564	0.008	0.659
Graduation diploma (D)	0.048	0.036	0.023	0.479
University degree (D)	0.044	0.042	**0.060**	0.027
Kids (D)	0.003	0.907		
Kids living in same house (D)	**-0.046**	0.012		
Separated or divorced (D)	-0.031	0.197		
Widowed (D)	-0.006	0.828		
East Germany (D)	0.000	0.984		
Female (D)	-0.006	0.697		
Job: blue collar (D)	0.036	0.130	0.032	0.224
Job: civil servant (D)	0.020	0.528	-0.013	0.784
Job: freelancer (D)	0.025	0.689	0.001	0.989
Job: self-employed (D)	**0.112**	0.002	**-0.081**	0.046
Retired (D)	**0.074**	0.013	-0.017	0.559
Work parttime (D)	0.016	0.559	-0.005	0.868
Work little (D)	-0.003	0.907	-0.009	0.809
Work not (D)	-0.038	0.129	-0.005	0.849
Unemployed (D)	-0.003	0.921	-0.036	0.326
Unemp.> 1 month (D)	-0.005	0.755	0.000	0.983
Unemp.> 6 months (D)	**-0.054**	0.006	0.016	0.534
Village (D)	-0.016	0.515		
Partner (D)	0.037	0.702		
Savings goal ahead (D)	**0.071**	0.004		
Business owner (D)	0.010	0.758		
Prob(inheritance)	0.054	0.104		
Occ. disab. insur. (D)	**0.058**	0.000		
Live shorter than av. (D)	-0.012	0.527		
Live longer than av. (D)	0.008	0.600		
Low financial risk	-0.018	0.209		
High financial risk	-0.019	0.581		
Saving goals (D)				
Low: buying a home	0.022	0.226		
High: buying a home	**0.051**	0.007		
Low: unforseen events	**-0.103**	0.000		
High: unforseen events	0.020	0.146		
Low: repaying debts	**0.074**	0.000		
High: repaying debts	-0.027	0.135		
Low: old-age provision	-0.015	0.437		
High: old-age provision	**0.040**	0.006		
Low: holidays	0.023	0.114		
High: holidays	-0.012	0.493		
Low: major purchases	-0.011	0.498		
High: major purchases	0.008	0.618		
Low: subsidizing offspring	-0.005	0.775		
High: subsidizing offspring	0.001	0.954		
Low: leaving bequests	**-0.030**	0.053		
High: leaving bequests	-0.007	0.735		
Low: getting tax subs.	**-0.029**	0.066		
High: getting tax subs.	**-0.009**	0.635		
Constant	-0.225	0.006		
Number of obs	1515			
LR chi2(68)	357.34			
$Prob > chi2$	0			
Pseudo R2	0.4834			
Log likelihood	-190.9217			

Table 5.22: Regression results: Savings motives and **financial wealth/permanent income**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**11.199**	0.000		
(Perm. Inc./ 10,000) sq.	**-11.166**	0.000		
Age / 10	**0.724**	0.011	0.616	0.109
Age / 10 sq.	**-0.046**	0.088	-0.053	0.166
Secondary school (D)	**0.358**	0.031	-0.325	0.118
Graduation diploma (D)	0.229	0.345	-0.148	0.668
University degree (D)	0.069	0.759	0.240	0.414
Kids (D)	0.100	0.643		
Kids living in same house (D)	**-0.707**	0.000		
Separated or divorced (D)	-0.126	0.589		
Widowed (D)	-0.105	0.700		
East Germany (D)	**-0.365**	0.025		
Female (D)	**-0.265**	0.076		
Job: blue collar (D)	-0.056	0.821	-0.015	0.959
Job: civil servant (D)	-0.266	0.465	-0.254	0.610
Job: freelancer (D)	**-1.140**	0.098	-0.713	0.385
Job: self-employed (D)	-0.482	0.212	-0.432	0.294
Retired (D)	0.251	0.381	-0.236	0.408
Work parttime (D)	-0.262	0.402	0.044	0.895
Work little (D)	-0.145	0.634	0.251	0.528
Work not (D)	**-0.458**	0.075	-0.040	0.890
Unemployed (D)	-0.332	0.247	0.170	0.642
Unemp.> 1 month (D)	0.111	0.546	0.079	0.769
Unemp.> 6 months (D)	**-0.555**	0.007	-0.091	0.755
Village (D)	-0.033	0.900		
Partner (D)	**-1.594**	0.097		
Savings goal ahead (D)	**0.593**	0.022		
Business owner (D)	**0.905**	0.005		
Prob(inheritance)	**0.811**	0.024		
Occ. disab. insur. (D)	0.236	0.207		
Live shorter than av. (D)	-0.115	0.550		
Live longer than av. (D)	-0.096	0.523		
Low financial risk	**-0.492**	0.001		
High financial risk	-0.318	0.408		
Saving goals (D)				
Low: buying a home	0.272	0.155		
High: buying a home	0.332	0.101		
Low: unforseen events	**-0.676**	0.002		
High: unforseen events	0.177	0.221		
Low: repaying debts	**0.457**	0.010		
High: repaying debts	**-0.389**	0.039		
Low: old-age provision	-0.039	0.832		
High: old-age provision	0.183	0.227		
Low: holidays	-0.073	0.620		
High: holidays	-0.196	0.272		
Low: major purchases	-0.033	0.835		
High: major purchases	0.238	0.170		
Low: subsidizing offspring	**-0.298**	0.086		
High: subsidizing offspring	-0.059	0.746		
Low: leaving bequests	-0.111	0.498		
High: leaving bequests	-0.331	0.120		
Low: getting tax subs.	**-0.361**	0.026		
High: getting tax subs.	-0.004	0.984		
Constant	-1.886	0.019		
Number of obs	1083			
LR chi2(68)	441.04			
$Prob > chi2$	0			
Pseudo R2	0.1188			
Log likelihood	-1636.0567			

Table 5.23: Regression results: Savings motives and **total wealth/permanent income**

	Respondent		Partner	
	Coef.	$P > t$	Coef.	$P > t$
Permanent income / 10,000	**12.861**	0.000		
(Perm. Inc./ 10,000) sq.	**-13.360**	0.000		
Age / 10	**0.905**	0.003	**0.975**	0.023
Age / 10 sq.	**-0.054**	0.070	-0.065	0.132
Secondary school (D)	**0.541**	0.004	-0.060	0.801
Graduation diploma (D)	**0.604**	0.024	-0.238	0.554
University degree (D)	**0.656**	0.014	0.299	0.396
Kids (D)	**-0.498**	0.037		
Kids living in same house (D)	-0.067	0.748		
Separated or divorced (D)	-0.286	0.279		
Widowed (D)	-0.388	0.194		
East Germany (D)	**-0.368**	0.041		
Female (D)	**-0.424**	0.012		
Job: blue collar (D)	0.404	0.157	-0.187	0.582
Job: civil servant (D)	-0.063	0.885	0.433	0.469
Job: freelancer (D)	-1.235	0.111	0.229	0.812
Job: self-employed (D)	-0.214	0.663	0.517	0.321
Retired (D)	-0.207	0.518	-0.403	0.233
Work parttime (D)	0.431	0.223	0.439	0.271
Work little (D)	**0.660**	0.056	-0.189	0.684
Work not (D)	**0.538**	0.060	0.507	0.133
Unemployed (D)	**-0.819**	0.006	-0.325	0.427
Unemp.> 1 month (D)	-0.076	0.718	-0.276	0.381
Unemp.> 6 months (D)	-0.321	0.163	0.390	0.243
Village (D)	**0.868**	0.005		
Partner (D)	**-3.514**	0.001		
Savings goal ahead (D)	**0.681**	0.019		
Business owner (D)	**1.964**	0.000		
Prob(inheritance)	0.643	0.120		
Occ. disab. insur. (D)	0.032	0.883		
Live shorter than av. (D)	**-0.438**	0.041		
Live longer than av. (D)	**-0.387**	0.024		
Low financial risk	**-0.414**	0.015		
High financial risk	-0.531	0.228		
Saving goals (D)				
Low: buying a home	**0.384**	0.072		
High: buying a home	**0.917**	0.000		
Low: unforseen events	**-0.652**	0.005		
High: unforseen events	0.066	0.682		
Low: repaying debts	0.204	0.305		
High: repaying debts	0.097	0.648		
Low: old-age provision	-0.202	0.321		
High: old-age provision	0.144	0.398		
Low: holidays	**0.382**	0.022		
High: holidays	-0.002	0.992		
Low: major purchases	-0.182	0.304		
High: major purchases	**0.341**	0.090		
Low: subsidizing offspring	0.225	0.252		
High: subsidizing offspring	**-0.436**	0.040		
Low: leaving bequests	-0.052	0.782		
High: leaving bequests	**0.547**	0.024		
Low: getting tax subs.	-0.229	0.213		
High: getting tax subs.	**-0.377**	0.094		
Constant	-1.991	0.021		
Number of obs	969			
F(33, 1100 / F(20, 661)	8.94			
$Prob > F$	0			
R squared	0.4032			
Adj. R sq.	0.3581			
Root MSE	2.0095			

Table 5.24: Regression results: All subjective variables for saving rates, relative financial wealth, relative total wealth

	Saving rates		Relative financial wealth		Relative total wealth	
	Coef.	$P > t$	Coef.	$P > t$	Coef.	$P > t$
Savings goal ahead (D)	**0.067**	0.006	**0.498**	0.046	**0.649**	0.022
Business owner (D)	0.009	0.770	**0.814**	0.010	**1.841**	0.000
Prob(inheritance)	0.031	0.366	**0.764**	0.028	**0.679**	0.097
Occ. disab. insur. (D)	**0.054**	0.001	0.234	0.201	-0.017	0.938
Live shorter than av. (D)	-0.009	0.627	-0.070	0.712	-0.328	0.127
Live longer than av. (D)	-0.004	0.797	-0.070	0.635	**-0.446**	0.009
Low financial risk (D)	-0.021	0.142	**-0.468**	0.001	**-0.318**	0.059
High financial risk (D)	-0.031	0.386	-0.452	0.224	-0.571	0.186
Saving goals (D)						
Low: buying a home	0.018	0.323	0.227	0.226	0.345	0.105
High: buying a home	**0.053**	0.007	**0.359**	0.068	**0.967**	0.000
Low: unforseen events	**-0.094**	0.000	**-0.641**	0.002	**-0.696**	0.002
High: unforseen events	**0.025**	0.067	0.208	0.140	0.052	0.742
Low: repaying debts	**0.066**	0.000	**0.438**	0.012	0.192	0.331
High: repaying debts	-0.027	0.133	-0.304	0.096	0.101	0.628
Low: old-age provision	-0.018	0.382	-0.039	0.827	-0.228	0.257
High: old-age provision	**0.034**	0.019	0.112	0.448	0.073	0.664
Low: holidays	**0.026**	0.078	-0.080	0.575	**0.360**	0.028
High: holidays	-0.008	0.670	-0.181	0.296	-0.005	0.981
Low: major purchases	-0.001	0.968	0.094	0.541	-0.093	0.595
High: major purchases	-0.001	0.967	0.223	0.189	0.306	0.123
Low: subsidizing offspring	-0.002	0.927	**-0.301**	0.077	0.228	0.241
High: subsidizing offspring	0.003	0.875	-0.051	0.773	**-0.467**	0.025
Low: leaving bequests	-0.022	0.160	-0.019	0.907	0.010	0.958
High: leaving bequests	-0.005	0.788	-0.330	0.111	**0.542**	0.023
Low: getting tax subs.	**-0.028**	0.079	**-0.307**	0.052	-0.271	0.136
High: getting tax subs.	-0.013	0.483	0.007	0.970	**-0.412**	0.063
Income variation						
Income development: pos. (D)	**0.026**	0.071	0.128	0.397	-0.083	0.633
Inc. dev.: highly volatile (D)	**-0.032**	0.074	-0.151	0.377	-0.066	0.728
Inc. dev.: slightly volatile (D)	-0.005	0.732	0.032	0.818	0.008	0.962
Variance of net income	0.000	0.351	0.000	0.904	0.000	0.661
Income risk (local unemp.rate)	**0.000**	0.024	-0.002	0.249	-0.001	0.649
Low: Germany's ec. situation	-0.020	0.140	0.181	0.155	0.133	0.365
High: Germany's ec. situation	0.028	0.414	-0.001	0.999	-0.640	0.134
Low: Own economic situation	**-0.093**	0.000	**-0.696**	0.000	**-0.493**	0.013
High: Own economic situation	0.024	0.138	0.162	0.309	**0.384**	0.037
Low: Own health situation	0.023	0.381	**-0.511**	0.036	-0.405	0.125
High: Own health situation	0.015	0.344	**-0.284**	0.065	-0.011	0.952
Low: Partner's health situation	-0.004	0.906	0.087	0.792	-0.119	0.749
High: Partner's health situation	-0.009	0.620	**0.344**	0.065	0.281	0.197

5.B Figures

Figure 5.4: Transformation of the wealth variable

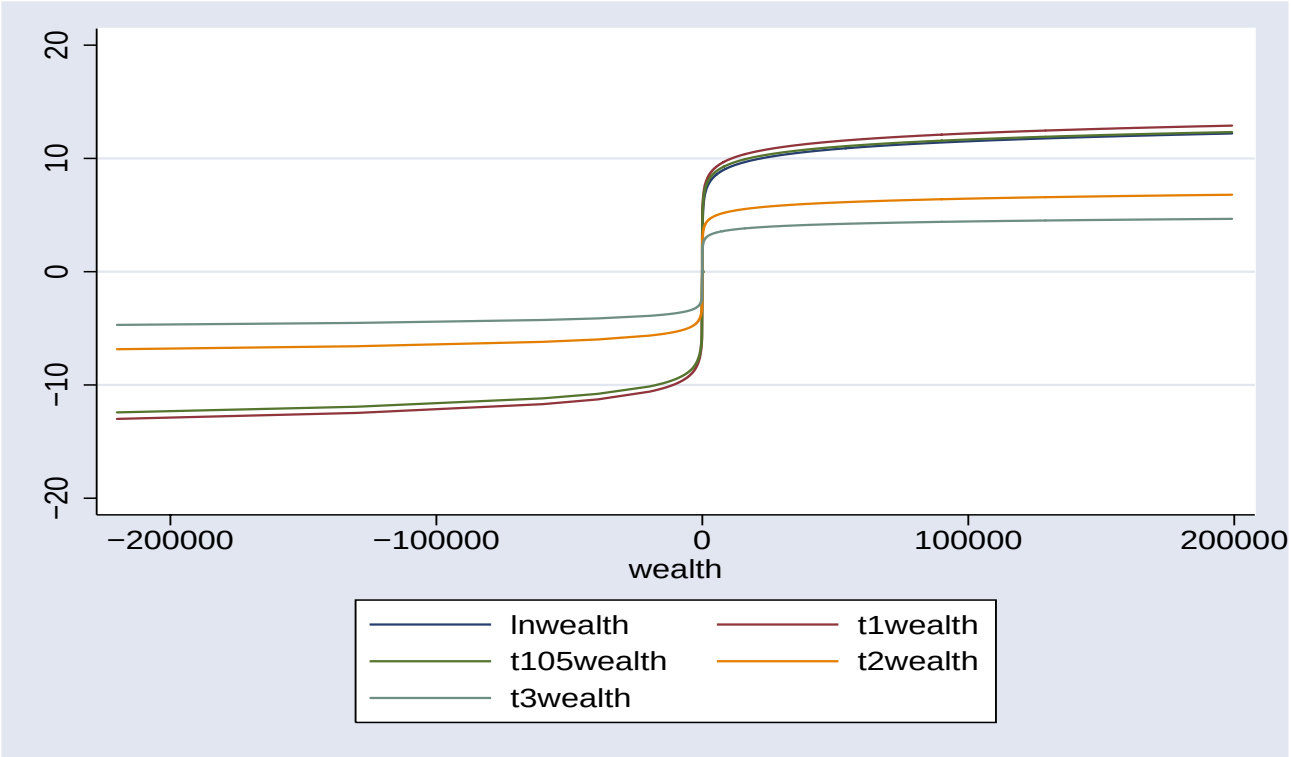
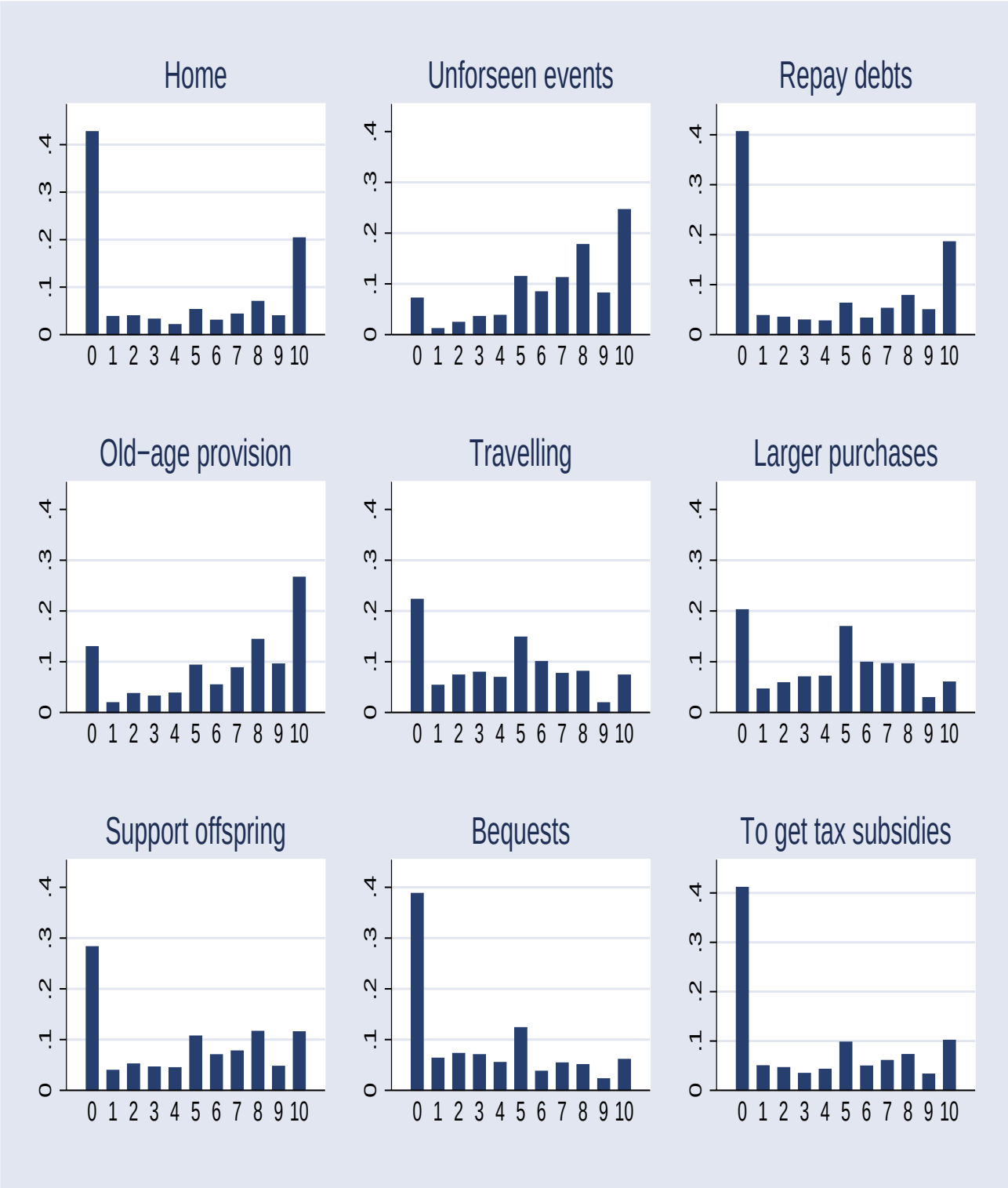
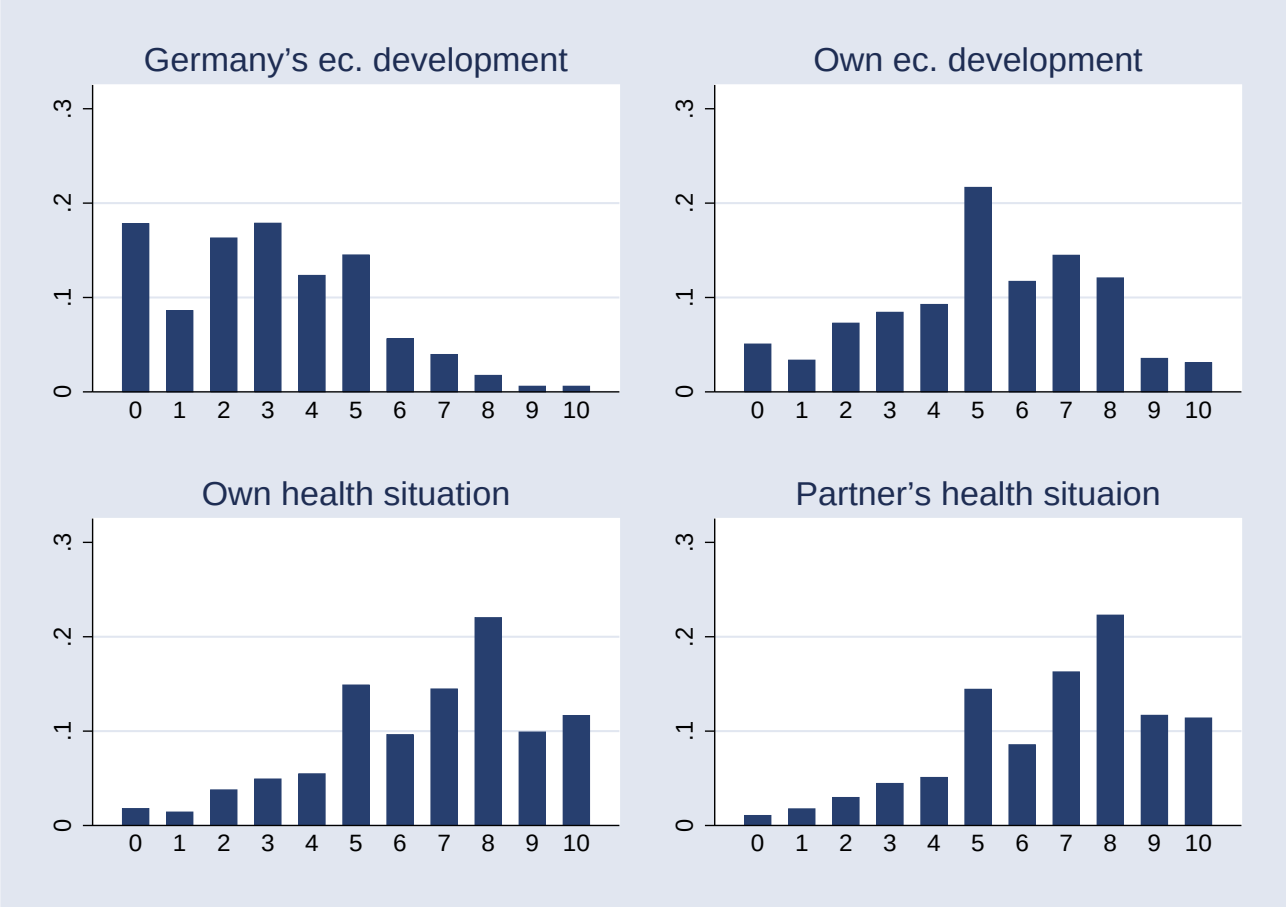


Figure 5.5: Histograms for saving reasons



Notes: Weighted values by subsample, hh income, and age. Answers are measured on a scale from 0 to 10, where 0 means 'totally unimportant' and 10 'very important'.

Figure 5.6: Histograms for development expectations



Notes: Weighted values by subsample, hh income, and age. Answers are measured on a scale from 0 to 10, where 0 means 'very negative' and 10 'very positive'.

Chapter 6

Imputing total expenditures from a non-exhaustive list of sub-items: An empirical assessment using the *SAVE* data set

6.1 Introduction

Income and expenditure surveys are available in many countries.[106] They provide a great deal of detailed and supposedly reliable consumption data, but these data only strive or totally lack other data necessary for many research questions. While general-purpose surveys like the German Socio-Economic Panel *GSOEP* or *SAVE* typically are a good source of socio-economic characteristics or other information relevant for savings analysis, they face the immanent problem of lacking detailed information on some other information like consumption. These problems are immanent since these surveys are typically not diary based; this would be way too costly and can usually only be justified by the need for central statistical offices to calculate weights for consumer price indices. The time restrictions for personal household interviews thus make it necessary to look for alternative ways to retrieve useful information.

Browning *et al.* (2003) therefore developed a method for getting a measure for total non-durable expenditures[107] of households without asking an exhaustive list of consumption items and sub-items. This saves time and makes asking consumption questions in general purpose surveys feasible. Furthermore, they also do not rely on a direct question for total consumption expenditures which has been shown to be measured with a large error, see Browning *et al.* (2003) or Battistin *et al.* (2003).

The idea of Browning *et al.* (2003) is to use the exhaustive information from an external survey with with exhaustive expenditure items to impute total expenditures in the non-exhaustive survey based on the sub-item. The external data will be a national income and expenditure survey, e.g. with expenditure information from a household diary. I apply this idea to the *SAVE* data set where only a few expenditure items are asked. The external data will be the German Income and Expenditure Survey (*EVS*). The aim of this chapter is to compare the direct measure of total expenditures in *SAVE* with the imputed values. Furthermore, I will test whether some expenditure sub-items can proxy total non-durable expenditures as well for German data, i.e. the *EVS*, as for other national surveys like

[106] E.g., the Income and Expenditure Survey (*EVS*) in Germany, the survey of Family Expenditures *FAMEX* in Canada, the Consumer Expenditure Survey *CEX* in the U.S., or the Survey of Family Budgets *SFB* in Italy.

[107] Browning and Crossley (2004) suggest that non-durable expenditures such as food may be differentially smoothed in response to shocks, and show that much more action appears in small durables.

those used in Browning *et al.* (2003). Finally, I use the imputed expenditure measure to compute household savings as a residual measure (difference of income and total non-durable expenditures).

Winter (2004) presents experimental evidence on how the choice of expenditure categories influences measures of household consumption. He interprets the findings that responses to one-shot questions on total monthly nondurable expenditures differ from the sum of disaggregated categories. Furthermore, he finds underreporting in the one-shot question even when considering that the answers to a detailed list of 35 categories might still also be subject to underreporting.

The structure of this chapter will be as follows: In Section 6.2, I will quickly mention three alternative methods of retrieving information on total expenditures in surveys, which are pointed out by Browning *et al.* (2003). In Section 6.3, descriptive summaries from the expenditure questions *SAVE* are displayed and compared to the expenditure values from the German *EVS*. It also shows the results from the imputation procedure based on the *EVS* for total expenditures in *SAVE* and compares the imputed to the recall values.

6.2 Asking expenditure questions in surveys

As pointed out in Section 8.1, running diary based surveys which might recover reliable detailed consumption values, is burdensome, time consuming and very costly. It would therefore practically be impossible to receive information on consumption and savings related topics from the observed household if we had to use diaries to get a good measure of expenditure data.

Browning *et al.* (2003) discuss three methods to gain information on total expenditure which will summarized below. In brief, the first one is to ask a single general total expenditure question, the second one asks for a detailed and exhaustive list of sub-items composing total expenditures while the third one is a nonexhaustive selected subset of the list of total sub-items. General-purpose surveys include retrospective or recall questions on consumption and expenditures. In contrast to the German *EVS* which delivers diary-based data on expenditures, other national expenditure surveys like the U.S. *CEX* or the Canadian *FAMEX* are partially based on interview recall questions. Beginning with the *SAVE* 2003 wave, the latter method, along with the first one, was applied.

General-purpose surveys include retrospective or recall questions on consumption and expenditures. In contrast to the German *EVS* which delivers diary-based data on expenditures, other national expenditure surveys like the U.S. *CEX* or the Canadian *FAMEX* are partially based on interview recall questions.

One-shot question for total non-durable expenditures At a first glance, it seems attractive to simply ask one total expenditure question in surveys: it is time saving, and the question can easily be understood. Thinking twice, this option appears far less appealing, since the question is very complex, and respondents tend to give a rough estimate which then is heavily loaded with noise.[108]

Battistin *et al.* (2003) compare the expenditure questions from the Bank of Italy Survey on Household Income and Wealth (*SHIW*) with the corresponding diary based survey (*SFB*). They develop a

[108] which might become quite clear if one tries to guess the own average consumption

model for the recall error process to correct for heaping and rounding in the recall values; still, the distribution of true[109] expenditures is different for total non-durable expenditure. They conclude that the *SHIW* reported non-durable expenditure measure diverge from values of the *SFB* and conclude that the recall error is more severe concerning total non-durable consumption than for subcategories like expenditures for 'food at home'.

