Essays in Behavioral Public Economics

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors

der Wirtschafts- und Gesellschaftswissenschaften

durch die

Rechts- und Staatswissenschaftliche Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

vorgelegt von

Felix Marklein

aus Rheinbach.

Bonn, 2009

Dekan:	Prof. Dr. Christian Hillgruber
Erstreferent:	Prof. Dr. Armin Falk
Zweitreferent:	Prof. Paul Heidhues, PhD

Tag der mündlichen Prüfung: 17.02.2009

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn unter http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Acknowledgments

This dissertation benefitted tremendously from the interaction with other researchers. First and foremost, I am grateful to my supervisor Armin Falk for encouraging me in my work and for teaching me the art of doing economic research. Moreover, I am grateful to my second advisor Paul Heidhues for the discussions about the central ideas contained in this thesis. My thanks also go to Urs Schweizer and Georg Nöldeke for all the energy, effort, and time they have invested in the Bonn Graduate School of Economics (BGSE). This Ph.D. program is a great institution and my work has been influenced heavily by the interaction with teachers and fellow students in the BGSE. I am indebted to Philipp Kircher, Wolfram Merzyn, Jörg Oechssler, Susanne Ohlendorf, and Burkhard Schipper for many insightful discussions.

Another important factor in the writing of this thesis has been the fruitful research environment at the Institute for the Study of Labor (IZA). In particular, I want to thank my co-authors Johannes Abeler, Thomas Dohmen, David Huffman, and Uwe Sunde for making work so much fun. Furthermore, Steffen Altmann and Matthias Wibral were great colleagues and gave me valuable comments on my papers. Last but not least, the visitors' program at IZA enabled me to talk to many excellent economists. Among them, Dan Hamermesh and Bentley MacLeod had a decisive impact on my research.

Moreover, I am grateful to Benny Moldovanu for enabling me to spend the second year of my graduate education abroad, at New York University (NYU). It has been a very intense experience to learn graduate economics and to conduct research at a top U.S. university, and I have benefitted a lot from it. At NYU, I want to thank in particular Ennio Stacchetti and Andrew Schotter. Their teaching gave me many new insights and they always had an open door when I wanted to discuss research ideas with them.

Several students I met in the BGSE and at NYU also had a direct or indirect impact on this thesis. I thank Moran Amir, Sophie Bade, Almut Balleer, Eugen Kovac, Sandra Ludwig, André Meier, Philippe Raab, Ariell Reshef, Julian Rode, and Julia Schwenkenberg for exciting discussions, not only about economics.

For financial support I want to thank the Deutsche Forschungsgemeinschaft (GK 629, "Quantitative Ökonomie"), the Studienstiftung des Deutschen Volkes (ERP-Stipendium), the BGSE, the IZA, and the Sonderforschungsbereich TR 15, "Governance und die Effizienz ökonomischer Systeme".

Finally, I owe a great thank you to my parents who always supported me in all my plans.

Contents

Acknowledgments i

List of Figures v

List of Tables vii

Introduction 1

Chapter 1. Economic Consequences of Biased Probability Judgment: Representative Evidence 5

1.1. Introduction 5

1.2. Data 10

1.3. Pervasiveness and Determinants of Biased Probability Judgment 12

1.4. Economic Consequences of Biased Probability Judgment 18

1.5. Conclusion 24

Chapter 2. The Non-Use of Bayes' Rule: Representative Evidence on Bounded Rationality 27

2.1. Introduction 27

2.2. Data 33

2.3. Results 33

2.4. Conclusion 44

Chapter 3. Fungibility, Labels, and Consumption: Experimental Evidence 47

3.1. Introduction 47

3.2. Experimental Design 50

3.3. Results 54

3.4. Vignette Survey 64

3.5. Discussion and Conclusion 68

Chapter 4. Who Reacts to Monetary Incentives? The Role of Extrinsic and Intrinsic Motivation among Volunteer Workers 71

4.1. Introduction 71
4.2. Theoretical Background 75
4.3. Survey Design 76
4.4. Results 78
4.5. Conclusion 88

Appendix A. Appendix to Chapter 1 91

Appendix B. Appendix to Chapter 2 107

Appendix C. Appendix to Chapter 3 113
C.1. Instructions for Laboratory Experiment 113
On-screen Instructions 118
Moral Obligation Scenarios 120
C.2. Wording of Vignette Survey 120

Appendix D. Appendix to Chapter 4 123

Bibliography 131

List of Figures

1.1 Measures of Cognitive Ability. 12

1.2 Probability Judgment. 14

2.1 Answers to Conditional Probability question. 34

3.1 Consumption Decision with Non-distortionary In-kind Benefit. 51

3.2 Consumption of the Subsidized Good in the Reference Stage. 55

3.3 Consumption of the Subsidized Good in the Subsidy Stage. 56

3.4 Consumption Increase of the Subsidized Good from Reference Stage to Subsidy Stage. 56

3.5 Cumulative Distribution of the Consumption of the Subsidized Good according to Mathematical Ability. 59

3.6 Cumulative Distribution of the Consumption of the Subsidized Good according to Feelings of Moral Obligation. 61

4.1 Proxies for Extrinsic and Intrinsic Motivation 79

4.2 Labor Supply in Response to Monetary Payments. 81

List of Tables

1.1 Sample Statistics. 11

1.2 Probability Judgment - Descriptive Statistics. 15

1.3 Probability Judgment - Determinants (I). 16

1.4 Probability Judgment - Determinants (II). 17

1.5 Economic Outcomes - Unemployment. 21

1.6 Economic Outcomes - Financial Situation. 23

2.1 Sample Statistics. 33

2.2 Bayesian Solution vs. Typical Judgment Biases. 35

2.3 Descriptive Statistics. 36

2.4 Determinants of Base Rate Neglect. 37

2.5 The Role of Academic Background. 38

2.6 Robustness Check - Multinomial Analysis (I). 40

2.7 Robustness Check - Multinomial Analysis (II). 41

2.8 Robustness Check - Linear Distance. 42

3.1 Payoff Functions Used in the Laboratory Experiment. 53

3.2 Consumption in the Laboratory Experiment. 58

3.3 Distance from Optimum in the Laboratory Experiment. 63

3.4 GRE Test Scores (Quantitative Section). 65

3.5 Consumption in the Vignette Survey. 67

4.1 Sample Statistics. 77

4.2 Labor Supply Schedule. 80

4.3 Labor Supply and Extrinsic vs. Intrinsic Motivation. 82

4.4 Crowding-in and Extrinsic vs. Intrinsic Motivation. 84

4.5 Labor Supply and Extrinsic vs. Intrinsic Motivation (Within-person). 86

4.6 Crowding-In (Within-person). 87

B.1 Probability Judgment - Detailed Descriptive Statistics. 108

B.2 Detailed Multinomial Analysis (I). 109

B.3 Detailed Multinomial Analysis (II). 110

B.4 Detailed Multinomial Analysis (III). 111

B.5 Detailed Multinomial Analysis (IV) 112

Introduction

Traditionally, an important tool in public policy analysis has been the principle of revealed preferences. Under the assumption that people choose optimally, the policy maker can infer what people want by observing their choices. Recent developments in behavioral economics have questioned this approach. In particular, if people are boundedly rational, policy makers can no longer rely on observed choices alone. Rather, the effectiveness of policy measures could be improved by taking into account that a substantial part of the population exhibits non-standard decision making patterns (Bernheim and Rangel 2007, Bertrand, Mullainathan, and Shafir 2004). Therefore, a new field of behavioral public economics has set out to identify domains in which incorporating insights from behavioral economics can improve policy making.

This dissertation aims at the identification of systematic behavioral biases that are relevant in the context of economic policy. In total, the dissertation consists of four chapters. In each chapter, we empirically analyze in a specific environment how individuals make economic decisions. We find in all chapters that, on average, human behavior is neither perfectly rational, nor completely random. On the one hand, there is always a relatively large share of people behaving in line with standard economic theory. On the other hand, we find in all our settings that a substantial share of people deviates from the predictions of standard theory. Most importantly, we find that deviations from the standard model are systematic. In all four chapters, a large share of people acts in accordance with behavioral theories that are well-known from the literature on bounded rationality and from cognitive psychology. Our empirical approach allows us to categorize the different types of behavior in the data. Taken together, the chapters in this dissertation provide an insight into behavioral patterns of decision-making in a number of economic domains. The findings often have straightforward implications for the design of public policy, as will be discussed in the respective chapters.

Chapter 1 and Chapter 2 share the same methodology: in both chapters we analyze a data set with roughly 1,000 observations that is representative for the adult population

living in Germany.[1] This approach allows us to elicit measures of the pervasiveness of certain cognitive biases in the population. As many economic decisions involve a substantial amount of uncertainty, it is crucial for economists to understand how people process probabilistic information. In Chapter 1, we address people's ability to form simple probability judgments. Our results show that almost a third of the respondents exhibits a cognitive bias. In particular, the biases we find are consistent either with the gambler's fallacy or with the hot hand fallacy. We find that in the setting under investigation the gambler's fallacy is more than twice as frequent as the hot hand fallacy. In a second step, we relate the observed biases to economic outcomes in two domains where they are likely to play an important role: job search and financial decision making. Our results show that the observed deviations from the standard model have a negative impact on individual economic outcomes in both domains. In the labor market, we find that people who exhibit the hot hand fallacy are significantly more likely to be long-term unemployed. In contrast, in financial decision making, people who are prone to the gambler's fallacy have a significantly higher probability to overdraw their bank account. These findings suggest that it is not biased probability judgment *per se* that determines economic outcomes. Rather, the specific form of a person's bias plays a decisive role, depending on the context in which economic decisions are made. Our findings have straightforward policy implications for the labor market and for debt counseling services.

Chapter 2 investigates the prevalence and determinants of probabilistic judgment biases in settings where subjects face conditional probabilities. In economic theory, the benchmark solution process for these environments is Bayes' rule. However, research in cognitive psychology has shown that people often have difficulties to solve conditional probability problems. In particular, biases such as base rate neglect have been widely documented in the literature (Tversky and Kahneman 1982). Our data set allows us to shed more light on the actual prevalence of base rate neglect in the general population. While we find that base rate neglect is rather prominent in the data, the share of people exhibiting this bias is by far lower than in previous studies with non-representative subject pools. Moreover, our data allows us to analyze the determinants of base rate neglect. Surprisingly, we find that both higher education and higher cognitive ability *increase* the probability that a person exhibits base rate neglect. This finding can also serve as an explanation for why our results regarding the prevalence of base rate neglect differ sharply

[1] These two chapters are based on joint work with Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde.

from non-representative subject pools: previous studies have been conducted mostly with university students and with medical professionals, i.e., with highly educated individuals.

Chapter 3 addresses the question whether people take the central principle of fungibility into account when making consumption decisions.[2] Recall that fungibility is an economic notion that implies that any unit of money is substitutable for another. If fungibility holds, the most important determinant for consumption should be a consumer's total wealth, whereas the specific composition of wealth should be irrelevant. We find in an incentivized laboratory experiment that many subjects do not treat money as fungible. Subjects are influenced by a label attached to a part of their budget and they change consumption in line with the suggestion of the label. Moreover, subjects with lower mathematical ability are more likely to ignore fungibility. A vignette survey confirms the external validity of our results. The findings lend support to behavioral models such as narrow bracketing and mental accounting. One implication of our results is that in-kind benefits distort consumption more than usually assumed.

In Chapter 4, we investigate the impact of monetary incentives on public-spirited behavior.[3] We analyze a survey in which volunteer workers are offered financial payments for their work and we elicit their labor supply reaction to these payments. The unique structure of the data set allows us to take into account the degree to which workers are motivated by extrinsic and intrinsic factors. We find that the impact of financial payments on labor supply depends crucially on the behavioral motives of the volunteers. Whereas high degrees of intrinsic motivation correspond to a low responsiveness to financial incentives, high degrees of extrinsic motivation are related to a significant increase in volunteers' labor supply when monetary rewards are introduced. Moreover, we find that the prevalence of crowding-out is rather low. Thus, offering financial incentives leads to a disproportionate crowding-in of extrinsically motivated workers.

[2]This chapter is based on joint work with Johannes Abeler and has been circulated as IZA discussion paper, see Abeler and Marklein (2008).

[3]This chapter is based on joint work with Armin Falk, Annette Meier, and Martin Schellhorn.

CHAPTER 1

Economic Consequences of Biased Probability Judgment: Representative Evidence

1.1. Introduction

Economic decisions often involve a significant amount of uncertainty. A thorough understanding of economic decision making is therefore only possible by investigating how people actually form probability judgments. Economic theory typically assumes decisions to be consistent with the basic laws of probability theory. In contrast, a large number of experimental studies has shown that humans frequently make mistakes in processing probabilistic information.[1] Despite the vast experimental evidence, three important questions have received little attention so far: first, do the cognitive biases we observe in laboratory subject pools also exist in the general population? Second, what are the determinants of biased probability judgment? And third, which impact do probability judgment biases have on economic decision making?

In this chapter we take a novel approach to the analysis of decision making under uncertainty by directly measuring the capability for probability judgment in a representative sample of more than 1,000 individuals. This procedure allows us to evaluate the pervasiveness of cognitive biases in the general population. In a second step, we study the determinants of biased probability judgment, with a particular focus on respondents' education level and cognitive ability. Finally, we evaluate the impact that biased probability judgment has on individual economic outcomes.

Respondents' abilities in probabilistic reasoning are elicited via a survey question that addresses a fundamental bias in probability judgment: the so-called gambler's fallacy. People who exhibit this bias expect random processes to have self-correcting properties, which is in direct contradiction to the principle of independence between random outcomes. The gambler's fallacy has been widely discussed in the literature, starting with a famous essay by Laplace (1820). In the chapter "Concerning Illusions in the Estimation of Probabilities", Laplace states:

[1] See for example Tversky and Kahneman (1971), Grether (1980), Charness and Levin (2005). A comprehensive overview is given by Conlisk (1996).

> *When a number in the lottery of France has not been drawn for a long time, the crowd is eager to cover it with stakes. They judge since the number has not been drawn for a long time that it ought at the next drawing to be drawn in preference to others. So common an error appears to me to rest upon an illusion by which one is carried back involuntarily to the origin of events. It is, for example, very improbable that at the play of heads and tails one will throw heads ten times in succession. This improbability which strikes us indeed when it has happened nine times, leads us to believe that at the tenth throw tails will be thrown.*

We designed our probability task such that respondents were confronted with the following series of eight tosses of a fair coin: tails - tails - tails - heads - tails - heads - heads - heads. Respondents then had to indicate the probability with which tails occurs in the next toss of the coin. The correct answer is 50%, as the coin tosses are independent of each other. Since the sequence ends with a streak of three heads, the gambler's fallacy would lead respondents to predict that tails comes up with a probability of more than 50%. The opposite bias, the so-called hot hand fallacy, would lead respondents to indicate a probability of less than 50%. The term hot hand derives from basketball, where players who make a shot are often believed to be more likely to hit the next shot, in contrast to players who miss a shot (see for example Gilovich, Vallone, and Tversky 1985). In our setting, the hot hand fallacy implies the belief that the streak of heads at the end of the sequence is likely to be continued.

The probability task was administered to a sample of more than 1,000 individuals, representative of the German population. Results show that 60.4% of the respondents give the correct answer. Thus, a majority of the respondents is aware of the independence between random outcomes. Among the incorrect answers, the gambler's fallacy is by far the most frequent bias: 21.1% of the respondents overestimate the probability for tails. In contrast, 8.8% answer in line with the hot hand fallacy, as they underestimate this probability. The answer "I don't know" is given by 9.6% of the sample. Our findings have two important implications: first, a substantial share of the population seems to lack basic knowledge about stochastic processes. Second, the biases we observe are systematic, as deviations from the normative solution are not distributed randomly. Rather, among people who make a mistake, the gambler's fallacy is the predominant bias.

In order to address the determinants of biased probability judgment, the survey elicited a number of background variables such as education, age, and gender. Our empirical analysis shows that more educated people are much more likely to answer correctly. Moreover, we find a gender effect, with women being less likely to give the correct answer. A unique feature of the data set is that it includes two measures of respondents' cognitive ability

that are very similar to measures from a widely used IQ-test. One measure is used as a proxy for general knowledge (also known as crystallized intelligence), whereas the other is a proxy for the speed of cognitive perception (also known as fluid intelligence).[2] These measures relate to respondents' answers in plausible ways: whereas the measure for general knowledge is positively correlated with answers to the probability question, the measure for perceptual speed has no explanatory power.[3] This finding is consistent with the view that a correct perception of the independence between random outcomes relies mainly on acquired knowledge, not on respondents' mechanical cognitive functions.

As uncertainty is a crucial factor in many economic decisions, one would expect that biased probability judgment has a detrimental effect on individual economic outcomes. The gambler's fallacy and the hot hand fallacy are likely to affect economic decision making in domains where people base their decisions on a sequence of realizations of a random process. In particular, these two biases should matter in situations where a streak of similar outcomes has occurred prior to the decision. Our data contain information about behavior in two domains where the specific form of judgment biases studied in this chapter is highly relevant: job search decisions by an unemployed person and consumption decisions by a cash-constrained consumer.

In the domain of job search, assume that a job-seeker who has received feedback on a number of applications has to decide whether to continue looking for a job. In this context, the employers' reactions can be interpreted as realizations of a random process: a job offer is a positive realization, whereas a rejection is a negative realization. As a thought experiment, consider the specific situation in which the job-seeker has received several rejections in a row. The hot hand fallacy then predicts that he is going to believe that this streak is likely to continue. As a result, he might stop searching for a job altogether and stay unemployed for a long time. Put differently, the hot hand fallacy can lead to persistent unemployment by exerting a negative influence on the job-seeker's perception of his job finding probability. In contrast, a job-seeker who is prone to the gambler's fallacy will continue his job search, out of the belief that a streak of rejections has made it *more* likely to receive a job offer in the future. As a consequence, persistence of unemployment should be less severe among job-seekers who are prone to the gambler's fallacy.

[2]See also the seminal contribution by Cattell (1963).

[3]In a related paper, Oechssler, Roider, and Schmitz (2008) find in a web-based experiment among university students that higher cognitive ability is associated with lower incidences of biased judgment.

In the empirical analysis, we find indeed that being prone to the hot hand fallacy is significantly related to long-term unemployment. Probit estimates indicate that exhibiting the hot hand fallacy is associated with a 6.1%-point increase in the probability of being long-term unemployed, controlling for background characteristics such as age, gender, years of schooling, and household wealth. This effect is relatively large compared to the share of long-term unemployed persons in the total sample of 6.8%. Additional support for our predictions comes from the finding that the gambler's fallacy, in turn, does not lead to a significantly higher risk of being long-term unemployed.

Another domain where a sequence of previously observed realizations might influence economic behavior is in the domain of consumption. As a thought experiment, assume that a cash-constrained consumer has to decide whether to overdraw his bank account in order to make a large purchase. In this context, the relevant series of random outcomes observed by the decision maker consists of personal income shocks that can be either negative or positive. Consider a situation in which the consumer experienced several negative income shocks in a row. Being prone to the gambler's fallacy predicts that he will be inclined to overdraw his bank account for short-term consumption, as he believes that a positive income shock is "due" in the near future. In contrast, the hot hand fallacy implies that the consumer will refrain from consumption and the bank account will not be overdrawn.

The regression results show that financial behavior is indeed related to people's ability to form probability judgments. We confirm our hypothesis that people who exhibit the gambler's fallacy are significantly more likely to overdraw their bank account. The effects are substantial: according to our probit estimates, being prone to the gambler's fallacy increases a person's probability for having an overdrawn bank account by 8.8%-points, while the share of people with an overdrawn bank account in the total sample is 16.6%. Our predictions are further supported by the finding that being prone to the hot hand fallacy does not increase the risk of having an overdrawn bank account.

Of course, one has to be very careful in interpreting the empirical findings of our analysis. Economic decisions in the domains of job search and consumption are highly complex and are influenced by many different factors. For instance, the institutions in which a decision takes place as well as a job-seeker's specific skills and preferences might have a large influence on success in the labor market. Still, on top of these factors, our findings allow us to speculate about a direct link between people's perception of probabilities and their actual economic decisions. The fact that the hot hand fallacy is important in job search decisions, whereas the gambler's fallacy is important in financial decisions suggests that

it is not biased probability judgment *per se* that determines economic outcomes. Rather, the specific form of a person's bias plays a decisive role, depending on the context in which economic decisions are made.

Our study offers an important methodological contribution, as the representative design of the sample allows us to draw conclusions about cognitive biases in the general population.[4] Several studies have argued that the effectiveness of policy measures could be improved by taking into account that a substantial part of the population exhibits non-standard decision making patterns (Bernheim and Rangel 2007, Bertrand, Mullainathan, and Shafir 2004). As a first step in this direction, the methodology employed in this chapter can inform policy makers about the actual prevalence of probability judgment biases and about the way in which these biases affect economic outcomes. From an applied perspective, our findings have straightforward implications for the design of policy measures on the labor market and in the domain of household debt counseling. We derive simple measures of debiasing that are in line with "libertarian paternalism", an approach that helps boundedly rational agents to make better decisions, without limiting the freedom of agents who decide optimally in the first place (Camerer, Issacharoff, Loewenstein, O'Donoghue, and Rabin 2003, Thaler and Sunstein 2003).

In the related literature, Rabin (2002) has shown from a theoretical perspective that the gambler's fallacy can be interpreted as people's tendency to exaggerate the degree to which a small sample reflects the properties of the underlying data generating process. In the empirical literature, a number of studies has used field data to investigate probability judgment biases. Clotfelter and Cook (1993) demonstrate that lottery players act in line with the gambler's fallacy: evidence from the Maryland state lottery shows that in the days after a winning number has been drawn, betting on this particular number drops significantly. Croson and Sundali (2005) investigate the betting behavior of roulette players in a casino in Reno, Nevada. They find that a long streak of the same outcome leads players to bet disproportionately on the opposite outcome. E.g., a streak of 5 times red in a row leads to significantly more bets on black. Terrell (1994) investigates field data from horse races and finds that betting behavior is consistent with the gambler's fallacy. Note that all these studies rely on aggregate data, as for instance the total number of bets placed on a particular number. In contrast, an important feature of our analysis is that we elicit individual data about respondents' perception of probabilities. Moreover,

[4]For instance, Hogarth (2005) argues in favor of a more representative design of empirical research in economics and a more careful assessment of the circumstances under which evidence from experiments can be generalized to the population at large.

our data set contains information about respondents' educational background, cognitive ability, and individual economic outcomes. Closely related to our analysis is a study by Stango and Zinman (2007) who show that cognitive biases can have an impact on financial decision making. The authors use U.S. data to investigate the so called *payment/interest bias*: consumers systematically underestimate the interest rate associated with a given loan principal and a repayment stream. Their empirical analysis shows that, for loans from non-bank finance companies, biased consumers do indeed hold loans with significantly higher interest rates.

The remainder of the chapter is structured as follows: Section 1.2 contains a description of the data. Section 1.3 presents evidence on the pervasiveness of cognitive biases in the population and addresses the determinants of these biases. The impact of biased probability judgment on economic outcomes is explored in Section 1.4, and Section 1.5 concludes.

1.2. Data

The data set under investigation consists of 1,012 observations and is a representative sample of the population living in Germany aged 16 years and older. The data were collected by the professional interview group TNS Infratest in June and July 2005. Households were contacted by interviewers according to the Random Route Method (see Fowler 2002) and one person per household was surveyed.[5] All interviewers used the CAPI procedure (Computer Assisted Personal Interview), administering questions and collecting answers with the help of a notebook computer. To elicit respondents' ability for probability judgment, we used the following question (translated from German):[6]

> *Imagine you are tossing a fair coin. After eight tosses you observe the following result: tails - tails - tails - heads - tails - heads - heads - heads. What is the probability, in percent, that the next toss is "tails"?*

We chose the sequence of outcomes such that the overall occurrence of tails in the sample is indeed 50% (4 out of 8), to avoid raising doubts among the respondents about the coin

[5]See Appendix A for a full overview of the original survey (in German).

