Daniel Schunk

Essays on Individual and Household Decision-Making: Experimental and Survey Evidence

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

vorgelegt im Sommersemester 2006

Referent:	Prof. Axel Börsch-Supan, Ph.D.
Korreferent:	Prof. Jörg Oechssler, Ph.D.
Dekan:	Prof. Konrad Stahl, Ph.D.
Tag der mündlichen Prüfung:	20.07.2006

Acknowledgments

I am indebted to many people for their invaluable help when writing this thesis. First, I would like to thank my supervisor Axel Börsch-Supan for his encouragement and support. He provided me with a truly excellent work environment and he boosted and challenged my interest in and understanding of research in economics. Joachim Winter also had a significant impact on my thinking and writing in economics, I owe him great intellectual debt. Similarly, I am indebted to Daniel Houser who taught me a lot about economics and with whom I had numerous intriguing discussions, in particular during my three months at George Mason University. I would like to thank Cornelia Betsch, the co-author of the first paper in this dissertation, from whom I learned much about psychology, as well as Arthur Kennickell and Arthur van Soest, who are a source of advice and support. I would also like to thank Jörg Oechssler for agreeing to be my second supervisor.

Without the stimulating research environment at the Mannheim Research Institute for the Economics of Aging (MEA), the Sonderforschungsbereich 504 (SFB 504), and the Center for Doctoral Studies in Economics and Management (CDSEM) at the University of Mannheim, I would not have gotten many ideas and opportunities, and – coming from a different scientific background – I would not have learned about economics so quickly and thoroughly. I am grateful to all my colleagues and friends at the MEA, the SFB 504, and the CDSEM. All of them – explicitly including Brunhild Griesbach, Isabella Nohe, and Petra Worms-Lickteig – contribute to the unique atmosphere here and make me enjoy my work a lot!

While working on this dissertation, I benefited greatly from discussions with, and from comments on my papers by Jonathan Baron, Barbara Berkel, Siegfried Berninghaus, Axel Börsch-Supan, Klaus Fiedler, Hans-Martin von Gaudecker, Uri Gneezy, Michael Haliassos, Florian Heiss, Daniel Houser, Michael Hurd, Hendrik Jürges, Arthur Kennickell, Alexander Ludwig, Jürgen Maurer, Anette Reil-Held, Susanne Rohwedder, Paul Ruud, Arthur van Soest, Mathias Sommer, Robert Sugden, Martin Weber, Matthias Weiss, Joachim Winter, and from feedback during seminar presentations at the Federal Reserve Board of Governers, Washington D.C.; George Mason University; University of Heidelberg; University of Mannheim; University of Munich; University College London; London School of Economics; and Tilburg University. Furthermore, I am very indebted to the student assistants that have been working with me. All of them have provided excellent work.

I gratefully acknowledge financial support from the Land Baden-Württemberg, the Deutsche Forschungsgemeinschaft via the SFB 504 and the CDSEM at the University of Mannheim, from the European Union (RTN project "Economics of Ageing in Europe"), and from the Gesamtverband der Deutschen Versicherungswirtschaft (GDV).

Last but not least, I want to affectionately thank my wife, Kirsten, and my family who are an endless source of emotional support and intellectual stimulus.

Introduction

> *„Human behaviour may be governed by rules, but it is possible that these rules simply encode preferences. […] Many psychologists argue that behaviour is far too sensitive to context and affect to be usefully related to stable preferences. However, if there are underlying preferences, then even if the link from preferences to rules is quite noisy it may be possible to recover these preferences and use them to correctly evaluate economic policies, at least as an approximation that is good enough for government policy work."*
>
> *Daniel L. McFadden. Nobel Prize Lecture 2002.*

Economics has always been concerned with the investigation of the decisions and the motivations of agents. The standard theory of decision-making under risk – Expected Utility Theory – was first proposed by Daniel Bernoulli (1738) and axiomatized by John von Neumann and Oskar Morgenstern (1947). Since then, this prescriptive and axiomatic framework has been the key building block in micro- and macroeconomic theory.

Over the last decades, the social sciences have developed numerous alternative axiomatizations and descriptive theories of choice under risk. The development of these theories has often been inspired by behavioral information in experimental and survey data. Overall, the scientific investigation of the decisions and motivations of human beings has been characterized by a rich interaction between theory and empirical evidence and between various disciplines of the social sciences, mainly economics and psychology. Despite or maybe because of all the ongoing discussions between economists and psychologists as well as between theoretical and empirical researchers, the development of knowledge about human decision behavior and the application of this knowledge to important problems in the real world has certainly been a success story of an intra- and interdisciplinary relationship. While the huge collection of findings is beginning to be organized under broader generalized frameworks and principles, many methodological – some would say: epistemological – differences between economics and psychology remain and warrant continuing, passionate, and fruitful interaction between the two disciplines.

Economists are generally more interested in a single theoretical approach that might be applied in various decision contexts, and they emphasize that an underlying theory is needed to interpret data in any structural way. Psychologists, on the other hand, tend to be more concerned with developing local theories that are designed for specific contexts, that are connected to a specific observed phenomenon, and that stem from less structural and more descriptive analyses of the underlying data. Rather than mapping information input to decisions based on a model that is built on a set of axiomatized behavioral assumptions – such as preference maximization – psychologists tend to focus on investigating the nature of individual traits and behavioral processes that are relevant given a specific decision context. Accordingly, they have developed a rich terminology that describes various factors and phenomena that are related to individual decision behavior. Reflecting many psychologists' opinion that existing knowledge about human behavior can hardly be codified into one parsimonious model, their terminology involves terms such as "*attitudes*", "*motives*", and "*modes*". While these terms denote individual-specific characteristics that are related to human behavior, they are not clearly defined based on an underlying behavioral model – in contrast to, e.g., the term "*preference*", which is central in economics. Overall, although methodological approaches and specific research priorities differ between the disciplines, both disciplines are united by their interest in gaining a better understanding of human behavior, and this involves an interest in exploring whether there exist individual characteristics that are relevant for decision behavior.

The existence of underlying individual characteristics or traits is an important question for social scientists: Discerning individual characteristics that are related to the enormous heterogeneity that is observable in human decision behavior eventually helps to build better models of human behavior and, thereby, to evaluate policies. Identifying these characteristics empirically is a challenging intellectual task. Ideally, the scientist would have a structural model that maps information input to decision behavior of an agent, given all relevant individual characteristics of the agent. This model would capture explicitly the underlying cognitive or mental process – e.g., preference maximization in a classical economic model. To test the descriptive validity of her model, the social scientist would search for situations and settings that are informative with regard to the model, and she wishes to observe controlled comparisons of different treatments in a natural environment, in which one variable X varies exogenously. Finally, she wants to obtain the complete and exact information required by her theory, such that she is able to test derived hypotheses and to refine her theories or develop alternative theories, if necessary.

Unfortunately, such natural experimental settings occur only very rarely. In most cases, finding the underlying determinants of human decision behavior is a difficult enterprise and the scientist inescapably faces several constraints. First, most empirical studies on human activity – regardless of whether they are based on experimental data or on survey data – have to deal with at least some degree of measurement error or incomplete information. This involves problems such as nonresponse to survey questions and uncertainty about answers on the part of the respondent. Second and related to the first issue, for many parameters of the underlying model, it is not clear how to measure them appropriately or they might not have a well-defined quantitative meaning. This issue is particularly apparent for psychometric measures, such as measures of individual attitudes. Furthermore, there are deeper problems involved: Assume that the researcher investigates a model that deals with stock market decisions and includes risk preferences as an individual-specific characteristic. If information on individual risk attitude is elicited using a series of lottery questions, then this measure of risk attitude is likely to be endogenous, since similar underlying and unobserved factors might drive stock market decisions and lottery decisions; briefly, in many cases it is not clear whether measured variation is really exogenous in a given context. Third, in real-world settings, many confounding and hardly controllable or measurable effects might have an impact on observed decision behavior. While controlled environments, such as laboratory environments, might be a remedy, they have their own limitations. For example, they are deliberately stylized and the sampled populations are far from being representative. Respecting all these constraints, the empirical scientist must evaluate which method from the universe of data collection and measurement methods is appropriate and practically feasible, given her interest in a certain phenomenon.

This dissertation presents a collection of papers which are inspired by my interest in the phenomenon that economic agents are heterogeneous in their decisions and which are concerned with the investigation of the relationship between individual-specific characteristics – such as preferences, modes, and motives – and economic decisions. All papers in this dissertation are united by the objective to obtain inference about determinants of economic behavior in empirical studies. I investigate behavior in distinct contexts, and I use data which have been collected in different ways: in a controlled laboratory setting and in a nationwide household survey. The first study, co-authored with Cornelia Betsch (Schunk and Betsch, 2006), combines elements of economics and psychology and investigates individual decision behavior in a static decision context. We use data from a laboratory experiment to analyze the link between psychometric measures of individual decision modes and individual utility functions, the backbone of modeling decision-making in the economic sciences. The second paper explores individual decision behavior in a dynamic decision context. I develop utility-based models of sequential decision behavior and design a laboratory experiment to investigate the relationship between individual preferences and behavior in the sequential decision tasks. The third paper is concerned with the saving decisions of German households. I use data from a representative socio-economic survey – the SAVE survey on household saving behavior – and analyze to what extent four co-existing saving motives are related to the saving decisions of private German households. While surveys are an important source of data for the study of household behavior, incomplete information, such as item and unit nonresponse, is a data problem that is particularly prevalent in surveys dealing with sensitive financial issues. Therefore, the last paper of this dissertation is concerned with designing an iterative statistical algorithm for dealing with the problem of incomplete information that results from item nonresponse to questions in the German SAVE survey. I find that there is heterogeneity in nonresponse behavior across different questions and that – in line with existing findings in cognitive psychology – item nonresponse is not occurring randomly but is related to the individual-specific characteristics of the survey participants. In a sense, this finding in the last paper – obtained using an involved statistical methodology – leads back to the first paper of this dissertation, since it underlines the importance of interdisciplinary collaboration: Cognitive psychologists have investigated the phenomenon of item nonresponse and have developed models for the processes that lead to nonresponse behavior. In a collaborative effort, these findings can help to understand and model the mechanisms of survey nonresponse better, and – a complementary approach to correcting data problems ex post based on statistical methods – they can help to address the problem of item nonresponse by improving the data collection methodology.

While all four papers presented in this dissertation are related to the empirical study of economic decision behavior, they touch different disciplines. Each of the papers is a self-contained study which can be read independently. In the remainder of this introduction, I describe the content and the results of each paper in more detail.

The first paper is motivated by the observation that in many economic decision situations people differ systematically in their preference for an intuitive or a deliberative decision mode – here I talk about preferences in the psychological sense of the word, i.e. this term denotes a subjective comparative judgment in a general sense, and is not used in the narrow sense that is connected to the economic concept of utility functions. In a laboratory setting we elicit subjects' utility functions, using a lottery-based elicitation method, and we obtain a psychometric measure of the relative preference for an intuitive and a deliberative decision mode. We then analyze the relationship between the curvature of the utility function and the psychometrically measured preference for an intuitive and a deliberative decision mode. The findings indicate that the curvature of the individual utility function is systematically related to the decision mode of the subjects: People who prefer the deliberative mode generally have a utility function that is more linear than the utility function of people who prefer the intuitive mode. We suggest the explanation that intuitive people's decisions in the lottery context mirror a feeling of risk and lead to behavior which is not risk neutral. They might have additionally integrated affective reactions towards the stimuli into the decision, influencing their decision towards the affective reaction. Deliberate decision makers seem to perform time-consuming cognitive operations – apparently not just calculation – leading to more risk neutral decisions and a more linear utility function. Overall, the findings suggest that individually stable traits, measured using a psychological questionnaire, might help explain people's decisions over monetary values; that is, they might be informative for understanding observed economic behavior, such as finance and insurance decisions.

The starting point of the second paper is the observation that individuals are heterogeneous with respect to their behavior in simple dynamic choice situations. In particular, their behavior does not correspond to the predictions of an optimal decision rule that is derived under the assumption of risk neutrality. A very elementary representative of a dynamic choice situation is a so-called search task. Search tasks are attractive for the experimental study of dynamic choice behavior: first, because of their simple sequential decision structure and, second, because this decision structure masks a complicated optimization problem that – comparable to sequential decision situations in our everyday lives – cannot be solved without a computer. Economic theory suggests that the heterogeneity observed in search behavior is reflected in the heterogeneity of individual preferences. Do individual preferences, which can be revealed using a simple preference elicitation mechanism, inform us on behavior in search situations? The paper first develops models for search behavior under the assumption of expected utility maximization and under the assumption of sequential updating of utility reference points during the decision task. Then, data from an economic laboratory experiment are used to investigate the link between individual preferences and search behavior. I find experimental evidence that supports the new reference point model: Though subjects could not make losses during the experiment, individual loss aversion is systematically related to the observed decision behavior in a way that is consistent with the predictions of a model that involves utility reference point updating. Measures of risk attitude, however, are not related to observed behavior. The finding that many people set reference points in sequential decision tasks has been obtained in a controlled laboratory setting and it is of primary interest for decision theory. It might also be of more general interest in, e.g., consumer economics, labor economics, and finance, and it might serve as a guide to theoretical and structural econometric specifications in applied search theory that explicitly allow for individual heterogeneity.

The third paper uses survey data from a random cross-section of German households to investigate household saving decisions. There are many different motives for saving a portion of one's income; these motives co-exist over the life-cycle and their relative importance changes. Existing research further emphasizes an enormous heterogeneity with respect to the household saving rate and the extent to which households plan their saving. The paper is concerned with linking heterogeneity in household saving behavior to four co-existing saving motives. First, the paper shows that the importance that households attach to the saving motives is related to how much households save at different life stages. The estimated effects are appropriate given the different stages of the households' life-cycle and they are broadly in line with existing findings in theoretical and empirical studies of life-cycle saving that focus on only one specific saving motive. Second, I classify the saver type of the households based on whether they engage in regular savings plans, or rather save irregularly and without a savings plan and I find that saving motives are related to the saver type of the household. Overall, the results indicate that heterogeneity with respect to the saving rate and the saver type is systematically related to the importance that households attach to different saving motives. This suggests that policy reforms that substantially change the importance of certain saving motives in the eyes of private households might alter household saving behavior in various ways.

As the third paper of this dissertation demonstrates, important empirical information on household behavior can be obtained from surveys. However, many interdependent factors that can only be controlled to a limited extent – such as privacy concerns, respondent uncertainty, cognitive burden of the question, and survey context – often lead to unit and item nonresponse. Missing data on certain items is a frequent source of difficulties in statistical practice, and it generally leads to biased inference. Therefore, the fourth paper is concerned with an imputation method for missing data. The purpose of the paper is to present and discuss the theoretical underpinnings and the practical application of an iterative multiple imputation method that has been developed for the German SAVE dataset. The developed algorithm essentially simulates the distribution of missing data and yields complete datasets that can be analyzed without discarding any observed values and that incorporate the uncertainty about which values to impute. The paper discusses properties of the iterative imputation algorithm, investigates the distribution of imputed values, and compares these findings with results from other imputation approaches. I find that there is heterogeneity in nonresponse behavior across different questions and that – in line with existing findings in cognitive psychology – item nonresponse is not occurring randomly but is related to covariates included in the imputation models.

References

Bernoulli, D. (1954): Exposition of a New Theory on the Measurement of Risk. (original 1738) *Econometrica*, 22, 23-36.

Neumann, J. v. and O. Morgenstern (1947): The Theory of Games and Economic Behaviour. 2nd edition. Princeton: Princeton University Press.

Schunk, D. and C. Betsch (2006): Explaining heterogeneity in utility functions by individual differences in decision modes. *Journal of Economic Psychology*, forthcoming.

Explaining heterogeneity in utility functions by individual differences in decision modes[1]

Abstract: The curvature of utility functions varies between people. We suggest that there is a relationship between preferred decision modes (intuition vs. deliberation) and the curvature of the individual utility function. In this study the utility functions of the subjects were assessed using a lottery-based elicitation method and related to the relative preference for intuition vs. deliberation. We found that people who prefer the deliberative mode have a utility function that is more linear than the utility function of people who prefer the intuitive mode. We suggest that intuitive people's decisions mirror a feeling of risk and lead to behavior which is not risk neutral. They may have additionally integrated affective reactions towards the stimuli into the decision influencing their decision towards the affective reaction. Deliberate decision makers seem to perform time consuming cognitive operations (apparently not just calculation) leading to more risk neutral decisions and a more linear utility function.

[1] This paper is joint work with Cornelia Betsch. It is forthcoming in the Journal of Economic Psychology in 2006.

1 Introduction

For decades economists and psychologists have investigated the relationship between stimuli and the perception and processing of stimuli. In psychophysics, for example, the relation between stimulus intensity (e.g., weight) and the related sensation (e.g., the perception of heaviness) is described in Fechner's law (Fechner, 1860). In the decision-making literature, prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992) describes the relation between varying amounts of money and its perceived utility. As a common denominator of this "psychophysical numbing" (Fetherstonhaugh et al., 1997, p. 297) we find curved, non-linear relationships between the variation of a stimulus and the subjective feelings towards the stimulus variation for most people. Value functions are typically concave (i.e., constant increments of scope yield successively smaller increments of value) and inversely s-shaped, which is usually interpreted as risk averse decision behavior when gambling for monetary gains and as risk seeking behavior in gambles with loss outcomes (Abdellaoui, 2000; Gonzalez and Wu, 1999; Tversky and Fox, 1995). However, the curvature of the value function varies as the subjective perception of the stimuli also varies between people. For example, Fetherstonhaugh and colleagues (1997) found that not all subjects had curved utility functions. "People […] exhibit diminished sensitivity in valuing lifesaving interventions against a background of increasing numbers of life at risk. […] Although psychophysical numbing was present in each study, its prevalence varied." (p. 283, 297). Considering individual differences could help explain why some people value saving 4,500 people independent of the number of threatened people (e.g., 11,000 or 250,000), and why others show dramatic differences.

The goal of our study is to link psychometric measures with individual utility functions, the backbone of all research on individual decision-making models in the economic sciences. We suggest that individually stable traits, measured based on a psychological questionnaire, might help explain observed economic behavior, such as finance and insurance decisions.

It is an established method to compare people's choices between risky monetary gambles to assess their utility function.[2] The gambles used are of the kind "win $\$x$ with

[2] Research in psychology emphasizes that the attitude towards risk cannot solely be captured by the curvature of the utility function. In most existing economic models, a person's attitude towards risk is exclusively captured by the curvature, i.e., the shape, of the person's utility function. We want to stress that

the probability of p vs. win \$$y$ with the probability of $(1 - p)$". Linearity of the utility function means that the utility of a risky monetary lottery is determined by the multiplication of the stated monetary value and its probability. If people place subjective values on the stated monetary outcomes (given the probability is held constant to exclude effects of probability weighting), the utility function becomes curved (i.e., it deviates from linearity). For example, your subjective feeling of the utility of \$5 depends on the reference total amount of money. It makes an affective difference to save \$5 when you buy a \$10 bottle of wine or when you buy a VCR for \$400. The increment of utility for the same amount of money is smaller as the scope increases. The stronger the influence of subjective values, the more the decision can deviate from the decision of an expected value maximizer.

In addition to the automatic subjective valuation by feelings, humans are able to use (meta)cognition, a deliberative, conscious reflection of the problem at hand. Previous research has shown that priming participants to use cognitive strategies makes the effect of subjective feelings disappear. For example, Bless et al. (1998) showed that the well-known framing effects in the gain and loss domain (Tversky and Kahneman, 1981) disappear when the problem is subtly framed as a statistical problem. Metacognitive comprehension, a deliberative mode of thinking, can overcome the automatic subjective feelings and would lead to a utility function that is not curved but instead approaches linearity. Hsee and Rottenstreich (2004) suggest that the value function differs depending on the decision mode. They label the two opposed modes as "valuation by calculation" vs. "valuation by feelings". They suggest "that concavity arises in part because most real-world valuations mix calculation and feeling. […] In such mixes, greater reliance on feeling yields greater concavity." (p. 28).

Although most real-world valuations might indeed mix deliberative and intuitive strategies, there is strong evidence that individuals differ in the way they *habitually* use the affective-intuitive or deliberative decision mode (e.g., Langan-Fox and Shirley, 2003). People with a preference for intuition base most of their decisions on affect, resulting in fast, spontaneous decisions, whereas people with a preference for deliberation tend to make slower, elaborated, and cognition-based decisions (Betsch, 2004). Intuitive processing means following instant, effortless evaluation processes (Hogarth, 2001)

in this paper, we are not concerned with discussing the distinction between different notions of risk aversion in detail. How broad one defines the concept of risk aversion is essentially related to the model one has in mind.

involving automatic, *affective (good vs. bad) reactions*. Various models capture the intuitive mode as a complementary concept to a *deliberative, effortful, planned and analytic* way of making decisions (e.g., Chaiken, 1980; Epstein, 1983; Hogarth, 2001). Intuitive people ask themselves, "How do I *feel* about it?", while deliberative people ask, "How do I *think* about it?" (for differences regarding this question format, see Verplanken et al., 1998).

Insights into the relationship between preferred decision modes and utility functions might be of particular relevance for understanding portfolio choice and stock market decisions. The question of whether or not there are stable individual differences in reasoning or decision-making competence has recently gained interest (see Parker and Fischhoff, 2005; Stanovich and West, 1998; 2000), for example in the context of investor overconfidence models (Glaser et al., 2004; Glaser and Weber, 2005).

We argue that the subjective assessment of intuitive people should be more influenced by affective reactions than the subjective assessment of deliberative people. According to the "risk as feelings" hypothesis (Loewenstein et al., 2001), probabilities and outcomes can directly evoke affect and impact behavior without cognitive mediation. We suggest that intuitive people use this feeling of risk to make their decisions. These decisions should mirror the feeling of risk and should lead to behavior which is *not* risk neutral. Thus, for intuitive people the utility function should be curved and not linear.

The subjective values of deliberatives (i.e., deliberative people) should correspond more closely to the stated monetary values presented. Although they might also have a sudden feeling of risk, their decision might be cognitively mediated and be a result of enhanced cognitive processing. Emotion leads to diminished sensitivity because the emotional response is relatively insensitive to quantity (or scope), once some change has been registered (Hsee and Rottenstreich, 2004). When people deliberate, however, they should pay more attention to quantity.

The individual preference for intuition and deliberation should therefore be related to the shape of people's value function. Concretely, we claim that the monetary utility function of people with a preference for intuition should reflect affect-based decision making and be curved (i.e., deviate from linearity). Conversely, the utility function of people with a preference for deliberative decision-making should be more linear than the one of non-deliberative decision makers.

2 Method

2.1 Overview

This hypothesis was tested in a lottery-based study that is presented in this paper. First, we assessed subjects' utility functions, based on a sequence of individually-adapted lottery questions in which the lottery probabilities were kept equal to avoid the potentially perturbing effect of individually-different probability weighting. Then, subjects filled in an inventory assessing their Preference for Intuition and Deliberation (PID, Betsch, 2004). Based on the lottery choices, we were able to estimate an index for the curvature of the utility function that we related to the individual preference for deliberation and intuition.

2.2 Subjects

A total of 200 students from the University of Mannheim participated in groups of at most 17 participants per session. The sample was obtained in two separate blocks (Sample 1 = 68 subjects, Sample 2 = 132); the procedure differed only minimally (see Procedure).

2.3 Procedure

Upon entering the lab, subjects were seated individually in front of a PC. In both samples the subjects were told that they would have to make many decisions regarding lotteries with two alternatives. The two lotteries (A and B) were presented simultaneously on the computer screen. Subjects were instructed to indicate their choice by clicking on the respective button for lottery A or B. After a selection was made, the next lottery appeared on the screen. Subjects were not constrained by time and answered all lottery questions at their leisure.

At the end of the procedure, the first sample answered the PID questionnaire by clicking on one of five radio buttons indicating their agreement with the statements. The second sample took part in 3 more unrelated studies before they answered the PID inventory, which was identical to the first sample. This order was chosen in attempt to prevent an influence of the value function elicitation procedure on the PID values. The time elapsed between the value function elicitation and the PID inventory was approximately 45 minutes. After the procedure, subjects from both samples were thanked, debriefed, and dismissed.

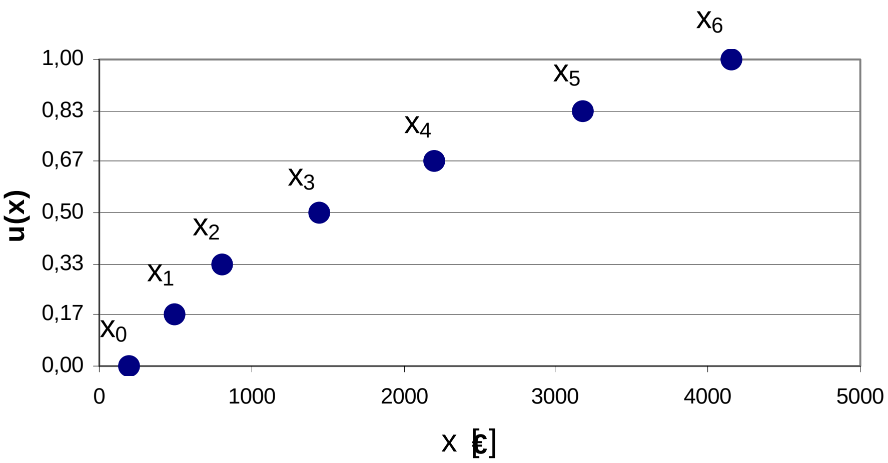
In order to provide incentives and to enhance motivation, one of the subjects in each session in the first sample was randomly selected to play for a real monetary pay-off based on his or her choices made in one of the lottery tasks. Since the outcomes of the lotteries were up to €6000, we informed the subjects that the randomly selected person played for 1% of the positive outcomes (i.e., the gains) presented in the lotteries. We dropped this procedure in the second sample and found no change in results. In the next section we describe the materials in more detail.

2.4 Materials

2.4.1 Value Function Elicitation

A value function assigns a subjective value, or utility, to a stated (objective) value. To approximate such a function, it is necessary to elicit a number of points of this function for every individual (for an illustration cf. Figure 1).

Figure 1: Utility function for gains for individual 1. The x_i are equally spaced in terms of their utility. This allows for the assessment of the curvature of the value function.

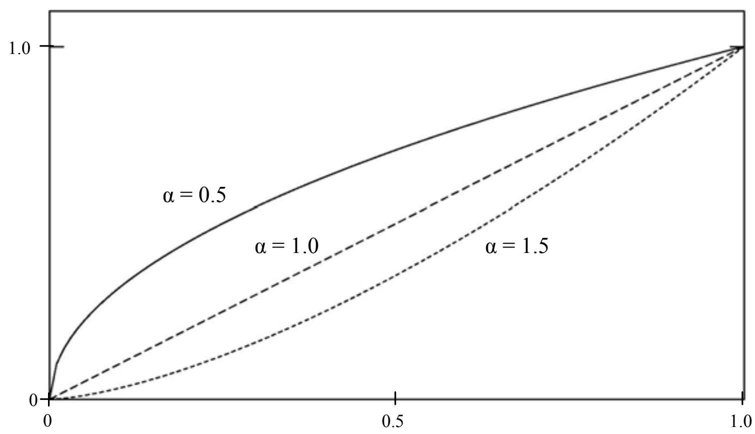


Various methods exist to construct individual value functions (i.e., to assess these points) from observed decisions in a series of monetary gambles (Farquhar, 1984). Our elicitation mechanism is based on a method proposed by Abdellaoui (2000) in which seven points are elicited separately for both the gain and loss domains {x_0 to x_6}. To elicit one single point x_i, subjects are required to make five decisions between lotteries. The lottery outcomes are adapted based on the prior decision of the subjects, in order to determine (after five iterations) an outcome x_i for which the subject is indifferent between the two lotteries, A and B. This indifference is achieved as follows: If the subject prefers lottery B to lottery A, then the value of x_i in lottery B is decreased such that lottery B is less attractive. Conversely, if the subject prefers lottery A to lottery B, then the value of x_i is increased such that lottery B becomes more attractive. These steps are repeated five times for all elicited points x_i. Based on the x values and the assumption of a utility function of a power form, it is possible to estimate two parameters, alpha (α) and beta (β). Alpha describes the utility function in the gain domain, and beta describes the function in the loss domain. Appendix A gives a detailed description of the method and calculation of α and β.

Alpha and beta characterize the risk attitude of the individuals in the sense of a measure of proportional risk attitude (Eisenführ and Weber, 2003). Standard nonlinear least squares regression is used to estimate α and β for gains and losses. A value of α and β equal to 1 denotes a linear utility function on gains and losses, respectively. If α is larger than 1, the utility function is convex and the individual is risk seeking for gains, if α is smaller than 1, the individual is risk averse for gains, since the utility function is concave (for β, vice versa).

We use the absolute difference between the risk parameters, α and β, and 1 as a *measure for the curvature of the utility function*; the higher the value is, the more the utility function is curved (i.e., the more it deviates from a linear function; see Figure 2). Therefore, we define $a = | 1-\alpha |$ and $b = | 1-\beta |$ as indices for curvature (i.e., for the deviation of the particular utility functions from a linear function).

Figure 2: The utility function for gains for various values of α. The absolute difference between the parameter α and 1 is a measure for the curvature of the utility function.



2.4.2 Individual Preference for Intuition and Deliberation (PID)

To assess preferences in making decisions intuitively or deliberatively, we use the Preference for Intuition and Deliberation scale (PID; Betsch, 2004). The measurement consists of 18 questions: nine items assessing the habitual preference for deliberation (PID-D) and nine items assessing the preference for intuition (PID-I). On a 5-point scale anchored at 1 (“I don’t agree.”) and 5 (“I totally agree.”), subjects answered questions regarding their decision-making habits. PID-D consists of items such as ”*I prefer making detailed plans to leaving things to chance*” or ”*I think before I act*.” PID-I includes items such as, “*With most decisions it makes sense to rely on your feelings*” or *”I carefully observe my deepest feelings”* (the complete PID inventory is included in Appendix B). In prior studies (total N > 2500; Betsch, 2004) the scale proved as reliable (Cronbach’s α for PID-D varied between 0.78 and 0.84, for PID-I between 0.78 and 0.81), and the 2-dimensional structure was confirmed. The inventory captures a habitual preference that is stable over time. A preference for a decision mode influences decision-making especially in unconstrained situations (e.g., no time pressure, enough resources, etc.).

People with high scores on deliberation have been shown to be conscientious perfectionists with a high need for structure (Betsch, 2004, Study 3). They aim at maximizing rather than satisficing their decision outcome. On the other hand, highly intuitive people are speedy decision-makers and tend to score high on social and emotion-bound personality dimensions like extraversion, agreeableness, and openness for experience.

3 Results

We tested the equality of means and variances between the samples for the parameters describing the utility functions (α and β), and for PID-I and PID-D. As the hypotheses of equality could not be rejected (all *F*- and *t-values* < 1.2), the data of the two samples were combined. From the total of 200 subjects, 15 were found to be outliers in terms of the standard error and were therefore deleted.[3]

For data analysis, we first calculated correlations between the curvature indices and the PID values. In line with previous findings (e.g., Betsch, 2004), the subjects in general had a significantly greater preference for deliberation (PID-D = 3.7, sd = 0.6) than preference for intuition (PID-I = 3.3, sd = 0.6), t (185) = -4.9, $p < 0.001$.

[3] The 15 excluded subjects were outliers in terms of the standard error of the coefficient estimates of the utility function: We excluded all subjects whose standard error of one of the coefficient estimates was more than one standard deviation larger than all other standard errors of coefficient estimates. High standard errors indicate unsystematic clicking, suggesting a lack of motivation. It is interesting to note that some of the deleted subjects were not only outliers in terms of the standard error of their coefficient estimates but also in terms of the time needed for the completion of the lottery questions: They needed considerably less time than all other subjects. There was no systematic relation between preference for intuition and deliberation and the occurrence of outliers.

The median of the coefficient estimates of the power function on gains (α) was 0.91, with a mean standard error (*se*) of the nonlinear least squares estimation of 0.06 ($M\alpha = 0.99$, $sd = 0.44$). In the loss domain, the median β equaled 0.90 ($se = 0.05$; $M\beta = 0.95$, $sd = 0.38$).[4] The coefficients of determination of the nonlinear regression approach 1 (the mean R^2 is 0.995 for α and 0.995 for β). In total, the results regarding subjects' risk attitudes are consistent with the predictions of prospect theory (Tversky and Kahneman, 1992) and subsequent work based on prospect theory.

3.1 Relationship between Preference for Intuition and Deliberation (PID) and the Curvature of the Utility Function

We hypothesized that high values of deliberation (PID-D) should coincide with a less curved utility function. Conversely, subjects with a greater degree of intuition (PID-I) should have more curved utility functions.

Based on this hypothesis, we expected that both curvature indices, $a = | 1-\alpha |$ and $b = | 1-\beta |$, would be positively correlated with a preference for intuition and negatively correlated with a preference for deliberation. This was supported by our data: A high preference for deliberation was found to be negatively and significantly related to the curvature of the utility function in the gain domain, r_a (Pearson's correlation coefficient) = -.20, $p < 0.01$, and in the loss domain, $r_b = -0.15$, $p < 0.05$. Similarly, a high preference for intuition was significantly positively correlated with the curvature index on both the gain ($r = 0.18$, $p < 0.05$) and the loss domain ($r = 0.22$, $p < 0.01$). Thus, more deliberative decision-makers had less curved, or more linear, utility functions, while more intuitive decision makers had more curved, or less linear, utility functions. This hypothesis found further support in an overall test.[5] Though the intuition and deliberation dimensions of the PID were not highly negatively correlated ($r = -.36$, $p < 0.001$), we defined $c = PID\text{-}I - PID\text{-}D$ as an overall measure for the preference for intuition ($Mc = -.37$, $sd = 1.0$). Higher values of c indicate a higher preference for intuition. Our hypothesis that c is positively correlated with the curvature indices a and b was strongly supported by the data (see Table 1, rows 1 and 2). Additionally, we performed a regression analysis to test whether the observed relationship deviates significantly from linearity, i.e., whether it is driven by extreme groups, the very intuitive and the very deliberative subjects. This is not the case, c

[4] Abdellaoui (2000) found 0.89 and 0.92 for the sample median of α and β, respectively, based on a study with 40 subjects in total.

[5] We are grateful to a reviewer suggesting this test to us.

is a significant predictor ($p < .01$) of both, a and b, but higher order terms of c, (i.e. c^2, $c^3, \ldots$) are not significant predictors.

Table 1: Overall test. Correlation between overall measure of preference for intuition (c) and the curvature indices (a, b) of the utility function. The table also presents results for various sample partitions.

	c
a ($N = 185$)	0.23 **
b ($N = 185$)	0.23 **
a ($\alpha >= 1$: $N = 74$)	0.41 ***
a ($\alpha <= 1$: $N = 136$)	0.14 $^+$
b ($\beta >= 1$: $N = 80$)	0.34 **
b ($\beta <= 1$: $N = 131$)	0.15 $^+$
a ($\alpha \neq 1$: $N = 160$)	0.20 **
b ($\beta \neq 1$: $N = 159$)	0.19 **

*Note: A higher c denotes a higher preference for intuition. A higher a or b value is associated with a more curved utility function. Correlations flagged with a + are significant on the 0.10-level, * on the 0.05 level, ** on 0.01, and *** on 0.001. c = PID-I – PID-D.*

3.2 Partitioning the Sample

Do our correlation findings reflect a relationship between habitual preferences for a decision mode and the curvature of the utility function, or do they rather stem from a systematic relationship between specific risk attitudes and the habitual preference for a certain decision mode? To investigate the robustness of our statistical findings, in particular to see whether the results are only driven by specific subgroups of the sample, we subdivided the sample into various partitions.

Table 1 presents the results from the sample partitioned based on the curvature estimates of the utility functions (i.e., based on their risk attitude). All correlation results held for the subgroups. If risk seeking subjects on the gain domain (i.e., we excluded the subjects with $\alpha \geq 1$) or subjects that were risk averse on the loss domain (i.e., we excluded the subjects with $\beta \geq 1$) were excluded, the correlation results are only marginally significant.

The results from the subgroup analysis provide further evidence that our data do *not* suggest a systematic relationship between the preference for a decision mode and a certain *risk attitude*. However, as we have hypothesized, there is instead a systematic relationship between the preference for a decision mode and the *degree of curvature* of individual utility functions. In sum, our findings are neither driven by only one specific subgroup of the sample, nor by differences between the PID extreme groups in their mean curvature estimates.

The last two rows of the subgroup analysis in Table 1 address an important interpretative point of our analysis: They show that the results still hold even if we delete all subjects with linear utility functions from the sample, that is, if we exclude all subjects whose behavior follows an expected-value calculation. This suggests that our results cannot be explained by proposing that the less intuitive subjects are simply calculating. Still, it is likely that they perform more complex cognitive operations than intuitive people as can be seen by the analysis of decision times.

3.3 Decision Times

Decisions based on affect should be faster compared to deliberative decisions because affect is quickly accessible (cf. affective primacy hypothesis, Zajonc, 1980) and cognitive operations are time consuming. We have correlated the total decision times of every individual with the individual overall measure for the preference for intuition, c. The findings ($r = -0.18$, $p < 0.05$) support the hypothesis that the more intuitive a subject is the less time the subject takes for completing both lottery tasks.

Our approach to classifying the curvature of the individual utility function is based on the assumption of the power functional form (see Appendix A and Fig. 2). Do intuitive and deliberative subjects differ systematically in the way the power function fits the elicited points of their utility function? As a measure of fit, we used the standard errors, se_α and se_β, of the coefficient estimates of α and β and we correlated them with c. Linear utility functions are the only utility functions that were, by construction of our mechanism, be fitted with a zero standard error of the coefficient estimate. Due to this fact those corner outcomes are excluded from the correlation. We found that c was not correlated with se_α ($r = 0.07$, $N = 161$), but it was correlated with se_β ($r = 0.17$, $p < 0.05$, N = 160). We conducted a mediation analysis with the loss domain data and regressed b on c in a first regression, se on c in a second regression and b on se and c in a third regression (Baron and Kenny, 1986). In the third regression both predictors were significant, indicating that se partially mediates the effect of c on b. A Sobel test revealed that the indirect effect of c on the b via the mediator se_β is nearly significantly different from zero (regression coefficients for c in the first step: 0.08, se = 0.03, in the third step: 0.05, se = 0.03, $z = 1.94$, $p < .06$).

That is, on the gain domain the fit of the power utility functions was unrelated to the habitual decision mode of the individual. On the loss domain, though, the more intuitive a person was, the worse was the fit of the power function, i.e., more intuitive subjects had given answers to the lottery questions that were less consistent with our parametric functional assumption of a regular and smooth utility function. The resulting higher standard error served as a mediator for the relation between the degree of intuition (c) and the degree of curvature in the loss domain (b). A possible explanation is that more intuitive subjects might have answered the lottery questions on the loss domain (which were asked after the 32 questions on the gain domain had been asked) in a rather erratic way and have made decision errors on the loss domain. An outlier analysis revealed that the correlation between se_β and c was driven by 3 subjects, who had $se_\beta > 0.35$, which was almost 6 standard deviations above the mean standard error. Excluding the outliers leads to a correlation between c and se_β of $r = 0.10$, $p > 0.20$, thus a mediation analysis is not necessary any more.

4 Discussion

In this study we showed that the curvature of individual value functions, assessed with an established elicitation method, is correlated with the individual preference for intuitive and deliberative decision-making. The more people preferred deliberative strategies, the more linear their utility functions were. Conversely, the more intuitive a person was, the more curved the utility function was. The effect was stable for various partitions of the sample.

This effect might have occurred because intuitive and deliberative decision-makers used different sources of information. While intuitive decision makers might have used the instant affect produced by the risky alternatives, deliberatives may have used rather the stated values as presented by the experimenter. Intuitives “go beyond the information given” (Bruner, 1957, p.41) and bias their judgment with additional affective information, while deliberates seemingly bias their judgment less with subjective evaluations. Several findings support this assumption.