This is also found by Browning *et al.* (2003) who compared the Canadian Out of Employment Panel *(COEP)* with the Canadian Family Expenditure Survey *(FAMEX)* and *SHIW* to *SFB*. For Italy, they find underreporting of total non-durable expenditures of 24%/30%, and for Canada 37%/32% (total expenditures) when comparing medians/means.

While, in addition to the measurement problem, one might wonder about nonresponse rates to total expenditures, this seems to be less of a problem. Browning *et al.* (2003) report a nonresponse rate of 6.0% for total expenditures, while Winter (2004) finds item nonresponse ('don't know' option) of about one third which in turn compares to the rate of 35.8% reported by Hurd *et al.* (1998). Nonresponse rates of that amount require analyses for response dependence on household and demographic characteristics. While there is evidence of significant demographic and other effects in Hurd et al., which raises the issue on sample selectivity, Winter finds evidence justifying the assumption of random non-response.

Exhaustive list of items The summaries of Browning *et al.* (2003) and Winter (2004) concerning the use of an exhaustive list of expenditure items give the advice that their inclusion in surveys is quite costly in interview time, which is costly in monetary terms as well as in terms of trade-off costs for other questions since there normally is a natural limit of interview time[110]. Apart from the time constraint, many of the items might be reported with noise as well.

Non-exhaustive list of items This paragraph splits into two sections. The first one investigates which sub-items are measured reliably and which ones proxy total expenditures reasonably well. The second one explains how the incomplete measures are used to get a reliable measures for total non-durable expenditure.

Browning *et al.* (2003) have shown that questions on expenditure on 'food at home' are not exposed to the same amount of noise as the total non-durable expenditure questions. Means and medians for the just mentioned comparison of survey to diary data are about the same, which is also true for the dispersion of the data. These encouraging results state that respondents are astonishingly well capable to give reliable responses to that question.

Since 'food at home' also represents a large budget item, it is useful for imputing total consumption. Browning *et al.* (2003) explain that even though they can not present evidence on its accuracy, a 'food outside home' question should be included, since it represents a substitute to food at home; it might also capture heterogeneity of the two budget shares for households having the same level of total expenditure. In addition to the food questions, Browning et al. advise us to also collect information

[109] assuming that the diary-based information reflect the truth

[110] after which interview abortion rates will rise dramatically

on utilities (or energy costs like water, fuel, electricity) and communication expenditures based on their analysis of the explanatory power of these variables for total expenditure.

The basic idea in the process of using only a sub-group of items to estimate total expenditure is the following. Using expenditure survey data with a precise measure of total expenditure x_{total}; based on a sum of all sub-item expenditures, one chooses a subset of goods $x_i, i = 1, 2...l$ and estimates the following

$$x_{total} = \alpha + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_l x_l + \epsilon \tag{6.1}$$

By interpreting the estimated coefficients β_i as weights, it is possible to use the $\hat{\beta}_i$ for predictions of $\hat{x}_{total}$ on the basis of the same goods $x_i, i = 1, 2...l$ using data from the general-purpose survey, e.g. *SAVE*:

$$\hat{x}_{total} = \hat{\alpha} + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_l x_l \tag{6.2}$$

Browning et al. suggest *not* to include income as an predictor for two reasons. The first one is that they suspect that income is an ill-measured variable, based on the results by Lusardi (1996). The second is that income introduces spurious relationships between income and the result of the imputation which then invalidates some uses of the imputed total expenditure measure.[111]

6.3 Descriptive findings in *SAVE* and in the *EVS*

The expenditure questions in the German *SAVE* data set were designed in the way suggested by Browning *et al.* (2003). In 2003, five questions were included to ask for four expenditure items[112], and total non-durable expenditures. The four items were 'food at home', 'food outside home', 'telecommunication services' and 'utilities' (heating and energy costs). The exact wording and survey implementation is shown in the Appendix, part 6.C. The inclusion of the consumption questions in that period was especially appealing since in the same year, a wave of the *EVS* survey was conducted.[113] Unfortunately, the 2003 scientific use file containing detailed information on expenditure components is not yet available. This is why for following analyses, the 1998 *EVS* wave is used.

6.3.1 Expenditure items in *SAVE*

This section describes empirical findings for the *SAVE* consumption expenditure data. As further explained in the Appendix, part 6.C, energy costs are asked in two questions where the first one collects information on the billing period, while the second one asks for the average costs per bill. The product of these two questions is then recalculated to obtain monthly heating expenditures. Table 6.1 lists methodological issues of the expenditure items.

[111] E.g., testing for excess sensitivity.

[112] One question controlling for the billing period of energy costs.

[113] The German Income and Expenditure Survey (*EVS*) started in 1978 and was repeated on a regular five year interval. It described as a series of cross section, although it contains a considerable true panel component of individuals willing to participate several times. Still, this panel information is *not* available due to the original data law agreements which makes it impossible to track households over time.

Table 6.1: Values for different expenditure items

	Food at home		Food outside home		Telecom. services		Water, fuel, electricity		Total non-durable exp.	
	N	Percent	N	Percent	N	Percent	N	Percent	N	Percent
Nonresponse	21	0.67	115	3.65	47	1.49	782[b]	24.79[b]	18	0.57
Outliers[a]	11	0.35	1	0.03	5	0.16	9	0.38	133	4.24
Zeros	183	5.84	622	20.47	75	2.41	158	6.66	225	7.17
Obs.	2939	93.18	2416	76.60	3027	95.97	2205	69.91	2778	88.08

[a] Outliers defined as values being larger than net household income if income is observed.

[b] Monthly energy costs are calculated from two questions, the billing period and costs per bill. From the 782 missings, 402 were due to the option 'included in monthly rent', 87 to 'heating billing period other than listed', 113 / 180 to nonresponse to the billing period / heating costs questions.

Source: All *SAVE* 2003 and 2004 subsamples

The items appear in the order they were asked in the survey, i.e. total non-durable expenditures were asked *after* all sub-items. At first glance, non-response rates seem to be rather low, especially for total non-durable expenditures. Respondents were obviously willing to give an answer despite having refused to answer to the previous questions. Still, the high number of zeros is disturbing: these might be hidden non-responses. I analyze the potential causes of the zero expenditures in probit estimations for any of the these expenditure items where the dependent variable was zeros vs. "positive values" on income, age and a set of demographic characteristics. However, one would suspect for example that 'food outside home'-expenditures strongly depend on income, schooling and employment status. This would indicate that at least some of the zero values are not hidden non-responses but might well be due to a lower living standard. Since this question also contains expenditures for food at canteens, the responses might be influenced by the work environment. Tables 6.8 and 6.9 shows the results of the Probit estimates. For 'food outside home' and 'telecom. services', the income polynomial is highly significant; the higher the income, the lower the probability for zero responses. The minimum effect is at about 5300/4700 €, respectively. This seems to support the hypothesis that not all zero values are due to respondents' uncertainty about the true value, as Winter (2004) proposes for the occurrence of zeros for total non-durable expenditures. Still, dummy variables for the *SAVE* subsamples are significant. While the TPI sample, which surveys skilled interviewees, contains no zero values for three of the items, the random subsample RR 2003 contains significantly more zero value respondents for three items including total non-durable expenditures.

It is not plausible that a household has zero 'food at home'-expenditures, so at least some of the zero expenditures can be attributed to non-response. Moreover, zero values are correlated: giving no answer in e.g. the 'food at home' or the 'total expenditure' question can explain almost half of the zero values of each other and the other two questions.

Given the results, it is not justifiable to drop zero value observations. For the high values, or outliers, shown in Table 6.1, a similar set of regression was done. The structure was much less stable; many variables predicted the output perfectly which is obvious given the low number of cases. For total non-durable expenditures, there is a clear income and schooling dependence: the higher income

and schooling, the lower the probability of outliers. This supports the often expressed reservation against asking total non-durable consumption recall questions. For the following analyses, I will drop outliers for each expenditure item, cf. Table 6.1.

Figures 6.1 shows histograms for the expenditure items, which are also depicted in Figure 6.2, excluding zero values. Both Figures exclude values which are higher than monthly net household income. The figures also show typical effects of recall questions: focal values are relatively frequent. This effect is also shown in Winter (2004) and Battistin *et al.* (2003); the latter also developed methods to account for this heaping and rounding. For total expenditures, 91.16% of the answers were multiples of 50 €, 76.0% multiples of 100, 32.0% of 500 and 20.2% at 1,000 €, not excluding zero values. These rounding effects are less frequent for 'food at home' (43% for multiples of 100 €), but the scale is of course finer.

As an executive summary for the *SAVE* expenditure data, two things are worth noting. First, nonresponse is by no means comparable to the findings by Winter (2004) or Battistin *et al.* (2003); in fact, nonresponse is completely ignorable concerning almost all expenditure items, including total non-durable consumption. Still, zero values are relatively frequent, and it is hard to tell whether these are hidden nonresponses or true zero expenditures. Compared to other data sources, the experiment including these expenditure questions seems very promising, and the next section will confront these values to the German diary based *EVS*.

6.3.2 Expenditures in the *EVS*

This study has been done whilst detailed data on different income and expenditure items were not yet available for the 2003 wave. Hence, I use the *EVS* 1998 and multiply all different consumption items being explained in the following by separate price indices[114] for every subgroup.[115]

Total non-durable consumption in the *EVS* was computed as the sum of the following sub-items. Included are all items of the group 'food, drinks, tobacco'; all items of the group 'clothing and shoes'; 'total costs of health care' (out of pocket health); 'total energy expenditures'; 'total education expenditures'; 'total expenditures for food outside home, drinks and lodging'; 'goods and services for housing'; 'traffic'; 'communications'; 'other goods and services[116]; 'expenditures for leisure, entertainment and cultural events'; 'other goods and services'. These categories have to be corrected, i.e. reduced, by expenditures for durable consumption goods: cars, bikes and motorcycles; phones and fax machines; TVs, VCRs, camera and camera equipment, and other durables like music instruments; bijou, watches.

When trying to impute the coefficients estimated for the four subcategories *food at home, food outside home, telecommunications* and *energy*, one must take account of the fact that in the 1998 *EVS*, food at home as well as food outside home, drinks are included as well. In order to replicate the measures used by Browning *et al.* (2003) as closely as possible, I use the 1993 *EVS* to calculate the

[114] *Source:* Federal Statistical Bureau of Germany

[115] and also, if measures were composite, for every subgroup within that group.

[116] For the latter category (other goods and services), it is recommended that all goods and services are included separately instead of using the whole group for allowing different price indices and has been done accordingly.

shares of food alone without drinks in these two subgroups, which is then used to approximate the corresponding 1998 expenditures.[117]

Table 6.2: Expenditure shares from income and expenditure surveys

	EVS (1998[a]/1993), N=49720			Canada	Italy	Spain
	Mean	Median	Std. Error	FAMEX (1996)	(SFB)	(ECPF)
Food at home	16.9%	16.3%	0.03%	22.1%	32.1%	57.4%
Food outside home	3.5%	2.8%	0.02%	6.3%	5.0%	0.3%
Telecommunications	2.5%	2.1%	0.01%	3.9%	3.3%	n.a.
Energy	8.3%	7.3%	0.02%	8.2%	8.0%	7.8%
Overall total	31.2%	31.1%	0.04%	40.5%	48.5%	49.3%

[a] In prices of 2003

Source: *EVS*: Own calculations; *FAMEX, SFB, ECPF*: Browning *et al.* (2003)

Table 6.2 compares the expenditure shares from the *EVS* with the corresponding expenditures in Canada, Italy, and Spain. The differences in the single four shares as well as in the overall totals suggest that total non-durable expenditures calculated from the other three samples is measured more restrictively as in the author's calculations.[118]

Comparison between *EVS* and *SAVE* expenditure data Table 6.3 compares the total non-durable expenditures from *SAVE* to the values of the *EVS*. The ratio of the mean/median expenditures, respectively, in both data sets are even lower than what is reported by Browning *et al.* (2003), which again might well be due to the fact that total non-durable consumption is calculated less restrictive here. This supports the general finding that recall-based total expenditures suffer significantly from underreporting.

Table 6.4 analogously compares the sub-item expenditures from *SAVE* to the values of the *EVS*. And again the results by Browning *et al.* (2003) can be mimicked: the ratio between the means and medians of both data sets are about 1, which supports the findings that measures for sub-items, especially 'food at home', work very well for recall-based surveys.

These findings allow to use the method discussed in Section 6.2 to impute total non-durable consumption for the *SAVE* data using as an external source the *EVS* data set.

6.3.3 Estimation of weights and imputation of total non-durable consumption

This section presents the estimation results from the method mentioned in Section 6.2. Different specifications were used to assign weights to different sub-items based on Equation 6.1. The estimated

[117] This implicitly assumes constant expenditure shares for drinks/foods over the five year gap. The shares are 86% for 'food at home' to 'food, drinks and tobacco at home' and 67% for the food expenditures outside home, respectively.

[118] Total nondurable expenditures in Browning *et al.* (2003) are defined as: food at home, food out, water, fuel, electricity, household operations, clothing, transportation (excluding car purchases) medical care, personal care, recreation (excluding purchases of recreational vehicles), reading material, educational expenses, alcohol and tobacco.

Table 6.3: Total expenditure measure

	CAPI 2003		RR 2003		CAPI 2004[a]		EVS 1998[a]
Mean	748.57	49.0%	830.20	54.3%	878.90	57.5%	1529.21
Median	650.00	47.1%	750.00	54.4%	750.00	54.4%	1379.93
Standard Error	28.44		13.97		60.40		3.34
N	483		2184		469		49720
only if expenditures > 0							
Mean	815.04	53.3%	904.97	59.2%	878.90	57.5%	
Median	700.00	50.7%	800.00	58.0%	750.00	54.4%	
Standard Error	29.01		14.08		60.40		
N	441		2001		469		

[a] In prices of 2003

Notes: All values weighted. Absolute numbers and relative values to *EVS* numbers shown.

Table 6.4: Sub-items of household expenditures

	EVS 98[a]		CAPI 2003		RR 2003		CAPI 2004		*SAVE* 2003 and 2004[b]	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Food at home	254.0	229.4	237.6	200	167.0	50	1165.5	400	232.8	175
Food outside h.	59.7	38.8	57.4	30	27.7	8	76.4	50	51.6	25
Telecom. services	35.3	29	48.1	30	29.1	10	165.2	63	40.8	25
Water,fuel,elec.	120.0	101	117.8	62.5	108.3	79.2	202.9	88	121.8	81

[a] In prices of 2003

[b] Conditional on each value being larger than zero and the income fraction being smaller than one

Note: All values weighted.

coefficients for the weights are then applied to the *SAVE* expenditure items to obtain the imputation measure for total non-durable consumption.

Estimation of weights Table 6.5 reports the results of five experiments. For each I report coefficient estimates and the R^2 for the regression.

The first column of Table 6.5 reports the results of regressing total non-durable expenditure on two items, only 'food at home' and 'food outside home'. 56% of total non-durable expenditures can be explained by these two predictors; this means that more than half of the variance of total non-durable consumption can be explained by food expenditures only. Including 'telecommunication services' and 'utilities', the explanatory power raises to 62% (2). Browning *et al.* (2003) argue that the assumption of linear Engel curves made in Equation 6.1 might be too hard a restriction. This is the reason to include squares and cross products of the four expenditure items in (3). Again, Browning et al.'s results are confirmed as the less restrictive functional form only adds minor additional explanatory

Table 6.5: Regression results for Equation 6.1

Specification	(1)	(2)	(3)	(4)	(5)
Food at home	2.695	2.189	2.089	1.908	1.596
Food outside home	4.057	3.717	3.871	3.665	3.777
Telecom. services		5.509	4.741	5.529	4.875
Water, fuel, electricity		1.560	1.538	1.358	1.173
Squares and cross products	N	N	Y	N	Y
Demographics	N	N	N	Y	Y
R squared	55.6%	62.4%	63.6%	63.3%	64.4%

[a] Demographics: Age and age sq., household size, home ownership (D)

Source: *EVS* 1998(1993)

Notes: Results from weighted regressions. t- and p-values not reported: all coefficients significant at the 0.1% level

power, which also is true for including demographic variables (4).[119] Finally, in specification (5), I combine all extensions of specification (1).

Table 6.6 compares the coefficient estimates and the R^2 from the second column in Table 6.5 (no squares and cross products, no demographics) with the results of the regressions from Browning et al. conducted for other national surveys. In order to get comparable results, Table 6.6 shows unweighted estimation results. While the R^2 is fairly lower, especially the 'food at home' weight factor is similar in all three data sets.

Table 6.6: Comparing the results from the *EVS* to *FAMEX* and *SFB*

	FAMEX	SFB	EVS
Food at home	2.190	2.220	2.231
Food outside home	3.280	2.327	3.731
Telecom. services	3.030	4.347	5.978
Water, fuel, electricity	2.720	1.489	1.578
Squares and cross products	N	N	N
Demographics	N	N	N
R squared	74.3%	63.4%	58.7%

Source: *EVS*: own calculations, results from unweighted regression; *FAMEX, SFB*: Browning *et al.* (2003)

Note: t- and p-values not reported: all coefficients significant at the 1% level

Imputation of total expenditures The specification chosen for the imputation in the *SAVE* data set was the one in the fourth column of Table 6.5, reestimated with the population weights from the *EVS* 1998; the four demographic variables were included, but not cross-terms and square products. The reason for *not* including them lies in the limited improvement over the specification without these terms and in the trade-off argument that outliers within the sub-items[120] could bear the risk

[119] In fact, I tried many different sets of demographic variables. The explanatory differences were very small, and I decided to use the one with the best trade-off of additional explanatory power and no data loss for missing observations in one of the demographic variables in *SAVE* .

[120] which were not controlled for if income was not smaller than each expenditure item

of imputing implausible values. The results are shown graphically in Figure 6.3. The dispersion of the imputed expenditure values in *SAVE* is not much larger than the one from the total non-durable expenditure measure in the *EVS*, and even more, dropping zero value sub-items does not affect the shape either.

The comparison between the imputed total non-durable expenditures, the sum of sub-items in *SAVE* and the values from the one-shot question is shown in Figure 6.4. The distribution for the one-shot question is shifted to the left compared to the imputed values' density; the mean of the one-shot question's values consistently is about one half of the mean of the imputed values. Comparing the simple sum of the sub-items to the one-shot question makes clear that indeed households do not forget to include other expenditures in the one-shot question which were not asked for before.[121] The mean again is just about one-half of the mean of the one-shot total expenditure question.

A possible threat to the use of the imputed consumption data is that nothing prevents imputed total monthly non-durable expenditures from being smaller than monthly household net income. In the present scheme, 27.4% of the imputed expenditures exceed income, which is a rather high number. Going back to Equation 6.1 and to Table 6.5, respectively, and including income (and, additionally, income and income squared) in the weights regression, this does not eliminate the problem. As Table 6.7 shows, values are still below income in about 20% of all imputed expenditure cases. For the one-shot question, this is applies only for about 5%. Additionally, including income in the imputation procedure would entail an endogeneity problem when analyzing e.g. saving rates, constructed as the residual of income and expenditures, in dependence of income. See also Section 6.2, Footnote 111.

Table 6.7: Expenditures below income

		Imputed expenditures		
	One-shot question	No income included	Income	Income and income squared
Below inome	133	590	446	422
All	2960	2155	2155	2155
Percent	4.5%	27.4%	20.7%	19.6%

Notes: Three different imputation specifications are compared: no income (only expenditure items plus demographic variables), income, and a second order polynomial for net income; (both in the estimation step in the *EVS* and in imputation step for *SAVE*).

6.4 Conclusions

Collecting diary-based detailed information on household expenditure is a costly and time consuming procedure. The use of one-shot expenditure questions, in contrast, has its limits, which was shown by many authors including the one from this chapter. Instead, a non-exhaustive list of sub-items which includes 'food at home', 'food outside home', 'utilities' and 'telecommunication services' can be used to impute total non-durable expenditure with expenditure allocation weights estimated by diary-based

[121] Remember that questions' order was first to ask for the sub-items and then for total expenditures.

income and expenditure surveys.[122] The *EVS* 1998 was used to compute the weights which were included in the *SAVE* 2003 and 2004 subsamples to impute total expenditures. Comparing absolute expenditures in the *EVS* and in *SAVE* for the used sub-items, this procedure seems justifiable, which then translates into similar total expenditure distributions. The drawback of this procedure is that nothing prevents expenditures from being larger than monthly revenues, since the estimation of weights cannot account for all household heterogeneity.

[122] It has been shown in the literature that these four items are good proxies of total non-durable expenditures.

6.A Tables

Table 6.8: Probit regression for zero values in expenditure items, part 1

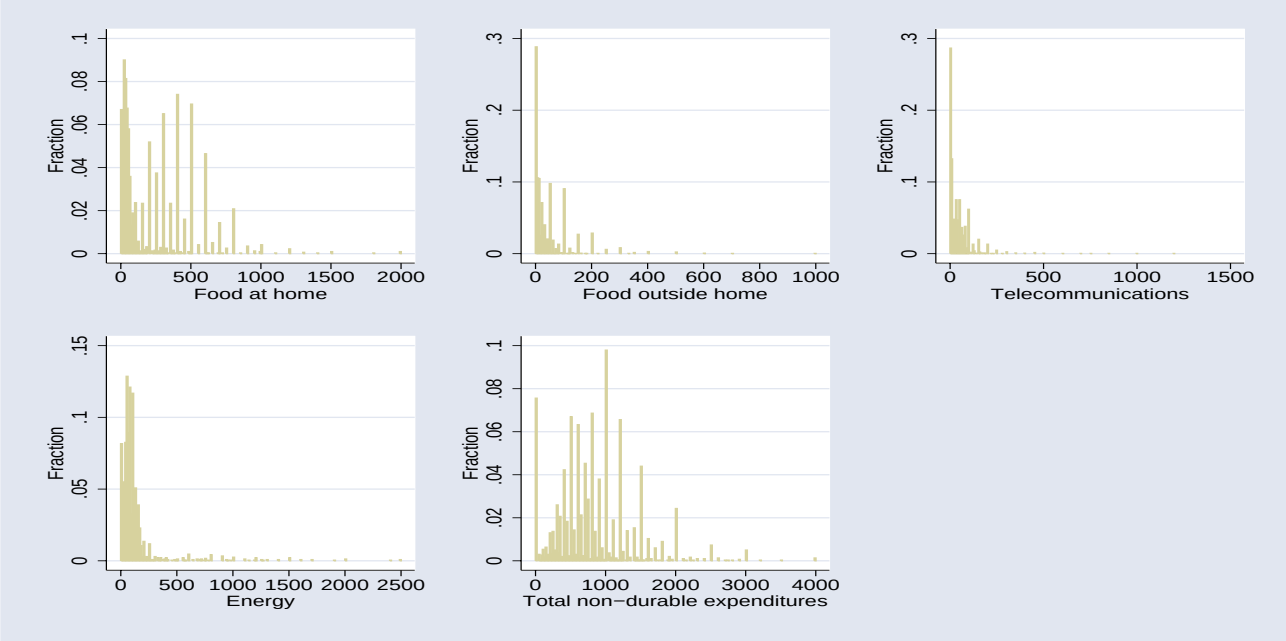
	Food at home		Food outside home		Telecom.	
	Coeff.	$P > z$	Coeff.	$P > z$	Coeff.	$P > z$
Net income / 10,000	-0.343	0.703	**-5.031**	0.000	**-3.352**	0.009
Net income / 10,000) sq.	1.134	0.263	**4.691**	0.000	**3.563**	0.018
Age/10	-0.067	0.686	0.198	0.109	0.247	0.263
Age/10 squared	0.010	0.548	-0.002	0.860	-0.022	0.333
Secondary school (D)	0.080	0.441	**-0.197**	0.009	0.045	0.745
Graduation diploma (D)	-0.030	0.842	**-0.288**	0.015	0.013	0.950
University degree (D)	-0.157	0.300	**-0.333**	0.002	-0.298	0.223
Household size	**0.177**	0.001	**0.210**	0.000	0.089	0.223
Kids (D)	-0.045	0.746	0.028	0.787	-0.221	0.215
Kids living in same house (D)	**-0.362**	0.018	0.024	0.827	-0.083	0.684
Job: blue collar (D)	**0.290**	0.065	0.185	0.139	0.137	0.533
Job: civil servant (D)	-0.105	0.670	-0.177	0.419	0.229	0.490
Job: freelancer (D)	0.423	0.194	-0.704	0.144	dropped	
Job: self-employed (D)	0.205	0.351	-0.113	0.582	dropped	
Retired(D)	-0.038	0.851	-0.248	0.076	-0.142	0.589
Work parttime (D)	0.066	0.735	-0.014	0.927	-0.074	0.796
Work little (D)	0.117	0.549	**0.253**	0.081	-0.361	0.276
Work not (D)	0.020	0.906	**0.371**	0.003	0.249	0.254
Unemployed (D)	0.148	0.476	0.076	0.570	-0.022	0.927
Past unemployment 1-6 months	**-0.231**	0.078	-0.121	0.205	0.021	0.905
Past unemp.> 6 months	-0.029	0.853	0.062	0.559	0.144	0.454
Partner	**-0.426**	0.003	-0.105	0.337	**-0.389**	0.030
Widowed (D)	-0.169	0.306	**0.203**	0.075	-0.179	0.366
Separated or divorced (D)	-0.111	0.581	0.132	0.348	-0.062	0.803
Widowd	0.008	0.938	0.095	0.192	0.115	0.397
East Germany (D)	**-0.202**	0.089	**0.275**	0.000	**-0.295**	0.059
Sample: RR 2003	**0.182**	0.071	**0.343**	0.000	-0.137	0.351
Sample: TPI 2004	dropped		**-0.347**	0.010	**-0.881**	0.017
Constant	-1.643	0.000	-1.869	0.000	-1.905	0.001
Obs.	2866		2792		2848	
LR chi2(24)	54.12		464.78		63.95	
Prob > chi2	0.002		0.000		0.000	
Pseudo R2	0.0479		0.1657		0.1019	

Table 6.9: Probit regression for zero values in expenditure items, part 2

	Energy		Total non-durable exp.	
	Coeff.	$P > z$	Coeff.	$P > z$
Net income / 10,000	0.765	0.478	0.021	0.980
Net income / 10,000) sq.	-0.594	0.670	0.821	0.401
Age/10	**-0.470**	0.009	**-0.343**	0.024
Age/10 squared	**0.045**	0.014	**0.031**	0.042
Secondary school (D)	0.143	0.198	-0.028	0.776
Graduation diploma (D)	0.078	0.610	-0.202	0.152
University degree (D)	-0.080	0.596	**-0.295**	0.039
Household size	0.087	0.134	0.056	0.283
Kids (D)	**-0.258**	0.080	-0.170	0.203
Kids living in same house (D)	0.053	0.741	0.063	0.661
Job: blue collar (D)	-0.228	0.195	0.112	0.444
Job: civil servant (D)	-0.158	0.467	0.168	0.393
Job: freelancer (D)	0.186	0.578	**0.543**	0.064
Job: self-employed (D)	-0.209	0.391	-0.382	0.172
Retired(D)	-0.208	0.340	-0.015	0.935
Work parttime (D)	-0.078	0.673	-0.155	0.386
Work little (D)	-0.161	0.443	0.089	0.607
Work not (D)	-0.005	0.976	0.008	0.958
Unemployed (D)	-0.168	0.445	-0.132	0.483
Past unemployment 1-6 months	-0.074	0.562	-0.070	0.548
Past unemp.> 6 months	-0.129	0.406	-0.010	0.942
Partner	**-0.449**	0.002	**-0.248**	0.060
Widowed (D)	0.196	0.208	0.077	0.602
Separated or divorced (D)	-0.074	0.734	-0.130	0.500
Widowd	0.078	0.452	0.118	0.209
East Germany (D)	0.143	0.213	**-0.234**	0.041
Sample: RR 2003	-0.007	0.943	**0.253**	0.009
Sample: TPI 2004	dropped		dropped	
Constant	-0.229	0.613	-0.745	0.052
Obs.	2199		2870	
LR chi2(24)	72.14		65.88	
Prob > chi2	0.000		0.000	
Pseudo R2	0.0668		0.0498	