[6]The exact wording was: "Nehmen Sie an, Sie werfen eine Münze, die gleichmäßig auf die eine oder die andere Seite fällt. Nach acht Würfen beobachten Sie folgendes Ergebnis: Zahl - Zahl - Zahl - Kopf - Zahl - Kopf - Kopf - Kopf. Wie hoch ist die Wahrscheinlichkeit, ausgedrückt in Prozent, dass der nächste Wurf 'Zahl' ist?".

Table 1.1: Sample Statistics

Variable	average	std. dev.	num. obs.
1 if female	0.537	(0.499)	1,012
Age	47.555	(18.371)	1,011
Years of schooling	10.328	(1.829)	965
Perceptual speed test (raw data)	28.339	(9.853)	902
Perceptual speed test (standardized)	0	(1)	902
Word fluency test (raw data)	22.892	(10.933)	849
Word fluency test (standardized)	0	(1)	849
Log net household income (per month)	7.521	(0.635)	940
Log net household wealth	7.177	(5.421)	772
1 if unemployed	0.114	(0.318)	1012
1 if long-term unemployed	0.068	(0.252)	1012
Account overdrawn	1.504	(0.885)	943
1 if account overdrawn at time of interview	0.166	(0.372)	1012

Table 1.1: Sample Statistics.

being fair. Respondents had to answer with a number between 0% and 100%. Alternatively, they had the possibility to answer "I don't know". The correct answer is 50%, as the coin is fair and the tosses are independent of each other.

In order to address the determinants of biased probability judgment, the data contain a number of socioeconomic background variables such as education, age, gender, income, and wealth. A novel feature of the survey is that it elicited measures of respondents' cognitive ability. The design of these measures is based on the two-component theory of cognitive ability, originating from research in developmental psychology (Lang 2005). According to this theory, cognitive ability can be broadly divided into cognitive mechanics and cognitive pragmatics (Baltes, Lindenberger, and Staudinger 2006, Lindenberger and Baltes 1997). The mechanics of cognition (fluid intelligence) reflect fundamental organizational properties of the central nervous system (Singer 1995). In contrast, the cognitive pragmatics (crystallized intelligence) reflect the knowledge-based domain of cognitive ability. Examples for the mechanics of cognition are the speed, the accuracy, and the coordination of cognitive processing operations. Examples for the pragmatics are reading and writing skills, educational qualifications, and professional skills.

Respondents' performance in the domain of cognitive mechanics was assessed via a digit-symbol test that has been designed to measure perceptual speed. For this test, respondents had to match the correct digit to symbols on the computer screen. They had to match as many digit-symbol pairs as possible within a time frame of 90 seconds. The

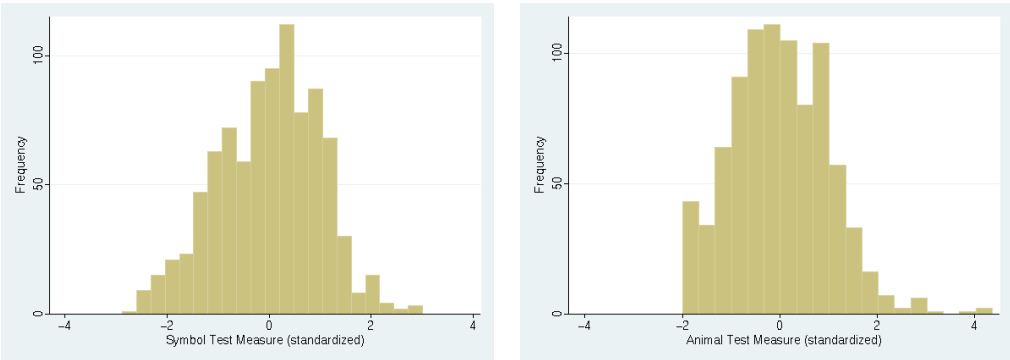


Figure 1.1: Measures of Cognitive Ability. Left panel: Perceptual speed test, measuring fluid intelligence. Respondents had to match as many digits and symbols as possible in 90 seconds of time. Right panel: Word fluency test, measuring crystallized intelligence. Respondents had to name as many distinct animals as possible in 90 seconds of time. Both measures are standardized.

CAPI method allowed us to measure performance directly, registering decisions through a software that was running in the background. In the area of cognitive pragmatics, a word fluency test was used to elicit a measure of respondents' general knowledge. The test asked participants to name as many distinct animals as possible in 90 seconds of time.

Regarding the impact of biased probability judgment on economic behavior, our research question leads us to focus on two specific domains: respondents' employment status and respondents' financial situation. With respect to employment status, the data contain information on whether respondents are registered as unemployed. Moreover, respondents have to indicate whether they are long-term unemployed, i.e., unemployed for 12 months and longer. To address respondents' financial situation, they are asked about the number of days (per year) their bank account is overdrawn, and whether their account is currently overdrawn at the time of the interview.

Sample statistics are shown in Table 1.1. About 53.7% of the sample are female, average age is 47.6 years, and respondents have on average 10.3 years of schooling. The measures for cognitive ability have a mean value of 28.3 (perceptual speed test) and 22.9 (word fluency test). For the empirical analysis, both cognitive ability measures are standardized to have variance one and mean zero. Histograms of the two standardized measures are presented in Figure 1.1. Note that the shape of both distributions resembles a normal distribution. With regard to the economic outcome variables, 11.4% of the respondents are unemployed, and 16.6% indicate that their bank account is overdrawn at the time of the interview.

1.3. Pervasiveness and Determinants of Biased Probability Judgment

In this section we document the pervasiveness of cognitive biases in a representative sample of the German population. We then proceed to address the determinants of biased probability judgment. Responses to the probability question are shown in the histogram in Figure 1.2. We find that 60.4% of the survey participants gave the correct answer of 50%. Thus, a majority of the respondents knows that the toss of a coin does not depend on outcomes of previous tosses. The remaining 39.6% seem to lack a basic understanding of probability theory. As can be seen from the histogram, the incorrect answers are spread out over the full range of the answer space, from 1% to 100%. The average estimate for tails to come up in the next coin toss is a probability of 54.2%. Recall that, as the sequence in our setting ends with a streak of three heads, the gambler's fallacy leads respondents to an estimate of more than 50%. In contrast, the hot hand fallacy implies an estimate of less than 50%. We find that the gambler's fallacy is exhibited by 21.1% of the respondents, whereas 8.8% of the respondents are prone to the hot hand fallacy. The answer "I don't know" is given by 9.6% of the respondents.

The fact that the gambler's fallacy is by far more pervasive than the hot hand fallacy indicates that respondents display *systematic* biases. This relates our results to an important debate in economics and in cognitive psychology (see Conlisk 1996, Stanovich and West 2000). According to one side of this debate, observing non-normative answers does not prove that people are boundedly rational. Theoretically, mistakes could be observed even if respondents were fully rational, e.g., due to lack of concentration or due to lack of motivation. However, one would expect such mistakes to be random noise, without a systematic pattern. If the mistakes in our sample were indeed random noise, the share of people who exhibit the gambler's fallacy and the share of people who exhibit the hot hand fallacy should be of approximately equal size. This is clearly not the case, as the gambler's fallacy is more than twice as frequent as the hot hand fallacy. Moreover, a Shapiro-Wilkinson Test rejects the null hypothesis of a Gaussian distribution of the deviations from the normatively correct answer at any conventional significance level ($p < 0.001$). Therefore, our findings indicate that the pattern of non-normative answers is systematic.

An explanation for the fact that the gambler's fallacy is the dominant bias in our setting is suggested in the work by Ayton and Fischer (2004). In their study, subjects were presented with sequences of binary outcomes that had either a high rate of alternations, or a high rate of streaks. Subjects then had to guess whether a given sequence was derived

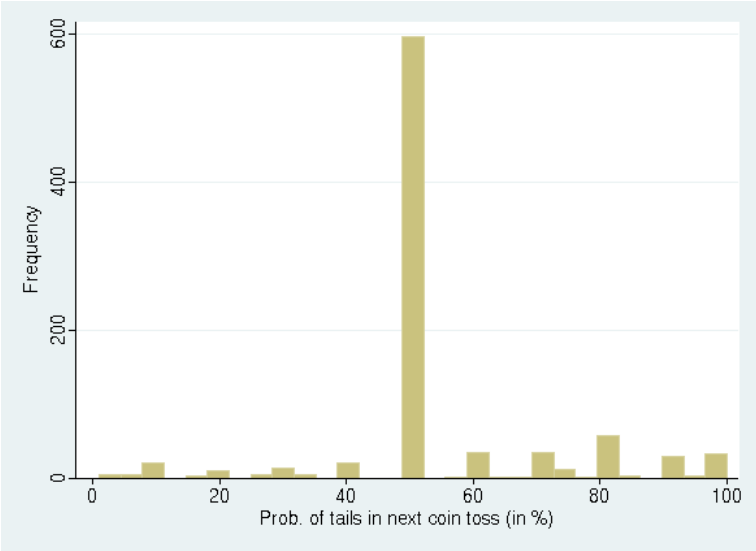


Figure 1.2: Probability Judgment. The normatively correct solution is a probability of tails of 50%. The 95 observations for the answer "I don't know" are not shown.

from human performance (e.g., hits and misses of a professional basketball player during a game), or from an inanimate chance process (e.g., heads and tails in the successive tosses of a fair coin). Their study demonstrates that subjects were more likely to attribute sequences with many streaks to human skilled performance. In contrast, sequences with high rates of alternation were attributed to inanimate chance processes. Our results are complementary to these findings, as we show that people who make predictions regarding the outcome of an inanimate chance process tend to overestimate the occurrence of alternations, whereas the belief in streaks is relatively infrequent.

Regarding the determinants of biased probability judgment, several basic insights become already apparent from the descriptive statistics. Table 1.2 provides a look at participants' answers, stratified by education, age, and gender. With respect to high school education, we see that people with more than 10 years of schooling have a relatively high propensity to answer correctly (72.6% vs. 54.5% of people with 10 years of schooling or less). Moreover, people with more than 10 years of schooling are very unlikely to either commit the hot hand fallacy (4.6%) or to answer "I don't know" (4.3%). Still, they frequently exhibit the gambler's fallacy: 18.5% of them estimate the probability of tails in the next toss to be higher than 50%. This finding suggests that the gambler's fallacy is prevalent even among highly educated individuals.

Looking at the control variables, we find that gender is an important factor: 65.6% of men give the correct answer, whereas only 56.0% of women do so. In particular, women

Table 1.2: Probability Judgment - Descriptive Statistics

	Hot hand	Correct answer	Gambler's fallacy	Don't know	Total sample
All	87 (8.8%)	596 (60.4%)	208 (21.1%)	95 (9.6%)	986 (100%)
Years of schooling > 10	15 (4.6%)	236 (72.6%)	60 (18.5%)	14 (4.3%)	325 (100%)
Years of schooling ≤ 10	72 (10.9%)	360 (54.5%)	148 (22.4%)	81 (12.3%)	661 (100%)
Female	51 (9.6%)	296 (56.0%)	119 (22.5%)	63 (11.9%)	529 (100%)
Male	36 (7.9%)	300 (65.6%)	89 (19.5%)	32 (7.0%)	457 (100%)
Age < 50	45 (8.1%)	349 (63.1%)	134 (24.2%)	25 (4.5%)	553 (100%)
Age ≥ 50	42 (9.7%)	247 (57.0%)	74 (17.1%)	70 (16.2%)	433 (100%)

Table 1.2: Probability Judgment - Descriptive Statistics. "Hot hand" refers to answers in the interval [0%,50%), "Correct answer" refers to answers equal to 50%, and "Gambler's fallacy" refers to answers in (50%,100%].

are much more likely to answer "I don't know". Younger people (below 50 years of age) are more likely to give the correct answer, but they are also more likely to commit the gambler's fallacy. Older people are much more likely to answer "I don't know" (16.2% vs. 4.5%).

In the following regression analysis we test whether these determinants are statistically significant and robust to controlling for background characteristics. Table 1.3 presents probit estimates with the dependent variable being equal to 1 if a respondent gives the correct answer of 50%.[7] It turns out that the effect of schooling is large and significant: an additional year of schooling is related to a 4%-point increase in the probability of giving the correct answer ($p < 0.01$), controlling for cognitive ability. As the baseline of correct answers in the total sample is 60.4%, this effect is quite sizeable. For the cognitive ability measures, we see that the coefficient for the word fluency measure is large and significant, whereas the coefficient for the perceptual speed measure is small and insignificant. These results suggest that the cognitive pragmatics (general knowledge) have a decisive impact on giving the correct answer, whereas mechanical cognitive ability (perceptual speed, quick comprehension) is not relevant to answering the question at hand. As the task does not allude to computational skills but is rather testing knowledge that is part of the general

[7]For simplicity, answering "I don't know" is categorized as a wrong answer in these regressions. If we exclude these observations from the analysis (i.e., categorize them as missing), the results are very similar. The coefficients for gender and years of schooling remain highly significant, only the word fluency measure turns out to be insignificant.

Table 1.3: Probability Judgment - Determinants (I)

Dependent variable: =1 if answer is 50%, =0 otherwise

	(1)	(2)	(3)	(4)
Word fluency measure	0.044**	0.039**		
	[0.017]	[0.018]		
Perceptual speed measure			0.003	-0.010
			[0.019]	[0.020]
Age	-0.001	-0.000	-0.002	-0.001
	[0.001]	[0.001]	[0.001]	[0.001]
1 if female	-0.090***	-0.092***	-0.084**	-0.090***
	[0.033]	[0.035]	[0.032]	[0.033]
Years of schooling		0.042***		0.046***
		[0.010]		[0.010]
N. Obs.	842	801	892	847
Prob > χ^2	0.001	0.000	0.016	0.000
Pseudo R^2	0.016	0.032	0.009	0.028

Table 1.3: Probability Judgment - Determinants (I). Probit estimates, marginal effects evaluated at the mean with standard errors in parentheses. Dependent variable is a dummy indicating whether a respondent gave the correct answer ("50%"). Observations with the answer "I don't know" are included as incorrect. Cognitive ability measures are standardized. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

education, this is a plausible finding. Regarding the control variables, we find a significant gender effect that persists even if we include age, cognitive ability and years of schooling in the regression. According to the estimates, the probability of giving the correct answer is about 9%-points lower for women ($p < 0.01$).[8]

To address the interplay between the nature of participants' cognitive biases and their background characteristics in more detail, we estimate a multinomial logit model (Table 1.4). This regression method allows us to determine whether the differences between the separate answer categories are statistically significant. The dependent variable indicates in which of the four categories a respondents' answer is located: either it is correct, or it is consistent with either the gambler's fallacy or the hot hand fallacy, or the respondent

[8] A similar gender effect has been shown by Charness and Levin (2005), who conducted a laboratory experiment to test under which circumstances individual behavior in a probabilistic decision making task is consistent with standard economic theory. Our data allows us to show in a representative sample that the gender effect persists when we control for background characteristics.

Table 1.4: Probability Judgment - Determinants (II)

Dependent variable: =1 if 50%, =2 if GF, =3 if HH, =4 if Don't know

	(1)			(2)		
	Gambler's fallacy	Hot hand	Don't know	Gambler's fallacy	Hot hand	Don't know
Word fluency measure	-0.058	-0.112	-0.638***			
	[0.094]	[0.137]	[0.169]			
Perceptual speed measure				0.145	-0.154	-0.049
				[0.106]	[0.143]	[0.180]
Years of schooling	-0.122**	-0.287***	-0.224***	-0.145***	-0.296***	-0.272***
	[0.053]	[0.080]	[0.087]	[0.051]	[0.076]	[0.095]
Age	-0.012**	-0.004	0.035***	-0.004	-0.003	0.035***
	[0.006]	[0.008]	[0.008]	[0.006]	[0.008]	[0.010]
1 if female	0.341*	0.320	0.637**	0.293*	0.388	0.748**
	[0.183]	[0.258]	[0.282]	[0.174]	[0.246]	[0.303]
Constant	0.547	1.018	-2.139**	0.471	1.026	-1.795
	[0.651]	[0.922]	[1.068]	[0.625]	[0.880]	[1.176]
N. Obs.		801			847	
Prob > χ^2		0.000			0.000	
Pseudo R^2		0.057			0.040	

Table 1.4: Probability Judgment - Determinants (II). Multinomial logit estimates, standard errors in parentheses. Dependent variable =1 if correct answer ("50%"), =2 if gambler's fallacy (> 50%), =3 if hot hand fallacy (< 50%), and =4 if "don't know". The reference category is the correct answer. Cognitive ability measures are standardized to have mean 0 and standard deviation 1. Significance at the 1%, 5%, and 10% level is denoted by ***, ** and *, respectively.

answered "I don't know". In the estimations, the reference group consists of those respondents who answered the probability question correctly. This allows us to analyze which factors determine whether a respondent exhibits a particular bias. The estimates confirm our earlier descriptive analysis. For instance, the effect of years of schooling is highly significant: schooling reduces the probability of making a mistake. This finding holds for each of the three possible mistakes, be it the hot hand fallacy, the gambler's fallacy, or the answer "I don't know". Remarkably, the effect of schooling is quite asymmetric: Whereas more schooling protects people from committing the hot hand fallacy (coefficient -0.287, $p < 0.01$), its impact on averting the gambler's fallacy is considerably weaker (coefficient

-0.122, $p = 0.02$). This is in line with the descriptive evidence which showed that the gambler's fallacy is quite common among highly educated individuals, whereas the hot hand fallacy is mostly confined to respondents with 10 years of schooling or less. With regard to the control variables, we find that women and older people are significantly more likely to answer "I don't know".

Taken together, our findings regarding the determinants of biased probability judgment all point in the same direction: more schooling increases the likelihood that a respondent gives the correct answer in the probability task. From a policy perspective, it is important to stress that we find the positive effect of schooling to be a *direct* effect. A competing hypothesis would be that schooling has only an indirect effect. For instance, one could presume that the only determinant of probability judgment is cognitive ability, which also affects the amount of schooling a given person obtains. To avoid this confound, we controlled in all regressions for cognitive ability by including measures of respondents' word fluency and perceptual speed. As the coefficient on years of schooling remains highly significant, the estimates suggest that schooling directly affects people's capability for probability judgment.

1.4. Economic Consequences of Biased Probability Judgment

Uncertainty is a crucial factor in many economic decisions. One would therefore expect that biased probability judgment can have a detrimental effect on individual economic outcomes. We propose two domains in which decisions are of a nature that closely resembles the structure of our probability judgment task: job search decisions by an unemployed person and consumption decisions by a cash-constrained consumer. In the following, we first develop the predictions of how the gambler's fallacy and the hot hand fallacy affect decision making in these domains. In a second step we take our hypotheses to the data.

1.4.1. Behavioral Predictions. The biases that are in the focus of this chapter are likely to affect economic decision making in domains where agents base their decisions on a sequence of realizations of a random process. A straightforward translation of this environment can be found in the domain of job search. As a thought experiment, assume that a person is looking for a job and sends out a number of applications. The relevant sequence of random outcomes consists then of the reactions that the job seeker receives on his applications: they can be either negative (a rejection) or positive (a job offer). After observing the realization of a sequence of outcomes, the job seeker has to decide whether to continue his search for a job or not.

Of course, many factors can play a role in the job finding process. For instance, the institutional environment, the particular skills of a job-seeker, as well as his previous work experience might have a large influence on success in the labor market. Still, our approach allows us to speculate about an additional influence that might play a role on top of these factors: the impact of a job-seeker's perception of probabilities on actual job search behavior. Given the nature of the cognitive biases we investigate, we are interested mainly in situations where a streak of similar outcomes has occurred prior to the decision. Consider the case in which a job seeker has received a streak of rejections. Then, the two probability judgment biases, gambler's fallacy and hot hand fallacy, lead to the following predictions: a person who is prone to the gambler's fallacy will believe that the streak of rejections is going to end, which implies that a job offer has now become *more* likely. As a consequence, he will be encouraged to continue his search for a job. In contrast, a person who is prone to the hot hand fallacy is going to believe that the streak of rejections is likely to continue. Given this belief, the worker will become discouraged and may give up his search for a job altogether.[9] We therefore predict that job-seekers who are prone to the gambler's fallacy face a high probability of leaving unemployment. In contrast, the hot hand fallacy can lead to prolongued unemployment by biasing the job-seeker's beliefs about his job finding probability such that they become too pessimistic.[10] Note that, due to the need for a streak of rejections to occur, our prediction is unlikely to affect persons who just became unemployed a short while ago. Rather, we would expect the detrimental effect of the hot hand fallacy to play a role for job seekers who have been unemployed for a long time.

Another domain where biased probability judgment might have a substantial effect on economic behavior is in the domain of consumption decisions. Assume that a cash-constrained consumer has to decide whether to make a large purchase that exceeds the amount of funds that is currently available in his bank account. Thus, in order to make the purchase, the consumer would have to overdraw his account. In this context, the sequence of random outcomes on which the decision is based can be thought of as unexpected income shocks that are either positive (e.g., finding a bank note on the sidewalk, winning money in a game of poker with friends) or negative (e.g., receiving a speeding ticket, having a bill to pay that is higher than anticipated). The consumption decision will then depend

[9]For related work on the role of a worker's personal job search history and the discouraged workers effect, see Falk, Huffman, and Sunde (2006a) and Falk, Huffman, and Sunde (2006b).

[10]In the opposite case (in which a streak of job offers has occurred) the theoretical predictions are less clear. Here, it is very likely that both a gambler's fallacy type and a hot hand fallacy type are going to accept one of the job offers and therefore stop searching for a job.

on the belief whether it is likely that a positive income shock occurs in the near future. If this probability is high, it is optimal to overdraw the bank account for the short period until the positive income shock realizes. If, instead, this probability is low, it is optimal to postpone the purchase until it can be made without overdrawing the account.

Again, we are interested in a situation where the decision maker has experienced a streak of similar outcomes. Consider the case in which a streak of negative income shocks has occurred. The probability judgment biases lead to the following behavioral predictions: if the consumer is prone to the gambler's fallacy, he will have the belief that the streak of negative income shocks is likely to end, such that a positive income shock will realize with a high probability in the near future. Thus, his inclination to overdraw the bank account in order to make the purchase will be high. In contrast, the hot hand fallacy will lead to the opposite prediction: if the consumer believes that the streak of negative income shocks is likely to continue, he will refrain from overdrawing his account and will not make the purchase.[11] In sum, a person who is prone to the gambler's fallacy is predicted to be more likely to have an overdrawn bank account, as the biased belief that a positive income shock is "due" can lead to persistent household debt. By contrast, a person who is prone to the hot hand fallacy will be less likely to have an overdrawn bank account.

1.4.2. Empirical Results. To test the predictions regarding employment status with our data, we run two sets of regressions. First, we estimate a probit model in which the dependent variable is a dummy equal to one if the respondent is registered as unemployed at the time of the interview. In a second set of regressions, the dependent variable is an indicator for whether the respondent is long-term unemployed, i.e., registered as unemployed for 12 months or more. As explanatory variables, we include dummies for the observed cognitive biases: the gambler's fallacy, the hot hand fallacy, and the answer "I don't know". The reference group consists therefore of those respondents who answered the probability judgment task correctly. Results from regressions with the unemployment dummy as dependent variable are presented in Table 1.5. In column (1) we control only for age and gender and find that both the hot hand dummy and the gambler's fallacy dummy are positive and weakly significant. Adding controls for education and for wealth renders both coefficients insignificant, see columns (2) and (3).