First, if intuitive subjects rely on quickly accessible affect, their reaction times should have been shorter compared to deliberative decision-makers who tend to reflect on their decisions. Indeed, this was the case in our sample and this is in line with findings from Betsch (2004): The time needed to finish the 64 lottery choices decreased, the more the decision maker preferred the intuitive over the deliberate decision mode. In Betsch's (2004) study, intuitive subjects indicated faster decision making than deliberative decision makers on a self-report scale. Furthermore, subjective evaluation happens automatically, but a meta-cognitive correction needs extra time, which might have caused the prolonged decision time for deliberative decision-makers. Second, as our analysis above shows, the faster decisions of intuitive decision makers were not generally a result of random clicking or a lack of motivation (though our results suggest that some very intuitive subjects might have answered the questions on the loss domain in a rather erratic way). Third, deliberative people tend to be maximizers of the objective expected values, which was demonstrated by the nearly linear shape of their utility functions. Again in line with findings by Betsch (2004), preference for deliberation (PID-D) correlated significantly with maximization (r = .27), a construct expressing the tendency to make optimal objective decisions as opposed to subjectively satisfying decisions (Schwartz et al., 2002). Maximizing is a highly cognitive process, involving conscious weighting, information search, for example, which requires more cognitive capacity than affective-intuitive, satisfying decisions.

Finally, in an unpublished pilot-study, we simply asked subjects after the utility elicitation procedure to what extent they relied on affect vs. calculation. Deliberatives reported that they calculated in 9% of the cases, whereas intuitive decision-makers reported that they calculated in only 5% of the cases. On the other hand, self-reports additionally showed that intuitive decision-makers (56%) used significantly more affect than deliberative decision makers (41%), the interaction effect was significant, $p < 0.05$. It seems unlikely that deliberatives actually "calculate" in the literal sense (also given the fact that the mean total time used for the 32 lottery decisions was max. 5 minutes). However, the self-report data on strategy use in addition to the decision time differences in this study indicate that deliberative decision makers did indeed perform more time-consuming cognitive operations.

As a limitation of this study we have to note that our explanation of the effect was not directly tested in this study. The basis of information used for the decisions was not manipulated. Johnson et al. (1988) found, for example, that the display of numbers (e.g., the probability of .9 as 9/10 or 513/570) elicits different strategies, namely calculation-based strategies vs. heuristic strategies. Such a method could be useful in future studies to further investigate the reported findings.

To summarize, in our study we found empirical evidence for the hypothesis that a habitual individual difference factor is able to account for the observed variation in the curvature of individual utility functions. This is another piece of evidence that both affective and deliberative processes play a role when people make decisions (cf. Loewenstein and O'Donoghue, 2004). On one hand, the findings in this study suggest that deliberative people use more cognitive strategies than intuitive people, and on the other hand, the data substantiates the speculation that the curvature of utility functions might come from affective evaluation and the integration of affect into the decision. This is especially the case for intuitive decision makers.

5 Conclusion

The degree of curvature of the utility function is interpreted as a measure of the risk attitude of a decision-maker. Psychologists claim that attitudes consist of affective, cognitive, and behavioral components (e.g., Breckler, 1984). One can argue that for intuitive subjects, the affective part of the attitude contributes more to the overall risk attitude compared to the cognitive part (vice versa for deliberative decision-makers). An explanation for our findings is that intuitive people use the affective risk information contained in the lotteries when making their decisions, which might lead to the risk attitude (i.e., a feeling of risk) becoming integrated in the judgment, resulting in risk averse or risk seeking behavior. Deliberative people, on the contrary, seem to base their decisions on the stated values rather than on affect. It seems unlikely that deliberative people do not have any affective reactions to the lotteries, but they might therefore abstract from this affective information and might discount or neglect it when making their judgments (a process that requires time).

This interpretation of the observed relationship between habitual decision modes and lottery choice behavior is in line with other research as well. In Kaufmann's (2003) study, people were presented with a list of return values for individual stocks, which differed in the total return and the variance of the return (i.e., the associated risk of the stock). People classified as intuitive, based on the PID scale, had a higher degree of sensitivity towards the risk of the individual stocks than the deliberatives. They preferred the shares with less variance in returns, thus they showed behavior which was not risk neutral which is in line with our findings. Similar to the findings in our study and consistent with the "risk as feelings hypothesis" (Loewenstein et al., 2001), the risky stocks seem to trigger a feeling of uncertainty that particularly affects intuitive people in their evaluations of the lotteries.

Although affect and risk perception are increasingly mentioned in the literature, the focus has mostly been on the influence of mood or affective states on risky decision-making (e.g., Isen et al., 1988; Mano, 1994; Wright and Bower, 1992). In this work we consider the impact of intuitive or deliberative decision-making based on the idea that the information used for a judgment varies with respect to the individually preferred habitual decision mode. While deliberative people rather use the stated information, intuitives seem to process not only the stated values but also their subjective feeling of how safe or how good a lottery is. People using affective information (i.e., people with a preference for intuition) may be more prone to the effects of mood on their decisions in risky situations. Future studies might attempt to control for mood effects to rule out this explanation.

Our results suggest that people differ systematically in the way they solve simple monetary risky decision problems. This study links psychometric measures with individual utility functions, the backbone of modeling individual decision-making in the economic sciences. We have identified a person variable – the individual preference for intuition and deliberation – that helps to explain heterogeneity in utility functions. The findings are further evidence that affective-intuitive and deliberative decision modes affect peoples' decisions in substantial ways. Further theoretical and empirical work on decision-making under risk and uncertainty will profit from considering different decision modes, for example by assessing the individual preference for intuition and deliberation.

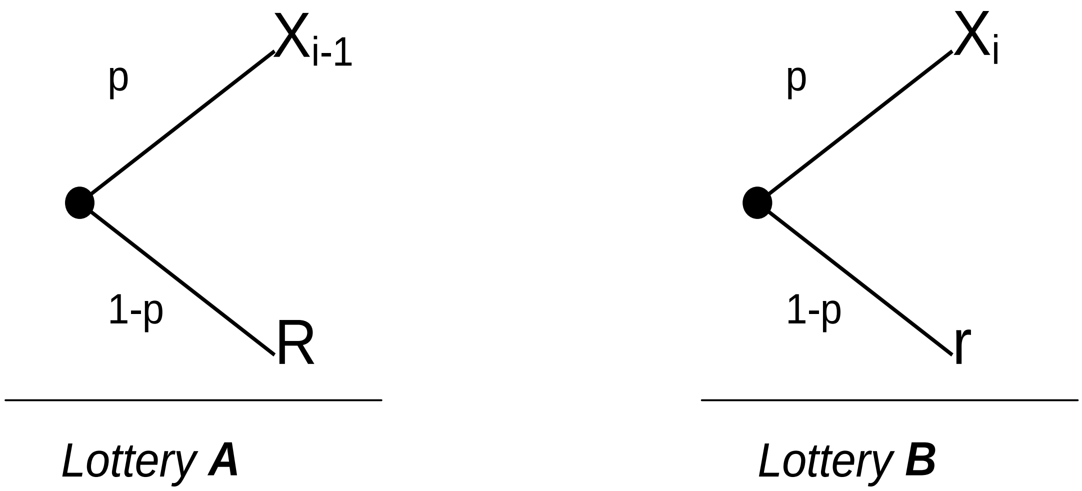
6 Appendix

6.1 Details of Eliciting the Value Function and Risk Attitude Parameters (α and β)

Individuals' utility functions for the gain and loss domains are elicited using a series of 64 individually adapted lottery choice questions presented by the computer.

The method of utility function elicitation is based on the construction of so-called standard sequences of outcomes, $\{x_0$ to $x_6\}$ (i.e., monetary outcomes that are equally spaced in terms of their utility). In our design, we use a five-step interval bisection procedure to determine an outcome x_1 for which the subject is indifferent between the lotteries A=(x_0, p; R, 1-p) and B=(x_1, p; r, 1-p) (see Figure 3), where x_0, R, x_1, and r denote monetary payoffs of the lottery and p and (1-p) denote the probabilities of the respective payoffs (see Figure 3). Here, we have $0 \leq r < R < x_0 < x_1$ with r, R and x_0 held constant. The answers to the first five presented lottery choice questions allow us to determine the desired x_1 that achieves indifference between lottery A and B.

Figure 3: An example of the two presented lotteries.



In the next step of this procedure (i.e., the next 5 presented lotteries) we determine, again based on bisection, an x_2 for which the subject is indifferent between the lotteries (x_1, p; R, 1-p) and (x_2, p; r, 1-p). We continue this method until we have determined an x_6, (that is, until we have $5 \cdot 6 = 30$ lottery choice questions in total, plus two consistency check questions). Another 32 questions that follow the same logic explained above are presented for the elicitation of the utility function for losses. Note that in our study we have set R to €100 and r to €0; x_0 has been set to €200.[6] These values are based on the suggestions of Abdellaoui (2000) and Wakker and Deneffe (1996). We start every five-step bisection procedure for the elicitation of a *new* x_i with a value of $x_i = x_{i-1}$ + €500. The interval within which we determine the new x_i via bisection is then [x_{i-1}, x_{i-1} + €1000]. Furthermore, p is set to 2/3 for all subjects and *for all lottery choices*, thus excluding the possibility of the perturbing effect of different individual probability weighting functions for the construction of the utility function.

Now, let $u(\cdot)$ denote the value- or utility-function on the gain or the loss domain and let $w(\cdot)$ denote the probability weighting function for the respective domain.[7] Then the constructed indifferences give pairs of equations of the following type:

$$w(p)\, u(x_i) + (1-w(p))\, u(R) = w(p)\, u(x_{i+1}) + (1-w(p))\, u(r) \quad (1)$$

$$w(p)\, u(x_{i+1}) + (1-w(p))\, u(R) = w(p)\, u(x_{i+2}) + (1-w(p))\, u(r) \quad (2)$$

From these two equations it follows:

$$u(x_{i+1}) - u(x_i) = u(x_{i+2}) - u(x_{i+1}) \quad (3)$$

That is, in terms of utility, the trade-off of x_i for x_{i+1} is equivalent to the trade-off of x_{i+1} for x_{i+2}. We obtain a standard sequence of outcomes, $\{x_0, x_1, x_2, x_3, x_4, x_5, x_6\}$, which is, by construction, increasing for gains and decreasing for losses and uniquely characterizes the individuals' utility function, since all x_i are equally spaced in terms of their utility (see Figure 1).

Following Tversky and Kahneman (1992), we assume a power utility function that is "by far the most popular form for estimating money value" (Prelec, 2000):

$$u(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -(-x)^{\beta} & \text{if } x < 0 \end{cases} \quad (4)$$

6 For the loss domain, we used the negative of the above values as R, r, and x_0, respectively.

7 That is, we implicitly assume that individual preferences can be represented by, for example, (Cumulative) Prospect Theory. Note, however, that the value function that we elicit is indeed a von-Neumann-Morgenstern utility function. Equation (3) holds also under Expected Utility Theory, which can be shown by substituting p for $w(p)$ in equations (1) and (2).

6.2 Items of the Preference for Intuition and Deliberation Scale (Betsch, 2004)

Item No.	Preference for deliberation	α =.76
1	Before making decisions I first think them through.	
3	Before making decisions I usually think about the goals I want to achieve.	
6	I think about myself.	
7	I prefer making detailed plans rather than leaving things to chance.	
10	I am a perfectionist.	
11	I think about a decision particularly carefully if I have to justify it.	
13	When I have a problem I first analyze the facts and details before I decide.	
14	I think before I act.	
16	I think more about my plans and goals than other people do.	
	Preference for intuition	α =.77
2	I listen carefully to my deepest feelings.	
4	With most decisions it makes sense to completely rely on your feelings.	
5	I don't like situations that require me to rely on my intuition.	
8	I prefer drawing conclusions based on my feelings, my knowledge of human nature, and my experience of life.	
9	My feelings play an important role in my decisions.	
12	When it comes to trusting people, I can usually rely on my gut feelings.	
15	I prefer emotional people.	
18	I am a very intuitive person.	
19	I like emotional situations, discussions, and movies.	

Note: Instructions: Please answer all the following questions about your life in general. Your answers should correspond to the way you generally make decisions. Circle the number that best represents your opinion. 1 means that you very much disagree; 5 means that you very much agree.

7 References

Abdellaoui, M. (2000): Parameter-free elicitation of utility and probability weighting functions. *Management Science*, 46, 1497-1512.

Baron, R. M., and D. A. Kenny (1986): The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.

Betsch, C. (2004): Präferenz für Intuition und Deliberation. Inventar zur Erfassung von affekt- und kognitionsbasiertem Entscheiden [Preference for Intuition and Deliberation (PID): An Inventory for Assessing Affect- and Cognition-Based Decision-Making]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 25, 179-197.

Bless, H., T. Betsch, and A. Franzen (1998): The impact of context cues on solutions to the "asian disease" problem. *European Journal of Social Psychology*, 28, 287-291.

Breckler, S. J. (1984): Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology*, 47, 1191-1205.

Bruner, J. S. (1957): Going beyond the information given. In: H. E. Gruber, K. R. Hammond, and R. Jessor (Eds.), *Contemporary approaches to cognition*, 41–69. Cambridge, MA: Harvard University Press.

Chaiken, S. (1980): Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality & Social Psychology*, 39, 752-766.

Eisenführ, F. and M. Weber (2003): Rationales Entscheiden (4th edition). Heidelberg, Germany: Springer.

Epstein, S. (1983): Aggregation and beyond: Some basic issues on the prediction of behavior. *Journal of Personality*, 51, 360-392.

Farquhar, P. H. (1984): Utility Assessment Methods. *Management Science*, 30, 1283-1300.

Fechner, G. T. (1860): Elemente der Psychophysik. Leipzig, Germany: Breitkopf und Härtel.

Fetherstonhaugh, D., P. Slovic, S. Johnson, and J. Friedrich (1997): Insensitivity to the value of human life: A study of psychophysical numbing. *Journal of Risk and Uncertainty*, 14, 283-300.

Glaser, M., M. Nöth, and M. Weber (2004): Behavioral Finance. In: D. J. Koehler and N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making*. Oxford: Blackwell.

Glaser, M., and M. Weber (2005): Overconfidence and Trading Volume. *Working Paper*, University of Mannheim.

Gonzalez, R. and G. Wu (1999): On the shape of the probability weighting function. *Cognitive Psychology*, 38, 129-166.

Hogarth, R. (2001): Educating intuition. Chicago: University Press.

Hsee, C. K. and Y. Rottenstreich (2004): Music, pandas, and muggers: on the affective psychology of value. *Journal of Experimental Psychology: General*, 133, 23-30.

Isen, A. M., T. E. Nygren, and F. G. Ashby (1988): Influence of positive affect on the subjective utility of gains and losses: It is just not worth the risk**.** *Journal of Personality & Social Psychology,* 55*,* 710-717.

Johnson, E. J., J. W. Payne, and J. R. Bettman (1988): Information displays and preference reversals. *Organizational Behavior and Human Decision Processes*, 42, 1-21.

Kahneman, D., and A. Tversky (1979): Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47, 263-291.

Kaufmann, M. (2003): Informationsintegration unter expliziter Einstellungsbildung [Information integration under explicit attitude formation]. *Unpublished manuscript*, University of Heidelberg.

Langan-Fox, J. and D. A. Shirley (2003): The Nature and Measurement of Intuition: Cognitive and Behavioral Interests, Personality and Experiences. *Creativity Research Journal*, 15, 207-222.

Loewenstein, G. and T. O'Donoghue (2004): Animal Spirits: Affective and Deliberative Processes in Economic Behavior. *Working Paper*, 5/4/2004. Carnegie Mellon University.

Loewenstein, G., E. Weber, C. K. Hsee, and E. Welch (2001): Risk as feelings. *Psychological Bulletin*, 127, 267-286.

Mano, H. (1994): Risk-taking, framing effects, and affect. *Organizational Behavior and Human Decision Processes*, 57, 38-58.

Parker, A. and B. Fischhoff (2005): Decision-making competence: External validation through an individual-differences approach. *Journal of Behavioral Decision-Making*, 18 (1), 1-27.

Prelec, D. (2000): Compound Invariant Weighting Functions in Prospect Theory. In: D. Kahneman, and A. Tversky (Eds.), *Choices, Values and Frames*. Cambridge, UK: Cambridge University Press.

Schwartz, B., A. Ward, J. Monterosso, S. Lyubomirsky, K. White, and D. R. Lehman (2002): Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality & Social Psychology*, 83(5), 1178-1197.

Stanovich, K. E. and R. F. West (1998): Individual differences in rational thought. *Journal of Experimental Psychology,* 127, 161-188.

Stanovich, K. E. and R. F. West (2000): Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23, 645-726.

Tversky, A. and C. Fox (1995): Weighting Risk and Uncertainty. *Psychological Review*, 102, 269-283.

Tversky, A. and D. Kahneman (1981): The framing of decisions and the psychology of choice. *Science*, 211, 453 - 458.

Tversky, A. and D. Kahneman (1992): Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5, 297-323.

Verplanken, B., G. Hofstee, and H. J. W. Janssen (1998): Accessibility of affective versus cognitive components of attitudes. *European Journal of Social Psychology*, 28, 23-36.

Wakker, P. P. and D. Deneffe (1996): Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Science*, 42, 1131-1150.

Wright, W. F. and G. H. Bower (1992): Mood effects on subjective probability assessment. *Organizational Behavior and Human Decision Processes*, 52, 276-291.

Zajonc, R. B. (1980): Feeling and thinking. Preferences need no inferences. *American Psychologist*, 35, 151–175.

Sequential Decision Behavior with Reference Point Preferences:

Theory and Experimental Evidence

Abstract: People are heterogeneous with respect to their behavior in sequential decision situations. This paper develops models for behavior in a simple sequential decision situation under the assumption of expected utility maximization and under the assumption of sequential updating of utility reference points during the decision task. I find experimental evidence that supports the new reference point model: Individual loss aversion is systematically related to the observed behavior in a way that is consistent with the predictions of the reference point model; that is, loss aversion helps to predict heterogeneity in behavior. Risk attitude is not related to observed behavior. The finding that many people set reference points in sequential decision tasks is of interest in, e.g., consumer economics, labor economics, finance, and decision theory.

1 Introduction

Sequential decision situations occur often in our everyday lives and people are very heterogeneous with respect to their behavior in these situations. Does information on individual preferences help us to predict how people behave in sequential decision situations? The goal of this paper is to investigate the relationship between individual preferences and sequential decision behavior based on a controlled laboratory experiment.
A very elementary representative of a sequential decision situation is a so-called search task. Search tasks are attractive for the study of sequential decision behavior: first, because of their very simple sequential decision structure that can be easily dealt with theoretically and empirically and that participants in a laboratory experiment understand easily; second, because this decision structure masks a complicated optimization problem that – comparable to sequential decision situations in our everyday lives – cannot be solved without a computer in most cases. Indeed, we face search tasks in many everyday situations, e.g., when we look for the best price of a certain product or when we search for a new job. In those tasks, we must essentially decide between committing resources to an attractive proposition or deferring the decision in the hope of receiving a better deal.

Behavior in economic search situations has been investigated both theoretically and experimentally in the fields of economics, mathematics, and psychology since the 1950's. Seminal theoretical work in the economic strand of this literature was done by Simon (1955) and by Stigler (1961). Since then, numerous authors have investigated variations of search problems, and they have focused on examining the huge heterogeneity in search behavior (e.g., Cox and Oaxaca, 1989; Harrison and Morgan, 1990; Hey, 1981; 1982; 1987; Houser and Winter, 2004; Kogut, 1990; Schunk and Winter, 2004; Sonnemans, 1998; 2000).

Economic theory suggests that this heterogeneity observed in sequential decision behavior is reflected in the heterogeneity of individual preferences. In this paper, a lottery-based preference elicitation mechanism is combined with a price search task in an economic laboratory experiment in order to investigate the link between individual preferences and individual search behavior based on utility-based search models. The underlying idea is that in price search tasks as well as in lottery tasks, people make financial decisions under risk, and they thereby reveal their preferences.

The contribution of this work is, first, the theoretical development of search models that are not based on the assumption of risk neutrality as well as the development of a model that involves reference point updating. Second, the paper provides experimental evidence that measures of loss aversion are a better predictor of search behavior than measures of risk aversion. An explanation for this finding is that subjects set utility reference points relative to which they evaluate the possible future outcomes in the search task. In particular, heterogeneity in individual search behavior might be better explained by a model that assumes sequential updating of utility reference points than by search models that are based on expected utility theory.

The findings are of interest for decision theory. They help to understand the determinants and properties of individual search behavior in markets (e.g., Zwick et al., 2003), and they serve as a guide to theoretical and structural econometric specifications that explicitly allow for individual heterogeneity in applied search theory. These specifications are being developed in many fields, including research on consumer search and job search (Eckstein and Van den Berg, 2006). Finally, the sequentially risky decision nature of the search problem makes the results interesting for theoretical and applied research in finance (Gneezy, 2003).

This paper first establishes links between search behavior and individual preferences by developing various search models, in particular the reference point model (section 2). Then, the experimental design (section 3) and the methodology to draw inference about search behavior and preferences based on the experimental data (section 4) are described. Next, the link between the elicited preferences and the observed search behavior is investigated (section 5): I first discuss descriptive information and a correlation analysis, and I finally present an analysis that exploits the discrete time-to-event nature and the panel nature of the data in order to investigate this link. The methodology and possible explanations for the findings are discussed in section 6; section 7 concludes.

2 Models of Search Behavior

In this section I first derive the optimal search behavior of an expected utility maximizer, both under risk neutrality (section 2.1) and without restrictions on individual risk attitude (section 2.2). For the derivation of the decision rules, two cases are considered: In the first case, the cost of each completed search step is treated as sunk costs; in the second case, I derive the finite horizon optimal stopping rule assuming that subjects do not treat past search costs as sunk costs. Finally, in section 2.3, I develop the reference point model.

2.1 Optimal Stopping in Search Tasks under Risk Neutrality

Assume that a searcher's goal is to purchase a certain good that she values at €100. The searcher sequentially observes any number of realizations of a random variable X, which has the distribution function $F(\cdot)$. In the current experiment, $F(\cdot)$ is a discrete uniform distribution with lower bound €75 and upper bound €150. Let the cost of searching a new location be € c. Assume that at some stage in the search process, the minimal value that the searcher has observed so far is € m.[1] Basic search theory assumes that individuals treat the cost of each search step, once completed, as sunk costs (Kogut, 1990;

[1] For the remainder of the derivation in this section, the currency units are skipped.

Lippman and McCall, 1976) and that they compare the payoff of one additional search step with the payoff from stopping.[2]

Then, subjects solve the problem based on a one-step forward-induction strategy and the expected gain from searching once more before stopping, $G(m)$, is generally given by:

$$G(m) = -\underbrace{[1 - F(m)]m}_{\otimes} - \underbrace{\int_{75}^{m} xdF(x)}_{\oplus} - c + m. \tag{1}$$

The term $\bigotimes$ accounts for the case in which a value larger than m is found with probability $(1 - F(m))$. In this case, m remains the minimum price. The term $\bigoplus$ stands for the case in which a lower value than m is found and computes the expected value in this case. There exists *a unique* value m^* with $G(m^*) = 0$, if $G(\cdot)$ is continuous and monotonic. Straightforward manipulation shows that the solution to this problem is identical to solving the following problem for m:

$$\pi(100 - m) = (1 - F(m))\pi(100 - m - c) + \int_{75}^{m} \pi(100 - x - c)dF(x) \tag{2}$$

Here, $\pi(\cdot)$ is the payoff-function from the search game. The payoff is truncated at €0 in the experiment:

$$\pi(x) = max\{0, x\} \tag{3}$$

The left-hand side of equation (2) is the payoff from stopping, and the right-hand side denotes the payoff from continuing the search. It is found that the optimal strategy is to keep searching until a value of X less than, or equal to, the optimal value m^* has been observed. For the search task considered in this paper, I find $m^* = 86$. That is, a risk-neutral searcher has the following decision rule: Stop searching as soon as a price less than or equal to €86 is found.

Now, consider that subjects do *not* treat search costs as sunk costs. That is, for their decision whether to stop or to continue the search, they consider the total benefits and costs of the search; the agent stops searching only if the stopping value is higher than the continuation value. It follows that the problem is treated as a finite horizon problem that is solved backwards. Define $S_t = \{t, m\}$ as the agents' state vector after t search steps.

After the agent has stopped searching, she will buy the item and she receives a total payoff:

$$\Pi(S_t) = max\{0, 100 - m - t \cdot c\}. \tag{4}$$

The agent stops searching only if the continuation value of the search is lower than the stopping value. The recursive formulation of the decision problem is therefore:

$$J_t(S_t) = max\{\Pi(S_t), E[J_{t+1}(S_{t+1})|S_t]\}. \tag{5}$$

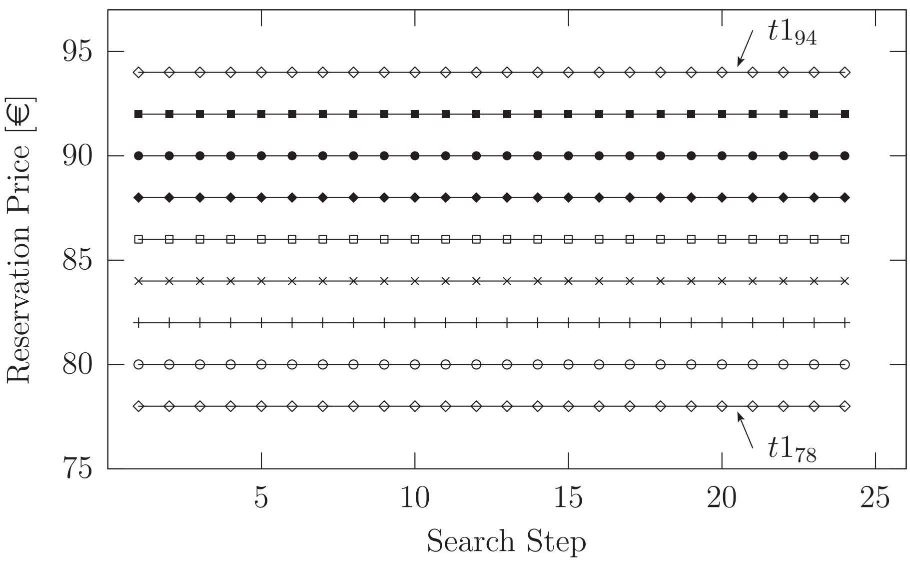
[2] Kogut's (1990) findings show that a certain proportion of subjects does not treat search cost as sunk. A model in which search cost are not treated as sunk cost is presented later in this section.

$E(\cdot)$ represents the mathematical expectations operator, and the expectation is taken with respect to the distribution of $S_{t+1}|S_t$. Again, this problem has the reservation price property at every t. The reservation price begins at €86, first stays constant, then starts decaying slowly, reaches €80 in the 19th round, and then decays at a rate of about one per round from that point forward.

2.2 Stopping Rules in Search Tasks Without Restrictions on Risk Attitudes

The derivations above are based on the assumption of a risk neutral searcher as in the existing literature on search behavior. Only for risk neutral subjects it is individually rational to use a risk neutral optimal stopping rule.

Figure 1: Constant reservation price path (type-1-rules) for different risk attitudes in, e.g., CARA or CRRA specifications of a utility function. The more risk averse a searcher is, the higher is her reservation price level.



As a more general case, I therefore develop models for a searcher with an arbitrary, monotone utility function $u(\cdot)$. If the searcher ignores sunk costs and takes her decisions based on a one-step forward-looking strategy, the equation that determines her reservation price m^* then has the following form that follows from (2):[3]

$$u(100 - m) = (1 - F(m))u(100 - m - c) + \int_{75}^{m} u(100 - x - c)dF(x) \tag{6}$$

Equation (6) is solved numerically for the reservation price $m^*(\eta)$, given the search environment and a utility function on gains that is parameterized entirely by a parameter

[3] This equation does *not* characterize the optimal solution to the search problem. It gives the optimal strategy for a searcher with arbitrary risk attitude, captured by $u(x)$, who ignores sunk costs, and who uses a one-step forward induction strategy.

(vector) η. The solution has the constant reservation price property, independent of the functional form of $u(\cdot)$. Figure 1 shows the constant reservation price decision rule for different risk attitude parameters of, e.g., a CRRA or a CARA utility function. The more risk averse the searcher is, the higher is her constant reservation price value. Henceforth, I refer to rules of this type as forward optimal rules, keeping in mind that this rule is only optimal *conditional* on the individual utility function and on the assumption of a one-step forward strategy that ignores sunk costs.

Analogous to the derivation of the optimal search rule in the risk neutral case, I now consider that subjects do not treat search costs as sunk costs. Again, this is a finite-horizon problem. After the agent has stopped searching, she buys the item and receives a total payoff:

$$\Pi^u(S_t) = max\{0, u(100 - m - t \cdot c)\}. \tag{7}$$

The recursive formulation of the dynamic discrete choice problem is:

$$J_t^u = max\{\Pi^u(S_t), E[J_{t+1}^u(S_{t+1})|S_t]\}. \tag{8}$$

This problem has, at every t, the reservation price property. The monotonically falling reservation price implies that the agent should not exercise recall, i.e. she should not recall previously rejected prices. Figure 2 plots the reservation price paths for a CRRA-utility function specification; figure 3 assumes a CARA-specification. Henceforth, I refer to rules of this type as backward optimal rules. These rules are optimal search rules conditional on the individual utility function.

Figure 2: Reservation price path for type-2-rules for different risk attitudes. CRRA specification of the utility function. The more risk averse a searcher is, the higher is her reservation price level.

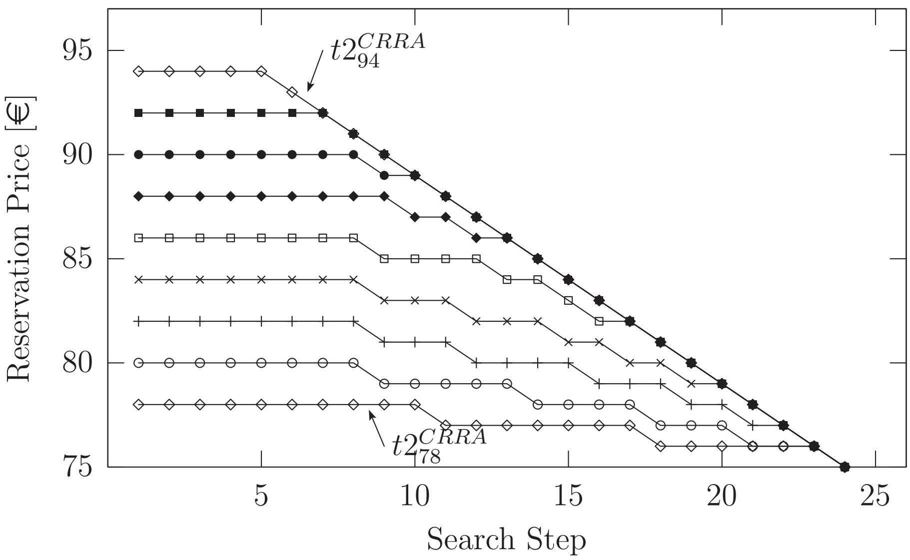
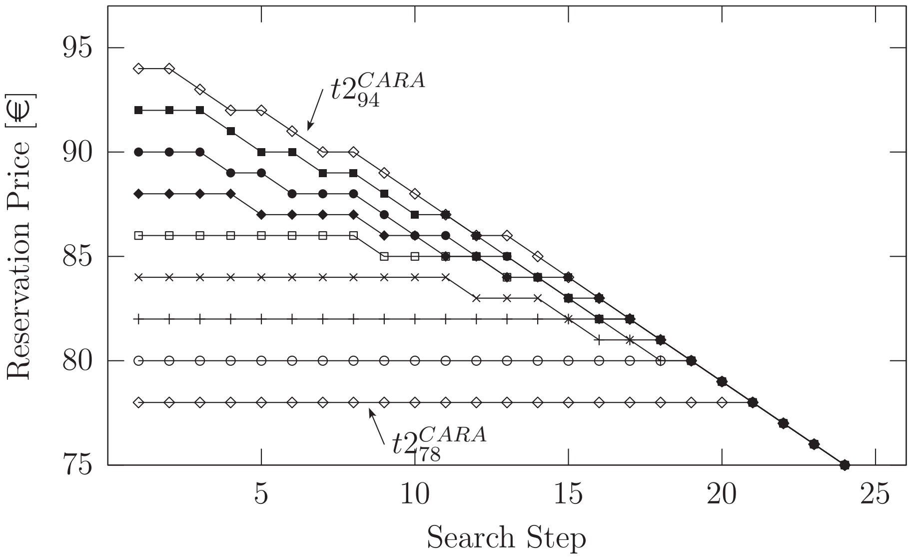


Figure 3: Reservation price path for type-2-rules and different risk attitudes. CARA specification of the utility function. The more risk averse a searcher is, the higher is her reservation price level.



From the theoretical deliberations so far it can be inferred that – regardless of what type of rule subjects use, forward or backward optimal rules – the more risk averse a person is, the earlier she should stop search, i.e. the higher is the reservation price that she uses.

2.3 The Reference Point Model

When talking to people that are actually facing a search situation, for example graduate students on the job market, many people talk about their decision situation as if they were comparing their possible future offers from continuing the search with the current best alternative that they have already been offered. They consider everything that is worse than the current best payoff that they have for sure as a loss relative to that sure payoff, and everything that is better is considered as a gain relative to the sure payoff. The model that I develop in this section, the reference point model (henceforth: *rp-model*), captures this idea that future possible payoffs are compared to a reference point. The model is based on a concept from the psychology of decision-making, the concept of loss aversion, which plays a central role in Kahneman and Tversky's (1979) descriptive theory of decision-making under risk. Loss aversion refers to the tendency of people to be more sensitive to reductions in their current level of well-being than to increases. The rp-model claims that during the search task, subjects set reference points relative to which the decision whether to stop or to continue the search is evaluated in terms of gains and losses. While the models based on EU-maximization (see previous subsection) implicitly assume that the reference point is always at zero payoff, the rp-model assumes a reference point which is always at the current best payoff.

To formalize these ideas, let $u(\cdot)$ be the individual utility function. Following Kahneman and Tversky (1979), I decompose the function into the utility function on gains, $u^+(\cdot)$, and the utility function on losses, $u^-(\cdot)$:

$$u(x) = \begin{cases} u^+(x) & x \geq 0 \\ u^-(x) & x < 0. \end{cases} \tag{9}$$

Subjects have to decide whether to stop or to continue the search at every search step t. The reference point at time t is the payoff that they get from stopping when they realize the best price draw, m_t, that they have in hand at time t. The utility from continuing the search is evaluated relative to this reference point:

If subjects find a price lower [higher] than $m_t - c$ in the next round $t+1$, they make a net gain [loss] relative to their current situation where they have m_t in hand – see the term $\bigotimes$ [$\bigoplus$] in (10).

The model implicitly assumes that subjects solve the problem based on one-step forward-induction. In the rp-model the *expected gain at time t from searching once more before stopping*, $G(m_t)$, is given by

$$\begin{aligned} G(m_t) &= \underbrace{\int_{-\infty}^{m_t-c} u^+(m_t - x - c)dF(x)}_{\bigotimes} \\ &+ \underbrace{\int_{m_t-c}^{m_t} u^-(m_t - x - c)dF(x) + (1 - F(m_t)) \cdot u^-(-c)}_{\bigoplus}. \end{aligned} \tag{10}$$

That is, the model assumes that people sequentially update their reference point in every time step. Model (10) is stationary in the same sense as the forward optimal model (6): The search behavior is independent of time t since subjects focus on the marginal gain or loss from the next step but not on the total payoff from the search. Identical with the prediction of the forward optimal search model (6), this model results in a *constant reservation price over time*. As in the forward optimal search model, the negligence of the sunk costs incurred during the search process is here responsible for the stationarity of the model.

I rewrite equation (10) for simplicity. For this purpose, define $p(x, m_t)$ as the rp-payoff-function, i. e. the function that determines individual payoff (relative to the reference point) in the framework of the rp-model (10), conditional on having the best offer m_t in hand at time t:

$$p(x, m_t) = \begin{cases} m_t - x - c & x \leq m_t \\ -c & x > m_t \end{cases} \tag{11}$$

With the help of (11), the rp-model (10) is equivalently written as:

$$\begin{aligned} G(m_t) &= \int_{-\infty}^{m_t - c} u^+(p(x, m_t))dF(x) + \int_{m_t - c}^{\infty} u^-(p(x, m_t))dF(x) \\ &= \int_{-\infty}^{\infty} u(p(x, m_t))dF(x). \end{aligned} \tag{12}$$

Several studies (e.g., Kogut, 1990; Sonnemans, 1998) find that many subjects also focus (to some extent) on total earnings from the search game, instead of only focusing on the marginal return of another draw. This translates into a reservation price that does not remain constant, but is falling when t increases, similar to the prediction of the backward optimal model (8).
In the framework of the rp-model, this means that subjects take into account that total payoff is left-truncated at €0. In other words, if subjects focus on total earnings, they take into account that when continuing the search, they do *not* risk losing money if their payoff at the current reference point is already €0. That is, the maximal loss that they can incur is the search cost (if the payoff at the reference point is higher than the search cost), or the payoff at the reference point (if the payoff at the reference point is less than the search cost).
This idea, namely that subjects also focus on total earnings instead of only focusing on the marginal return of another draw, is translated into the framework of the rp-model by a modification of the rp-payoff-function.
For this purpose, I first define two functions $q(\cdot)$ and $v(\cdot)$:

$$q(y) = \begin{cases} q(y) = y & y \geq 0 \\ 0 & y < 0 \end{cases} \tag{13}$$

$$v(y) = \begin{cases} v(y) = y & y \geq -c \\ 0 & y < -c \end{cases} \tag{14}$$

The modified rp-payoff-function $p(x, m_t, t)$ now has the following form.[4]

$$p(x, m_t, t) = \begin{cases} q(100 - c \cdot t - x - c) & m_t \geq 100 - c \cdot t \\ v(m_t - x - c) & m_t < 100 - c \cdot t \wedge x \leq m_t \\ v(m_t - (100 - c \cdot t)) & m_t < 100 - c \cdot t \wedge x > m_t \end{cases} \tag{15}$$

[4] A detailed derivation of the function $p(x, m_t, t)$ is given in the appendix of the paper.

With the modified version of the rp-payoff-function, the rp-model (12) is written as follows:

$$\begin{aligned} G(m_t) &= \int_{-\infty}^{m_t - c} u^+(p(x, m_t, t))dF(x) + \int_{m_t - c}^{\infty} u^-(p(x, m_t, t))dF(x) \\ &= \int_{-\infty}^{\infty} u(p(x, m_t, t))dF(x). \end{aligned} \tag{16}$$

I have now developed two search models, (16) and (10), that assume that subjects update their reference points during the search process. The EU-based models presented in the previous subsection are based on only one branch of the utility function, $u^+(\cdot)$, that is, in these models, behavior can essentially be captured by a one-parameter functional form. However, both rp-models are based on two independent branches of the utility function, $u^+(\cdot)$ and $u^-(\cdot)$. In line with existing empirical studies on loss aversion (e.g., Benartzi and Thaler, 1995; Schmidt and Traub, 2002; Tversky and Kahneman, 1992), I therefore assume the following one-parameter form of the reference point utility function:

$$u(x) = \begin{cases} u^+(x) = x & x \geq 0 \\ u^-(x) = \lambda \cdot x & x < 0 \end{cases} \tag{17}$$

This functional form is a strong assumption since it imposes that individuals are risk neutral and that only the kink at the utility reference point plays a role for observed search behavior. The assumption is introduced to reduce the reference point model to one preference parameter which can be identified with a standard experimental method, and I essentially impose the same assumption for the identification of loss aversion measures based in the experiment. The measurement of a parameter that characterizes loss aversion is, of course, always connected to a functional form assumption and there is much disagreement over the definition and empirical measurement of an index for loss aversion (see Johnson et al., 2006; Koebberling and Wakker, 2005).[5] I will come back to this important issue later in this paper (in particular in the discussion section) and I will discuss that the main conclusions from this paper are robust to this assumption. Based on utility specification (17), the crucial parameter that determines individual search behavior is now the individual loss aversion parameter λ. The *stationary rp-model* (10) implies a constant reservation price search rule; the level of the reservation price path is a function of loss attitude λ.[6]

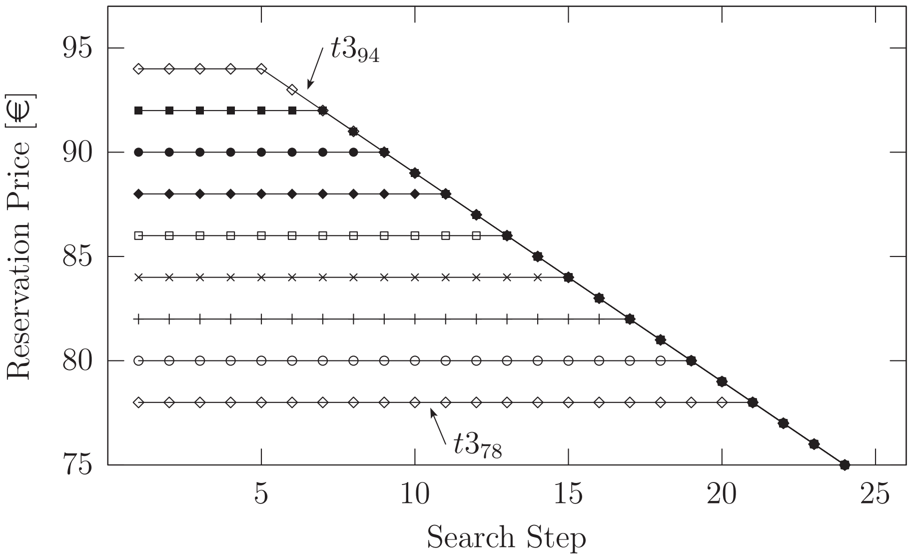
The *non-stationary rp-model* (16) implies, in line with the stationary rp-model (10), a

[5] For example, it is under debate, whether and how to define a global measure of loss aversion. Equation (17) implies that λ is a global measure of loss aversion, but it is obvious that under the assumption of more flexible functional forms for $u^+(\cdot)$ and $u^-(\cdot)$, loss aversion can also be defined *locally*, i.e. as a function of x. As section 3, which describes the experimental design, shows, I estimate *five different measures of loss aversion* (based on five different x-values) for each subject and the results of this paper are investigated for all five measures. This serves to underline that the findings are independent of one specific local measure of loss aversion.