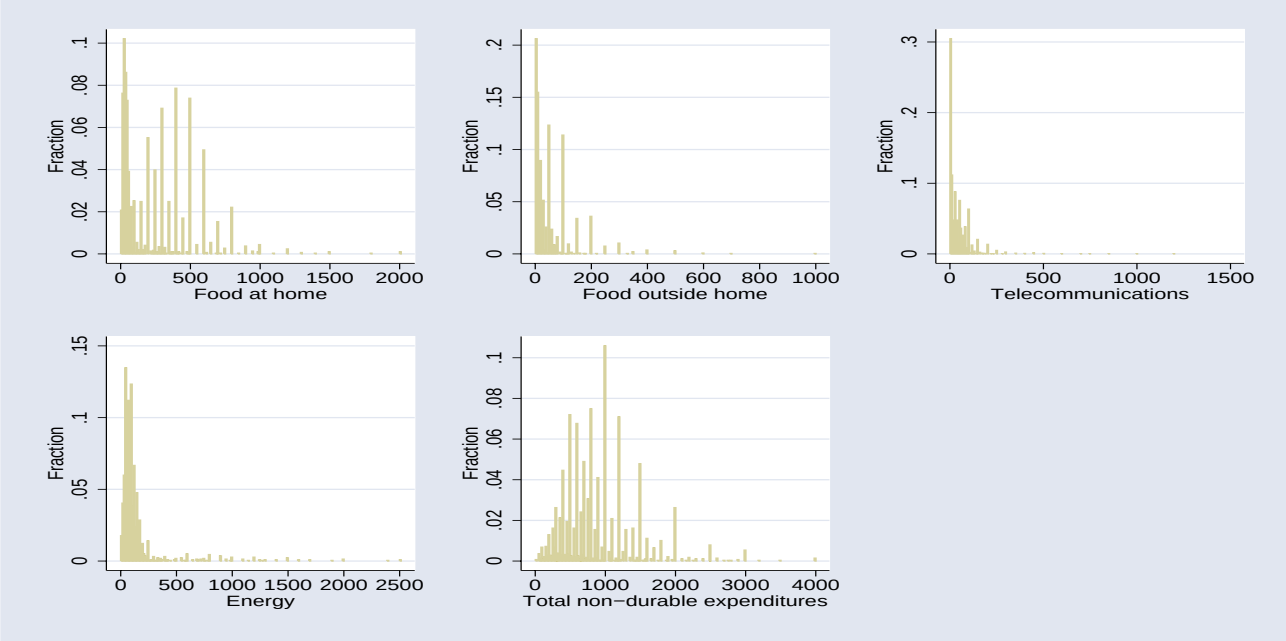
6.B Figures

Figure 6.1: Distribution of expenditure values



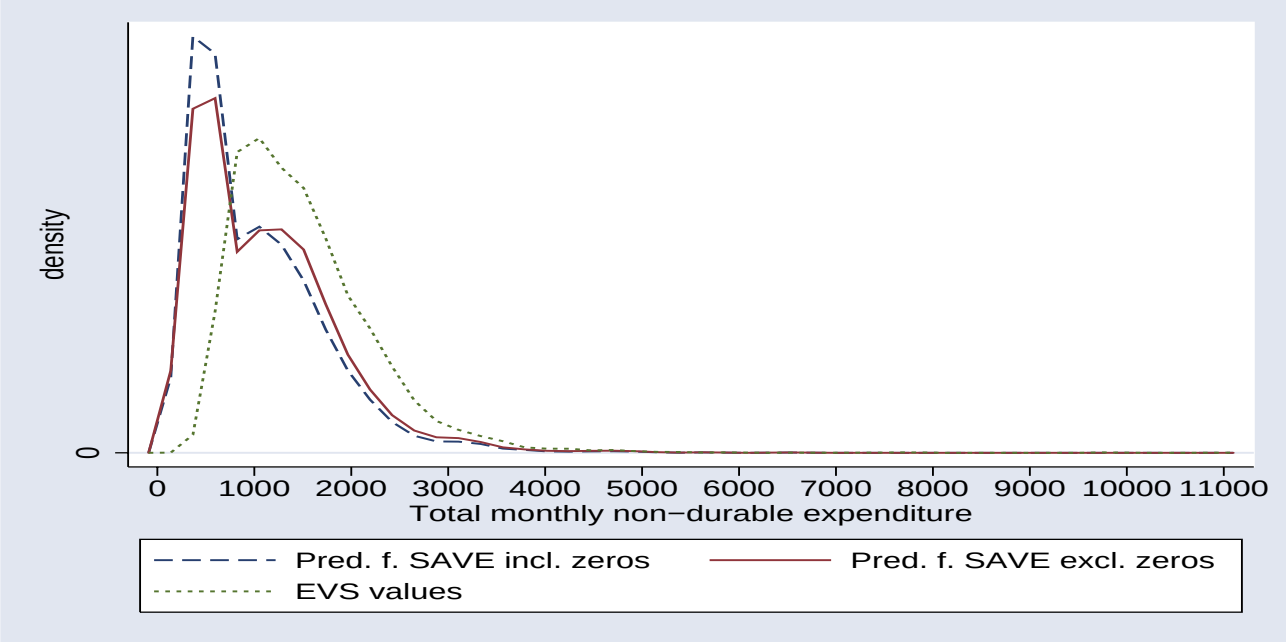
Note: Unweighted values.
Source: *SAVE* 2003 and 2004 subsamples.

Figure 6.2: Distribution of expenditure values excluding zeros



Note: Unweighted values.
Source: *SAVE* 2003 and 2004 subsamples.

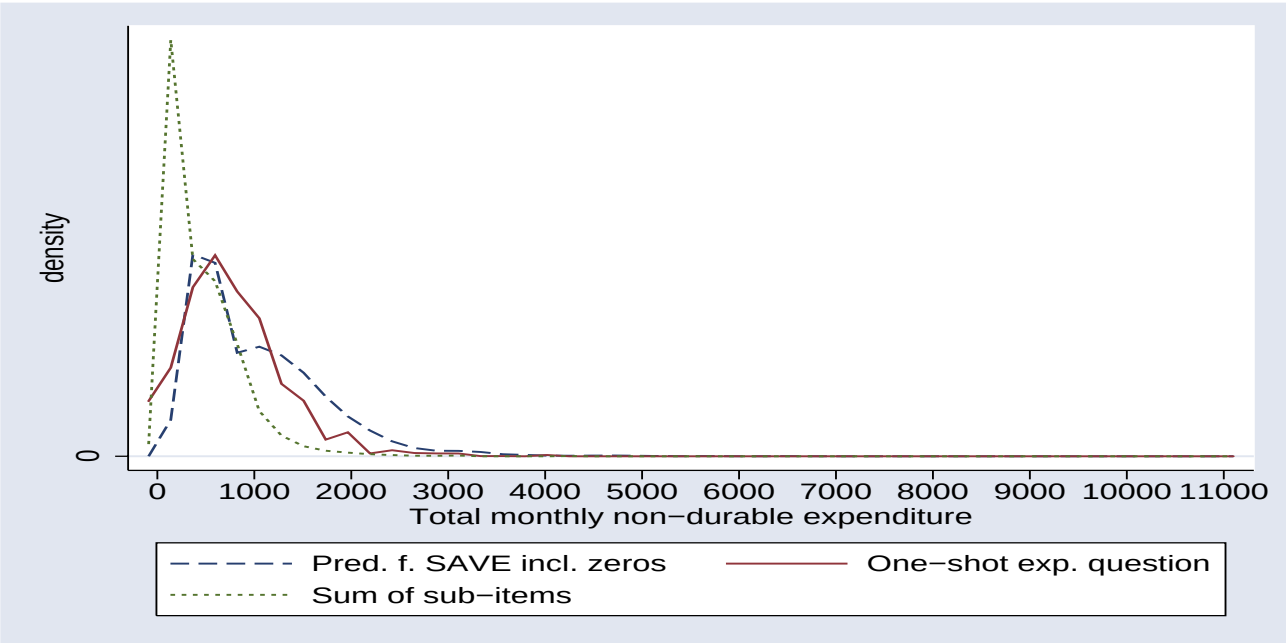
Figure 6.3: Distributions of total monthly nondurables expenditures measures



Source: *SAVE* 2003 and 2004 subsamples and *EVS* 1998

Notes: All values weighted. 1998 *EVS* values in prices of 2003. Kernel density estimates, using the Epanechnikov kernel and optimal bandwith selection.

Figure 6.4: Distributions of imputed expenditures, values from SAVE and the sum of sub-items



Source: *SAVE* 2003 and 2004 subsamples

Notes: All values weighted. Kernel density estimates, using the Epanechnikov kernel and optimal bandwith selection.

6.C Expenditure questions in *SAVE*

Consumption questions were, if not asked within the full P&P interview environment of the TPI subset, part of the CAPI component of the *SAVE* questionnaire. This applies to the CAPI 2003 and the RR 2003 subsamples. The block of consumption questions appeared after the savings questions. In addition to the interviewers reading the questions, respondents were handed 'showcards' which additionally defined the questions more precisely.
Heating costs were asked in a two-step process since first the period was asked for and then the corresponding amount of the bill; this sum, therefore, has to be recalculated on a monthly basis.
The expenditure questions were asked in the way as depicted on the following page.

- *Think of the year 2002(2003). About how much did your household spent on an average month for food you consumed at home?*
  Showcard:
  Count: food and alcohol-free beverages purchased in grocery stores, supermarkets and similar stores
  Don't count: Expenditures for alcoholic beverages like beer, wine and liquors
- *Think of the year 2002(2003): About how much did your household spent on an average month for food outside home, e.g. in restaurants?*
  Showcard:
  Count: meals taken in restaurants, canteens, bars etc.
  Don't count: Expenditures for dropping by bars when nothing was eaten, and expenditures for celebrations like weddings, birthdays etc.
- *Think of the year 2002(2003): About how much did your household spent on an average month for telecommunications, cell phones and internet connections?*
  Showcard:
  Count: Basic and variable fees for fixed networks and cell phones, including text messages; royalties for private internet connections (AOL, MSN)
  Don't count: Purchases of phones and cell phones
- *What is the time period for your heating cost settlement?*
  Weekly; Monthly; every two/three/six months; once a year?
- *What was your last heating bill?*
  Showcard:
  Don't count: Costs for electricity not used for heating (illumination, cooking etc.)?
- *Think again of the year 2002(2003): About how much did your household spent in an average month all in all for all goods and services including purchases in supermarkets, meals in restaurants, telecommunications, heating etc.?*
  Showcard:
  This is the sum of all household's expenses for daily use.
  Count: First four items plus expenses for clothing and health care; royalties for private internet connections (AOL, MSN)
  Don't count: Rent and large irregularly purchases like homes, cars, furniture and large electronic tools like stoves or refrigerators lasting many years.

Chapter 7

Measures for savings and saving rates in the German *SAVE* data set

7.1 Introduction

Savings and saving rates are a key element for the analysis of household behavior. They present the foundation for many different research areas, as e.g., how well households are prepared for old-age (the old-age provision motive), what measures they take to insure against unknown shocks (the precautionary motive), how important consumption smoothing is (the intertemporal substitution motive), among many others, which already have been almost completely by Keynes (1936), to which Browning and Lusardi (1996) added the downpayment motive to complete the list.

For an empirical analysis, micro data on households (or individuals) are needed to get an insight into peoples' actual behavior. The crucial variable is a precise assessment of households' savings. This sounds a lot easier than it can actually be implemented in a survey not only collecting savings and income, but also many additional variables inevitable for a broad behavioristic analysis.[123]

This paper goes to the very roots of respondents' understanding of savings. Do they realize savings just as the residual left-overs on their accounts after subtracting all monthly expenditures from all monthly net income flows? If so, then savings will heavily be underestimated. E.g., credit repayments in that sense would be expenditures but clearly are a savings component[124]. The same applies for monthly contributions to private savings and pension plans, whole life insurances or building society contracts, which might be perceived as an additional income tax even though these are discretionary contributions to build up private wealth. Even more complicated to assess are employers' contributions to occupational pension schemes since these contracts normally only tell the employee the future benefits of this kind of pension scheme.

Four kinds of measures for household savings are typically calculated. The first one is the first difference in net wealth since all savings have to be allocated to any form of financial or real investment. Not looking at the stock of wealth but at the contributions and withdrawals gives the second measure for savings (flow measure). The third one is the residual, or *epsilon* measure of savings which subtracts

[123] To visualize the complexity of a simple savings question, just ask yourself: "How much, on average, do *I* save per month?'

[124] Even though, loan and mortgage repayments have been incorrectly included in the one-shot question on total non-durable expenditures in the Canadian Out of Employment Panel (*COEP*), see Browning and Crossley (2002).

all expenditures[125] from all compositions of net income[126]. Concerning these measures, one has to be aware of the symmetry of positive and negative values, and to adjust flow and stock measures. This is the subject of Börsch-Supan *et al.* (1999)'s work. See also Brugiavini and Weber (2003) for a discussion on these saving concepts. The fourth one is to directly include a question on household savings in the questionnaire, leaving it to the respondents to deliver a reasonable assessment of private savings.

Much research has been done for almost all of the above mentioned savings motives, but they are far from being consistent even within a research area restricted to only one of the topics. Apart from differences in the econometric assessment of the topics and variations in their specifications, the data base used is frequently different. But not only the the data base itself, also the dependent variable used differs. From which of the four measures are savings calculated from? In a very recent study, Alessie *et al.* (2004) compare savings and saving rates for the Italian Survey on Household Income and Wealth *(SHIW)* and the Dutch Socio-Economic Panel *(SEP)* using different saving concepts. For the flow measure, they compute the saving rate as $\frac{Y-C}{C}$ where Y is the sum of personal incomes of parents and child, and C is non-durable consumption. The procedure to the saving rate relative to consumption instead of income was proposed by Attanasio (1998) to avoid the problems of outliers and zero income observations. Alessie *et al.* (2004) also implicitly equal consumption and expenditures. This might sound tautological, but as Aguiar and Hurst (2004) very recently pointed out, the dramatic decline in expenditures at the time of retirement is matched by an equally dramatic rise in time spent on home production. This argument goes back to Becker (1986), who states that consumption is the output of a "home production" function that uses both expenditure and time as inputs. The innovation of that paper is that the authors empirically disentangle changes in actual consumption from changes in expenditures. Taken together, the results highlight how direct measures of consumption distinguish between anticipated and unanticipated shocks to income, while using expenditure alone obscures this difference and leads to false rejections of the PIH.

This paper presents concisely the main findings for savings and saving rates estimated with the *SAVE* survey. Section 7.2 quickly summarizes the definition of savings and shows which of the four saving measures can be computed with *SAVE*, while Section 7.3 concentrates on the assessment and measurement of the one-shot savings question in *SAVE*. Section 7.4 discusses the results from different correction measures for the one-shot savings question, and Section 7.5 summarizes this chapter's issues.

7.2 Saving measures in *SAVE*

The *SAVE* survey bears the possibility to calculate savings in three of the four ways described in Section 7.1. The first one is the first difference in financial wealth components, while the second one would be the *epsilon* measure of savings as the difference between all components of household income

[125] Expenditures then are typically defined as total non-durable expenditures, excluding contributions to any saving accounts.

[126] Income is the sum of all different income components, like wage and capital income, transfer income, and social security benefits.

and total non-durable expenditures, which are asked in *SAVE*. The third one is a one-shot saving question asking for total savings a household achieved in the last year, without explicitly guiding the respondent by using a comprehensive list of components which typically should be accounted for.

This section quickly depicts and explains potential problems for each these three savings measures.

7.2.1 First difference in wealth

Financial wealth is asked for the beginning and the end of the calender year previous to the survey field time. A savings measure often found in the empirical literature on savings and savings behavior is to calculate savings as the difference between financial wealth at the end and the beginning of the preceding year. Ignoring changes in the credit situation or real wealth allocation, including homes, this shows even conceptually a massive flaw. Any reallocations from real to financial wealth or the other way around would be registered as savings/dissavings, even though this is clearly without effect on the total wealth situation. The same is true for short selling, or credit financed investments. Other problems are of empirical nature. Values for the total amount in certain wealth categories like saving accounts, stock, or mutual funds, certainly are as much due to recall error as the one-shot savings or expenditure questions. Additionally, it is impossible to distinguish between 'active' and 'passive' savings (putting money into that financial wealth category, or does market appreciation/depreciation account for higher/lower values at the end of the year?).

7.2.2 Residual measure

A third savings measure would be to compute savings as the residual measure (difference between savings and expenditures). This savings measure requires two variables to be reliable. The first one is income; see Chapter 2 for a discussion on income values in *SAVE* .

The other variable, expenditures, has been proven to be a rather imprecise measure for households' expenditures, cf. Browning *et al.* (2003) or Winter (2004). Knowing these results before including the expenditure questions, there were also four sub-items included from which is known that they are typically remembered very well; based on these sub-items, Browning *et al.* (2003) proposed a method to impute total non-durable expenditures in a way to fitting factual values much better than the one-shot question's values. This is done in Chapter 6 for the *SAVE* data using the *EVS* 1998 as an external data source.

Even if expenditures were assessed correctly by this procedure, there is still a conceptual problem. Durable consumption goods are difficult to assess. First of all, there are no data on durables in *SAVE*. Second, durables affect savings not by the time they are purchased, since they represent at that time wealth in equal size, neglecting transaction costs. More relevant is the useful economic life to calculate the depreciation rate, which about represents the periodical consumption value. But these data are extremely difficult to assess, for which not even diary-based consumption and expenditure data are a reliable data source.

7.2.3 One-shot saving questions

The *SAVE* questionnaire also contains a direct one-shot question to yearly savings. This might well be subject to the same problems being discussed in the context of the one-shot expenditure question.

The following section will explain more deeply the possible problems and correction methods for this one-shot savings questions.

7.3 Reliability of one-shot savings measure

This section argues that filtering might cause an underestimation of savings, and discusses the treatment of credit repayments, contributions for life insurances and different pension plans. I will also compare the values to the official numbers from the German Federal Reserve bank.

7.3.1 Problems due to filtering

The direct question for total last year's savings in *SAVE* follows a set of preliminary questions. These include who in the household actually makes the financial decisions, whether respondents talk with other persons inside/outside the household about financial concerns, and who files for the income tax declaration. Next, respondents were asked about the reception of one-off payments like inheritances, tax refunds etc. as well as about the allocation of these one-off incomes. Questions checking whether the household pursuits a certain savings goal and by what time that goal shall be reached follow. In addition, a 'meet-ends' question is included as well as a filter question for a self-assessment of actual savings behavior[127]. The one-shot question was asked as: 'And finally: Could you tell us how much money you and your partner together have saved in the year 2000?', only for respondents who did not choose the fourth or fifth item of the filter question[128] and did not refuse to answer to the filter question.

Table 7.1 shows the different stages which lead to zero values of savings (be these zeros correct or incorrect). Due to filtering, 19.5% of all respondents were not asked for their savings.

To see what the filtering implies for the resulting data, I apply the statistical error types I (rejecting a true hypothesis) and II (not rejecting a false hypothesis) to the filtering process. Filtering bears the chances of a type-II error (household does not save, but should do so according to the filter process; this is depicted in Table 7.2) as well as of a type-I error (household saves but was cancelled out in the filtering process). The type-I error will lead to an underestimation of savings and saving rates since for these households zero savings are assumed.

[127] 'I/we save a fixed amount regularly,', 'I/we put something aside each month but I/we decide on the amount according to the financial circumstances', 'I/we put something aside when we have something left over to save', 'I/we do not save because we do not have enough scope financially to do so', '/we do not save because we would prefer to enjoy life now'

[128] "can't save / don't want to save"

Table 7.1: Origin of nonpositive numbers for one-shot savings question

	TPI 2001	CAPI 2001	CAPI 2003	RR 2003	TPI 2004
N	660	1169	486	2184	483
obs. lost due to filterin	11.2%	17.0%	18.4%	24.2%	17.0%
ref. in filter	1.4%	0.6%	0.6%	0.6%	1.0%
# HH for one-shot savings question	586	970	396	1655	401
does not apply / did save nothing	20.3%	21.2%	20.3%	19.5%	22.0%
refusals	4.3%	17.2%	18.5%	15.5%	3.2%
positive values for savings	442	597	241	1076	301
imputed zero saving values					
in % of all obs. for savings	29.4%	40.0%	40.8%	43.6%	34.5%

Source: All *SAVE* subsamples

Table 7.2: Display of the type-II error

	TPI 2001	CAPI 2001	CAPI 2003	RR 2003	TPI 2004
regular savings					
N	391	417	186	769	249
% of zeros in savings question	14.3%	10.1%	9.1%	10.9%	16.1%
% refusals	4.1%	18.7%	21.5%	14.3%	3.6%
flexible saings					
N	57	277	110	437	52
% of zeros in savings question	14.0%	14.1%	16.4%	14.7%	19.2%
% refusals	3.5%	19.9%	20.9%	19.2%	3.9%
saves occasionally					
N	89	276	98	449	84
% of zeros in savings question	53.9%	45.3%	45.9%	38.8%	39.3%
% refusals	5.6%	12.3%	10.2%	14.0%	2.4%

Source: All *SAVE* subsamples

Notes: The share of zero values should be rather low for the first two categories. The difference of the shares is very small for the filter-categories "saves regularly a fixed amount" and " saves flexibly". For the third filter category "saves occasionally", the share of zeros is very high (43.5%).

Households claiming not being able (or willing) to save show the following pattern which is depicted in Figure 7.2. The dependent variable refers to the filter question and takes on the value 1 if the respondent answered not being able / willing to save and 0 otherwise.

Significant variables which raise the probability for not being able/willing to save are dummies for: retirement; unemployment and partial unemployment; whether the respondent is partially, little ('geringfügig'), or not employed; separated/divorced; and a dummy for credit repayments. Negative significant dummy variables are: self-employed; females; households living in Eastern Germany; and dummy variables for the ownership of financial wealth categories all financial wealth categories except bonds. Second-order polynomials for income and age are also negative significant, but not depicted in Figure 7.2.

Insignificant variables, which are not depicted in Figure 7.2 are: dummies for schooling; dummies for the job type except self-employed; dummy variables for kids and kids living in the same house; a partner dummy; and dummies for the different *SAVE* subsamples.

An important insight from the analysis of the savings ability is that households might *not* consider credit repayments being savings. This is shown by the positively significant coefficient of credit repayments (since the dependent variable takes on the value one if not being able /willing to save). Not adding credit repayments to savings therefore would lead to an underestimation of savings.

When looking at the values given for the one-shot savings question, one can see that about one fifth of respondents answered to having zero savings or below[129], further 13% refused to give values. In the following analyses for zero values and refusals will be separated.

Figure 7.3 shows marginal effects of different significant regressors for the probability to give zero values. Next to the results in Figure 7.2, one can see that values for savings ability are consistent with the probability to give zero values: households saving regularly a fixed amount answered significantly less frequently with zero values than households who save only occasionally. While credit uptakes influence saving measures consistently, credit repayments are not understood as savings (as shown before); again, the coefficient is positive significant.

The corresponding analysis of refusals for saving values will only be mentioned here. Only four variables from the set of 39 explanatory variables were significant: unemployed, occasional savers, households from the two TPI subsamples and households with credit repayments have a lower probability to refuse the answer. For occasional savers and credit repayers this can be explained by the fact that these types have a higher probability of answering with zero values.[130] Income, age, schooling, and job variables have no influence on refusals.

Looking at the values from the one-shot saving question, one can observe in 27 cases saving rates which exceed net monthly income when savings is recalculated for a mensual measure.[131] For three of these 27 cases high savings can be explained by high one-off income sources.[132] For the remaining 24 cases, saving rates are up to 1100% what in turn can heavily affect means. Four reasons for these values can be thought of. (1) Respondents gave wrong numbers which were recorded "correctly" (in CAPI interviews). (2) Respondents gave correct numbers which were recorded incorrectly (in CAPI interviews). (3) Respondents thought of the correct values but gave wrong numbers concerning one or two decimal places; these values were then recorded "correctly" (in CAPI interviews). (4) Respondents thought of and registered correct values which were then incorrectly scanned and transferred to computer data by the survey agency (in P&P interviews). There is a fifth possibility which also occurred with the data at hand, but was recognized by the author and corrected accordingly by the

[129] Cf. Table 7.1. Negative savings are censored to zero due to the question ("did not save / have dipped into savings")

[130] Running a multinomial logit regression with the three choices "Positive values", "zero/negative savings", "refusal", the results are in line with the findings from the binary model.

[131] For the procedure for checking income values, see Chapter 2.

[132] *SAVE* asks for one-off payments like inheritances, lottery gains, or tax refunds. 78% of all surveyed households declared not having received any one-off income sources; the most common income source of the group which received any or several one-off payments tax refunds was the most common case.

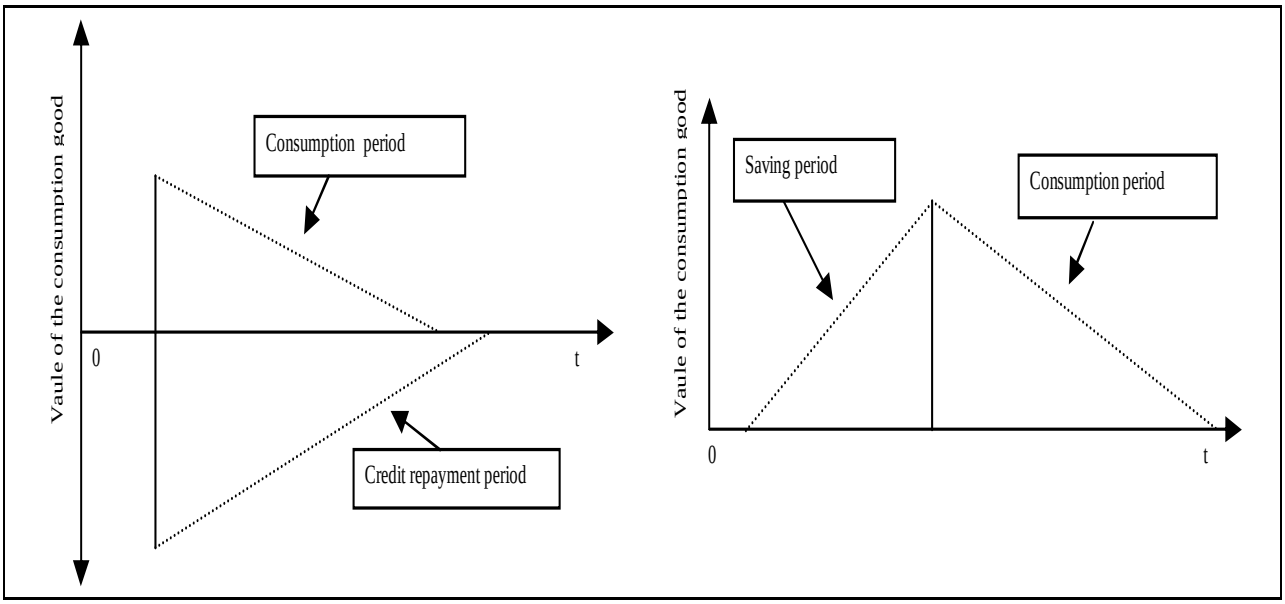
survey agency: variables can be coded incorrectly, which gives implausible values for any observation or no variation between observations. This is the most unproblematic case.

The 24 implausible cases distribute to the modes P&P with 10 cases (1.4%) and CAPI with 14 cases (0.8%). There are two ways of dealing with these 24 cases. First, all of these cases can be coded as missings. Second, if it is believed that errors in the transmission process are responsible for the high savings / saving rates and if the error lies in the imprecision of one decimal place, savings can be divided by 10. The latter way was applied here.

7.3.2 Perception of savings

Figure 7.1 might be helpful to illustrate the challenge of precisely measuring savings. A household is planning the acquisition of a consumption good which, since depicted as a durable good, he will consume for a certain time period. This acquisition can be financed in two ways. First, he can take up a credit (Figure 7.1, left hand side), or he can save the financial means needed (Figure 7.1, right hand side). When only looking at the latter case, one would neglect households with a higher time preference who prefer to consume now rather than after the saving period. Credit repayments thus are savings as well as the accumulation of assets.

Figure 7.1: Financing a consumption good



Notes: The depicted consumption good is a durable; its consumption period about equals the depreciation period. If consumption is immediate (voyages, dinners etc.) the consumption period reduces to 0, such that the downward sloping line of the consumption period would become a vertical line. The value of the credit is larger than the value of the consumption good if interest rates are positive, as depicted.

Credit repayments

Credit repayments only compare to savings if they are not produced by wealth reallocations. Balance sheets would be shortened if credits were repaid through the liquidation of wealth. This is why I only

add credit repayments to savings if they do not exceed 50% of net income.[133] If zero saving values were given, savings just equal credit repayments. This procedure is costly in the sense of data loss due to consistency reasons if observations are dropped if values for savings *or* credit repayments were refused.