[11]Predictions in the opposite case (in which a streak of positive income shocks has occurred) are ambivalent: if the consumer is prone to the hot hand fallacy, he will believe that the positive income shocks are going to continue and he will decide to make the large purchase. If the respondent is prone to the gambler's fallacy instead, he will expect that a negative income shock is likely to occur in the near future. Still, if the streak of positive income shocks has led to a large amount of funds available in his account, he might nevertheless make the purchase.

Table 1.5: Economic Outcomes - Unemployment

Dependent variable:	1 if unemployed			1 if long-term unemployed		
	(1)	(2)	(3)	(4)	(5)	(6)
1 if gambler's fallacy	0.045*	0.032	0.031	0.029	0.019	0.025
	[0.028]	[0.027]	[0.031]	[0.023]	[0.022]	[0.025]
1 if hot hand	0.089**	0.058	0.056	0.089***	0.064**	0.061*
	[0.046]	[0.043]	[0.046]	[0.042]	[0.038]	[0.040]
1 if "don't know"	0.050	0.031	0.031	0.060*	0.044	0.052
	[0.045]	[0.043]	[0.053]	[0.040]	[0.037]	[0.046]
Age	-0.002***	-0.003***	-0.002***	-0.001**	-0.002***	-0.001
	[0.001]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
1 if female	-0.026	-0.026	-0.033	-0.021	-0.019	-0.022
	[0.020]	[0.020]	[0.024]	[0.016]	[0.016]	[0.018]
Years of schooling		-0.022***	-0.019***		-0.015***	-0.013**
		[0.006]	[0.007]		[0.004]	[0.005]
Household wealth			-0.012***			-0.008***
			[0.002]			[0.002]
N. Obs.	985	938	733	985	938	733
Prob > χ^2	0.000	0.000	0.000	0.017	0.000	0.000
Pseudo R^2	0.035	0.066	0.117	0.029	0.058	0.119

Table 1.5: Economic Outcomes - Unemployment. Probit estimates, marginal effects evaluated at the mean with standard errors in parentheses. In columns (1) to (3), dependent variable =1 if respondent is registered as unemployed on the day of the interview, =0 otherwise. In columns (4) to (6), dependent variable =1 if respondent is registered as unemployed for more than 12 months, =0 otherwise. In all specifications, the reference category is the group of people who answered the probability question correctly ("50%"). Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

As the predictions we developed in Section 1.4.1 are especially relevant in addressing long-run effects, we concentrate our analysis in the next step on persons who have been looking for a job for an extended period of time. Columns (4) to (6) present results from regressions where the dependent variable is a dummy equal to one for respondents who are long-term unemployed. The explanatory variables are the same variables as before. Our estimates show that, controlling for age and gender, the hot hand fallacy is positive and significant at the 1%-level. Thus, the predictions regarding the effect of the hot hand fallacy on long-term unemployment are supported by the data. If we control in addition for respondents' educational background, the coefficient for the hot hand fallacy remains significant at the 5%-level. Even if we control for education and wealth, the coefficient for the hot hand fallacy dummy remains significant at the 10%-level and the marginal effect indicates that the probability of being long-term unemployed is increased by 6.1%-points.[12] Given that the baseline of long-term unemployment in the sample is 6.8%, this is a considerable effect. We therefore conclude that our predictions regarding the detrimental effect of the hot hand fallacy are supported in case of long-term unemployed job seekers. Moreover, we find for all specifications that the gambler's fallacy has no significant effect on employment status, which is again in line with our predictions.

Next, we empirically test the predictions regarding consumption decisions of cash-constrained consumers. To this aim, we analyze the interplay of probability judgment biases and the decision to overdraw one's bank account. In a first set of regressions, the dependent variable indicates how many days per year a respondent's bank account is overdrawn. The variable can take on four distinct values, as the answer space was split up into four intervals (0 days, 1 to 30, 31 to 90, more than 90 days). As explanatory variables we include again the dummies for whether a person exhibits a bias in probability judgment, with the reference group being those respondents who gave the correct answer to the probability task. Results of ordered probit regressions are presented in columns (1) to (3) of Table 1.6. The estimates show that the dummy for whether a respondent is prone to the gambler's fallacy is positive and significant at the 1%-level even if we control for age, gender, and education. Adding controls for net household income and net household wealth leaves the coefficient virtually unchanged and significant at the 5%-level. This gives support to our prediction: people who are prone to the gambler's fallacy have a higher number of days per year on which their bank account is overdrawn.

[12] We deliberately chose not to control for household income in the unemployment regressions, in order to avoid endogeneity problems.

Table 1.6: Economic Outcomes - Financial Situation

Dependent variable:	Account overdrawn (days per year)			1 if account overdrawn (on day of interview)		
	(1)	(2)	(3)	(4)	(5)	(6)
1 if gambler's fallacy	0.277***	0.248**	0.247**	0.088***	0.087***	0.088**
	[0.103]	[0.105]	[0.114]	[0.034]	[0.036]	[0.040]
1 if hot hand	0.084	0.065	0.098	0.059	0.050	0.068
	[0.152]	[0.153]	[0.158]	[0.049]	[0.049]	[0.054]
1 if "don't know"	-0.249	-0.220	-0.152	-0.026	-0.022	-0.005
	[0.186]	[0.192]	[0.209]	[0.044]	[0.048]	[0.059]
Age	-0.020***	-0.019***	-0.016***	-0.004***	-0.003***	-0.002***
	[0.003]	[0.003]	[0.003]	[0.001]	[0.001]	[0.001]
1 if female	-0.044	-0.023	-0.074	0.037	0.042	0.035
	[0.085]	[0.087]	[0.095]	[0.024]	[0.026]	[0.029]
Years of schooling	-0.022	-0.021	-0.017	-0.019***	-0.018**	-0.014
	[0.024]	[0.025]	[0.027]	[0.007]	[0.008]	[0.009]
Net monthly household income		-0.037	0.090		-0.035*	0.013
		[0.068]	[0.082]		[0.020]	[0.026]
Net household wealth			-0.030***			-0.011***
			[0.010]			[0.003]
N. Obs.	881	840	695	938	879	716
Prob > χ^2	0.000	0.000	0.000	0.000	0.000	0.000
Pseudo R^2	0.048	0.045	0.049	0.050	0.049	0.064

Table 1.6: Economic Outcomes - Financial Situation. In columns (1) to (3), ordered probit estimates, standard errors in parentheses. Dependent variable takes four different values: =1 if account overdrawn 0 days per year, =2 if 1 to 30 days per year, =3 if 31 to 90 days per year, and =4 if more than 90 days a year. In columns (4) to (6), probit estimates, marginal effects evaluated at the mean with standard errors in parentheses. Dependent variable =1 if account overdrawn at the day of the interview, =0 otherwise. In all specifications, the reference category is the group of people who answered the probability question correctly ("50%"). Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

As it might be relatively complicated for respondents to assess the number of days per year on which their account is in the negative, the survey also included the straightforward question of whether a person's account was overdrawn at the time of the interview. A look at the raw data reveals large effects: out of the respondents who answer the probability question correctly, 14.3% have an overdrawn bank account. In contrast, this figure is 24.5% among the group of people who are prone to the gambler's fallacy. Results of a probit estimation with this simple measure as dependent variable are presented in columns (4) to (6). The estimates are very similar to our earlier findings: in most specifications, the coefficient on the gambler's fallacy dummy is positive and significant at the 1%-level. The marginal effects indicate that being prone to the gambler's fallacy increases the probability of having an overdrawn bank account by about 8.8%-points ($p = 0.020$), controlling for age, gender, education, income, and wealth. This is a sizeable effect, as the share of respondents with an overdrawn bank account is 16.6% in the total sample. Again, these findings are in line with the predictions: the gambler's fallacy has a substantial impact on consumers' decision to overdraw their bank account.

Taken together, our results support the behavioral hypotheses that were developed in Section 1.4.1. In particular, we find that the gambler's fallacy affects financial decision making, whereas the hot hand fallacy has an impact on job search decisions. These findings suggest that it is not biased probability judgment *per se* that affects economic outcomes. Rather, depending on the context in which economic decisions are made, the specific form of a person's probability judgment bias plays a decisive role.

1.5. Conclusion

This chapter has addressed three closely related research questions. First, it investigated people's ability to make simple probability judgments. For this purpose, we used a specifically designed probability judgment task that was administered to a representative sample of the German adult population. The results showed that more than a third of the respondents was unable to answer the probability question correctly, indicating that a substantial part of the population has difficulties with making simple probability judgments. Among the incorrect answers, by far the most frequent bias was the gambler's fallacy, i.e., the tendency to overestimate the occurrence of alternations in random sequences. Second, we have addressed the determinants of biased probability judgment. Our results have shown that education (years of schooling) and a knowledge-based measure of cognitive ability are positively related to performance in the probability judgment task.

The third part of the chapter has explored the relation between the observed probability judgment patterns and respondents' behavior in two domains: job search and financial decision making. The hot hand fallacy, i.e., the tendency to overestimate the occurrence of streaks in random sequences, was shown to be significantly related to a higher probability of being long-term unemployed. In contrast, the gambler's fallacy corresponds to a higher probability of overdrawing one's bank account.

Having established that biased probabilistic reasoning translates into inferior economic outcomes, our finding regarding the impact of schooling becomes important from a policy perspective. Our estimates have shown that schooling has a large and significant impact on reducing people's cognitive biases. The fact that the estimates were positive and significant even if we controlled for cognitive ability suggests that the knowledge obtained in school mitigates probability judgment biases in a direct way. Thus, it may be worthwile to put a stronger focus on teaching simple probabilistic reasoning already in the early grades of high school. More generally, an increased dissemination of basic knowledge about random processes might help people to make better decisions in the economic domains of life.

From an applied perspective, our findings have straightforward implications for the design of policy measures that lead to a debiasing of decision makers. With respect to labor market policy, our results suggest that job centers should offer courses that teach job-seekers about the probabilities that play a role in the application process. In particular, it could be helpful to point out to job-seekers that they should not rely on their personal search history in order to form a belief about the probability with which future applications will be successful. Rather than believing that their individual streak of unsuccessful applications will be continued, job-seekers should learn to focus on more general economic measures such as the overall labor market conditions. Similar policy implications can be derived for the domain of household finances. Here, counselors who give advice to over-indebted households could inform consumers who exhibit the gambler's fallacy that they should avoid to overestimate the probability of a positive income shock to occur in the near future.

Of course, the analysis in this chapter has limitations. For instance, we were not able to identify the exact channels through which a cognitive bias influences behavior. While some of our results are in line with the gambler's fallacy being a sign of overoptimism (as in the overdrawn bank account measure), this behavior might as well be related to time-inconsistent preferences (cf. Laibson 1997). In future research, we plan to incorporate incentivized measures of people's time and risk preferences into our analysis, in

order to identify the channels through which biased probability judgment affects economic behavior.

CHAPTER 2

The Non-Use of Bayes' Rule: Representative Evidence on Bounded Rationality

2.1. Introduction

Bayes' rule is an important cornerstone of economic theory. It prescribes that agents make decisions under uncertainty by correctly processing new information conditional on the prior information they have. Experimental evidence suggests however that people often have difficulties with making judgments that involve conditional probabilities (Tversky and Kahneman 1971, Grether 1980, Charness and Levin 2005). The consequences of flawed probabilistic reasoning can be severe: for instance, an insufficient understanding of the importance of base rates has resulted in suicides of blood donors who were mistakenly tested positive for HIV in the 1980s, and incorrect use of probabilistic information has led to wrongful convictions in criminal cases (for detailed accounts, see Gigerenzer 2003). Despite an increased interest in behavioral deviations from the Bayesian updating model, insights about the determinants of biased probability judgment remain scarce. Moreover, little is known about the prevalence of judgment biases in the general population. Existing studies have been restricted to selected groups, such as university students, chemical workers, or medical doctors (Viscusi and O'Connor 1984, Hoffrage, Lindsey, Hertwig, and Gigerenzer 2000).

In this chapter, we measure people's ability for Bayesian reasoning in a representative sample of the German population, consisting of more than 1,000 observations. This approach enables us to evaluate the actual prevalence of cognitive biases that are described in the experimental literature. In addition, we analyze the determinants of biased probability judgment by relating respondents' probability estimates to individual background characteristics such as education, cognitive ability, gender, and age.

The question we use to elicit respondents' ability in assessing conditional probabilities is framed as a judgment about weather conditions. This framing embeds the task in a context in which respondents are likely to be familiar with probabilities, as weather forecasts in TV and radio often refer to probabilistic measures. Participants received all information in form of natural frequencies, which has been shown to enhance participants'

performance (Gigerenzer and Hoffrage 1995). The exact wording of the question was as follows (translated from German):[1]

> *Imagine you are on vacation in an area where the weather is mostly sunny and you ask yourself how tomorrow's weather will be. Suppose that, in the area you are in, on average 90 out of 100 days are sunny, while it rains on 10 out of 100 days. The weather forecast for tomorrow predicts rain. On average, the weather forecast is correct on 80 out of 100 days. What do you think is the probability, in percent, that it is going to rain tomorrow?*

Respondents had to answer with a number between 0% and 100%. Alternatively, they could give the answer "I don't know".

The normative solution to this question is a rain probability of about 31%.[2] Although the problem may seem complicated at first, an easy way to solve it is by finding the correct information partition. Out of 100 days, the sun shines on 90 days and it rains on 10 days. For the 90 sunny days, the weather forecast predicts sunshine on 72 days and rain on 18 days. For the 10 rainy days, it predicts sunshine on 2 days and rain on 8 days. Thus, given that the forecast predicts rain, the probability that it will actually rain is $8/(18+8) \approx 31\%$. The structure of our question is very similar to the "cab problem", introduced by Kahneman and Tversky (1972). However, in order to identify correct answers without ambiguity, we modified the base rate in our task such that the Bayesian solution is neither close to the base rate (10%), nor to the middle of the answer space (50%), which might be a focal point.[3]

Our results show that only a very small proportion of answers is close to the Bayesian solution. In particular, we find that *none* of the respondents gives the correct answer of 31%. A probability of 30% is indicated by 2.8% of the respondents, and the interval 31% ± 10%-points contains only 5.9% of all answers. The most frequent bias we find in the data is consistent with the so-called base rate neglect, a phenomenon that is well-established in the cognitive psychology literature (Bar-Hillel 1980, Tversky and Kahneman 1982, Koehler 1996). In terms of our task, this bias implies that respondents ignore the

[1]The original wording was: "Stellen Sie sich vor, Sie sind im Urlaub in einer Gegend, in der meistens schönes Wetter ist und Sie fragen sich, wie wohl morgen das Wetter sein wird. Angenommen an Ihrem Urlaubsort scheint im Durchschnitt an 90 von 100 Tagen die Sonne und an 10 von 100 Tagen regnet es. Für morgen lautet der Wetterbericht der Wettervorhersage dass es regnen wird. Die Wettervorhersage sagt das Wetter für den nächsten Tag im Durchschnitt an 80 von 100 Tagen richtig voraus. Was glauben Sie: Wie hoch ist die Wahrscheinlichkeit, ausgedrückt in Prozent, dass es morgen regnet?".

[2]Bayes' rule: $\Pr(R|RF) = \frac{\Pr(RF|R)\Pr(R)}{\Pr(RF)} = \frac{0.8*0.1}{0.9*0.2+0.1*0.8} = 0.308$.

[3]In the cab problem, the probability with which the additional information is correct is 80% (as it is in our task), but the base rate is 15% instead of 10%, leading to a Bayesian posterior of 41%. This can make interpretation of the results problematic, as the correct solution is then relatively close to the focal point of 50%.

information about the base rate of rainy days and concentrate only on the probability with which the weather forecast is correct. The resulting estimate of 80% is given by 24.2% of the sample. Thus, in the setup under investigation, base rate neglect is over four times as frequent as answers that are located in the 10%-point interval around the Bayesian solution. Another bias that is well-known from the psychological literature is the tendency to answer with the pure base rate. In our setup it leads respondents to an estimate that is equal to the prior probability of rain, ignoring the information provided by the weather forecast. The corresponding answer, an estimate of 10%, is given by 10.8% of the sample. Taken together, this set of results has two important implications: first, in a representative sample of the population, only a very small fraction of answers is close to the normatively correct answer. Second, deviations from the Bayesian solution appear to be systematic, as certain mistakes are by far more frequent than others.

The main goal of this chapter is an assessment of the prevalence of probability judgment biases and an analysis of their determinants. We deliberately do not address the question of whether a certain bias leads to "good" or "bad" estimates, as this depends crucially on the specific parameters of the decision environment. For instance, when looking at the bias of base rate neglect, it follows directly from Bayes' rule that the difference between the Bayesian posterior and the base rate neglect estimate becomes very large for base rates close to 0%, whereas it becomes negligible for base rates close to 50%. To illustrate this point, assume the base rate of rainy days in our weather task was not 10%, but 5%. The resulting posterior would be 17.4%, which is far away from the estimate of 80% that is obtained by base rate neglect. In contrast, a base rate of 45% would lead to a Bayesian posterior of 76.6%, which is very close to the base rate neglect estimate. Thus, from a decision-making perspective, base rate neglect can be a useful heuristic in environments where the base rate of an event is close to 50%. The question of whether respondents realize under which circumstances a certain heuristic is more useful than another is beyond the scope of this chapter.

In the economic literature, a common explanation for the existence of judgment biases is bounded rationality: people decide rationally under the constraint that they have cognitive limitations which may prevent them from finding the normatively correct solution (for an overview, see Conlisk 1996). Therefore, when analyzing the determinants of biased probability judgment, proxies for respondents' cognitive ability are of particular interest. A unique feature of our data set is that it includes two simple measures of cognitive ability which are based on two sub-modules from one of the most widely used IQ tests. Moreover,

the data contain information on respondents' years of schooling and on whether they hold a university degree. As control variables, respondents' age and gender are included.

Our empirical analysis of the determinants of biased probability judgment leads to surprising results: a high level of education does not increase the likelihood of giving an answer that is close to the Bayesian solution. Rather, we find that highly educated respondents are *more* likely to exhibit base rate neglect. These effects are particularly pronounced for people with an academic background: among respondents who hold a university degree, 37.9% give the base rate neglect estimate of 80%, and 2.6% give an answer that is in the 10%-point interval around the Bayesian solution. In contrast, among respondents without a university degree, 22.4% are prone to base rate neglect, and 6.3% give an answer in the Bayesian interval. Stratifying the sample by years of schooling leads to very similar results, with the share of base rate neglect being much higher among respondents with more than 10 years of schooling (31.5% vs. 20.6%). We confirm the statistical significance of our findings in a number of regressions, controlling for respondents' background characteristics. In particular, marginal effect estimates from probit regressions reveal that the probability to commit base rate neglect is roughly 15%-points higher for respondents who hold a university degree, controlling for cognitive ability, age, and gender. In a similar vein, one additional year of schooling leads to a 2%-point increase in the probability to commit base rate neglect. Furthermore, our estimates show that the impact of cognitive ability is very similar to the impact of education: a higher score in the cognitive ability measures significantly increases the probability to commit base rate neglect and leads to estimates that are further away from the normative solution. A series of robustness checks shows that our results continue to hold when estimating multinomial logit models and when using a linear measure of respondents' biases.

An interpretation of our results in light of standard theories of human decision making seems difficult, as all our measures for education and for cognitive ability have a significant detrimental effect on respondents' probability judgment. Even approaches that allow for bounded rationality, e.g., by incorporating deliberation cost (Smith and Walker 1993), cannot explain our findings. In a cognitive task as complex as the one we use in this chapter, one would expect deliberation cost to be relatively high for people with less formal education. In contrast, for highly educated people deliberation cost should be relatively low. Other things equal, this reasoning would imply that more educated people perform better in assessing conditional probabilities. Our results indicate the contrary, as education, in particular university education, increases the likelihood that respondents are

led astray in the probability judgment task. An identification of the exact channels which are responsible for the detrimental effect of education and cognitive ability on Bayesian judgment is beyond the scope of this chapter.

From a methodological viewpoint, the design of our study allows us to draw conclusions about the general population that cannot always be obtained with non-representative subject pools. Unfortunately, an adequate comparison of our findings to results from previous research is difficult, as most psychological studies do not report the full distribution of answers. Rather, their focus is either on the share of base rate neglect in the data, or on the share of Bayesian answers, depending on the respective research question. In a seminal study by Casscells, Schoenberger, and Graboys (1978), 60 students and staff from Harvard Medical School were asked to estimate the probability that a patient actually has a disease, given a positive result in a diagnostic test. Parameters of the question were chosen such that the normatively correct answer is approximately 2%, whereas base rate neglect leads to an estimate of 95%. Results show that about 18% of the participants give the correct answer. In contrast, almost half of the respondents exhibit base rate neglect. The average estimate in the sample is 56%. Eddy (1982) reports from a similar study that among 100 physicians, 95 exhibit base rate neglect. In an overview article, Tversky and Kahneman (1982) state that base rate neglect is typically both the modal answer and the median answer in Bayesian updating tasks. Unfortunately, their article neither contains figures regarding the exact share of respondents committing base rate neglect, nor does it report the typical share of normatively correct answers. Gigerenzer and Hoffrage (1995) investigate 15 distinct Bayesian updating tasks in order to show that presenting information in natural frequencies instead of probabilities enhances respondents' tendency to apply Bayesian reasoning. The experiments were conducted with 60 undergraduates (mainly psychology majors) from the University of Salzburg, Austria. In the updating task that is most similar to the one we use in this chapter, the authors find that about 15% of the sample give the Bayesian answer. However, no information is reported about the distribution of non-normative answers. In particular, nothing is said about the share of subjects who exhibit base rate neglect or who answer with the pure base rate. A study that reports a full picture of the answers in a Bayesian updating task is provided by Cosmides and Tooby (1996). The task used in their study is a standard medical diagnosis problem, where the base rate of the disease is 0.1% and the false positive rate of the diagnostic test is 5%. The Bayesian posterior for having the disease, conditional on a positive test result, is approximately 2%. In their sample of 25 Stanford undergraduates, base rate

neglect is exhibited by 14 subjects (56% of the sample), whereas 3 subjects (12%) give the Bayesian answer, and another 3 subjects (12%) answer with the pure base rate. Closest to our analysis is a study by Stanovich and West (1998) who analyze the relationship between respondents' cognitive ability and their assessment of probabilities in a sample of 211 psychology undergraduates. We address their findings in detail when discussing our results in Section 2.3.4.

In comparison to previous research, the share of base rate neglect in our sample is much lower than it is in studies with non-representative subject pools. Taking our analysis of the determinants of probabilistic judgment biases at face value, one reason for this difference are the different education levels of participants in our study and participants in previous research. The existing studies have mostly been conducted with academically trained individuals. In contrast, only 12% of the respondents in our representative data set hold a university degree. Finding a much higher share of base rate neglect in samples that consist entirely of academically trained subjects is in line with our empirical analysis of the determinants of biased probability judgment, as our results have shown that the likelihood of being prone to this particular bias rises significantly with respondents' education level. These findings are important: one might be tempted to conclude from previous studies that base rate neglect must be an extremely widespread phenomenon, since tests with highly sophisticated students and professionals should constitute a lower bound for the prevalence of flawed decision making in the general population. Our results are not consistent with this view. In light of our findings, it is likely that previous studies with highly educated subjects resulted in an *upper* bound for the prevalence of base rate neglect.

When considering the share of Bayesian answers and the share of answers that give the pure base rate, generalizing the results from student subject pools to the population at large seems much more appropriate. In these two domains, differences between our findings and the existing literature are relatively small. This is again in line with our empirical analysis of the underlying determinants, as we find that respondents' level of education has neither a significant impact on their likelihood to give a Bayesian estimate, nor on their likelihood of answering with the pure base rate.

The remainder of this chapter is organized as follows: Section 2.2 contains a description of the data, Section 2.3 presents the results, while Section 2.4 concludes.