[6] Algebraic transformations show that under (17) the rp-model (10) is identical to the classical risk neutral forward induction model (2) under the assumption that $\lambda = 1$.

reservation price path that varies systematically with the loss aversion parameter λ: the higher loss aversion, the higher the reservation price. However, in contrast to the stationary rp-model, the reservation price starts falling after a certain number of time-steps (see figure 4).

Figure 4: Reservation price path for type-3-rules: Non-stationary reference point model under risk neutrality. The more loss averse a searcher is, the higher is her reservation price level.



The stopping rules derived from the reference point models (10) and (16) are comparable to two classical search models that are based on EU-maximization:
– The stationary rp-model (10) predicts the same search behavior as the EU-based forward optimal search model (equation (6)), and both models assume that subjects ignore sunk costs.
– Similar to the EU-based model (8), the non-stationary rp-model (16) predicts that the reservation price is first constant and starts falling after a certain number of time steps. In both models, subjects do not ignore sunk costs.

While EU-based models and the rp-model predict very similar search behavior, the *explanation* for the search behavior is different: In the rp-model, loss aversion explains the level of the reservation price path, whereas in the EU-models, risk aversion explains this level. The rp-model is built on the idea that "loss aversion [...] provides a direct explanation for modest-scale risk aversion" (Rabin, 2000, p. 1288). Due to the similar predictions of the models, distinguishing between these preference-based explanations for search behavior requires *independent measures of individual preferences*, which I elicit in the experiment for each subject using standard lottery procedures.[7] I come back to this point in section 5.2. The following section describes the experimental design.

[7] It is tempting to find a parametrization of the decision environment (i.e., search cost c and price distribution $F(\cdot)$), in which the (empirical) identification of the underlying preferences based on *only* the

3 Experimental Design

The experiment consisted of three parts (A, B, and C) that were presented to the subjects in fixed alphabetical order. Parts A and C of the experiment served to elicit parameters that characterize subjects' preferences, and part B consisted of a series of repeated price search tasks used to elicit subjects' search behavior.

Note at this point that the decision in the price search task (part B), namely whether to stop (**s**) or to continue (**c**) the search, corresponds conceptually to the choice between a sure payoff (**s**) and a lottery (**c**) with several consequences. In order to create similar decision situations in both, the search task (part B) and the preference elicitation parts (part A and C), the certainty equivalent method (e.g., Wakker and Deneffe (1996) for risk aversion, and Tversky and Kahneman (1992) for loss aversion) has been used for preference elicitation. This way, subjects also deal with decision situations involving the comparison between a sure payoff (**s**) and a lottery (**c**) in the preference elicitation part of the experiment. Various methods for the elicitation of risk and loss attitudes exist, in particular the multiple price list design (Andersen et al., 2005; Holt and Laury, 2002), and parameter-free methods that have been developed in the decision-theoretic strand of the literature (e.g., Abdellaoui, 2000). The certainty equivalent method was used in this experiment since the decision situation used in this method is most similar to the decision situation in search tasks, as mentioned above.[8] For all participating subjects, the method has been used with various starting parameters, both for the case of loss and risk aversion, to get an estimate of the robustness of the results.[9]

observed search behavior (and *without* additional and independent preference information) would be easier. This would require finding an environment, in which the models do *not* yield similar predictions over observed ranges of the underlying preference parameters. Simulation studies, obtainable from the author upon request, show that identification is not easier in other environments. The key issue is that at the search step where the models yield *different* predictions, most subjects have already stopped searching. For example, consider a risk neutral searcher. This person has a constant reservation price of €86 for 8 search steps (see figure 2 and figure 3), but – if she were behaving according to the rp-model (see figure 4) – she has a constant reservation price of €86 for 13 search steps. Now, the probability that I observe her searching for more than 8 steps (which would allow for discrimination between the two models) is only $(1 - (\frac{12}{76}))^8 \approx 25\%$, and this percentage – one possible measure for how easy one can discriminate between EU-based models and the rp-model based on observed search behavior alone – does not change much if the search environment is modified, i.e. if I use different price distributions or search cost. I make a similar argument in the appendix, section 8.4, in the discussion about the assumptions that underlie the empirical analysis of search behavior.

[8] Wakker and Deneffe (1996) used - also at the University of Mannheim where the present experiment takes place and at various other places - the same elicitation method with the same stimuli, identical number of iterations, and identical probability parameters, but with different starting values. This method has the disadvantage that it does not use incentives, its advantage is that subjects are exposed to only few lottery decisions and it is the only established preference elicitation method that involves comparisons between sure payoffs and a lottery.

[9] To further motivate the usage of the certainty equivalent method, note that – compared to multiple price list methods – the method in use avoids using probabilities other than 50-50-probabilities. 50-

The descriptions of the experimental design will begin with part C, continue with part A, and end with part B. This makes some details of the design clearer.

3.1 Part C: Risk Attitude

In part C, the certainty equivalent method (e.g., Wakker and Deneffe, 1996) is used to elicit individual risk attitude. That is, subjects are presented with a two-outcome lottery (**c**) and a sure payoff (**s**) and they are asked to enter one missing value such that they are indifferent between the sure payoff and the participation in the lottery. In total, only three lotteries are presented to the subjects.

Two values, x_{min} =€0 and x_{max} =€24, are defined. The subject is asked to enter a sure payoff, the certainty equivalent $s_{0.50}$, that is as attractive to her as the participation in the lottery $(x_{min}, p; x_{max}, (1-p))$.[10] In the second question, the subject is asked to enter the sure payoff $s_{0.25}$ that is as attractive to her as the lottery $(x_{min}, p; s_{0.5}, (1-p))$. Finally, in the last question, the subject is asked to reveal indifference between the lottery $(s_{0.5}, p; x_{max}, (1-p))$ and a sure payoff by stating the sure payoff $s_{0.75}$.

The values €0, €$s_{0.25}$, €$s_{0.5}$, €$s_{0.75}$, and €24 are equally spaced in terms of their utility, which allows for the estimation of the individual utility function, thereby obtaining a risk attitude index for each subject in the domain between €0 and €24.[11]

3.2 Part A: Loss Attitude

Part A consists of two blocks, *(A-1)* and *(A-2)*, that are presented in *random order*, such that a direct order effect on the behavior in the search task can be excluded. In block *(A-1)* I use a method by Tversky and Kahneman (1992). Subjects are again presented with a 50-50-gamble $(x, 50\%; y, 50\%)$ and a sure outcome (**s**). In all five presented lottery tasks the sure consequence (**s**) has the value €0. One consequence of the two-outcome lottery has a value of $x \in$ {€-1,€-10,€-25,€-50,€-100}. These values are presented in random order. Subjects are asked to enter the monetary value y of the other outcome of this 50-50-lottery such that the lottery and the sure payoff of €0 are equally attractive to them (i.e., they have to adjust a mixed prospect to acceptability).[12]

In block *(A-2)*, subjects are presented with three pure certainty-equivalent lotteries of the same type as in part C, but with x_{min} =€1 and x_{max} =€9.[13]

50-probabilities have the advantage that they are well-known to most decision-makers through events such as throwing coins.

[10] The value of p was set to 50% for all subjects, i. e. $p = 1 - p$.

[11] Note that the search task is designed such that subjects earn at least €0 and at most €24.

[12] Please see the appendix, figure 5, for the graphical presentation of the lotteries.

[13] The lottery with x_{max} =€24 as one consequence was intentionally *not* shown in part A (but only in part C) in order to exclude an effect on the behavior in the search task, in which the maximum possible

3.3 Part B: Search Behavior

In part B subjects perform a sequence of search tasks. Each subject's goal is to purchase a certain good that she values at €100. The good is sold at infinitely many locations[14], and visiting a new location costs €1. Subjects are informed that the integer price at each location is drawn independently from a uniform price distribution with a lower bound of €75 and an upper bound of €150. After each price draw, subjects can stop and choose any price encountered so far, or they can continue their search at the incremental cost of another euro. The outcome of each search task is calculated as the evaluation of the object (€100) minus the price at the chosen location minus the accumulated search cost. To ensure that subjects are experienced with the task and to minimize the observation of learning behavior, subjects are allowed to perform an unlimited number of practice search tasks before performing a sequence of 15 tasks that determine their payoff for part B of the experiment. Finally, after the experiment is completed, one of these 15 rounds is selected randomly to determine the payoff.

The search-model question.
After the search task is finished, there is one additional lottery question (henceforth referred to as the search-model question), worded as follows:[15]
You have now dealt with lottery tasks and a price search task. Perhaps you have realized that the decision in the search task (to stop or to continue the search) is similar to the decision between the lotteries presented to you:
If you stop your search, you obtain a sure payoff, but if you decide to continue the search, you essentially play a lottery with a risky outcome.
Which of the two lotteries, I or II, is most similar to the lottery that you play when you continue the search from your point of view?
Lottery I: (€**A**, p%;€**B**, (100-p)%)
Lottery II: (€**X**, p%;€ −**Y**, (100-p)%)
(A, B, X, and Y denote arbitrary positive numbers, and p is a (percentage) number between 0 and 100).

---

gain is €24. The main purpose of the lotteries in block *(A-2)* in the framework of this design, in which the order of *(A-1)* and *(A-2)* was randomized, was to exclude systematic effects of *(A-1)* on the search task. Note that the data from *(A-2)* can still be used to check the validity of the analyses presented in this paper: I find that the conclusions of this paper are independent of which risk attitude parameters, those stemming from part C or those stemming from part *(A-2)*, are used in the analysis.

[14] In other words, subjects are not prevented from searching as long as they want. It is not reasonable, however, to search for more than 25 steps, because, given the payoff structure, every search task lasting for more than 25 rounds ends with a zero payoff. No subject has searched for more than 25 steps.

[15] The graphical presentation of the two lotteries I and II presented in the search-model question is *identical* with the graphical presentation of all other lotteries. Furthermore, the two lotteries, I and II, are presented in random order.

This question is of importance: Search models that are based on expected utility theory (henceforth: EU-theory) assume that subjects evaluate the next search step as a *pure* lottery (cf. lottery I). In contrast, the new rp-model assumes that subjects evaluate the next search step as a *mixed* lottery (cf. lottery II). Therefore, the answer to the search model question allows for subdividing the subject sample into two groups: subjects behaving in a manner consistent with an EU-based model and subjects behaving in a manner consistent with a model in which subjects set utility reference points.

A few final remarks on the experimental design: First, the purpose of including both mixed *(A-1)* and pure *(A-2)* lottery tasks in the first part is to have subjects get used to both tasks *before* they have to answer the search-model question. Second, to make sure that subjects have sufficient experience with the search task *and* have been exposed to pure and mixed lotteries, the search-model question is presented directly after they have performed the search task. Third, since subjects are informed on the instruction sheet about the properties of the search experiment (i.e., they are aware that their minimum payoff is €0 and that their maximum payoff is €24), the certainty-equivalent method with the values x_{min} =€0 and x_{max} =€24 is used *after* they have answered the search-model question (i.e. in part C). This avoids the potential influence of an exposure to lotteries with x_{min} =€0 and x_{max} =€24 on the answer to the search-model question.

3.4 Administration and Payoffs

The study was conducted in the Summer and Fall of 2004 in the experimental laboratory of the SFB 504, a national research center at the University of Mannheim. In eight sessions 119 students of the University of Mannheim participated in the experiment which was run entirely on computers using software written by the author. The instruction sheet presented full information about the search task (i.e., as in section 3.3) and I want to stress that it was emphasized that (i), subjects' payoff was truncated at €0 (i.e., they could *not incur losses* from the search task) and that, (ii), they would not earn a show-up fee (i.e., no reference point was induced).

4 Inference about Preferences and Search Behavior

This section first presents and discusses how risk and loss attitude is estimated from the data obtained in the lottery tasks of the experiment. Then, I describe how individual search behavior is classified based on the data obtained in the search experiments and the search models developed above.

4.1 Estimation of Risk Attitude

I estimate individual risk attitude based on a parametric approach allowing for a specification of both constant relative and constant absolute risk aversion (CRRA and CARA, respectively). For both functional forms, the utility function is estimated from the data obtained in part C using nonlinear least squares.

Utility functions of the power form (e.g., Abdellaoui, 2000; Tversky and Kahneman, 1992) assume that subjects have constant relative risk aversion (CRRA):

$$u(x) = \left(\frac{x - x^G_{min}}{x^G_{max} - x^G_{min}}\right)^{(\alpha+1)} \tag{18}$$

x^G_{max} is the largest elicited value of x in the gain domain, i.e. €24; x^G_{min} is the smallest elicited x-value in the gain domain, i.e. €0. The estimated coefficient α characterizes each subject's risk attitude under the CRRA-assumption. If $\alpha > 0$, the subject is risk seeking; if $\alpha < 0$, the subject is risk averse.

Utility functions of the exponential form (e.g., Currim and Sarin, 1989; Pennings and Smidts, 2000) assume that subjects have constant absolute risk aversion (CARA):

$$u(x) = \frac{1 - e^{-\gamma(x - x^G_{min})}}{1 - e^{-\gamma(x^G_{max} - x^G_{min})}} \tag{19}$$

For $\gamma = 0$ the function is defined to be linear, i.e. the subject is risk neutral. In the CARA-specification, the estimated coefficient γ characterizes each subject's risk attitude in the sense of an Arrow-Pratt-measure of risk attitude (Pratt, 1964), that is: $-u''(x)/u'(x) = \gamma$. If $\gamma < 0$, the subject is risk seeking; if $\gamma > 0$, the subject is risk averse.

4.2 Estimation of Loss Attitude

Based on the subjects' responses in part A of the experiment, an individual-specific index for loss aversion is calculated. The statistic $\lambda_x = -y/x$ is a measure of individual loss aversion, where $x \in \{$€-1,€-10,€-25,€-50,€-100$\}$ and y is the response to the corresponding lottery given in part A. This method of estimating a coefficient of loss aversion is the method used in Tversky and Kahneman (1992) and its idea is to obtain a simple measure that captures the tradeoff between gains and losses. Note that because of the five x-values that are used for the elicitation, I essentially elicit five different measures of loss aversion, and I will use all measures in the subsequent analyses.

4.3 Classification of Decision Rules Used in the Search Task

The next step of the analysis is to determine the decision rule used by each subject in the search task. In order to do so, a fixed set of candidate decision rules is specified, the "universe of search rules", and the decision rule that fits observed behavior best is attributed to each subject. Since the utility-based search models developed in section 2 establish a relationship between preference parameters and decision rules, I can assign preference parameters to the subjects based on the attributed search rules.[16]

The Universe of Search Rules

For the investigation of the relationship between individual preferences and search behavior, I use as candidate decision rules all those search rules that can be derived from the search models developed in section 2. The universe of search rules (i.e., the set of candidate search rules that are used in this paper to characterize search behavior) consists of the following 51 rules:

The *first class* of these decision rules, henceforth referred to as **type-1-rules**, share the constant reservation price property (see figure 1). These rules are either based on the assumption that subjects use the forward optimal search rule (equation (6), the EU-based model that neglects sunk costs), or the stationary rp-model (equation (10), the rp-model that neglects sunk costs). Each rule says that the subject uses a reservation price $r \in \{€78, ..., €94\}$ which is constant during the complete search round. The universe contains 17 *type-1-rules* denoted by $t1_{78}, t1_{79}, ..., t1_{94}$. Every rule corresponds to a certain risk attitude parameter α^{search} and γ^{search}.[17]

The *second class* of decision rules is based on the finite horizon search model (i.e., the backward optimal search rules developed in section 2). According to these **type-2-rules**, the reservation price is a function of the search step t and of individual risk attitude. I assume again 17 different *type-2-rules*, denoted by $t2_{78}^{CRRA}, t2_{79}^{CRRA}, ..., t2_{94}^{CRRA}$, derived based on the assumption of a CRRA-specification of the utility function: For the first rule, the reservation price at $t = 1$ is €78, for the second rule, it is €79, etc., and for the last rule it is €94 (see figure 2). Each reservation price path corresponds to a certain α-interval. The 17 price paths $t2_{78}^{CRRA}, t2_{79}^{CRRA}, ..., t2_{94}^{CRRA}$ correspond to a decreasing sequence of 17 α-intervals taken from the interval $[-0.973, 25.20]$.
Alternatively, the 17 *type-2-rules* can be derived based on the assumption of a CARA-specification of the utility function (see figure 3). Then, each reservation price path corresponds to a certain γ-interval, and the 17 paths correspond to an increasing sequence of γ-intervals taken from $[-2.028, 0.837]$. In the paper, it will always be clear from the context whether the particular *type-2-rules* are derived based on either a CRRA- or a CARA-specification of the utility function. Conditional on the assumption that a certain subject uses a finite horizon search model, risk coefficients α^{search} and γ^{search} can be

[16] I can attribute only small intervals of preference parameters and not exact point-values, since the prices presented in the price search task are discrete.

[17] Under risk neutrality, one finds a constant reservation price of €86. The set of 17 constant reservation price rules, $t1_{78}, t1_{79}, ..., t1_{94}$, is sufficiently large to classify all observed behavior (see figure 6 in the appendix), no subject is assigned a lower or higher reservation price, if I allow for a larger universe of rules.

attributed to her. These coefficients are the risk attitudes that explain best the observed search behavior.

Finally, the **type-3-rules** are based on the non-stationary rp-model (16), the rp-model developed under the assumption that subjects focus on total payoffs from searching. The reservation price is a function of the search step t and of individual loss aversion λ (see figure 4). Again, 17 different rules are considered, $t3_{78}, t3_{79}, ..., t3_{94}$: For the first rule, the reservation price at $t = 1$ is €78, for the second rule, it is €79, etc., and for the last rule it is €94. The rules correspond to a decreasing sequence of λ-intervals taken from the interval $[0.042, 3.392]$. Based on the *type-3-rules*, I attribute to every individual a loss coefficient λ^{search}. The assigned loss attitude coefficient best explains the observed search behavior conditional on the assumption that the subject uses the non-stationary rp-model.

Classification Procedure

To classify search behavior, I determine for each subject the proportion of choices consistent with each decision rule and I maximize this proportion over the set of all candidate decision rules (i.e., a subject is assigned the decision rule that generates the largest fraction of correct predictions). It is assumed that each subject follows exactly one of the decision rules in the universe of candidate rules and that she uses the same rule in each of the 15 payoff tasks. This assumption seems reasonable in view of the fact that all subjects are experienced when they begin the 15 payoff relevant tasks (see section 3.3).

Formally, the classification procedure is described as follows: Each search rule $c_i \in \mathcal{C}$, where $\mathcal{C}$ is the *universe of search rules* described above, is a unique map from subject i's information set S_{it} to her continuation decision $d_{it} \in \{0, 1\} : d_{it}^{c_i}(S_{it}) \rightarrow \{0, 1\}$. Now, let d_{it}^* denote the observed decision of subject i in period t. Then, define the indicator function:

$$X_{it}^{c_i}(S_{it}) = 1(d_{it}^* = d_{it}^{c_i}(S_{it})) \tag{20}$$

Let T_i be the number of decisions that are observed for subject i. I attribute to each subject the search rule that maximizes the likelihood of being used by that subject:

$$\hat{c_i} = \arg\max_{c_i \in \mathcal{C}} \sum_{t=1}^{T_i} X_{it}^{c_i}(S_{it}) \tag{21}$$

5 Results

This section starts with self-contained descriptions of the findings from the utility function elicitation (part A and part C) and of the search task (part B). The main contribution of this section is the combination of the data on individual preferences and on search behavior, such that correlations on the subject level can be analyzed. I test whether basic hypotheses on the relationship between search behavior and preference parameters, derived from the search models developed in section 2, are supported by the data.

5.1 Part C and Part A: Risk and Loss Attitude

Of the 119 subjects that participated in the experiment, 13 were excluded.[18] From the data in part C two indices of risk attitude, an index α (derived from a CRRA specification) and an index γ (derived from a CARA specification), were estimated *for each* subject.[19] From the data obtained in part (*A-1*), five indices of loss attitude, λ_1, λ_{10}, λ_{25}, λ_{50}, and λ_{100}, were calculated *for each* subject.

Table 1: Estimation results of the CRRA and CARA utility function specification and classification of subjects according to their risk attitude.

	Functional specification	
	CRRA(α)	**CARA (γ)**
Minimum coefficient estimate	-0.457	-0.153
Maximum coefficient estimate	2.345	0.093
Median coefficient estimate	0.000	0.000
Mean R^2 of all estimates	0.998	0.998
Proportion risk averse	37%	37%
Proportion risk neutral	37%	37%
Proportion risk seeking	26%	26%

Table 1 reports results of the nonlinear least squares estimation of the risk coefficients α and γ, including the mean coefficient of determination R^2 for those two estimations. The coefficients of determination are close to 1 for all nonlinear regressions. The proportions of different risk attitudes in the sample are independent of the functional form assumption of the utility function. The proportions of subjects in the particular categories is in general agreement with findings in other experimental studies. For example, the proportion of risk seeking subjects is higher than the proportions reported by studies based on, e.g., multiple-price list elicitation methods (e.g., Harrison et al., 2005; Holt and Laury, 2002), and it is lower than the proportions reported by, e.g., Abdellaoui (2000).

[18] In contrast to all other subjects, the utility functions derived from the answers of these 13 subjects are not strictly monotone. This is evidence that they did not understand the lottery tasks correctly or did not take it seriously.

[19] Alternatively, the data from part *(A-2)* can be used for the estimation of risk attitude. The data from part C are preferable, since in part C, the risk attitude index has been elicited in a monetary domain which is identical to the payoff domain of the search task. (Part *(A-2)* has in fact only been included to avoid order effects, see section 3.2). Therefore, only the results from part C are reported here. The conclusions of this paper are identical if the data from part *(A-2)* are used. The corresponding analyses can be obtained from the author upon request.

Table 2: Results of the loss aversion lottery questions and classification of subjects according to their loss attitude.

	x-values				
	−100	**−50**	**−25**	**−10**	**−1**
Minimum λ	1	.9	.96	.9	.5
Maximum λ	10	16	20	20	20
Median λ	1.7	1.6	1.6	1.9	1.5
Loss averse	70%	69%	69%	69%	61%
Loss neutral	30%	30%	30%	30%	37%
Loss seeking	0%	1%	1%	1%	2%

Table 3: Pearson correlation between the different elicited loss aversion coefficients. All correlations are statistically significant at the 1%-level.

		x-values				
		−100	**−50**	**−25**	**−10**	**−1**
x-values	**−100**	1.00				
	−50	0.88	1.00			
	−25	0.82	0.95	1.00		
	−10	0.80	0.94	0.96	1.00	
	−1	0.66	0.73	0.72	0.74	1.00

Table 2 shows the results of the loss aversion elicitation part of the experiment. Across all five elicited loss aversion questions subjects were predominantly loss averse in their choices. I find median loss aversion coefficients that are significantly higher than 1, and the values are higher than those reported in Schmidt and Traub (2002), but lower than the median values reported in Tversky and Kahneman (1992). As expected, there is a high and statistically significant degree of correlation between the individual answers to the various loss aversion questions (see table 3). In fact, 39% of the subjects exhibited constant loss aversion, that is, their loss aversion coefficient is identical for all loss aversion questions.[20]

5.2 Part B: Search Behavior

Search behavior differs considerably across individuals, for more information, see the appendix (section 8.3). Overall, I find a preponderance of early stoppers compared to behavior under the risk neutral stopping rules; this confirms results from earlier experimental studies (e.g., Hey, 1987; Sonnemans, 1998).

[20] Several empirical studies confirm the predominance of loss averse choices (e.g., Fishburn and Kochenberger, 1979; Tversky and Kahneman, 1992; Schmidt and Traub, 2002; Pennings and Smidts, 2003).

Considering (a) the universe of 51 search rules (see figures 1, 2, 3, and 4), (b) the rather low average number of search steps compared to the optimal strategy, and (c) the fact that only a finite number of search rounds per individual (namely 15 rounds) is observed, it is clear that discrimination between very similar reservation price paths, that is *across* search rule types (e.g., between $t1_{80}, t2_{80}$, and $t3_{80}$), is hardly possible. Individual search rule *types* are not (empirically) identified.[21] In contrast, the identification *within* a certain rule type is clear: For example, there is significant difference in whether a subject's behavior is more consistent with, for example, $t1_{80}$ rather than with $t1_{81}$.[22] In other words, individual risk attitude or loss attitude parameters can be attributed to a subject based on her behavior in the search task, *conditional on the assumption* that the subject uses a specific model. But, as I have already discussed at the end of section 2.3, this model itself cannot be identified based on the observation of the search behavior alone; independent measures of preferences that are elicited in part A and part C are additionally needed.

5.3 The Search-Model Question: Subdividing the Sample

As I have explained above (see section 3.3), the search-model question is used to subdivide the sample into P^R and P^C: 39 subjects answered that they see a similarity between the search task and lotteries with gains and losses and were categorized into group P^R; 67 subjects think about lotteries with only gains and were categorized into group P^C. Descriptive statistics on individual preferences and search behavior *by subgroup* are reported in table 4.

Table 4: Descriptive statistics for the complete sample and the subgroups P^R and P^C.

	Complete Sample		$\mathbf{P^R}$		$\mathbf{P^C}$	
	Median	**Std.Dev.**	**Median**	**Std.Dev.**	**Median**	**Std.Dev.**
$\lambda_{\mathbf{1}}$	1.5	2.57	1.5	1.72	1.5	2.96
$\lambda_{\mathbf{10}}$	1.9	2.63	1.7	1.41	2.0	3.12
$\lambda_{\mathbf{25}}$	1.6	2.38	1.5	1.02	1.6	2.86
$\lambda_{\mathbf{50}}$	1.6	2.27	1.5	1.21	1.8	2.68
$\lambda_{\mathbf{100}}$	1.7	2.10	1.5	1.84	2.0	2.23
α	0.0	0.38	0.0	0.26	0.0	0.44
γ	0.0	0.04	0.0	0.03	0.0	0.04
Search steps (ss)	80.49	18.05	81.79	18.99	79.73	17.57

[21] Asymptotically, that is if an infinite number of search rounds per individual is observed, individual search rules types are, of course, identified.

[22] For more details about the classification, please see the appendix, section 8.3.

5.4 Analyzing Search Behavior and Individual Preferences

As mentioned above, observed search behavior alone is not sufficient to identify "users" of the reference point model. However, in order to discriminate between subjects that use the rp-model and subjects that use one of the classical EU-based models, I can derive *hypotheses on the relationship between search behavior and individual preferences* that are testable based on the information gained in parts A, B, and C of the experiment. Essentially, it is hypothesized that for subjects from P^R, individual loss aversion is systematically related to search behavior, while for subjects from P^C, risk aversion is systematically related to search behavior. Specific hypotheses are stated below:

Conditional on the assumption that a population P^R of subjects uses the rp-model, the rp-model predicts that:
(H1) The more loss averse – measured as λ_x in part A – a subject from P^R is, the fewer search steps (denoted by *ss*) this subject should do in the search task.
(H2) For subjects from P^R, the index of loss aversion λ_x – elicited in part A – should be positively correlated with the index of loss aversion, λ^{search}, elicited in the search task, part B.
Conditional on the assumption that a population P^C of subjects does not use the rp-model but one of the classical models (either the forward optimal search model or the backward optimal search model), it is claimed that:
(H3) The more risk averse – measured as α and γ in the preference elicitation part C – a subject from P^C is, the fewer steps *ss* this subject should do in the search task.
(H4) For subjects from P^C, the indices of risk attitude – measured as α and γ in the preference elicitation part C – should be positively correlated with the particular indices of risk attitude γ^{search} and α^{search}, respectively, revealed through the search behavior.

In the remainder of this section, I study the correlation between preference parameters and search parameters in the sample. Then, I develop a duration model that investigates which of the risk and loss aversion measures has better explanatory power for the observed search duration. But before these analyses, it is helpful to compare descriptive statistics on preference estimates (see table 1 and table 2) with the theoretical findings on the relationship between preference parameters and search behavior (see section 4.3). This gives a first impression of the relationship between the empirical findings and the theory.
Risk attitude (CRRA-specification): Table 1 shows that all estimates for α lie in the interval $[-0.457, 2.345]$. From the developed search models follows that these estimates correspond to reservation price paths that start between €83 (for $\alpha = 2.345$) and €87 (for $\alpha = -0.457$). That is, essentially only the following search rules are compatible with the preference estimates: $\{tX_{83}, ..., tX_{87}\}$ for $X \in \{1, 2\}$.
Risk attitude (CARA-specification): Table 1 shows that all estimates for γ lie in the in-

terval $[-0.153, 0.093]$. These estimates correspond to reservation price paths that start between €84 (for $\gamma = -0.153$) and €87 (for $\gamma = 0.093$). That is, only the following search rules are relevant: $\{tX_{84}, ..., tX_{87}\}$ for $X \in \{1, 2\}$.
Loss attitude: The estimated λ_x-values lie in the interval $[0.5, 20]$, see table 2. This corresponds to reservation price paths that start between €83 (for $\lambda = 0.5$) and €98 (for $\lambda = 20$).[23] From the universe of search rules, the following rules apply: $\{tX_{83}, ..., tX_{94}\}$ for $X = 3$.

The first finding from this descriptive analysis is: The variance in the degree of curvature of the utility function is not sufficient to explain the heterogeneity in the observed search behavior. As figure 6 (see appendix) suggests and I have already argued, the complete universe of search rules is needed but also sufficient to describe the search behavior of all observed individuals. The second finding from the descriptive analysis is that although the estimated loss aversion coefficients are generally compatible with a wider range of different search rules than the estimated risk aversion coefficients, the variation observed in loss aversion is also not sufficient to capture the observed heterogeneity in search behavior.

Correlation Analysis

Table 5 reports the results of an investigation of the above mentioned hypotheses (H1)-(H4) based on a rank correlation analysis between observed preference and search parameters. A clear pattern emerges: For the complete sample P, there are negative correlations of marginal significance between most estimates for individual loss aversion and the number of search steps (ss); this is consistent with (H1). In contrast, the estimates for individual risk attitude are not correlated with the number of search steps.
For the subgroup P^R, I find strong support for (H1): There are significantly negative correlations between all estimates for individual loss aversion and the number of search steps (ss). Additionally, results from these analyses support (H2): The estimates for individual loss aversion derived from the lottery questions, λ_x, and the estimates derived from the observed search behavior, λ^{search}, are correlated at the 10%-level (λ_1 and λ_{10}), or even at the 5%-level (λ_{25}, λ_{50}, and λ_{100}). For P^C, no significant correlations are found, suggesting that none of the hypotheses (H3) and (H4) for this group is supported. The hypotheses (H3) and (H4) are not supported by any of the considered subgroups either.

[23] It was always the same subject who is responsible for the maximum value for λ in all five cases (see table 2). Without this subject, reservation price paths between €83 and €95 would correspond to all estimated λ_x-values.

Table 5: Spearman correlations between preferences and search parameters for the (sub)samples.

[P (106 individuals)]

	Search steps (ss)		$\lambda^{\mathbf{search}}$		$\alpha^{\mathbf{search}}$		$\gamma^{\mathbf{search}}$	
	ρ	**p-value**	ρ	**p-value**	ρ	**p-value**	ρ	**p-value**
$\lambda_{\mathbf{1}}$	-0.10	0.29	0.04	0.65				
$\lambda_{\mathbf{10}}$	-0.17	0.08	0.11	0.28				
$\lambda_{\mathbf{25}}$	-0.17	0.08	0.08	0.39				
$\lambda_{\mathbf{50}}$	-0.16	0.10	0.10	0.29				
$\lambda_{\mathbf{100}}$	-0.16	0.10	0.11	0.28				
α	-0.02	0.87			0.03	0.78		
γ	-0.01	0.88					0.06	0.63

[P^R (39 individuals)]

	Search steps (ss)		$\lambda^{\mathbf{search}}$		$\alpha^{\mathbf{search}}$		$\gamma^{\mathbf{search}}$	
	ρ	**p-value**	ρ	**p-value**	ρ	**p-value**	ρ	**p-value**
$\lambda_{\mathbf{1}}$	-0.32	0.05	0.28	0.09				
$\lambda_{\mathbf{10}}$	-0.40	0.01	0.30	0.06				
$\lambda_{\mathbf{25}}$	-0.40	0.01	0.30	0.05				
$\lambda_{\mathbf{50}}$	-0.38	0.02	0.32	0.05				
$\lambda_{\mathbf{100}}$	-0.41	0.01	0.33	0.04				
α	-0.10	0.56			0.00	1.00		
γ	0.10	0.56					0.00	0.99

[P^C (67 individuals)]

	Search steps (ss)		$\lambda^{\mathbf{search}}$		$\alpha^{\mathbf{search}}$		$\gamma^{\mathbf{search}}$	
	ρ	**p-value**	ρ	**p-value**	ρ	**p-value**	ρ	**p-value**
$\lambda_{\mathbf{1}}$	0.03	0.80	-0.09	0.48				
$\lambda_{\mathbf{10}}$	-0.03	0.82	-0.01	0.95				
$\lambda_{\mathbf{25}}$	-0.04	0.77	-0.04	0.75				
$\lambda_{\mathbf{50}}$	-0.03	0.83	-0.01	0.91				
$\lambda_{\mathbf{100}}$	-0.02	0.89	-0.02	0.90				
α	0.07	0.55			0.03	0.81		
γ	-0.07	0.57					0.00	1.00

Duration Analysis

A further analysis of the relationship between individual preferences and search duration controls for the simultaneous influence of risk and loss attitude on search behavior. Furthermore, the inclusion of unobserved effects for each observed search round captures possible behavioral differences that stem from the particular sequence of price draws that the subjects face in each single round. The analysis also exploits the discrete time-to-event-nature and multiple-spell-nature: The event is the stopping of the search, the duration is measured discretely as the number of search steps, and 15 spells (= search rounds) per subject were observed.

For one specific search round, let $T \geq 1$ denote the search duration that has some distribution in the population. From the distribution function of T, I derive the hazard function $h_0(t)$ for T. The discrete time hazard gives the probability of stopping the search in the next time step, conditional on not having stopped so far:

$$h_0(t) = P(T = t \mid T \geq t) \tag{22}$$

Assuming that the subjects in the population use a constant reservation price rule, the hazard function $h_0(t)$ is constant. That is, the stopping events are generated from a process without memory and $h_0(t) = h_0$, leading to a geometric duration distribution.[24] To account for the finite horizon nature of the search problem (i.e., subjects stop their search in time step 25 if they have not been successful until then), a piecewise constant hazard function is used:

$$h_0(t) = \begin{cases} h_1 & t < 25 \\ h_2 & t = 25. \end{cases} \tag{23}$$

To investigate the hypotheses derived above, I test whether the hazard, i.e. the conditional probability of stopping in the next time step, can be explained by individual preference parameters. Therefore, two covariates X are used in the hazard function: one covariate that characterizes risk attitude (α or γ) and one covariate characterizing loss attitude ($\lambda_1, \lambda_{10}, \lambda_{25}, \lambda_{50}$, or λ_{100}). The idea of a proportional hazard is adopted (i.e., the conditional individual probability of stopping the search differs proportionately based on a function of the covariates). For discrete time data, this leads to the complementary log-logistic model (Clayton and Hills, 1993) and the discrete time hazard can be written as:

$$h_i(t, X) = 1 - exp[-exp(\beta' X_i + \delta_1 h_1 + \delta_2 h_2)], \tag{24}$$

where, i = 1,..., 106. β is a parameter vector, h_1 and h_2 characterize the baseline hazard.

[24] The assumption of a constant hazard can be motivated based on theoretical deliberations and based on the empirical finding that non-constant reservation price paths did not perform significantly better than constant reservation price paths in the classification procedure. Please see the appendix, section 8.4, for a discussion of the constant hazard assumption.

Table 6: Duration analysis. Estimation results for various preference specifications and (sub)samples. I use two covariates in each duration regression: One covariate for loss attitude ($\lambda_1, \lambda_{10}, \lambda_{25}, \lambda_{50}$, or λ_{100}) and one covariate for risk attitude (α or γ). That is, for each considered sample, 10 duration regressions are presented.

[P (106 individuals)]

CRRA			**CARA**		
Regressor	**Coefficient**	**p-value**	**Regressor**	**Coefficient**	**p-value**
$\lambda_{\mathbf{1}}$	0.02	0.08	$\lambda_{\mathbf{1}}$	0.02	0.07
α	0.02	0.73	γ	-0.59	0.34
$\lambda_{\mathbf{10}}$	0.02	0.04	$\lambda_{\mathbf{10}}$	0.02	0.03
α	0.04	0.57	γ	-0.76	0.23
$\lambda_{\mathbf{25}}$	0.02	0.03	$\lambda_{\mathbf{25}}$	0.02	0.02
α	0.02	0.75	γ	-0.69	0.28
$\lambda_{\mathbf{50}}$	0.02	0.04	$\lambda_{\mathbf{50}}$	0.03	0.02
α	0.02	0.77	γ	-0.68	0.28
$\lambda_{\mathbf{100}}$	0.02	0.12	$\lambda_{\mathbf{100}}$	0.02	0.11
α	0.02	0.74	γ	-0.58	0.35

[P^R (39 individuals)]

CRRA			**CARA**		
Regressor	**Coefficient**	**p-value**	**Regressor**	**Coefficient**	**p-value**
$\lambda_{\mathbf{1}}$	0.07	0.01	$\lambda_{\mathbf{1}}$	0.07	0.01
α	0.17	0.28	γ	-1.67	0.15
$\lambda_{\mathbf{10}}$	0.09	0.00	$\lambda_{\mathbf{10}}$	0.09	0.00
α	0.18	0.26	γ	-1.80	0.12
$\lambda_{\mathbf{25}}$	0.14	0.00	$\lambda_{\mathbf{25}}$	0.14	0.00
α	0.20	0.21	γ	-1.91	0.11
$\lambda_{\mathbf{50}}$	0.09	0.01	$\lambda_{\mathbf{50}}$	0.09	0.01
α	0.15	0.35	γ	-1.61	0.16
$\lambda_{\mathbf{100}}$	0.04	0.09	$\lambda_{\mathbf{100}}$	0.04	0.08
α	0.18	0.26	γ	-1.72	0.14

[P^C (67 individuals)]

CRRA			CARA		
Regressor	**Coefficient**	**p-value**	**Regressor**	**Coefficient**	**p-value**
$\boldsymbol{\lambda_1}$	0.02	0.17	$\boldsymbol{\lambda_1}$	0.02	0.18
α	-0.02	0.83	γ	0.11	0.89
$\boldsymbol{\lambda_{10}}$	0.02	0.17	$\boldsymbol{\lambda_{10}}$	0.02	0.18
α	0.00	1.00	γ	-0.09	0.91
$\boldsymbol{\lambda_{25}}$	0.02	0.09	$\boldsymbol{\lambda_{25}}$	0.02	0.10
α	-0.01	0.94	γ	-0.04	0.96
$\boldsymbol{\lambda_{50}}$	0.02	0.08	$\boldsymbol{\lambda_{50}}$	0.02	0.08
α	-0.01	0.94	γ	-0.06	0.94
$\boldsymbol{\lambda_{100}}$	0.03	0.12	$\boldsymbol{\lambda_{100}}$	0.03	0.13
α	-0.01	0.86	γ	0.10	0.90

Now, recall that every subject had to play 15 search rounds. All prices were drawn from a uniform distribution. The series of price draws are different across rounds but they are identical across individuals. Therefore, I expect an unobserved effect for each search round. To account for this unobserved heterogeneity, a random effect that is common to all observations from a certain search round j ($j = 1, ..., 15$) is included. The following model is considered:

$$h_{i,j}(t, X) = 1 - exp[-exp(\beta' X_i + \delta_1 h_1 + \delta_2 h_2 + u_j)] \tag{25}$$

where u_j is supposed to be normally distributed with mean zero.