At the same time, credit uptakes as such would typically not affect savings, as in most cases the uptaking of a credit would be related to the acquisition of a durable good of similar size (equal size + interest rate markup + adminstration fees). This would result in a 'longer balance' since the liquidity inflow faces an expenditure position (capital asset) of equal size (e.g. durable consumption goods like furniture, cars, technical equipment etc.; housing). Credit uptakes therefore will not be subtracted from savings; only repayments will be added.[134] In total, 32% of respondents declared to repay credits, and 5% to have new credits taken up. 64% of all home owners are indebted, and 38% of tenants.

Contributions to life insurance

Having controlled for credit repayments, it is still unclear what exactly households have in mind when thinking about savings. Is it the sum which remains on the account at the end of the month (the balance of labor, capital and transfer income and total expenditures)? This certainly would underestimate savings, since it neglects credit repayments, as argued. Additionally, other expenditures are not completely consumption expenditures. Regular contributions to certain capital investments (whole life insurances, saving plans, building society contracts) could well be perceived as additional taxes on income. Heckman selection regressions for gross saving rates[135] on an income and age polynomial and a set of household and individual characteristics and a set of dummy variables for the ownership of each financial wealth category and for credit repayment show that five of the six financial wealth dummies are positively significant except for whole life insurances. This might be an evidence that contributions to whole life insurances are possible not included in respondents' saving values. This section will explain how this can be remedied.

Whole life insurance represents the Germans' most important instrument of private old-age provision, cf. Walliser and Winter (1999). 52% of all male respondents answered to own a life insurance contract as well as 34% of all respondents (weighted; unweighted: 38%). This share is below the values observed in the *EVS* 2003 where the share for all respondents is at 46%. The ownership of life insurance contracts strongly differs between the *SAVE* subsamples.

Table 7.3 shows the differences in ownership rates between the subsamples. It should be taken into account that the first line only shows values for the whole subsample; thus the observation that ownership rates dropped between 2001 and 2004 in the TPI subsamples by 23% is not valid. This is shown in the panel comparison where, of course, only values are shown for households being observed in both years. For these households, some differences occurred between 2001 and 2004: 33% claimed

[133] This rather high share of net income can be reasonable since high debit interest rates might force households to quick repayment.

[134] There is no way in the data to tell whether the credit uptake is used to finance a durable or non-durable consumption good, and thus I implicitly assume the usage for a durable good.

[135] The original values given to the one-shot question

Table 7.3: Life insurances: ownership changes

	TPI 2001	TPI 2004	CAPI 2001	CAPI 2003	RR 2003	all
Ownership rates	50.60%	27.50%	39.20%	32.40%	26.40%	33.50%
Panel comparison		$\longleftrightarrow$		$\longleftrightarrow$		
only in 2001 / withdr.		33.10%		17.20%		
neither in 2001 nor in 2003/2004		40.30%		52.40%		
in both 2001 and 2003/2004		22.50%		24.50%		
onlyin 2003/2004 / new contr.		4.10%		6.00%		

Source: All *SAVE* subsamples.
Note: Weighted values.

to having owned life insurances only in the first year, only 4% in the year 2004 if they had none in 2001.

The decline in life insurance ownership has also been documented by Braun *et al.* (2002), Figure 22b, with values measured with the *EVS*, but the 2003 wave reports even lower values than predicted by Braun *et al.* (2002) (52% predicted for 2003, 45% reported in the *EVS* 2003[136]). Similar results are reported by the Gesamtverband der deutschen Versicherungswirtschaft (GDV) (2003). For the past year, trends for life insurances are summarized as follows: (1) significant growth of contributions for new contracts where one-off payments grew higher than the regular contributions; (2) significant reduction of new contracts; (3) growth of new contributions, and (4) growing contract payments (Leistungsausahlungen).

Contributions for life insurance contracts are normally paid on a regular monthly basis and therefore have a similar character to contributions for the public pension system (GRV), which might lead households to the thinking not to include the contributions into savings, but to view them as some sort of tax.

When considering wealth data, one should keep in mind how they were asked in the surveys. *SAVE* is, as most general-purpose surveys, not diary-based, but collects recall information. Values are thus much more due to response error. When looking at the wealth information closely, one can see that in 8 cases, wealth at the beginning and the end of last year differ by the factor 10, 100, 1000 and 10,000 (ex.: wealth at the beginning of the year = 12,000 €, wealth at the end of last year: 12 €), so it must presumed that households think of the same value in both cases but the observed value is wrong by one of the four errors mentioned above. These eight values have been corrected accordingly.[137]

Contributions to life insurances approximately calculated by the difference in life insurance wealth at the end and the beginning of the year. If numbers were correct, the difference includes the internal rate of return as well as the reduction of the cash-in penalty. One has to take account of the possibility that wealth was removed in that period (negative contributions), and further that this difference,

[136] Own calculations

[137] This procedure was done accordingly for all components of financial wealth. In *SAVE* , financial wealth consists of six categories in 2001; in 2003, private old-age provisions was asked in a more detailed manner, which raises the number of financial wealth categories to eight. If only one value was available for the beginning or the end of the last year due to refusal, the missing value was replaced by the observed one thereby assuming no wealth difference.

recalculated for a monthly basis, does not exceed a certain income share. If this would be the case, then this wealth difference might only be realized by the reallocation of wealth and not by savings from current disposable income. This share of income which is to be assumed with $\frac{1}{3}$, which is less than the assumed share for credit repayments with $\frac{1}{2}$ since the contributions would take place on a regular basis. High debit interest rates might force higher repayment rates which is the reason to allow for higher shares.

Old-age provision

Occupational pension schemes In 2001, the *SAVE* questionnaire included the financial wealth category "Other contractually defined private old-age provision, e.g. special old-age provision assets or private pension policies". After 2001, this question was split up into three separate questions: (1) occupational pension schemes[138]; (2) fiscally subsidized private old-age provisions ('Riester-Rente')[139] and (3) other contractual private old-age provisions[140]. Along with the split, the design was slightly changed. Individuals are not only asked for the pension wealth at the beginning and the end of each year, but also for employer and employee (own) contributions separately for each pension scheme.

If the three sub-items for private pensions are aggregated in 2003 and 2001 and compared to the year 2001, the following results can be seen.

Table 7.4: Private old-age provision: ownership changes

	TPI 2001	TPI 2004	CAPI 2001	CAPI 2003	RR 2003	all
Ownership rates	19.40%	24.40%	11.70%	26.60%	20.20%	19.10%
Panel comparison		$\longleftrightarrow$		$\longleftrightarrow$		
only in 2001 / withdrawal		4.10%		11.90%		
neither in 2001 nor in 2003/2004		65.60%		63.70%		
in both 2001 and 2003 / 2004		13.40%		8.50%		
only in 2003/2004 / new contr.		16.90%		15.90%		

Source: All *SAVE* subsamples
Note: Weighted values.

The results from Table 7.4 can be divided into the three pension schemes, which is presented in Table 7.5.

When we only look at fully dependently employed, the ownership rates of occupation pension schemes rises to 20.4%. This is less than half the values Kortmann (2003) finds (44% for men and 39% for women) using data from private and public employers, pension fonds and 'Pensionskassen', and all administrators of private pension schemes in the public service[141]

[138] This comprises defined pension plans, pension fonds and 'Direktzusagen'

[139] E.g. 'staatlich geförderte und zertifizierte Sparanlagen, die nicht vor Ruhestandseintritt auflösbar sind.'

[140] Which in Germany is: 'Altersvorsorge-Sondervermögen oder private Rentenversicherungsverträge, die nicht staatlich gefördert werden bzw. abgeschlossen wurden, bevor es solche Fördermöglichkeiten gab'.

[141] 'Träger öffentlicher Zusatzversorgungsleistungen.

Table 7.5: Ownership rates of different old-age provisions

	CAPI 2003	RR 2003	TPI 2004	all
Occupational pensions	13.30%	11.20%	14.10%	12.00%
Fiscally subsidized old-age provisions	5.80%	4.70%	5.90%	5.10%
Other old-age provisions	13.20%	7.60%	9.80%	8.90%
conditional on full employment[a]				
Occupational pensions	19.50%	17.50%	20.40%	18.40%
Fiscally subsidized old-age provisions	7.60%	7.90%	9.20%	8.10%
Other old-age provisions	19.00%	13.30%	14.70%	14.50%

[a] For couples: if male partner is fully employed.
Source: All *SAVE* 2003 and 2004 subsamples
Note: Weighted values.

Following the arguments from the last section, one should also consider to impute private old-age pensions and add the contributions to savings. There are two problems associated with that procedure. First, nonresponse is very high.[142] This clearly is the consequence of the problems immanent in the complexity of this topic. Employee and especially employer contributions for occupational pension schemes are certainly less well known than own contributions to other private pension schemes or life insurances. Second, answers are frequently inconsistent: the difference of stocks equals the contributions in many cases, but the question asked for monthly contributions. It can be assumed that respondents gave a crude approximation for the pension stocks, and took the year's differences as contribution measures. These values have been divided by 12, accordingly. Still, the refusal level is an issue; a correction of savings by these contributions is associated with a high loss of observations. An alternative would be an imputation of contributions, measured as certain percentage of net income for households, who own the the respective private pension category. The contributions to occupational pensions relative to *net* income are measured as 3.1%/3.0% (mean/median; weighted values) for the *SAVE* 2003 and 2004 data; N=372, after imputation.

Unfortunately, there is no possibility to check these numbers by an external data source; the *EVS* represents in this respect no outside data source, and numbers for occupational pension scheme contributions have never been estimated relative to net income. The best differentiated data source for occupational pensions so far is Kortmann (2003).

Fiscally subsidized private pension schemes The contribution rates for 'Riester'-pensions can be imputed relatively easy: the so-called 'Riester'-stairs gives an lower and upper bound for private contributions. In 2002, 1% of last year's gross income are the necessary investment to receive full fiscal subsidies. It rises gradually rises to 4% in 2008. Own contributions and subsidies accrue the pension plan. The full regulation is shown in Table 7.6.

For the current contracts, the following steps were made: (1) 'Riester'-contributions refer to gross income, but *SAVE* only collects net income. The ratio of gross to net income for the average income

[142] Conditional nonresponse is at over 70% for stocks and 66% for contributions for occupational pensions. For the other two private pension schemes, values are comparable.

Table 7.6: Minimum and maximum contribution rates to 'Riester' saving plans

Investment year	Minimum contribution rate in % of gross income	Maximum contribution for fiscal subsidies	
2002 and 2003	1 %	525 €	
2004 and 2005	2%	1,050 €	reduced by basic
2006 and 2007	3%	1,575 €	and child allowances
2008 and later	4%	2,100 €	

Note: In principal, lower contributions than the minimum are possible, but subsidies will reduce accordingly and thereby the incentives.

earner was 64%, which was taken as a proxy for all households. Contributions therefore have to be multiplied by the inverse value of about 1.6. (2) Maximum contributions are given by the values shown in Table 7.6. (3) For *SAVE* , the years relevant to impute current contributions are the years 2003 and 2004, only.

If one is interested in projecting values over the life cycle, minimum contribution rates have to be calculated as weighted averages of the numbers in Table 7.6 with the weights being the years to retirement: the younger and thereby the longer retirement age lies ahead, the higher is the weight for the 4% value for calculating *average* contribution rate. Example: minimum contributions are 3.5% for a today 40-year-old entering retirement at age 65, 3.2% for a 50 year old, 2.0% for a 55-year-old, and 2% for a 60 year old.

Other private old-age provisions Like for the situation of occupational pension schemes, external data sources are limited, which requires to take another look at the data in *SAVE*. 281 respondents claimed to own private pensions in 2003 and 2004. In the data for stocks and contributions, we only observe 21 cases containing full information. The procedure is comparable to dealing with occupational pensions.

1. For all observations which include stocks as well as contributions: If contributions equal the difference of year's end and year's beginning, they were in sum (employer and employee's contributions) or, if only one is available, separately divided by 12 since monthly contributions were mixed up with annual contributions.

2. If either employer or employee's contributions were refused, the other value was *not* imputed since I assumed that other private contributions were made without employers' contributions (unlike for occupational pensions).

3. After that, contributions were compared to net income. If that share was higher than 20%[143], the contributions were calculated as the difference between the two stock values, divided by 12.

[143] This is assumed a lower share than for life insurances and credit repayments since the regular contributions to old-age provisions are typically not a high share of monthly net income

4. In the next control stage, the mean of the corrected shares was calculated conditional for shares being smaller than 30% of net income; this conditional mean was used to impute the missing contributions relative to the net income.

5. Contributions being larger than 30% were replaced by the conditional mean share as well.

In total, the mean contribution share (of all original and imputed values) is at 5.2% of net income.

7.3.3 Imputed rent

If comparing the wealth situation at old age, it seems advisable to take the housing wealth into account since this typically represents the household's largest wealth asset. This can be done in two ways. First, one could treat housing wealth as if it was payed out as an annuity over the rest of the life cycle. The second alternative is to use the imputed rent of the owner-occupied housing. This construct compares the housing wealth to comparable market rent payments a household had to pay if he would sell the house and rent a similar object.

Estimates for the imputed rent relative to the worth of the concerning housing are about 3.9% p.a. (median) and 4.7% (mean) based on weighted values in the *EVS* 1998. These values are slightly below those of the Ring Deutscher Makler (RDM).[144]

Comparing values from the *EVS* and the RDM, I presume a value of 5% p.a., or 0.42% per month.

The difference of the imputed rent and the annuity method lies in the time horizon. While it is infinite for the first one, the second one depends on the duration of the annuity payments which is in the case at hand given by the year of death. A simple example would make this clear. Assume a house worth 500,000 €. If it was sold and the sum annuitized for a monthly rent payment with the duration given by the difference of life expectancy minus the retirement entry age, which is assumed to, say, 30 years, it would pay a rent of 2,071.21 € per month if real interest rate was 2.8% percent p.a. The same housing wealth assumed as monthly imputed rent would then be 2083.33 €, nearly exactly the same amount. Of course, this comparison is highly age-dependent: A shorter life expectancy clearly would raise the annuity, while the imputed rent delivers the higher value for later years of death.[145]

There is another way to retrieve an additional income flow from home ownership: reverse annuity mortgages. The reverse mortgage pays a regular income, and typically is available regardless of current income. The amount one can borrow depends on age, the current interest rate, other loan fees, and the appraised value of the home. Generally, the more valuable the home is, the older one is, the lower the interest, the more one can borrow. No payments are needed, because the loan is not due as long as the house is one's principal residence. The loan is repaid when the occupation of the home ceases (by death, or by selling the home).

[144] Values are shown separated for houses/flats which were built before/after 1940. The RDM publishes these numbers for 256 cities and regions. In Eastern Germany, these values are much higher than in Western Germany. This means that in the Eastern part, purchase prices for houses are much lower. Munich, for example, shows very low imputed rent values which can be translated in the belief that housing worth will not significantly deteriorate and can collect rent payments for a long time horizon, or otherwise, the market is assumed to be less volatile.

[145] Assuming that the annuity depends on individual subjective life expectancies and not on standardized life tables.

7.4 Computing savings and saving rates

After the discussed steps of calculating different possibly unconsidered monthly savings, the question arises: what kind of difference does it make? Does it only affect households who have rather high monthly savings anyway, or does this also significantly change the picture of savings we get for poorer households? I will also quickly mention the results for the other two savings measures in *SAVE* (difference in financial wealth and residual meausue).

7.4.1 Corrected one-off savings measure

The effects of four different saving measures are shown in Figures 7.4 and 7.5. Differences between net and gross savings [146] are 1.4%/0.2% (mean/median), N=2391. The means and median values are summarized in Table 7.7.

Table 7.7: Saving rates with different savings components

	Gross saving rates	Net saving rates	Net + LIC[a]	Net + LI + POAC[b]
Mean	9.4%	10.8%	11.4%	12.3%
Median	3.5%	3.7%	4.2%	5.1%
Standard error	23.7%	30.7%	31.1%	31.5%
Share of HH with 0 savings	43.1%	42.6%	40.4%	35.6%

[a] LIC = Contributions to whole life insurances.

[b] POAC = Contributions to private old-age pension schemes

Source: *SAVE* 2003 and 2004.

Notes: Weighted values. N=2391 → only values if all measures were observed for each observation. The official statistics of the German Federal Reserve bank show values of an average of 10.8% for saving rates in 2003.

The results of Table 7.7 make one point very clear. Adding different contributions, debt repayments etc. to gross savings does not change much for households which do not save at all. The share of non-savers goes down from a high rate of 43% to 35% after adding five flow measures (credit repayments, contributions to life insurances, occupational, fiscally subsidized and other private old-age provisions). This means that the inequality of the savings distributions will not decrease but rather increase after correcting the one-shot savings question. The largest drop in non-saver rates is achieved when accounting for all private old-age provision types which is an additional hint that respondents might systematically ignore certain saving components which require contributions on a regular basis.

The results from Table 7.7, though, must be seen in the light that the one-shot savings question probably suffers heavily from underreporting. A similar analysis for the one-shot total non-durable expenditure question can be found in Chapter 6. But unlike response rates for the expenditure question, direct and indirect refusals for the savings question are much more frequent. So in addition to the probably underreported values itself, there is an information loss due to these refusals, as well

[146] Gross savings are defined as the values retrieved from the one-shot savings question, while net savings are gross savings plus credit repayments

as through the type-I error in the filtering process. In addition, many of the zero values might well be no true zeros, as shown in Section 7.3.

7.4.2 Other savings measures

Calculating saving rates from first differences in financial wealth shows values of 5.2%/0% (mean/median) with 2123 observations for *SAVE* 2003 and 2004 with 8.2% observations being below zero, and 52% being exactly zero, which supports the previous findings that respondents tend to repeat values (year's beginning values for year's end).

The residual measure of savings, no matter whether the one-shot or the imputed non-durable expenditure measure is used for the computation, proves to be extremely noisy. As is shown in Chapter 6, about one fourth of the imputed non-durable values are higher than net income. The mean/median saving rates are 22%/33% [147]; using the one-shot expenditure values, saving rates rise to 49%/60%. Values reach implausible values, e.g. the one-shot expenditure question takes on values up to 17 times the corresponding monthly income. The question explicitly only asked for *normal* non-durable expenditures, supported by show cards.

Computing savings by any of these two means seems little promising. Stocks of financial wealth are very often the same for the beginning and the end of the year, or implausibly remote. Non-durable expenditure values are also problematic. Even though the imputed values do a slightly better job, they entail other problems, cf. Section 6.3.3.

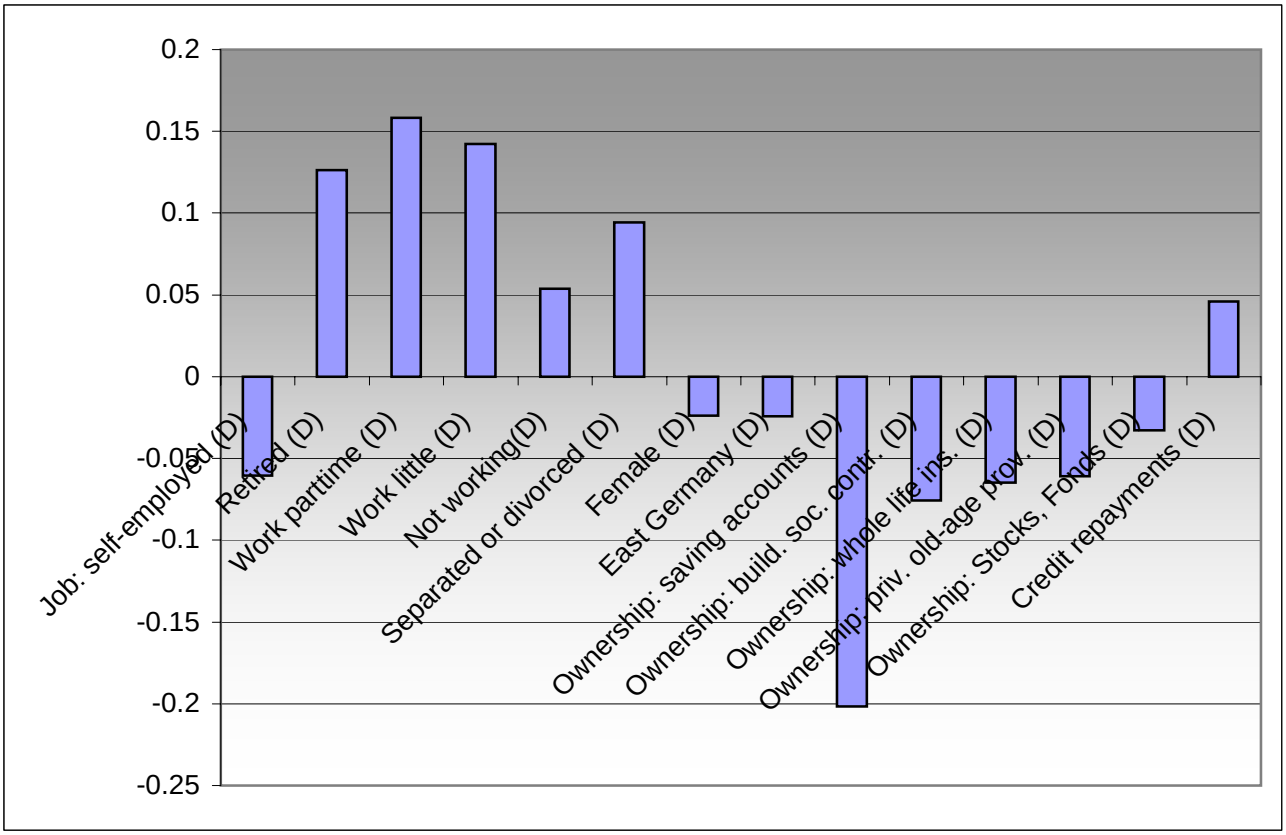
7.5 Conclusions

This chapter shows potential flaws of and limited possibilities to correct a one-shot saving question. Any measure to 'correct' the values given by the respondents is associated with rather untestable hypotheses. Hints are shown that respondents indeed do not include contributions to whole life insurances. For that reasons, the one-shot measures are corrected for credit repayments (=net savings), contributions to whole life insurance, and to old-age pension schemes. Mean values rise from 9.4% from the original one-shot question values to 11.4% when including credit repayments and imputed contributions to life insurances, and to 12.3% when additionally including contributions to all three types of private old-age provision. It is unclear, though, which of these items is already part of the values and are therefore accounted for twice. The share of non-savers, in contrast, does not change remarkably when including credit repayments and contributions to life insurances; the drop is larger when adding contributions to private old-age provisions.

[147] If income is not used in the estimation and imputation step, saving rates are 12.0%/28%; cf. Table 6.7.

7.A Figures

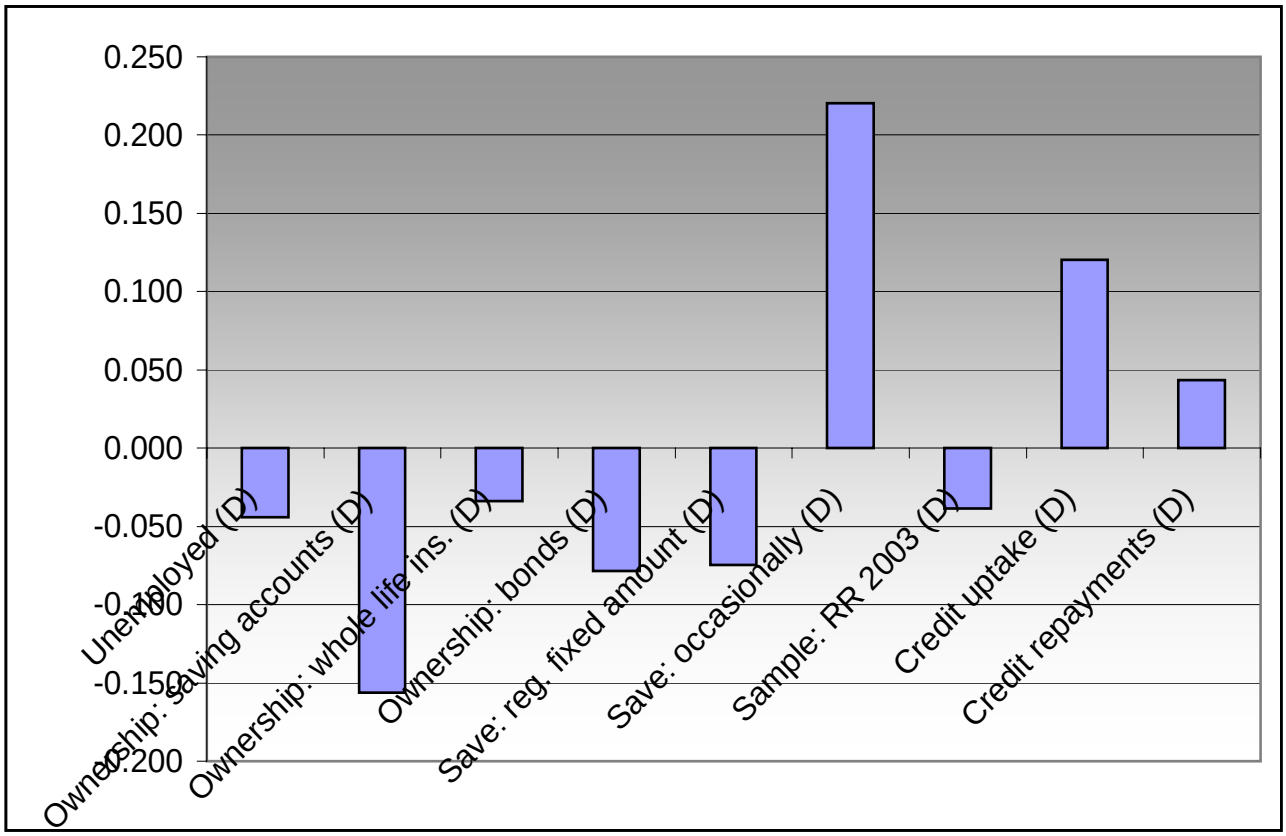
Figure 7.2: Marginal effects: influence of household- and individual characteristics on the savings ability



Source: All *SAVE* subsamples.

Notes: Dependent variable takes on the value 1 if respondent answered not being able / willing to save and 0 otherwise (one of the first three categories of filter question). Only significant coefficients displayed. Income and age not displayed: negative influence up to 23.000 €/months and up to age 57 years. Columns show the marginal effect of a change from 0 to 1. Ex.: Owners of saving accounts have 21% lower probability not being able to save. The brighter the figure, the higher the probability for savings capability.

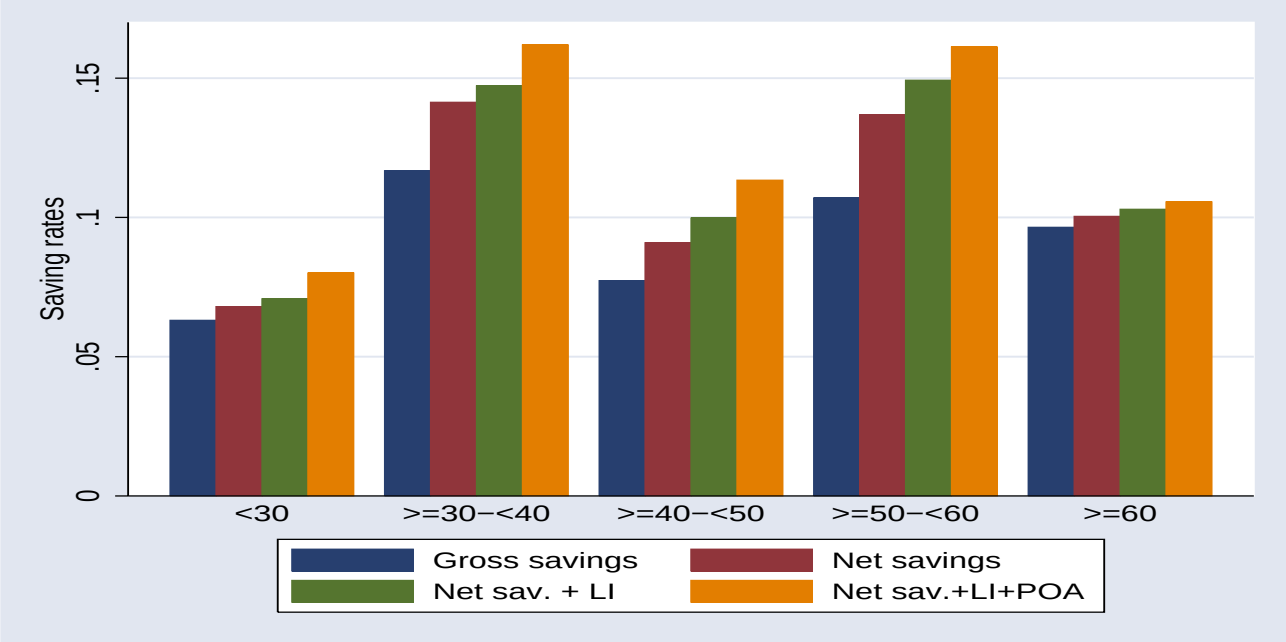
Figure 7.3: Marginal effects: influence of household- and individual characteristics on the savings ability



Source: All *SAVE* subsamples.