Table 2.1: Sample Statistics

Variable	average	std. dev.	num. obs.
1 if female	0.537	(0.499)	1,012
Age	47.555	(18.371)	1,011
Word fluency test (raw data)	22.893	(10.933)	849
Word fluency test (standardized)	0	(1)	849
Perceptual speed test (raw data)	28.339	(9.853)	902
Perceptual speed test (standardized)	0	(1)	902
Years of schooling	10.328	(1.829)	965
1 if university degree	0.118	(0.322)	1,012

Table 2.1: Sample Statistics.

2.2. Data

The data set under investigation is a representative sample of the population living in Germany aged 16 years and older. It is the same data set as in Chapter 1.[4] For details on the way the data were collected see Section 1.2. In total, 988 participants answered the probability question. In addition, the data contain a number of background characteristics such as age, gender, and education. A novel feature of the survey is that it elicited measures of respondents' cognitive ability. One measure elicits respondents' perceptual speed through a test in which participants have to find correspondences between symbols and digits. The other measure is a word fluency test that provides a proxy of respondents' general knowledge. The detailed design of these two measures has been described in the previous chapter (see Section 1.2).

Sample statistics are shown in Table 2.1. About 53.7% of the sample are female, average age is 47.6 years, and respondents have on average 10.3 years of schooling. The measures for cognitive ability have a mean value of 28.3 (perceptual speed test) and 22.9 (word fluency test). For the empirical analysis, both cognitive ability measures are standardized to have mean zero and standard deviation one. Histograms of the two standardized measures have been presented in Figure 1.1.

2.3. Results

2.3.1. Prevalence of Biased Probability Judgment. A first look at the results shows that respondents' estimates regarding the probability of rain are spread out over the

[4]A full overview of the original survey has been given in Appendix A.

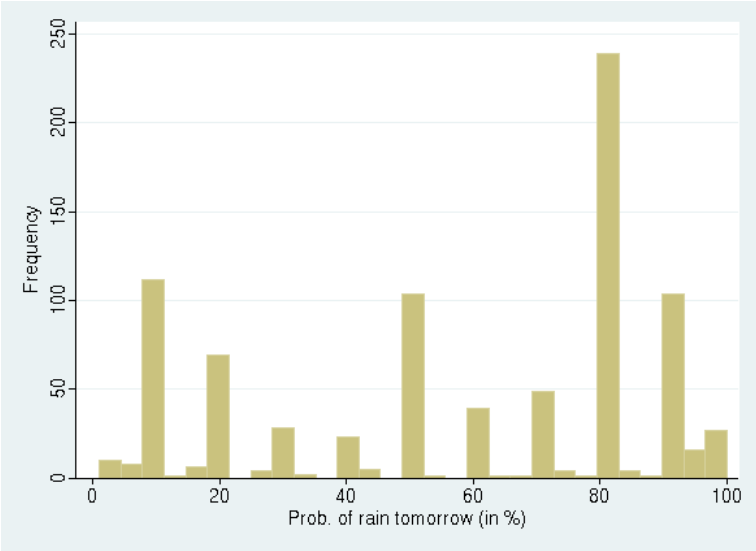


Figure 2.1: Answers to Conditional Probability Question. The normatively correct solution is a rain probability of about 31%. The 129 observations for the answer "I don't know" are not shown.

full range of the answer space (Figure 2.1). Still, we find that none of the 988 respondents gives the Bayesian answer of 31%. A rain probability of 30% is indicated by 2.8% of the respondents, and only 3.0% of answers are in the interval 31% ± 5%-points. Even the interval 31% ± 10%-points contains only a small share (5.9%) of all answers. Taken together, these findings show that a large majority of the respondents has difficulties with making Bayesian judgments in the setting at hand, as their estimates are far away from the normatively correct solution.

Next, we analyze the non-normative answers in the sample. Table 2.2 provides an overview of typical judgment biases and shows their prevalence in our data. Recall that a common bias described in the literature on conditional probability judgment is base rate neglect. We find that, in our setting, the corresponding estimate of 80% is given by 24.2% of the respondents. Thus, almost a quarter of the respondents chooses in line with the most prominent prediction for biased probability judgment. A second bias that is often reported in the literature is respondents' tendency to focus only on the base rate, thereby ignoring the information provided in addition to the base rate probability. The corresponding estimate of 10% is given by 10.8% of the sample, and the 10%-point interval around the pure base rate estimate contains 20.7% of all answers.

In sum, we find that in our setup the share of respondents who exhibit base rate neglect is about four times as large as the share of respondents whose answer is located in the Bayesian interval. Moreover, the pure base rate interval contains more than three times

Table 2.2: Bayesian Solution vs. Typical Judgment Biases

	Solution process	Resulting answer	% of answers in sample
Bayesian:	$\frac{\Pr(RF\|R)*\Pr(R)}{\Pr(RF)} = \frac{0.8*0.1}{0.9*0.2+0.1*0.8} = 0.308$	30% (20%,40%]	2.8% 5.9%
Pure base rate:	$\Pr(R) = 0.1$	10% [0%,20%]	10.8% 20.7%
Base rate neglect:	$\Pr(RF\|R) = 0.8$	80%	24.2%
Other:			49.2%

Table 2.2: Bayesian Solution vs. Typical Judgment Biases. Category "Other" consists of all answers in the intervals (40%,60%], (60%,80%), (80%,100%], and "I don't know". A full overview of the data is given in Appendix B.

as many answers as the Bayesian interval. These results are a clear indication that the pattern of incorrect answers is non-random. It is very unlikely that the observed deviations from the normative solution are due to an error source like lack of concentration or lack of motivation. Rather, the data pattern indicates that respondents' answers are driven by systematic biases in probabilistic judgment.

2.3.2. Determinants of Biased Probability Judgment. If probability judgment biases occur due to boundedly rational behavior, we should expect a significant relationship between respondents' probability estimates and proxies for their cognitive ability. We therefore proceed to investigate the interplay between respondents' answers on the one hand and their performance in the cognitive ability tests and their education levels on the other. The descriptive statistics in Table 2.3 reveal a striking finding: stratifying the sample according to years of schooling indicates that highly educated people are much more likely to commit base rate neglect, whereas they do not have a higher likelihood to give an answer that is close to the Bayesian solution. We find that 31.5% of those with more than 10 years of schooling exhibit base rate neglect, and 4.9% give answers in the interval that contains the Bayesian solution. In contrast, among people with at most 10 years of schooling, only 20.6% commit base rate neglect, and 6.3% give an answer in the Bayesian interval. The effect of education is even more pronounced if we focus our analysis on respondents with a university education: among those who hold a university degree,

Table 2.3: Descriptive Statistics

	Bayesian	Pure base rate	Base rate neglect	Other	Total
All	58 (5.9%)	205 (20.7%)	239 (24.2%)	486 (49.2%)	988 (100%)
Years of schooling > 10	16 (4.9%)	68 (21.0%)	102 (31.5%)	138 (42.6%)	324 (100%)
Years of schooling ≤ 10	42 (6.3%)	137 (20.6%)	137 (20.6%)	348 (52.4%)	664 (100%)
University degree	3 (2.6%)	30 (25.9%)	44 (37.9%)	39 (33.6%)	116 (100%)
No univ. degree	55 (6.3%)	175 (20.1%)	195 (22.4%)	447 (51.3%)	872 (100%)
Word fluency > 0	26 (4.6%)	114 (20.2%)	151 (26.7%)	274 (48.5%)	565 (100%)
Word fluency ≤ 0	32 (7.6%)	91 (21.5%)	88 (20.8%)	212 (50.1%)	423 (100%)
Perceptual speed > 0	24 (4.2%)	112 (19.8%)	148 (26.2%)	281 (49.7%)	565 (100%)
Perceptual speed ≤ 0	34 (8.0%)	93 (22.0%)	91 (21.5%)	205 (48.5%)	423 (100%)
Female	31 (5.8%)	117 (22.0%)	116 (21.8%)	268 (50.4%)	532 (100%)
Male	27 (5.9%)	88 (19.3%)	123 (27.0%)	218 (47.8%)	456 (100%)
Age < 50	27 (4.9%)	107 (19.3%)	158 (28.5%)	262 (47.3%)	554 (100%)
Age ≥ 50	31 (7.1%)	98 (22.6%)	81 (18.7%)	224 (51.6%)	434 (100%)

Table 2.3: Descriptive Statistics. Table shows number of answers in each answer category. "Bayesian" are answers in the interval (20%,40%] around the Bayesian solution of 30.8%. "Pure base rate" are answers in the interval [0%,20%] around the pure base rate of 10%. "Base rate neglect" are answers equal to 80%. Category "Other" consists of all answers in the intervals (40%,60%], (60%,80%), (80%,100%], and "I don't know". Cognitive ability measures are standardized.

37.9% are prone to base rate neglect, and 2.6% give an answer in the Bayesian interval. In contrast, among people without university degree, 22.4% exhibit base rate neglect, and 6.3% give an answer in the Bayesian interval. Very similar results obtain when we split the sample at the median according to the cognitive ability measures.[5] A look at the control variables gender and age reveals that there is a gender effect, with men being more likely to commit base rate neglect (27.0% vs. 21.8%). A similar pattern can be seen for age, with younger people being more likely to do base rate neglect (28.5% vs. 18.7%).

In the following, we use regression analysis in order to control for respondents' background characteristics and to assess whether our findings are statistically significant. In a first step, we analyze the determinants of base rate neglect, as this is both the most commonly described phenomenon in the literature and the most frequent bias in our sample. We estimate a probit model with the dependent variable being equal to one if a

[5]In particular, in the group with high cognitive ability, the share of people who give an answer in the Bayesian interval is about 4%, whereas it is about 8% in the group of people with lower cognitive ability.

Table 2.4: Determinants of Base Rate Neglect

Dependent variable: =1 if answer is 80% (base rate neglect), =0 otherwise

	(1)	(2)	(3)	(4)	(5)	(6)
Word fluency measure	0.053***	0.048***	0.038**			
	[0.015]	[0.015]	[0.016]			
Perceptual speed measure				0.051***	0.046***	0.040**
				[0.015]	[0.017]	[0.018]
1 if female		-0.053*	-0.058*		-0.045	-0.047
		[0.030]	[0.031]		[0.029]	[0.030]
Age		-0.001	-0.001		-0.000	-0.000
		[0.001]	[0.001]		[0.001]	[0.001]
Years of schooling			0.019**			0.023***
			[0.009]			[0.008]
N. Obs.	844	843	802	895	894	850
Prob > χ^2	0.000	0.001	0.000	0.001	0.002	0.000
Pseudo R^2	0.013	0.018	0.024	0.012	0.014	0.024

Table 2.4: Determinants of Base Rate Neglect. Probit estimates, marginal effects evaluated at the mean with standard errors in brackets. Dependent variable is a dummy equal to one if respondent's answer is 80% (base rate neglect). Cognitive ability measures are standardized to have mean 0 and standard deviation 1. Significance at the 1%-, 5%-, and 10%-level is denoted by ***, **, and *, respectively.

respondent makes the mistake of base rate neglect. As explanatory variables we include the measures of cognitive ability, age, gender, and years of schooling. The results are presented in Table 2.4. We find that the significant determinants of whether a respondent is prone to base rate neglect are cognitive ability and education. Remarkably, all these measures have positive coefficients, meaning that people with better cognitive ability and with more education are more likely to neglect the base rate information. For instance, the estimates in column (6) of Table 2.4 show that an increase in the measure for perceptual speed by one standard deviation leads to an increase in the probability to commit base rate neglect by about 4%-points ($p < 0.05$), whereas an additional year of schooling leads to an increase in this probability by another 2%-points ($p < 0.01$). In contrast, in most specifications the effect of age and gender is not significant once we control for cognitive ability and education.

Table 2.5: The Role of Academic Background

Dependent variable: =1 if answer is 80%, =0 otherwise

	(1)	(2)	(3)	(4)
1 if university degree	0.155*** [0.047]	0.170*** [0.048]	0.144*** [0.052]	0.167*** [0.050]
1 if female		-0.051* [0.027]	-0.052* [0.030]	-0.045 [0.029]
Age		-0.003*** [0.001]	-0.002* [0.001]	-0.001 [0.001]
Word fluency measure			0.038** [0.017]	
Perceptual speed measure				0.042*** [0.016]
N. Obs.	986	985	843	894
Prob $> \chi^2$	0.001	0.000	0.000	0.000
Pseudo R^2	0.011	0.025	0.027	0.027

Table 2.5: The Role of Academic Background. Probit estimates, marginal effects evaluated at the mean with standard errors in brackets. Dependent variable is a dummy equal to one if respondent's answer is 80% (base rate neglect). Cognitive ability measures are standardized to have mean 0 and standard deviation 1. Significance at the 1%-, 5%-, and 10%-level is denoted by ***, **, and *, respectively.

The descriptive part of the analysis has already stressed that base rate neglect is especially widespread among respondents who hold a university degree. This is an important finding, as previous research on probability judgment has relied heavily on subject pools consisting of university students and of academically trained professionals (e.g., in medicine). In contrast, the representative nature of our study entails that only 12% of the respondents have an academic background. The data set therefore allows us to assess whether the behavior of people with a university education can be generalized to the behavior of the population at large. To study this question in a regression framework, we define a dummy variable that is equal to one if the respondent holds a university degree, and equal to zero otherwise. Results from probit regressions with the base rate neglect dummy as dependent variable are shown in Table 2.5. The results are striking: controlling for gender, age, and cognitive ability, people with a university degree are between 14 and 17%-points more likely to commit base rate neglect. Given that in the total sample the

share of people who commit base rate neglect is about 24%, this is a substantial effect that is difficult to be reconciled with existing theories of human decision making.

We have seen that, next to base rate neglect, a common judgment bias is respondents' tendency to answer with the pure base rate. In our sample, 20.7% of all answers are in the interval [0%,20%] where the pure base rate of 10% is located. Still, it is not clear what the determinants of this bias are. As the descriptive statistics in Table 2.3 have shown, stratifying the sample according to cognitive ability, age, gender, and education results only in very little variation in the share of respondents whose estimate is located in the pure base rate interval. For instance, this share is 21.0% among respondents with more than 10 years of schooling, whereas it is 20.6% among respondents with 10 years of schooling or less. Similarly, for all other measures the share of respondents in the pure base rate interval is always relatively close to 20%, indicating that these measures cannot explain what the determinants of giving an estimate close to the pure base rate are. Analogous to the previous analysis for base rate neglect, we performed probit regressions where the dependent variable is a dummy equal to one for answers in the pure base rate interval and equal to zero otherwise. In these regressions, none of the explanatory variables has a significant impact. The same is true for probit regressions with the dependent variable being equal to one if a respondent's answer is exactly 10% (results not reported).

2.3.3. Robustness Checks. In this subsection we perform two sets of robustness checks. First, in order to address the interplay of respondents' answers and their background characteristics in more detail, we estimate multinomial logit models with the same four answer categories we investigated earlier in the descriptive analysis.[6] These regressions allow us to test whether differences between the answer categories are statistically significant. Second, we estimate regressions in which the dependent variable is the linear distance between a respondent's answer and the Bayesian solution. Using this continuous measure enables us to analyze the determinants of biased probability judgment by taking into account whether deviations are either closer or further away from the correct solution. This might be relevant, e.g., in settings where estimates that are further away from the correct solution entail a higher cost for the decision maker.

In the multinomial regression analysis, the dependent variable indicates in which of the four categories a respondents' answer is located: Bayesian, pure base rate, base rate neglect, or "other". The base category in the regressions is "other". This allows us to

[6]All of our results are independent of the specific categorization we choose. For multinomial logit regressions with a more detailed grid of seven answer categories, see Appendix B.

Table 2.6: Robustness Check - Multinomial Analysis (I)

Dependent variable: =1 if "Other", =2 if BRN, =3 if PBR, =4 if Bayesian

	(1)			(2)		
	Base rate neglect	Pure base rate	Bayesian	Base rate neglect	Pure base rate	Bayesian
Word fluency measure	0.185**	-0.027	-0.108			
	[0.090]	[0.098]	[0.185]			
Perceptual speed measure				0.191*	0.004	-0.228
				[0.104]	[0.108]	[0.182]
Years of schooling	0.113**	0.080	-0.264**	0.128***	0.068	-0.257***
	[0.049]	[0.052]	[0.104]	[0.048]	[0.051]	[0.095]
1 if female	-0.258	0.172	0.019	-0.224	0.096	-0.005
	[0.175]	[0.187]	[0.335]	[0.170]	[0.181]	[0.311]
Age	-0.004	0.005	0.003	-0.000	0.007	0.002
	[0.005]	[0.005]	[0.009]	[0.006]	[0.006]	[0.010]
Constant	-1.480**	-1.972***	0.192	-1.831***	-1.901***	0.237
	[0.616]	[0.661]	[1.157]	[0.602]	[0.649]	[1.081]
N. Obs.		802			850	
Prob > χ^2		0.000			0.000	
Pseudo R^2		0.019			0.020	

Table 2.6: Robustness Check - Multinomial Analysis (I). Multinomial logit estimates, standard errors in parentheses. Dependent variable =1 if "Other" (answer either in the intervals (40%,60%], (60%,80%), (80%,100%], or "Don't know"), =2 if Base rate neglect (answer equal to 80%), =3 if Pure base rate (answer in [0%,20%]), and =4 if Bayesian (answer in (20%,40%]). The reference category is "Other". Cognitive ability measures are standardized to have mean 0 and standard deviation 1. Significance at the 1%, 5%, and 10% level is denoted by ***, ** and *, respectively.

address the factors that determine whether a respondent gives a Bayesian estimate or whether he displays a typical judgment bias. The estimation results support all the earlier findings from the descriptive analysis (see Table 2.6). For the specification in panel (1), the coefficient for years of schooling indicates that respondents with more education have a significantly higher probability of doing base rate neglect ($p < 0.05$) and a significantly lower probability of giving an answer that is located in the Bayesian interval ($p < 0.05$). Note that these results were obtained from a regression that included the knowledge-based measure of cognitive ability as a control variable. If we control for perceptual speed instead,

Table 2.7: Robustness Check - Multinomial Analysis (II)

Dependent variable: =1 if "Other", =2 if BRN, =3 if PBR, =4 if Bayesian

	(1)			(2)		
	Base rate neglect	Pure base rate	Bayesian	Base rate neglect	Pure base rate	Bayesian
Word fluency measure	0.181**	-0.055	-0.298*			
	[0.088]	[0.097]	[0.169]			
Perceptual speed measure				0.152	-0.033	-0.457***
				[0.099]	[0.104]	[0.158]
1 if university degree	0.924***	0.707**	-0.137	1.014***	0.696**	-0.282
	[0.261]	[0.285]	[0.630]	[0.250]	[0.275]	[0.625]
1 if female	-0.224	0.203	0.009	-0.207	0.135	-0.043
	[0.171]	[0.184]	[0.307]	[0.167]	[0.178]	[0.289]
Age	-0.008	0.002	-0.005	-0.005	0.004	-0.009
	[0.005]	[0.005]	[0.008]	[0.006]	[0.006]	[0.009]
Constant	-0.274	-1.128***	-1.951***	-0.424	-1.162***	-1.673***
	[0.254]	[0.279]	[0.444]	[0.280]	[0.302]	[0.454]
N. Obs.		843			894	
Prob > χ^2		0.000			0.000	
Pseudo R^2		0.019			0.021	

Table 2.7: Robustness Check - Multinomial Analysis (II). Multinomial logit estimates, standard errors in parentheses. Dependent variable =1 if "Other" (answer either in the intervals (40%,60%], (60%,80%), (80%,100%], or "Don't know"), =2 if Base rate neglect (answer equal to 80%), =3 if Pure base rate (answer in [0%,20%]), and =4 if Bayesian (answer in (20%,40%]). The reference category is "Other". Cognitive ability measures are standardized to have mean 0 and standard deviation 1. Significance at the 1%, 5%, and 10% level is denoted by ***, ** and *, respectively.

the impact of education becomes even stronger (see panel (2) of Table 2.6). Again, more years of schooling increase the probability of exhibiting base rate neglect ($p < 0.01$) and reduce the probability of giving a Bayesian answer ($p < 0.01$). In addition, higher cognitive ability makes it more likely to commit base rate neglect. Conditional on education and cognitive ability, the estimates for age and gender are not significant.

Next, we address our earlier finding that people with an academic background have a strong tendency to commit base rate neglect. We therefore conduct multinomial logit regressions similar to the ones in Table 2.6. The only difference in the specification is

Table 2.8: Robustness Check - Linear Distance

Dependent variable: distance from 30.8% (in %-points)

	(1)	(2)	(3)	(4)
Word fluency measure	2.032***	1.487**		
	[0.700]	[0.719]		
Perceptual speed measure			2.746***	1.826**
			[0.756]	[0.801]
1 if female	-0.546	-0.573	-0.143	-0.141
	[1.364]	[1.385]	[1.317]	[1.339]
Age	-0.101**	-0.109***	-0.060	-0.087*
	[0.039]	[0.042]	[0.043]	[0.047]
Years of Schooling		0.808**		0.929**
		[0.383]		[0.371]
Constant	40.393***	32.590***	38.388***	30.171***
	[2.046]	[4.787]	[2.219]	[4.658]
N. Obs.	750	715	803	765
Prob > F	0.000	0.000	0.000	0.000
Adjusted R^2	0.021	0.026	0.029	0.034

Table 2.8: Robustness Check - Linear Distance. OLS estimates, standard errors in brackets. Dependent variable is the linear distance (in %-points) between a respondent's answer and the Bayesian solution of 30.8%. By definition, all observations where a respondent answered "I don't know" are excluded from the following analysis. Cognitive ability measures are standardized to have mean 0 and standard deviation 1. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively. Respondents who answered "I don't know" are not included.

that we now replace the measure for years of schooling by a dummy equal to one if the respondent holds a university degree (see Table 2.7). The results confirm our earlier finding: academic education leads to a significant increase in both the probability to give a base rate neglect estimate ($p < 0.01$) and in giving the pure base rate estimate ($p < 0.05$).

In a final step, we present results from a series of OLS regressions where the dependent variable is the linear distance from the normative solution (in %-points). This approach allows us to take the cost of an error into account. In particular, one might assume that estimates that are further away from the correct solution entail a higher cost to the decision maker. Estimation results are shown in Table 2.8. We find that both the cognitive ability measure and years of schooling have positive coefficients and are highly significant in all specifications. This implies that, in the setup under investigation, higher cognitive ability

and more education lead respondents to estimates that are further away from the Bayesian solution. For instance, the estimates in column (4) indicate that a one standard deviation increase in the measure for perceptual speed translates into an increase in the respondent's bias by 1.8%-points ($p < 0.05$), whereas an additional year of schooling increases this bias by another 0.9%-points ($p < 0.05$). Notably, in all regressions the measure for perceptual speed has a stronger detrimental effect than the word fluency measure. Among the control variables, age has a significantly negative impact in most specifications, indicating that, on average, older people tend to be further away from the Bayesian solution.

2.3.4. Related Work. To put our results into perspective, it is useful to address the related work by Stanovich and West (1998) who analyze a sample of 211 psychology undergraduates from a medium-sized U.S. university. The authors try to investigate the relationship between respondents' cognitive ability and their assessment of probabilities. The study had subjects solve two conditional probability questions: the cab problem and a standard medical diagnosis task. As a proxy for cognitive ability, respondents were asked to report their test scores in the Scholastic Aptitude Test (SAT), a standardized test for college admissions in the United States.[7] The parameters of the cab problem are chosen such that the base rate is 15% and the additional info is correct with 80%, leading to a Bayesian posterior of 41%. In contrast to our study, subjects were not asked for a precise estimate but had to choose their answer from six alternative intervals: less than 10%, 10-30%, 30-50%, 50-70%, 70-90%, more than 90%. To report the findings, the authors construct three categories: [0%,30%) for answers that rely too heavily on the base rate (n=38), [30%,70%] for Bayesian answers (n=45) and (70%,100%] for answers that tend to neglect the base rate (n=128, see Table 4 of their paper).[8] The results show that the average SAT scores for the three answer categories are rather similar: 1096, 1129, and 1145, respectively. None of the SAT score differences is significant at conventional levels. In contrast, outcomes of the medical diagnosis task lead to significant results. Here, subjects who score higher on measures of cognitive ability are more likely to exhibit base rate neglect, which seems at first sight to be in line with the findings of our study. Unfortunately, the medical diagnosis task from which this claim is derived does not allow for a

[7]Additional proxies for cognitive ability were elicited through specifically designed problems from the psychology literature, e.g., the so-called Raven Matrices and the Nelson-Denny comprehension measure. For simplicity, we focus on the results that relate to the SAT test score measure. Results for the other measures of cognitive ability were very similar.