Table 6 presents estimation results for the complete sample and for the subgroups. In all estimations a likelihood ratio test suggests that the included unobserved effect is highly statistically significant. For the complete sample of subjects, P, (H1) is supported: An increase in individual loss aversion is related to a significant increase in the conditional probability of stopping the search, i.e. to a decrease in search duration. This effect is significant at the 5%-level for λ_{10}, λ_{25}, and λ_{50}; it is marginally significant for λ_1, regardless of the specification of the risk attitude coefficient (α or γ). In all specifications, risk attitude has no significant explanatory power for search duration.
Considering the subsample P^R, even stronger support for (H1) is found: Apart from λ_{100}, all estimates for individual loss aversion have explanatory power for search duration at least at the 2%-significance level. Again, individual risk attitude is always insignificant.
In the subgroup P^C, no preference parameter has significant explanatory power. I have performed all analyses presented in table 6 including higher order terms of the risk and loss aversion measures in order to test for possible nonlinear relationships. The same conclusions as those reported above are obtained. While the risk attitude terms are never

jointly significant, the loss attitude terms are jointly significant in the same cases as reported in table 6.[25]

According to the analyses presented above, all estimated loss attitude coefficients have better explanatory power for individual search behavior than the estimated risk attitude coefficients, which are insignificant in all analyses, regardless of whether CARA- or CRRA-specifications of the utility function are considered. The findings concerning the explanatory power of loss aversion do not hold for the subgroup P^C, but they are very strong for the subgroup P^R, suggesting that members of the two groups behave differently when "solving" the search task.

6 Discussion

This paper focuses on the development and experimental testing of various search models, in particular the reference point model (rp-model). The results suggest that the rp-model is similar to EU-based models in its predictions about reservation price paths, but it is better than EU-based models in reconciling the experimental data on individual preferences with the data on individual search behavior. Combined with established empirical results on individual preferences (such as the empirical distribution of loss aversion in a population, see, e.g., Johnson et al., 2006; Pennings and Smidts, 2003; Schmidt and Traub, 2002; Tversky and Kahneman, 1992), the rp-model is consistent with existing findings on search behavior, for example the large heterogeneity of search rules and the predominance of early-stopping in the population (Cox and Oaxaca, 1989; Hey, 1987; Sonnemans, 1998).

To further investigate individual heterogeneity, I hypothesize that at least a specific subgroup, P^R, of the subjects uses the proposed rp-model. Since identification of this subgroup – merely based on the observed search behavior of subjects – is in practice not possible, the subgroup P^R is identified with the help of the search-model question. Under the assumption that subjects understand this question correctly and that they are able to relate this question to their actual search behavior, the question is likely to be an instrument for dividing the complete sample into the particular subgroups P^R and P^C. The main empirical result of this paper – namely that individual loss aversion is systematically related to search behavior, whereas risk aversion is not related to search behavior – is independent of this search-model question. Nevertheless, all specific results concerning the subgroup differences in search behavior are built on the assumption that information obtained in this question is valid.

Several issues have to be kept in mind when interpreting the results from this study. First, the presented experimental setup is based on one specific search environment which is

[25] The robustness of the results from the duration analysis has been checked further. Please see the appendix, section 8.4, for a brief discussion of different specifications of the duration model.

characterized by the price distribution, the search cost, and the option to recall previously rejected offers. It is conceivable that subjects behave differently in a different search environment (e.g., in an environment, in which the price distribution is not known). In particular, the effect of loss aversion on search behavior might become more or less pronounced if higher losses (i.e., higher search cost) are involved. In the context of a search environment with a considerably longer time between the search steps, the observed effect of loss aversion might rather be called an endowment effect (Huck et al., 2005; Kahneman et al., 1991): If a person holds an object that she may keep for sure for a certain time, she might consider this object as an endowment, and the next step in the search is evaluated only relative to this endowment.

A further issue to be addressed is the recall option of the search task. To my knowledge, no search model or rule that explicitly predicts the recall option has been investigated in the literature so far; the rp-model is not able to predict recall decisions either. Indeed, 2.4% of all stop-or-go decisions in the sample are decisions to stop and recall a price that had been rejected in an earlier search step.

Finally, a limitation is the linearity assumption of the reference point model. This assumption was introduced in order to reduce the reference point model to one underlying preference parameter which can be identified with a standard experimental lottery method. While this assumption is critical since it essentially "assumes away" that the curvatures of $u^+(\cdot)$ and $u^-(\cdot)$ are a potential explanation for the observed effect, the relaxation of this assumption would not compromise the key empirical finding of this study, namely that the observed tradeoff between gains and losses – measured at five different values – is related to search characteristics such as search duration. However, a design that is preferable from a theoretical point of view would avoid relying on measures of loss aversion, and would instead elicit the complete utility function (i.e., on gains and losses), estimate an appropriate functional form (that does *not* impose risk neutrality) for every individual, and then investigate the relationship between each individuals' utility function and her decision rule, based on a reference point model that does not impose risk neutrality. Ideally, in a nested modeling framework, the question would then be whether the combination of a gain and loss utility function ($u^+(\cdot)$ and $u^-(\cdot)$) is better able to explain observed search behavior than the restriction to only a utility function on gains, $u^+(\cdot)$. Of course, this would require a more involved econometric methodology, as well as a more complicated experimental design. The separate elicitation of a utility function on gains and on losses is *not* sufficient, since this does not allow for estimating the tradeoff between gains and losses; subjects would have to answer many more lottery questions than in the current experimental design, both mixed lotteries as well as pure lotteries in the gain and in the loss domain.

A further way to improve in particular the statistical methodology of the paper is to model the joint distribution of individual preference parameters (estimated in the lottery part) and decision rules (obtained from the search task) in a full maximum likelihood framework; this is one way to put more structure on the analysis and to be able to account for decision errors.

It remains unclear whether these modifications in the experimental design and the statistical methodology would substantially alter the key finding from this study: Five measures that capture the tradeoff between gain and loss outcomes in a lottery are significantly related to characteristics of human search behavior, although the design of the search experiment prevented the subjects from making any loss during the search task. In contrast, measures of the tradeoff between lotteries that involve only gains are not related to characteristics of search behavior.

In sum, the analyses lend support to the claim that subjects use different strategies when "solving" the search task. The contribution of this paper is to combine information on sequential decision behavior and on individual preferences and to investigate correlations at the subject level. I do not find evidence supporting the classical EU-based search models in the sample. Controlling for the effect of risk attitude, I do, however, obtain support for the hypothesis that loss aversion is related to search behavior.

There are two principal explanations for these findings: First, loss aversion could simply be measured better than risk aversion, and would therefore be a better explanatory variable. This explanation is in line with Rabin (2000), who provides theoretical arguments that loss aversion can account better for observed decision behavior over modest stakes than the standard notion of risk aversion. A second explanation, which is in line with results on myopic loss aversion in dynamic decision tasks, such as stock market decisions (e.g., Benartzi and Thaler, 1995), is that people set reference points during their search. It is important to note that no *direct* empirical support for the specific reference point updating assumption that underlies the presented form of the reference point model can be found. The observed relationship between loss aversion and search behavior is also consistent with subjects that set one constant utility reference point at a payoff higher than €0 and evaluate the outcomes during the search relative to this fixed reference point. This explanation cannot be rejected by the experimental data, but from a behavioral point of view there is no argument that would favor this model over the reference point model that is developed in this paper.

7 Conclusions

Subjects are heterogeneous with respect to their sequential decision behavior. Using data obtained in a controlled laboratory experiment that involved a search task as a simple representative of sequential decision situations, I have shown that – to some extent – this heterogeneity can be linked to heterogeneity in individual preferences.

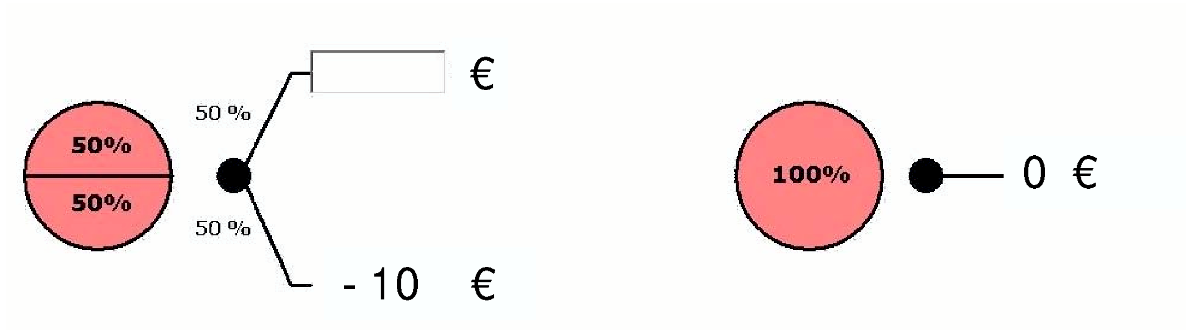
Considering the entire sample, I found that the answers to simple lottery questions that involve *gains and losses*, combined with a standard parametric assumption for the identification of a loss aversion index, had significant explaining power for observed sequential decision behavior. In contrast, the answers to lottery questions that involve *only gains* had no explaining power – regardless of whether a CARA or a CRRA specification was assumed. I consider this as evidence that loss aversion was systematically related to search behavior, while risk attitude was not related to search behavior. The experimental design idea of the test for loss aversion was straightforward: Subjects were not given a show-up fee such that no reference point was induced and all outcomes of the sequential decision task were only realized in the positive domain, i.e. the experience of losses was avoided. Given this design, the findings suggest that people set utility reference points during their search relative to which they evaluate potential future outcomes. Overall, the proposed reference point search model describes observed behavior better than search models derived from expected utility theory.
To further investigate heterogeneity in search behavior, the answer to a question on search behavior was used as an instrument to subdivide the sample into two subgroups. It is found that for the subjects from one subgroup, P^C, there was no relationship between individual preferences and search behavior. However, for the other subgroup, P^R, individual preferences and search behavior were strongly related in a way that is consistent with the predictions of the reference point model. This means, in addition to heterogeneity in individual preferences, there might also be heterogeneity in the way people solve the search task: Some people set reference points in sequential decision tasks, while others do not set reference points. The two subgroups of the sample use different models for solving the search task, and with the help of an instrument to separate these two subgroups, individual search behavior is predictable to a certain degree, provided that information on individual preferences, specifically on loss aversion, is available.

I have discussed that an alternative explanation for the observed effect is that loss aversion can simply be measured better than risk aversion. However, the fact that about a third of the subjects answered to the search-model question that they think about losses when doing the search task as well as findings in other studies about decision-making in dynamic choice tasks (Benartzi and Thaler, 1995; Odean, 1998), are further evidence that loss aversion and reference point setting might indeed play a role for many people in sequential decision situations. The finding that people set reference points in sequential decision tasks is of interest for recent theoretical and applied research in many fields, e.g., decision theory, marketing science (Zwick et al., 2003), labor economics (Eckstein and van den Berg, 2006), and finance (Gneezy, 2003).

8 Appendix

8.1 Graphical Presentation of the Lotteries on the Computer Screen

Figure 5: Graphical presentation of the lotteries on the computer screen.



8.2 On the Function $p(x, m_t, t)$ in the RP-Model

The form of the rp-payoff-function $p(x, m_t, t)$ becomes clear under a rigorous case differentiation with respect to possible price draws. $q(\cdot)$ and $v(\cdot)$ are defined as in section 2.3, i.e.

$$q(y) = \begin{cases} q(y) = y & y \geq 0 \\ 0 & y < 0 \end{cases} \tag{26}$$

$$v(y) = \begin{cases} v(y) = y & y \geq -c \\ 0 & y < -c \end{cases} \tag{27}$$

The following cases are possible:

Case 1

The price draw is better than the best price in hand minus the search cost: $x < m_t - c$

- $m_t \geq 100 - c \cdot t$
  $\Rightarrow p(x, m_t, t) = 100 - c \cdot t - x - c = q(100 - c \cdot t - x - c)$
- $m_t < 100 - c \cdot t$
  $\Rightarrow p(x, m_t, t) = m_t - x - c = v(m_t - x - c)$

Case 2

The price draw is worse than the best price in hand minus the search cost: $x \geq m_t - c$

- $m_t \geq 100 - c \cdot t$
  $\Rightarrow p(x, m_t, t) = 0 = q(100 - c \cdot t - x - c)$

- $m_t < 100 - c \cdot t$

  - $m_t - c \leq x \leq m_t$
    $\Rightarrow p(x, m_t, t) = m_t - x - c = v(m_t - x - c)$

  - $m_t < x$
    - $m_t \leq 100 - c \cdot t - c$
      $\Rightarrow p(x, m_t, t) = -c = v(m_t - (100 - c \cdot t))$

    - $m_t > 100 - c \cdot t - c$
      $\Rightarrow p(x, m_t, t) = m_t - (100 - c \cdot t) = v(m_t - (100 - c \cdot t))$

8.3 Details on Search Behavior

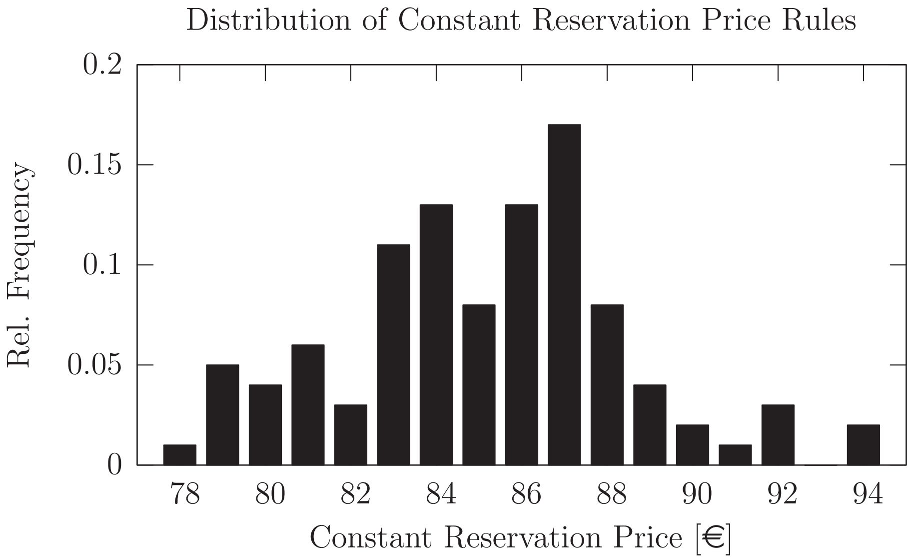
Descriptive Findings

In total, 8532 stop-or-go-decisions were observed in the experiment. The mean number of search steps for all 15 search rounds was 80.5, with a minimum of 49 steps, a maximum of 135 steps and a standard deviation of 18.1 steps. The mean number of search steps per search round was 5.4, with a minimum of 1, a maximum of 25 and with a standard deviation of 3.4 steps. The mean number of search steps was significantly lower than the expected number of search rounds under the assumption of risk neutrality: The expected number of search rounds for an individual that uses the forward optimal search rule (i.e., a constant reservation price of €86) is 6.3 steps. Under a finite horizon model, 7.2 steps are expected. Figure 6 shows the distribution of constant reservation price rules in the sample, conditional on the assumption that all subjects use such a rule.

Classification of Search Behavior

This brief section presents some results of the classification procedure:
If the universe of search rules is limited to the 17 type-1-rules – the constant reservation price rules – 92.8% of all observed stop-or-go-decisions can be explained. When limited to the type-2-rules, 93.0% are explained under the CARA-specification and 92.7% under the CRRA-specification. Finally, the type-3-rules explain 92.8% of all decisions. Under the CARA-specification, all 3 decision rules (type-1, type-2, and type-3) explain observed behavior equally well for 83 (78%) of the subjects (i.e., I cannot discriminate between the 3 rule types for 83 subjects). Under the CRRA-specification, all 3 decision rules (type-1, type-2, and type 3) explain observed behavior equally well for 89 (84%) of the subjects.
In this context, it is important to note that the main purpose of the classification method is *not* to determine a minimal universe of decision rules that best describes the behavior of all subjects in the sample but to estimate the preference parameters that best describe

Figure 6: Imposing a constant reservation price rule on every subject, I obtain the following distribution of constant reservation price rules in the sample. The lowest observed reservation price is €78, the highest reservation price is €94.



observed search behavior. Therefore, the encountered problems of over-fitting, reflected in the lack of discrimination between different search rules, are not a problem for the analysis presented in this paper. In that, the presented method is akin to estimating other preference parameters from experimental data.[26]

The findings presented here have again made clear that it is impossible to attribute search models to the subjects merely based on their revealed search behavior unless we have much more observations per subject; i.e., discrimination *across* search rule types is infeasible. Since I can clearly discriminate *within* a certain rule-type – i.e., I can discriminate between, e.g., rule $t1_p$ and rule $t1_q$ (for $p, q \in \{78, ..., 94\}$ and $p \neq q$) – I am able to attribute preference parameters (risk or loss attitude, depending on the search model) to the subjects.

8.4 On the Duration Analysis

The Assumption of a Constant Hazard

The main motivation for the constant hazard assumption is the finding in section 5.2 and further detailed in the appendix (section 8.3) that a discrimination between the different search rule-types is hardly possible, since all search rules have a similar rate of consistency

[26] Schunk and Winter (2004) use the same classification procedure. More sophisticated statistical methods for the joint determination of the universe of decision rules and the classification of decision rules that allow for errors are used by Houser and Winter (2004) and Houser *et al.* (2004).

with the observed search behavior. It follows that the assumption of a constant reservation price, that is a *type-1-rule*, is generally a good proxy for the observed search behavior. A constant reservation price, in turn, implies a constant hazard in the duration model, as the reservation price path is interpreted as a hazard function in a duration model.
A glance at figures 1, 2, 3, and 4 reveals that all of the rules in the universe of search rules consist of an initial part that has a constant reservation price. What rule is least consistent with the assumption of a constant reservation price that is used for the duration analysis? Figure 3 reveals that if a subject uses a CARA-finite horizon rule *and* if the subject is very risk averse, it might be using the worst rule in terms of consistency with the constant hazard assumption. The subject might then have a reservation price of €94 at $t = 1$ and $t = 2$, and the price starts falling already from $t = 3$ on. The probability that this individual does not search for more than two steps is $1 - (1 - \frac{20}{76})^2 \approx 46\%$. That is, even in this "worst case", the constant hazard assumption is correct in 46% of all cases, and this "worst case" characterizes only very few subjects, as figure 6 reveals.
Since a certain reservation price path in figure 1, 2, 3, or 4 can be interpreted as the hazard function of the particular individual that is using the corresponding search rule, a modeling approach that is nonparametric concerning the individual hazard function would effectively require the identification of reservation price paths. With the data at hand, this is practically impossible without further restrictions on the hazard function, given the identification problems encountered in section 4.3, which stem from the low number of observations per subject.

Robustness

Various alternative specifications for the duration model have been considered:

(a) It is tempting to include a random effect for *each subject* instead of including an effect for *each search round*. In this specification the unobserved effect term is highly insignificant. However, *all* results presented in this paper also hold in this specification, although in some cases they are statistically weaker.
(b) If the unobserved effect is left out from the estimated model, results are obtained that are virtually identical with results that are obtained based on the random effect specification for each subject (see specification (a) above).
(c) The hazard h_1 is highly significant in all estimations, but the drop-out term h_2 for time-step 25 is in general not significant, suggesting a specification without h_2 (i.e., a constant hazard instead of a piecewise constant hazard). All results are very similar to those reported in the paper; the effect of the loss aversion coefficient on search duration is even stronger than in the results reported in the paper.
In sum, the findings from alternative specifications support the conclusions that are drawn in this paper.

References

Abdellaoui, M. (2000): Parameter-free elicitation of utility and probability weighting functions. *Management Science*, 46(11), 1497−1512.

Andersen, S., G. W. Harrison, M. Lau, and E. E. Rutström (2005): Preference heterogeneity in experiments: Comparing the field and lab. *CEBR Discussion Paper No. 2005-03.*

Benartzi, S. and R. H. Thaler (1995): Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, 110, 73−92.

Clayton, D. and M. Hills (1993): *Statistical Models in Epidemiology.* Oxford: Oxford University Press.

Cox, J. C. and R. L. Oaxaca (1989): Laboratory experiments with a finite-horizon job search model. *Journal of Risk and Uncertainty*, 2, 301−330.

Currim, M. and R. K. Sarin (1989): Prospect versus utility. *Management Science*, 35, 22 −41.

Eckstein, Z. and G. J. van den Berg (2006): Empirical labor search: A survey. *Journal of Econometrics*, forthcoming.

Fishburn, P. and G. A. Kochenberger (1979): Two-piece Von Neumann-Morgenstern utility functions. *Decision Sciences*, 10, 503−518.

Gneezy, U. (2003): Updating the reference level: Experimental evidence. *The University of Chicage Graduate School of Business. Working Paper.*

Harrison, G. W., E. Johnson, M. M. McInnes, and E. E. Rutström (2005): Risk aversion and incentive effects: Comment. *American Economic Review*, 95(3), 897−901.

Harrison, G. W. and P. Morgan (1990): Search intensity in experiments. *Economic Journal*, 100, 478−486.

Hey, J. D. (1981): Are optimal search rules reasonable? and vice versa? (and does it matter anyway?). *Journal of Economic Behaviour and Organization*, 2, 47−70.

Hey, J. D. (1982): Search for rules for search. *Journal of Economic Behaviour and Organization*, 3, 65−81.

Hey, J. D. (1987): Still searching. *Journal of Economic Behaviour and Organization*, 8, 137 −144.

Holt, C. A. and S. K. Laury (2002): Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644−1655.

Houser, D., M. Keane, and K. McCabe (2004): Behavior in a dynamic decision problem: An analysis of experimental evidence using a Bayesian type classification algorithm. *Econometrica*, 72(3), 781−822.

Houser, D. and J. Winter (2004): How do behavioral assumptions affect structural inference? Evidence from a laboratory experiment. *Journal of Business and Economic Statistics*, 22(1), 64–79.

Huck, S., G. Kirchsteiger, and J. Oechssler (2005): Learning to like what you have: Explaining the endowment effect. *The Economic Journal*, 115, 689–702.

Johnson, E. J., S. Gächter, and A. Herrmann (2006): Exploring the nature of loss aversion. IZA-Discussion Paper No. 2015, Institute for the Study of Labor.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1991): The endowment effect, loss aversion, and status quo bias: Anomalies. *Journal of Economic Perspectives*, 5, 193–206.

Kahneman, D. and A. Tversky (1979): Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.

Koebberling, V. and P. P. Wakker (2005): An index of loss aversion. *Journal of Economic Theory*, 122, 119–131.

Kogut, C. A. (1990): Consumer search behaviour and sunk costs. *Journal of Economic Behaviour and Organization*, 14, 381–392.

Lippman, S. A. and J. J. McCall (1976): The economics of job search: A survey. *Economic Inquiry*, 14, 155–189.

Odean, T. (1998): Are investors reluctant to realize their losses. *The Journal of Finance*, 53, 1775–1798.

Pennings, J. M. E. and A. Smidts (2000): Assessing the construct validity of risk attitude. *Management Science*, 46(10), 1337–1348.

Pennings, J. M. E. and A. Smidts (2003): The shape of utility functions and organizational behavior. *Management Science*, 49, 1251–1263.

Pratt, J. W. (1964): Risk aversion in the small and in the large. *Econometrica*, 32, 122–136.

Rabin, M. (2000): Risk aversion and expected utility theory: A calibration theorem. *Econometrica*, 68(5), 1281–1292.

Schmidt, U. and S. Traub (2002): An experimental test of loss aversion. *Journal of Risk and Uncertainty*, 25(3), 233–249.

Simon, H. A. (1955): A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69, 99–118.

Sonnemans, J. (1998): Strategies of search. *Journal of Economic Behavior and Organization*, 35, 309–332.

Sonnemans, J. (2000): Decisions and strategies in a sequential search experiment. *Journal of Economic Psychology*, 21, 91–102.

Stigler, G. J. (1961): The economics of information. *Journal of Political Economy*, 69, 213–225.

Tversky, A. and D. Kahneman (1992): Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.

Wakker, P. P. and D. Deneffe (1996): Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Science*, 42, 1131–1150.

Zwick, R., A. Rapoport, A. K. C. Lo, and A. V. Muthukrishnan (2003): Consumer sequential search: Not enough or too much? *Marketing Science*, 22(4), 503–519.

What Determines the Saving Behavior of German Households? An Examination of Saving Motives and Saving Decisions

Abstract: Many motives for saving a portion of one's income co-exist and their relative importance changes over the life-cycle. This paper is concerned with linking heterogeneity in household saving behavior to four co-existing saving motives. First, I find that the importance that households attach to the saving motives is related to how much households save at different life stages. Second, I classify the saver type of the households based on whether they engage in regular savings plans, or rather save irregularly and without a savings plan. I find that saving motives are related to the saver type of the household. The results show that heterogeneity with respect to the saving rate and the saver type – which has been emphasized in recent studies – is systematically related to the importance that households attach to different saving motives. This suggests that policy reforms that substantially change the importance of certain saving motives in the eyes of private households might alter household saving behavior in various ways.

1 Introduction

For a typical household, many different considerations influence saving decisions over the life-cycle. For example, households save to finance consumption after retirement. They save in order to insure against various economic, biometric, and political risks that they are exposed to over the life-cycle. Households might also engage in saving for supporting their children or grandchildren, e.g. during their education, or for leaving a bequest to them. Finally, many households are interested in saving for purchasing real estate at some point in their life. Many of these considerations and circumstances imply explicit saving targets and they require specific forms of saving, such as long-term and planned saving for retirement.

Briefly, various saving motives co-exist over the life-cycle, and different motives might be associated with different forms of saving. Understanding what motives drive saving behavior over different stages of the life-cycle and how the relative contribution of these motives changes over the life-cycle will help us to understand differences in saving rates among households as well as past and future trends in saving behavior. As underlined by various authors (e.g., Börsch-Supan and Lusardi, 2003), this understanding is of utmost policy relevance, since reforms of the social security systems directly interact with household saving as a private insurance. E.g., the currently ongoing reform of the German pension system is essentially concerned with the trade-off between public and private old-age saving: The reform moves the rather monolithical and very generous system that provides almost all retirement income within a single public pay-as-you-go-framework to a three-pillar system, in which private and occupational pensions will have an increasingly important role. Accordingly, the importance of private saving for old age has increased in recent years. Understanding the motives for private saving is also important from the perspective of taxation: For instance, the taxation of bequests and inheritances is non-distortionary if intergenerational transfers are accidental but may have efficiency costs if bequests are intentional (see, e.g., Bernheim, 2002). In summary, private saving is an important determinant of household economic security as well as social and economic well-being.

The point of departure of this paper – the observation that co-existing motives determine saving behavior – is an idea that goes back to Keynes (1936). While there is an extensive body of empirical literature on saving motives, which I review briefly in a later section, only few empirical studies take into account that different saving motives co-exist over the life-cycle; most studies focus on only one motive and make simplifying assumptions about the other motives such that those can be relegated to the background. A consistent finding in the literature is that there is considerable heterogeneity in household saving behavior, a point that is emphasized by numerous authors, e.g., Alessie et al. (1997), Browning and Lusardi (1996), and Kurz (1985). In addition, many studies recognize explicitly that the contributions of saving motives to household saving might change over the life-cycle (e.g., Horioka and Wanatabe, 1997; Kennickell and Lusardi, 2005). So far, however, there has been only little interest in the investigation of how co-existing saving motives whose contribution might change over stages of the life-cycle help to explain the observed *heterogeneity in how much households save*. This shortcoming is criticized by, e.g., Alessie and Lusardi (1997), Samwick (2006) and Wärneryd (1999, p. 264). Furthermore, extending the vast literature that seeks to explain how much households save, a recently emerging literature emphasizes *heterogeneity in the extent to which households plan their saving or choose specific forms of saving*, *such as savings plans* (e.g., Ameriks et al., 2003; Lusardi and Mitchell, 2006; Sourdin, 2005). While Ameriks et al. (2003) relate heterogeneity in the propensity to plan to the general household budgeting behavior as well to a household's general attitudes and skills, Lusardi and Mitchell (2006) and Sourdin (2005) focus on planning and old-age provision. Overall, recent findings, obtained from studies that mostly focus on one specific saving motive, suggest that the heterogeneity in household socio-economic characteristics, in household preferences, and in household saving motives is associated with heterogeneity in saving behavior with respect to two – not necessarily independent – dimensions, namely how much households save and whether they plan their saving.

This paper focuses on the question to what extent heterogeneity in saving behavior can be explained by the importance that households attach to four potentially co-existing saving motives: The old-age provision motive, the precautionary motive, the bequest motive, and the motive to purchase a house (henceforth: housing motive). The paper finds that the importance attached to certain saving motives is related to heterogeneity in each of the two dimensions of saving behavior. More specifically, the paper first estimates the relationship between the saving motives and the saving rate. I find that information on saving motives is related to the household saving rate, and that the relative contribution of the saving motives to household saving changes over age classes. Second, the paper investigates whether saving motives help to explain what type of savers households are, e.g., whether they engage in regular savings plans, or rather save irregularly and without a savings plan. I find evidence for a relationship between the information on certain saving motives and the saver type of the households, i.e. the households' propensity to plan their saving.

To identify which of the saving motives are operative I use explicit data, i.e. answers to survey questions about the importance that households attach to the considered saving motives, henceforth referred to as "subjective" measures.[1] On the one hand, subjective measures can generally be criticized for being more prone to misreporting than other measures, for instance in the case when certain answers are socially desired. Furthermore, in the specific context of this paper the reported saving motives themselves can cause estimation bias since they are endogenous to the saving behavior of households. On the other hand, the subjective measures used in this study have considerable advantages: First, their cognitive burden is very low and the item nonresponse rate is negligible. Second, they provide an alternative way to measure the strength of the precautionary motive, which does not restrict attention to income risk only – a limitation in existing studies of precautionary saving that is criticized in the literature (e.g., Hurst et al., 2005).[2] The subjective measure for the strength of the precautionary motive that is considered in this study includes other risks, such as health risks, longevity risk, and interest rate risk.[3]

Overall, the paper presents empirical evidence that the importance that households attach to various saving motives is associated with observed saving behavior. The findings suggest that policy reforms that substantially change the importance of certain saving motives in the eyes of private households might indeed alter household saving behavior in various ways and with differential effects over households' life stages.

---

[1] Subjective data on saving motives have been used in existing studies. For example, information about bequest intentions has been used to learn about the existence of a bequest motive in studies by Alessie et al. (1999), Jürges (2001), and by Mirer (1979); Essig (2005b) studies the precautionary motive. Alessie and Kapteyn (2001) provide a detailed discussion about the usefulness of subjective data in research on saving behavior.

[2] The theory of precautionary saving predicts that households with higher income risk have higher accumulation, and most studies investigate the relationship between a measure for income risk and a stock or flow measure of saving without considering or controlling for other sources of risk. Essig (2005b), Palumbo (1999), Carroll and Samwick (1997), and Kennickell and Lusardi (2005) are exceptions, they consider further sources of risk.

[3] Theoretical studies have shown the relevance of these risks for savings behavior, see, e.g., Yaari (1965) and Leung (1994) for uncertainty about lifetime. Palumbo (1999) presents a theoretical model that includes uncertainty about medical expenses – i.e., health risks – estimates its parameters based on data from the U.S. Panel Study of Income Dynamics (PSID), and finds that uncertain medical expenses represent an important motive for precautionary saving.

The remainder of this paper is structured as follows: Section 2 provides information on the data, describes how the principal variables used in this study are measured, and presents basic descriptive statistics. Section 3 provides an overview of studies on saving motives and saving behavior and relates the current paper and its empirical framework to the existing literature. In section 4, the empirical analysis of the relationship between information on saving motives, the saving rate, and household saver types is presented, and the findings are discussed. Section 5 concludes.

2 Data and Descriptive Statistics

2.1 The SAVE Survey

2.1.1 Overview

Departing from the Dutch CentER Panel and the U.S. Health and Retirement Study (HRS) as an example, researchers of the University of Mannheim have cooperated with the Mannheim Center for Surveys, Methods and Analyses (ZUMA), TNS Infratest (Munich), Psychonomics (Cologne) and Sinus (Heidelberg) to produce a questionnaire on households' saving and asset choice. The SAVE dataset records detailed information on both, financial variables such as income, saving, and asset holdings as well as on sociological and psychological characteristics of households. Great care was taken that the interviewer talks to the member of the household who knows about income, wealth and saving behavior whom we henceforth refer to as the household head.

2.1.2 The Sample

A first wave of the SAVE study, which was based on quota sampling, was fielded in the summer of 2001. The findings from this study were used to investigate the impact of different survey modes on response behavior (see Essig and Winter, 2003). The next wave benefited from the methodological findings of the 2001 wave and was conducted in summer 2003. The 2003 wave, which is used for the analysis presented in this paper, is a random sample of 2184 households.

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 (Bundeslaender), districts and community types. Further sampling details are presented in Heien and Kortmann (2003).

2.1.3 Data-Quality, Item Nonresponse, and Multiple Imputation

Essig (2005a) discusses various methodological aspects of the SAVE dataset, in particular the questionnaire, interviewer and interviewee motivation, and the representativeness of the survey. He compares the 2003 random sample and the German microcensus 2002 with respect to the distributions of age, household net income, and household size, and he concludes that the SAVE random sample "fits the German microcensus extremely well". He also confirms that various financial measures, such as income and financial wealth, are in line with findings from a related German survey, the German Socio-Economic Panel 2003 (GSOEP). Finally, Essig concludes that unit and item nonresponse rates are very similar to related other surveys in Germany or other countries.

Table 1: Demographic characteristics of the random sample of 2184 households.

Characteristic	(%)
Age	
18-34	21.4
35-49	29.7
50-64	23.7
65+	25.2
Marital Status	
Currently married	59.7
Previously married	20.9
Not married	19.4
Education	
Haupt-/Volksschule or below	40.9
Mittlere Reife, Fachhochschulreife	37.8
Allgemeine/fachgebundene Hochschulreife	21.3
Employment Status	
Retired	35.2
Blue collar	16.0
White collar	22.6
Civil servant	4.2
Self-employed	6.0
Unemployed	7.0
Education/Apprenticeship/Military service/Parental leave	9.0
Number of children	
0	24.5
1	22.0
2	32.2
3	13.4
4+	7.9

Item nonresponse to sensitive questions about household financial circumstances is documented and discussed in Essig and Winter (2003) and in Schunk (2006). To prevent biased inference based on an analysis of only complete cases, an iterative multiple imputation procedure has been applied to the SAVE data (Schunk, 2006). Multiple imputation simulates the distribution of missing data and allows for a more realistic assessment of variances in subsequent analyses than single imputation. The procedure uses a Markov-Chain Monte-Carlo method to replace missing data by draws from an estimate of the conditional distribution of the data. The multiple imputation algorithm generates five data sets with all missing values replaced by imputed values. For all descriptive statistics and all estimation results presented in this paper, the five imputed datasets are analyzed separately, and the results of the five analyses are then combined based on methods derived by Rubin (1987). The use of these methods assures that the missing data uncertainty is reflected in all findings presented in this paper.

2.2 Basic Demographic Characteristics

Table 1 shows basic demographic characteristics of the households in the 2003 random sample. Statistics concerning the age, marital status, number of children, education, and employment status of the household head are tabulated. Table 1 and all other statistics and estimations presented in this paper are not weighted.

2.3 Measuring Household Saving Behavior

2.3.1 Saving Motives

The SAVE survey asks directly about saving motives. Households are asked how important they rate the considered saving motives in their own view. Each reason for saving has to be rated on a scale from 0 ("*of absolutely no importance*") to 10 ("*of highest importance*"). To mitigate interpersonal differences in the response behavior to this question, a common approach is to classify the answers on a more coarse symmetric scale: All answers from 0 to 3 are in the lowest category (which I denote as "*unimportant*"), answers from 4 to 6 are in the middle category ("*important*"), and answers from 7 to 10 are in the highest category ("*very important*").

Table 2 shows the distribution of the answers across the four age classes that are considered in this study. Many households rate "saving as a precaution" and "saving for old age" as very important motives, whereas the bequest and the housing motive are overall of much less importance in all age classes.[4] These findings are in line with findings in Alessie et al. (1999) which are based on an analysis of binary measured saving motives.

Table 2: Descriptive statistics on the question about households' saving motives.

		Old-age provision motive			Precautionary motive		
		(1)	(2)	(3)	(1)	(2)	(3)
	All	22%	19%	59%	14%	24%	62%
Age	<35	20%	21%	59%	15%	25%	60%
Age	35-49	14%	20%	66%	11%	27%	62%
Age	50-64	20%	13%	67%	14%	22%	64%
Age	≥65	35%	20%	45%	18%	21%	61%

		Bequest motive			Housing motive		
		(1)	(2)	(3)	(1)	(2)	(3)
	All	49%	31%	20%	54%	10%	36%
Age	<35	54%	26%	20%	34%	18%	48%
Age	35-49	43%	38%	19%	48%	11%	41%
Age	50-64	53%	30%	17%	61%	8%	31%
Age	≥65	50%	28%	22%	71%	5%	24%

Note: (1) Unimportant, (2) Important, (3) Very important.

2.3.2 Annual Saving

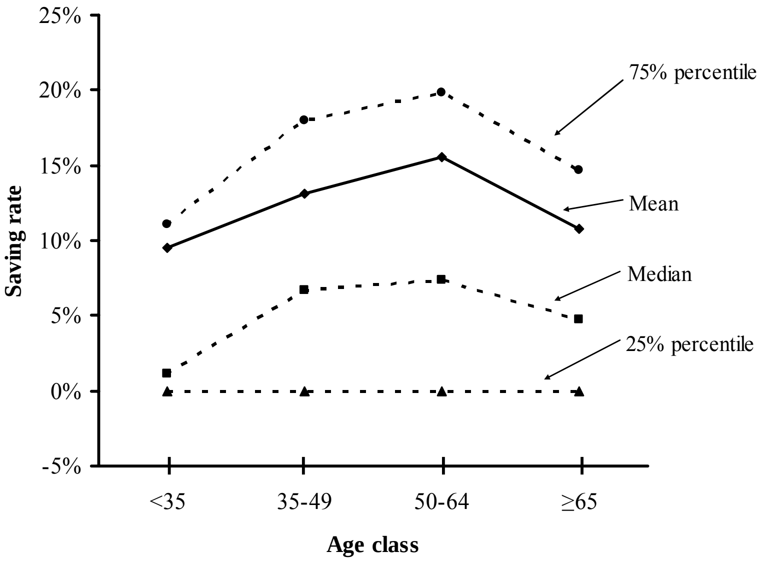
After a number of questions that introduce to household finances and saving, respondents are directly asked for their saving in the previous year 2002 ("*Can you tell me how much money you and your partner saved in total in the year 2002*?"). Households that did not have any positive saving marked that they had zero saving or dipped into their saving; i.e., the answers are left-censored at zero. Repayments of all recorded types of housing debt (excluding the interest paid) are then added in order to obtain a measure for *active* saving

[4] In this paper, the measure for the bequest motive captures the intention to leave assets to heirs after death *and* the intention to transfer money to children or grandchildren inter vivos (see, e.g., Reil-Held, 2006). The measure is calculated as the arithmetic mean of the importance ratings to the question about "leaving bequests to children/grandchildren" and the question about "support/education of children/grandchildren". The scale is classified from 1 to 3 *after* computing the arithmetic mean. The conclusions from this study do not change, if I only use the classical bequest motive without inter vivos transfers. This suggests that saving for education/support of children/grandchildren alone is not a very important saving motive in Germany.

in 2002.[5,6] This study is concerned with the relationship between saving motives and active saving decisions, therefore, any passive saving flows are not taken into account in the considered saving measure.[7]

Figure 1 shows the mean and quartile saving rates for the 2003 cross section in each of the age classes that are considered in this study. The cross-sectional data exhibit two main features that are broadly in line with findings by Börsch-Supan et al. (2003) based on cross-sections of the German Income and Expenditure Survey (EVS) in various years: First, the saving rate has a hump shape and, second, median saving rates are positive even for elderly respondents. The appendix gives further information on the distribution of wealth and income across age classes in the SAVE sample.