Notes: Conditional estimation for respondents who save according to the filter question. Dependent variable takes on the value 1 if respondent answered having nothing saved in the previous calender year or even dipped into savings. Only significant dummy variables displayed. Income is negative significant up to 19.500 €. Columns represent hight of marginal effects from a change from zero to one of the independent variables. Ex: saving accounts owners have a 16% lower chance of having zero or negative savings in the previous year. or even dipped into savings. The brighter the figure, the higher the probability for positive saving values.

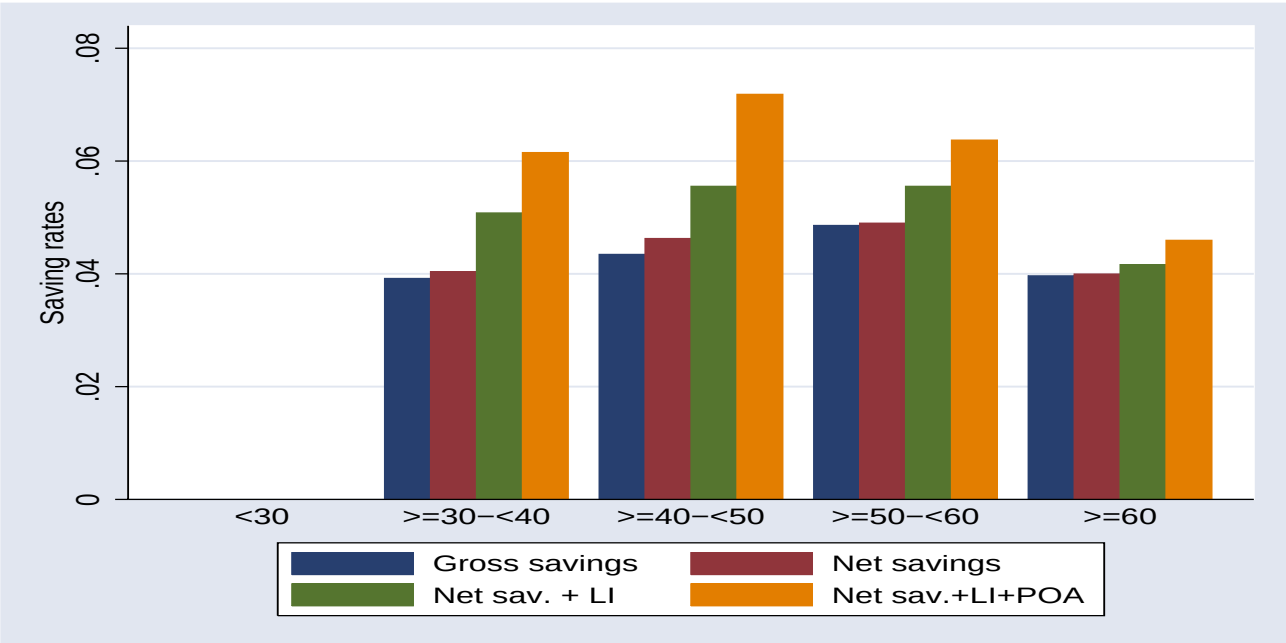
Figure 7.4: Different measures for saving rates by age classes: means



Source: *SAVE* 2003 and 2004

Notes: Weighted values. Only observations shown if no variable is missing. LI = Contrib. to whole life insurances; POA = Contrib. to priv. old-age pension schemes

Figure 7.5: Different measures for saving rates by age classes: medians



Source: *SAVE* 2003 and 2004

Notes: Weighted values. Median values for households < 30 years are zero. Only observations shown if no variable is missing. LI = Contrib. to whole life insurances; POA = Contrib. to priv. old-age pension schemes

Chapter 8

Personal assets and wealth ownership: How well are the Germans prepared?*

8.1 Introduction

The process of demographic change, and the fact that the benefits of a growing proportion of pensioners must be financed by fewer and fewer contributors, poses major problems for the German pay-as-you-go pension scheme. For this reason, the 2001 pension reform entailed a reduction in the level of statutory pensions and strengthened the funded second and third pillars of old-age pension provision. Financial incentives aimed at encouraging the accumulation of additional pension provision were introduced to enable contributors and pensioners to maintain current levels of old-age pension provision. Private, funded pension provision is less sensitive to demographic developments than pay-as-you-go state systems and, what is more, the accumulation process allows burdens to be spread more evenly across the generations than in the PAYG system. This means that the baby-boom generation will be in a position to prefinance part of their old-age pension provision. The calculations performed by the Rürup Commission on the basis of revised assumptions revealed that demographic developments will have a far more serious impact on the financial standing of the pension system than originally supposed at the time the Riester reform was introduced.[148] As a result, an increasingly important role is set to be played by supplementary private provision in the future.

The aim of partially replacing part of the first pillar of pension provision with voluntary private supplementary provision is an entirely new approach to the pension issue in Germany. At this juncture, so soon after the introduction of the Riester pension, there is little evidence available about how well accepted or how effective this measure has been. In the first year after the Riester pension scheme was introduced, around 5 million policies subject to state bonus were concluded, of which three million were private and two million occupational pension policies. Only one in five of manual and white-collar workers in the core 20 to 45 age group have so far taken out a policy and only a further 18 percent plan to do so (Schnabel (2003)). These figures have so far failed to come up to expectations; see also Chapter 7, Section 7.3.2. However, experience with similar pension products (such as Individual Retirement Accounts in the USA) in other countries suggests that similar supplementary pension products are not successful from one day to the next but require after a lengthy introductory period

* This chapter is joint work with Axel Börsch-Supan.

[148] Commission on the long-term financial viability of the German social security system, referred to in brief in the following as the Rürup Commission.

extending over several years. Nevertheless, there are a number of deficits in the current framework of private pension provision which probably impede the rapid acceptance of such policies on a large scale; discussion of reform proposals to address these problems is already underway.[149]

Two risks are inherent in the introduction of supplementary private pensions. On the one hand, the voluntary nature of supplementary pension provision represents a risk. It is by no means certain that households - and particularly low-income households - are willing and/or able to set aside additional savings for old age on a consistent basis. Empirical findings confirm that lower-income groups are less willing and able to make additional savings for their old-age pensions, and that this is exacerbated by these groups being less well informed about financial matters (Bulmahn (2003)). However, it is precisely the households in this group - a group which as a rule can expect to receive a relatively low state pension, in many cases a very modest additional occupational pension, and which also has very few significant assets at its disposal - which would be most in need of additional sources of income in old age. The high rates of cancellation of building society contracts and whole life insurance policies - despite the high losses which such cancellations imply owing to low surrender values and up-front sales commission - suggest that there is also a very real risk that private old-age pension schemes may also be terminated.[150] While they are less damaging than cancellations, periods in which people find themselves in straitened financial circumstances - owing to unemployment, loss of earnings, or if they stay at home to bring up children - and consequently suspend payments to such schemes for a period of time are also of great significance as well as being highly probable.

A further risk - the reverse side of the opportunities involved - is the rate of return on policy contributions. While the rate of return on contributions to the pay-as-you-go system corresponds to the growth rate for total wages and the population, the rate of return on payments to private pension schemes is determined by interest rates on the capital market, whereby individual capital returns depend on the success or otherwise of specific investment vehicles. The critics of funded private pension provision emphasize the risk to which financial investments are subject, particularly in view of the recent performance of the capital market. One cause of concern is the so-called "asset meltdown" hypothesis according to which demographic developments will result in a significant decrease in demand for financial assets and consequently in the capital returns on such assets. The pertinent calculations do, however, demonstrate that the portents for a demographically-induced fall in rates of return are by no means as gloomy as predicted in the popular press (Börsch-Supan *et al.* (2003)).[151] Nonetheless, the situation on capital markets since 2001 provides an unmistakable warning that lengthy periods of below-average, or even negative, capital returns are certainly a danger to be reckoned with.

---

[149] Refer, for example, to the discussion by Fehr *et al.* (2003) "Die Riester Rente - ein Flop?" in ifo-Schnelldienst 5/2003, the Bertelsmann Stiftung Pension Report (2003), the proposals submitted by the Independent Expert Commission on Tax Reform and the proposals of the Rürup commission.

[150] A survey in 2002 revealed, in Germany, that only 50 percent of policies are maintained through to the agreed term. Policies are most frequently cancelled shortly before people - in the 55-64 age group - enter retirement. The most frequently cited reasons are debts (26%), divorce (16%) and unemployment (13%) (Bertelsmann-Stiftung (2003)).

[151] Owing to demographic factors, overall capital market returns will fall by around one percentage point, assuming diversification within the EU region (Börsch-Supan *et al.* (2003)).

Both of the risks referred to above have been investigated by Essig and Reil-Held (2003), who came to the conclusion that consistent savings towards private pension provision will be essential in the future in order to maintain living standards in retirement.

This paper examines whether households are in a position to close the gap in provision created by the reduction in statutory retirement pensions without changing their current patterns of behavior, i.e. by continuing to save as they are doing at present. In this context, expectations regarding life span and retirement age play a decisive role; both these factors determine the requisite payout volume that needs to be covered by accumulated savings. The age of retirement coincides with the end of the savings phase available to households.

This paper is structured as follows: Section 8.2 examines anticipated life spans. This section compares subjective life expectancy with the latest mortality tables published by the Federal Statistical Office and checks the validity and consistency of the results. Section 8.3 examines household assets according to their composition and volume as well as the level of assets available to such households at the onset of retirement. Section 8.4 calculates the statutory pension entitlements of households before and after the two reforms of 2001 and 2004. Section 8.5 presents the key results of our analysis: the potential payout from private assets is compared with the benefits provided by the public retirement insurance system. The degree to which the new pension gap can be closed over a pensioner's entire period of retirement is then assessed. Section 8.6 highlights the results of this study in summarized form and presents recommendations for economic policy.

In the *SAVE* TPI 2004 subsample, the questionnaire design was slightly changed to enable the measurement of precise subjective life expectancy. While the structure of the *SAVE* questionnaire is discussed in length in Chapter 2, the changed design asking for subjective life expectancy is depicted in Appendix 8.A. This paper draws on two of the *SAVE* sub-samples depicted in Figure 2.1 of Appendix 2.A in Chapter 2: TPI 2004[152], which contains the date on subjective life expectancy, and Random Route (RR) 2003[153], which represents the largest *SAVE* subsample.

SAVE is a household survey. Values such as savings, assets and income are therefore assigned to households as a whole rather than to individuals. This requires yet another assumption if individual variables such as age, age of retirement or life expectancy are to be related to household size. The assumption made is that, except in the case of households headed by a single woman, the head of household is always a man. This means that calculations at the household level are based on the life expectancy and anticipated age of retirement of the male head of household.

[152] As a reminder: The abbreviation is derived from a subsidiary of NFO Infratest, the Wetzlar-based test panel institute which maintains a permanent panel ("access") from which this subsample was taken.

[153] Again, Random Route refers to a method in which households are selected at random from a sampling frame with a specific starting point and continuing along a specified route. A frequently used method for the selection of sampling households.

8.2 Life expectancy and anticipated age of retirement

A key component in the analysis of individual savings behavior is subjective life expectancy. Individual subjective life expectancy is approximated statistically in a large number of microeconometric studies. This imposes the restriction on the approximations that individuals adapt their behavior in line with observable statistical variables, such as survival probability in a life cycle model of consumption. There are two objections to this approximation. On the one hand, one might expect changes in life expectancy which cannot be found in the mortality tables as the latter, owing to their design, respond only slowly to environmental changes. On the other hand, the assumption of constant life expectancies for each cohort and sex ignores differences between individuals which certainly can be reflected in subjective life expectancy and which can lead to a corresponding change in behavior. This means that it is easier to explain individual behavior if these subjective differences can be integrated in the estimates; cf. Hurd and McGarry (1997). In order to take this into account, *SAVE* ascertains the individual life expectancy of the respondent and their partner by deploying a three-step question (cf. Appendix 8.A). The following analysis looks at the extent to which these subjective results are comparable with the mortality tables of the Federal Statistical Office and how the results of various questions asked in 2001 and 2003 compare with those for the year 2004. The factors influencing individual life expectancy are also determined.

Owing to its complexity this topic needs to be dealt with in considerably more detail. As referred to at the beginning of this paper, life expectancies play an important role in old-age pension pro-vision. This issue will be discussed in Section 8.5.

8.2.1 A comparison of *SAVE* IV descriptives and official statistics

Asking respondents about their subjective life expectancy puts such people in the unpleasant situation of having to think about their own mortality. A striking feature of the post-interview comments offered by respondents is that these questions are felt to be more personal and disquieting than questions about their wealth and assets. While this is not reflected in people's willingness to respond - which at over 96% is in fact extremely high - their comments do suggest that this subject needs to be approached with a degree of caution. For this reason a multi-phase process was used to ascertain to what age respondents and their partners believe they will live.e. The first question asked was to what age the respondents believed men and women in the same cohort lived on average[154] The next question was whether the respondents believed they themselves would live longer or shorter than the average for their cohort, followed by the request to express this difference in years. Respondents were then presented with four possible explanations for why they thought they would live longer or shorter lives than the average.[155]

Individual life expectancy is derived from the average age at death stated by respondents and the difference in years between this average age and the age to which they expected to live.

[154] Cf. Figure 8.1, and Table 8.1 which shows the average anticipated life expectancy of men and women.

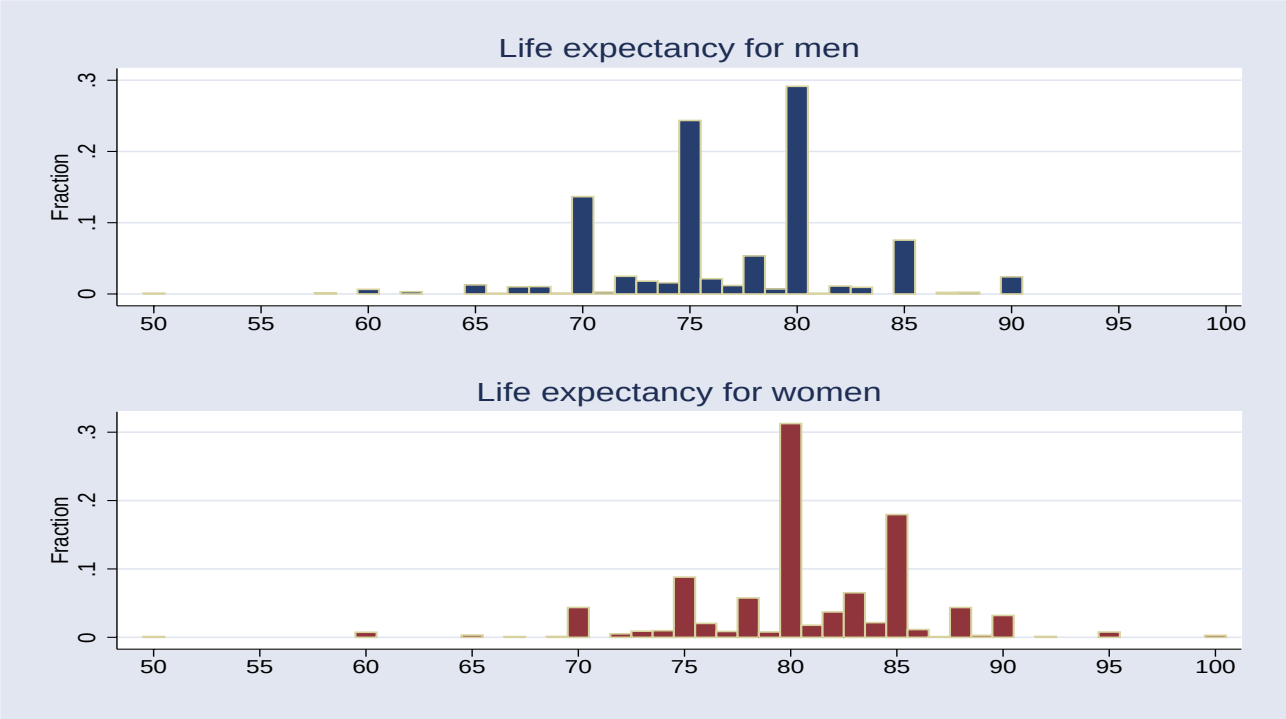
[155] One option was an open field in which other reasons could be entered.

The interview procedure was repeated in the same way with the respondent's partner.

The three steps for assessing subjective life expectancy are summarized in Appendix 8.A.

Figure 8.1 shows the differences in the anticipated average age at death of men and women. Most of the responses are concentrated around the so-called focal points[156] between age 70 and 85; in the case of women, there is a discernible shift in the age estimates to the right. The modal value is 80 in both distributions, however. Table 8.1 summarizes the results in descriptive form.

Figure 8.1: Average life expectancy by sex



Note: Weighted values.
Source: *SAVE* TPI 2004

Table 8.1: Average life expectancy by sex

	Men	Women
Mean value	76.71	80.7
Median	76	80
Standard error	0.25	0.25
Mortality table values 2000/2002	75.38	81.22

Note: Weighted according to age and income.
Source: *SAVE* TPI 2004.

The mean values for average life expectancy by sex in *SAVE* are fairly close to the values in the mortality tables drawn up by the Federal Statistical Office; with a difference of 0.5 years, the life expectancy of women approximates very closely in both sets of data. The difference in life expectancy

[156] Values which represent a multiple of a specific number, such as 10 or 5.

between men and women is underestimated, however. The difference of 3.99 years is 1.85 years less than that in the mortality tables. It is striking, however, that the differences in the absolute values between the individual and mortality table values are as small as they are given that 75.38 or 81.22 years of age are the life expectancy values for people born in the year 2002, while the *SAVE* survey asked for the anticipated life expectancy of people of the same age. In combination, Table 8.1 and Table 8.2 illustrate what households understand by individual life expectancy. The longevity figures which typically find their way into the media reflect the projected life expectancy of the recently born, and because these life expectancy values are based on period mortality tables they are entirely fictitious, with the life expectancy of newborns being thrown into the equation with the same weight as that of a centenarian.[157] Mortality in 100 years is unlikely to be same as it is today.[158] At the same time, lifespan and remaining life expectancy figures which are based on mortality tables suggest that life expectancy actually increases as people grow older.[159] This is due to the interaction of two opposing effects. On the one hand, the cohort effect - which implies lower life expectancy for older birth cohorts and thus for people who are already in the older age groups - means that life expectancy declines in old age. On the other hand, by the age of 50, for example, people have already survived a number of risks (infant mortality, the risks of various illnesses). People who have already lived longer than the average expected lifespan, for example, will also have significantly higher overall life expectancy.

There are therefore two effects which lead to an underestimation of individual life expectancy. Firstly, life expectancy figures which are based on mortality tables merely describe current population mortality. Secondly, respondents underestimate the effect of having already survived specific risks by the time they reach a certain age. Table 8.2, which shows individual life expectancy[160] according to age group, clearly demonstrates how the latter effect is neglected. Younger households anticipate living to a significantly older age than do older households.

This underestimation of individual life spans can lead to serious errors being made in the financing of retirement income. In this context, refer to Section 8.5.

Another interesting question is the extent to which the respondents believe that their income position exercises an influence on their individual life expectancy. Von Gaudecker (2004), for example, identifies socioeconomic status as the key factor explaining differences in mortality in Baden-Württemberg, and Reil-Held (2000) uses Socio-Economic Panel data to show how income and life expectancy are linked. Table 8.3 shows that the differences between four income groups are negligible, however. In other words, the respondents do not associate an improved income situation with a lower mortality risk.

---

[157] Life expectancy at birth states the average number of years a specific group (men/women) will live if mortality rates remained unchanged throughout a newborn's life. As a result, the construction of this value means that it en-compasses all age groups.
Cf. CIA factbook, http://www.cia.gov/cia/publications/factbook/docs/notesanddefs.html#2102

[158] Refer to Oeppen and Vaupel (2002) who demonstrate a roughly linear increase in average life expectancy of around 0.25 years per year over the last 160 years.

[159] Refer to Von Gaudecker (2004) for a discussion of various ratios and concepts for measuring life expectancy.

[160] This is average life expectancy ± individual difference.

Table 8.2: Individual life expectancy (own and that of partner) by sex

	Men					
	<30	30-39	40-49	50-59	60 and older	All
Mean value	81.41	77.86	75.18	73.51	78.86	76.8
Median	80	80	75	75	80	78
Standard error	2.551	0.753	0.68	0.627	0.593	0.343
N	9	110	127	112	106	464
Mortality table values 2000/2002[a]	75.97	76.66	77.33	78.64	83.81	78.37
	Women					
Mean value	87.74	80.87	80.1	80.13	81.49	80.84
Median	85	80	80	80	81	80
Standard error	2.422	0.619	0.64	0.54	0.547	0.291
N	5	98	112	110	101	426
Mortality table values 2000/2002[a]	81.69	82.03	82.42	83.22	87.06	87.06

[a] The mortality table values increase as people grow older because they have already managed to survive to a specific age. The life expectancy of a centenarian, for example, is not 75.4 (as negative probabilities, like time machines, do not exist), but 101.96.

Note: Weighted values.

Source: *SAVE* TPI 2004.

Table 8.3: Individual life expectancy according to income quartiles

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
		Men		
Mean value	76.58	75.72	77.16	77.03
Median	77	75	76	76
Standard error	0.76	0.84	0.87	0.85
Number	102	79	79	81
		Women		
Mean value	81.8	79.84	80.3	81.82
Median	85	80	80	80
Standard error	0.97	1.71	1.25	0.65
Number	28	23	37	36

Note: Weighted values.

Source: *SAVE* TPI 2004.

Another interesting question is whether there was a divergence in response behavior between the years 2001 and 2004. While the question about average general life expectancy was changed (and respondents were asked for a specific figure rather than age groups), the question about relative individual life expectancy was identical at the time of both observations. Table 8.3 shows the figures produced by households or individual respondents in both years only in response to the question on relative life expectancies for the years 2001 and 2004. There may be two reasons for changes in the information provided about respondents' relative life situation. Certain things may well have actually changed in a respondent's life which represent new information for the estimation of a person's individual life expectancy. On the other hand, responses may simply vary owing to the fact that a period of three years lies between both interviews. All in all, Table 8.4 shows that the assumptions regarding relative life expectancy are essentially stable over time. Zaller (1992), for example, examined

how opinions changed over a period of four months. The responses - selected from a choice of 5 categories (one of which, however, was a 'don't know' category) were identical in 48% of cases (cf. in this context the diagonals in Table 8.4.)

Table 8.4: Stability of results: Relative life expectancy in 2001 and 2004[a]

	Relative life expectancy 2001			
Relative LE 2004[a]	Shorter	Exactly as long	Longer	Total
Shorter	32	31	1	64
	55.20%	8.90%	1.40%	13.40%
Exactly as long	25	278	41	344
	43.10%	79.70%	58.60%	72.10%
Longer	1	40	28	69
	1.70%	11.50%	40.00%	14.50%
Total	58	349	70	477
	100.00%	100.00%	100.00%	100.00%

[a] The precise wording of the question was: "If you think of your own situation and the state of your health, do you think that, in comparison to the men (if subject male) / women (if subject female) of your age group, you will live shorter, approximately as long as, or longer than the average?"

Note: Weighted according to age and income.

Source: *SAVE* TPI 2001 and 2004; panel structure.

The descriptive results and the discussion of the implications of results generated by mortality tables show that households tend to underestimate their own life expectancy. This is likely to prove a pleasant surprise in old age for some households - the surprise may however be accompanied in some cases by a recognition that one's plans have been too short sighted. Again, cf. Section 8.5 for this issue.

8.2.2 Regression results

The questions on precise subject life expectancy were asked for the first time in the *SAVE* survey in the TPI 2004 subsample. This subsample is about a quarter of the size of the RR 2003 subsample, although the latter only surveyed general anticipated lifespans in relation to general life expectancy categories. This subsample can still be used if certain prerequisites are met, however, by imputing individual life expectancy data. This section therefore presents the regression results from the TPI 2004 subsample; the dependent variable is the subjective life expectancy, explanatory variables are a set of observable individual and household characteristics. The estimated coefficients of these regression results are therefore used in the same set of independent variables for subsample RR 2003 in order to impute individual life expectancies in the latter. Forecasting life expectancy would only be justified by the high explanatory power of the regression. The following analysis will show whether this can be reliably assumed.

The multivariate regression analysis broadens the previous bivariate tables to enable the influence of a number of different factors to be measured at the same time. Our analysis includes potential

influencing factors such as permanent income[161], age, educational and occupational variables, as well as further explanatory variables such as optimism, health expectations, and current or previous smoking behavior. The results are detailed in the Appendix, Table 8.22 (for respondents) and Table 8.23 (for respondents' partners). Respondent variables were also used for the regression of the life expectancy of the partner, as the data was given by the respondent on behalf of the partner and the characteristics of the respondent could consequently influence the data provided for the partner.

The results can be summarized as follows. The age effect is significant in both regressions, with age difference[162] also being significant for the partner regression. Interestingly enough, income and education variables are not significant. This contradicts the findings of Von Gaudecker (2004), for example, according to which the influence of socio-economic status on mortality appeared to be strengthened or even caused by the level of educational attainment. Women have a considerably longer life expectancy in both regressions, even more so in the regression for the partner (4.8 vs. 3.4 years). While smokers are obviously not yet fully aware of the potential curtailing of their lifespan by their habit, the message seems to have got home to former smokers who anticipate living 1.2 years less than average. Also significant in the partner regression are the expectations regarding the development of respondents' own and their partners' health situation. The effect of self-assessed optimism is also significant: the higher this value is (on a scale from 0 to 10), the higher life expectancy is. Four dummy variables offering possible explanations of relative life expectancy are also included, each with a high explanatory value of, on average, plus or minus four years. In summary, the variables very effectively map individual life expectancy (50% for a small cross-section). Cohort effects of life expectancy are recognized, as shown in the overall negative significant age effect (with a minimum of 45 years), but not the opposite effect of risks which have already been survived by people who have reached a ripe old age (cf. Section 8.2.1). As referred to at the beginning of this section, imputed life expectancy in another subsample or data record is only legitimate if the regression has sufficient explanatory power. For a small cross-sectional analysis, the explanatory variable R^2 is - at over 50% - very high. As explained before, the basis of the data used will be broadened to calculate the asset position, pension claims (and gaps in provision) at retirement age by additionally drawing on the *SAVE* Random Route Sample 2003 subsample.

8.2.3 Anticipated age of retirement and replacement rates

This section looks at the second variable determining the length of a person's retirement - the anticipated age of retirement. The age of retirement has been a protracted topic of discussion in Germany over the last twelve months as more and more observers have pointed out that the average age of retirement is too low. Table 8.5 shows the relevant figures from the Socio-Economic Panel and VDR, both of which demonstrate an age of retirement of 60 - significantly earlier than the state pension age

[161] The construct 'permanent income' is intended to eliminate potential transitory fluctuations in gross or net monthly income and to obtain a better planning variable for households. The method of estimating permanent income is based on the proposals of King and Dicks-Mireaux (1982) and is described in more precise terms for the *SAVE* data in Chapter 5, Section 5.2.2.

[162] This may reflect the hope that an older partner does not die before oneself.

of 65. The anticipated ages of retirement in *SAVE* are also shown, these being significantly higher than the current figures.

Table 8.5: Actual and anticipated age of retirement

	SOEP 1999[a]		VDR 2001[b]	SAVE[c]					
				TPI 2001[d]		TPI 2004[d]		RR 2003	
	Men	Women	Men and women	Men	Women	Men	Women	Men	Women
Mean value	59.7	60.7	60.2	63.1	61.7	64.4	62.9	63.3	63
Median	N/A	N/A	N/A	65	60	65	64	65	65
Std. dev.	N/A	N/A	N/A	3.19	3.01	2.55	3.14	4.09	3.79
Std. error	N/A	N/A	N/A	0.21	0.31	0.18	0.34	0.18	0.14
Number				223	92	203	84	547	702

[a] Observed values Source: Berkel and Börsch-Supan (2003)

[b] Observed values. Source: Own calculations / with grateful acknowledgement to Christina Benita Wilke for the data provided.