[8]Unfortunately, the intervals used in the data analysis are relatively large, and answers are not reported in more detail. For instance, for an appropriate comparison of their results to our findings it would have been interesting to know the share of subjects in the interval 30-50%, as these are closest to the Bayesian solution of 41%.

distinction between Bayesian answers and answers that are equal to the pure base rate: the parameters are chosen such that the base rate is 0.1% and the additional information is correct with 95%. The resulting Bayesian posterior is roughly 2%. Again, respondents could choose their answer from the same six intervals used in the cab problem. For the data analysis, the authors construct three categories: answers in [0%,10%) are interpreted as Bayesian (!) estimates (n=36), answers in [10%,90%] are defined as "intermediate" (n=57), and answers in (90%,100%] are said to neglect the base rate (n=118, see Table 5 of their paper). The results show that average SAT scores for the three categories are 1103, 1109, and 1153, indicating that test scores of subjects in the Bayesian interval and in the intermediate interval are very similar, whereas test scores in the base rate neglect interval are significantly higher ($p < .025$). However, a serious drawback of this analysis is that Bayesian answers ("2%") and answers that give only the pure base rate ("0.1%") are pooled together in the same category. Thus, it is not warranted to link the low average SAT score for this group exclusively to respondents who give a Bayesian estimate. It could as well be that average SAT scores are lower for people who answered with the pure base rate. In fact, in the cab problem mentioned above, the average SAT score is indeed lowest in the pure base rate category. Therefore, the findings of Stanovich and West (1998) should be taken with a grain of salt.

2.4. Conclusion

In this chapter, we have investigated people's ability for probabilistic judgment in a representative sample of the German population. In particular, we have used a task that has been specifically designed to test whether respondents are able to form conditional probability estimates. Our results have shown that only very few answers are in the vicinity of the Bayesian solution. Deviations from the normatively correct answer appear to be systematic and the two most frequently described biases in the literature, base rate neglect and pure base rate, feature prominently in the data. The analysis of the determinants of biased probability judgment has led to surprising results: highly educated respondents do not have a higher likelihood of giving a correct answer. Rather, they are significantly more likely to commit the mistake of base rate neglect. In a similar vein, measures for cognitive ability have a positive and highly significant impact on the probability that people commit base rate neglect.

An important implication of our findings is that the existing research on decision making with conditional probabilities does not translate one-to-one to the general population,

as previous studies have mostly been conducted with university students and with academic professionals. It has to be noted, however, that our results may not be independent of the task structure. In our task, base rate neglect is a bias that leads to estimates that are far away from the Bayesian solution. In contrast, in situations where the base rate neglect outcome is closer to the normatively correct solution, it might be a useful heuristic to neglect the base rate. Taking our results at face value one would therefore expect that highly educated respondents perform better than other respondents in settings where the base rate neglect estimate is close to the Bayesian solution.

CHAPTER 3

Fungibility, Labels, and Consumption: Experimental Evidence

3.1. Introduction

Fungibility of money is a central principle in economics. It implies that any unit of money is substitutable for another. In the analysis of consumer choice, for example, fungibility prescribes that consumption decisions are based exclusively on the consumer's total wealth—the composition of wealth is irrelevant (Modigliani and Brumberg 1954). Fungibility is assumed throughout most of economic theory. Some empirical findings, however, cast doubt on the generality of this concept. A number of studies have demonstrated that assuming a lack of fungibility explains several empirical phenomena that are hard to reconcile with standard economic assumptions. Odean (1998), for example, analyzes stock market behavior of individual investors and finds that they sell winning stocks too soon and keep losing stocks too long. This finding is consistent with the assumption that investors evaluate their stock holdings separately (i.e., treat them as non-fungible) and are loss averse in each stock holding. Similar examples come from the fields of asset pricing (Benartzi and Thaler 1995, Barberis, Huang, and Santos 2001), stock market participation (Barberis, Huang, and Thaler 2006), stock trading (Barberis and Huang 2001), and life-cycle saving (Shefrin and Thaler 1988). Yet, surprisingly little is known about the degree to which individual decision-making is in line with the notion of fungibility.

In this chapter, we experimentally test whether consumers treat different forms of wealth as fungible. We combine a laboratory experiment and a randomized vignette survey which both have the same general design: participants make a simple two-goods consumption decision with one good being subsidized in a particular way. Subjects have a cash budget and an additional lump-sum subsidy at their disposal. In the *Cash treatment*, the subsidy is given in cash. In the *Label treatment*, the subsidy is given as an in-kind benefit, i.e., it has to be spent on the subsidized good. The crucial feature of our design is that the amount of the subsidy is lower than the amount that is spent on the subsidized good in the optimum. Therefore, treatments merely differ in the label attached to the subsidy; a rational consumer should not be influenced by whether the subsidy is given

in cash or in kind. In contrast, a consumer who does not treat different income sources as fungible will spend the in-kind subsidy disproportionately on the subsidized good. Standard theory predicts consumption to be the same in both treatments. If we find that consumption of the subsidized good is significantly higher in the Label treatment, we can conclude that behavior is not consistent with the principle of fungibility.

The results of both studies show that many subjects do not act in line with fungibility. In the laboratory experiment, average consumption of the subsidized good is close to optimal in the Cash treatment, whereas it is significantly higher in the Label treatment. Moreover, our findings document a substantial impact of subjects' mathematical ability, as subjects with relatively strong mathematical skills tend to act in accordance with standard economic theory. By contrast, in the group of subjects with weaker mathematical skills, the treatment difference is large and significant.[1] This supports the view that the violation of fungibility occurs for cognitive reasons which relates our study to recent work by Frederick (2005), Benjamin, Brown, and Shapiro (2006), Casari, Ham, and Kagel (forthcoming), and Dohmen, Falk, Huffman, and Sunde (2007). These studies show that people with higher cognitive skills are more likely to behave in line with standard economic theory, whereas people with lower cognitive skills tend to act in accordance with theories of boundedly rational behavior.

We complement the laboratory analysis by conducting a vignette survey. The general structure of the survey is analogous to the structure of the experiment. The survey is designed such that the setting is as close as possible to participants' everyday decisions, and such that the choice set is very small. The results indicate that participants in the Label treatment consume significantly more of the subsidized good. Moreover, this effect is very pronounced among respondents with a relatively weak mathematical background. In sum, the survey demonstrates that the findings of the laboratory experiment continue to hold in a less abstract and very simple setting, confirming the external validity of our experimental results.

Taken together, this chapter shows that consumers do not always treat money as fungible, which has important implications for several areas of economic research. First, our findings lend support to field studies that explain behavior of stock market investors by assuming that investors are loss averse and do not treat different wealth components as fungible (e.g., Odean 1998, Benartzi and Thaler 1995). These studies assume that investors

[1] An alternative source of the treatment effect could be that subjects feel morally obliged to comply with the label on the subsidy and thus increase their consumption of the subsidized good. We try to measure the moral attitudes of our subjects and find that these do not influence whether fungibility is violated or not.

evaluate each component of their portfolio (e.g., stock vs. bond holdings) separately. We provide a direct test of this assumption. Second, our results are important for theories of life-cycle saving which usually rely on the assumption of fungibility. Since we find that even in a very simple setup fungibility does not hold for all subjects, our evidence supports savings models in which the assumption of fungibility is relaxed (see, e.g., Shefrin and Thaler 1988, Barberis and Huang 2001). Finally, the specific design of our study allows us to give a rationale for the observed behavior of benefits recipients. Assuming a lack of fungibility could explain why housing benefits have such a strong effect on market rents as shown by Susin (2002), Gibbons and Manning (2006), and Fack (2006). Taking our results at face value, even non-distortionary housing benefits can induce tenants to spend a higher share of their income on housing. If landlords are aware of this behavior, they can increase rents accordingly. We discuss this issue in more detail in Section 3.5.

Why should people treat money as non-fungible? Tversky and Kahneman (1981) suggest that decision makers often do not decide globally but rather evaluate parts of a decision separately. This phenomenon has been called "narrow framing" (Kahneman and Lovallo 1993) or "narrow bracketing" (Read, Loewenstein, and Rabin 1999). For the allocation of a budget coming from different sources, making separate decisions implies a violation of fungibility. Since assessing the decisions separately is cognitively less demanding, our finding that subjects with lower mathematical skills are more likely to ignore fungibility also points to narrow bracketing as a potential explanation for the treatment effect. Closely related to the consumption setup we analyze is the concept of "mental accounting" (Thaler 1985, 1999). Mental accounting proposes that consumers use a set of heuristics to deal with their day-to-day financial decisions. An important assumption is that consumers have mental budgets for different expenditure categories or for different investment categories, thereby constraining the fungibility of money. In this framework, a label can influence consumption choice if it determines to which mental budget the consumer assigns the benefit payment.[2]

Most empirical studies about fungibility rely on non-incentivized surveys (e.g., Heath and Soll 1996, O'Curry 1997, Prelec and Loewenstein 1998, White 2006). Only few papers investigate fungibility in incentivized laboratory experiments or in field settings. Gneezy and Potters (1997), Thaler, Tversky, Kahneman, and Schwartz (1997), and Rabin and

[2]Note that mental accounting can be an advantageous heuristic, e.g., when it helps the decision maker to overcome self-control problems. In the setup we analyze in this study, however, deciding according to mental accounting or narrow bracketing is a mistake that leads to a lower payoff.

Weizsäcker (2007) find in incentivized experiments that subjects evaluate subsequent gambles separately. Rockenbach (2004) examines investors' behavior in a binomial option-pricing task and finds that investors do not realize that the value of a call option in such a setup is independent of the probability with which the good state of the world occurs. This behavior is not in line with fungibility, but with investors evaluating safe and uncertain investments separately. Arkes, Joyner, Pezzo, Nash, Siegel-Jacobs, and Stone (1994) show that spending out of an unanticipated windfall gain is higher than out of an anticipated payment and Epley, Mak, and Idson (2006) find that framing such a windfall gain as "bonus" instead of "rebate" increases spending even further.

The remainder of the chapter is organized as follows: The design of the experiment is described in Section 3.2. Section 3.3 presents experimental results and analyzes explanations for our findings, whereas Section 3.4 reports design and results of the vignette survey. Section 3.5 discusses our findings and concludes.

3.2. Experimental Design

3.2.1. General Setup. The goal of this chapter is to experimentally test whether individual behavior is in line with fungibility. We examine this question in a simple two-goods consumption case where one good is subsidized in a particular way. Assume that a consumer has a cash budget of amount R and a subsidy of amount S at his disposal. In the *Cash treatment*, the subsidy is a lump-sum payment in cash. In the *Label treatment*, the subsidy is of the same amount, but it is given in form of an in-kind benefit, i.e., it has to be spent on the subsidized good. The crucial feature of our design is that the amount of the subsidy is lower than the amount that is spent on the subsidized good in the optimum. Thus, a rational consumer should not be influenced by whether the subsidy is given in cash or in kind. The only difference between treatments is that the in-kind subsidy has a label attached to it, whereas the cash subsidy has not.

Consider the indifference curve diagram in Figure 3.1, where the subsidized good (s) is on the horizontal axis and the other good (o) is on the vertical axis. For simplicity, the price of the subsidized good is normalized to $p_s = 1$. When the consumer has only R at his disposal, the optimal consumption bundle is A. In the Cash treatment, the budget constraint is shifted to the right (dashed line) and the optimal consumption bundle is B. In the Label treatment, the subsidy is paid in kind and the consumer faces a kinked budget constraint (solid line). However, the kink does not affect optimal decision making, as the amount of S is lower than the amount s^B spent on the subsidized good in the optimum.

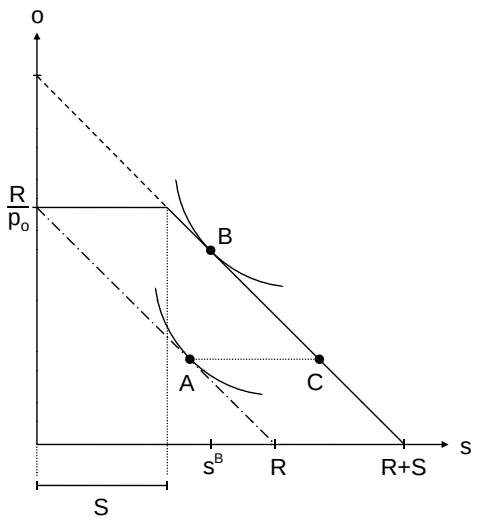


Figure 3.1: Consumption Decision with Non-distortionary In-kind Benefit. The subsidized good (s) is on the horizontal axis, the other good (o) is on the vertical axis. The dashed line is the budget constraint when the subsidy is given in cash. The solid line is the budget constraint when the subsidy is given in kind.

Thus, the subsidy is *non-distortionary* and the first-best choice B is feasible. Under the assumption that fungibility holds, consumption should be identical across treatments.

Now consider a consumer who does not treat different income sources as fungible. A violation of fungibility implies that the consumer has cognitive or mental sub-budgets between which money cannot be shifted.[3] In our setup, the consumer thus has a sub-budget for the subsidized good, and a sub-budget for the other good. We assume that a labeled payment is posted to the sub-budget for the corresponding good, whereas a cash payment is allocated optimally to the two sub-budgets. In the Cash treatment, we would still expect such a consumer to choose bundle B. As both income components are cash, the consumer can optimally allocate the funds to the sub-budgets for the two goods and thus choose the optimal consumption bundle. The difference to the standard model occurs in the Label treatment. Here, the consumer will still allocate the cash endowment optimally (bundle A). The subsidy, however, will be allocated to the sub-budget for the subsidized good. In the extreme case where sub-budgets are completely non-fungible, the entire subsidy will be spent on the subsidized good, resulting in consumption bundle C where $s^C = s^A + S$ (see Figure 3.1). If both goods are normal, $s^C \geq s^B$.[4] Therefore,

[3]A closely related concept that includes non-fungible sub-budgets as a building block is the concept of "mental accounting" (Thaler 1985).

[4]Note that this reasoning depends on the assumption that the consumer decides first about the cash budget and then about the subsidy payment. In contrast, if the consumer spent the subsidy first, he would be able to allocate the cash budget so as to reach bundle B. In the experiment, we are therefore testing the joint hypothesis of fungibility and order of spending.

if some subjects do not treat money as fungible, we expect average consumption of the subsidized good in the Label treatment to be higher than in the Cash treatment. This does not exclude the possibility that some subjects act in line with fungibility, or that others are only influenced to a certain extent by the label attached to the subsidy.

3.2.2. Laboratory Experiment. In the experiment, subjects made two subsequent consumption decisions. The first decision stage, which we call *reference stage*, serves to yield a reference transaction to which decisions in the second stage can be compared. The second stage, called *subsidy stage*, is our main treatment stage. In both consumption decisions, subjects allocated their budget on two goods. For each good we induced a standard microeconomic utility function by specifying monetary payoffs for the possible consumption levels. Subjects' total payoff was the sum of the payoffs for each of the two goods in both stages.

In the reference stage, subjects were endowed with a cash budget of 50 money units which they could allocate freely on the two goods. In the subsidy stage, subjects had again an endowment of 50 money units at their disposal, and additionally received a subsidy payment of 30 money units. The only difference between the two treatments is the form of this subsidy—the reference stage is identical in both treatments. In the *Cash treatment* (CT), the subsidy was given as an unconditional cash grant. The CT serves as our baseline treatment. In the *Label treatment* (LT), the subsidy was given as an in-kind benefit. Importantly, we chose the parameters such that the in-kind benefit is non-distortionary. By shifting the remainder of their budget appropriately, subjects could reach the same optimal consumption level as in the CT.

The exact specification of the payoff function is presented in Table 3.1. Subjects could buy up to 25 units of each good. Payoff is increasing in consumption, and marginal payoff is weakly decreasing. For clarity, we call the two goods *subsidized good* and *other good* throughout the chapter, although strictly speaking the subsidized good was only subsidized in the second stage of the Label treatment. Prices per unit were $p_s = 3$ for the subsidized good and $p_o = 2$ for the other good. Payoff function and prices remained constant over the course of the experiment. Budget left over in any stage could neither be saved nor did it yield any payoff. There was no time limit for decisions. For the chosen parameters, the consumption bundles (s, o) shown in Figure 3.1 are as follows: optimal consumption in the reference stage is $A = (12, 7)$, the optimal bundle in the subsidy stage is $B = (13, 20)$, and the consumption that corresponds to complete non-fungibility is $C = (22, 7)$.

Consumption (in units)	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Payoff subsidized good	0	36	70	102	132	160	186	210	232	252
Payoff other good	0	30	57	81	102	120	135	147	157	166

Consumption (in units)	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
Payoff subsidized good	270	286	299	310	316	322	328	333	338	343
Payoff other good	175	184	192	200	208	216	223	230	237	244

Consumption (in units)	**20**	**21**	**22**	**23**	**24**	**25**
Payoff subsidized good	347	351	355	358	361	364
Payoff other good	251	256	261	266	271	276

Table 3.1: Payoff Functions Used in the Laboratory Experiment. "Subsidized good" denotes the good that is subsidized in the subsidy stage of the Label treatment.

Subjects earned their initial endowment in a real-effort task. Before the start of the reference stage, subjects had to count the number of zeros in large spreadsheets that consisted of zeros and ones. When they managed to determine the correct number of zeros in a given amount of time they earned 100 money units that were later split in half for the two consumption decisions.[5] We chose this rather boring activity to minimize the intrinsic motivation subjects could have for the task and thus to strengthen their perception of really having earned the money (cf. Cherry, Frykblom, and Shogren 2002).

Subjects were students from the University of Bonn studying various majors except Economics. Treatments were assigned randomly and no subject participated in more than one treatment. At the beginning of the experiment, instructions were read aloud and subjects had to complete a number of example questions to ensure that they understood the task.[6] Detailed instructions for the two stages were given later on the computer screen. This allowed us to have subjects of both treatments in the same session and thus to align the delivery of the two treatments as much as possible. At the end of the experiment, subjects completed a questionnaire. The experiment was computerized using the software z-Tree (Fischbacher 2007). 92 subjects participated in the experiment, of whom one subject failed to complete the real-effort task. This leaves us with 45 observations in the Cash treatment and 46 observations in the Label treatment. Payoff points (cf. Table 3.1) were converted at a rate of 100 points = 1 Euro. In addition to their earnings from the consumption decisions, subjects received a show-up fee of 2.50 Euro. On average, subjects

[5]The precise rules were as follows: subjects worked on 8 large tables with 300 entries each. To complete the task, they had to count the correct number of zeros on four sheets within 15 minutes and submit this number to the computer. If subjects did not complete the task, they only received an endowment of 10 money units.

[6]For an English translation of the instructions, see Appendix C.

earned 12.20 Euro (∼14.80 USD at the time of the experiment). Sessions lasted between 60 and 70 minutes.

3.2.3. Vignette Survey. To test the external validity of our findings, we complement the laboratory analysis by conducting a vignette survey. The general structure is analogous to the structure of the experiment: participants decide twice, in a reference stage and a subsidy stage, and there are two treatments, a Cash treatment and a Label treatment. The consumption decision was presented through a scenario (a *vignette*) which was framed as the search for an apartment. Participants had to decide between two apartments, a small and cheap one and a large and expensive one. All parameters in the scenario were chosen to resemble participants' actual housing decisions in life.

The survey differs from the experiment in two dimensions: (i) we do not induce a utility function but rely on participants' actual housing preferences, and (ii) the choice set is very small, as participants had only two options to choose from. The survey is thus a robustness check whether results of the experiment continue to hold in a simpler and less abstract environment. A total of 525 students participated in the survey. None of the respondents participated in the laboratory experiment. More details on the survey design will be presented together with the results in Section 3.4.

3.3. Results

In this section we report results from the laboratory experiment. First, we show that giving a labeled subsidy instead of a cash grant increases consumption of the subsidized good. Then we present evidence that this effect is stronger for subjects with lower mathematical ability. Finally, we demonstrate that subjects' moral concerns cannot explain the treatment effect.

3.3.1. Consumption in the Experiment. Before we turn to the subsidy stage, we look at consumption decisions in the reference stage. Note that the design of the reference stage is exactly the same in both treatments. In particular, subjects are not aware of the fact that there will be two different treatments in the subsidy stage. Figure 3.2 shows a histogram of consumption choices for the (later to be) subsidized good. Choices in the Cash treatment are represented by grey bars, whereas choices in the Label treatment are represented by black bars. We find that choices are very similar: the modal choice in both treatments is the optimum of 12 units.[7] Average consumption is close to the

[7]For ease of exposition, we report only the consumption of the subsidized good. Consumption of the other good can then be readily calculated, as almost all subjects choose a consumption bundle that is on the Pareto frontier. Our results do not change when we confine the analysis to the Pareto optimal choices.

optimum, with 11.0 units in the Label treatment and 11.6 units in the Cash treatment.[8] These findings indicate that subjects have understood the experimental task and took the consumption decision seriously.

Next, we focus on outcomes in the subsidy stage. Our first result concerns the impact of the labeled subsidy on consumption choice.

> **Result 1:** *In the Label treatment, consumption of the subsidized good is significantly higher. The marginal propensity to consume out of the subsidy is twice as large as in the Cash treatment.*

A histogram of consumption choices is shown in Figure 3.3. Recall that the experiment is designed such that the same optimal consumption bundle can be reached in both treatments. If all subjects acted in line with fungibility, there should be no treatment difference. Results for the Cash treatment show that the modal choice is to consume the optimal amount of the subsidized good (13 units), and average consumption is 14.4 units. By contrast, we can see that the modal choice in the Label treatment is a consumption of 20 units, and only a relatively small share of subjects choose the payoff-maximizing amount of 13 units. Overall, subjects in the Label treatment buy too much of the subsidized good, consuming 16.7 units on average. The treatment difference is highly significant (U-test, $p = 0.006$).[9] In Table 3.2 we provide OLS estimates of the treatment effect. The estimates in columns (3) and (4) show that the label dummy is positive and highly significant. The coefficient has a value of about 2.3 consumption units ($p = 0.006$) and remains unaffected when controlling for age and gender.[10]

The two-stage design of our experiment enables us to compute marginal propensities to consume out of the subsidy payment by comparing decisions in the subsidy stage to decisions in the reference stage. A histogram of the intra-person change in consumption is shown in Figure 3.4. On average, the consumption increase in the Cash treatment is +2.8 units. In contrast, the average consumption change in the Label treatment is +5.7 units. This difference is highly significant ($p = 0.001$). As the subsidy payment has a value of 10 units of the subsidized good, the resulting marginal propensity to consume out of the subsidy is 57.4% in the Label treatment vs. 28.0% in the Cash treatment.

[8] An OLS regression indicates that the treatment difference is not statistically significant (see columns (1) and (2) of Table 3.2). A U-test between treatments is however significant at the 10%-level ($p = 0.066$). We are not worried about this marginal difference because consumption is slightly *lower* in the Label treatment. Thus, if there is any inertia in subjects' decisions, results from the reference stage work against a potential treatment effect in the subsidy stage.

[9] All significance levels in this section are obtained from two-sided Mann-Whitney U-tests.

[10] All regression results in Table 3.2 remain virtually unchanged if we perform tobit regressions instead of OLS, adjusting for the fact that subjects could not consume more than 25 units of each good.