Figure 1: Mean and median saving rates for different age classes, SAVE 2003.



Note: Data points are connected to facilitate readability

---

[5] Household saving(s) can be measured and defined in different ways. For a discussion of micro data measures for household saving(s) and the corresponding statistical and methodological issues, see, e.g., Alessie et al. (1997), Börsch-Supan et al. (1999), Brugiavini and Weber (2003), and Kennickell and McManus (1994).

[6] For 98 households I find that the repayments of housing debt are positive while the answer to the direct saving question is zero. For these households, I count the repayments of housing debt as total *active* saving of the household. The conclusions from this study do not change if these 98 households are excluded from the analysis.

[7] First, note that this contains a behavioral assumption. Second, note that in the SAVE questionnaire, the question about the importance of saving motives is asked in the context of a series of questions about active saving decisions; that is, in this respect, the respondents are framed to think about active savings when they answer the questions about the importance of saving motives.

2.3.3 Saver Types

SAVE elicits information on whether households save in a planned or regular manner, or whether households save irregularly and without a savings plan. The following question is asked:

Which sentence best describes the personal saving behavior of you and your partner?

☐ *I/we save a fixed amount regularly, for instance in a savings plan, in a savings account, in shares or in a life insurance scheme.* [1]

☐ *I/we put something aside each month, but I/we decide on the amount according to the financial circumstances.* [2]

☐ *I/we put something aside when I/we have something left over to save.* [3]

☐ *I/we do not save because I/we do not have enough financial scope to do so.* [4]

☐ *I/we do not save because I/we would prefer to enjoy life now.* [5]

The questionnaire asks households explicitly to choose only the one behavioral pattern that characterizes best their behavior. Clearly, the fact that one of the categories has been chosen does not rule out that actual saving behavior is more complicated and consists of several patterns. Nevertheless, the answers to this question are informative concerning the predominant saving pattern of the household. According to the answers given to this question, I classify households into four different saver types: Households that plan their saving or engage in some sort of savings plan that is associated with fixed regular saving (category [1]); households that save regularly, but do not engage in a savings plan (category [2]); households that save irregularly (category [3]); and households that do not save (category [4] and [5] combined).

Table 3 cross-tabulates the answers to this question with age classes and shows key financial statistics for each saver type. The table shows in particular that a very large proportion of households plans their saving and saves a fixed amount regularly. This proportion is significantly lower for households in the highest age class; further investigation reveals that there is also a significant difference between retired and non-retired households. Furthermore, table 3 shows that the average saving rate is highest for the group of households that engages in fixed regular saving, and decreases across saver types.

Table 3: Descriptive statistics on household saver types.

		Household Saver Type			
		1 *Regular, planned*	2 *Regular*	3 *Irregular*	4 *No saving*
	All	35%	20%	21%	24%
Age	<35	34%	14%	19%	33%
Age	35-49	47%	16%	16%	21%
Age	50-64	40%	18%	21%	21%
Age	≥65	20%	32%	27%	21%
	Mean saving rate	18.9%	15.2%	10.3%	1.7%
	Std. err.	1.1%	1.1%	0.9%	0.3%
	Mean financial wealth [€]	40,147	25,050	16,749	9,895
	Std. err. [€]	3,917	2,209	2,340	3,604
	Mean total wealth [€]	201,074	187,800	114,104	75,635
	Std. err. [€]	20,654	18,648	11,063	11,133

3 Saving Motives and Existing Literature

In this section, the existing literature is discussed in the context of the four considered saving motives and it is then related to the study presented in this paper.

Classical life-cycle theory goes back to Modigliani and Brumberg (1954) and Friedman (1957) and derives consumption and saving behavior from a well-defined intertemporal optimization problem that assumes rational and forward-looking agents who face a deterministic income path and smooth the utility of consumption over their life-cycle. Under standard assumptions about the utility function and combined with the fact that income is usually substantially lower after retirement than before, classical life-cycle theory thereby essentially captures an *old-age provision motive*. While the original intuition of the classical life-cycle model – that households save during their working years to accumulate assets which they use to sustain consumption after they retire – has been confirmed by numerous empirical studies over the years, there is also vast evidence that a large fraction of elderly households do not use up their wealth as predicted by the classical model; Mirer (1980) and Menchick and David (1983), for instance, are among the earliest of these studies. Alessie et al. (1999) show in a panel study that many elderly households even continue to accumulate wealth.

The basic model has been extended to include specific saving motives. To present an extension that includes a precautionary saving motive, I follow the prominent example of Carroll (1992, 1997). Consider a household who faces a risky labor income path and maximizes the discounted value of future utility from consumption up to time T, his time of death:

$$\max_{\{C_t\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t E[U(C_t)] \quad . \tag{1}$$

The household faces an intertemporal budget constraint:

$$X_{t+1} = R(X_t - C_t) + Y_{t+1} \, . \tag{2}$$

And the household faces a borrowing constraint:

$$X_t - C_t \geq 0 \qquad \text{for all t.} \tag{3}$$

Here, C_t is consumption, X_t is cash-on-hand at the beginning of the period, Y_t is labor income which is assumed to follow a stochastic path, β^t is the subjective discount rate, and R is the constant gross interest rate.

This model illustrates that, in the absence of complete insurance, expected shocks in disposable income lead prudent agents to save for smoothing the consumption path; i.e. under the given assumptions, savings do not only serve to finance consumption after retirement but also to insure households against income shocks. Simulations of (partially) calibrated versions (and various extensions) of the model predict that *savings for precautionary motives* can explain a large share of total wealth accumulation (see, e.g., Caballero, 1991; Carroll, 1997; Gourinchas and Parker, 2002). Most of the empirical work on precautionary saving focuses on income risk as the origin for precautionary wealth accumulation and estimates the relationship between various measures for income risk and wealth accumulation. Evidence on the precautionary motive based on micro data yields mixed results and ranges from little or no evidence (e.g., Guiso et al., 1992; Skinner, 1988) to evidence for substantial precautionary accumulation (e.g., Carroll and Samwick, 1998; Gourinchas and Parker, 2002). In the context of this variety which might be due to numerous reasons such as country and measurement differences, two shortcomings of existing studies are being emphasized in the recent literature. First, Fuchs-Schündeln and Schündeln (2005) who find considerable precautionary savings in Germany based on data from the German Socio-Economic Panel (GSOEP), argue that the extreme differences observed in existing empirical studies of precautionary saving might stem from the fact that many empirical studies fail to control for self-selection into occupations, since they do not include measures for the risk attitude of the households. Second, it is argued that the total amount of saving for precautionary accumulation might have been underestimated because risks other than income risks are not considered in most studies (e.g., Hurst et al., 2005; Kennickell and Lusardi, 2005). The present empirical study intends to circumvent the former shortcoming by including a measure for risk attitude in the multivariate estimation framework; the latter shortcoming is approached by using a measure for the importance of the precautionary motive that does not restrict attention to income risk only.

The basic version of the life-cycle hypothesis has also been extended to include a *housing motive*. Extensions that include a housing motive have been analyzed theoretically by Artle and Varaiya (1978) and by Hayashi et al. (1988). They find that in a world with downpayment constraints, the desire to purchase a house leads to additional saving for the purpose of financing home purchase. Emphasizing the role of downpayment constraints in the Italian housing market, Guiso et al. (1994) present evidence from micro data that the desire to finance housing purchase has an effect on the consumption profile of Italian households. Similarly, Moriizumi (2003) uses household data to investigate the presence of a housing motive in Japan and reports that wealth accumulation for housing purchase increases household saving and suppresses consumption for younger households. The degree of housing financial market imperfections in Italy and Japan might play an important role for the estimated effects in those studies, but it should be noted that German housing markets are also far from being perfect (Chiuri and Jappelli, 2003), suggesting that a housing motive might also have an effect on saving behavior in Germany.

Parents might not only care about themselves but also about the well-being of their children. Hurd (1987) extends the life-cycle hypothesis such that it includes a *bequest motive*. Again, the evidence on the presence and strength of an altruistic bequest motive is mixed (see, e.g., Jürges (2001) and Reil-Held (1999) for an overview and examinations of the bequest motive with the German SOEP data). The observed positive saving rates among many elderly – which contradict the simple form of life-cycle theory – do *not* prove the existence of an altruistic bequest motive. Bequests might also be purely selfish or they might be accidental (see Hurd (1990) and Kotlikoff (2001) for reviews of related literature), in which case they might stem from, e.g., uncertainty about the time of death (e.g., Davies, 1981), or from an unanticipated lack of capacity to consume (Börsch-Supan, 1992; Börsch-Supan and Stahl, 1991). Therefore, it is impossible to identify an operative bequest motive from saving rates or the shape of the wealth profile in the presence of co-existing saving motives that a study does not control for. Since the present study includes *explicit* measures for the saving motives, it identifies whether there is an overall contribution of an *intentional* (vs. an accidental) bequest motive; it is not possible to identify the separate contributions of strategic vs. altruistic intentional bequests.

While the above-mentioned studies are representative of the vast literature that focuses on only one specific saving motive and estimates the contribution of one motive versus the potential contributions of all other motives, only few studies have focused on co-existing motives. An early series of these studies was inspired by Kotlikoff and Summers (1981) (and is reviewed in Kotlikoff (1988) and in Kessler and Masson (1989)) and has been explicitly interested in the relative contribution of co-existing motives to the *stock* of accumulated wealth. Three more recent empirical studies investigate the importance of various co-existing saving motives for the *flow* of household saving using micro data sets. First, Horioka and Wanatabe (1997) calculate the contribution of net saving to the flow of household saving for a large number of saving motives. They compute this contribution from direct questions about the hypothetical amount of current wealth that a household would hold for a specific motive, from questions about the household's hypothetical wealth target for that motive, and from questions about the hypothetical number of years until the household's planned realization date of that motive. Horioka and Watanabe find that the old-age provision motive, the precautionary motive and the housing motive are clearly the three most important motives in Japan. Second, in the context of a detailed analysis of wealth holdings, income and savings in the Netherlands, Alessie et al. (1997) report descriptive statistics on a set of binary questions on whether certain saving motives exist at different stages of the life-cycle. They find that the precautionary motive is the predominant motive over the life-cycle, a housing motive is indicated by many young households but only by few older households, saving for children is particularly important at older age, and the existence of an old-age provision motive is generally indicated by only very few households in the Netherlands. Third, Alessie et al. (1999) focus on saving after retirement and report descriptive statistics on subjective importance ratings of saving motives; they find that the precautionary motive is the most important motive among retired households.

While these studies dealing with co-existing saving motives are based on descriptive statistics of survey questions concerning different saving motives, many studies that focus on one specific motive use multivariate reduced form models, in which the saving rate or accumulated household wealth is regressed on a number of socio-economic and financial household characteristics, and – if available – household preferences and expectations enter the equation additively (see, e.g., Fuchs-Schündeln and Schündeln, 2005; Kennickell and Lusardi, 2005). The present paper is also based on a reduced form equation for explaining saving behavior. While this assures comparability with a huge body of existing work, it neglects that some regressors might be endogenous to the process of wealth formation; I present different specifications to show the sensitivity of the results with respect to the potential endogeneity of measures for household wealth. Generally, the selection of the included regressors is guided by extended versions of the classical life-cycle model that emphasize the role of households' expectations about the future (see, e.g., Lusardi, 1999).

4 Empirical Analysis

The empirical analysis consists of three parts. In the first subsection, the relationship between saving motives and the *saving rate* is investigated based on different specifications of a semiparametrically estimated saving regression. The second subsection uses an almost identical multivariate specification but is concerned with the association between co-existing saving motives and the *saver type* of the household based on a multinomial model. The last subsection discusses the findings.

4.1 Saving Rate and Saving Motives

The estimation is based on the following specification:

$$y_i = \frac{S_i}{I_i} = \beta_0 + \beta_I I_i + \beta_{I^2} I_i^2 + \beta_W W_i + \beta_{W^2} W_i^2 + \beta_Z' Z_i + \beta_{risk} riskpref_i + \beta_{fut}' fut_i + \beta_{mot}' motives_i + \varepsilon_i \quad (4)$$

$$(i = 1,..., N)$$

Here, S is annual household saving as described in section 2.3.2., I is net household income, and W is household financial wealth or household total wealth, depending on the specification that is used for the analysis. Z is a vector of household characteristics: *age*, *age²*, *age³* of the household head, her/his gender, household size, the number of children of the household head or family, homeownership, educational status, and various job characteristics. The variable *riskpref* captures self-assessed risk attitude of the household head. The inclusion of measures for expectations concerning the future has been motivated in a section above; equation (4) refers to the included controls, such as expectations about income uncertainty and about the future development of the German economic situation, as *fut.* Finally, *motives* stands for the measures for the four saving motives (see section 2.3.1). These four measures are interacted with dummies for the four age classes (*< 35 years*, *35-49 years*, *50-64 years*, *≥ 65 years*) that are considered in this study. All included regressors are described in more detail in the appendix.[8]

[8] To see that the findings concerning saving motives are meaningful, note also that in *each* single age class and for *each* considered saving motive, the importance ratings of the saving motives are non-degenerately

As is clear from section 2.3.2, the dependent variable in the saving regression is left-censored at zero. A censored regression model is used to explain the saving rate *y* for all $i = 1,..., N$:

$$y_i^* = \beta' X_i + \varepsilon_i, \qquad y_i = \max(y_i^*, 0) \tag{5}$$

Tobit estimates will generally be inconsistent if the error terms are heteroscedastic or non-normal (e.g., Goldberger, 1983; Hurd, 1979). For all specifications that I consider, the assumptions of homoscedasticity and normality of the error term are rejected in the present censored model at the 5% level based on the corresponding Lagrange Multiplier tests for censored models (Chesher and Irish, 1987). I therefore use Powell's (1984) semiparametric censored least absolute deviations (CLAD) estimator, which is consistent and asymptotically normal even if errors are heteroscedastic. In contrast to the assumption of homoscedastic and normal errors, which is imposed in the Tobit model, CLAD imposes the following conditional median restriction:

$$Med(\varepsilon_i \mid X_i) = 0 \tag{6}$$

The CLAD-estimator requires the minimization of a nondifferentiable function, Buchinsky's (1994) iterative linear programming algorithm (ILPA) is used. The properties of CLAD with respect to the degree of censoring and the sample size have been investigated in various simulation studies (Deaton, 1999; McDonald and Xu, 1996; Paarsch, 1984). Both the degree of censoring and the sample size of the considered estimation in this paper, fall well beyond the limits that are specified in those studies and therefore strongly advocate the use of the CLAD estimator rather than Tobit estimation. Standard errors for the CLAD estimates are computed using 150 bootstrap replications.

The existing stock of wealth might be a substitute for, e.g., precautionary or retirement wealth accumulation, that is, it might be endogenous to the saving decision. To investigate the sensitivity to the inclusion of wealth, I use three specifications: Specification (a) excludes the wealth variables, specification (b) includes financial wealth only, and specification (c) uses total net wealth of the household.

---

distributed over the three importance rating categories (see table 2). As well, the saving rate has considerable and very similar variation in *each* age class. This is important in order to ensure that the effect of saving motives on the saving rate is identified. If, for example, *all* respondents in a certain age class would rate a certain saving motive as "very important", the saving motive could be operative, although the estimation would not find a significant coefficient for the motive in the particular age class. Note further that the results presented in this study are robust to the choice of the symmetric scale in section 2.3.1.

Table 4: CLAD estimation of three different specifications of the saving regressions.

	(a)			(b)			(c)		
savings rate	Coef.		Std. Err.	Coef.		Std. Err.	Coef.		Std. Err.
age	0.263	***	0.099	0.198	***	0.076	0.255	***	0.097
age2	-0.048	**	0.019	-0.037	**	0.014	-0.047	**	0.018
age3	0.003	**	0.001	0.002	**	0.001	0.003	**	0.001
partner	0.027	**	0.013	0.019	*	0.011	0.024	*	0.013
hhsize	-0.010	*	0.005	-0.010	**	0.005	-0.009	*	0.005
children	0.001		0.004	0.003		0.003	0.001		0.004
female	-0.015	*	0.009	-0.013		0.008	-0.012		0.009
highschool	0.030	**	0.012	0.019	*	0.011	0.027	**	0.014
civilservant	0.032	*	0.019	0.034	**	0.017	0.035	*	0.018
selfemployed	0.042		0.025	0.021		0.025	0.038		0.026
unemployed	-0.079	**	0.033	-0.063	***	0.020	-0.077	***	0.030
homeowner	0.080	***	0.010	0.068	***	0.009	0.070	***	0.012
retired	0.009		0.020	0.010		0.017	0.005		0.018
unemp_prob	-0.020		0.020	-0.014		0.021	-0.019		0.022
heritage_prob	0.019		0.027	0.012		0.028	0.014		0.027
earnings_var	0.000		0.001	0.000		0.000	0.000		0.001
dev_ger_econ_sit	0.006	***	0.002	0.005	**	0.002	0.006	***	0.002
lifeexpect	-0.002		0.004	-0.001		0.004	-0.002		0.004
dev_health_sit	0.002		0.002	0.002		0.002	0.002		0.002
riskpref	0.005	**	0.002	0.002		0.002	0.004	**	0.002
netinc	-0.065		0.089	-0.111		0.051	-0.070		0.095
netinc2	0.005		0.064	0.009		0.036	0.006		0.086
financialwealth				0.125	***	0.040			
financialwealth2				-0.008		0.015			
wealth							0.004		0.003
wealth2							0.000		0.000
mot_oldage1	0.033	*	0.019	0.025		0.017	0.034	*	0.018
mot_oldage2	0.027	**	0.011	0.021	**	0.010	0.030	***	0.011
mot_oldage3	0.005		0.010	0.010		0.011	0.006		0.010
mot_oldage4	-0.004		0.010	0.000		0.009	0.000		0.010
mot_precaution1	0.003		0.017	0.010		0.015	0.005		0.016
mot_precaution2	0.010		0.011	0.011		0.010	0.008		0.012
mot_precaution3	0.025	**	0.011	0.014		0.011	0.026	**	0.010
mot_precaution4	0.024	*	0.013	0.022	**	0.010	0.026	**	0.012
mot_homepurchase1	0.025	*	0.015	0.021	*	0.012	0.022		0.015
mot_homepurchase2	0.004		0.008	0.005		0.007	0.006		0.008
mot_homepurchase3	0.005		0.010	0.005		0.009	0.007		0.010
mot_homepurchase4	-0.011		0.011	-0.013		0.009	-0.014		0.011
mot_bequest1	-0.011		0.017	-0.011		0.014	-0.011		0.017
mot_bequest2	0.002		0.010	0.004		0.009	0.001		0.010
mot_bequest3	0.017		0.011	0.013		0.011	0.016		0.011
mot_bequest4	0.022	**	0.011	0.017	*	0.010	0.017		0.012
constant	-0.476	***	0.172	-0.347	***	0.130	-0.465	***	0.168
# obs.	2184			2184			2184		
Pseudo R2	0.069			0.091			0.071		

*Note: *** : 1% significance level; ** : 5% significance level; * : 10% significance level.*

Table 4 presents the results of the CLAD estimation, and I report on results that are significant at the 10%-level in the text.[9]

Each of the three age variables is significant in all specifications, and the three age variables are jointly significant in all specifications. The high school dummy which indicates whether the household head and/or her/his partner have senior high school education (the German "(Fach-)Abitur"), the dummy for civil servants, for unemployed household heads and for households that own their currently occupied house or apartment are all significant in the three specifications: Households in which at least one of the partners has high school education, have on average a saving rate which is about 3 percentage points higher than the saving rate of households for which this is not the case.[10] Households with unemployed household heads have a saving rate, which is about 8 percentage points lower than households whose household head is working, and civil servants have a saving rate, which is about 3 percentage points higher on average. The coefficient of the home-ownership dummy is positive and significant, suggesting that households that own their occupied house or apartment have a saving rate that is about 8 percentage points higher than the saving rate of households that are not homeowners.

Turning to the main variables of interest, the saving motives, it is first found that despite the many included covariates, some of the interactions between saving motives and age classes still have significant predictive power. The coefficients of those interactions are a measure for the change in the saving rate in percentage points that is associated with a one unit increase in the importance rating of a certain saving motive for a certain age class in the considered cross-section. That is, on average, a household in the oldest age group that rates the precautionary saving motive as "*very important*" has a

---

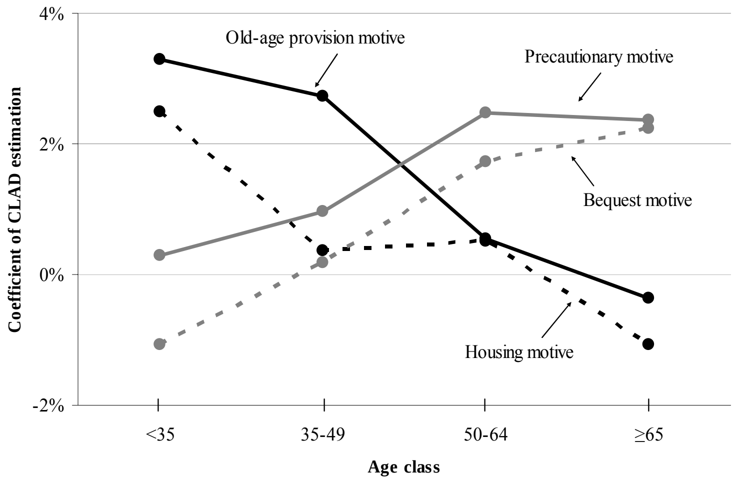
[9] Please refer to the tables for more detailed information on the significance levels.
Two findings stands out in table 4: First, despite the inclusion of many explanatory variables, more than 90% of the variation in the saving rate remains unexplained. This is common in most studies of this type (see, e.g., Lusardi, 1999, p. 103-109). Note that the value of R^2 even decreases further if I follow a common approach, transform zero saving rates to a very small value and then log-transform the data for the savings rate. This suggests that the linear specification (4) in combination with the CLAD estimation which is robust to outliers (the presence of which is unavoidable in data of this type) should be preferred to the log-transformation in the present case. Second, while most reported coefficients do not vary much across specifications, specification (b) differs somewhat from specifications (a) and (c) – a finding that also shows up in the following sections of this paper and that is due to the correlation between financial wealth and the dependent variable.

[10] All numerical examples that I use for illustrating the results of the CLAD-estimations refer to specification (a).

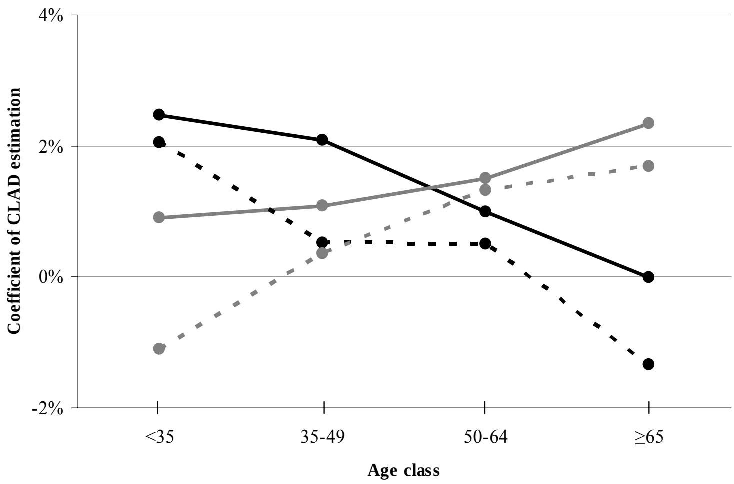
saving rate that is 2.4 percentage points higher than the saving rate of a household with identical covariates that rates the precautionary motive as "*important*". Figure 2 shows the coefficients of the four saving motives by the age group of the household head for the three considered specifications. All figures show a similar pattern and illustrate how the association between saving motives and the saving rate varies over age groups.

Figure 2: Coefficients of the CLAD estimation for four saving motives and age classes. The coefficients of the CLAD estimation denote the change in the saving rate in percentage points due to a change in the subjective rating of a certain saving motive by one unit.

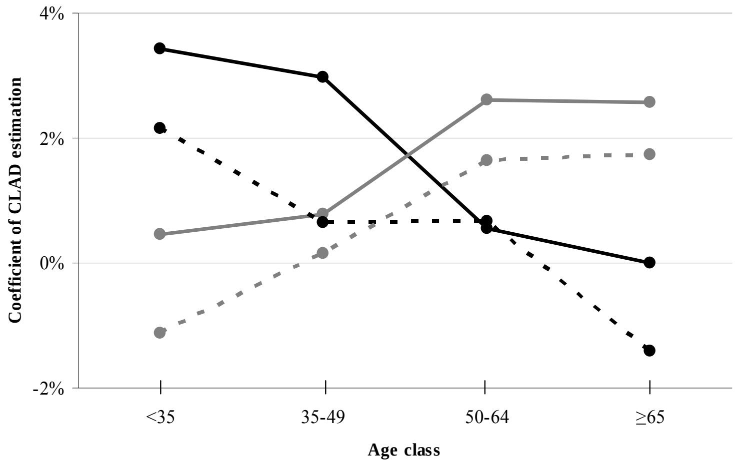
Specification (a):



Specification (b):



Specification (c):



Note: Data points are connected to facilitate readability

The findings from this analysis are informative in two respects: First, concerning the subjective information on saving motives that is elicited in the SAVE study and, second, concerning the question which saving motives are operative at what life stage.

Concerning the subjective information on saving motives, I find that while the descriptive statistics on the importance ratings of the single saving motives (see section 2.3.1) do not show a significant trend over all age classes (with the exception of the housing motive), the multivariate analysis does find that saving motives change significantly over age groups in their explanatory power for actual saving behavior. An explanation for the finding that trends over life stages vary between the pure descriptive statistics and the multivariate analysis is that households answer the subjective question about the importance of the saving motives by just stating their *general* importance rating of the saving motives.[11] The multivariate analysis, however, estimates whether information on a single motive is indeed related to actual saving behavior at a certain life stage and under the assumption of co-existing saving motives.

[11] I want to give two examples: First, almost every sixth childless household in the oldest age class rates the bequest motive as important or very important, although the corresponding question explicitly talks about children or grandchildren as the recipients. Second, Table 2 reveals that almost 30% of the households in the oldest age class think that the housing motive is an important or very important saving motive; however, the age, the financial resources, and the answer to a specific question about the savings goal suggest clearly that almost all of these households will most likely not purchase a house in the future.

Concerning the question which motive is operative at what life stage, table 4 shows that the old-age provision motive and the housing motive both are significantly related to the saving rate in early life stages. While the presence of a housing motive in the youngest age class of German households is of interest for itself, the finding that there is a particularly strong effect of the old-age motive for the youngest age class deserves some more explanation as it might be connected to the increased public debate about the German pension system which started in the late 1990s and which was associated with marketing and information campaigns by insurance and bank companies. These campaigns have especially targeted younger households, which will be affected stronger by the reforms than older cohorts. Börsch-Supan et al. (2004) provide evidence for a recent increase in the awareness about the fact that one effect of the pension reform will be a decrease in pension levels, and young households are particularly aware of these facts.[12]

Table 4 further reveals that in contrast to the old-age provision motive and the housing motive, the bequest motive and the precautionary motive are particularly operative for older age groups. Both findings are comparable with existing studies that focus on only one specific saving motive. First, in his study that focuses exclusively on the bequest motive, Jürges (2001) also finds an operative bequest motive among the elderly. He reports consistently and significantly different wealth trajectories for elderly households that declare that they have a bequest motive compared to households that declare not to have a bequest motive. Second, the effect of the precautionary motive is in line with findings on precautionary wealth accumulation by Kazarosian (1997) and Lusardi (1998, 2000), who investigate older workers, as well as by Carroll and Samwick (1998) and by Kennickell and Lusardi (2005). An explanation for the increase in the precautionary motive with age are the increased health risks that older people face, i.e. risks associated with considerable health costs. Indeed, even controlling for many

[12] Furthermore, the great majority of household heads in the SAVE sample are dependent employees (see table 1), for whom participation in the German pay-as-you-go system is mandatory, and for many of whom private old-age provision has only recently become an important issue, given that a large proportion had completely relied on publicly funded old-age provision provided by the traditionally fairly generous German pension system. The German retirement insurance system has a high replacement rate, generating net retirement incomes that have been about 70 percent of pre-retirement net earnings for a dependent employee with a 45-year earnings history and average life-time earnings in the late 1990s. Overall, public pensions constitute more than 80 percent of the income of households headed by persons aged 65 and older, while funded retirement income, such as asset income or firm pensions, plays a much smaller role than, e.g., in the Netherlands or the Anglo-Saxon countries (Börsch-Supan et al., 2003).

household characteristics, I find that with increasing age expectations concerning the development of the health situation get worse, whereas expectations about the future economic situation are not significantly related to the age of the household, and subjective expectations about future earnings variance decrease with an increase in age (see appendix, section 6.3).

4.2 Saver Types and Saving Motives

The previous section shows that information on saving motives helps to explain *how much* households save. Do saving motives also help to explain *how* households save, i.e. whether they engage in regular savings plans, or rather save irregularly and without a savings plan? The goal of this section is to relate heterogeneity in the degree of planning and regularity of saving behavior to households' saving motives in a multivariate framework that includes the saving motives as in the previous section. The results are informative as to whether certain motives for saving are crucial in determining the saver type of a household.

Authors that are concerned with heterogeneity in the extent to which households plan their saving (e.g., Lusardi, 1999; Lusardi and Mitchell, 2006; Venti, 2006) underline that numerous behavioral and psychological factors interfere with the ability to compute optimal plans that would follow conventional theory, or to simply make a plan and execute it. Most research on financial planning and saving behavior is specifically concerned with long-term planning for retirement. For example, Lusardi (1999) controls for numerous variables such as age and lifetime income and finds that those who have given "little or no" thought to retirement have financial wealth that is significantly lower than the financial wealth of those who have given the subject more thought.

Conventional life-cycle theory assumes that households formulate savings plans based on expectations about the future and is concerned with the amount of household saving, but the theory neither models psychological factors that are relevant in this respect, nor does it take a stand on the regularity and contractual form of household saving and its relationship to saving motives. However, given certain income paths, life-cycle theory has some implications: For example, consider a household with an extremely volatile income path that regularly drops below the expenditure and consumption path and with only a small stock of financial and liquid wealth. This household might well have precautionary savings, which have been accumulated in periods with higher income and which are needed to finance consumption in unforeseen low income periods (see, e.g., Carroll and Samwick (1998), who provide simulations based on the buffer stock model). But in the presence of borrowing constraints, intertemporal consumption smoothing implies that this household would not engage in regular or in contractual saving: The household would not be a regular saver because of the dramatic income shocks that occur from time to time, and the household would not engage in contractual saving since the money should not be bound contractually, in order to be able to finance consumption in unexpected low-income periods. In turn, high-income civil servants[13], for instance, would probably save very regularly to provide for unforeseen events for which liquid wealth is needed or to provide for old age. Given the attractiveness of certain savings contracts, in particular considering existing state subsidies for certain long-term savings plans, it might also be rational for high-income civil servants to engage in contractual saving. First, these deliberations serve to illustrate that while the life-cycle model is informative concerning the saver type for specific income paths, it is generally rather silent about the relationship between the form of saving and saving motives. Second, they suggest that any study that is concerned with the identification of the relationship between saving motives and household saver types should include proxies for the income uncertainty of the household; the present study includes dummies for the type of employment and a subjective measure for future earnings variance.

I investigate the relationship between saver type and saving motives using discrete choice models. The same explanatory variables as in the analysis in section 4.1 enter the estimation. The only difference is that the saving motives are not interacted with age classes, since there is no a-priori hypothesis that the effect of saving motives on the saver type should vary by age class. Furthermore, the sample for this analysis is restricted to the non-retired population, since life-cycle theory predicts that retired households dissave. In particular, there should not be an old-age provision motive any more for retired households, i.e. those households do not save for an income drop due to retirement. In fact, the data show, first, a sudden decrease in the saving rate after retirement and a significant increase in left-censored observations with the corresponding saving rate being less or equal to zero.

[13] In Germany, civil servants can expect a non-declining income path until retirement. A civil servant can only be transferred to a new position if her wage does not decline due to the transfer. Furthermore, a civil servant can only be dismissed is she is sentenced to a certain period in prison for any criminal charge or for charges associated with treason.

Table 5: Multinomial logit estimation for three different specifications. Base category: Irregular savers.

	(a)		(b)		(c)	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Regular						
age	-1.424	2.920	-1.551	2.906	-1.444	2.922
age2	0.357	0.752	0.383	0.748	0.364	0.753
age3	-0.028	0.062	-0.031	0.061	-0.029	0.062
partner	-0.070	0.266	-0.086	0.266	-0.070	0.266
hhsize	-0.087	0.116	-0.082	0.115	-0.092	0.116
children	-0.029	0.114	-0.022	0.115	-0.021	0.115
female	-0.205	0.208	-0.194	0.208	-0.205	0.208
highschool	0.039	0.245	0.019	0.247	0.038	0.246
civilservant	0.931 *	0.480	0.909 *	0.483	0.929 *	0.480
selfemployed	0.397	0.377	0.396	0.378	0.372	0.381
unemployed	-0.732 **	0.341	-0.709 **	0.342	-0.721 **	0.342
homeowner	0.274	0.220	0.199	0.224	0.215	0.244
unemp_prob	-0.121	0.376	-0.100	0.379	-0.114	0.377
heritage_prob	-0.916	0.575	-0.956 *	0.580	-0.921	0.576
earnings_var	-0.004	0.003	-0.004	0.003	-0.004	0.003
dev_ger_econ_sit	0.003	0.045	0.000	0.045	0.004	0.045
lifeexpect	0.024	0.084	0.018	0.084	0.022	0.084
dev_health_sit	0.032	0.055	0.029	0.055	0.033	0.055
riskpref	-0.013	0.043	-0.019	0.043	-0.014	0.043
netinc	1.354 **	0.585	1.074 *	0.591	1.277 **	0.595
netinc2	-0.110	0.068	-0.091	0.078	-0.105	0.069
financialwealth			0.857 *	0.487		
financialwealth2			-0.077	0.085		
wealth					0.030	0.055
wealth2					0.000	0.001
mot_oldage	0.248 *	0.143	0.238 *	0.144	0.250 *	0.143
mot_precaution	-0.129	0.164	-0.138	0.165	-0.131	0.165
mot_homepurchase	0.182	0.116	0.167	0.117	0.182	0.116
mot_bequest	-0.081	0.141	-0.088	0.141	-0.082	0.141
constant	1.074	3.589	1.379	3.578	1.113	3.592
Regular, planned, contractual						
age	1.538	2.535	1.402	2.538	1.490	2.538
age2	-0.223	0.649	-0.200	0.650	-0.209	0.650
age3	0.003	0.053	0.001	0.053	0.002	0.053
partner	-0.103	0.226	-0.143	0.226	-0.104	0.226
hhsize	-0.030	0.100	-0.033	0.100	-0.037	0.100
children	-0.103	0.100	-0.083	0.101	-0.092	0.101
female	-0.157	0.173	-0.138	0.174	-0.161	0.173
highschool	0.155	0.207	0.077	0.210	0.151	0.208
civilservant	0.878 **	0.441	0.841 *	0.446	0.877 **	0.441
selfemployed	-0.026	0.340	-0.081	0.346	-0.061	0.343
unemployed	-0.532 **	0.258	-0.470 *	0.258	-0.516 **	0.258
homeowner	0.498 ***	0.183	0.373 **	0.186	0.420 **	0.204
unemp_prob	-0.151	0.315	-0.104	0.320	-0.139	0.316
heritage_prob	0.060	0.431	-0.057	0.440	0.059	0.431
earnings_var	-0.002	0.001	-0.001	0.001	-0.002	0.001
dev_ger_econ_sit	0.089 **	0.038	0.088 **	0.038	0.091 **	0.038
lifeexpect	0.033	0.068	0.015	0.069	0.030	0.068
dev_health_sit	-0.016	0.045	-0.020	0.045	-0.015	0.045
riskpref	-0.002	0.036	-0.013	0.036	-0.004	0.036
netinc	0.777	0.544	0.206	0.539	0.676	0.551
netinc2	-0.046	0.056	0.004	0.058	-0.037	0.057
financialwealth			1.400 ***	0.408		
financialwealth2			-0.089 ***	0.029		
wealth					0.039	0.048
wealth2					0.000	0.001
mot_oldage	0.790 ***	0.127	0.749 ***	0.127	0.789 ***	0.127
mot_precaution	-0.294 **	0.140	-0.300 **	0.141	-0.292 **	0.140
mot_homepurchase	-0.122	0.097	-0.129	0.098	-0.119	0.097
mot_bequest	0.172	0.115	0.165	0.116	0.168	0.115
constant	-2.829	3.157	-2.331	3.161	-2.750	3.161
# obs.	1066		1066		1066	
Pseudo R2	0.068		0.077		0.068	

Second, there is a highly significant difference in the distribution of households across saver types between the retired and the non-retired sample, and only mild and mostly insignificant differences in the distribution between different age classes of the non-retired sample. And, third, the analysis presented above shows that the old-age provision motive has no significant predictive power for households in the highest age class.

The relationship between the saver type classification and saving motives is first investigated using a multinomial logit model for three alternatives.[14] Table 5 presents estimation results using the type "irregular saver" (category [3]) as a base category. For reasons stated above, I present again the three different specifications that have been used in the previous section.

Table 5 reveals that the estimated coefficients and standard errors do not differ very much across specifications; therefore, the following interpretation of the results does not distinguish between specifications. Focusing on the type of households that plan their saving and engage in some sort of regular savings plan (type 1), it is first found that civil servants are significantly more likely to be of this type, and unemployed households are significantly less likely to be of this type relative to the base category, type 3. While the bequest and the housing motive are not significantly related to the relative probability ratios, an increase in the subjective importance rating of the precautionary motive is associated with a significant decrease in the probability of being of this saver type (type 1) relative to being an irregular saver (type 3). More specifically: Relative to the base alternative, an increase of the precautionary saving rating from "unimportant" to "important" is associated with a 26% smaller probability of being in the group of households that plan their saving and engage in some sort of regular savings plan. Conversely, an increase in the importance rating of the old-age provision motive comes along with an increase in the relative probability of being in this group. The model estimates a 120% higher probability relative to the base alternative if the old-age provision motive is increased by one unit. For the group of regular savers that do not engage in fixed saving (saver type 2), no significant relationship at the 10% level is found except from the

[14] In multinomial logit models, the odds ratio between any two choices does not depend on the other choices, this property is termed the independence of irrelevant alternatives (IIA). A Hausman-McFadden test (Hausman and McFadden, 1984) suggests that for all specifications that I consider, the IIA assumption cannot be rejected.

result that an increase in the importance of the old-age provision motive is positively associated with the probability of being a regular saver relative to the base alternative.

An important underlying assumption of the multinomial logit estimation is the assumption of independence of irrelevant alternatives (IIA). This assumption implies a certain pattern of substitution across alternatives. If substitution actually occurs in this way and if the model is specified correctly, then the multinomial logit model is appropriate. While the IIA property that gives rise to the proportional substitution pattern of the multinomial logit model was not rejected in the present case by a Hausman-McFadden test (see footnote 14), it has been noted that this test has low power under many circumstances (see, e.g., McFadden, 1987). Therefore, I have also estimated a multinomial probit model that relaxes the IIA assumption by allowing for correlation across choices in the unobserved components. The findings from the multinomial probit model are in line with the conclusions presented above, and they are detailed in the appendix, section 6.4. Finally, I have also investigated the relationship between saver type and saving motives based on binary logit models for all three specifications.[15] In the binary choice models the probability of being of a certain saver type is compared to the probability of being in *any* of the other groups. The findings support all conclusions from the multinomial choice analysis.

The analyses in this section present descriptive evidence that there is a relationship between importance ratings of saving motives and the household saver type. First, I found that an increase in the importance attached to precautionary reasons for saving is associated with a decrease in the probability of being of saver type 1 relative to saver type 3, and – as revealed by further investigation – to a decrease of the probability of being of saver type 1 relative to type 2. An explanation is that households with a strong precautionary motive are aware that they might need their savings at some particular but unknown point in time, and they therefore decide that their savings should not be bound in a savings plan or in shares by that unknown point in time. Causality might also go in the opposite direction: Households have a strong precautionary motive, because they can only save irregularly when there is some money left over and because they might need the saved money in periods when nothing is left over.