[c] Estimated values

[d] Panel comparison: Only households / respondents interviewed in both waves

Notes: Weighted values in *SAVE*; figures exclude freelance professionals and the self-employed.

The first *SAVE* survey was carried out before the introduction of the 'Riester' pension scheme. Discussion of the pension system came to the attention of large sections of the population in 2003 in the wake of the work of the 'Rürup' Commission. The quintessence of the discussion is also reflected in the data: the average age of retirement for men in 2004 is estimated to be 1.5 years later, while the average expected pension replacement rate fell by 5 percentage points. Table 8.5 shows the expected age of retirement and replacement rates according to age group. While the age of retirement differed only insignificantly between the group of under 30-year-olds and the group of 50-59-year-olds in the year 2001 among re-interviewed male respondents, the difference of 3.6 years in the year 2004 reached the significant 5% mark. However, the anticipated pension replacement rate for this group, which also increased, is unrealistic.

Table 8.21 shows the results of the multivariate regression analysis for the age of retirement and relative pension levels. Education variables are positively significant, which could mean one of two things. People with higher educational qualifications may really retire later, or they may simply be more aware of the normal age of retirement of the future. The dummy variables for the Random Route 2003 subsamples and TPI 2004 are also significant. On average, respondents for this subsample anticipated retiring between 0.6 and 1.2 years later. Particularly for the TPI 2004 variable, this is clearly a result specific to this particular year, during which pension reform was a focal point of public discussion. The significant age effect[163] is positive with a maximum of 71 years. This means that younger households expect to receive relatively lower value pensions than older households - again probably due to the current pension reform discussion.

[163] Age and quadratic age are jointly significant.

Table 8.6: Anticipated age of retirement and pension replacement age according to age

	TPI 2001[a]				TPI 2004[a]				2003 RR			
	Men		Women		Men		Women		Men		Women	
Age	Ret. age	Rate[b]	Ret.a.t	Rate	Ret.a.	Rate	Ret.a.	Rate	Ret.a.	Rate	Ret.a.	Rate
Under 30	63.8	58.2%	63.4	54.0%	66	65.0%	62		63.6	50.9%	63.7	48.0%
30-39	63.7	59.9%	61.2	58.7%	65.2	53.3%	63.2	53.1%	64	57.6%	62.8	52.2%
40-49	62.6	64.9%	61.4	62.3%	64	56.5%	62.8	43.9%	62.9	59.2%	62.8	59.5%
50-59	62.1	63.6%	60.9	57.1%	63.4	64.0%	62.3	53.9%	62.8	62.2%	62.8	61.2%
60 and older	63	67.0%	62.5	59.0%	64	52.6%	63.9		63.8	60.6%	63.2	60.7%

[a] Panel comparison: Only households / respondents interviewed in both waves

[b] Pension replacement rate in relation to last income received prior to retirement

Notes: Weighted according to age and income in *SAVE*; figures exclude freelance professionals and the self-employed.

8.3 Household assets

This section describes the asset situation of households based on the data provided by both *SAVE* subsamples TPI 2004 and Random Route 2003. We begin by presenting the current asset situation and go on to describe how the asset situation can be calculated at the age of retirement.

8.3.1 Descriptives for *SAVE* IV and *SAVE* RR 2003

Table 8.7 shows the value of households' financial, real estate (owner-occupied and other real estate) and total assets. While the mean values are very close to each other, the underlying distribution of assets differs in both subsamples. The median for total assets in TPI 2004, for example, is thirteen times the value of RR 2003. This is partly to do with the fact that 15% of households in TPI 2004 and 32% of households in RR 2003 have 0 or negative assets. A median of 0, for example, for real estate means that at least half of all households do not own any residential property at all.

Table 8.7: Financial, real estate, and total assets

	Financial assets		Real estate assets		Total assets	
	TPI 2004	RR 2003	TPI 2004	RR 2003	TPI 2004	RR 2003
Mean value	23805.1	21062.73	151863.3	130270	166507.2	140014.5
Median	4000	2300	100000	0	119000	9000
Std. error	2791.49	2791.49	8955.31	5930.8	12768.12	14105.88
Number	306	1266	469	1901	234	1109

Note: Weighted values.

Tables 8.7 and Table 8.8 show which asset categories go to make up total assets and demonstrate that owner-occupied property makes up by far the largest element of people's assets. This is certainly noteworthy in the light of people's asset position upon entering retirement. Residential property is seldom divisible. In other words, it is not possible to sell off a house or flat bit by bit.[164] Nor is post-retirement communal living or flat-sharing likely to be to everyone's taste. It is of course possible

[164] Excluding the renting out of accessory apartments, for example.

to sell one's property and to reshuffle one's assets by converting real estate into financial wealth. However, this also entails a substantial increase in the household's consumption expenditure in the form of rental payments for alternative rented property. Bearing in mind the value attached to home ownership in Germany, this latter option obviously cannot be in the interests of owner-occupiers.

Table 8.8: Relative value of assets:[a] Financial, real estate, and total assets. SAVE TPI 2004

	Owner-occupied property	Other real estate	Business assets	Financial assets	Loans	Other assets
			Absolute values			
Mean	60.4%	4.9%	0.2%	33.3%	40.7%	1.3%
Median	79.0%	0.0%	0.0%	11.4%	0.0%	0.0%
Std. error	2.8%	1.0%	0.1%	2.8%	17.8%	17.8%
Number	211	211	211	211	211	211
			Conditional values for possession of asset category			
Mean	81.7%	35.0%	8.5%	39.1%	99.9%	10.9%
Median	93.5%	30.5%	4.6%	17.6%	20.0%	4.7%
Std. error	1.9%	3.6%	3.1%	3.0%	38.2%	3.8%
Number	150	39	6	187	109	32

[a] Value of specified assets divided by the value of gross total assets (total assets - loans).
Note: Weighted according to age and income in *SAVE*

Table 8.9: Relative value of assets:[a] Financial, real estate, and total assets. SAVE RR 2003

	Owner-occupied property	Other real estate	Business assets	Financial assets	Loans	Other assets
			Absolute values			
Mean value	43.2%	3.6%	1.6%	50.1%	38.7%	1.5%
Median	28.6%	0.0%	0.0%	29.4%	0.0%	0.0%
Std. error	1.6%	0.5%	0.3%	1.6%	12.5%	0.3%
Number	787	787	787	787	787	787
			Conditional values for possession of asset category			
Mean value	83.7%	35.5%	32.4%	56.9%	165.7%	16.8%
Median	91.3%	30.0%	25.7%	77.1%	25.6%	7.5%
Std. error	1.0%	2.6%	4.6%	1.7%	51.4%	2.9%
Number	404	79	40	694	192	73

[a] Value of specified assets divided by the value of gross total assets (total assets - loans).
Note: Weighted according to age and income in *SAVE*

8.3.2 Assets on retirement

Assets on retirement will differ from current assets owing to the influence of two factors. On the one hand, assets will bear interest at a nominal rate i. On the other, account must be taken of general inflation. The purchasing power of 1 euro will not be the same in ten years' time as it is today. In this respect, the required rate of return r equals the difference between the nominal rate of interest i and the rate of price increase p. We assume that all assets are subject to a constant effective interest

rate and base our calculations on the required rate of return of 2.8% per annum used by the Rürup Commission.

The interest period is the difference between the expected age of retirement and the present age of the head of household. We assume that no asset consumption takes place prior to retirement. Compound interest effects therefore generate a substantial increase in assets during the period prior to retirement. At an effective rate of interest of 2.8%, for example, the value of assets doubles over a period of 25 years.

Current flows of savings in each period up to retirement are added to existing assets as described in the following. *SAVE* surveys total household savings for the previous year. These annual savings are then recalculated and spread evenly across the year as monthly savings figures. We also assume that these monthly savings remain at a constant level right up to retirement and that they also yield 2.8% p.a. interest.

The entire asset position of the household on entering retirement is therefore determined by the assets with accrued interest added plus the monthly savings and interest on such savings.

Appendix 8.C summarizes the relevant financial equations.

8.4 Claims on the public retirement insurance system

This section describes households' claims on pension benefits from the public retirement insurance system before and after the reform proposals of the Rürup Commission adopted for 2005 and which involve an incremental reduction, as a response to demographic trends, in relative pension levels.

8.4.1 Calculation method

We calculate claims on the public retirement insurance system on the basis of individual data relative to the entitlement position of a benchmark pensioner ("Eckrentner").[165]

Two values are needed in order to calculate individual entitlements from the public retirement insurance system: the number of insurance years (calculated on the basis of the anticipated age of retirement and estimated age at which people's working lives begin[166]) and average earned income over an entire working life.

[165] A benchmark pensioner has worked and paid pension contributions for 45 years - precisely in accordance with the average for all contributors - and has consequently had one earnings point per year credited to his or her pension account. Average income in 2002 was 28,949, or net monthly average earnings of 1,531.92. 45 earnings points currently entitle a benchmark pensioner to a gross monthly pension of 1,175.85 euros in the western Germany and 1033.65 in the new eastern states or, after deducting health and long term care insurance contributions of 89.77, a net monthly pension of 1,086.08. Refer to the publications of the German Federal Social Insurance Office for Salaried Employees (BfA).

[166] People are assumed to begin their working lives at the following ages: age 16 for those with a lower or intermediate secondary school leaving certificate, aged 20 for those with the 'Abitur' upper secondary school leaving certificate, and age 25 for those with university or polytechnic degrees. Actual average numbers cannot be easily received, but looking at the values found, the *current* job entry age, or the school finishing age, are higher than the values assumed here. But the German pension system guarantees additional 'Entgeltpunkte' for school/university attendance.
If the head of household has been unemployed for a period longer than one month, an additional year of work is deducted.

The pension entitlement RA per household HH for the year 2004 is therefore as follows:

$$RA_{HH} = \frac{\mathrm{E}[REA_{HH}] - EEA_{HH}}{45\,years} \cdot \frac{\varnothing EK_{HH}}{\varnothing EK_{POP}} \cdot RA_{ER} \cdot AF \tag{8.1}$$

where

$\mathrm{E}[REA_{HH}]$ = anticipated age of retirement of the household HH
EEA_{HH} = Age at which household begins its working life HH
$\varnothing EK_{HH,POP}$ = Average earned income of household HH or of the population
RA_{ER} = Benchmark pensioner's net pension entitlement (1,062.40 € in 2002)
AF = Adjustment factor used to calculate deductions or additional credits in the case of earlier or later retirement

We draw on the standard age of retirement of 65, disregarding total and partial disability rules. The earliest possible age of retirement is currently age 63 and will be reduced by the 1999 Pension Reform Act in two-monthly steps to the age of 62[167] by 2010 to 2011. Monthly deductions/credits amount to 0.3/0.5% per month or 3.6/6% per year. While there is no statutorily defined upper age limit for retirement, the actuarially unfair credits do not, however, offer any incentive for people to postpone their retirement decisions.[168]

Equation 8.1 describes the deductions or credits received by a household compared with a benchmark pensioner depending on relative income, relative number of insurance years and anticipated age of retirement.

We calculate the pension gap as follows.

Gross pension levels prior to the Riester reform and the introduction of the sustainability factor, as well as the values after both reforms had taken full effect, were calculated using the MEA-PENSIM model (for a description of the simulation model, refer to Wilke (2004)). Post-reform pension value is estimated at 86.6% of the pre-reform gross pension value where the 2004 Pension Reform Act provisions have taken full effect, which will be the case from the year 2030 onwards. Calculating net pension levels is somewhat more complicated. The new pension taxation rules make it very difficult to make any general assumptions in this respect. For the sake of simplicity we therefore assume that the percentage decrease in gross pension levels is reflected analogously in net pension levels. Based on these assumptions the reduction factor, RF, will therefore be specified in the following way:

$$RF = f(\text{cohort, pension entry age, life expectancy}) \tag{8.2}$$

See Table 8.10 for clarifying examples of (8.2).

The gap in provision, DL, is thus calculated as follows:

$$PG_{HH} = RA_{HH} \cdot RF = RA_{HH} \cdot (1 - RF) \tag{8.3}$$

[167] Which implies a pension benefit deduction of 10.8%

[168] Börsch-Supan *et al.* (2004c) calculated the incentive-neutral deductions and credits at around 7% p.a. Refer also to Börsch-Supan (2004a) and Börsch-Supan (2004b) for a definition of incentive-neutral deductions.

Table 8.10: Pension levels after the year 2003 relative to pension levels before the pension reforms 2001 and 2004

	Age				
Life expectancy	40	45	50	55	60
70	85.7%	87.7%	89.7%	91.6%	93.2%
75	85.0%	86.9%	88.7%	90.6%	92.4%
80	84.5%	86.3%	87.8%	89.7%	91.5%
85	84.3%	85.7%	87.1%	88.8%	90.6%

Note: Assumed retirement age of 65 years. Early retirement leads to slightly *higher* values since the relative pension level declines with time, and with earlier retirement one would get the higher values. Ex.: At retirement with 61 years of a today 40-year-old with life expectancy of 70 years, shown values would be higher by 0.5%, for a 60-year-old by about 1%. These values are only relative numbers and are contrasted by the pension reduction factor for early retirement.
Source: MEA-PENSIM.

where PG_{HH} indicates the pension gap of household HH due to the pension reforms. (8.3) suggests that the pension gap is equally great in percentage terms for all households, as was intended by the Commission on the long-term financial viability of the German social security system ("Rürup Commission").

8.4.2 Specification of individual variables

The anticipated age of retirement of a household HH, $\mathrm{E}[REA_{HH}]$, is derived from the *SAVE* data record.[169] In this context we draw on the earliest possible retirement age under future legislation - age 62, with a pension deduction of 10.8%.[170] As there is no statutorily stipulated upper age limit, higher figures are not corrected. The deductions and credits referred to in Section 8.4.1 are taken into account to determine the deduction factor AF.

The age at which people started their working life EEA_{HH} is not surveyed in the *SAVE* study. We therefore assume the following three labor market entry ages, in relation to educational attainment: age 16 for those with a lower or intermediate secondary school leaving certificate, age 20 for those with the 'Abitur' upper secondary school leaving certificate and 25 for those with a university or polytechnic degree.

The length of a person's working life, and thus the number of years during which contributions are paid, the difference between $\mathrm{E}[REA_{HH}]$ and EEA_{HH}, is reduced by one year if the employment history of the head of household includes a period of unemployment of longer than six months.[171]

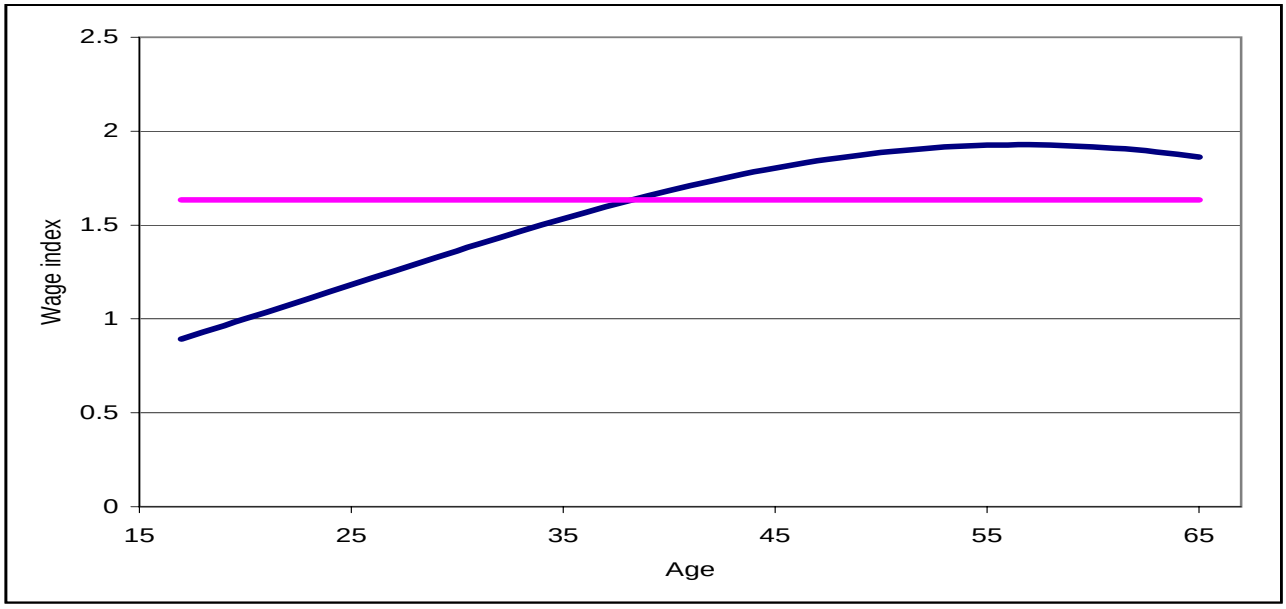
[169] "What are your expectations? - At what age do you expect to retire or receive an old-age pension?"

[170] In fact, observations for values under 62 are allowed, but the deduction rate remains constant at 10.8%. This allows for early disability retirement.

[171] This ultimately simplifies the actual statutory rules. In fact, unemployed people in receipt of benefits from the labor office are also covered by the pension insurance scheme. Contributions to the scheme are paid by the labor office. The relevant contributions for those in receipt of unemployment benefit are levied on the basis of 80% of their last earned income. The contributions for unemployment assistance claimants are only levied on the basis of the assistance amount. People drawing old-age pensions or receiving total disability benefits are no longer entitled to benefits from the labor office, even if their entitlement period has not yet expired. The situation is quite different for those in

The household's average earned income, $\emptyset EK_{HH}$, is determined by estimating the age profile of the income received according to the estimate of permanent household income. The household's earned income on retirement is thus determined and compared with the average value over the working life. The estimates reveal that earned income is 5% higher than average income at the age of retirement (cf. Figure 8.2).

Figure 8.2: Estimated earnings curve over the entire working life



Notes: Normalized values: Index is 1 at age 20. The average figure is 12.3% below that at age 65 (normal age of retirement).
Source: All *SAVE* data without retired, self-employed freelancers.

The retirement ages elicited in the survey were in some cases considerably lower than the future earliest state pension age of 62. Deductions are frozen at 10.8% for these cases. If no pension entry age was given by respondents, an entry age of 61 was assumed.[172] Households are typically subject to larger uncertainties and exogenous fluctuations at the beginning of their working lives than they are later on.[173] All the results in Sections 8.4 and 8.5 therefore relate to households in the age 40 or older age group and are not yet retired.

8.4.3 Results

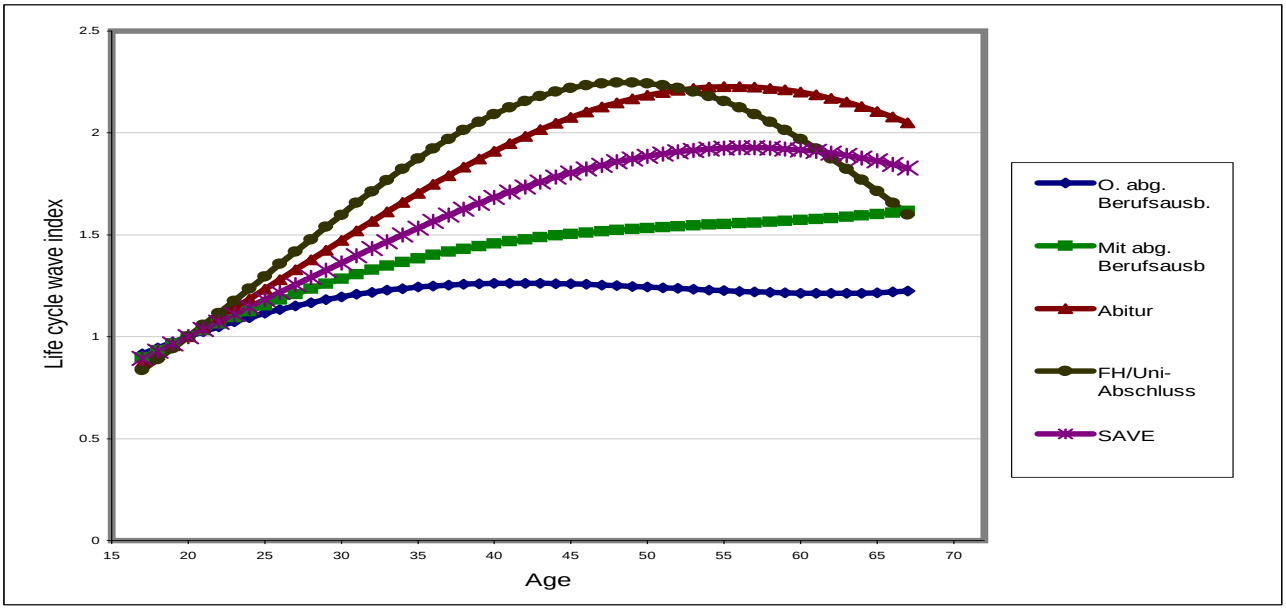
The actual meaning of the figures and how they should be interpreted are now briefly discussed before the results themselves are presented. Projected future pension benefits can basically be regarded as nominal, real or economic status-preserving euro amounts. These three different presentation options

receipt of partial disability benefits. In this case the labor office continues to pay unemployment benefit in addition to disability benefits, but only up to the additional earnings limit.

[172] If the respondent is already older than 61 and not yet retired: pensions entry age ≡ age + 1

[173] Bearing in mind factors such as choice of profession, choice of partner or the decision to start a family or not.

Figure 8.3: Comparison with results from Fitzenberger *et al.* (2001)



Notes: Normalized values: Index is 1 at age 20. Wage projections except *SAVE* refer to results from Fitzenberger *et al.* (2001).

differ by the way in which future inflation or, additionally, wage trends are accounted for.[174] It is also possible to present possible (future) pension benefits in relation to (future) income. While this enables explicit adjustments for inflation to be avoided, the disadvantage is that it is necessary to know what percentage rates mean, which is unfortunately not necessarily the case for many households.

The status-preserving variables for pension data are consequently values which specify how much the household would receive if it were to enter retirement today, taking account of all other outstanding future contributions. This does away with the need for any further mental acrobatics for general or wage inflation adjustments and the variables can therefore be compared directly with today's income.

8.5 Private pension provision: how the pension gap can be closed

This section compares the post-retirement income sourced from annuitized assets (cf. Section 8.3) with claims on the public retirement insurance system, and considers in particular whether private asset positions can compensate for the pension gap which arises as a consequence from the pension reforms 2001 and 2004.

[174] Cf. Börsch-Supan *et al.* (2004b), who discuss the presentation options in considerable detail and emphasize the importance of how each of them are used. The way these are presented can have a huge influence on the way a person's individual pension future is perceived as is the case, for example, with pension notifications sent by the BfA. The latter present a picture of future nominal pension entitlements and include a statement that no adjustments have been made for inflation. The values as stated are therefore higher than many people expect, which can in turn result in misguided economic behavior.

Table 8.11: Pension levels prior to and after the impact of the 2004 pension reform

	Values before Riester reform		after Rürup		Pension gap
	pension level	relative to pred.last net inc.	pension level	relative to pred.last net inc.	in today's Euros
Meant	1393.9	58.6%	1247.6	52.2%	156
Median	1352.8	57.7%	1211.5	51.1%	144.4
Std. Error	17.1	0.3%	15.2	0.3%	2.5
Obs.	897	897	875	875	875

Note: Weighted values.
Source: *SAVE* RR 2003 and TPI 2004. Self-employed and freelancers excluded.

Section 8.5.1 calculates a household's hypothetical annuities[175] based on assets saved prior to entering retirement. The focus of the analysis lies in using financial wealth and savings, since housing wealth, which typically represents the largest portfolio share, cannot be liquidated in a comparable way to financial assets. Additionally, selling a house is associated with a status change from renter to tenant which in turn would necessitate rent payments. Cf. Chapter 7.3.3 for a discussion, where also an alternative procedure is mentioned: reverse annuity mortgages. These payments also depend on the age of the household (minimum age: typically 62 years), the value of the home, and interest rates. They become more and more popular in the U.S., even for richer households, but play a very minor role in Germany.

Section 8.4.3 argues that status-preserving variables offer an appropriate means of presenting future income and assets positions. Similarly we make adjustments to the development of assets in order to ensure that personal assets are not overestimated in comparison with pension developments. Future assets are therefore adjusted for wage growth g by dividing them by $(1+g)^{retirement\ age_{HH}}$ or using the approximative required rate of return $cr = r - g$.

Section 8.5.2 compares private savings accumulated by the age of retirement with the pension gap and examines to what extent households are in a position to close this gap. The starting point in the analysis is the subjective life expectancy. But since it can be presumed that households systematically underestimate their life expectancy, results from alternative values are also shown, including the official life tables.

After the sensitivity analysis for life expectancy values, Section 8.5.3 shows the results from Section 8.5.2 in dependence of the parameters r, the real interest rate, and g, the real wage growth rate.

Section 8.5.4 shows the ability to fill the pension gap in dependence of three household characteristics: age, income, and schooling.

[175] The term annuity as used here refers to individual life expectancy and is not based on the life expectancy calculated by insurance companies using individualized mortality tables. In this respect the term differs from its usual actuarial definition in that there is no longevity risk if the household reaches precisely its anticipated life expectancy

8.5.1 Monthly life annuities from provisions for old age and personal assets

In addition to pension benefits provided by the public retirement insurance system, all house-holds could hypothetically receive annuities from their saved personal assets. In this case we make the following assumptions. The household does not include any inheritances in its financial planning; these only arise if members of the household die earlier than expected and therefore do not consume their entire savings. The relevant points in time for calculating annuities are the anticipated date of retirement (in order to project savings and assets based on the effective rate of interest r = 2.8% p.a.) and life expectancy (the point in time up until which annuity payments will be made). The payout period is thus the difference between these two points in time.

In *SAVE*, Gross financial wealth[176] is constructed by eight different wealth categories. For consistency reasons, only full information observations are used. This means that for a household indicating the ownership of a wealth category but refuses the answer for the corresponding wealth value for any of these categories, financial cannot be constructed. This procedure entails a large data loss. To mitigate this data loss, financial wealth is imputed in the following way. In a semi-logarithmic specification, relative financial wealth (financial wealth / net income) was regressed on a polynomial on net income and age; sociodemographics; and dummy variables indicating the ownership for any of the six[177] wealth categories. Solving the predicted values for financial wealth (exponent of ln(financial wealth/ net income) times net income) raises the number of observations by 50%. The relevant sample was the later used (age 40 to 65, no retired HH, no self-employed and freelancers). The regression results are shown in Table 8.24. Imputed values are only used for households for which computing wealth was not possible. Households indicating no possession for every financial wealth group have a total financial wealth, accordingly.

Table 8.12: Monthly life annuities from predicted savings, future financial and housing wealth, based on subjective life expectancy

	Savings	Financial wealth[a]	Housing wealth	Conditional housing wealth
Mean	348.0 €	299.6 €	1,548.7 €	2,960.4 €
Median	108.4 €	99.6 €	384.9 €	1,834.5 €
Std. error	26.7 €	39.8 €	176.6 €	316.0 €
Obs.	662	673	768	417

[a] Net financial wealth (total financial wealth reduced by consumption, family and other short-run credits)

Note: Weighted values.

Source: *SAVE* RR 2003 and TPI 2004. Only households with head being between 40-65. Self-employed and freelancers excluded.

The means of monthly life annuities are relatively high. The future values of today's financial net wealth reaches a monthly life annuity of 300 €. Assuming that today's savings will be continued in the future and annuitized at the pension entry age, an additional second monthly life annuity of 348

[176] financial wealth without credits

[177] The three forms of private old-age provisions of any form were aggregated.

€ will be received. Finally, if it would be liquidated, housing wealth would provide a third annuity of 1548 €.

Means of wealth values are misleading, since the wealth distributions typically are skewed to the right. This can be read by the median values which are significantly lower than the means; the median values are about 1/3 for savings and financial wealth. Additionally, some households do not hold financial assets or do not even save, which leads to an annuity of 0. E.g., the conditional[178] annuity for housing wealth provides an annuity of 2960 € which is about twice as high as the unconditional value.

Additionally, the annuities presented in Table 8.12 are calculated using subjective life expectancies. These are, as shown in Section 8.2, significantly lower projections from current life tables. Table 8.13 shows the same values as in Table 8.12, but assumes a more realistic life expectancy. Hence, monthly private pensions from accrued savings, financial and housing wealth drop by about 20%.