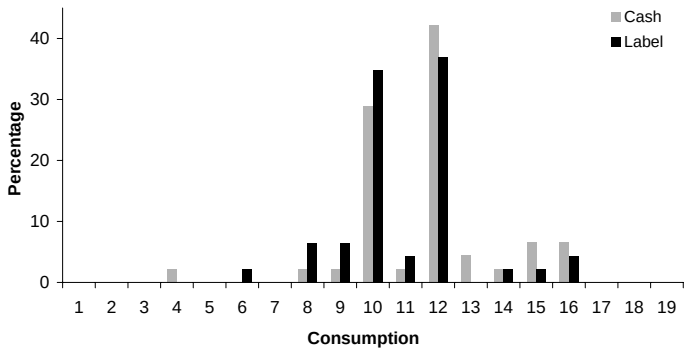


Figure 3.2: Consumption of the Subsidized Good in the Reference Stage. In this stage, the optimum is to buy 12 units of the subsidized good.

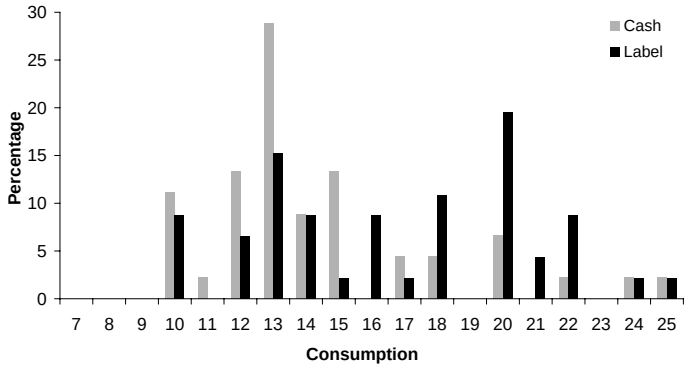


Figure 3.3: Consumption of the Subsidized Good in the Subsidy Stage. In this stage, the optimum is to buy 13 units of the subsidized good.

Our next result documents the considerable heterogeneity we observe in behavior across individuals.

Result 2: *A substantial part of the treatment difference is driven by subjects who increase their consumption by the full amount of the subsidy.*

As can be seen from Figure 3.4, the most frequent choice in the Cash treatment is a consumption increase by either 1 or 2 units. In contrast, the modal choice in the Label treatment is a consumption increase by 10 units. This decision implies that the entire subsidy is spent on the subsidized good, on top of the consumption from the reference stage. Subjects who treat the income sources as completely non-fungible will do exactly this (cf. bundle C in Figure 3.1). In the Label treatment, 10 out of 46 subjects spend the whole subsidy on the subsidized good, while this is true for only 1 out of 45 subjects in

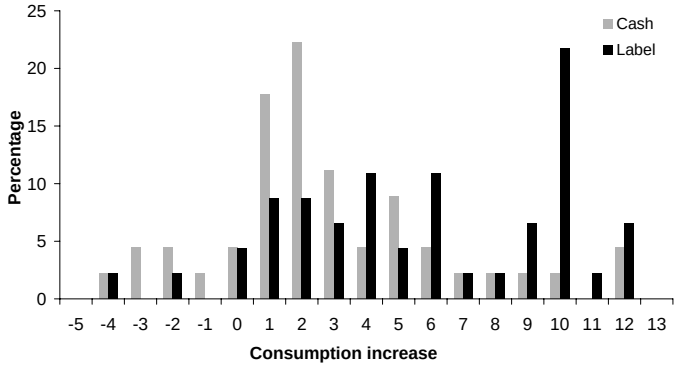


Figure 3.4: Consumption Increase of the Subsidized Good from Reference Stage to Subsidy Stage. The subsidy is worth 10 units of the subsidized good.

the Cash treatment. These subjects drive a substantial part of the treatment effect, but not all of it. If we exclude these 11 subjects from the analysis, the treatment difference in absolute consumption remains, but it becomes considerably smaller (1.4 units, previously 2.3 units; $p = 0.089$). The same is true for the treatment difference in consumption change (1.9 units, previously 2.9 units; $p = 0.002$). Interestingly, subjects who spend the entire subsidy on the subsidized good decide much faster than the remaining subjects. They need on average 107 sec for their decision, whereas the other subjects need 234 sec, more than twice as long. This difference suggests that spending the subsidy fully on the subsidized good is the result of a simple decision heuristic, rather than being derived from extensive deliberations.

3.3.2. Impact of Mathematical Ability. A consumer who does not treat different income components as fungible reduces the complexity of his consumption decision. Thus, ignoring fungibility can be seen as a form of "narrow bracketing" (Read, Loewenstein, and Rabin 1999).[11] Subjects with lower cognitive and mathematical skills will have a larger gain from reducing the complexity of the decision. We therefore expect these subjects to violate fungibility more often and, as a consequence, to be influenced more by the treatment manipulation. In fact, Read, Loewenstein, and Rabin (1999) conjecture that narrow bracketing is negatively correlated with cognitive ability. Our next result supports their intuition.

[11]Thaler (1985) argues that mental accounting, a concept similar to narrow bracketing, serves as a heuristic to overcome problems of limited self-control. In our experimental setup, however, limited self-control plays no role.

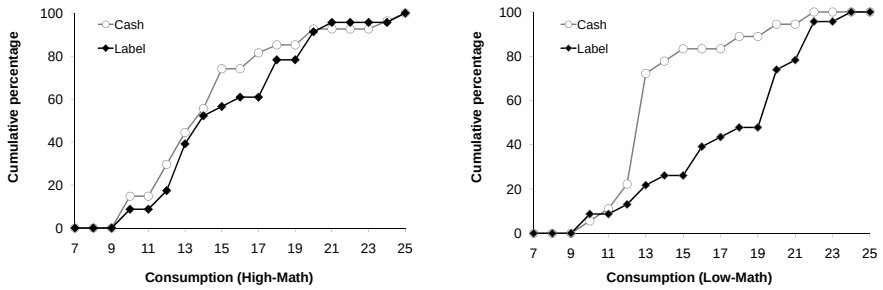


Figure 3.5: Cumulative Distribution of the Consumption of the Subsidized Good according to Mathematical Ability. Consumption decisions in the High-Math group (left panel) and in the Low-Math group (right panel). The black line depicts consumption in the Cash treatment, the grey line consumption in the Label treatment. In order to form a High-Math group ($n = 50$) and a Low-Math group ($n = 41$) we elicited the math grades subjects obtained in their high school completion exam and split the sample at the median.

Result 3: *The treatment difference in consumption is larger for subjects with lower mathematical ability.*

To analyze the interplay of cognitive ability and the treatment effect we use subjects' math grade in their high school completion exam as a proxy. The grades were elicited in the post-experimental questionnaire. In the German high school system, grades rank from 1 (excellent) to 5 (fail). We split the sample at the median, leading to a "High-Math" group ($n = 50$) of subjects who had a math grade of either 1 or 2, and a "Low-Math" group ($n = 41$) of subjects who had math grades of 3 or below.[12] Since it could be that mathematical ability impacts not only the decision in the Label treatment but also in the Cash treatment, we compare the treatment difference within the High-Math group to the difference within the Low-Math group.

Figure 3.5 presents cumulative percentages of consumption decisions in both treatments. In the High-Math group (left panel), cumulative distributions are very close to each other, suggesting that the label has no impact on consumption choices in this group. In contrast, in the Low-Math group (right panel), there is a wide gap between the two distributions, especially in the range between 13 units (the optimum) and 20 units. Statistical analysis shows that the treatment effect in the High-Math group is 0.9 units ($p = 0.357$). In contrast, the treatment effect in the Low-Math group is 3.8 units ($p = 0.004$). Regression analysis confirms that the difference between treatment effects is statistically significant (see columns (5) and (6) in Table 3.2). The dependent variable is the consumption of the

[12]One particularity of the German high school system is the existence of two types of math courses: intensive course and basic course. The results remain very similar when we take the type of math course into account.

Table 3.2: Consumption in the Laboratory Experiment

	Dependent Variable: consumption of subsidized good (in units)							
	Reference Stage		Subsidy Stage					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1 if Label Treatment	-0.666	-0.652	2.273***	2.270***	0.911	0.869	2.091*	2.088*
	[0.438]	[0.439]	[0.808]	[0.818]	[1.081]	[1.103]	[1.069]	[1.081]
1 if Low-Math Group					-0.796	-0.861		
					[1.160]	[1.191]		
Label * Low-Math					2.883*	2.960*		
					[1.615]	[1.653]		
1 if High-MO Group							-0.970	-0.989
							[1.157]	[1.186]
Label * High-MO							0.240	0.254
							[1.655]	[1.676]
Age		-0.077		0.021		0.047		-0.010
		[0.075]		[0.140]		[0.141]		[0.144]
1 if Female		-0.402		0.132		0.068		0.115
		[0.458]		[0.852]		[0.843]		[0.858]
Constant	11.622***	13.605***	14.422***	13.870***	14.741***	13.663***	14.875***	15.045***
	[0.311]	[1.769]	[0.574]	[3.293]	[0.733]	[3.259]	[0.790]	[3.511]
N. Obs.	91	91	91	91	91	91	91	91
Prob $> F$	0.132	0.270	0.006	0.057	0.010	0.046	0.036	0.132
Adjusted R^2	0.014	0.011	0.071	0.051	0.091	0.071	0.062	0.040

Table 3.2: Consumption in the Laboratory Experiment. OLS estimates, standard errors in parentheses. Dependent variable is the consumption of the subsidized good. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

subsidized good. As explanatory variables we include a treatment dummy (=1 if Label treatment), a Math dummy (=1 if Low-Math) and the interaction of the two. Thus, the reference category in these regressions is the group of High-Math subjects in the Cash treatment. The estimates show that the label dummy becomes small (0.9 units) and insignificant, indicating that subjects in the High-Math group do not exhibit a significant treatment effect. In contrast, the coefficient of the interaction term is large (2.9 units) and significant at the 10%-level. Controlling for age and gender leaves the significance level unaffected, and the coefficient increases slightly, to 3.0 units.

We obtain similar results when we ignore the Cash treatment and directly compare consumption of High-Math and Low-Math subjects in the Label treatment: subjects in the Low-Math group consume on average 2.1 units more of the subsidized good ($p = 0.066$). Recall that our earlier analysis has shown that subjects who spend the entire subsidy on the subsidized good account for a substantial share of the treatment effect. Among these subjects, the relation between mathematical ability and consumption behavior is extremely pronounced: 82% of them are in the Low-Math group, compared to 45% of subjects in the total sample. In sum, all our findings indicate that the treatment effect is significantly larger for subjects with relatively low mathematical ability.

3.3.3. Moral Obligation as an Alternative Explanation? So far we have attributed the treatment difference to cognitive limitations that prevent subjects from treating the two income components as fungible. However, one could also imagine that receiving a benefit payment causes a feeling of moral obligation to spend the money in accordance with the benefit giver's intention. In response to the intention that is conveyed by the label, recipients might increase their consumption of the subsidized good above the level they would have chosen if they had received the same amount as an unconditional cash payment. The next result shows that behavior in the experiment is not related to subjects' attitudes towards moral obligation.

> **Result 4:** *A feeling of moral obligation to comply with the label on the subsidy cannot explain the treatment difference in consumption.*

To measure attitudes towards moral obligation, we included a scenario in the post-experimental questionnaire in which subjects had to judge the behavior of a fictitious person. The scenario read as follows (translated from German):[13]

[13]The exact wording was: "Herr und Frau Müller haben zwei Kinder (5 und 8 Jahre) und verdienen zusammen pro Monat 2000 Euro netto. Sie bekommen zusätzlich pro Kind 180 Euro Kindergeld, also insgesamt 360 Euro im Monat. Sie geben normalerweise jeden Monat ca. 300 Euro für die Kinder aus

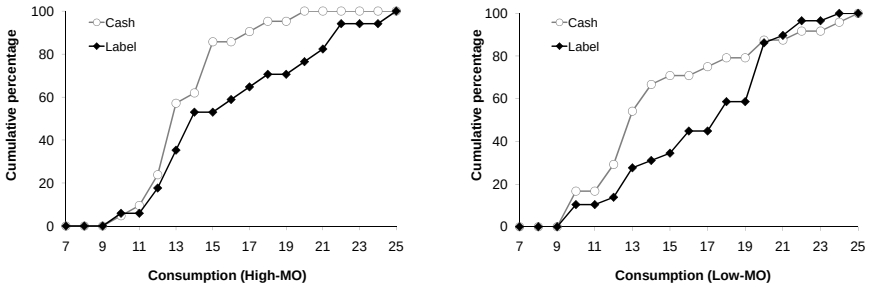


Figure 3.6: Cumulative Distribution of the Consumption of the Subsidized Good according to Feelings of Moral Obligation. Consumption decisions in the High-MO group (left panel) and in the Low-MO group (right panel). The black line depicts consumption in the Cash treatment, the grey line consumption in the Label treatment. To form a High-MO group ($n = 38$) and a Low-MO group ($n = 53$) we elicited subjects' moral concerns and split the sample at the median.

> *Mr and Mrs Miller have two children (5 and 8 years old). They earn a total amount of 2000 Euro per month, after taxes. Additionally, they receive 180 Euro child benefit per child, i.e., a total of 360 Euro per month. Usually, they spend about 300 Euro per month for their children (child clothing, toys, etc.). They spend the rest of the child benefit on other things (e.g., their own hobbies).*

Subjects had to indicate on a point scale from 1 to 6 how they judge the fictitious persons' behavior, 1 indicating "not justified at all" and 6 indicating "fully justified".[14] The decision situation described in the scenario above is very similar to the consumption decision in our experiment. In both situations, the intended use of the subsidy is obvious but the subsidy is not binding, i.e., a rational decision maker should not be influenced by the label attached to the subsidy.[15] Analogous to our analysis of mathematical ability, we split the sample at the median. This results in a group with strong moral obligation ("High-MO", $n = 38$) including all subjects who answered with a value of 3 or below, and a group with weak moral obligation ("Low-MO", $n = 53$) consisting of the remaining subjects.

Cumulative percentages for both groups are shown in Figure 3.6. The distributions look very similar: in both panels there is a wide gap between the Cash treatment and the Label treatment that starts at 13 units (the optimum) and extends until about 20 units are reached. The similarity of the two distributions indicates that the effect of the label

(für Kinderkleidung, Spielsachen etc.). Das restliche Kindergeld geben Sie für eigene Zwecke (Hobbys etc.) aus."

[14]The exact wording was: "Ich finde das Verhalten der Müllers: (1) nicht gerechtfertigt, ... , (6) völlig gerechtfertigt."

[15]The questionnaire included two other scenarios concerning (i) a person claiming student support provided by the state although not being entitled to it, and (ii) a person temporarily claiming unemployment benefits although having a new job already on the horizon. The results for all three scenarios are very similar. We therefore decided to focus on the child benefit scenario as it is closest to the decision in the experiment.

does not depend on subjects' moral concerns. The treatment effect is 2.3 units in the High-MO group ($p = 0.059$), whereas it is 2.1 units ($p = 0.028$) in the Low-MO group. Thus, the effect is essentially the same in both groups. A regression analysis confirms this result. Columns (7) and (8) of Table 3.2 include a dummy equal to one if a subject is in the High-MO group and an interaction term between the label dummy and the High-MO dummy. The reference category in these regressions is therefore the group of Low-MO subjects in the Cash treatment. We find that, in contrast to the regressions in columns (5) and (6), the coefficient on the label dummy remains large (2.1 units) and significant at the 10%-level. This indicates that subjects having relatively weaker moral obligation still exhibit a treatment effect that is very close to the overall treatment effect. Moreover, the interaction term is small (0.2 units) and far from being significant, showing that the treatment effect is basically the same in both groups.[16] Taken together, these results are not in line with the hypothesis that moral obligation drives the treatment effect in the experiment.

3.3.4. Robustness Check: Distance from Optimum. The analysis we conducted so far has concentrated on subjects' choices regarding the absolute consumption level of the subsidized good. To test the robustness of our findings, we now analyze an additional measure of performance: the distance from the optimal consumption level. This allows us to test whether the magnitude of the error subjects make depends on the form of the subsidy.

The regression specifications are completely analogous to our earlier analysis. The only difference is that the dependent variable is now the linear distance between subjects' consumption choice and the optimal consumption level. Results are presented in Table 3.3. Columns (1) and (2) show that in the reference stage distance to the optimal consumption is virtually the same in both treatments. Subjects in the Label treatment are about 0.1 units further away from the optimum, which is not significant. In contrast, results from the subsidy stage indicate that there is a large and significant treatment effect: consumers in the Label treatment are about 1.9 units further away from the optimal consumption level, see columns (3) and (4).

Concerning the relationship between math grades and the size of the treatment effect estimates are very similar to the results presented so far. The treatment effect for the High-Math group is small (0.4 units) and insignificant. In contrast, the coefficient on

[16]The answers to the two other vignettes are also not systematically related to the consumption of the subsidized good. The same holds true if we take the average answer of each subject to the three vignettes as our proxy for moral obligation.

Table 3.3: Distance from Optimum in the Laboratory Experiment

	Dependent Variable: distance from optimal consumption level (in units)							
	Reference Stage		Subsidy Stage					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1 if Label Treatment	0.120	0.117	1.903***	1.891***	0.422	0.336	1.530*	1.505*
	[0.334]	[0.337]	[0.667]	[0.670]	[0.882]	[0.890]	[0.873]	[0.945]
1 if Low-Math Group					-1.204	-1.351		
					[0.945]	[0.961]		
Label * Low-Math					3.204**	3.361**		
					[1.317]	[1.333]		
1 if High-MO Group							-1.458	-1.422
							[0.945]	[0.962]
Label * High-MO							0.627	0.697
							[1.352]	[1.361]
Age		-0.006		0.089		0.125		0.050
		[0.058]		[0.115]		[0.114]		[0.117]
1 if Female		-0.327		0.704		0.633		0.695
		[0.351]		[0.698]		[0.680]		[0.696]
Constant	1.489***	1.829	2.444***	-0.008	2.926***	-0.227	3.125***	1.545
	[0.238]	[1.356]	[0.474]	[2.698]	[0.598]	[2.628]	[0.646]	[2.850]
N. Obs.	91	91	91	91	91	91	91	91
Prob > F	0.721	0.802	0.005	0.027	0.003	0.008	0.013	0.039
Adjusted R^2	-0.010	-0.023	0.074	0.068	0.117	0.117	0.085	0.076

Table 3.3: Distance from Optimum in the Laboratory Experiment. OLS estimates, standard errors in parentheses. Dependent variable is the distance to the optimal consumption of the subsidized good. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

the interaction term is large (3.2 units) and significant at the 5%-level, see columns (5) and (6). Finally, we analyze the impact of subjects' moral concerns. The estimates are shown in columns (7) and (8). As before, we include the moral obligation dummy and its interaction with the label dummy in the regression. We find that the label dummy remains sizeable (1.5 units) and significant at the 10%-level. This indicates that the group of subjects with relatively weak moral obligation exhibits a substantial treatment effect. The interaction term is small (0.7 units) and insignificant, indicating that subjects with relatively high moral obligation exhibit a similar treatment effect. Thus, our earlier findings are confirmed: also for the distance from the optimum, the label has a significant impact on choices. Moreover, the treatment effect is significantly correlated with subjects' math grades, whereas there is no such correlation for subjects' attitudes towards moral obligation.

3.4. Vignette Survey

In this section we report results from the vignette survey which we conducted as a complement to the laboratory experiment. As in the experiment, we find a significant treatment effect. This effect is much stronger for subjects with a relatively weak mathematical background.

3.4.1. Design. The survey has the same general structure as the experiment. Participants decide twice, in a reference stage and in a subsidy stage, and there are two treatments, Cash treatment (CT) and Label treatment (LT). The survey presented participants with a scenario (a *vignette*), in which the consumption decision was framed as the search for a new apartment. In both stages, survey participants decided hypothetically which of two apartments to rent, either a small and cheap one (250 Euro per month) or a large and expensive one (410 Euro per month). In the reference stage, participants' only income source was a basic income of 900 Euro per month. In the subsidy stage, participants faced the same apartment choice again, now with their budget being increased by an additional payment of 200 Euro per month. While the reference stage was exactly the same in both treatments, the subsidy stage differed in the way the additional payment was labeled: in the CT, participants' initial budget was increased by a scholarship payment of 200 Euro, whereas in the LT, their budget was increased by a housing subsidy of 200 Euro. The crucial feature of the design is that the housing subsidy is non-distortionary, as

the amount of the payment is lower than the rent of either apartment. The participants had to indicate for each monthly budget which apartment they would choose.[17]

We constructed the figures in the vignette to be close to actual average budgets and rent expenditures of German students. According to a representative survey of the German student population conducted on behalf of the Federal Ministry of Education and Research, the average monthly income of a student is 823 Euro, and students who rent an apartment pay a monthly rent of 300 Euro on average (Isserstedt, Middendorff, Weber, Schnitzer, and Wolter 2004).[18] Treatments were assigned randomly and nobody participated in more than one treatment. Participants were students of the universities of Bonn and Cologne. Overall, we collected the choices of 525 participants, 264 in the Cash treatment and 261 in the Label treatment. None of the participants took part in the laboratory experiment. In addition to the apartment choices, we asked participants to indicate their age, gender, and field of study.

3.4.2. Consumption Decisions. Results in the reference stage were very similar across treatments. In the Cash treatment, 20.8% of participants chose the large apartment, whereas this share was 23.8% in the Label treatment ($p = 0.242$, Fisher Exact test).[19] This finding indicates that the randomization process worked well: there is no significant difference in the housing preferences of participants in either treatment. A probit regression confirms the result: in the reference stage, the treatment dummy is small (0.029) and far from being significant (see columns (1) and (2) in Table 3.5).

Now we focus on outcomes of the subsidy stage. The main research question is whether the form of the subsidy payment influences consumption decisions. If participants treat the two income sources as fungible, the share of participants choosing the large apartment should be identical across treatments. In contrast, if some participants violate fungibility, we expect average housing consumption to be higher in the Label treatment. Choices in the subsidy stage show that this is indeed the case.

Result 5: *In the vignette survey, participants in the Label treatment chose the large apartment significantly more often, compared to participants in the Cash treatment.*

[17]See Appendix C.2 for the precise wording of the questionnaire (translated from German).

[18]These figures refer to the average 26-year-old student in Germany. Among our survey participants, average age is 26.8 years. As the price level in the Cologne-Bonn region is relatively high compared to the rest of Germany, actual monthly income of students in our sample is likely to be higher than 823 Euro.

[19]Unless noted otherwise, all significance levels in this Section derive from two-sided Fisher Exact tests.

Table 3.4: GRE Test Scores (Quantitative Section)

Field of Study	Mean GRE Score	N. Obs.
Physics	744	35
Mathematical Sciences	733	6
Economics	706	26
Computer Sciences	704	51
Chemistry	696	46
Philosophy	636	5
Geology	630	11
Biology	615	51
Business	595	55
Geography	578	12
Foreign Languages	573	37
History	556	10
Medical Sciences	552	6
Sociology	545	11
Psychology	543	9
Law	539	85
Education	534	10
Communications	533	7

Table 3.4: GRE Test Scores (Quantitative Section). "Mean GRE Score" refers to the mean score in the quantitative section of the GRE that test takers in the years 2001–2004 achieved, according to field of study. "N. Obs." refers to the number of participants in our sample who are enrolled in the respective field of study. For ease of exposition, only fields of study with at least 5 observations are included in this table.

In the Cash treatment 51.5% of participants chose the large apartment, whereas in the Label treatment 62.8% of participants did so. The difference is highly significant ($p = 0.012$). Again, this finding is confirmed by a probit regression. Columns (3) and (4) in Table 3.5 report the results. We see that the label dummy is significant at the 5%-level, the marginal effect is estimated to be 11.3 %-points. The coefficient remains virtually unchanged when we control for age and gender.