Second, I find that an increase in the importance of the old-age provision motive is associated with a significantly higher probability of engaging in regular and planned saving. This finding might have several explanations. One explanation is that households

[15] The results can be obtained from the author upon request.

that want to save for retirement react to the incentives of banks and insurance companies as well as the subsidies by the government and use the more attractive longer-term savings plans in order to save for long-term saving goals. A recent study by Reil-Held and Schunk (2006) reveals that – controlling for co-existing saving motives – there is indeed a statistically significant association between the importance attached to an old-age provision motive and the probability of buying state-promoted and long-term savings plans, such as a so-called Riester-pension, life-insurance schemes, or other private pension schemes. A further plausible explanation is that households indicating a high importance of old-age provision exercise self-commitment: Savings that are planned for retirement should remain untouched during work-life and are therefore made in the form of fixed contractual savings.

Through allowing for the co-existence of various saving motives, the presented results on the significance of the old-age provision motive add well to findings about saving behavior and future planning. First, combined with the descriptive result in table 3 that households that save regularly and in a savings plan also have a higher saving rate on average, the findings are in line with the above-mentioned findings by Lusardi (1999) concerning a relationship between retirement planning and wealth accumulation. Second, they complement findings by Ameriks et al. (2003), who report direct evidence that households with a high propensity to plan their long-term future save more, are better able to exercise self-control, and self-commit to a certain behavior.

4.3 Discussion

The presented estimations include an extensive set of variables. This serves to show that the measures for saving motives correlate with saving behavior even after controlling for the rich information about households available in the SAVE survey. The fact that three different specifications generally lead to similar results underlines the robustness of the results. Of course, the direction of the causality as well as the presence of third factors is debatable in the given context; the presented methodology does not address the question of causation, and any causal interpretation of the results would depend on the underlying model and its underlying assumptions.[16] In the given context, accumulated wealth itself could have an effect on the importance that households attach to certain saving motives. Additionally, it is important to note that the cross-sectional data that are used for this study

[16] An example is the basic assumption that people are forward-looking: If people were not forward-looking, the saving motives would not play any role for explaining their savings behavior, people would simply save what is left over after consumption, without having any specific saving motive in mind.

do not allow to control for cohort effects. The cross-sectional data neither permit the estimation of structural models that account for endogeneity and dynamics. But since the dependent variable in the analysis of saver types characterizes a stable behavioral rule rather than one single observed saving decision, the analysis of saver types is not very sensitive to dynamic shocks that might have an impact on the findings.

A limitation of this study is that through providing two independent analyses, I implicitly make the behavioral assumption that households face two independent decisions: They decide how much they save, and they decide whether to engage in savings plans, save regularly, or rather save irregularly. These two decisions are not necessarily independent, as the descriptive statistics in table 3 indicate. Another model would be that households decide first about how much they save and then – conditional on the amount that they want to save – they decide about how regular they save or whether they engage in a savings plan. It is not clear which is the correct model for the decision-making process in this case. Further multinomial choice analyses of the saver type, in which I include the saving rate as an additional covariate, reveal that the saving rate is significantly and positively associated with the relative probability of being a regular saver (type 2) and a saver who engages in savings plans (type 1); however, the coefficients of the saving motives are hardly affected by the inclusion of the saving rate, indicating that the established relationships still hold.

Finally, the measures for the saving motives themselves could be related to other included variables – such as risk preferences or future expectations – or to unobserved factors that are relevant for decision-making but that the study does not control for, e.g. psychological traits of the respondent. Given that there is no testable theory that relates the psychological traits measured in SAVE to saving motives and saving decisions and that would guide a further analysis of their relationship to savings behavior, a straightforward way to learn more about the potential impact of those factors on the presented results is to simply include those psychometric variables additively in the regressions. As an example, consider that optimism rather than classical preference measures may be linked to major economic decisions, as is claimed by various scholars (Gervais and Goldstein, 2004; Rigotti et al., 2004; Puri and Robinson, 2005). Following this idea, a self-reported measure for optimism has been included in the analysis. As expected, this measure correlates with most elicited measures for future expectations. Nevertheless, the inclusion of the measure for optimism into the analyses presented in this paper does not have a considerable effect on the coefficient estimates for the saving motives, i.e., it does not alter the conclusions from this paper. In the SAVE survey, the household head is also asked to provide a self-assessment concerning her happiness, her self-assuredness, and she is asked to what degree she considers herself a creature of habit or a person that is open to change.[17] The inclusion of all these subjective measures in the analyses does not have an impact on the conclusions of this paper. As well, SAVE elicits alternative measures for risk preferences than the one considered in the presented analysis;[18] after including these alternative risk measures, still the same relationship between saving motives and how much and in what form households save is found. These findings underline the robustness of the results.

Overall, the results – established in a framework that controls for the co-existence of different saving motives – show that the subjective assessment of the importance of saving motives is significantly related to two dimensions of household saving behavior. If these relationships are insensitive to a wide range of policy changes and to changes in micro- and macro-economic circumstances, then targeted information campaigns or policy reforms that substantially change the importance of certain saving motives in the eyes of private households might indeed have various effects on the saving behavior of those households. These findings are of particular interest in the context of current policy reforms in Germany, which directly interact with private household saving, and therefore require an understanding of whether and how households react to the desired reforms and the associated information campaigns. Particularly helpful for policy would be the question whether the relative saving contributions of different motives compete with each other. Given that (most) households are constraint in their budget, I argue that saving motives compete, and – as suggested by the findings in this paper – that a different set of motives competes at different life stages. From a policy perspective it is of interest to understand the precise nature of this competition better. Is the old-age provision motive competing with the housing motive, with the precautionary motive, or with both motives? How does the nature of this competition change over the life-cycle? The present study illustrates that indeed many motives whose relative contribution changes over age classes

[17] For all these above-mentioned measures (i.e., optimism, self-assuredness, etc.), respondents are asked on a scale from 0 to 10 whether a statement of the form "*I am optimistic*", *"I am a self-assured person"*, etc. *"does not apply at all*" (0), or "*applies very well*" (10).

[18] More specifically, respondents are asked about their willingness to take risks with respect to their health, their career, leisure time and sports, and car drving on a scale from 0 to 10.

are simultaneously associated with saving decisions and must be taken into account when discussing the effect of policy reforms on household behavior. However, the present study is neither concerned with exactly estimating the relative contribution of each single motive for specific groups of the population; nor does the study investigate precisely to what extent specific motives compete with each other.

5 Conclusion

This paper has investigated household saving behavior based on a random sample of German households. The data contain rich information on household financial and socio-demographic characteristics and they offer the opportunity to investigate saving behavior under the assumption of co-existence of various saving motives which are elicited based on subjective importance ratings.

The results of this study support the view that households' saving decisions are influenced by different saving motives that co-exist over age classes, but whose relative contribution to household saving changes with age. Households' reported importance of various saving motives is related to heterogeneity in saving behavior with respect to two dimensions: First, it is related to heterogeneity in the household saving rate at different life stages. The effects of various saving motives are generally appropriate given the different stages of the households' life-cycle. Second, both the old-age provision motive and the precautionary motive are related to heterogeneity in the saver type, i.e. related to a classification of households based on whether they engage in savings plans, save regularly, or irregularly. While I have discussed that it is debatable how the decision-making process concerning the choice of the saver type works, the latter findings suggest that for many households the decision whether to save in a savings plan is related to the purpose of their saving. For instance, according to the estimations, households indicating a high importance of old-age provision have a high probability of saving regularly and in savings plans. At the same time, these relationships can be driven by a wish to exercise self-control on the part of those households that are concerned about their retirement saving. How this relationship works precisely, how psychological determinants and institutional incentives influence the process of wealth accumulation and how the process of wealth accumulation itself might feed back onto the relevant psychological determinants of saving behavior are very interesting and important open questions.

The finding of a significant relationship between the importance that households attach to different saving motives and their actual behavior suggests that policy reforms – e.g., the current German pension reform – that substantially change the importance of certain saving motives in the eyes of private households might indeed alter household saving behavior in different ways, and with differential effects over the life stages. This study has not focused on how large the contributions to the single saving motives are for specific groups of the population and in which way the relative contributions of the motives are competing. An extension of this study that investigates the relationship between saving motives and the flow of household saving to various specific financial assets – such as pension plans, building society contracts etc. – serves to assess better the contribution to each single saving motive and is on the agenda.

6 Appendix

6.1 Wealth and Income

Table A.1: Distribution of wealth and income of German households in SAVE 2003.

		Financial wealth in 2002 [€]		
		Mean	Std. err.	Median
	All	25,125	1,771	7,986
Age	<35	9,252	922	1,200
	35-49	31,778	4,417	10,500
	50-64	32,852	2,551	14,100
	≥65	23,490	3,920	9,000
		Total wealth in 2002 [€]		
		Mean	Std. err.	Median
	All	150,833	9,005	25,486
Age	<35	48,215	6,346	2,000
	35-49	168,627	23,103	40,000
	50-64	206,210	17,545	74,681
	≥65	164,889	14,582	37,250
		Net income in 2002 [€/month]		
		Mean	Std. err.	Median
	All	2,476	92	1,866
Age	<35	2,215	194	1,500
	35-49	2,945	158	2,315
	50-64	2,832	273	1,990
	≥65	1,810	71	1,500

Note: The difference in standard errors is often due to a few extremely large values, for instance the standard error of household net income in age class 50-64.

6.2 Covariates

Table A.2: Description of covariates included in the estimated models.

Variable	Description
age, age2, age3	*age* is the age (in years) of the household head divided by 10, i.e.: *age* = (age of household head)/10. *age2* is squared *age*, and *age3* is cubic *age.*
partner	Dummy: 1 if the household is married and/or lives permanently with a partner in his/her household.
hhsize	Total number of people living in the household.
children	Total number of children and children-in-law of the household.
female	Dummy: 1 if household head is female.
highschool	Dummy: 1 if the household head and/or his/her partner have a general senior high school leaving certificate or a comparable certificate for University of Applied Sciences ("(Fach-)Abitur").
civilservant	Dummy: 1 if the household head is a civil servant (see also footnote 13).
selfemployed	Dummy: 1 if the household head is running a business or any other self-employed activity.
unemployed	Dummy: 1 if the household head is currently unemployed.
homeowner	Dummy: 1 if a household member owns the currently occupied house/apartment.
retired	Dummy: 1 if the household head is retired.
unemp_prob	Subjective probability of becoming unemployed in the year of the survey. If living with partner and both partners are working: Subjective probability that at least one of the partners becomes unemployed.
heritage_prob	Subjective probability of inheriting a substantial amount or receiving a gift in the next two years. The probability is included only for those respondents who answer in the follow-up question that the inheritance or the gift or will improve the financial situation significantly.
earnings_var	Subjective earnings variance. The measure of subjective earnings variance is calculated from the subjective unemployment probability of both partners, from net income, and from the replacement rate, as in Lusardi (1998).
dev_ger_econ_sit	Expectation about future development of German economic situation, 0 for very negative expectation, 10 for very positive expectation.

lifeexpect	Subjective life expectancy of the household head, 7 brackets: < 65, 65-70, 71-75, 76-80, 81-85, 86-90, > 90
dev_health_sit	Expectation about future development of health situation, 0 for very negative expectation, 10 for very positive expectation about future health situation.
riskpref	Risk attitude: Willingness to take risks with respect to money matters. 0: no willingness to take risks, 10: high willingness to take risks.
netinc	Net income of the household, divided by 10,000 €.
netinc2	*netinc* · *netinc*
wealth	Total net wealth of the household (i.e., savings investments, savings bonds, share- and real-estate bonds, occupational and private pension schemes, real estate, business wealth etc.), divided by 100,000 €.
wealth2	*wealth* · *wealth*.
financialwealth	Financial wealth of the household (i.e., savings investments, savings bonds, share- and real-estate bonds, occupational and private pension schemes etc.), divided by 100,000 €.
financialwealth2	*financialwealth* · *financialwealth*.

Additionally, subjective importance ratings of the four saving motives are included as covariates. In the CLAD-estimation, these measures are interacted with dummies for the four age classes that are considered in this study. In the regression output, “1” stands for the lowest age class (<*35*), “2” refers for the age class *35-49*, “3” indicates age class *50-64*, and the oldest age class is denoted by “4”. That is, “mot_oldage1” refers to the old-age provision motive in the lowest age class. In total, 4 · 4 = 16 interacted variables for the saving motives are included in the regressions.

6.3 Expectations and Age

Table A.3: Linear regression of future expectations on age and further household characteristics.

	dev_health_sit			*dev_ger_econ_sit*			*earnings_var*		
	Coef.		Std. Err.	Coef.		Std. Err.	Coef.		Std. Err.
age	-0.391	***	0.046	0.054		0.049	-1.948		1.385
partner	0.342	***	0.114	-0.002		0.122	-8.452	**	3.461
hhsize	-0.024		0.053	0.069		0.056	-1.555		1.595
kids_no	-0.172	***	0.040	-0.028		0.043	-0.437		1.211
female	-0.259	***	0.091	-0.066		0.097	-6.709	**	2.758
highschool	0.328	***	0.122	0.503	***	0.130	-6.709	*	3.677
civilservant	0.226		0.251	0.252		0.267	-18.702	**	7.580
selfemployed	0.464	**	0.223	0.032		0.238	-8.067		6.771
unemployed	-0.568	***	0.153	-0.662	***	0.165	2.346		4.657
homeowner	0.302	***	0.105	-0.010		0.112	-0.868		3.181
retired	-0.651	***	0.163	-0.310	*	0.173	8.094	*	4.883
riskattitude	0.008		0.020	0.104	***	0.021	-0.154		0.601
netinc	0.284		0.206	0.525	**	0.220	152.712	***	6.791
netinc2	-0.028		0.025	-0.036		0.027	-10.096	***	0.808
wealth	-0.007		0.020	-0.022		0.021	-3.567	***	0.603
wealth	0.000		0.000	0.000		0.000	0.028	***	0.007
constant	8.804	***	0.245	2.491	***	0.260	-2.971		7.373
# obs.	2184			2184			2184		
R2	0.227			0.046			0.286		

Note: This table presents a regression of subjective expectations concerning the health situation, concerning the German economic situation and concerning the variance of future earnings on household characteristics. The table shows in particular that an increase in age is associated with significantly worse expectations concerning the development of the health situation. The findings from this regression – a strong negative effect of the age-variable on the expectation concerning the development of future health, no significant effect of the age-variable for the expectations concerning the development of the German economic situation, and a positive but insignificant effect for expectations concerning earnings variance – remain the same if I include higher order terms of the age variable (age^2 and age^3) and test for joint significance.

6.4 Multinomial Probit Model for Saver Types

The multinomial probit model allows to relax the assumption of independence of irrelevant alternatives by estimating the variance-covariance parameters of the latent-variable errors, instead of imposing that errors are independently and identically distributed according to a type 1 extreme value distribution. I have not motivated the multinomial choice analysis in section 4.2 based on an additive random utility choice framework, since I consider the underlying econometric model less as a behavioral model of choice in this context but rather as a descriptive analysis of the statistical association between saver types and saving motives. In this line, the purpose of the multinomial probit analysis presented in the appendix is not to claim that a different behavioral structure describes this association better, but only to show that even if I relax the IIA assumption by allowing for correlation between the latent-variable errors, the conclusions from this paper still hold. The multinomial probit model assumes that the stochastic error terms have a multivariate normal distribution. As described by Train (2003), the model requires normalization since both the location and scale of the latent variable are irrelevant. To normalize location, I choose – as in the multinomial logit model – saver type 3 (irregular savers) as the base alternative. To normalize for scale, I fix the diagonal elements to 1. While this still imposes some structure on the covariance matrix that is necessary for identification since the model does not include alternative specific variables, it still allows for correlation between the error terms of saver type 1 and saver type 2, which the multinomial logit model does not do.

The results (see table A.4), which are estimated by maximum simulated likelihood, confirm the role of the precautionary and the old-age provision motive that is discussed in the paper. If other categories are chosen as base categories, e.g. saver type 1, and the model allows for correlation between the error terms of other saver types, I find similar results.

Table A.4: Multinomial probit estimation for three different specifications. Base category: Irregular savers.

	(a)		(b)		(c)	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Regular						
age	-0.497	2.495	-1.540	2.483	-0.526	3.015
age2	0.152	0.507	0.347	0.562	0.158	0.619
age3	-0.014	0.038	-0.025	0.043	-0.015	0.035
partner	-0.052	0.190	-0.051	0.201	-0.050	0.191
hhsize	-0.057	0.087	-0.063	0.087	-0.061	0.086
children	-0.030	0.096	0.001	0.087	-0.024	0.097
female	-0.145	0.145	-0.143	0.153	-0.148	0.144
highschool	0.042	0.200	-0.019	0.185	0.037	0.200
civilservant	0.666 *	0.347	0.576 *	0.327	0.660 *	0.353
selfemployed	0.217	0.323	0.363	0.277	0.196	0.335
unemployed	-0.533 **	0.237	-0.480 *	0.256	-0.523 **	0.240
homeowner	0.235	0.231	0.068	0.169	0.190	0.231
unemp_prob	-0.088	0.286	-0.033	0.288	-0.079	0.284
heritage_prob	-0.472	0.594	-0.814 *	0.441	-0.484	0.593
earnings_var	-0.002	0.002	-0.003	0.003	-0.002	0.002
dev_ger_econ_sit	0.015	0.048	-0.020	0.034	0.015	0.049
lifeexpect	0.023	0.058	0.015	0.061	0.022	0.057
dev_health_sit	0.015	0.043	0.027	0.041	0.016	0.041
riskpref	-0.005	0.030	-0.010	0.032	-0.006	0.030
netinc	0.795 *	0.411	0.861 **	0.429	0.731 *	0.430
netinc2	-0.062	0.049	-0.081	0.070	-0.057	0.051
financialwealth			0.508	0.367		
financialwealth2			-0.144	0.095		
wealth					0.020	0.037
wealth2					0.000	0.001
mot_oldage	0.285	0.318	0.043	0.112	0.277	0.323
mot_precaution	-0.121	0.160	-0.047	0.126	-0.119	0.159
mot_homepurchase	0.079	0.144	0.167 *	0.088	0.084	0.143
mot_bequest	-0.015	0.147	-0.109	0.105	-0.021	0.146
constant	0.245	2.643	1.310	3.092	0.288	2.734
Regular, planned, contractual						
age	1.019	2.763	1.738	2.057	1.040	2.679
age2	-0.140	0.498	-0.302	0.502	-0.143	0.467
age3	0.001	0.036	0.011	0.040	0.001	0.033
partner	-0.071	0.176	-0.096	0.176	-0.071	0.177
hhsize	-0.028	0.083	-0.021	0.078	-0.032	0.083
children	-0.075	0.081	-0.076	0.079	-0.068	0.082
female	-0.115	0.139	-0.090	0.133	-0.120	0.138
highschool	0.112	0.160	0.052	0.160	0.108	0.160
civilservant	0.647 *	0.345	0.543 *	0.298	0.641 *	0.350
selfemployed	-0.021	0.303	-0.193	0.255	-0.059	0.308
unemployed	-0.461 **	0.232	-0.353 *	0.210	-0.445 *	0.234
homeowner	0.367 ***	0.143	0.306 **	0.142	0.312 **	0.157
unemp_prob	-0.115	0.251	-0.073	0.251	-0.105	0.248
heritage_prob	0.023	0.445	0.133	0.338	0.029	0.452
earnings_var	-0.001	0.001	0.000	0.001	-0.001	0.001
dev_ger_econ_sit	0.063 *	0.033	0.077 ***	0.029	0.065 **	0.033
lifeexpect	0.027	0.053	0.019	0.052	0.026	0.052
dev_health_sit	-0.008	0.039	-0.022	0.036	-0.008	0.038
riskpref	-0.001	0.028	-0.006	0.027	-0.002	0.028
netinc	0.495	0.476	-0.117	0.360	0.401	0.492
netinc2	-0.027	0.047	0.027	0.040	-0.019	0.048
financialwealth			0.862 ***	0.233		
financialwealth2			-0.055 ***	0.018		
wealth					0.027	0.034
wealth2					0.000	0.001
mot_oldage	0.588 ***	0.121	0.623 ***	0.102	0.590 ***	0.118
mot_precaution	-0.216 *	0.114	-0.241 **	0.110	-0.216 *	0.111
mot_homepurchase	-0.080	0.104	-0.138 *	0.075	-0.079	0.105
mot_bequest	0.118	0.103	0.160 *	0.089	0.116	0.103
constant	-1.787	4.731	-2.865	2.598	-1.807	4.919
# obs.	1066		1066		1066	
Log Likelihood	-996.902		-987.599		-996.342	

7 References

Alessie, R. and A. Kapteyn (2001): New Data for Understanding Saving. *Oxford Review of Economic Policy*, 17 (1), 55-69.

Alessie, R. and A. Lusardi (1997): Saving and income smoothing: Evidence from panel data. *European Economic Review*, 41, 1251-1279.

Alessie, R., A. Lusardi, and T. Aldershof (1997): Income and Wealth over the Life Cycle: Evidence from Panel Data. *Review of Income and Wealth*, 43(1), 1-32.

Alessie, R., A. Lusardi, and A. Kapteyn (1999): Saving after Retirement: Evidence from Three Different Surveys. *Labour Economics,* 6, 277-310.

Ameriks, J., A. Caplin, and J. Leahy (2003): Wealth Accumulation and the Propensity to Plan. *Quarterly Journal of Economics*, 118 (3), 1007-1048.

Artle, R. and P. Varaiya (1978): Life-cycle consumption and homeownership. *Journal of Economic Theory*, 18 (1), 38-58.

Bernheim, D. (2002): Taxation and Saving. In: A. Auerbach and M. Feldstein (Eds.), *Handbook of Public Economics,* Volume 3, 1173-1249. Amsterdam: North-Holland.

Ben Akiva, M., D. McFadden, D. Bolduc, A. Börsch-Supan, P. Delquié, T. Gärling, D. Gopinath, O. Larichev, T. Morikawa, A. Polydoropoulou, and V. Rao (1999): Extended Framework for Modeling Choice Behavior. *Marketing Letters*, 10, 187-203.

Börsch-Supan, A. (1992): Saving and Consumption Patterns of the Elderly. The German Case. *Journal of Population Economics*, 5, 289-303.

Börsch-Supan, A. and A. Lusardi (2003): Saving: A Cross-National Perspective. In: A. Börsch-Supan (Ed.), *Life-Cycle Savings and Public Policy: A Cross-National Study in Six Countries*, 1-32. New York: Academic Press.

Börsch-Supan, A. and K. Stahl (1991): Life-Cycle Savings and Consumption Constraints. *Journal of Population Economics*, 4, 233-255.

Börsch-Supan, A., F. Heiss, and J. Winter (2004): Akzeptanzprobleme bei Rentenreformen: Wie die Bevölkerung überzeugt werden kann. Köln: Deutsches Institut für Altersvorsorge.

Börsch-Supan, A., A. Reil-Held, and R. Schnabel (2003): Household Savings in Germany. In: A. Börsch-Supan (Ed.), *Life-Cycle Savings and Public Policy: A Cross-National Study in Six Countries*, 57-99. New York: Academic Press.

Börsch-Supan, A., A. Reil-Held, R. Rodepeter, R. Schnabel, and J. Winter (1999): Ersparnisbildung in Deutschland: Meßkonzepte und Ergebnisse auf Basis der EVS. *Allgemeines Statistisches Archiv*, 83, 385-415.

Browning, M. and A. Lusardi. (1996): Household Saving: Micro Theories and Micro Facts. *Journal of Economic Literature,* 34, 1797-1855.

Brugiavini, A. and G. Weber (2003): Household Saving: Concepts and Measurement. In: A. Börsch-Supan (Ed.), *Life-Cycle Savings and Public Policy: A Cross-National Study in Six Countries*, 33-55. New York: Academic Press.

Buchinsky, M. (1994): Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression. *Econometrica,* 62 (2), 405-459.

Caballero, R. (1991): Earning Uncertainty and Aggregate Wealth Accumulation. *American Economic Review*, 81, 859-871.

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, 1–56.

Carroll, C. D. and A. A. Samwick (1997): The nature of precautionary wealth. *Journal of Monetary Economics*, 40, 41-71.

Carroll, C. D. and A. A. Samwick (1998): How Important is Precautionary Saving? *Review of Economics and Statistics,* 80 (3), 410-419.

Chesher, A., and M. Irish (1987): Residual analysis in the grouped and censored normal linear model. *Journal of Econometrics*, 34, 33-61.

Chiuri, M. and T. Jappelli (2003): Financial market imperfections and homeownership: a comparative study. *European Economic Review*, 47, 857-875.

Davies, J. B. (1981): Uncertain Lifetime, Consumption, and Dissaving in Retirement. *Journal of Political Economy*, 89, 561-577.

Deaton, A. (1999): Econometric issues for survey data. In: A. Deaton (Ed.), *The analysis of household surveys. A microeconometric approach to development policy.* Baltimore: Johns Hopkins University Press.

Essig, L. (2005a): Methodological aspects of the SAVE data set. *Mea-Discussion-Paper 80-05,* MEA – Mannheim Research Institute for the Economics of Aging. University of Mannheim.

Essig, L. (2005b): Precautionary saving and old-age provisions: Do subjective saving motive measures work? *Mea-Discussion-Paper 84-05,* MEA – Mannheim Research Institute for the Economics of Aging. University of Mannheim.

Essig, L. and J. Winter (2003): Item Nonresponse to Financial Questions in Household Surveys: An Experimental Study of Interviewer and Mode Effects. *MEA-Discussion Paper 39-03*, MEA – Mannheim Research Institute for the Economics of Aging. University of Mannheim.

Friedman, M. (1957): A Theory of the Consumption Function. Princeton: Princeton University Press.

Fuchs-Schündeln, N. and M. Schündeln (2005): Precautionary Savings and Self-Selection: Evidence from the German Reunification "Experiment". *Quarterly Journal of Economics*, 120(3), 1085-1120.

Gervais, S. and I. Goldstein (2004): The effects of biased self-perception in teams. *Working Paper.* Duke University and the Wharton School.

Goldberger, S. (1983): Abnormal selection bias. In: S. Karlin, T. Amemiya, and L. Goodman (Eds.), *Studies in Econometrics, Time Series, and Multivariate Statistics.* New York: Academic Press.

Gourinchas, P. O. and J. Parker (2002): Consumption Over the Life-Cycle. *Econometrica*, 70, 47-89.

Guiso, L., T. Jappelli, and D. Terlizzese (1992): Earnings Uncertainty and Precautionary Savings. *Journal of Monetary Economics,* 30, 307–337.

Guiso, L., T. Jappelli, and D. Terlizzese (1994): Housing finance arrangements, intergenerational transfers and consumption: The Italian experience. *Economic Modelling,* 11 (2), 145-155.

Hausman, J. and D. McFadden (1984): A specification test for the multinomial logit model. *Econometrica*, 52, 1219–1240.

Hayashi, F., T. Ito, and J. Slemrod (1988): Housing finance imperfections, taxation, and private saving: A comparative Simulation analysis of the United States and Japan. *Journal of the Japanese and International Economies*, 2 (3), 215-238.

Heien, T. and K. Kortmann (2003): Spar- und Finanzanlageverhalten privater Haushalte (SAVE II). Methodenbericht. München: Infratest Sozialforschung.

Horioka, C. Y. and W. Watanabe (1997): Why do people save? A Micro-Analysis of Motives for Household Saving in Japan. *The Economic Journal*, 107, 537-552.

Hurd, M. (1979): Estimation in truncated samples when there is heteroscedasticity. *Journal of Econometrics*, 11, 247-258.

Hurd, M. (1987): Savings of the Elderly and Desired Bequests. *American Economic Review*, 77, 298-312.

Hurd, M. (1990): Research on the Elderly: Economic Status, Retirement, and Consumption & Saving. *Journal of Economic Literature*, 28, 565-637.

Hurst, E., A. Lusari, A. Kennickell, and F. Torralba (2005): Precautionary Savings and the Importance of Business Owners. *Working Paper No. 11731.* National Bureau of Economic Research (NBER).

Jürges, H. (2001): Do Germans Save to Leave an Estate? An Examination of the Bequest Motive. *Scandinavian Journal of Economics,* 103 (3), 391-414.

Kazarosian, M. (1997): Precautionary Saving – A Panel Study. *Review of Economics and Statistics*, 79, 241-247.

Kennickell, A. and D. McManus (1994): Multiple Imputation of the 1983 and 1989 Waves of the SCF-panel. Proceedings of the Section on Survey Research Methods, *Annual Meetings of the American Statistical Organization*, Toronto, Canada.

Kennickell, A. and A. Lusardi (2005): Disentangling the Importance of the Precautionary Saving Motive. *Working Paper. Board of Governers of the Federal Reserve System. Washington D.C.*

Kessler, D. and A. Masson (1989): Bequest and wealth accumulation: are some pieces of the puzzle missing? *Journal of Economic Perspectives*, 3 (3), 141-152.

Keynes, J. M. (1936): The general theory of employment, interest and money. London: MacMillan.

Kotlikoff, L. J. (1988): Intergenerational transfers and savings. *Journal of Economic Perspectives*, 2 (2), 41-58.

Kotlikoff, L. J. (Ed.) (2001): Essays on Saving, Bequests, Altruism, and Life-Cycle Planning. Cambridge, MA: The MIT Press.

Kotlikoff, L. J. and L. H. Summers (1981): The role of intergenerational transfers in aggregate capital accumulation. *Journal of Political Economy*, 89 (4), 706-732.

Kurz, M. (1985): Heterogeneity in Savings Behavior: A Comment. In: K. Arrow and S. Honkapohja (Eds.), *Frontiers of Economics.* Oxford: Basil Blackwell.

Leung, S. F. (1994): Uncertain Lifetime, the Theory of the Consumer, and the Life-Cycle Hypothesis. *Econometrica,* 62, 1233-1239.

Lusardi, A. (1998): On the Importance of the Precautionary Saving Motive, *American Economic Review*, 88 (2), 449-453.

Lusardi, A. (1999): Information, Expectations, and Savings. In: H. Aaron (Ed.), *Behavioral Dimensions of Retirement Economics*, New York: Brookings Institution/Russell Sage Foundation.

Lusardi, A. (2000): Precautionary Saving and the Accumulation of Wealth. *Working Paper*. Dartmouth College.

Lusardi, A. and O. S. Mitchell (2006): Financial Literacy and Planning: Implications for Retirement Wellbeing. *Working Paper*. Dartmouth College.

McDonald, J. and Y. J. Xu (1996): A comparison of semi-parametric and partially adaptive estimators of the censored regression model with possibly skewed and leptokurtic error distributions. *Economics Letters*, 51 (2), 153-159.

McFadden, D. (1987): Regression-Based Specification Tests for the Multinomial Logit Model. *Journal of Econometrics*, 34, 63-82.

McFadden, D. (1999): Rationality for Economists. *Journal of Risk and Uncertainty*, 13, 73-105.

Menchik, P. and M. David (1983): Income Distribution, Lifetime Saving and Bequests. *American Economic Review,* 73, 672-690.

Mirer, T. (1979): The Wealth-Age Relation among the Aged. *American Economic Review,* 69, 435-443.

Mirer, T. (1980): The Dissaving Behavior of the Retired Aged. Southern Economic Journal 46, 1197-1205.

Modigliani, F. and R. Brumberg (1954): Utility analysis and the consumption function: An interpretation of cross-section data. In: J. H. Flavell and L. Ross (Eds.), *Social Cognitive Development Frontiers and Possible Futures*. Cambridge, NY: University Press.

Moriizumi, Y. (2003): Targeted saving by renters for housing purchase in Japan. *Journal of Urban Economics*, 53 (3), 494-509.

Paarsch, H. (1984): A Monte Carlo comparison of estimators for the censored regression model. *Journal of Econometrics*, 24, 197-213.

Palumbo, M. (1999): Uncertain Medical Expenses and Precautionary Saving Near the End of the Life-Cycle. *Review of Economic Studies*, 66, 395-421.

Powell, J. L. (1984): Least absolute deviations estimation for the censored regression model. *Journal of Econometrics*, 25, 303-325.

Puri, M. and D. Robinson (2005): Optimism and Economic Choice. *Working Paper No. 11361.* National Bureau of Economic Research (NBER).

Reil-Held, A. (1999): Bequests and Aggregate Wealth Accumulation in Germany. *The Geneva Papers on Risk and Insurance*, 24 (1), 50-63.

Reil-Held, A. (2006): Crowding out or crowding in? Public and private transfers in Germany, Forthcoming: *Journal of European Population.*

Reil-Held, A. and D. Schunk (2006): Pension Reform and Private Old-Age Provision: First Experiences from Germany. Mimeo. Mannheim Research Institute for the Economics of Aging. University of Mannheim.

Rigotti, L., M. Ryan, and R. Vaithianathan (2004): Entrepreneurial Innovation. *Working Paper*. Duke University.

Rubin, D. B. (1987): Multiple Imputation for Nonresponse in Surveys. New York: Wiley.

Samwick, A. A. (2006): Retirement is Only One Aspect of Complex Saving Behavior. *Business Economics*, 41 (1), 21-27.

Schunk, D. (2006): An Iterative Multiple Imputation Procedure for Dealing with Item Nonresponse in the German SAVE Survey. Mimeo. Mannheim Research Institute for the Economics of Aging. University of Mannheim.

Skinner, J. (1988): Risky Income, Life-Cycle Consumption and Precautionary Saving. *Journal of Monetary Economics*, 22, 237-255.

Sourdin, P. (2005): Pension Contribution as a Commitment Device: Evidence of Sophistication among Time-inconsistent Households. Working Paper 2005-17.University of Adelaide - School of Economics.

Train, K. (2003): Discrete Choice Methods with Simulation. Cambridge: Cambridge University Press.

Venti, S. (2006): Choice, Behavior and Retirement Saving. Forthcoming in: G. Clark, A. Munnell, and M. Orzsag (Eds.), *Oxford Handbook of Pensions and Retirement Income*. Oxford: Oxford University Press.

Wärneryd, K.E. (1999): The Psychology of Saving. Cheltenham, UK: Edward Elgar Publishing.

Weil, D. N. (1991): What determines savings? A review essay. *Journal of Monetary Economics*, 28 (1), 161-170.

Yaari, M. (1965): Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *Review of Economic Studies*, 32, 137-150.

An Iterative Multiple Imputation Procedure for Dealing with Item Nonresponse in the German SAVE Survey

Abstract: Important empirical information on household behavior is obtained from surveys. However, various interdependent factors that can only be controlled to a limited extent lead to unit and item nonresponse, and missing data on certain items is a frequent source of difficulties in statistical practice. This paper presents the theoretical underpinnings of a Markov Chain Monte Carlo multiple imputation procedure and applies this procedure to a socio-economic survey of German households, the SAVE survey. I discuss convergence properties and results of the iterative multiple imputation method and I compare them briefly with other imputation approaches. Concerning missing data in the SAVE survey, the results suggest that item nonresponse is not occurring randomly but is related to the included covariates. The analysis further indicates that there might be differences in the character of nonresponse across asset types. Concerning the methodology of imputation, the paper underlines that it would be of particular interest to apply different imputation methods to the same dataset and to compare the findings.

1 Introduction

Important empirical information on household behavior is obtained from surveys. However, various interdependent factors that can only be controlled to a limited extent, such as privacy concerns, respondent uncertainty, cognitive burden of the questions, and survey context, lead to unit nonresponse and item nonresponse. Unit nonresponse is the lack of any information for a contacted survey participant and as such is the strongest type of refusal. The phenomenon that only a subset of the information is missing, e.g. only the response to the question on household income, is referred to as item nonresponse.

The general phenomenon of item nonresponse to questions in household surveys as well as problems of statistical analysis with missing data have been analyzed by various authors, beginning with the work by Ferber (1966) and Hartley and Hocking (1971); see Beatty and Herrmann (2002) for a review. Recent examples for Germany, focusing on income, saving, and asset choice, are Biewen (2001), Riphahn and Serfling (2005), and Schräpler (2003) who work with data from the German Socio-Economic Panel (GSOEP). Finally, Essig and Winter (2003) describe and analyze nonresponse patterns to financial questions in the first wave of the German SAVE survey. They exploit that this first wave has included a controlled experiment specifically designed to analyze the effects of interview mode and question format on answering behavior.

The German SAVE study focuses on details of households' finances, as well as households' sociological and psychological characteristics. For the large majority of variables in SAVE, item nonresponse is not a problem. For example, there is hardly any nonresponse to detailed questions about socio-demographic conditions of the household, to questions about households’ expectations and about indicators of household economic behavior. Mainly due to privacy concerns and cognitive burden, though, there are significantly higher item nonresponse rates for detailed questions about household financial circumstances than for other less private and less sensitive questions. Tables 1 and 2 show that these questions can have a missing rate of over 40%. Similar missing rates for questions about financial circumstances have been documented in various socio-economic survey contexts (e.g., Bover, 2004; Hoynes et al., 1998; Juster and Smith, 1997; Kalwij and van Soest, 2005).

Table 1: Response rates for monthly net income and for the question about total annual saving.

	Value	Bracket	Unknown
Net income	69%	25%	6%
Annual saving	88%		12%

Note: Calculations are unweighted and based on the 2003/2004 wave of the SAVE data.

Table 2: Response rates for financial and real wealth items.

	Have item			Value reported for
	Yes	No	Unknown	those having the item
Savings/term accounts	56%	36%	8%	74%
Building society savings agreements	26%	66%	8%	67%
Whole life insurance policies	28%	64%	8%	57%
Bonds	8%	84%	8%	57%
Shares & real-estate funds	18%	74%	8%	61%
Owner occupied housing	47%	49%	4%	96%

Note: Calculations are unweighted and based on the 2003/2004 wave of the SAVE data.

For studies that use the detailed financial information in the SAVE study, missing information on one of those variables is a problem. It is tempting and still very common to simply delete all observations with missing values. But deleting observations with item nonresponse, i.e. relying on a complete-case analysis, might lead to an efficiency loss due to a smaller sample size and to biased inference when item nonresponse is related to the variable of interest.[1] Particularly for multivariate analyses that involve a large number of covariates, case deletion procedures can discard a high proportion of subjects, even if the per-item rate of missingness is rather low.

The purpose of this paper is to present and discuss the theoretical underpinnings and the practical application of an iterative multiple imputation method that has been developed for the German SAVE dataset. Missing item values are imputed controlling for observed characteristics of nonrespondents and respondents in order to preserve the correlation structure of the dataset as much as possible. The method yields a multiply imputed and complete data set that can be analyzed by the public using standard software packages without discarding any observed cases. In contrast to single imputation, multiple imputation allows the uncertainty due to imputation to be reflected in subsequent analyses of the data (see, e.g., Rubin, 1987; Rubin, 1996; Rubin and Schenker, 1986).

[1] See, e.g., Rubin (1987) and Little and Rubin (2002) for discussions about efficiency and bias in a missing data context.

Iterative multiple imputation methods have recently been applied to other large-scale socio-economic survey data (Bover, 2004; Kennickell, 1998). The imputation method for the U.S. Survey of Consumer Finances, developed by Arthur Kennickell, has been applied to the Spanish Survey of Household Finances (Bover, 2004), and it has also inspired the development of the imputation method that is presented in this paper. The convergence properties of such an iterative procedure have so far only been analyzed systematically on simulated datasets and small datasets with only few variables (e.g., Schafer, 1997); as well, in the context of survey data, the effects of imputation on the resulting distributions of imputed variables have only been documented and compared based on non-iterative imputation methods that focus on specific variables such as income (Frick and Grabka, 2005), or wealth items (Hoynes et al., 1998). The specific contribution of this paper is to investigate the convergence properties of an iterative imputation method that is applied to a large socio-economic survey, the German SAVE survey, and to analyze the resulting distributions of various imputed financial survey items. The latter gives insights about item nonresponse behavior of the survey participants and about the bias that would result from a complete-case analysis.

The paper is organized as follows: Section 2 gives an overview of the SAVE survey, section 3 describes the theoretical underpinnings of the iterative imputation algorithm, develops and documents the application of this algorithm to the SAVE survey, and describes its relationship to existing work on imputation in large surveys. Section 4 investigates the convergence properties of the algorithm and compares imputed and observed data. Section 5 discusses the presented algorithm and concludes the paper.