Table 8.13: Monthly life annuities from predicted savings, future financial and housing wealth, based on subjective life expectancy plus 3 years

	Savings	Financial wealth[a]	Housing wealth	Conditional housing wealth
Mean	311.0 €	248.8 €	1,222.6 €	2,324.2 €
Median	94.8 €	84.6 €	378.8 €	1,580.4 €
Std. error	31.3 €	33.6 €	85.4 €	139.1 €
Obs.	672	683	781	427

[a] Net financial wealth (total financial wealth reduced by consumption, family and other short-run credits)

Note: Weighted values.

Source: *SAVE* RR 2003 and TPI 2004. Only households with head being between 40-65. Self-employed and freelancers excluded.

8.5.2 Values from monthly private pensions in relation to the pension gap

As the next step, it will be shown how the private pension wealth from Tables 8.12 and 8.13 compare to the pension gap shown in Table 8.11.

Table 8.14 summarizes the results as percentaged filling of the pension gap. It also compares three different measures of financial wealth plus savings.

The first one is net financial wealth (computed as the sum of eight different financial wealth categories minus short-run consumption, family and other credits) and net savings (computed as gross savings[179] plus credit repayments plus contributions to life insurances), 'Riester' plans and occupational pension plans, assuming that households do not think of these categories when trying to recall total last year's savings. Evidence for this hypothesis is provided in Chapter 7. The second one is gross financial wealth (all financial wealth categories ignoring short-run credits) and net savings. The third is net financial wealth and gross savings (without credit repayments and imputed contributions). The numbers are quite similar and do not change the qualitative statements. This becomes clear when

[178] Conditional on home ownership

[179] as the values given to the one-shot savings question

looking at the short-run credit volume which is rather low with a mean of 1850€ and a median of 0 (weighted values). The difference between the gross and net medians of monthly savings measure is about 15 €. In the following, the first set will be used for the rest of the analysis (net savings, net financial wealth).

Numbers higher than 100% means that the private pensions which a household can retrieve from private wealth, suffice to fill the pension gap, numbers below 100% indicate that the pension gap cannot be covered. It shall be emphasized again at this point that the *current* savings behavior is projected, and therefore it is not accounted for possible behavioral changes possibly induced by the pension reforms 2001 and 2004. The research interest is: do household save enough today to fill the future gap?

Table 8.14: Households' ability to fill the pension gap in dependence of the life expectancy

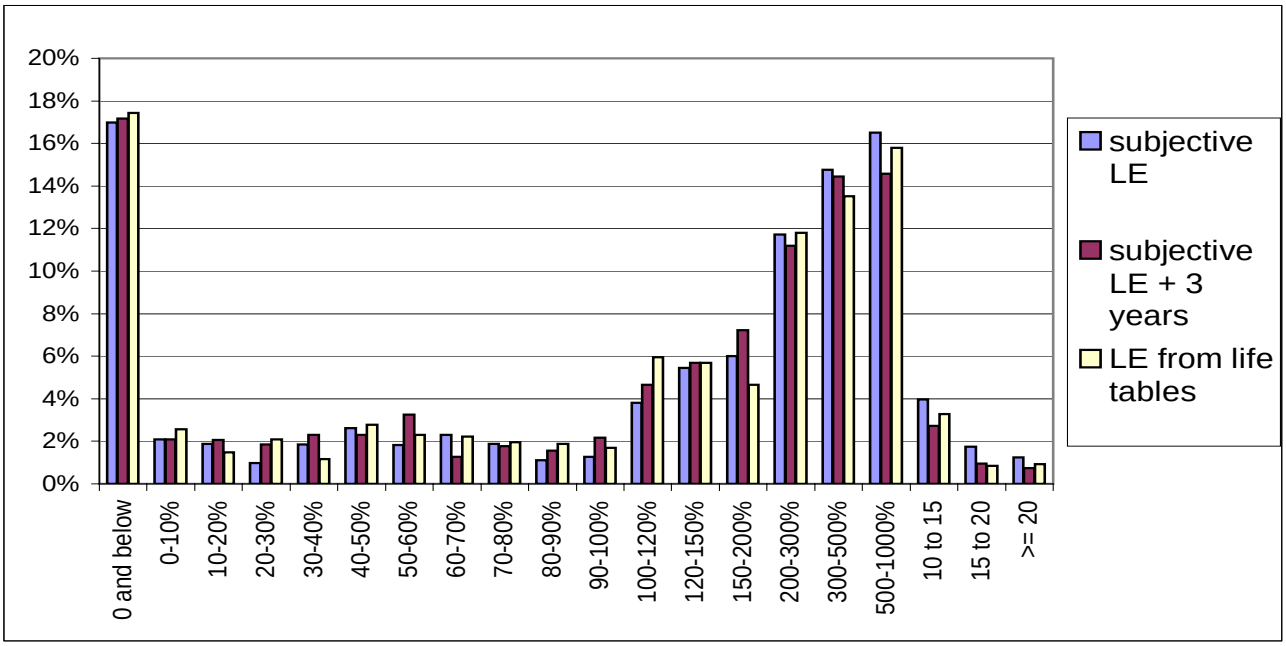
	own LE	own LE + 3	own LE + 5	own LE + 10	LE from life tables
	Net financial wealth including consumption credits and accrued savings, including contributions to Riester plans, life insurances and credit repayments				
Mean	362.3%	290.1%	262.0%	216.9%	305.0%
Median	198.8%	157.2%	141.6%	117.6%	161.0%
Std. error	25.6%	19.5%	17.4%	14.1%	19.1%
Obs.	575	579	579	579	599
Gap in provision	32.4%	35.8%	37.3%	43.7%	35.1%
	Gross financial wealth without credits and accrued savings, including contributions to Riester plans, life insurances and credit repayments				
Mean	379.1%	303.4%	273.6%	226.1%	316.8%
Median	204.2%	168.9%	147.0%	121.0%	168.3%
Std. error	25.6%	19.5%	17.3%	14.0%	18.9%
Obs.	593	597	597	597	619
Gap in provision	30.5%	34.0%	35.9%	42.0%	33.4%
	Net financial wealth including consumption credits and accrued gross savings				
Mean	328.2%	262.0%	237.1%	196.7%	275.7%
Median	174.5%	146.4%	131.9%	107.5%	144.1%
Std. error	23.3%	17.8%	15.9%	12.9%	17.0%
Obs.	609	614	614	614	638
Gap in provision	35.5%	38.8%	40.6%	45.6%	37.8%

Note: Weighted values. Share of households with zero or negative savings and financial wealth: 16%.

The same arguments pointed out for Tables 8.12 and 8.13 also apply for Table 8.14: mean values are quite high and definitely high enough to fill the gap. Median values, in contrast, come close to one when assuming a 10 year longer life expectancy, and about one third of the households in the sample will not be able to fill the gap. This percentage grows to 44% assuming a 10 year higher life expectancy. The values' dependence on life expectancy is has a simple reason. Individual longevity risk is not covered by personal assets where this risk is not borne by other insured persons as is the case, for example, in the public retirement insurance system. The last column of the table shows the filling of the gap assuming values of current life tables; results compare about to underestimating the subjective life expectancy by three years. But the current life tables ignore all medical progress in

the future and of the past[180]. The Rürup- commission, for example, anticipates an increase in life expectancy of 2.5 years over the next 30 years. This forecast is about 2/3 lower than the estimates of Oeppen and Vaupel (2002). If the more optimistic of these scenarios - a long-term increase in life expectancy of 0.25 years per year - proves to be correct, the public pension insurance system will be confronted by further financing problems which will articulate themselves once again in the form of rising contribution rates and lower pension benefits. The non-coverage of the gap thus affects about 36 to 40% of all households, while the median value is able to fill the gap by more than 100%.

Figure 8.4: Distribution of the ability to fill the pension gap



Note: Weighted values.

The skewness of the distribution of savings and wealth is thus also reflected in Figure 8.4 which shows the distribution of the ratios of monthly life annuities to the pension gap. The distribution also shows that obviously, there are two major groups of households: the ones not being able at all to fill the gap at 0% and below, and the other ones who can easily absorb the financial task with 200-1000% of the gap.

8.5.3 Robustness of the results

This section checks the effects of different interest and growth rate scenarios on the results found so far. The shown numbers base on a scenario with real interest rate of 2.8% p.a., which corresponds to a nominal interest rate of 4.3% p.a. assuming an inflation rate of 1.5% which is aspired by the European Central Bank. The current interest rate is below the assumed real interest rate, but the average rate since the 1970s is higher than this value. The same is true for the current growth rate; its average value since the 1970s is also higher than the assumed wage growth rate of 1.5%. Current reforms

[180] affecting the survivors of each cohort in the life tables

aim to strengthen productivity and regain higher growth rates; on the other hand, the demographic change might weaken economic growth and reduce the probability to reach a long-run growth rate of 1.5% or more.

The procedure here is thus as follows. First, the growth rate is assumed being constant, while the interest rate will be varied (r = 2,0% / 2,8% / 3,5%, cf. Table 8.15), while in the second comparison, interest rates will be held constant and growth rates are varied (g = 1,0% / 1,5% / 2,0%, cf. Table 8.16). In a third step, a pessimistic scenario is contrasted by an optimistic one (r=2,0% and g=1,0% versus r=3,5% and g=2,0%, cf. Table 8.17).

Table 8.15: Filling the pension gap in dependence of the capital market return

Low interest rate r = 2.0%					
	Net financial wealth and accrued savings				
	own LE	own LE + 3	own LE + 5	own LE + 10	LE from life tables
Mean	323.6%	256.1%	229.7%	186.9%	270.0%
Median	174.1%	142.1%	126.5%	100.4%	139.2%
Std. error	23.4%	17.5%	15.4%	12.2%	16.8%
Obs.	575	579	579	579	599
Gap in provision	35.1%	38.7%	41.6%	47.5%	38.7%
Medium interest rate r = 2.8%					
	Net financial wealth and accrued savings				
	own LE	own LE + 3	own LE + 5	own LE + 10	LE from life tables
Mean	362.3%	290.1%	262.0%	216.9%	305.0%
Median	198.8%	157.2%	141.6%	117.6%	161.0%
Std. error	25.6%	19.5%	17.4%	14.1%	19.1%
Obs.	575	579	579	579	599
Gap in provision	32.4%	35.8%	37.3%	43.7%	35.1%
High interest rate r = 3.5%					
	Net financial wealth and accrued savings				
	own LE	own LE + 3	own LE + 5	own LE + 10	LE from life tables
Mean	400.2%	323.5%	294.0%	246.9%	339.4%
Median	216.8%	175.6%	158.1%	132.4%	179.4%
Std. error	28.0%	21.7%	19.4%	16.0%	21.4%
Obs.	575	579	579	579	599
Gap in provision	31.1%	32.8%	34.2%	39.2%	33.2%

Note: Weighted values. Wage growth rate g = 1.5%.

A higher interest rate clearly rises the level of coverage of the pension gap, while lower interest rates makes this task more difficult. The effects are also not strong enough to change the number of households not being able to fill the gap in a large scope; the ratio remains by about 1/3.

Table 8.16: Filling the pension gap in dependence of the wage growth rate

Weak wage growth rate g = 1.0%					
	Net financial wealth and accrued savings				
	own LE	own LE + 3	own LE + 5	own LE + 10	LE from life tables
Mean	379.8%	303.8%	274.4%	227.2%	319.9%
Median	207.0%	165.3%	147.9%	122.2%	169.0%
Std. error	26.8%	20.5%	18.2%	14.8%	20.2%
Obs.	575	579	579	579	599
Gap in provision	31.3%	34.0%	36.4%	42.1%	33.7%
Medium wage growth rate g = 1.5%					
	Net financial wealth and accrued savings				
	own LE	own LE + 3	own LE + 5	own LE + 10	LE from life tables
Mean	362.3%	290.1%	262.0%	216.9%	305.0%
Median	198.8%	157.2%	141.6%	117.6%	161.0%
Std. error	25.6%	19.5%	17.4%	14.1%	19.1%
Obs.	575	579	579	579	599
Gap in provision	32.4%	35.8%	37.3%	43.7%	35.1%
Strong wage growth rate g = 2.0%					
	Net financial wealth and accrued savings				
	own LE	own LE + 3	own LE + 5	own LE + 10	LE from life tables
Mean	345.8%	277.1%	250.4%	207.3%	291.0%
Median	186.2%	152.1%	137.8%	111.4%	150.1%
Std. error	24.6%	18.7%	16.6%	13.4%	18.1%
Obs.	575	579	579	579	599
Gap in provision	33.7%	36.8%	39.6%	45.1%	36.4%

Note: Weighted values. Interest rate r = 2.8%.

A stronger wage growth rate affects the relative position of retirees. Since economic status-preserving values are compared here, the relative value of savings is negatively affected by a higher wage growth rate. This is why a higher growth rate would lower the share of coverage of the pension gap.

Table 8.17: Filling the pension gap: two different scenarios

Optimistic scenario (g = 2.0%, r = 3.5%)					
	Net financial wealth and accrued savings				
	own LE	own LE + 3	own LE + 5	own LE + 10	LE from life tables
Mean	381.6%	308.8%	280.7%	235.7%	323.4%
Median	210.7%	166.8%	151.4%	126.4%	169.5%
Std. error	26.7%	20.7%	18.5%	15.2%	20.3%
Obs.	575	579	579	579	599
Gap in provision	31.5%	33.7%	35.6%	41.6%	34.1%
Pessimistic scenario (g = 1.0%, r = 2.0%)					
	Net financial wealth and accrued savings				
	own LE	own LE + 3	own LE + 5	own LE + 10	LE from life tables
Mean	338.8%	267.9%	240.2%	195.5%	282.9%
Median	184.7%	146.0%	130.4%	106.4%	148.0%
Std. error	24.3%	18.2%	16.1%	12.7%	17.7%
Obs.	575	579	579	579	599
Gap in provision	34.1%	37.8%	40.8%	45.9%	37.7%

Note: Weighted values.

The difference between the optimistic and the pessimistic scenario are, given the results from Tables 8.15 and 8.16, relatively small. Scenarios for interest and wage growth rate thus do not change the results qualitatively.

8.5.4 Household types and the ability to fill the pension gap

This part examines household types by three characteristics, schooling, age and income, and checks whether differences in any of these three variables influences the household's ability to fill the pension gap. The procedure is to divide households for each of these three variables into three groups. For age and income, terciles are built to uniformly distribute households.

Table 8.18: Filling the pension gap in dependence of age

	Age		
	40 to 45	46 to 52	53 bis 65
Mean	349.65%	360.63%	375.25%
Median	197.86%	220.68%	181.54%
Std. error	37.83%	37.93%	54.58%
Obs.	192	192	191
Gap in provision	32.29%	31.77%	32.98%
Relation of monthly pensions	86.51%	89.12%	91.50%
Pension values before 'Riester' (Median)	1,480.47 €	1,342.89 €	1,191.72 €
Pension gap (Median)	196.19 €	145.91 €	102.32 €
Monthly life annuities (Median)	376.79 €	335.71 €	168.13 €

Note: Weighted values.

The share of households being able to fill the gap is rather constant over age classes. On a first glance, this is surprising since older households are not as hard affected by the pension reform as younger households (cf. Table 8.10). But younger households have a larger amount of financial wealth which will accrue to the pension entry age from which they will receive higher monthly life annuities.

Table 8.19: Filling the pension gap in dependence of income

	Income		
	up to 1800€	1800 up to 2750 €	more than 2750 €
Mean	193.08%	356.07%	621.43%
Median	71.65%	233.66%	395.34%
Std. error	22.92%	38.85%	64.51%
Obs.	217	163	194
Gap in provision	53.00%	28.83%	12.37%
Relation of monthly pensions	90.06%	88.53%	88.55%
Pension values before 'Riester' (Median)	977.56 €	1,479.84 €	1,859.53 €
Pension gap (Median)	94.32 €	169.36 €	201.65 €
Monthly life annuities (Median)	71.37 €	358.85 €	806.33 €

Note: Weighted values.

When looking at pension gaps and income classes, one can find that the calculated pension reduction factors[181] are distributed relatively evenly over income classes. This is due to the age distribution which is relatively even between income classes. Differences arise concerning the ability to fill the pension gap since households in the upper third of the income distribution are better prepared due to a larger financial provision *relative to their income*. This is important to mention since, of course, the pension gap of these households is also higher, according to their income and their pension entitlements.

[181] which take account of the affect of the pension reforms

Table 8.20: Filling the pension gap in dependence of schooling levels

	Schooling		
	Hauptschule/ Mittlere Reife	Abitur/Fach-. hochschulabschl.	FH/Studium-abschluss
Mean	297.84%	361.30%	624.98%
Median	151.87%	220.68%	375.91%
Std. error	22.87%	57.97%	97.20%
Obs.	413	50	112
Gap in provision	38.74%	30.00%	9.82%
Relation of monthly pensions	89.12%	87.72%	88.59%
Pension val. before 'Riester' (Median)	1,280.10 €	1,600.82 €	1,632.66 €
Pension gap (Median)	134.10 €	189.17 €	179.41 €
Monthly life annuities (Median)	187.13 €	411.19 €	630.78 €

Note: Weighted values.

Similar to income classes, the separation for schooling classes reveals an even distribution for the pension reduction factor, as age is evenly distributed between the schooling classes. But concerning the ability to fill the pension gap, Table 8.20 shows large differences. The share of households not being able to fill the gap is much smaller for the group with a college degree. One might presume that this is due to a higher associated income, but that is not the case as can be at the pension entitlements. They are much the same for the second and the third group, but their financial wealth and savings is not. The share of households with zero savings is much smaller in the group with the highest schooling. This can have two explanations. The first one is that households with a college degree are more disciplined and self-controlled concerning their financial planning and foresight. The other one is that that their response behavior might differ from households with a lower schooling, which might more likely to tend to escape the effort to answer to 'annoying' questions for savings and wealth.

8.6 Conclusion

This chapter analyzes how many households are already prepared to fill the upcoming pension gap, assuming no changes of of financial behavior. Considering accrued[182] financial wealth and monthly savings as the basis for calculating monthly life annuities, it can be shown that about 1/3 of all households in the sample will not be able to fill the pension gap they will have to face, even if they use all their financial wealth and savings. Even the median household would nearly loose all degree of freedoms for other financial allocation choices. If a household is forced to consume all the financial wealth to reach to the pension level known today, it practically eliminates the possibilities reaching a *higher* pension level which is closer to the income flows before retirement.

The values of subjective life expectancy, which is a crucial variable in this analysis, are shown to be assumed independent of age which is counterfactual. Therefore, the subjective life expectancy is compared to the influence of more realistic values. This affects additional 3 to 10% of the households in the sample which, assuming the realistic values, will not be able to fill the pension gap.

The lessons to be learned are that the long-term savings rate will need to increase if we are to master the challenges posed by the demographic trends which, in the final analysis, are the reason for the introduction of the sustainability factor. Policymakers would be well advised to draw attention to these developments.

Despite low uptake and acceptance, the introduction of the Riester pension has at the very least increased peoples' awareness of the problems the future holds. Households now reflect considerably more on the provision they are making for their old age than was the case prior to the reform. The task must be to reinforce this trend.

[182] at the time of the households' pension entry age

8.A Design of the subjective life expectancy questions

Subjective life expectancy is asked in a three step question to keep the level of concern about this subject as low as possible.

The wording of the questions was:

1. *What do you belief: up to what age will men and women of your age live?*
2. *When thinking about your living and health situation. What do you belief: compared to persons of your gender and age, will you live shorter, about as long, longer?*
   followed by the question for the number of how many years longer or shorter that might be.
3. If answer to last question was 'shorter' or 'longer': *Why do you belief to live shorter / longer than the average?*
   followed by a list with four possible reasons respondents might think of and one open field.

The interview procedure was repeated in the same way with the respondent's partner.

When calculating subjective life expectancies, one has to be aware of the gender of the person to which the calculation applies (respondent/partner) since the first question asks not for the average of persons of the *same* gender but for both, men and women.

8.B Tables

Table 8.21: Regression results: Age of retirement and replacement rate

	Age of retirement		Replacement rate	
	Coeff.	$P > t$	Coeff.	$P > t$
Permanent net income / 10,000	-3.09	0.352	0	0.121
(Perm. income/10,000) squared	2.87	0.508	-0.343	0.156
Age / 10	-0.638	0.218	0.056	0.047
(Age / 10) squared	0.051	0.383	-0.004	0.219
Intermediate secondary school leaving certificate (D)	**0.646**	0.003	**-0.017**	0.162
Upper secondary school leaving certificate (D)	**0.538**	0.057	**-0.027**	0.106
University/polytechnic degree (D)	**1.261**	0.000	**-0.031**	0.052
Children (D)	**0.698**	0.032	**0.005**	0.781
Children living in same household (D)	-0.243	0.347	0.003	0.814
Job: Employee (D)	-0.002	0.993	0.002	0.890
Job: Civil servant (D)	**-1.888**	0.000	**0.076**	0.000
In part-time employment (D)	0.064	0.842	-0.025	0.175
In marginal part-time employment (D)	**0.747**	0.035	**-0.066**	0.003
Not gainfully employed (D)	0.244	0.422	-0.078	0.000
Unemployed (D)	-0.31	0.378	0.041	0.065
Unemployed for more than one month	-0.031	0.895	0.004	0.732
Unemployed for more than six months	0.02	0.943	-0.011	0.479
Partner (D)	-0.2	0.556	-0.021	0.289
Separated or divorced (D)	-0.106	0.767	-0.021	0.332
Widowed (D)	0.332	0.350	0.012	0.563
Gender: female (D)	**-0.972**	0.000	**0.002**	0.881
Eastern Germany (D)	0.64	0.014	-0.027	0.054
Subsample: RR 2003	**1.269**	0.000	**-0.051**	0.002
Subsample: TPI 2004	**-0.306**	0.200	**-0.011**	0.415
Constant	**64.582**	0.000	**0.429**	0.000
Number of observations	1856		941	
F(33, 1100 / F(20, 661)	4.87		6.63	
Prob > F	0		0	
R sq.	0.06		0.148	
Adj. R sq.	0.048		0.126	

Table 8.22: Regression results: Respondents' life expectancy

	Coeff.	$P > t$
Permanent income / 10,000	9.056	0.53
(Permanent income / 10,000) squared	-5.14	0.813
Age / 10	**-0.714**	0.004
(Age / 10) squared	**0.008**	0.002
Intermediate secondary school leaving certificate (D)	-0.426	0.555
Upper secondary school leaving certificate (D)	0.068	0.942
University/polytechnic degree (D)	-1.092	0.259
Children (D)	0.572	0.603
Children living in same household (D)	**-1.873**	0.015
Job: Employee (D)	-0.576	0.542
Job: Civil servant (D)	1.537	0.238
Job: Freelancer (D)	3.368	0.240
Job: Self-employed (D)	-0.667	0.589
Pensioner (D)	**-4.222**	0.007
In part-time employment (D)	-0.004	0.997
In marginal part-time employment (D)	**4.488**	0.001
Not gainfully employed (D)	**2.511**	0.067
Unemployed (D)	**-2.926**	0.073
Unemployed for more than one month	1.14	0.126
Unemployed for more than six months	**-2.346**	0.014
Partner (D)	**-2.157**	0.128
Separated or divorced (D)	-0.077	0.953
Widowed (D)	1.728	0.568
Gender: female (D)	**3.399**	0.000
Eastern Germany (D)	-0.835	0.335
Smoker (D)	-0.842	0.243
Former smoker (D)	**-1.212**	0.075
Expectations regarding health status	0.229	0.121
Self appraisal: optimist	**0.217**	0.077
Live less long owing to: Illness (D)	**-5.088**	0.000
Live less long owing to: Life circumstances (D)	-1.133	0.426
Live less long owing to: Early death of family member (D)	**-5.811**	0.001
Live less long owing to: Other reasons (D)	**-3.884**	0.097
Live longer owing to: Health status (D)	**3.583**	0.010
Live longer owing to: Life circumstances (D)	**3.842**	0.001
Live longer owing to: Longevity of family members (D)	**3.341**	0.011
Live longer owing to: Other reasons (D)	**4.884**	0.009
Self assessment of risk: Health	0.002	0.982
Constant	**89.902**	0.000
Number of observations	430	
F(33, 1100 / F(20, 661)	10.6	
Prob > F	0	
R sq.	0.5074	
Adj. R sq.	0.46	

Table 8.23: Regression results: Respondents' life expectancy

	Respondent		Partner	
	Coeff.	$P > t$	Coeff.	$P > t$
Permanent income / 10,000	-14.04	0.454		
(Permanent income / 10,000) squared	25.554	0.344		
Age / 10	**-0.905**	0.003		
(Age / 10) squared	**0.01**	0.002		
Age diff. to partner ($Age_P - Age_{BP}$)			**0.208**	0.022
Secondary school (D)	-0.104	0.902	-0.406	0.616
Graduation diploma (D)	-0.106	0.926	-0.186	0.873
University degree (D)	-1.455	0.216	**3.315**	0.004
Kids (D)	-0.571	0.682		
Kids living in same house (D)	-1.081	0.265		
Job: blue collar (D)	**-2.162**	0.051	-0.859	0.563
Job: civil servant (D)	-0.516	0.750	-1.549	0.366
Job: freelancer (D)	2.978	0.388	-3.426	0.286
Job: self-employed (D)	-1.35	0.332	0.05	0.986
Retired(D)	-2.721	0.120	-0.356	0.798
Work parttime (D)	1.565	0.292	1.06	0.341
Work little (D)	**2.91**	0.066	-1.614	0.209
Work not (D)	-0.685	0.662	0.035	0.977
Unemployed (D)	-2.9	0.124	-1.465	0.378
Work parttime (D)	0.744	0.388	0.4	0.645
Work little (D)	-0.96	0.392	0.489	0.627
Work not (D)	2.285	0.234		
Widowed (D)	2.121	0.722		
Female (D)	**-4.887**	0.000		
Eastern Germany (D)	0.113	0.914		
Expectations regarding health status	**-0.373**	0.099	**0.904**	0.000
Self appraisal: optimist	**0.473**	0.001		
Live shorter: Illness (D)	-1.227	0.435		
Live shorter: Life circumstances(D)	**2.953**	0.068		
Live shorter: Early death of family members (D)	-3.433	0.115		
Live shorter: Other reasons (D)	-3.416	0.227		
Live longer: Health status (D)	2.196	0.152		
Live longer: Life circumstances(D)	**2.298**	0.080		
Live longer: Longevity of family members (D)	-1.43	0.328		
Live longer: Other reasons (D)	**4.48**	0.081		
Constant	**96.64**	0		
Number of observations		365		
F(33, 1100) / F(20, 661)		3.9		
Prob >F		0		
R sq.		0.3831		
Adj. R sq.		0.2849		

Table 8.24: Regression results: logarithm of relative financial wealth

	Coef.	P>t
Net income / 10,000	-0.287	0.876
(Net income/10,000) squared	0.710	0.735
Age	0.059	0.002
Age squared	0.000	0.135
Household size	-0.059	0.599
Children (D)	-0.095	0.689
Children living in same household (D)	-0.087	0.721
Eastern Germany (D)	-0.489	0.009
Village (D)	0.159	0.622
Secondary school (D)	0.162	0.330
Graduation dipl. (D)	-0.227	0.380
University (D)	0.328	0.118
Job: Employee (D)	-0.062	0.711
Job: Civil servant (D)	-0.203	0.331
Work parttime (D)	0.055	0.924
Work little (D)	-0.955	0.122
Work not (D)	-0.361	0.461
Unemployed (D)	0.273	0.581
Female (D)	-0.210	0.112
Widowed (D)	0.240	0.596
Separated or divorced (D)	-0.360	0.225
Partner (D)	-0.454	0.138
FW: Saving accounts (D)	0.604	0.000
FW: Build. soc. contr. (D)	0.331	0.014
FW: Whole life ins. (D)	1.063	0.000
FW: Priv. old-age prov. (D)	0.274	0.073
FW: Bonds (D)	0.821	0.000
FW: Stocks (D)	0.679	0.000
Home owner (D)	0.240	0.103
Constant	-1.137	0.131
Number of observations	354	
Prob > F	0	
F(33, 1100 / F(20, 661)	7.92	
R squared	0.4149	
Adj. R squared	0.3625	

8.C Calculation formulas

A number of methods used in financial mathematics to present future value and annuities were applied in Sections 8.3 and 8.4. The equations used to calculate individual values are presented in brief here.