Note that the estimates presented so far understate the effect of the label on consumption choice, since a substantial share of participants chose the large apartment already in the reference stage (approximately 22% of the sample). For these participants we cannot observe whether the treatment manipulation has an effect on choices, as they cannot increase their housing consumption further when having the larger budget in the subsidy stage. To get a more precise picture of the impact of the subsidy, we now look at the sample of participants who chose the small apartment in the reference stage ($n = 408$). In

Table 3.5: Consumption in the Vignette Survey

Dependent Variable: 1 if large apartment is chosen						
	Reference Stage		Subsidy Stage			
	(1)	(2)	(3)	(4)	(5)	(6)
1 if label treatment	0.029	0.030	0.113***	0.108**	-0.022	-0.020
	[0.036]	[0.037]	[0.043]	[0.043]	[0.077]	[0.077]
1 if low-math group					-0.096	-0.097
					[0.067]	[0.067]
Label * low-math					0.217**	0.212**
					[0.086]	[0.087]
Age		-0.002		0.001		0.002
		[0.004]		[0.005]		[0.006]
1 if female		0.021		0.045		0.013
		[0.039]		[0.045]		[0.049]
N. Obs.	525	523	525	523	498	497
Prob > χ^2	0.4212	0.7478	0.009	0.057	0.003	0.017
Pseudo R^2	0.001	0.002	0.010	0.011	0.020	0.020

Table 3.5: Consumption in the Vignette Survey. Probit estimation, marginal effects evaluated at the mean with standard errors in parentheses. Dependent variable is 1 if participant chooses the large apartment. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

this sample, 39.2% of participants in the CT switched to the large apartment in the subsidy stage. In contrast, 52.7% of participants in the LT switched to the large apartment, which implies that the label on the subsidy increased the probability to switch by 13.5 %-points. This marginal effect of 34.4% is highly significant ($p = 0.008$). The marginal propensity to consume out of the subsidy jumps from 31.4% in the CT to 42.2% in the LT.[20] These numbers are also interesting from a public policy perspective, as they suggest that the label on the housing subsidy leads to a substantial increase in the share of total income that recipients are willing to spend on housing.[21]

3.4.3. Impact of Mathematical Ability. In the experiment, the treatment effect was stronger among subjects with lower mathematical ability. The same is true for the survey.

Result 6: *In the vignette survey, the treatment effect is stronger for participants with a relatively weak mathematical background.*

[20]Switching from the small to the large apartment increases housing consumption by 160 Euro (410 Euro - 250 Euro), and the subsidy is worth 200 Euro.

[21]A precise estimate of this share is however difficult to obtain in our simple setup, as the housing decision in the survey is a choice from only two options. It is not clear how participants would decide if they were offered additional options where apartment size and apartment price are in the intermediate range between the two options in the survey.

To address whether the size of the treatment effect is correlated with mathematical ability, we take the major field of study as a proxy for participants' mathematical background. To rank majors accordingly, we use data on the performance of test takers in the Graduate Record Examination (GRE) Test.[22] In the quantitative section of the GRE, test takers have to answer a large number of mathematical questions in a short amount of time. As a proxy for the participants' mathematical background we use average quantitative GRE scores of all test takers in the years 2001-2004, according to major field of study. Table 3.4 lists majors present among our participants ranked by GRE score. In order to divide our sample into a High-Math group and a Low-Math group we perform a cluster analysis. The results of this analysis point to a GRE of 680 as the optimal break point. This cutoff leads to a High-Math group ($n = 167$) consisting of participants who study Physics, Mathematics, Economics, Computer Science, and Chemistry. In contrast, the Low-Math group ($n = 331$) consists of participants studying Philosophy, Geology, Biology, Business, Geography, Languages, History, Medical Science, Sociology, Psychology, Law, Education, and Communications.

Results for the High-Math group indicate that the share of participants who chose the large apartment is 58.1% in the Cash treatment, whereas it is 55.9% in the Label treatment ($p = 0.890$). These numbers suggest that participants in the High-Math group are not affected by the treatment manipulation. In contrast, in the Low-Math group, 48.3% of the participants chose the large apartment in the Cash treatment, whereas 68.4% did so in the Label treatment ($p = 0.001$). Probit estimates confirm that the treatment effect between the two groups is large and significant (see columns (5) and (6) of Table 3.5). When we include the Low-Math dummy and its interaction with the Label dummy in the regression, the Label dummy becomes small (-0.022) and insignificant. In contrast, the coefficient on the interaction term is large (21.7 %-points) and significant at the 5%-level. These results confirm that participants in the High-Math group do not exhibit a treatment effect, whereas the treatment effect in the Low-Math group is substantial.

In sum, results from the vignette survey are very similar to results from the laboratory experiment: we observe a sizeable labeling effect, indicating that many subjects do not act in line with fungibility. Moreover, this effect is particularly strong among subjects with relatively low mathematical skills. As the survey relied on participants' actual housing preferences, and as participants had only two options to choose from, these results indicate

[22]Prospective graduate students have to take the GRE test before applying to Ph.D. programs in the U.S., the UK, or Canada. For details concerning the GRE test, see www.gre.org. Data was obtained from ETS, the institution organizing the test.

that even in a very natural and relatively simple environment, many people do not act in line with fungibility.

3.5. Discussion and Conclusion

In this chapter we have conducted an incentivized laboratory experiment and a randomized vignette survey to test whether consumers treat different income sources as fungible. Experiment and survey yield the same results: even in our simple setup, many subjects do not act in line with fungibility. This effect is stronger for persons with lower mathematical skills. Differences in preferences, e.g, concerning the moral obligation to comply with the intention of the subsidy giver, do not drive our results.

An important application of our findings concerns the effect of housing benefits on market rents. Empirical studies for the U.S. and for France have shown that a rise in housing benefits for low-income tenants has lead to drastic rent increases (see, e.g., Susin 2002, Fack 2006).[23] The standard explanation for this phenomenon is that the subsidy causes an increase in housing demand which is met by an inelastic supply. In contrast, our findings suggest that this is only part of the story. Taking our results at face value, tenants who receive housing benefits are willing to spend a higher share of their budget on housing, compared to tenants who receive the same amount as a cash grant. If landlords anticipate this effect they are able to increase the rent accordingly. Laferrère and Le Blanc (2004) present evidence from France in support of this view: controlling for apartment and neighborhood characteristics, landlords discriminate between non-assisted tenants and tenants who receive housing assistance, charging the latter group significantly higher rents. As a result, housing benefits do not necessarily make the recipients better off, but may constitute a transfer payment from taxpayers to landlords. In our view, this problem can be mitigated by linking housing benefits less saliently to rent payments. For instance, the periodicity of benefit payments could be chosen such that it differs from the periodicity of rent payments. Moreover, one could design the benefit system such that the exact amount of the subsidy depends on variables which the landlord cannot observe.[24]

Due to the non-distortionary nature of the subsidy in our experiments, our findings are also related to the study of unconditional transfer payments. The recipients of child benefits, for instance, are not restricted in their use of these funds and only the name

[23]Similarly, Gibbons and Manning (2006) show for the U.K. that a reduction in housing benefits has lead to lower rents.

[24]Similarly, housing benefits should be paid directly to the tenant and not to the landlord. In the French housing benefit system, for example, the subsidy is often paid directly to the landlord. This might additionally affect the bargaining process by creating a "moral property right" of the landlord (see Gächter and Riedl 2005).

of the benefit payment marks it as a separate income component. Still, if recipients do not treat money as fungible, they will spend the subsidy disproportionately on child-related goods. Indeed, the effects we identify in the laboratory are also present in the field: Kooreman (2000) analyzes data from the Netherlands and shows that the marginal propensity to consume child clothing out of child benefits is much higher than it is out of other income.[25] Munro (2005) finds that the unconditional "winter fuel allowance" in the UK has a positive effect on heating expenditures. Taken together, our results demonstrate in a controlled environment that many consumers tend to comply with the label attached to a subsidy. Public policy can therefore influence consumption behavior in a simple way by explicitly stating the intended use of a transfer payment. However, whether the policy maker's goal to raise recipients' consumption of the subsidized good is reached depends crucially on the market environment at hand. Spending a higher share of one's budget on a good only translates into higher consumption if the price of the good does not increase by too much.

Our results do not imply that everybody in every situation will violate fungibility. In our study, treating money as non-fungible is linked to cognitive ability and not to preferences, suggesting that this behavior is a mistake.[26] Once the rational solution becomes obvious to subjects, e.g., by learning, they will probably regret their decision and choose the optimal solution. Moffitt (1989) studies the food stamps program in Puerto Rico and finds no difference between (non-distortionary) in-kind benefits and cash grants. People in his study had considerable experience with food stamps and were apparently able to learn the rational behavior. By contrast, most people make only a few housing decisions in their life; here, the scope for non-rational behavior is likely to be larger.

[25]However, Blow, Walker, and Zhu (2004) analyze data from the U.K. and find such a labeling effect only for some of their specifications.

[26]Note that our subject pool consisted exclusively of university students, so it is safe to assume that our subjects have higher cognitive skills than the population average. Since results were stronger for subjects with lower cognitive ability, we probably underestimate the true effect.

CHAPTER 4

Who Reacts to Monetary Incentives? The Role of Extrinsic and Intrinsic Motivation among Volunteer Workers

4.1. Introduction

Many people contribute to society's welfare without receiving any financial compensation for it: for instance, they give money to charity, they donate blood, or they volunteer their time and their work effort.[1] An area of particular interest in this context is the involvement of senior citizens in prosocial activities. Most OECD countries are facing the demographic phenomenon of aging societies. In order to mitigate the loss of human capital that ensues when elderly workers leave the workforce, public policy might want to encourage senior citizens to stay active as volunteers and transmit their skills and know-how to the younger generations. From a policy perspective, it is therefore important to understand how the level of public-spirited behavior is influenced by the environment in which people interact. In particular, a crucial question is whether monetary incentives are a valid instrument to foster prosocial behavior.

Economic research has offered contradicting views on the consequences of introducing monetary incentives into non-monetary environments. According to standard theory, introducing monetary rewards should lead to a (weakly) higher supply of the desired activity. In contrast, it has been suggested that there are circumstances under which material rewards have a detrimental effect on behavior. In a seminal paper, Titmuss (1970) argued that offering monetary compensation to blood donors might actually *reduce* the total supply of blood donated.[2] In a similar vein, a literature on the so called "motivation crowding-out" phenomenon assumes that the introduction of extrinsic incentives destroys intrinsic motivation (for a survey, see Frey and Jegen 2001). An influential paper by Bénabou and Tirole (2006) has provided a theoretical explanation for the potentially detrimental effect of monetary rewards by analyzing a model that combines people's concerns for their reputation with heterogeneity in individual responsiveness to extrinsic and intrinsic incentives.

[1] Freeman (1997) has pioneered the empirical analysis of volunteering behavior.

[2] For a critical comment on this hypothesis, see Arrow (1972).

In this chapter, we empirically analyze the effect of an introduction of monetary payments on the labor supply of volunteers. We use a unique data set that allows us to investigate directly the role of individual heterogeneity with respect to extrinsic and intrinsic motivation. The data set contains observations on more than 500 volunteers who work for a charitable organization in Germany. The volunteers are retired professionals who continue to work in their old job as unpaid consultants for small and medium-sized firms, often in developing countries. In the main survey question, respondents had to indicate their labor supply response if they were offered monetary rewards. In particular, they had to indicate how their willingness to work would change if they were offered payments of 100 Euro, 500 Euro, and 1000 Euro per week, respectively. The proxies for respondents' intrinsic and extrinsic motives are elicited via simple personality questions. To proxy for the degree to which participants are motivated by *intrinsic* factors, they were asked how important it is for them to "be there for others". This question captures the essence of public-spirited behavior, where a person cares not only about herself, but also about her contribution to others' well-being. In contrast, the role of *extrinsic* sources of motivation is proxied by asking respondents how important it is for them to "be able to afford things". With this question we can identify the degree of importance respondents attach to material goods. Note that, while our approach relies on hypothetical survey data, it still gives us the advantage that we can elicit behavior that would be extremely difficult to obtain with field data. In particular, we are able to construct a labor supply schedule for a large number of volunteers who are being offered financial payments. Our work is therefore a complement to empirical studies with field data.

Our study yields three main results. First, we find that a large majority of the respondents acts in line with standard economic theory: 84% of the respondents would react to the introduction of financial incentives with a (weak) increase in their labor supply. In contrast, the prevalence of crowding-out is rather low, as only a relatively small share of respondents would work less when financial incentives were introduced. Second, we show that the degree to which workers are motivated by extrinsic factors has a decisive impact on their reaction to the introduction of financial incentives. Our results indicate that volunteers who score high in the extrinsic proxy do indeed react to the introduction of financial incentives by working more than the other volunteers. Furthermore, the degree to which intrinsic factors matter to the volunteers has the opposite effect: intrinsically motivated people tend to show a weaker reaction to financial incentives than their counterparts. Thus, two simple personality measures help to predict how individuals are going

to adjust their work effort in response to financial incentives. The third result reveals that the differences in labor supply we find on the aggregate level are mostly driven by a crowding-in effect that affects the workers disproportionately. Volunteers who are motivated by extrinsic factors are significantly more likely to be crowded-in when they are offered money for their work. The effect is large and significant, with a unit increase in the extrinsic proxy corresponding to an increase in the probability of being crowded in by 6%-points, with a baseline probability of 23.8% at the medium payment level. Moreover, also the degree of intrinsic motivation affects the crowding-in of labor supply. Our estimates show that a unit increase in the intrinsic proxy implies a reduction in the probability of being crowded-in by 8%-points. Thus, we find strong evidence indicating that intrinsically motivated volunteers are more likely to keep their work load constant, even when being offered large monetary payments. When analyzing the crowding-out of labor supply, neither extrinsic nor intrinsic motives have a significant impact on the volunteers' decision. This is not surprising as only very few volunteers would reduce their labor supply due to the introduction of financial incentives.

Robustness checks show that our results continue to hold when using an alternative measure that addresses the concern that different respondents might use different scales when rating how important extrinsic and intrinsic factors are to them. We construct a very simple within-person measure: volunteers to whom being able to afford things is more important than being there for others are categorized as "extrinsic", volunteers for whom the opposite is true are categorized as "intrinsic", and volunteers to whom both factors are of equal importance are categorized as "neutral". This alternative measure therefore relies only on the assumption that a respondent uses the same scale when answering each of the two questions. Using these category dummies in labor supply regressions that are otherwise identical to our earlier analysis demonstrates that the "extrinsic" dummy is positive and significant, whereas the "intrinsic" dummy is negative and significant, with "neutral" being the reference category.

In sum, our study provides evidence that the effect of financial incentives on public-spirited behavior depends crucially on the behavioral motives of the volunteers. High degrees of intrinsic motivation correspond to a low responsiveness to financial incentives. In contrast, high degrees of extrinsic motivation are related to behavior in accordance with standard economic theory, where workers increase their labor supply when offered monetary rewards. Thus, in the setting we investigate, the introduction of financial incentives leads to a disproportionate crowding-in of extrinsically motivated workers.

The low overall prevalence of crowding-out in our sample is in line with Prendergast (1999), who reviews the empirical evidence on the effects of financial incentives in labor contexts. Regarding the idea of material incentives destroying intrinsic motivation he argues that "while this idea holds some intuitive appeal, it should be noted that there is little conclusive empirical evidence (particularly in workplace settings) of these influences".[3] In contrast, evidence on the negative effect of financial rewards is often derived from the context of charitable activities. Gneezy and Rustichini (2000) show in controlled experiments that people's work effort in a task where monetary rewards are absent is significantly higher than it is in exactly the same task with low, performance-contingent payments. In particular, the task used in some of their experiments was to collect money for a charitable organization. A recent blood donation experiment by Mellström and Johannesson (2005) has demonstrated a negative effect of financial rewards on the blood donation rate.

Our findings point to a fundamental difference between public-spirited deeds such as donating blood or collecting money on the one hand, and deeds such as the provision of volunteer labor on the other. Collecting money or donating blood could be considered rather dull or even painful activities that do not generate an immediate utility gain for the volunteer. Therefore, the utility derived from these activities lies probably in the reputational benefits associated with them. In contrast, for volunteers who work in activities that are inherently satisfactory, one may be hard pressed to argue that they decrease their involvement as soon as financial incentives are introduced. Thus, for activities such as volunteer work in a task that is both interesting and inherently satisfactory (e.g., volunteer work within the old job, as it is the case in our sample), the research focus should be rather on the question whether people keep their work load constant as opposed to working more.

In the related literature, several recent experiments have analyzed the interplay of financial incentives and reputational concerns. Ariely, Bracha, and Meier (2007) analyze both a lab experiment and a field experiment in which people have to perform a real effort task in order to earn money for a charity. In the experiment, subjects work either paid or unpaid and either in a public setting or in a private setting. The results show that, when working without rewards for a good cause (American Red Cross), subjects put forth significantly more effort in the public condition than in the private condition. Moreover, introducing rewards leads to a significant increase in effort in the private condition, whereas it has no effect in the public condition. Linardi and McConnell (2008) analyze a

[3]See also Lazear (2000) who provides extensive empirical evidence on the positive effects of monetary incentives in labor contexts.

very similar setup in which subjects work for a charitable organization under controlled laboratory conditions. Again, subjects work either paid or unpaid and either in a public setting or in a private setting. All effects are qualitatively similar to Ariely, Bracha, and Meier (2007), albeit on a lower magnitude. Closest to our study is a paper by Carpenter and Myers (2007) who investigate field data on the volunteer behavior of firefighters in Vermont. The data set includes proxies for individuals' degree of altruism and reputational concerns. The results show that the decision to volunteer is highly correlated with altruistic motives and with reputational concerns. Moreover, it is shown that the positive effect of financial rewards declines with peoples' reputational concerns, which is in line with the hypothesis that extrinsic incentives can crowd out prosocial behavior.

The remainder of the chapter is structured as follows: Section 4.2 gives an overview of the theoretical background that informs our empirical analysis, Section 4.3 describes the survey design. The empirical results are presented in Section 4.4, while Section 4.5 concludes.

4.2. Theoretical Background

In order to predict the effect of financial rewards on behavior, a number of theories might be relevant for the setting we investigate. According to standard economic theory, agents who decide whether to carry out a certain task face a simple tradeoff: they choose to carry out the task only if the benefits from it are higher than the cost. In a labor market setting this would imply that, e.g., workers accept a job only if the rewards are higher than the cost that is caused by the work effort. Similarly, it seems plausible that people who volunteer derive non-monetary utility from their volunteer work that makes it worthwile to carry out this activity. Therefore, standard theory offers a clear prediction regarding the effect of financial rewards: the introduction of financial payments for a certain activity should always have a (weakly) positive effect. If it was attractive to carry out an activity in the absence of rewards, it should become even more attractive to do so if a reward is introduced.

Early studies on the potentially detrimental effects of monetary payments emerged from the field of cognitive psychology. Researchers following Deci (1971) have found that the introduction and subsequent removal of material incentives can undermine intrinsic motivation for a task as soon as the reward is being removed.[4] The so-called cognitive evaluation theory (CET) predicts that tangible rewards undermine intrinsic motivation if

[4]A comprehensive meta-analysis reviewing 128 empirical studies is provided by Deci, Koestner, and Ryan (1999), another survey article is Ryan and Deci (2000).

they are made contingent on task performance. In economics it has been suggested that the detrimental effect of financial incentives is rooted in the preferences of individuals: some people may just dislike being paid for an activity that they were ready to carry out for intrinsic reasons in the first place (Frey and Jegen 2001). In this framework, two opposing effects operate as soon as financial incentives are introduced: on the one hand, a price effect makes the activity under study more attractive; on the other, intrinsic motivation suffers from the introduction of explicit incentives. The overall prediction for observed behavior is then ambiguous, as it depends on the relative strength of the two effects.

A recent paper by Bénabou and Tirole (2006) has suggested that the detrimental effect of monetary incentives might result from people's reputational concerns. In their model, agents decide whether to engage in a prosocial activity. Agents' utility function consists of three components: a valuation for behaving in a public-spirited way (intrinsic motivation), a valuation for material rewards (extrinsic motivation), and a concern for having a prosocial reputation. The reputational concerns are such that agents wish to be perceived as having a high valuation for public-spirited behavior and as having a low valuation for material rewards. The agents' types are private information, such that a signal-extraction environment arises: by observing the degree to which a person engages in the prosocial activity, the other agents can draw inferences about the person's "altruism" and "greed", respectively. In the equilibrium of this model, the introduction of monetary rewards creates doubts about the true motive for which people engage in public-spirited activities, which can lead to a crowding-out of prosocial behavior through extrinsic incentives.[5] While our data set is not rich enough to capture all the parameters that are relevant in the elaborate setting of Bénabou and Tirole (2006), our empirical analysis still accounts for the idea that volunteers are heterogenous with respect to their responsiveness to extrinsic and intrinsic incentives.

4.3. Survey Design

The data were collected via a survey among volunteers who work for the Senior Expert Service (SES), a non-profit organization in Germany. The SES sends retired professionals on assignments both in Germany and abroad. Typically, these "Senior Experts" share their

[5]Seabright (2004) and Ellingsen and Johannesson (2007) have proposed related models that address the interplay of financial incentives and altruistic motives, but neither of these studies takes the two-dimensional heterogeneity of people with respect to their degree of extrinsic and intrinsic motivation into account. Bénabou and Tirole (2003) have explored the impact of extrinsic vs. intrinsic motivation in a model with an informed principal, where the agent has to perform a task and where the principal has private information on the agent's ability for this task.

Table 4.1: Sample Statistics

Variable	Mean	Std. Dev.	Median	Min	Max	N. Obs.
age	65.489	4.477	66	45	77	583
female	0.073	0.261	0	0	1	587
log(hours)	2.441	0.741	2.708	1.609	3.555	579
log(income)	8.027	0.491	8.006	6.214	10.309	533
be able to afford things	2.696	0.674	3	1	4	570
be there for others	3.088	0.575	3	1	4	578

Table 4.1: Sample Statistics. The variable log(hours) is the natural logarithm of the average hours the respondent works per week when being on an assignment for the SES. The variable log(income) is the natural logarithm of the net household income per month. The variable being able to afford things is derived from a question that asked participants how important it is for them to be able to afford things. Answers ranged from 1 (=not important) to 4 (=very important). The variable be there for others is defined analogously.

professional experience with small and medium sized companies, mostly in the fields of industrial production and infrastructure, but also in trade and agriculture. The experts do not receive remuneration for their work. Travel and accommodation costs are borne either by the client or by the German Ministry for Economic Cooperation and Development, in case the client is unable to cover these costs. In 2005, Senior Experts were sent to a total of 1,477 assignments, most of them in Europe (42%), Asia (40%), and Latin America (9%). The average duration of assignments in developing and emerging countries was five weeks, whereas it was about three weeks for assignments in Europe.[6] The survey was mailed out in November 2004 to 750 Senior Experts who had at least one assignment in the time period between January and August 2004. We obtained 587 answers, which corresponds to a response rate of 78.3%. This response rate is very high, indicating that the volunteers take their work for SES seriously.

The main questions in the survey confronted respondents with an introduction of financial incentives to the non-monetary environment they are working in. Participants had to indicate how they would adjust their work effort for three different payment levels: 100 Euro per week, 500 Euro per week, and 1000 Euro per week. The question read: "How would your willingness to work for the SES change compared to the current situation if you were paid 100 Euro [500 Euro; 1000 Euro] per week of assignment?". Participants answered on a scale from 1 ("I would stop working for the SES") to 5 ("I would definitely want to work more").[7] These questions provide us with a direct measure of respondents'

[6]For more information, see SES (2005) and the SES webpage at www.ses-bonn.de.