2 The SAVE Survey – An Overview

In Germany, there has been no dataset available that records detailed data on both financial variables such as income, savings, and asset holdings and on sociological and psychological characteristics of households. The German Socio-Economic Panel (*German SOEP*) has rich data on household behavior and records indicators of saving and asset choices; in 1988 and 2002, the quantitative composition of households' assets was covered in much more detail. Another representative survey, *Soll und Haben,* records detailed data on the composition of various financial assets, but it only has qualitative indicators and does not quantify asset holdings. Finally, the official budget and expenditure survey (*Einkommens- und Verbrauchsstichprobe, EVS*), conducted every five years by the Federal Statistical Office, has very detailed information on the amount and composition of income, expenditure, and wealth, but information on other household characteristics is very limited, in particular in the most recent waves in 1998 and in 2003. Taking as a basis the Dutch CentER Panel and the U.S. Health and Retirement Study (HRS), researchers of the University of Mannheim have cooperated with the Mannheim Center for Surveys, Methods and Analyses (ZUMA), NFO Infratest (Munich), Psychonomics (Cologne) and Sinus (Heidelberg) to produce a questionnaire on households' saving and asset choice; see Börsch-Supan and Essig (2005). The questionnaire has been designed in such a way that the interview should not exceed 45 minutes and was first fielded in 2001 using a quota sampling design. The first random sample was drawn in 2003.[2] The questionnaire consists of six parts (see table 3).

Table 3: Structure of the questionnaire of the SAVE Survey.

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

The first, relatively short part explains the purpose of the study and describes the precautions that have been taken with respect to confidentiality and data protection. Part 2 lasts about 15 minutes and contains questions on the socio-economic structure of the household, including age, education and labor-force participation of the respondent and his or her spouse. Part 3 of the questionnaire contains qualitative and simple quantitative questions on saving behavior and on how households deal with income and assets, including hypothetical choice tasks and questions on saving motives; questions are also asked on financial decision processes, rules of thumb, and attitudes towards consumption and money. Part 4 is the critical part of the questionnaire. It contains a comprehensive financial review of the household and therefore the most sensitive questions in financial items such as income from various sources and holdings of various assets. 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 5 contains questions about psychological and social variables. It includes the social environment,

[2] A description of SAVE and further details on methodological aspects of the SAVE survey are found in Schunk (2006).

expectations about income, the economic situation, health, life expectancy and general attitudes to life. The interview ends with open-ended questions about the interview situation, and a question that asks whether the respondent would be willing to participate in a similar survey in the future (part 6).

3 A Multiple Imputation Method for SAVE

3.1 Motivation and Theoretical Underpinnings

To deal with item nonresponse, one can resort to a complete-case analysis, to model-based approaches that incorporate the structure of the missing data, or one can use imputation procedures. A complete-case analysis may produce biased inference, if the dataset with only complete observations differs systematically from the target population; weighting of the complete cases reduces the bias but generally leads to inappropriate standard errors. Additionally, a complete-case analysis leads to less efficient estimates, since the number of individuals with complete data is often considerably smaller than the total sample size.[3] Formal modeling that incorporates the structure of the missing data involves basing inference on the likelihood or posterior distribution under a structural model for the missing-data mechanism and the incomplete survey variables, where parameters are estimated by methods such as maximum likelihood. Multiple imputation essentially is a way to solve the modeling problem by simulating the distribution of the missing data (Rubin, 1996). Ideally, the imputation procedures control for all relevant observed differences between nonrespondents and respondents, such that the results obtained from the analysis of the complete dataset are less biased overall and estimates are more efficient than in an analysis based on complete cases only.

The goal of imputation is not to create any artificial information but to use the existing information in such a way that public users can analyze the resulting complete dataset with standard statistical methods for complete data. It is often seen as the responsibility of the data provider to provide the imputations: First, because imputation is a very resources-consuming process that is not at the disposal of many users. Second, because some pieces of information which are very useful for the imputation, such as information on interviewer characteristics, are not available to the public. Users are free to ignore the imputations, all imputed values are flagged.

[3] Rubin (1987) and Little and Rubin (2002) illustrate and discuss biased inference and efficiency losses based on complete-case analyses and weighted complete-case analyses.

Assumptions

Many different statistical imputation methods exist and are applied in a variety of data contexts. Examples are mean or median imputation, hotdeck imputation and regression-based imputation. Hotdeck is a very frequently used nonparametric method (e.g., in the RAND-HRS). For hotdeck, only very few conditioning variables can be used, even when the dataset is very large. Regression-based imputations need parametric assumptions. Since regression-based methods allow for conditioning on many more variables than hotdeck methods, they are better than hotdeck methods in preserving a rich correlation structure of the data, provided that an appropriate parametric assumption is made.

Ideally, to impute the missing values, a statistical model should be explicitly formulated for each incomplete survey variable and for the missing-data mechanism. The parameters should then be estimated from the existing data (and from potentially available further information, such as information about the interview process) by methods such as maximum likelihood. Identifying the probability distributions of the variables under study is often very hard and requires weakly motivated assumptions, since the mechanisms of nonresponse are often very complex (Manski, 2005).

Clearly, imputation methods have to make some statistical assumption about the nonresponse mechanism and about the distribution of the data values in the survey.[4] For the imputation method presented in this paper, the underlying assumption about the way in which missing data were lost is that missing values are *ignorable*. To define the ignorability assumption, let us first define *missing at random* (MAR):[5]

Suppose that Y is a variable with missing data and X is a vector of always observed variables in the dataset. Then, formally:

$$Y \text{ is } MAR \Longrightarrow P(Y \text{ is observed} \mid X, Y) = P(Y \text{ is observed} \mid X)$$

That is, after controlling for information in X, the probability of missingness of Y is unrelated to Y.[6] MAR implies that the imputation method should condition on all

---

[4] The Bayesian nature of the presented imputation algorithm also requires specification of a prior distribution for the parameters of the imputation model. In practice, unless the data are very sparse or the sample is very small, a noninformative prior is used (see Schafer (1997) for details). Based on Schafer (1997), it can be concluded that the data in the SAVE survey are neither sparse, nor is the sample small. Consequently, I do not make any assumption about the prior distribution of parameters.

[5] Note that the MAR assumption cannot be tested from available data (Cameron and Trivedi, 2005).

[6] MAR does not imply that the missing values are a random subsample of the complete dataset. This latter condition is much more restrictive and is called ‘missing completely at random’ (MCAR). See Little and Rubin (2002) for further discussions.

variables that are predictive of the missingness of Y, since MAR may no longer be satisfied if variables that determine the nonresponse are not included as conditioning variables (Schafer, 1997).

The missing data mechanism is said to be *ignorable*, if, (a), the data are MAR and, (b), the parameters for the missing data generating process are unrelated to the parameters that the researcher wants to estimate from the data.[7] Ignorability is the formal assumption that allows one to, first, estimate relationships among variables between observed data and, then, use these relationships to obtain predictions of the missing values from the observed values.

Of course, for these relationships to yield unbiased predictions, one would need the correct model for the observed and missing values. The imputation method presented in this paper relies on simple parametric assumptions for all core variables with high rates of missingness[8] and the method uses nonparametric hotdeck methods for discrete variables with only few categories and with very low rates of missingness. The fact that data have been multiply imputed increases robustness to departures from the true imputation model considerably compared to single imputation approaches that are based on the same imputation model. This has been demonstrated in simulation studies (Ezzati-Rice et al., 1995; Graham and Schafer, 1999; Schafer, 1997). Furthermore, using simulated and real datasets from different scientific fields and with varying rates of item nonresponse, existing research emphasizes the robustness of multiple imputation to the specifically chosen imputation model, given that appropriate conditioning variables are available in the dataset (e.g., Schafer, 1997; Bernaards et al., 2003).

The imputation method used for SAVE aims at capturing all relevant relationships between variables in order to preserve the correlation structure between the variables. The method therefore conditions on as many relevant and available variables as possible in the imputation of each single variable. All possible determinants of the variable to be imputed are included as predictors of that variable. Additionally, as has been argued above, including all variables that are potential predictors of missingness makes the MAR-assumption more plausible, because this assumption depends on the availability of

[7] In the literature, MAR and "*ignorability*" are often treated as equivalent under the assumption that condition (b) for ignorability is almost always satisfied (Cameron and Trivedi, 2005).

[8] In line with other iterative or non-iterative and regression-based imputation methods for survey data, e.g. Bover (2004), Frick and Grabka (2005), and Kennickell (1998), I generally assume a linear model for the imputation of continuous variables with high missingness.

variables that can explain missingness and that are correlated with the variable to be imputed.[9]

Multiple Imputation

Single imputation does not reflect the true distributional relationship between observed and missing values and it does not allow the uncertainty about the missing data to be reflected in the subsequent analyses. Estimated standard errors are generally too small (see also appendix, section 6.2), and even if an appropriate imputation model is chosen, single imputation is more prone to generate biased estimates than multiple imputation. These defects – documented and discussed in, e.g., Li et al. (1991) and Rubin and Schenker (1986) – can seriously affect the subsequent interpretation of the analyses.

In multiple imputation, M>1 plausible data sets are generated with all missing values replaced by imputed values. All M complete datasets are then used separately for the analysis and the results of all M analyses are combined such that the uncertainty due to imputation is reflected in the results (see appendix, section 6.2). Briefly, multiple imputation simulates the distribution of missing data and the resulting overall estimates then incorporate the uncertainty about which values to impute. This involves two types of uncertainty: Sampling variation assuming the mechanisms of nonresponse are known and variation due to uncertainty about the mechanisms of nonresponse (Rubin, 1987).

Unless the fraction of missing data is extremely large, it is sufficient to obtain a relatively small number M of imputed datasets, usually not more than five, which is the choice for M in the SAVE imputation method.[10] The relative gains in efficiency from larger numbers are minor under the rates of missing data that are observed in surveys such as the SAVE survey.[11]

Markov Chain Monte Carlo Simulation

Tanner and Wong (1987) present an iterative simulation framework for imputation based on an argument that involves the estimation of a set of parameters from conditioning information that is potentially unobserved. I review briefly their arguments to motivate the iterative imputation method that is used for the SAVE study:

---

[9] Details about the inclusion of conditioning variables in the SAVE imputation method are discussed in section 3.2.4.

[10] Both, the Spanish Survey of Household Finances (Bover, 2004) and the U.S. Survey of Consumer Finances (Kennickell, 1998) also provide 5 imputations.

[11] Rubin (1987) and Schafer (1997) define efficiency in the context of multiply imputed datasets and discuss the choice of M and its impact on efficiency in detail.

Let x_u be unobserved values of a larger set x and let $x_o=x\backslash x_u$. X_u is the sample space of the unobserved data, θ is a set of parameter values to be estimated for which the parameter space is denoted by Θ. The desired posterior distribution of the parameter values, given the observed data, can be written as:

$$f(\theta \mid x_o) = \int_{X_u} f(\theta \mid x_o, x_u) f(x_u \mid x_o) dx_u \quad (1)$$

Here, $f(\theta \mid x_o, x_u)$ is the conditional density of θ given the complete data X, and $f(x_u \mid x_o)$ is the predictive density of the unobserved data given the observed data. The predictive density of the unobserved data given the observed data can be related to the posterior distribution that is shown above as follows:

$$f(x_u \mid x_o) = \int_{\Theta} f(x_u \mid \phi, x_o) f(\phi \mid x_o) d\phi \quad (2)$$

The basic idea of Tanner and Wong is that the desired posterior is intractable based on only the observed data, but it is tractable after the data are augmented by unobserved data x_u in an iterative framework. The suggested iterative method for the calculation of the posterior starts with an initial approximation of the posterior. Then, a new draw of x_u is made from $f(x_u \mid x_o)$ given the current draw from the posterior $f(\theta \mid x_o)$, and this draw is then used for the next draw of $f(\theta \mid x_o)$. Tanner and Wong show that under mild regularity conditions, this iterative procedure converges to the desired posterior.

In an imputation framework, the target distribution is the joint conditional distribution of x_u and θ, given x_o. Based on the ideas of Tanner and Wong, the iterative simulation method is summarized as follows: First, replace all missing data by plausible starting values. Given certain parametric assumptions, θ can then be estimated from the resulting complete data posterior distribution $f(\theta \mid x_o, x_u)$. Let now θ^t be the current value of θ. The next iterative sample of x_u can then be drawn from the predictive distribution of x_u given x_o and θ^t:

$$x_u^{t+1} \sim f(x_u \mid x_o, \theta^t) \qquad \text{[Imputation step (I-step)]} \quad (3)$$

The next step is again to simulate the next iteration of θ from the complete data posterior distribution:

$$\theta^{t+1} \sim f(\theta \mid x_o, x_u^{t+1}) \qquad \text{[Prediction step (P-step)]} \quad (4)$$

Repeating steps (3) and (4), i.e. sequential sampling from the two distributions, generates an iterative Markovian procedure $\{(\theta^t, x_u^t) : t = 1,2,...,N\}$. For the purpose of imputation, this procedure yields a successive simulation of the distribution of missing values, conditioned on both, observed data and distributions of missing data previously simulated. The set of conditioning variables in this algorithm is not necessarily the entire set of all possible values (Tanner and Wong, 1987). Geman and Geman (1984) apply a similar procedure in the field of image processing and show that the stochastic sequence is a Markov chain that has the correct stationary distribution under certain regularity conditions. Li (1988) presents an additional formal argument that the process moves closer to the true latent distribution with each iteration and finally converges. The method is called Markov Chain Monte Carlo (MCMC) because it involves simulation and the sequence is a Markov Chain. Formally, the method is also related to Gibbs sampling (Hastings, 1970), and in the missing data literature, it is often referred to as data augmentation. This method has been used in many statistical applications (e.g., Bover 2004; Kennickell, 1998; Schafer 1997). Sequential simulation algorithms of the MCMC-type can be modified and implemented in different ways, I briefly come back to this issue in section 5.

3.2 The MIMS-Model

3.2.1 Variable Definitions

The multiple imputation method for SAVE (MIMS) distinguishes between core variables and non-core variables. The core variables have been chosen such that they cover the financial modules of the SAVE survey that involve all questions related to income, saving(s), and wealth of the household. The non-core variables include socio-demographic and psychometric variables, as well as indicator variables for household economic behavior. Except for the participation questions of the core variables (e.g., "Did you or your partner own asset X?") and the question about the value of owner-occupied housing, all core variables have missing rates of at least 6%. The non-core variables have considerably lower missing rates, in almost all cases much less than 2%. The following variables (grouped into three categories) are defined as core-variables:

- *Income variables (E):* 40 binary variables indicating income components, 1 continuous variable for monthly net income, and 1 ordinal variable indicating net income in follow-up brackets.
- *Savings variables (S):* 1 binary variable indicating whether the household has a certain savings goal, 1 continuous variable indicating the amount of this savings goal, and 1 continuous variable indicating the amount of total annual saving.
- *Asset variables (A):* 48 binary variables indicating asset ownership and credit, 44 continuous variables indicating the particular amounts.

All other variables in the dataset are non-core variables.

3.2.2 Algorithmic Overview

MIMS is a multiple imputation procedure that is based on the idea of a Markovian process that I have described in the previous subsection. The general algorithmic structure of MIMS is similar to the FRITZ imputation method that is used for the multiple imputation of the Survey of Consumer Finances and for the Spanish Survey of Household Finances (Kennickell, 1998; Bover, 2004). To set the stage for a more detailed discussion of MIMS in the next section, this section gives a brief algorithmic overview of MIMS.

For this purpose, all variables are categorized as follows:

- All variables that are not core variables are called other variables, **O**.
- **P** is a subset of **O**, the subset of all variables that is used as conditioning variables or predictors for the current imputation step.
- The union of all variables from **P** and all core variables that are used as conditioning variables for the current imputation step is referred to as the set **C** (= conditioning variables). In the following algorithmic description, **C** always contains the updated information based on the most recent iteration step. It contains, in particular, the imputed core variables that have been obtained in the last iteration step.

The complete imputation algorithm for the SAVE data works as follows:

- Impute all variables using logical imputation, whenever possible.

Outer Loop *– REPEAT 5 times, j = 1,..., 5 (= Generate 5 datasets)*

  - Impute variables from O using (sequential) hotdeck imputation, obtain complete data ***O****.

  - Impute the income variables E using ***P*,*** *obtain complete data* ***E*.***

  - Impute the savings variables S using ***P**** *and* ***E*,*** *obtain complete data* ***S****.

  - Impute the asset variables A using ***P*, E*,*** *and* ***S*,*** *obtain complete data* ***A****.

  Inner Loop *– REPEAT N times (= Iterate N times)*

    - Impute the income variables E using ***C****.*

    - Impute the savings variables S using ***C****.*

    - Impute the asset variables A using ***C****.*

  Inner Loop *– END*

Outer Loop *– END*

The five repetitions in the outer loop generate one imputed dataset each. After the complete algorithm, five complete datasets are obtained, which I henceforth refer to as implicates. The algorithm generates an additional flag-dataset which contains binary indicators that identify for each value whether it has been imputed or observed.

3.2.3 Description of MIMS

As the algorithmic description shows, MIMS follows a fixed path through the dataset. The first step of the procedure consists of logical imputation. In many cases, the complex tree structure of the SAVE survey or cross-variable relationships allow for the possibility to logically impute missing values. The following path through the dataset is guided by the knowledge of the missing item rates and by cross-variable relationships. The path starts with variables with low missing rates, such that those variables can subsequently be used as conditioning variables for variables with higher missing rates. For example, among the core variables, the net income variable is imputed first, since its missing rate is generally lower than the missing rates of other core variables.[12] The algorithmic description shows that as soon as the iteration loop starts, all variables are already imputed, i.e. starting values for the iteration process have been obtained, and all variables can be used as conditioning variables during the iteration.

Each variable is imputed based on one of the following three general methods:[13]

(1) For all *categorical or ordinal variables* with only few categories and with a low missing rate, a hotdeck procedure with several conditioning variables is used.

(2) For all *binary, categorical, or ordinal core variables,* binomial or ordered Probit models are used.

(3) For all c*ontinuous or quasi-continuous variables,* randomized linear regressions with normally distributed errors are used. This regression procedure, in particular the handling of constraints and restrictions, follows Bover (2004) and Kennickell (1998). First, the conditional expected value is estimated and an error term, drawn from a symmetrically censored normal distribution, is added. This normal distribution has mean zero and its variance is the residual variance of the estimation. The error term is always restricted to the central three standard deviations of the distribution in order to avoid imputing extreme

[12] The lower missing rate for the net income variable is – at least partly – due to the survey design. The net income question was presented using an open-ended format with follow-up brackets for those who did not answer the open-ended question. The imputation of the bracket answers is described later in this paper.

[13] These methods and their application to binary, categorical, ordinal and (quasi-)continuous variables with high and low missing rates are illustrated and discussed in more detail in Little and Rubin (2002).

values. In few cases, logical or other constraints require that the error term has to be further restricted; examples are non-negativity constraints. The imputed value is also restricted to lie in the observed range of values for the corresponding variable. That is, in particular, imputed values will not be higher than observed values for a certain variable.

Due to the skip patterns in the questionnaire, the SAVE data have a very complex tree structure that imposes a logical structure and that has to be accounted for in the imputation process. Further constraints stem from these logical conditions of the data, from the ranges provided (e.g., bracket respondents), from cross-relationships with other variables, or from any prior knowledge about feasible outcomes. For several variables, the specification of all relevant constraints is the most complex part of the imputation software. If necessary, the procedure draws from the estimated conditional distribution limited to the central three standard deviations, until an outcome is found that satisfies all possible constraints that apply in the particular case.

Two remarks are important at this point to gain an understanding of key procedures of the algorithm.

(1) Ownership and amount imputations

For certain quantities, e.g. the amount of assets held by a household, the SAVE survey uses a two-step question mode: In step one, households are asked about ownership of assets from a certain asset category and a binary variable records the answer. In step two, those households that have reported that they own assets from the particular category are asked about the exact value of the corresponding assets. From a modeling point of view, this is a corner solution application. Following Bover (2004) and Kennickell (1998), a hurdle model is used in MIMS to impute the missing values in these two steps: First, a Probit model is estimated for the binary ownership variable, and missing information is predicted. Then, as described above, randomized linear regressions with normally distributed errors are used for imputing continuous amounts. These regressions are estimated based on all observations that own the asset. Alternatively, Tobit models or sample-selection models might be appropriate. Tobit models are less attractive for the given problem, since they include the implicit assumption that the model governing selection and the model governing the estimation of the amounts are the same. Heckman selection models are theoretically attractive, but cause estimation problems in practice: First, the necessary exclusion restrictions differ substantially across asset categories, but there is no theoretical reason why they should differ. Second, in most cases, strong exclusion restrictions are needed to ensure identification and convergence of the Heckman procedure in each iteration step of MIMS. This means that in practice only a very small set of conditioning variables can be used for the estimation of the second step of the Heckman model. Under these circumstances and given that the goal of the multiple imputation method is to simulate the distribution of amounts conditional on ownership and conditional on a maximally large set of potentially correlated variables, MIMS uses hurdle models for ownership and amount imputations.

(2) Net income variables

To alleviate the problem of item nonresponse to income questions (see, e.g., Juster and Smith, 1997), the survey question on monthly net income was presented using an open-ended format with follow-up brackets for those who did not answer the open-ended question. That is, there are two types of income information available: Exact (in the sense of point data) income information for households that answered the open-ended question, and interval information on household income for those who only answered the bracket question. To make best possible use of all the available income information, the imputation procedure uses a maximum-likelihood estimation procedure. The likelihood is a mixture of discrete terms (for the interval information) and continuous terms (for the point data information). After prediction of the missing income values and the addition of the randomized error term, a nearest neighbor approach is used to determine the imputed amount for household net income.[14] The procedure works as follows: First, an income bracket is predicted for all complete nonrespondents to both (i.e., open-ended *and* bracket) income questions. Now, all observations have either exact income information (if they have reported this information) or bracket information (either they have reported this information, or it has been imputed in the preceding step). Then, each observation i for whom an exact net income value has to be imputed and whose net income lies in bracket j is matched with the continuous reporter r from bracket j whose predicted net income value is closest to the predicted value of respondent i. The net income value assigned to observation i is then the reported continuous income value of the respondent r.[15]

---

[14] Nearest neighbor methods have been motivated in a statistical missing data context by Little et al. (1988) and they have subsequently used in the context of bracketed follow-up questions by, e.g., Hoynes et al. (1998) in the AHEAD.

[15] In contrast to this procedure, Hoynes et al. (1998) impute the brackets for the full nonrespondents using an ordered Probit model that is estimated using *only* those respondents that have provided bracket answers. The chosen procedure in MIMS has the advantage of making better use of the available information (since it uses the information from bracket respondents *and* from contiuous, i.e. open-ended, respondents) and it

3.2.4 Selection of Conditioning Variables

As is clear from the descriptions above, each regression or hotdeck method is tailored specifically to the variable to be imputed.[16] Of particular importance are the conditioning variables which have been selected individually for every single variable with missing information according to the following guidelines:

(A) Hotdeck imputations: *Hotdeck imputations,* which have been used for discrete variables with very low missing rates, allow for only few and discrete conditioning variables due to the quickly increasing number of the corresponding conditioning cells. The conditioning variables have first been selected based on theoretical relationships if available and, second, based on the strength of a correlation with the variable to be imputed; those correlations have been systematically explored. As an example for the latter, consider the question which asks respondents to rate their expectation concerning the future development of their own health situation on a scale from 0 (negative) to 10 (positive), which has a missing rate of 0.6%. As conditioning variables, the respondents' age (subdivided into five age classes), self-assessed information on the respondents' current health status (rated on a scale from 0 to 10 and subdivided into three classes), and self-assessed information on how optimistic the respondent generally is (rated on a scale from 0 to 10 and subdivided into three classes) are used.[17] All these conditioning variables are significantly correlated with the variable to be imputed, both individually, as well as jointly in a multiple regression. In some cases, it would be desirable to include core variables as additional conditioning variables in the hotdeck imputations. For example, net income is clearly expected to be correlated with educational status. Generally, the pattern of nonresponse makes this impossible, since the set of nonrespondents to the qualitative questions is in almost all cases a subset of the set of nonrespondents to the relevant core questions.

---

circumvents the practical problem in SAVE that the subsample of bracket respondents is too small to be able to include much conditioning information into the estimation of an ordered Probit model. Hoynes et al. (1998) motivate their procedure by arguing that full nonrespondents are more similar to bracket respondents than to continuous reporters. Note, however, that the evidence on the similarity between nonrespondents, bracket respondents and continuous respondents is mixed (Kennickell, 1997).

[16] A spreadsheet with information on the specific imputation methods for each imputed variable in SAVE (e.g., hotdeck, various regression techniques), as well as information on the used conditioning variables can be obtained from the author upon request.

[17] Note that these three conditioning variables already correspond to $5 \cdot 3 \cdot 3 = 45$ different cells.

(B) Regression-based imputations: In theory, every *regression-based imputation* should use all relevant variables in the dataset, as well as higher powers and interactions of those terms as conditioning variables (see section 3.1 and Little and Raghunathan, 1997). The imputation procedure should, in particular, attempt to preserve the relationships between all variables that might be jointly analyzed in future studies based on the imputed data (Schafer, 1997). In practice, a limit to the number of included conditioning variables is imposed by the degrees of freedom of the regressions. Additionally, there must not be collinearity between conditioning variables, which can easily arise in some cases due to the tree structure of the questions. Due to these constraints concerning the inclusion of conditioning variables, it is of particular importance to select these variables following certain guidelines such that best possible use is made of the available information. For that purpose, the variables used in the regression-based imputations of the core variables have been classified into three non-disjoint categories:

(B-1) *Determinants of the nonresponse.*

Research in psychology, economics, and survey methodology has investigated the relationship between observed respondent and household characteristics and item nonresponse behavior in various survey contexts (for an overview, see Groves et al., 2002). Findings from empirical studies that focus particularly on financial survey items suggest that certain variables might be useful predictors of nonresponse to wealth and income questions (Hoynes et al., 1998; Riphahn and Serfling, 2005). Following these findings, MIMS considers the following variables as determinants of nonresponse to the core variables: Age (as well as squared and cubic age), gender, dummy variables for educational achievement and employment status, as well as household size. Riphahn and Serfling (2005) and Schräpler and Wagner (2001) provide evidence that it is not only the individual respondent's characteristics that may be associated with item nonresponse to financial variables, but also the combination of interviewer and respondent characteristics. In this spirit, the following variables that capture the relationship between interviewer and interviewee characteristics are also considered as determinants of nonresponse to the core financial variables in SAVE: Dummies for whether the interviewer is older than the interviewee, for her/his educational status relative to the interviewee, for the interviewer's gender, and for the gender combination of interviewer and interviewee.

(B-2) *Variables that are related to the variable to be imputed based on different economic models.*

This category contains essentially all core variables, since financial characteristics of households, e.g. saving(s), income and asset categories, are all interrelated. Certain qualitative variables on household socio-economic and financial characteristics that are not already part of the variables in *(B-1)* are also included, for example an indicator for marital status. Variables that measure individual preferences, such as measures for risk attitude, are further included into this category.

(B-3) *Other variables that might be related to the variables to be imputed.*

This category includes variables that are correlated with the variables to be imputed but this relationship is not captured in any formal established economic theory that the author knows of. An example is the smoking habit of the respondent: While there is no formal theory that *directly* relates smoking habits to economic characteristics of a household, there is abundant evidence for a statistically strong association between smoking habits and economic characteristics (e.g., Hersch, 2000; Hersch and Viscusi, 1990; Levine et al., 1997).

The selection of the conditioning variables for the regression is based on the following procedure: First, since the goal is to include as many conditioning variables as possible, all variables from categories *(B-1)*, *(B-2)*, and *(B-3)* are included for each imputation regression. If necessary – because of multicollinearity or insufficient degrees of freedom – variables are removed in the following order: First, variables from *(B-3)* are removed. Then, variables from *(B-2)* are aggregated if possible: E.g., instead of including information on the value of owner-occupied housing and on other real estate as two separate conditioning variables, these two variables can be combined to form a variable for total real estate wealth. In a few cases, notably variables with very low variability, such as the measure of wealth in "other contractually agreed private pension schemes", further conditioning variables from category *(B-2)* have to be removed. In this case, the decision is based on the significance of the variables in the regression. Generally, psychometric variables are removed first and credit variables are removed subsequently, since those variables have the lowest variability and the highest missing rate among the core variables.

4 Results

MIMS has been applied to the 2003/2004 wave of the SAVE survey which contains 3154 observed households and all statistics presented in this section are based on this wave. This section discusses the convergence properties of the algorithm and presents descriptive analyses of the imputed and the observed data. The presented analyses serve to illustrate the differences between the five implicates, the impact of imputation on the distribution of values in the complete dataset, and they are informative concerning the differences in the character of nonresponse across various financial survey items.

4.1 Convergence of MIMS

Assessing convergence of the sequence of draws to the target distribution is more difficult than assessing convergence of, e.g., EM-type algorithms, since there is no single target quantity to monitor, like the maximum value of the likelihood. In this subsection, I first develop a convergence criterion that is based on a measure for the average change in the values of a certain variable vector between two consecutive iteration steps. I then use a standard convergence criterion that is also mentioned in Bover (2004) and which is defined with respect to measures of position and dispersion of the distribution of the variable to be imputed. Both convergence criteria are used for assessing convergence of three core variables of the SAVE survey.

Let us assume first that there is missing information on only one variable Y in the dataset. That is, all conditioning variables are complete data vectors without missing values. Let $Y_{i,t}$ be the imputed value of the variable of interest for household i in iteration step t, and let I be the total number of imputed observations for variable Y in the dataset. Then, the squared change in the value of variable Y between iteration step t and $t-1$ is:

$$s(t) = \frac{1}{I}\sum_{i=1}^{I}(Y_{i,t} - Y_{i,t-1})^2 \qquad (5)$$

If the procedure has converged, the parameters θ that characterize the distribution of the imputed variable have stabilized.[18] That is, after convergence has been achieved, there is no systematic component in the change of Y over iterations steps any more; only a non-

[18] Note: This suggests a further way to assess convergence: One can investigate the degree of serial dependence of a certain parameter value over iteration steps by analyzing the autocorrelation function. Ideally, this has to be done for *all* parameters of the particular imputation model, and it is preferred for datasets with only few variables and a correspondingly small set of conditioning variables and parameters (Schafer, 1997).

systematic component remains. $Y_{i,t}$ and $Y_{i,t-1}$ can then be assumed to be draws from the same distribution. This implies that – as soon as convergence has been achieved – we have:

$$s(t) = \frac{1}{I}\sum_{i=1}^{I}(Y_{i,t} - Y_{i,t-1})^2 = Var(Y_{i,t} - Y_{i,t-1}) = Var(Y_{i,t}) + Var(Y_{i,t-1}) = 2Var(Y_{i,t}) \qquad (6)$$

Indeed, if the procedure has converged, the distribution of the remaining non-systematic component is well known, since it is characterized by the distribution of the simulated error term that is added to the particular predicted value of in each iteration step. I.e., $Var(Y_{i,t})$ can be calculated as the variance of the simulated error term: This error term, ε, is drawn from a normal distribution, the variance of which is – by construction – the residual variance of the particular estimation (see section 3.2.3). This normal distribution is then double censored to the central three standard deviations. I derive the variance of a double censored variable ε in the appendix (see section 6.1).

From these deliberations follows: If the process has converged, $s(t)$, calculated based on the imputed values of the variable $Y_{i,t}$ and $Y_{i,t-1}$, should be equal to $e(t) = 2Var(\varepsilon_t)$, i.e. it should be equal to two times the variance of the simulated error term in iteration step *t*. Furthermore, if convergence has been achieved, $s(t)$ and $e(t)$ are stationary, i.e. they should not have any trend over iterations steps and the sample autocorrelation function for $s(t)$ and $e(t)$ should not indicate autocorrelations at any lag.

In real world data-sets, such as in the SAVE data, it is rarely the case that all conditioning variables are non-missing, as I have assumed for the derivation above. In particular, this condition will not be satisfied in MIMS, since – for reasons given above (see section 3.2.4) – MIMS conditions on as many core variables as possible which have rather high missing rates themselves. But even if the conditioning variables themselves have been imputed, the parameters θ that characterize the distribution of imputed variables should, of course, have stabilized if the process has converged. That is, if the process converges, $s(t)$ and $e(t)$ are stationary, i.e. they should not have any trend over iterations steps, and the corresponding autocorrelation functions should not indicate any autocorrelations. Therefore, displaying $s(t)$ and $e(t)$ over time provides an intuitive graphical way to investigate convergence of the process.[19] Note, however, that the fact that the conditioning variables are also imputed has the effect that $s(t)$ should be in fact larger than $e(t)$ even if the process has converged, since the imputed conditioning variables themselves are drawn from the corresponding posterior distribution in the particular iteration step.

Figure 1 shows $s(t)$ and $e(t)$. Five different iteration runs are shown for $t = 1,\ldots, 30$ and one additional run is shown for $t = 1,\ldots, 100$ in the last row of the figure. The runs are displayed for three variables that are used to assess convergence, one from each category of the core variables.[20] In all simulation runs, $e(t)$ quickly resembles a horizontal line. As expected due to the sample size, $s(t)$ is very volatile. It lies above the value $e(t)$, and after few iterations, it does not exhibit any trend over the following iteration steps.[21] The results indicate quick convergence in the first few iteration steps for *net income* and for *annual saving*. For the *net income* variable, $s(t)$ is lower than $e(t)$[22]; this is due to the nearest neighbor algorithm and the available bracket information for many nonrespondents which reduces variability of a certain imputed value over iteration steps.

A further investigation of the sample autocorrelation functions of $s(t)$ and $e(t)$ does not reveal any correlations. The corresponding autocorrelation data and figures can be obtained from the author upon request.

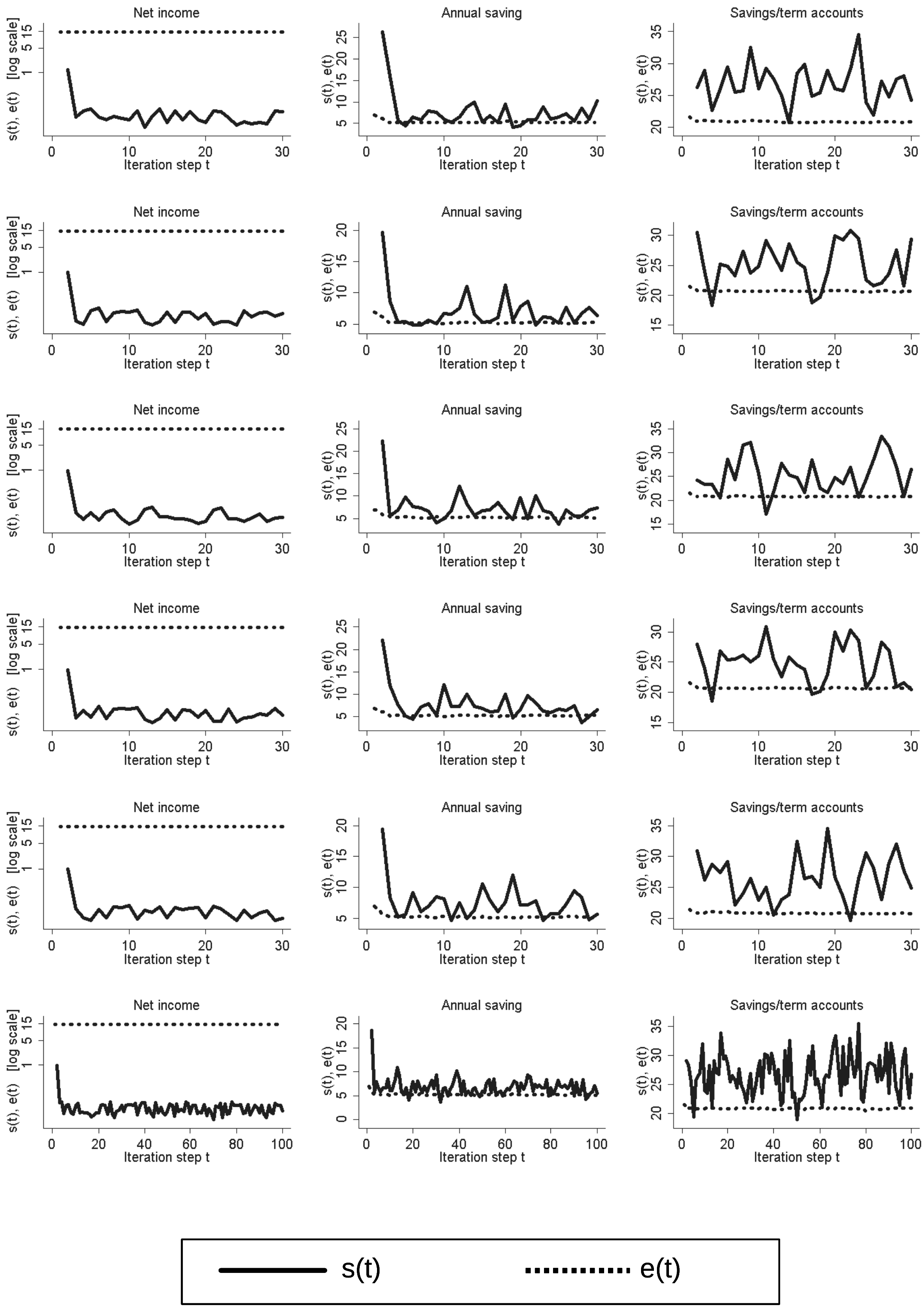
[19] The purpose of these derivations is to suggest a simple graphical convergence diagnostic for an MCMC-method that is applied to a large dataset and that uses a very large set of conditioning variables. I do *not* claim an equivalence result: While convergence of the algorithm would imply that *s(t)* and *e(t)* do not exhibit any downward or upward trend, the converse is not true; i.e. stationarity of *s(t)* and *e(t)* does *not* imply convergence of the algorithm.

[20] Note that only those values for whom no further constraints apply in all iteration steps (e.g., neither non-negativity constraints nor maximum-value constraints), are used for the calculation of *s(t)* and *e(t)*.

[21] If the calculation of *s(t)* is restricted to those observations for which the conditioning variables are almost complete, i.e. non-missing, then the plot reveals that *s(t)* fluctuates around *e(t)*, as predicted. However, the number of observations is even smaller in this case.

[22] Note, that *s(t)* and *e(t)* are plotted on a logarithmic scale for the net income variable in order to be able to plot both variables in one graph.

Figure 1: Convergence diagnostics: *s(t)* and *e(t)* displayed for three key variables.



Note: For net income, s(t) and e(t) are divided by 1,000,000, for annual saving and savings/term accounts, s(t) and e(t) are divided by 10,000,000.

A common criterion for assessing the convergence of a distribution, also suggested in Bover (2004), is to compare (functions of) quantiles, e.g., the median and the interquartile range, resulting from successive iterations of the variable *Y*:

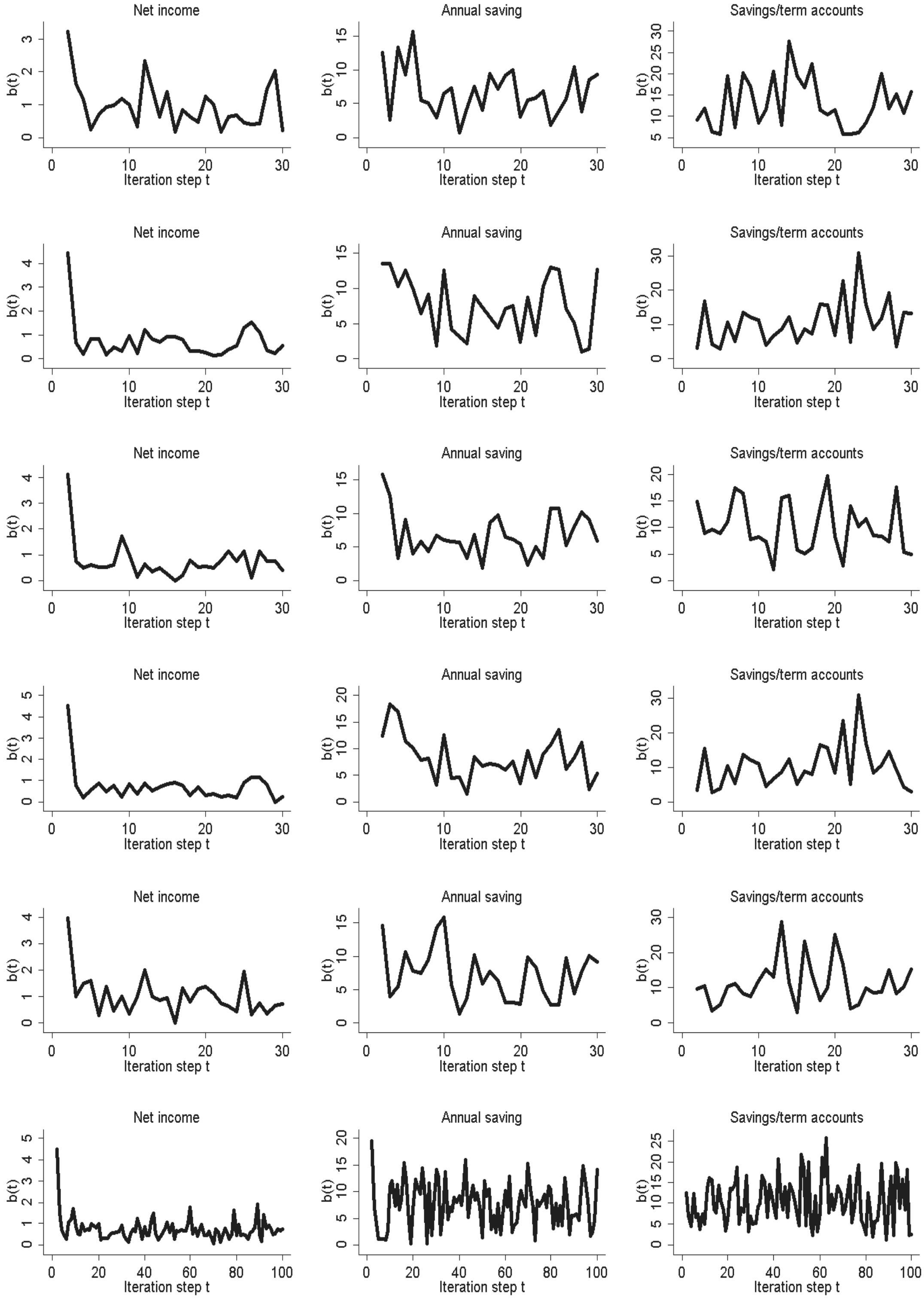
$$b(t) = \sqrt{(Q50_t^Y - Q50_{t-1}^Y, (Q75 - Q25)_t^Y - (Q75 - Q25)_{t-1}^Y)} \cdot \sqrt{(Q50_t^Y - Q50_{t-1}^Y, (Q75 - Q25)_t^Y - (Q75 - Q25)_{t-1}^Y)'} \tag{7}$$

Here, Q25, Q50, and Q75 denote the 25th, 50th, and 75th quantile, respectively, of the particular distribution of imputed values. As long as the process converges, b_t has a downward trend. As soon as the process has converged, $b(t)$ should not exhibit a trend any more. Figure 2 shows b_t for the three variables that are used for convergence diagnosis. As before, five iteration runs are shown for *t = 1,..., 30* and one run is shown for *t = 1,..., 100*. The figures reveal convergence for the *net income* variable, and some indication for convergence of the *annual saving* variable, which is, however, not really convincing.