The future value of a current asset is determined by

$$FV(W_{retirement\,age_{HH}}) = (1+gr)^{retirement\,age_{HH}} \tag{8.4}$$

where

$$\begin{aligned} FV(W_{retirement\,age_{HH}}) &= \text{the future value of an asset } W \text{ at the} \\ &\quad \text{age of retirement of the household } HH \\ gr &= \text{the effective growth rate} \approx r - g \end{aligned}$$

We assume that savings remain constant at today's levels every year until retirement. The future value of savings at the age of retirement is therefore:

$$FV(S_{retirement\,age_{HH}}) = S^t_{HH} \cdot \frac{(1+gr)^{(retirement\,age_{HH} - age_{HH})} - 1}{gr} \tag{8.5}$$

where

$S_{retirement\,age_{HH}}$, S^t_{HH} represents the household's HH savings in year t or on retirement

The fraction in (8.5) forms the inverse of the annuity equation as, in this case, constant contributions are invested over a specific period of time.

Assuming that contributions are not paid in annually but on a monthly basis, the following modification must be made to (8.5):

$$V(S_{retirement\,age_{HH}}) = \frac{S^t_{HH}}{12} \cdot \frac{gr}{(1+gr)^{\frac{1}{12}-1}} \frac{(1+gr)^{(retirement\,age_{HH} - age_{HH})} - 1}{gr} \tag{8.6}$$

The annuity arising from all the savings accumulated up to the age of retirement, including the accrued assets and accrued interest, is calculated as follows:

$$MLA^m_{HH} = FVTA \cdot \frac{r}{1-(1+r)^{-retirement\,years+1}} \cdot \frac{(1+r)^{\frac{1}{12}-1}}{r} \tag{8.7}$$

where

$$\begin{aligned} MLA^m_{HH} &= \text{the annuity of a household in month } m \text{ following retirement} \\ FVTA &= \text{the future value of the entire assets} \\ retirement\,years &= \#\text{ of years during which person continues to live after retirement} \\ &= \text{subjective life expectancy - pension entry age} \\ r &= \text{the real interest rate} \end{aligned}$$

List of figures

List of tables

Bibliography

Abdellaoui, M. (2000): Parameter-free elicitation of utility and probability weighting functions. *Management Science*, 46(11), 1497–1512.

Aguiar, M. and E. Hurst (2004): Consumption vs. expenditure. NBER Working Paper No. 10307.

Alessie, R., A. Brugiavini, and G. Weber (2004): Household saving and cohabitation. Mimeo, presented at ISOM 2004.

Alkemade, P., C. Biancotti, and I. Faiella (2003): Preliminary report on the LWS institutional database. Luxembourg Income Study Report.

Andersen, R., J. Kasper, M. R. Frankel, and Associates (1979): *Total Survey Error*. San Francisco: Jossey-Bass.

Arrondel, L. (2002): Risk management and wealth accumulation behavior in France. *Economics Letters*, 74(2), 187–194.

Attanasio, O. (1998): A cohort analysis of saving behavior by U.S. households. *Journal of Human Resources*, 33(3), 575–609.

Attanasio, O. P. (1999): Consumption. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, 741–812. Amsterdam: Elsevier.

Barro, R. J. (1974): Are government bonds net wealth? *Journal of Political Economy*, 82(6), 1095–1117.

Barro, R. J. (1978): *The Impact of Social Security on Private Savings - Evidence from the U.S. Time Serie*. Washington, D.C.: American Enterprise Institute.

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro (1997): Preference parameters and behavioral heterogeneity: An experimental approach in the Health and Retirement Study. *Quarterly Journal of Economics*, 112(2), 537–579.

Battistin, E., R. Miniaci, and G. Weber (2003): What do we learn from recall consumption data? *Journal of Human Resources*, 38(2), 354–385.

Baumol, W. J. and R. E. Quandt (1964): Rules of thumb and optimal imperfect decisions. *American Economic Review*, 54(2), 23–46.

Beatty, P. and D. Herrmann (2002): To answer or not to answer: Decision processes related to survey item nonresponse. In R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little (Eds.), *Survey Nonresponse*, 71–85. New York: Wiley.

Becker, G. (1986): A theory of the allocation of time. *Economic Journal*, 75(299), 493–508.

Berkel, B. and A. Börsch-Supan (2003): Pension reform in Germany: The impact on retirement decisions. NBER Working Paper No. 9913.

Bertelsmann-Stiftung (2003): Vorsorgereport. Private Alterssicherung in Deutschland.

Biewen, M. (2001): Item non-response and inequality measurement: Evidence from the German earnings distribution. *Allgemeines Statistisches Archiv*, 85(4), 409–425.

Börsch-Supan, A. (1992): Population aging, social security design, and early retirement. *Journal of Institutional and Theoretical Economics*, 148, 533–557.

Börsch-Supan, A. (2000): Incentive effects of social security on labor force participation: Evidence in Germany and across Europe. *Journal of Public Economics*, 78(1), 25–49.

Börsch-Supan, A. (2001): Incentive effects of social security under an uncertain disability option. In D. A. Wise (Ed.), *Themes in the Economics of Aging*, 281–303. Chicago, IL and London, UK: University of Chicago Press.

Börsch-Supan, A. (2004a): Faire Abschläge in der Rentenversicherung. *Sozialer Fortschritt*. Forthcoming.

Börsch-Supan, A. (2004b): Sozialversicherung. In E. Gaugler, W. A. Oechsler, and W. Weber (Eds.), *Handwörterbuch des Personalwesens*, 1800–1809. Stuttgart: Schäffer-Poeschel.

Börsch-Supan, A. and L. Essig (2002): *Sparen in Deutschland: Ergebnisse der ersten SAVE-Studie*. Köln: Deutsches Institut für Altersvorsorge.

Börsch-Supan, A. and L. Essig (2003): Household savings in Germany: Results of the first SAVE study. NBER Working Paper No. 9902.

Börsch-Supan, A., F. Heiss, and J. Winter (2004a): *Akzeptanzprobleme Bei Rentenreformen*. Köln: Deutsches Institut für Altersvorsorge.

Börsch-Supan, A., A. Ludwig, and A. Reil-Held (2004b): Hochrechnungsmethoden und Szenarien für gesetzliche und private Renteninformationen. MEA Discussion Paper 49-04.

Börsch-Supan, A., A. Ludwig, and M. Sommer (2003): *Demographie und Kapitalmärkte. Die Auswirkungen der Bevölkerungsalterung auf Aktien-, Renten-, und Immobilienvermögen*. Köln: Deutsches Institut für Altersvorsorge.

Börsch-Supan, A., A. Reil-Held, R. Rodepeter, R. Schnabel, and J. Winter (1999): Ersparnisbildung in Deutschland: Meßkonzepte und Ergebnisse auf Basis der EVS. Working Paper No. 99-02, Sonderforschungsbereich 504, Universität Mannheim.

Börsch-Supan, A., A. Reil-Held, R. Rodepeter, R. Schnabel, and J. Winter (2001): The German savings puzzle. *Research in Economics*, 55(1), 15–38.

Börsch-Supan, A., R. Schnabel, S. Kohnz, and G. Mastrobuoni (2004c): Modelling of retirement choices in germany. In J. Gruber and D. Wise (Eds.), *Social Security Programs and Retirement Around the World: Micro-Estimation*, 285–343. Chicago, IL and London, UK: University of Chicago Press.

Börsch-Supan, A. and K. Stahl (1991): Life cycle savings and consumption constraints: Theory, empirical evidence, and fiscal implications. *Journal of Population Economics*, 4(3), 233–255.

Börsch-Supan, A. and C. B. Wilke (2003): The German public pension system: How it was, how it will be. MEA Discussion Paper No. 34-2003.

Bound, J., C. Brown, and N. Mathiowetz (2001): Measurement error in survey data. In J. J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5, 3705–3843. Amsterdam: Elsevier.

Braun, R., F. Burger, M. Miegel, U. Pfeiffer, and K. Schulte (2002): *Erben in Deutschland*. Köln: Deutsches Institut für Altersvorsorge.

Braun, R., M. Miegel, and U. Pfeiffer (2000): *Vermögensbildung unter neuen Rahmenbedingungen*. Köln: Deutsches Institut für Altersvorsorge.

Brennan, M., J. Hoek, and C. Astridge (1991): The effects of monetary incentives on the response rate and cost-effectiveness of a mail survey. *Journal of the Market Research Society*, 33(3), 229–241.

Bromiley, P. and S. Curley (1992): Individual differences in risk taking. In J. Yates (Ed.), *Risk-Taking Behavior*. New York: John Wiley.

Browning, M. and T. F. Crossley (2002): Recall expenditure questions in the Canadian Out of Employment Panel. Draft.

Browning, M. and T. F. Crossley (2004): Shocks, stocks and socks: Smoothing consumption over a temporary income loss. CAM Working Paper No. 2004-05.

Browning, M., T. F. Crossley, and G. Weber (2003): Asking consumption questions in general purpose surveys. *Economic Journal*, 113, F540–F567.

Browning, M. and A. Lusardi (1996): Household saving: Micro theories and micro facts. *Journal of Economic Literature*, 34(4), 1797–1855.

Brugiavini, A. and G. Weber (2003): Household saving: Concepts and measurement. In A. Börsch-Supan (Ed.), *Life-Cycle Savings and Public Policy*, 33–55. San Diego, CA, and London, UK: Academic Press.

Bulmahn, T. (2003): Zur Entwicklung der privaten Altersvorsorge. Vorsorgebereitschaft, Vorsorgeniveau und erwartete Absicherung im Alter. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 55(1), 29–54.

Burbidge, J. B., L. Magee, and A. L. Robb (1988): Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Associations*, 83(401), 123–127.

Caballero, R. J. (1991): Earnings uncertainty and aggregate wealth accumulation. *American Economic Review*, 81(4), 859–871.

Cagetti, M. (2003): Wealth accumulation over the life cycle and precautionary savings. *Journal of Business & Economic Statistics*, 21(3), 339–353.

Carroll, C. D. (1992): The buffer stock theory of saving: Some macroeconomic evidence. *Brookings Papers on Economic Activity*, 2, 61–135.

Carroll, C. D. (1997): Buffer-stock saving and the life cycle-permanent income hypothesis. *Quarterly Journal of Economics*, 112(1), 1–55.

Carroll, C. D., K. E. Dynan, and S. D. Krane (2003): Unemployment risk and precautionary wealth: Evidence from households' balance sheets. *Review of Economics and Statistics*, 85(3), 586–604.

Carroll, C. D. and A. A. Samwick (1997): The nature of precautionary wealth. *Journal of Monetary Economics*, 40(1), 41–71.

Carroll, C. D. and A. A. Samwick (1998): How important is precautionary saving? *Review of Economics and Statistics*, 80(3), 410–419.

Dardanoni, V. (1991): Precautionary savings under income uncertainty: A cross-sectional analysis. *Applied Economics*, 23(1), 153–160.

Deaton, A. (1992): *Understanding Consumption*. Oxford, UK: Clarendon Press.

Diamond, P. A. and J. A. Hausman (1984): Individual retirement and savings behavior. *Journal of Public Economics*, 23, 81–114.

Dillman, D. A. (2000): *Mail and Internet Surveys: The Tailored Design Method*. New York, NY: Wiley.

Drèze, J. and F. Modigliani (1972): Consumption decisions under uncertainty. *Journal of Economic Theory*, 5(3), 308–335.

Dynan, K. (1993): How prudent are consumers? *Journal of Political Economy*, 101(6), 1104–1113.

Engen, E. and J. Gruber (2001): Unemployment insurance and precautionary saving. *Journal of Monetary Economics*, 47(3), 545–579.

Essig, L. and A. Reil-Held (2003): Chancen und Risiken der Riester-Rente. MEA Discussion Paper 35-03.

Essig, L. and J. Winter (2003): Item nonresponse to financial questions in household surveys: An experimental study of interviewer and mode effects. MEA Working Paper 39-03.

Farquhar, P. H. (1984): Utility assessment methods. *Management Science*, 30, 1283–1300.

Fehr, H., D. Kiesewetter, and M. Myßen (2003): Die Riester-Rente - ein Flop? *ifo-Schnelldienst*, 5.

Feldstein, M. S. (1974): Social security, induced retirement, and aggregate capital accumulation. *Journal of Political Economy*, 82(5), 905–926.

Feldstein, M. S. (1978): Reply. In R. J. Barro (Ed.), *The Impact of Social Security on Private Savings - Evidence from the U.S. Time Serie*. Washington, D.C.: American Enterprise Institue.

Ferber, R. (1966): Item nonresponse in a consumer survey. *Public Opinion Quarterly*, 30(3), 399–415.

Fitzenberger, B., R. Hujer, T. E. MaCurdy, and R. Schnabel (2001): Testing for uniform wage trends in West Germany: A cohort analysis using quantile regression for censored data. *Empirical Economics*, 26(1), 41–86.

Frey, D., T. Greitemeyer, S. Schulz-Hardt, and E. Traut-Mattausch (2002): Erwartungsgeleitete Wahrnehmung bei der Einführung des Euro: Der Euro ist nicht immer ein Teuro. *Wirtschaftspsychologie*, 4, 22–28.

Friedman, M. (1957): *A Theory of the Consumption Function*. Princeton, NJ: Princeton University Press.

Gabler, S., J. H. P. Hoffmeyr-Zlotnik, and D. Krebs, Eds. **(1994):** *Gewichtung in der Umfragepraxis*. Opladen: Westdeuscher Verlag.

Gabler, S., J. H. P. Hoffmeyr-Zlotnik, and D. Krebs, Eds. **(1997):** *Stichproben in der Umfragepraxis*. Opladen: Westdeutscher Verlag.

Gentry, W. and G. Hubbard (2000): Entrepreneurship and household saving. NBER Working Paper No. 7894.

Gesamtverband der deutschen Versicherungswirtschaft (GDV) (2003): Jahresgutachten 2003.

Gourinchas, P.-O. and J. A. Parker (2002): Consumption over the life cycle. *Econometrica*, 70(1), 47–89.

Groves, R. M. (1989): *Survey Errors and Survey Costs*. New York, NY: Wiley.

Groves, R. M. and M. P. Couper (1998): *Nonresponse in Household Interview Surveys*. New York: Wiley.

Groves, R. M., D. A. Dillman, J. L. Eltinge, and R. J. A. Little (2002): *Survey Nonresponse*. New York: Wiley.

Guiso, L., T. Jappelli, and D. Terlizzese (1992): Earnings uncertainty and precautionary saving. *Journal of Monetary Economics*, 30(2), 307–337.

Heaton, J. and D. Lucas (2000): Portfolio choice and asset prices: The importance of entrepreneurial risk. *Journal of Finance*, 55(3), 1163–1198.

Heien, T. and K. Kortmann (2003): *Spar- und Finanzanlageverhalten privater Haushalte (SAVE II). Methodenbericht*. München: Infratest Sozialforschung.

Horowitz, J. L. and C. F. Manski (1995): Identification and robustness with contaminated and corrupted data. 63(2), 281–302.

Horowitz, J. L. and C. F. Manski (1998): Censoring of outcomes and regressors due to survey nonresponse: Identification and estimation using weights and imputations. 84, 37–58.

Hox, J. and E. de Leeuw (2002): The influence of interviewers' attitude and behavior on household survey nonresponse: An international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little (Eds.), *Survey Nonresponse*, 103–120. New York: Wiley.

Hubbard, R. G., J. Skinner, and S. P. Zeldes (1995): Precautionary saving and social insurance. *Journal of Political Economy*, 103(2), 360–399.

Hurd, M. D. (1990): Research on the elderly: Economic status, retirement, and consumption and saving. *Journal of Economic Literature*, 28, 565–637.

Hurd, M. D., F. T. Juster, and J. P. Smith (2003): Enhancing the quality of data on income: Recent innovations from the HRS. *Journal of Human Resources*, 38(3), 758–772.

Hurd, M. D., D. McFadden, H. Chand, L. Gan, A. Merrill, and M. Roberts (1998): Consumption and savings balances of the elderly: Experimental evidence on survey response bias. In D. Wise (Ed.), *Frontiers in the Economics of Aging*, 353–387. Chicago, IL: University of Chicago Press.

Hurd, M. D. and K. McGarry (1997): The predictive validity of subjective probabilities of survival. NBER Working Paper No. 6193.

Ifo Institut (2002): Euro – Teuro ? ifo Standpunkt Nr. 33.

Jappelli, T. and F. Modigliani (1998): The age-saving profile and the life-cycle hypothesis. *CSEF Working Paper No. 4*.

Jappelli, T. and M. Pagano (1994): Personal saving in Italy. In J. Poterba (Ed.), *International Comparisons of Household Saving*, 237–268. Chicago: University of Chicago Press.

Johnson, N. L. (1949): Systems of frequeny curves generated by methods of translation. *Biometrika*, 36, 149 –176.

Jürges, H. (2001): Do Germans save to leave an estate? An examination of the bequest motive. *Scandinavian Journal of Economics*, 103(3), 391–414.

Juster, F. T., J. P. Lupton, and H. Cao (2002): Ensuring time-series consistency in estimates of income from wealth. Michigan Retirement Research Center Working Paper 2002-030, University of Michigan.

Juster, F. T. and J. P. Smith (1997): Improving the quality of economic data: Lessons from the HRS and AHEAD. *Journal of the American Statistical Association*, 92(440), 1268–1278.

Kahneman, D. and A. Tversky (1979): Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.

Kapteyn, A., R. Alessie, and A. Lusardi (2004): Explaining the wealth holdings of different cohorts: Productivity growth and Social Security. *European Economic Review, forthcoming.*

Kazarosian, M. (1997): Precautionary saving - a panel study. *Review of Economics and Statistics*, 79(2), 241–247.

Kennickell, A. and A. Lusardi (2004): Disentangling the importance of the precautionary saving motive. Dartmouth College Working Paper. Unpublished.

Keynes, J. M. (1936): *The General Theory of Employment, Interest and Money.* London: Macmillan.

Kimball, M. S. (1990): Precautionary saving in the small and in the large. *Econometrica*, 58(1), 53–73.

King, B. F. (1983): Quota sampling. In W. G. Madow, I. Olin, and D. B. Rubin (Eds.), *Incomplete Data in Sample Surveys*, Volume 2, chapter 6, 63–71. New York: Academic Press.

King, M. and L. D. Dicks-Mireaux (1982): Asset holdings and the life cycle. *Economic Journal*, 92(366), 247–67.

Klein, S. and R. Porst (2000): Mail Surveys. Ein Literaturbericht. Technischer Bericht Nr. 10/2000, ZUMA, Mannheim.

Kommission zur Nachhaltigkeit der Finanzierung der sozialen Sicherungssysteme (2003): Bericht der kommission. http://www.bmgs.bund.de/deu/gra/themen/sicherheit/kommission/index.cfm.

Kortmann, K. (2003): *Situation und Entwicklung der betrieblichen Altersversorgung in Privatwirtschaft und öffentlichem Dienst 2001 - 2003. Endbericht. Studie im Auftrag des Bundesministeriums für Gesundheit und Soziale Sicherung (BMGS).* München: Infratest Sozialforschung. http://www.bmgs.bund.de/download/broschueren/F304.PDF.

Kotlikoff, L. J. (1989): *What Determines Savings?* Cambridge, MA: MIT Press.

Laibson, D. (1997): Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2), 443–477.

Laibson, D. and R. Zeckhauser (1998): Amos Tversky and the ascent of behavioral economics. *Journal of Risk and Uncertainty*, 16(1), 7–47.

Leland, H. E. (1968): Saving and uncertainty: The precautionary demand for saving. *Quarterly Journal of Economics*, 82(3), 465–473.

Little, R. J. A. and D. B. Rubin (1987): *Statistical Analysis with Missing Data*. New York: John Wiley & Sons.

Lusardi, A. (1996): Permanent income, current income, and consoumtion: Evidence from two panel data sets. *Journal of Business and Economic Statistics*, 14(1), 81–90.

Lusardi, A. (1997): Precautionary saving and subjective earnings variance. *Economics Letters*, 57(3), 319–326.

Lusardi, A. (1998): On the importance of the precautionary saving motive. *American Economic Review Papers and Proceedings*, 88(2), 449–453.

Lusardi, A. (2000): Precautionary saving and the accumulation of wealth. Dartmouth College Working Paper.

MacKinnon, J. G. and L. Magee (1990): Transformin the dependent variable in regression models. *International Economic Review*, 31(2), 315–339.

Meghir, C. (2004): A retrospective on Friedman's theory of permanent income. *The Economic Journal*, 114(496), 293–306.

Meyer, B. D. (1995): Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, 13(2), 151–161.

Modigliani, F. and R. Brumberg (1954): Utility analysis and the consumption function: An interpretation of cross-section data. In K. K. Kurihara (Ed.), *Post-Keynesian Economics*, 388–436. New Brunswick, N.J.: Rutgers University Press.

Moore, J. C., L. L. Stinson, and E. J. Welnik, Jr. (1999): Income reporting in surveys: Cognitive issues and measurement error. In M. G. Sirken, D. J. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur, and R. Tourangeau (Eds.), *Cognition and Survey Research*, 155–174. New York: Wiley.

Murata, K. (2003): Precautionary savings and income uncertainty: Evidence from Japanese micro data. *Monetary and Economic Studies*, 21(3), 21–52.

Nicoletti, C. and F. Peracchi (2001): Sample particpation and income nonresponse in the ECHP. Unpublished manuscript, Tor Vergata University, Rome.

O'Donoghue, T. and M. Rabin (1999): Doing it now or later. *American Economic Review*, 89(1), 103–124.

Oeppen, J. and J. W. Vaupel (2002): Broken limits to life expectancy. *Science*, 296, 1029–1030.

Pennings, J. M. E. and A. Smidts (2003): The shape of utility functions and organizational behavior. *Management Science*, 49(9), 1251–1263.

Philipson, T. (1997): Data markets and the production of surveys. *Review of Economic Studies*, 64(1), 47–72.

Philipson, T. (2001): Data markets, missing data, and incentive pay. *Econometrica*, 69(4), 1099–1111.

Philipson, T. and A. Malani (1999): Measurement errors: A principal investigator-agent approach. 91, 273–298.

Porst, R. (1996): Ausschöpfungen bei sozialwissenschaftlichen Umfragen: Die Sicht der Institute. Arbeitsbericht Nr. 96/07, ZUMA, Mannheim.

Reil-Held, A. (2000): Einkommen und Sterblichkeit in Deutschland: Leben Reiche länger? Discussion Paper No. 00-14, Sonderforschungsbereich 504, Universität Mannheim.

Rice, S. A. (1929): Contagious bias in the interview: A methodological note. *Journal of Sociology*, 35, 420–423.

Riphahn, R. T. and O. Serfling (2002): Item non-response on income and wealth questions. Discussion Paper No. 573, IZA Bonn.

Rodepeter, R. (1999): Identifikatikon von Sparprofilen im Lebenszyklus. Dissertation, Universität Mannheim.

Rodepeter, R. and J. Winter (1999): Rules of thumb in life-cycle saving decisions. Discussion Paper No. 99-81, Sonderforschungsbereich 504, Universität Mannheim.

Sandmo, A. (1970): The effect of uncertainty on saving decisions. *Review of Economic Studies*, 37(3), 353–360.

Schmidt, P. (1995): *Die Wahl des Rentenalters*. Bern, New York, Paris, Wien: Lang.

Schnabel, R. (2003): *Die Rentenlücke. Das Problem wächst*. Köln: Deutsches Institut für Altersvorsorge.

Schnell, R. (1997): *Nonresponse in Bevölkerungsumfragen*. Opladen: Leske+Budrich.

Schräpler, J.-P. (2003): Gross income non-response in the German Socio-Economic Panel: Refusal or don't know? *Schmollers Jahrbuch*, 123, 109–124.

Schunk, D. and C. Betsch (2004): Explaining heterogeneity in utility functions by individual differences in preferred decision models. Discussion Paper No. 04-26, Sonderforschungsbereich 504, Universität Mannheim.

Schwarz, N. (1999): Self-reports: How the question shapes the answers. *American Psychologist*, 54(2), 93–105.

Siddiqui, S. (1997): The pension incentive to retire: Empirical evidence for West Germany. *Journal of Population Economics*, 10(4), 463–486.

Singer, E. (2002): The use of incentives to reduce nonresponse in household surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little (Eds.), *Survey Nonresponse*, 163–177. New York: Wiley.

Skinner, J. (1988): Risky income, life-cycle consumption and precautionary savings. *Journal of Monetary Economics*, 22(2), 237–255.

Slovic, P., B. F. and S. Lichtenstein (1986): *The psychometric study of risk perception.*, 3–24. Plenum Press: New York.

Slovic, P. (1997): Trust, emotion, sec, politics, and science: Surveying the risk-assessment battlefield. In B. M., D. Messick, A. Tenbrunsel, and K. Wade-Benzoni (Eds.), *Psychological Perspectives to Environmental and Ethical Issues in Management*, 277–313. San Francisco, CA.: Jossey-Bass.

Starmer, C. (2000): Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38(2), 332–382.

Starr-McCluer, M. (1996): Health insurance and precautionary savings. *American Economic Review*, 86(1), 285–295.

Stocké, V. (2001): Socially desirable response behavior as rational choice: The case of attitudes towards foreigners. Discussion Paper No. 01-24, Sonderforschungsbereich 504, Universität Mannheim.

Sudman, S., N. M. Bradburn, and N. Schwarz (1996): *Thinking About Answers*. San Francisco, CA: Jossey-Bass.

Thaler, R. H. (1990): Savings, fungibility, and mental accounts. *Journal of Economic Perspectives*, 4(1), 193–205.

Thaler, R. H. and H. M. Shefrin (1981): An economic theory of self control. *Journal of Political Economy*, 89(2), 392–406.

Tourangeau, R. (1984): Cognitive sciences and survey methods. In T. Jabine, E. Loftus, M. Straf, J. Tanur, and R. Tourangeau (Eds.), *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*, 73–100. Washington, DC: National Academic Press.

Tourangeau, R., L. J. Rips, and K. Rasinski (2000): *The Psychology of Survey Response*. New York, NY and Cambridge, UK.

Tourangeau, R. and T. W. Smith (1996): Asking sensitive questions: The impact of data collection mode, question format, and question context. 60, 275–304.

Tversky, A. and D. Kahneman (1992): Advances in prospect theory: Cumulatitive representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.

Van Soest, A. and M. D. Hurd (2004): Models for anchoring and acquiescence bias in consumption data. NBER Working Paper No. 10461.

Von der Heyde, C. (1994): Gewichtung am Beispiel: Einwohnermeldeamt versus Random Route. In S. Gabler, J. H. P. Hoffmeyer-Zlotnik, and D. Krebs (Eds.), *Gewichtung in der Umfragepraxis*, 141–151. Opladen: Wesdeutscher Verlag.

Von Gaudecker, H.-M. (2004): Regionale Mortalitätsunterschiede in Baden-Württemberg. MEA Discussion Paper 46-04.

Wakker, P. P. and D. Deneffe (1996): Eliciting Von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Science*, 42, 1131–1150.

Walliser, J. and J. Winter (1999): Tax incentives, bequest motives and the demand for life insurance: Evidence from Germany. Discussion paper No. 99-28, Sonderforschungsbereich 504, Universität Mannheim.

Wansbeek, T. and E. Meijer (2000): *Measurement Error and Latent Variables in Econometrics*. Amsterdam: Elsevier.

Weber, E. U., A.-R. Blais, and N. E. Betz (2002): A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, 15(4), 263–290.

Wilke, C. B. (2004): Ein Simulationsmodell des Rentenversicherungssystems: Konzeption und ausgewählte Anwendungen von MEA-PENSIM. MEA Discussion Paper 48-04.

Winter, J. (2002a): Bracketing effects in categorized survey questions and the measurement of economic quantities. Discussion Paper No. 02-34, Sonderforschungsbereich 504, Universität Mannheim.

Winter, J. (2002b): Understanding response behavior in surveys of the elderly: A natural experiment on the role of response uncertainty. Unpublished manuscript, University of Mannheim.

Winter, J. (2004): Response bias in survey-based measure of household consumption. *Economics Bulletin*, 3(9), 1–12.

Wooldridge, J. M. (2001a): Asymptotic properties of weighted M-estimators for standard stratified samples. *Econometric Theory*, 17, 451–470.

Wooldridge, J. M. (2001b): *Econometric Analysis of Cross Section and Panel Data*. Camebridge, MA: MIT Press.

Zaller, J. (1992): *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.

Zeldes, S. P. (1989): Optimal consumption with stochastic income: Deviations from certainty equivalence. *Quarterly Journal of Economics*, 104(2), 275–298.

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass ich die Dissertation selbständig angefertigt und mich anderer als der in ihr angegeben Hilfsmittel nicht bedient habe, insbesondere, dass aus anderen Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche gekennzeichnet und mit Quellenangaben versehen sind, nicht stattgefunden haben.

Mannheim, 10. April 2005 Lothar Essig