Overall, the findings from the two convergence diagnostics presented above suggest relatively quick convergence of the algorithm on the *net income* variable, and mixed evidence for the *annual saving* variable. The convergence properties of the algorithm have been investigated on all other core variables. No indication for divergent behavior or long-term drift has been found, in all cases, $s(t)$, $e(t)$, and $b(t)$ are stationary after few iteration steps and no autocorrelation is present in $s(t)$, $e(t)$, and $b(t)$. However, $s(t)$ and $b(t)$ do *not* exhibit a clear downward trend for many variables in the early iteration steps; that is, they are stationary from the first iteration step on. The variable s*avings and term accounts* which is displayed in the presented figures, is an example of such a variable. This result, which is also mentioned by Kennickell (1998), suggests that those variables have essentially converged in the first iteration step; i.e. convergence has already been achieved in the first prediction step which has served to generate the starting values for the iteration.

Note, that iteration runs with t = 300, which are not displayed graphically in this paper, have also been analyzed for both suggested convergence criteria; as well, the corresponding autocorrelation functions have been investigated. The findings show that even longer iteration procedures do not achieve better convergence results based on the presented diagnostics; in particular, no autocorrelation at longer lags is found.

Figure 2: Convergence diagnostics: *b(t)* displayed for three key variables.



Note: Values b(t) are divided by 100.

Overall, the results are in line with findings based on the iterative algorithm implemented for the imputation of the Survey of Consumer Finances (Kennickell, 1998). Kennickell reports quick convergence on key variables, the algorithm is run for 6 iteration steps overall.[23] Given the findings about convergence in this section, MIMS is run for 20 iteration steps, this takes about 2 days per implicate.

4.2 Observed, Imputed, and Complete Data

This subsection has two main purposes: First, the reader should get an impression of the differences across the five imputed and across the five complete data implicates. For this reason, the following tables report descriptive statistics of key financial variables for all five implicates. Second, the section presents and briefly discusses differences between the distributions of observed and imputed data. The section ends with a graphical comparison between observed and imputed data.

The following table 4 reports descriptive statistics for the observed data, for the five imputed implicates, and for the five complete data implicates (complete data implicates consist of observed *and* imputed data, i.e. the full rectangular data matrix). The means of all variables vary across complete data implicates and across imputed data implicates. Medians of all variables vary only across imputed data implicates, not across complete data implicates.[24] I first turn to the financial wealth variables. The table shows a consistent pattern for all financial wealth variables and for the saving variable: The mean of the imputed data is considerably higher than the mean of the observed data. This finding deserves further investigation.

[23] A comparison with similar iterative imputation methods, described in Bover (2004) and Kennickell (1998), would be informative. Bover (2004) and Kennickell (1998) do not present graphical or numerical evaluations of the convergence properties of their imputation method.

[24] The fact that summary distributional characteristics, such as mean values, are similar across implicates is in line with our finding that the imputation for all 5 implicates – which have all started with different initial values for the imputed variables – have indeed converged, and not diverged. Again, longer simulations lead to similar results.

Table 4: Descriptive statistics for the observed data, for the 5 imputed implicates, and for the 5 complete data implicates.

	Observed data	Imputed data Implicate No.					Complete data Implicate No.				
		1	2	3	4	5	1	2	3	4	5
						Net income [€]					
Mean	2,554	2,382	2,390	2,400	2,386	2,388	2,501	2,504	2,507	2,502	2,503
Median	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
Min.	25	25	25	25	25	25	25	25	25	25	25
Max.	120,000	20,000	20,000	23,333	20,000	20,000	120,000	120,000	120,000	120,000	120,000
						Annual saving [€]					
Mean	2,624	5,453	5,336	5,624	5,553	5,784	2,948	2,940	2,971	2,970	2,994
Median	1,000	3,929	3,738	3,895	3,946	3,772	1,000	1,000	1,000	1,000	1,000
Min.	0	0	0	0	0	0	0	0	0	0	0
Max.	150,000	78,206	56,586	98,161	55,435	112,878	150,000	150,000	150,000	150,000	150,000
						Savings/term accounts [€]					
Mean	8,174	12,155	12,272	11,755	12,274	12,129	9,068	9,094	8,978	9,094	9,062
Median	500	10,784	11,176	10,360	11,628	10,436	2,000	2,000	2,000	2,000	2,000
Min.	0	0	0	0	0	0	0	0	0	0	0
Max.	1,000,000	88,897	95,767	116,545	81,713	143,290	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
						Building society savings agreements [€]					
Mean	1,775	3,917	3,873	3,755	3,726	3,907	2,124	2,117	2,098	2,093	2,122
Median	0	1,844	1,805	1,528	1,671	1,972	0	0	0	0	0
Min.	0	0	0	0	0	0	0	0	0	0	0
Max.	100,000	54,442	64,057	58,061	66,725	68,408	100,000	100,000	100,000	100,000	100,000

Note: All calculations are unweighted.

Table 4 (continued)

	Observed data	Imputed data Implicate No.					Complete data Implicate No.				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
					Whole life insurance policies [€]						
Mean	5,042	13,881	14,333	14,393	13,981	13,821	6,813	6,904	6,916	6,833	6,801
Median	0	9,970	10,793	10,380	9,840	9,922	0	0	0	0	0
Min.	0	0	0	0	0	0	0	0	0	0	0
Max.	500,000	196,235	189,196	224,734	198,699	203,240	500,000	500,000	500,000	500,000	500,000
					Bonds [€]						
Mean	1,644	10,237	10,459	11,364	11,291	10,915	2,625	2,650	2,754	2,745	2,702
Median	0	0	0	0	0	0	0	0	0	0	0
Min.	0	0	0	0	0	0	0	0	0	0	0
Max.	1,000,000	316,511	349,122	345,260	403,173	380,301	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
					Shares & real-estate funds [€]						
Mean	3,857	8,511	8,350	8,618	8,291	8,460	4,555	4,531	4,571	4,522	4,547
Median	0	526	558	962	925	845	0	0	0	0	0
Min.	0	0	0	0	0	0	0	0	0	0	0
Max.	18,000,000	250,392	264,187	249,577	270,813	277,270	1,800,000	1,800,000	1,800,000	1,800,000	1,800,000
					Owner occupied housing [€]						
Mean	123,280	44,800	43,388	40,108	38,672	43,639	111,710	111,501	111,018	110,806	111,538
Median	0	0	0	0	0	0	0	0	0	0	0
Min.	0	0	0	0	0	0	0	0	0	0	0
Max.	5,000,000	934,811	1,159,067	1,243,168	876,550	1,248,193	5,000,000	5,000,000	5,000,000	5,000,000	5,000,000

Note: All calculations are unweighted.

For this purpose, table 5 gives information on the imputation of asset variables by showing the results of the ownership imputation. The first column of the table shows the asset ownership rates for those who answer the ownership question, the following columns show imputed ownership rates for all implicates. It is found that except for the item *savings and term accounts*, ownership rates among nonrespondents are in fact lower than ownership rates among respondents. Both findings, namely that the imputation overall leads to higher means for financial asset variables (table 4) but at the same time generates lower ownership rates for financial assets (table 5) is in line with findings by Hoynes et al. (1998) who use a *non*-iterative regression-based single imputation method.[25]

Table 5: Percentage of households owning assets: Observed values and 5 imputed implicates

	Observed data	Imputed data Implicate No. 1	2	3	4	5
Savings/term accounts	60.8	70.5	70.5	70.5	70.5	70.5
Building society savings agreements	27.8	15.5	16.7	15.1	14.7	15.9
Whole life insurance policies	30.4	17.1	16.7	16.7	17.5	17.5
Bonds	8.8	2.0	2.4	1.6	2.4	2.0
Shares & real-estate funds	19.8	9.2	9.2	10.0	9.6	9.2
Owner occupied housing	48.6	35.9	37.6	36.8	35.0	36.8

Note: All calculations are unweighted.

It can be concluded that, for most financial asset items, the included conditioning variables shift the distribution to higher values for financial wealth on average, compared to the original distribution of observed values, which would simply be replicated if no conditioning variables were used. The findings by Smith (1995), who reports that the

[25] Hoynes et al. (1998) find higher mean values for all complete nonrespondents on all comparable financial asset variables. They also find lower imputed ownership rates than observed ownership rates on all financial asset variables, except from the item "bonds" and the item "checking and savings accounts". For these items, they find imputed ownership rates that are similar to the observed rates. A more detailed comparison with results from other imputation procedures would be of high interest at this point. To the author's knowledge, however, a systematic evaluation of the effect of the imputation on the distribution of different wealth components is only presented in the paper by Hoynes et al. (1998). Further methodological insights about the impact and relevance of an iterative procedure could be obtained from comparing an application of the Hoynes et al. (1998)-procedure and the MIMS-procedure to the same dataset.

effect of follow-up brackets to open-ended financial wealth questions in the HRS is a substantial increase in mean wealth, go into the same direction.

In contrast to the findings concerning the financial wealth variables, table 4 shows that the mean of imputed values of owner occupied housing are lower than observed values. How are home ownership and owner-occupied housing values distributed across observed and imputed values? Table 5 has already shown that – according to the imputation – the fraction of homeowners, i.e. households with a positive value for owner-occupied housing[26], is considerably lower among nonrespondents than among respondents. Table 6 serves to further investigate the difference between the observed and the imputed distribution of the value of owner occupied housing. Each column of the table gives the percentage distribution of home values for homeowners across four categories. The table shows that households that did not answer the corresponding question are more likely to occupy real estate with a low value. Interestingly, the results on home-ownership and owner-occupied housing values are again in line with findings by Hoynes et al. (1998), who report that those with incomplete responses on the housing questions have characteristics that make them more likely to be renters, and – given that they are homeowners – it makes them more likely to have low values for real estate.[27]

Table 6: Distribution of owner-occupied housing values for homeowners (percent).

	Observed data	**Imputed data Implicate No.**				
Range (1,000 €)		**1**	**2**	**3**	**4**	**5**
0 - 49.9	9.0	12.9	13.7	14.9	14.1	14.9
50 - 99.9	8.3	12.9	16.8	17.0	18.5	8.5
100 - 199.9	29.3	26.9	27.4	31.9	26.1	34.0
> 200	53.4	47.3	42.1	36.2	41.3	42.6

Note: All calculations are unweighted.

[26] Of course, one can argue that the fraction of homeowners is not equal to the fraction of households with a positive value for owner-occupied housing, since it can also be the case that respondents own real estate and answer that its value is zero. In fact, about 5% of the respondents that report owning real estate give a value of zero in the follow-up question. In all tables above, these respondents are counted as homeowners.

[27] While the purpose of this paper is not to investigate the relationship between item nonresponse to certain questions and socio-economic characteristics, the above findings are interesting in this respect: They suggest that nonrespondents to questions about housing might have other socio-economic characteristics than nonrespondents to the financial wealth questions. A multivariate analysis indeed finds some evidence for this hypothesis.

Finally, I turn to the findings for the *net income* variable. Though medians are identical for imputed and observed values, the mean of monthly net income is lower for the imputed than for the observed values (table 4). For further investigation, table 7 compares the distribution of net income values between imputed and observed data. No substantial difference in the net income distributions of both groups is observable. The reason for the finding that the mean of monthly net income is lower for the imputed than for the observed values are a few extreme values in the observed distribution of monthly net income: If the observed distribution of monthly net income values is trimmed such that the top 0.5-percentile is left out (corresponding to 10 observations that reported having a net income between 26,000 € and 120,000 € per month), a mean monthly net income value of 2,306 € is found. This value is lower than the mean monthly net income of the imputed observations of all five implicates (see table 4); on average by about 83 €.

Table 7: Distribution of monthly net income (percent)

	Observed data	**Imputed data Implicate No.**				
Range (1,000 €)		**1**	**2**	**3**	**4**	**5**
0 - 0.9	13.3	14.1	13.9	14.2	13.8	14.0
1 - 1.99	34.4	33.7	33.3	33.1	33.8	33.1
2 - 2.99	28.3	29.4	29.8	29.3	29.5	30.0
3 - 3.99	13.9	11.8	12.3	12.2	12.1	11.9
4 - 4.99	4.8	6.2	5.9	6.4	6.2	6.2
5 - 6.99	2.8	2.4	2.3	2.2	2.1	2.4
> 7	2.5	2.4	2.5	2.6	2.5	2.4

Note: All calculations are unweighted.

Overall, it is found that MIMS does not have a strong effect on the distribution of income values in SAVE. In contrast, findings from a regression-based single imputation procedure of annual income variables for the SOEP suggest that item nonresponse on income appears to be selective with respect to both tails of the income distribution (Frick and Grabka, 2005); the overall effect of their imputation is an increase in the mean of after-tax income by 1.7%.

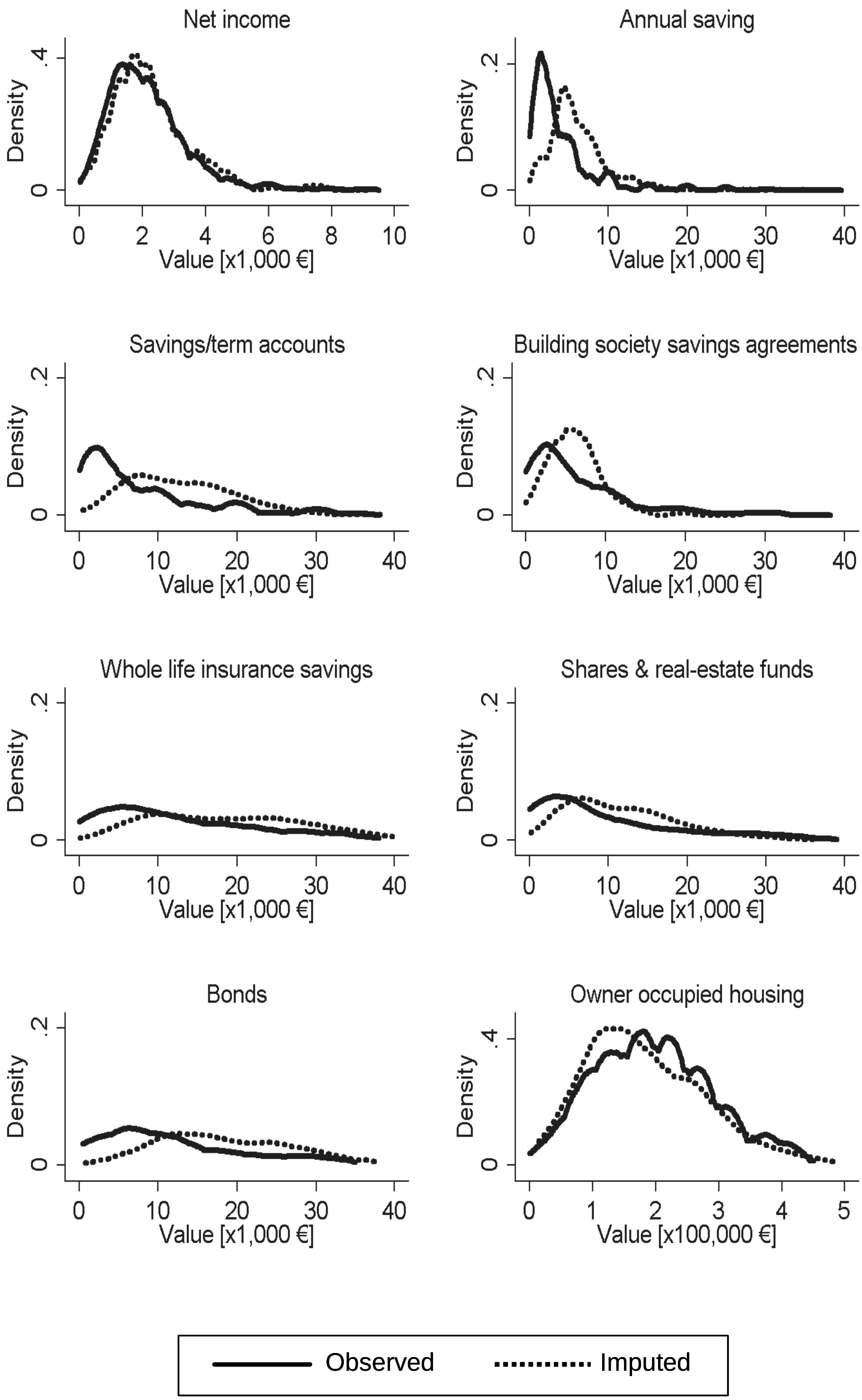
To further illustrate the effects of imputation, figure 3 presents kernel density estimates of observed and imputed values for the above mentioned financial variables. The kernel density is estimated for positive values of the variables that have been analyzed above, an Epanechnikov kernel and Silverman’s rule of thumb (Silverman, 1986) for bandwidth selection have been used. Kernel density estimates for the imputed data are usually obtained using Rubin's (1987) method to combine the data from the five implicates before the density estimation. According to Rubin (1987)[28], the overall imputed value $\overline{Y_i}$ of variable Y for a certain observation i is simply the average over the individual five imputed values, $m = 1,..., 5$, that is:

$$\overline{Y_i} = \frac{1}{5}\sum_{m=1}^{5} Y_{i,m}. \qquad (8)$$

In addition to the discussed findings concerning mean financial wealth differences between imputed and observed values, the figures illustrate nicely that the inclusion of covariates has a substantial effect on the distribution of asset holdings, a conclusion that is also emphasized by Hoynes et al. (1998). For the variables *annual saving* and *owner occupied housing*, the effect of focal point answers on the density is clearly visible: For example, the leftmost spike in the distribution of *annual saving* is due to the large amount of households reporting a total amount of annual saving of exactly 1,000 €. The second "spike" (or better: "plateau") stems from all households reporting 5,000 €. This multimodality is not replicated by the distribution of the imputed data, and it is debatable whether it should be replicated. One way of replicating multimodality would be to additionally use a nearest neighbor procedure after the regression-based imputation. For reasons given above, MIMS uses a nearest neighbor procedure only for variables that have follow-up brackets.

[28] Rubin (1987) derives general methods for combining the information from multiply imputed datasets. A brief summary of these methods, given in the appendix of this paper, section 6.2, informs the reader about how to work with the multiply imputed SAVE data.

Figure 3: Density functions of observed and imputed values.



Note: All calculations are unweighted.

5 Discussion and Conclusion

Except for controlled experimental settings, survey studies about human past and intended behavior rarely generate complete information. For several reasons that have been discussed in this paper, it is however desirable to provide users with a complete dataset in which all missing values have been imputed.

Missing values are rarely known with certainty. To be able to reflect the uncertainty of missing data in subsequent analyses, multiple imputation is used for the SAVE survey. This goal of this paper is to present the key theoretical underpinnings of a Markov Chain Monte Carlo multiple imputation algorithm, to describe and document the practical application of such a multiple imputation algorithm to the SAVE data, and to present and discuss properties of the algorithm as well as the resulting imputed datasets.

The Markov Chain Monte Carlo technique that is used for the algorithm presented in this paper is similar to the method presented in Schafer (1997) who uses smaller datasets with few conditioning variables, and it is similar to the method presented in Bover (2004) and in Kennickell (1998), who apply an iterative method to data from two large scale socio-economic surveys. It is important to note that modifications of this implementation are conceivable and should be explored: For example, the sequential simulation algorithm can be modified such that each draw from a certain conditional distribution depends not only on the conditional distribution estimated in the preceding iteration step, but also on conditional distributions estimated in earlier iteration steps (Cameron and Trivedi, 2005). Alternatively, in each iteration step the distribution of unobserved values can be simulated a certain number of times p, and the parameter values for the next iteration step can then be estimated from all p simulated distributions; this means that multiple versions of the unobserved data are generated from the predictive distribution in one iteration step. A comparison of convergence properties between these different ways of implementing the data augmentation algorithm would certainly be helpful. Considering the fact that the method proposed in this paper is based on the assumption of ignorable missing data, future research efforts should also be directed towards modeling the missing data mechanism explicitly and eventually a model should be formulated for each incomplete survey variable and for the corresponding mechanism of missingness. Particularly given the complexity of the nonresponse patterns in SAVE, this constitutes a substantial effort. A comparison with the results obtained from MIMS would be of highest scientific interest.

So far, convergence properties of MCMC methods have only been systematically analyzed on simulated datasets and datasets with fewer variables compared to the large household survey that is analyzed in this paper (see, e.g., Schafer, 1997). The findings of the present study suggest that the algorithm converges in only few iteration steps. For most variables, the process is stationary after not more than about 5-10 iterations steps. For all other variables, it is stationary from the first iteration step on, suggesting that the algorithm has already converged in the first iteration step – a phenomenon that is also reported by Kennickell (1998). It is certainly worth investigating the convergence properties of MCMC algorithms in the context of large surveys or large simulated datasets in a collaborative effort and with standardized methods. This will further contribute to a more comprehensive evaluation of the relevance of MCMC methods for survey research.

Finally, the comparison of the imputed implicates has revealed some insights that are of interest from a methodological point of view as well as for the practitioner. First, it has been shown that variable means differ across implicates; this reflects the uncertainty about the imputed values. A comparison between imputed and observed values has further revealed that the use of covariates in the imputation process has a substantial effect on the distributions of individual asset holdings. This finding suggests that item nonresponse is not occurring randomly but is related to the included covariates. The analyses have also indicated that there might be differences in the character of nonresponse across asset types. The results indicate interesting directions for future research on the relationship between socio-economic characteristics and nonresponse to specific items. Furthermore, the presented brief comparison with findings from other studies suggests that – from the point of view of imputation methodology – it would be of particular interest to apply different imputation methods, e.g. iterative and non-iterative methods, to the same dataset and to compare the resulting effects.

Eventually, the SAVE survey will be a longer panel survey. This offers additional possibilities for the imputation of each cross section. So far, the panel consists of only two waves. Of those households that are part of the random sample in 2003, 646 have participated again in 2005. From the point of view of imputation, the interesting observations are those that did not provide an answer to a certain question in one of the two waves but did answer in the respective other wave. As an example, I consider the monthly net income variable. The proportion of households that did not provide an answer to the income question in one of the two waves and for whom the relevant socio-demographic characteristics, for example employment and marital status, have not changed between 2003 and 2005, is 6%. I find that for two thirds of the households for which this condition applies, the imputed value in the wave in which the nonresponse had occurred lies in the +/-25%-band around the value given by the same household in the respective other wave. While a bit more than half of those households had given bracket information in one of the waves, the others did not provide any information about net income in one of the waves. This is informative in two respects: First, the fraction of imputed values that lie in this band can be used as one possible measure to evaluate the imputation procedure (although one should be cautious since it is based on only small numbers). It is found that – on average – a much lower fraction of imputed values would lie in this +/-25%-band if the procedure simply drew from the observed distribution of net income values, instead of conditioning the imputation on a large set of variables. Second, this finding shows that with respect to imputation one cannot learn too much from a panel that consists of only two waves, since the proportion of households who answer in only one of the two waves and for whom relevant socio-demographic characteristics have not changed between the two waves is rather low – an observation that is confirmed for all other core variables in SAVE. Exploiting the panel dimension for imputation makes more sense with a longer panel, since a longer panel increases the probability of having additional information on item nonrespondents in at least one of the waves. This information which is available in waves other than the one that is currently being imputed can then be used as further conditioning information for the hotdeck and regression-based methods. This is one important direction for the further development of MIMS.

6 Appendix

6.1 Derivation of the Variance of a Normally Distributed Random Variable that is Symmetrically Censored

Consider a normally distributed random variable y^* with mean zero and standard deviation σ:

$$y^* \sim N(0,\sigma) \tag{A1}$$

Alternatively, with $\varphi(\cdot)$ being the density function of the standard normal distribution, we can write:

$$y^* \sim \frac{1}{\sigma}\varphi\left({y^*}/{\sigma} \right) \tag{A2}$$

We now define a new random variable y, which is obtained from the original one, y^*, by symmetrically censoring the variable y:

$$y = \begin{cases} -a & if \quad y^* < -a \\ y^* & otherwise \\ a & if \quad y^* > a \end{cases} \tag{A3}$$

This variable has the following density function:

$$f(y) = \begin{cases} 0 & if \quad \left| y^* \right| > a \\ \Phi\left(-{a}/{\sigma}\right) = 1 - \Phi\left({a}/{\sigma}\right) & if \quad \left| y^* \right| = a \\ \frac{1}{\sigma}\varphi\left({y^*}/{\sigma} \right) & if \quad \left| y^* \right| < a \end{cases} \tag{A4}$$

Here, $\Phi(\cdot)$ is the cumulative distribution function of $\varphi(\cdot)$.

This distribution is a mixture of discrete and continuous parts. It is the variance of the random variable y that we want to calculate as a function of the censoring value a.

In order to do so, I use the variance decomposition formula:

$$Var(y) = E(Var[y \mid a]) + Var(E[y \mid a]) \tag{A5}$$

I compute the first term on the right-hand side, then the second term on the right-hand side, and then combine the two results.

(a) Computation of $E(Var[y \mid a])$*:*

The expected value of the conditional variance of *y*, given the censoring value *a*, can be decomposed as follows:

$$E(Var[y \mid a]) = 2\Phi\left(-a/\sigma\right) \cdot Var\,[y \mid y = a] + \left[1 - 2\Phi\left(-a/\sigma\right)\right] \cdot Var\left[y \mid |y| < a\right] \qquad \text{(A6)}$$

It is obvious that $Var[y \mid y = a] = 0$.

That is, $Var\left[y \mid |y| < a\right]$ remains to be computed, and it is known that $Var\left[y \mid |y| < a\right] = Var\left[\, y^* \mid |y^*| < a\right]$.

$Var\left[\, y^* \mid |y^*| < a\right]$ can be decomposed as follows:

$$\begin{aligned} Var\left[y^* \mid |y^*| < a\right] &= \int_{-a}^{a} {y^*}^2 \varphi(y^*)\,dy^* \\ &= \int_{-\infty}^{\infty} {y^*}^2 \varphi(y^*)\,dy^* - \int_{-\infty}^{-a} {y^*}^2 \varphi(y^*)\,dy^* - \int_{a}^{\infty} {y^*}^2 \varphi(y^*)\,dy^* \\ &= \sigma^2 - 2\int_{a}^{\infty} {y^*}^2 \varphi(y^*)\,dy^* \\ &= \sigma^2 - 2Var\left[y^* \mid y^* > a\right] \end{aligned} \qquad \text{(A7)}$$

$Var\left[y^* \mid y^* > a\right]$ is the variance of a truncated normally distributed variable. This variance is computed as follows (see Johnson and Kotz, 1970):

$$Var\left[y^* \mid y^* > a\right] = \sigma^2(1 - \delta(a)), \qquad \text{(A8)}$$

where

$\delta = \lambda(a)\left(\lambda(a) - a/\sigma\right)$, and $\lambda(a) = \dfrac{\varphi\left(a/\sigma\right)}{1 - \Phi\left(a/\sigma\right)}$.

It follows:

$$\begin{aligned} Var\left[y^* \mid |y^*| < a\right] &= \sigma^2 - 2\sigma^2(1 - \delta(a)) \\ &= \sigma^2(1 - 2 + 2\delta(a)) \\ &= \sigma^2(2\delta(a) - 1) \end{aligned} \qquad \text{(A9)}$$

And therefore:

$$E(Var[y \mid a]) = \left[1 - 2\Phi\left(-a/\sigma\right)\right] \cdot \sigma^2(2\delta(a) - 1) \qquad \text{(A10)}$$

(b) Computation of $Var(E[y \mid a])$*:*

We find:

$$Var(E[y \mid a]) = 2\Phi\left(-a/\sigma\right)\cdot\{a - E[y]\}^2 + \left[1 - 2\Phi\left(-a/\sigma\right)\right]\cdot\{E[y \mid |y| < a] - E(y)\}^2$$
$$= 2\Phi\left(-a/\sigma\right)\cdot a^2, \qquad \text{(A11)}$$

since $E[y] = a$, $E[y \mid |y| < a] = 0$, and $E(y) = 0$ by symmetry arguments.

Combining the results of *(a)* and *(b)* finally yields the expression for the variance of a symmetrically censored normally distributed variable, with mean zero, standard deviation σ and censoring value *a*:

$$Var(y) = 2\Phi\left(-a/\sigma\right)\cdot a^2 + \left[1 - 2\Phi\left(-a/\sigma\right)\right]\cdot\sigma^2\left(2\delta(a) - 1\right) \qquad \text{(A12)}$$

6.2 Rules for Inference Based on Multiply Imputed Datasets

The 5 implicates of the SAVE data can be analyzed using standard complete data methods. Every model has to be estimated 5 times, once for each complete and imputed dataset. The results across these estimations vary, this reflects the missing-data uncertainty. Rubin (1987) has derived a method for combining the results from a data analysis performed *M* times, once for each of *M* imputed data sets, to obtain a single set of results: Suppose that $\hat{Q}_m$ is the scalar point estimate of interest, obtained from data set *m*. Suppose further that $\hat{U}_m$ is the standard error associated with $\hat{Q}_m$ .The overall estimate is then the average of the individual estimates,

$$\overline{Q} = \frac{1}{M}\sum_{m=1}^{M}\hat{Q}_m. \qquad \text{(A13)}$$

For the overall standard error, one must first calculate the within-imputation variance,

$$\overline{U} = \frac{1}{M}\sum_{m=1}^{M}\hat{U}_m \qquad \text{(A14)}$$

and the between-imputation variance,

$$\overline{B} = \frac{1}{M-1}\sum_{m=1}^{M}(\hat{Q}_m - \overline{Q})^2. \qquad \text{(A15)}$$

The total estimated variance of the multiple-imputation point estimate is then

$$T = \overline{U} + (1 + \frac{1}{M})\overline{B}. \qquad \text{(A16)}$$

Single imputation underestimates the standard errors of the estimates because it has zero between imputation variance.

Additional methods for combining the results from multiply imputed data that hold under certain special assumptions about the data are presented in Schafer (1997).

7 References

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.

Bernaards, C. A., M. M. Farmer, K. Qi, G. S. Dulai, P. A. Ganz, and K. L. Kahn (2003): Comparison of Two Multiple Imputation Procedures in a Cancer Screening Survey. *Journal of Data Science,* 1 (3), 293-312.

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. and L. Essig (2005): Household saving in Germany: Results of the first SAVE study. In: D. A. Wise (Ed.), *Analyses in the Economics of Aging*, 317-352. Chicago: The University of Chicago Press.

Bover, O. (2004): The Spanish Survey of Household Finances (EFF): Description and Methods of the 2002 Wave. *Documentos Ocasionales N. 0409*. Banco de Espana.

Cameron, A. C. and P. K. Trivedi (2005): *Microeconometrics. Methods and Applications*. New York: Cambridge University Press.

Essig, L. and J. Winter (2003): Item Nonresponse to Financial Questions in Household Surveys: An Experimental Study of Interviewer and Mode Effects. *MEA-Discussion Paper 39-03*, MEA – Mannheim Research Institute for the Economics of Aging. University of Mannheim.

Ezzati-Rice, T. M., W. Johnson, M. Khare, R. J. A. Little, D. B. Rubin, and J. L. Schafer (1995): Multiple imputation of missing data in NHANES III. *Proceedings of the Annual Research Conference*, U.S. Bureau of the Census, 459-487.

Ferber, R. (1966): Item nonresponse in a consumer survey. *Public Opinion Quarterly*, 30 (3), 399-415.

Frick, J. R. and M. M. Grabka (2005): Item nonresponse on income questions in panel surveys: Incidence, imputation and the impact on inequality and mobility. *Allgemeines Statistisches Archiv*, 90 (1), 49-62.

Geman, S. and D. Geman (1984): Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6 (6), 721-741.

Graham, J. W. and J. L. Schafer (1999): On the performance of multiple imputation for multivariate data with small sample size. In: R. Hoyle (Ed.), *Statistical Strategies for Small Sample Research*, 1-29, Thousand Oaks, CA: Sage.

Groves, R. M., D. A. Dillman, J. L. Eltinge, and R. J. A. Little (2002): Survey nonresponse. New York: Wiley.

Hartley, H. O. and R. R. Hocking (1971): The analysis of incomplete data. *Biometrics*, 27, 783-808.

Hastings, W. K. (1970): Monte Carlo Sampling Methods Using Markov Chain and Their Applications. *Biometrika*, 57, 97–109.

Hersch, J. (2000): Gender, Income Levels, and the Demand for Cigarettes. *Journal of Risk and Uncertainty*, 21, 263-282.

Hersch, J. and W. K. Viscusi (1990): Cigarette Smoking, Seatbelt Use, and Differences in Wage-Risk Tradeoffs. *The Journal of Human Resources*, 25(2), 202-227.

Hoynes, H., M. Hurd, and H. Chand (1998): Household Wealth of the Elderly under Alternative Imputation Procedures. In: D. A. Wise (Ed.), *Inquiries in the Economics of Aging*, 229-257. Chicago: The University of Chicago Press.

Johnson, N. and S. Kotz (1970): Distributions in Statistics – Continuous Univariate Distributions. Vol. 2. New York: Wiley.

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.

Kalwij, A. and A. van Soest (2005): Item Non-Response and Alternative Imputation Procedures. In: A. Börsch-Supan and H. Jürges (Eds.), *The Survey of Health, Ageing and Retirement in Europe – Methodology,* 128-150*.* Mannheim: Mannheim Research Institute for the Economics of Aging.

Kennickell, A. B. (1997): Using range techniques with CAPI in the 1995 Survey of Consumer Finances. Board of Governors of the Federal Reserve System, Washington, D.C.

Kennickell, A. B. (1998): Multiple Imputation in the Survey of Consumer Finances. *Proceedings of the 1998 Joint Statistical Meetings*, Dallas TX.

Levine, P. B., T. A. Gustafson, and A. D. Velenchik (1997): More Bad News for Smokers? The Effects of Cigarette Smoking on Wages. *Industrial and Labor Relations Review*, 51, 493-509.

Li, K. (1988): Imputation Using Markov Chains. *Journal of Statistical Computing and Simulation*, 30, 57-79.

Li, K., T. Raghunathan, and D. Rubin (1991): Large sample significance levels from multiply-imputed data using moment-based statistics and an *F* reference distribution. *Journal of the American Statistical Association,* 86, 1065–1073.

Little, R. J. A. and D. B. Rubin (2002): Statistical Analysis with Missing Data. New York: Wiley.

Little, R. J.A. and T. Raghunathan (1997): Should Imputation of Missing Data Condition on All Observed Variables? *Proceedings of the Section on Survey Research Methods*, Joint Statistical Meetings, Anaheim, California.

Little, R. J. A., I. G. Sande, and F. Scheuren (1988): Missing-data adjustments in large surveys. *Journal of Business and Economic Statistics,* 6 (3), 117-131.

Manski, C. (2005): Partial Identification with Missing Data: Concepts and Findings. *International Journal of Approximate Reasoning,* 39 (2-3), 151-165.

Riphahn, R. and O. Serfling (2004): Item Non-response on Income and Wealth Questions. *Empirical Economics,* 30 (2), 521-538.

Rubin, D. B. (1987): Multiple Imputation for Nonresponse in Surveys. New York: Wiley.

Rubin, D. B. (1996): Multiple Imputation After 18+ Years. *Journal of the American Statistical Association*, 91 (434), 473-489.

Rubin, D. B. and N. Schenker (1986): Multiple Imputation for Interval Estimation from Simple Random Samples with Ignorable Nonresponse. *Journal of the American Statistical Association*, 81 (394), 366-374.

Schafer, J. L. (1997): Analysis of incomplete multivariate data. London: Chapman & Hall.

Schunk, D. (2006): Overview and documentation of the German SAVE survey. Mimeo. Mannheim Research Institute for the Economics of Aging. University of Mannheim.

Schräpler, J.-P. (2003): Gross income non-response in the German Socio-Economic Panel: Refusal or don't know? *Schmollers Jahrbuch*, 123, 109-124.

Schräpler, J.-P. and G. G. Wagner (2001): Das Verhalten von Interviewern - Darstellung und ausgewählte Analysen am Beispiel des "Interviewerpanels" des Sozio-Ökonomischen Panels. *Allgemeines Statistisches Archiv,* 85, 45-66.

Silverman, B. W. (1986): Density Estimation for Statistics and Data Analysis. London: Chapman and Hall.

Smith, J. P. (1995): Racial and Ethnic Differences in Wealth. *Journal of Human Resources,* 30, 158-183.

Tanner, M. A. and W. H. Wong (1987): The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, 82 (398), 528-550.

Lebenslauf

Persönliche Daten:	Daniel Schunk Mannheim Research Institute for the Economics of Aging (MEA) Universität Mannheim L 13, 17 68131 Mannheim Tel.: +49-(0) 621-181-3448 E-Mail: mail@daniel-schunk.de Geburtsdatum / -ort: 31.12.1976 / Koblenz
Aktuelle Tätigkeit:	Wissenschaftlicher Mitarbeiter Mannheim Research Institute for the Economics of Aging (MEA) Universität Mannheim

Ausbildung:

02 – 05 / 2005	Studium/Forschung am Interdisciplinary Center for Economic Science, George Mason University, VA, USA
Seit 10 / 2002	Promotion Universität Mannheim, Fakultät für Volkswirtschaftslehre und Center for Doctoral Studies in Economics and Management (CDSEM) Professor Axel Börsch-Supan, Ph.D.
09 / 2002	Abschluß als Diplom-Systemwissenschaftler, Fachbereich Mathematik und Informatik, Universität Osnabrück.
2000 – 2001	Studium „Systems Science“, University of Illinois at Urbana-Champaign, IL, USA
1997 – 2002	Studium der „Angewandten Systemwissenschaft“, Universität Osnabrück
1987 – 1996	Bischöfliches Cusanus-Gymnasium, Koblenz, Abschluß: Abitur

Berufserfahrung:

01 – 03 / 2002	Praktikum, McKinsey&Company Inc., Köln
05 – 08 / 2001	Praktikum, AEC/Centrec-Consulting, Savoy, IL, USA
1999 – 2002	Wissenschaftliche Hilfskraft, Institut für Umweltsystemforschung, Universität Osnabrück
1996 – 2000	Leitung und Organisation von Weinproben, Weingut Anton Hammes, Alken/Mosel
1996 –1997	Wehrdienst, Ernst-Rodenwaldt-Institut, Koblenz, Abteilung „Ergonomie“

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die Dissertation selbständig angefertigt und mich anderer als der in ihr angegebenen 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, den 10. April 2006