[7]For the original wording of the questions (in German), see questions 43-45 in Appendix D.

reaction to the introduction of monetary incentives. In addition, the survey asked for socio-demographic background variables such as age, gender, and monthly household income.

To measure the impact of extrinsic and intrinsic factors on the volunteers' labor supply decision, the survey contained two personality-related questions. In order to proxy the degree to which respondents are motivated by extrinsic incentives, they were asked: "How important is it for you to be able to afford things?". In contrast, a proxy for the degree to which intrinsic motivation matters for the respondent was obtained through the question: "How important is it for you to be there for others?". Participants answered on a scale from 1 ("not important") to 4 ("very important").[8] The survey was constructed such that these proxies were obtained together with a total of ten personality-related questions about, e.g., the importance of owning a house, of having a happy marriage, and of having children. This design was chosen in order to minimize the probability that participants establish a connection between the two personality questions of interest and the question about adjusting their labor supply in face of monetary incentives.

An overview of the sample statistics is given in Table 4.1. The median age in the sample is 66 years. The share of women is 7.3%, reflecting the fact that professionals in the fields in which SES is active are predominantly male. The median of hours worked during an assignment is 15 hours per week. The median net household income of the volunteers in our sample is 3,000 Euro per month. Regarding the measure for extrinsic motivation, the average answer of respondents is 2.7, with a standard deviation of 0.674. The average answer for the intrinsic motivation measure is slightly higher, at 3.1, with a standard deviation of 0.575. Histograms of the two measures are shown in Figure 4.1. The upper graph shows the distribution of the extrinsic proxy. A share of 8.4% of the participants finds it very important to be able to afford things, and 56.5% find it important. In contrast, 31.4% indicate that being able to afford things is less important to them, and to 3.7% it is not important at all. The lower graph shows the distribution of the proxy for intrinsic motivation. A share of 20.4% of the volunteers finds it very important to be there for others, while 68.9% find it important. The share of participants finding it less important is 9.9% and only 0.9% find it not important at all.

4.4. Results

4.4.1. Labor Supply Decision. In a first step, we analyze volunteers' individual labor supply schedules (see Table 4.2). Most respondents (51.0%) keep their labor supply unchanged for all three payment levels. Moreover, 15.0% do not react to the lowest

[8]For the original wording of the questions (in German), see question 12 in Appendix D.

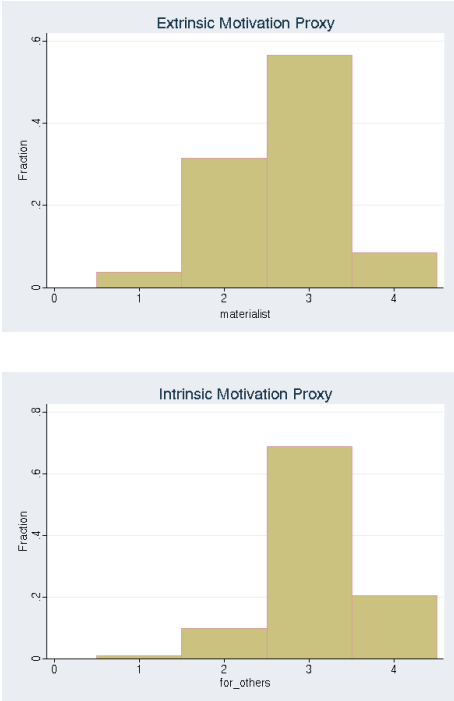


Figure 4.1: Proxies for Extrinsic and Intrinsic Motivation. Upper graph shows histogram of answers to the question "How important is it for you to be able to afford things?" Lower graph shows histogram of answers to the question "How important is it for you to be there for others?" Answers are coded as follows: 1="not important", 2="less important", 3="important", 4="very important".

payment but would work more when offered the medium or the highest payment, while 11.7% of the respondents would keep their labor supply constant for the low payment and the medium payment and would only work more if offered the high payment of 1000 Euro per week. Finally, 6.6% of the respondents would work more for all payment levels, even for the low payment of 100 Euro. Taken together, these findings suggest that for about 4 out of 5 respondents (84.3%) the introduction of financial payments leads to a (weakly) higher labor supply. In contrast, the occurrence of crowding-out due to the introduction of financial payments is much less frequent. The most common pattern is that crowding-out occurs at the medium and/or at the high payment level (5.9% and 3.5%, respectively). Only 1.4% of the respondents would reduce their labor supply for all payment levels. These findings are captured in our first result:

> **Result 1 (Payment Levels and Labor Supply):** *A large majority of the sample acts in line with standard economic theory: the introduction of financial payments leads for about 84% of the respondents to a (weakly) higher labor supply. In comparison, crowding-out of labor supply is observed only for a relatively small share of the respondents.*

Table 4.2: Labor Supply Schedule

	Weekly payment:			N. Obs.
	100 Euro	500 Euro	1000 Euro	
Labor supply:	unchanged	unchanged	unchanged	292 (51.0%)
	unchanged	work more	work more	86 (15.0%)
	unchanged	unchanged	work more	67 (11.7%)
	work more	work more	work more	38 (6.6%)
	unchanged	work less	work less	34 (5.9%)
	unchanged	unchanged	work less	20 (3.5%)
	work more	unchanged	unchanged	11 (1.9%)
	work less	work less	work less	8 (1.4%)
	unchanged	work more	unchanged	4 (0.7%)
	work more	work more	unchanged	4 (0.7%)
	unchanged	work more	work less	3 (0.5%)
	work less	unchanged	unchanged	2 (0.3%)

Table 4.2: Labor Supply Schedule. This table shows the within-person labor supply schedule for all three payment levels, sorted by the number of observations in each cell. For ease of exposition, "work less" consists of answers 1 (=stop working) and 2 (=tend to work less) and "work more" consists of answers 4 (=tend to work more) and 5 (=definitely work more). Schedules chosen only by a single respondent are not shown (4 observations in total).

A closer look at the data reveals that respondents' labor supply differs sharply according to whether the offered payment is either high or low. When faced with the low payment of 100 Euro per week, a large majority of the respondents (88.3%) would keep their labor supply unchanged, and 10.0% of the sample would work more (Figure 4.2). Evidence on a crowding-out effect at this payment level is very weak, as only 1.7% of the sample would work less than before. For higher payment levels, a quite different picture emerges: an increase in the offered payment to 500 Euro per week and to 1000 Euro per week leads to an increase in *both* the crowding-in and the crowding-out of work effort. The share of respondents who would work more than before increases to 23.8% (500 Euro) and 33.5% (1000 Euro), whereas the share of respondents who would work less than in the case of a zero wage increases to 7.4% (500 Euro) and to 11.6% (1000 Euro). The finding of crowding-out especially at high levels of payment is a puzzling phenomenon that is at odds with the existing theories of prosocial behavior. These theories typically predict crowding-out to be highest at *low* payment levels, where the loss of reputation is large and the wage payment is too low to compensate for this loss. Thus, it seems difficult to find an explanation for our findings in light of the theory. One might speculate that this result is based on the preferences of the volunteers in our sample: it could be that the volunteers who are crowded out have a specific notion of what constitutes an appropriate payment for their work. If these volunteers consider a payment of 100 Euro per week as appropriate,

Figure 4.2: Labor Supply in Response to Monetary Payments. Histograms of the labor supply in response to the introduction of a payment of 100 Euro, 500 Euro, and 1000 Euro per week, respectively. Answers are coded as follows: 1=stop working, 2=tend to work less, 3=unchanged, 4=tend to work more, 5= definitely work more.

whereas higher payments of 500 and 1000 Euro are considered to be inappropriately high, a pattern similar to the pattern in our data might arise.

Next, we address the interplay between the introduction of financial incentives and the volunteers' behavioral motives. We use regression analysis to assess whether the impact of extrinsic and intrinsic motivational factors on a person's labor supply decision is statistically significant, controlling for background characteristics. Results of an ordered probit regression are presented in Table 4.3. The dependent variable is the respondents' labor supply decision. Control variables are age, gender, current hours worked and household income. We estimate separate regressions for the payment levels of 100, 500, and 1000

Table 4.3: Labor Supply and Extrinsic vs. Intrinsic Motivation

	Dependent Variable: Labor Supply					
Payment:	100 Euro	500 Euro	1000 Euro	100 Euro	500 Euro	1000 Euro
	(1)	(2)	(3)	(4)	(5)	(6)
be able to	0.123	0.138*	0.151**	0.070	0.115	0.148**
afford things	[0.098]	[0.076]	[0.071]	[0.104]	[0.079]	[0.073]
be there	0.009	-0.168*	-0.147*	-0.088	-0.206**	-0.189**
for others	[0.112]	[0.087]	[0.083]	[0.121]	[0.091]	[0.086]
age	-0.019	-0.011	-0.000	-0.013	-0.005	0.010
	[0.014]	[0.011]	[0.010]	[0.016]	[0.012]	[0.011]
female	-0.580**	-0.297	-0.420**	-0.605*	-0.232	-0.319
	[0.283]	[0.206]	[0.194]	[0.325]	[0.223]	[0.208]
log(hours)	-0.046	0.010	-0.026	-0.100	0.011	-0.031
	[0.088]	[0.069]	[0.064]	[0.097]	[0.074]	[0.069]
log(income)				-0.516***	-0.142	0.043
				[0.155]	[0.111]	[0.104]
N. Obs.	550	546	546	502	498	498
Prob > χ^2	0.093	0.051	0.010	0.013	0.110	0.021
Pseudo R^2	0.018	0.011	0.011	0.035	0.011	0.012

Table 4.3: Labor Supply and Extrinsic vs. Intrinsic Motivation). Ordered probit estimates, standard errors in parentheses. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively. The dependent variable is the respondents' labor supply on a scale from 1 to 5, derived from a question that asked participants how their willingness to work would change if they were offered a payment of either 100 Euro, or 500 Euro, or 1000 Euro (per week). Answers are coded as follows: 1=stop working, 2=tend to work less, 3=unchanged, 4=tend to work more, 5= definitely work more. The variable be able to afford things is derived from a question that asked participants how important it is for them to be able to afford things. Answers ranged from 1 (=not important) to 4 (=very important). The variable be there for others is defined analogously. The variable log(hours) is the natural logarithm of the average hours the respondent works per week when being on an assignment for the SES. The variable log(income) is the natural logarithm of the net household income per month.

Euro. For each payment level, two different specifications are estimated, with and without a control for household income.

The estimation results show that the proxy for being extrinsically motivated has positive coefficients in all specifications. However, only 3 out of 6 coefficients are statistically significant at conventional levels. The weakest results obtain for the 100 Euro payment setting, where the coefficient is never significant, see columns (1) and (4). In contrast, the strongest results are found for the 1000 Euro payment setting, where the coefficient is significant at the 5%-level even if we control for household income, see column (6).

Turning to the proxy for intrinsic motivation, we find again that coefficients are small and insignificant in the 100 Euro setting. In contrast, for the 500 Euro and 1000 Euro settings, all 4 coefficients are negative and significant at the 5%-level and the 10%-level, respectively. Taken together, these findings lead to the following result:

> **Result 2 (The Role of Extrinsic and Intrinsic Motivation):** *Volunteers to whom extrinsic sources of motivation are important react to the introduction of financial incentives by working more than the other volunteers. In contrast, volunteers to whom intrinsic sources of motivation are important tend to offer less labor than the other volunteers. These findings are more pronounced the higher the amount of the financial incentives is.*

Note that in the 100 Euro payment setting, the control for household income is highly significant (column (4)). Thus, when offered 100 Euro per week, labor supply by richer people is lower than labor supply by poorer people. This finding is reassuring, as it is in line with the standard assumption of a diminishing marginal utility of money. In contrast, the income measure has no explanatory power for the higher payments of 500 and 1000 Euro. This indicates that respondents' reaction to relatively high levels of payment is independent of their financial background. The control variables age and current amount of hours worked have no impact on the reaction towards the introduction of financial incentives. The coefficient of the gender dummy is however negative and marginally significant in some of the specifications. This finding is in line with the empirical study by Mellström and Johannesson (2005): in their field experiment on blood donations, crowding-out of prosocial behavior was observed only among women.

4.4.2. Crowding-in vs. Crowding-out. The findings we have obtained so far could in principle come from different underlying sources. For instance, in line with the motivation crowding-out literature, one group of people might find it optimal to reduce their work effort when facing monetary incentives, whereas others do not (Frey and Jegen 2001). On the other hand, it might as well be that one group of people keeps their work effort constant, whereas others choose to work more. To discriminate between different explanations for the overall labor supply outcome, we investigate the effect of financial incentives and motivational measures separately according to whether we observe a crowding-in or a crowding-out of labor supply in the data.

The crowding-in effect is analyzed in Table 4.4. We estimate probit regressions with the dependent variable being equal to one if the respondent chooses to work more in face of financial incentives, and zero otherwise. The regressions include the proxies for intrinsic and extrinsic motivation and the same control variables as before. Again, we find that for

Table 4.4: Crowding-in and Extrinsic vs. Intrinsic Motivation

	Dependent Variable: Crowding-In (1= work more, 0=otherwise)					
Payment:	100 Euro (1)	500 Euro (2)	1000 Euro (3)	100 Euro (4)	500 Euro (5)	1000 Euro (6)
be able to afford things	0.010 [0.020]	0.063** [0.029]	0.069** [0.032]	0.005 [0.019]	0.060** [0.029]	0.069** [0.032]
be there for others	0.000 [0.022]	-0.074** [0.032]	-0.087** [0.036]	-0.012 [0.022]	-0.080** [0.033]	-0.098*** [0.038]
age	-0.005* [0.003]	-0.002 [0.004]	0.004 [0.005]	-0.003 [0.003]	-0.002 [0.004]	0.008 [0.005]
female	-0.077 [0.031]	-0.074 [0.070]	-0.102 [0.079]	-0.073 [0.030]	-0.041 [0.078]	-0.042 [0.091]
log(hours)	-0.007 [0.017]	0.018 [0.026]	-0.022 [0.028]	-0.009 [0.018]	0.031 [0.027]	-0.006 [0.030]
log(income)				-0.067** [0.029]	-0.019 [0.040]	0.064 [0.046]
N. Obs.	550	546	546	502	498	498
Prob > χ^2	0.184	0.038	0.010	0.173	0.075	0.012
Pseudo R^2	0.021	0.020	0.022	0.028	0.021	0.026

Table 4.4: Crowding-in and Extrinsic vs. Intrinsic Motivation. Probit estimates, marginal effects evaluated at the mean with standard errors in parentheses. Coefficients are marginal effects. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively. The dependent variable is a dummy indicating whether the respondent's willingness to work is higher when offered a payment of either 100 Euro, or 500 Euro, or 1000 Euro (per week). The remaining variables are defined as in Table 4.3.

the payment of 100 Euro the motivational proxies are not significant. In the 500 Euro and 1000 Euro regressions, a different picture emerges: here, both motivational proxies are highly significant. The extrinsic proxy has a positive coefficient, and the marginal effect indicates that a unit increase in the extrinsic proxy is equivalent to an increase in the probability of crowding-in to occur of between 6.0%-points and 6.9%-points. Recall that the baseline probability for crowding-in to occur at these payment levels is 23.8% (500 Euro) and 33.5% (1000 Euro). Next, we turn to the proxy for intrinsic motivation. The estimates are negative and significant in the 500 Euro and 1000 Euro regressions for all specifications. Thus, volunteers with high levels of intrinsic motivation are less likely to be crowded-in by financial incentives. The marginal effect indicates that for a unit increase

in the intrinsic proxy, the reduction in the probability of crowding-in is between 7.4%-points and 9.8%-points. Finally, we are going to address the crowding-out of labor supply due to the introduction of financial incentives. We find for all payment levels that the motivational proxies do not have any significant effects (regressions not reported). This is likely due to the fact that we have overall only very few observations of crowding-out in the sample, as described in Section 4.4.1. These findings are summarized in the next result:

> **Result 3 (Crowding-in vs. Crowding-out):** *The crowding-in of labor supply is higher among those volunteers who tend to be more extrinsically motivated than others. A unit increase in the extrinsic measure corresponds approximately to a 6%-points increase in the probability of crowding-in to occur. In contrast, volunteers who are more intrinsically motivated than others are less likely to be crowded-in. A unit increase in the measure for intrinsic motivation corresponds approximately to a 8%-points decrease in the probability of crowding-in to occur. Regarding the crowding-out of labor supply, neither the degree of extrinsic motivation nor the degree of intrinsic motivation has a significant impact on volunteers' decisions.*

Taken together, our findings have shown that the variation we found in the analysis of total labor supply is not due to disproportionate crowding-out, but due to a crowding-in effect that affects separate subsets of the volunteers in different ways. In particular, extrinsically motivated volunteers are significantly more likely to work more when offered financial payments. In contrast, intrinsically motivated volunteers are significantly more likely to keep their work load unchanged, even if they are offered substantial financial payments.

4.4.3. A Within-Person Measure of Motivation. The results presented in the previous subsection relied implicitly on the assumption that all respondents were using the same ordinal scale when answering the personality questions. In this section we relax this assumption by constructing a very simple within-person measure of extrinsic and intrinsic motivation which relies solely on the relative weight people give to being able to afford things compared to being there for others. This new measure sorts respondents into three separate categories: it takes the value "extrinsic" for people to whom the ability to afford things is more important than being there for others, the value "neutral" if the two proxies have the same level of importance for the respective person, and the value "intrinsic" if being there for others is more important than the ability to afford things. The only assumption needed is therefore that the scale is the same *within respondents* when answering each of the two questions.

Table 4.5: Labor Supply (Within-person Measure)

	Dependent Variable: Labor Supply					
Payment:	100 Euro	500 Euro	1000 Euro	100 Euro	500 Euro	1000 Euro
	(1)	(2)	(3)	(4)	(5)	(6)
1 if extrinsic	0.141	0.218	0.414***	0.110	0.239	0.445***
	[0.213]	[0.165]	[0.157]	[0.234]	[0.176]	[0.168]
1 if intrinsic	-0.027	-0.221**	-0.090	-0.060	-0.221**	-0.110
	[0.138]	[0.107]	[0.100]	[0.151]	[0.112]	[0.105]
age	-0.021	-0.012	-0.001	-0.013	-0.006	0.009
	[0.014]	[0.011]	[0.010]	[0.016]	[0.012]	[0.011]
female	-0.602**	-0.332	-0.473**	-0.639**	-0.272	-0.381*
	[0.280]	[0.204]	[0.192]	[0.322]	[0.220]	[0.205]
log(hours)	-0.052	0.018	-0.024	-0.103	0.021	-0.030
	[0.088]	[0.069]	[0.064]	[0.098]	[0.074]	[0.069]
log(income)				-0.519***	-0.142	0.038
				[0.154]	[0.111]	[0.104]
N. Obs.	550	546	546	502	498	498
Prob > χ^2	0.137	0.026	0.003	0.015	0.075	0.008
Pseudo R^2	0.016	0.012	0.013	0.034	0.012	0.014

Table 4.5: Labor Supply and Extrinsic vs. Intrinsic Motivation (Within-person). Ordered probit estimates, standard errors in parentheses. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively. The dependent variable is the respondents' labor supply, derived from a question that asked participants how their willingness to work would change if they were offered a payment of either 100 Euro, or 500 Euro, or 1000 Euro (per week). Answers are coded as follows: 1=stop working, 2=tend to work less, 3=unchanged, 4=tend to work more, 5=definitely work more. The variable extrinsic is a dummy equal to one if the respondent rated the importance of being able to afford things higher than the importance of being there for others. The variable intrinsic is a dummy equal to one if the respondent rated the importance of being able to afford things lower than the importance of being there for others. The omitted category is neutral, consisting of the respondents who rated both factors as equally important. The variable log(hours) is the natural logarithm of the average hours the respondent works per week when being on an assignment for the SES. The variable log(income) is the natural logarithm of the net household income per month.

According to this measure, in the total sample 63 respondents (11.1%) are classified as extrinsic, whereas 267 (47.1%) are neutral, and 237 (41.8%) are intrinsic. In the following we estimate ordered probit regressions with the dependent variable indicating the labor supply response. The control variables are the same as before. The only difference to our earlier analysis is that the motivational measures are now replaced by two dummies: "extrinsic" is equal to one if the respondent falls into the category extrinsic (defined above)

Table 4.6: Crowding-in (Within-person Measure)

	Dependent Variable: Crowding-In (1=work more, 0=otherwise)					
Payment:	100 Euro (1)	500 Euro (2)	1000 Euro (3)	100 Euro (4)	500 Euro (5)	1000 Euro (6)
1 if extrinsic	0.010	0.092	0.142**	0.011	0.116*	0.158**
	[0.043]	[0.066]	[0.072]	[0.045]	[0.071]	[0.076]
1 if intrinsic	-0.008	-0.089**	-0.074*	-0.014	-0.084**	-0.078*
	[0.027]	[0.038]	[0.043]	[0.027]	[0.040]	[0.045]
age	-0.005**	-0.003	0.003	-0.003	-0.002	0.007
	[0.003]	[0.004]	[0.005]	[0.003]	[0.004]	[0.005]
female	-0.077**	-0.087	-0.124*	-0.074**	-0.057	-0.069
	[0.030]	[0.066]	[0.075]	[0.029]	[0.074]	[0.087]
log(hours)	-0.007	0.021	-0.020	-0.008	0.033	-0.005
	[0.017]	[0.026]	[0.028]	[0.018]	[0.027]	[0.030]
log(income)				-0.067**	-0.019	0.063
				[0.028]	[0.040]	[0.045]
N. Obs.	550	546	546	502	498	498
Prob > χ^2	0.190	0.029	0.010	0.169	0.054	0.012
Pseudo R^2	0.021	0.021	0.022	0.028	0.023	0.026

Table 4.6: Crowding-In (Within-person). Probit estimates, marginal effects evaluated at the mean with standard errors in parentheses. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively. The dependent variable is a dummy indicating whether the respondent's willingness to work is higher when offered a payment of either 100 Euro, or 500 Euro, or 1000 Euro (per week). The remaining variables as well as the omitted category are defined as in Table 4.5.

and is zero otherwise. The dummy "intrinsic" is defined analogously. The base category is the group of "neutral" respondents, i.e., respondents to whom both extrinsic and intrinsic sources of motivation are of equal importance. The results of the labor supply regression are shown in Table 4.5. We see that the extrinsic dummy is positive and highly significant ($p < 0.01$) in the 1000 Euro regression, whereas it is not significant in the regressions for the lower payment levels. In contrast, the intrinsic dummy is negative and significant in the 500 Euro regressions ($p < 0.05$). Thus, the proxy that is obtained by using intra-person ratings of the relative importance given to extrinsic vs. intrinsic factors is only partially successful in explaining the variation in labor supply in our data.

As a second robustness check we investigate the occurrence of crowding-in analogously to our earlier analysis. To that aim we estimate a probit model with a crowding-in dummy

as the dependent variable. The results are shown in Table 4.6. The specification is analogous to the earlier specifications, with the only difference that the original proxies are again replaced by dummies that indicate which motivational factor is more important for the respondent. As before, the reference group are the participants to whom both factors are of equal importance. The results confirm our earlier findings. For almost all of the 500 Euro and 1000 Euro regressions, the extrinsic proxy is positive and significant, whereas the intrinsic proxy is negative and significant. Moreover, the definition of our within-person measure allows a straightforward interpretation: among people to whom the extrinsic factor is more important, a payment of 500 Euro increases the probability of crowding-in by 11.6%-points ($p < 0.10$), compared to the reference group. In contrast, among those persons to whom the intrinsic factor is more important, the probability of crowding-in is decreased by about 8.4%-points ($p < 0.05$). These effects are quite sizeable, as the occurrence of crowding-in is 25.0% in the reference group of people to whom both factors are of equal importance. For the 1000 Euro payment we find similar effects: the probability of crowding-in among people to whom the extrinsic factor is more important is increased by 15.8%-points ($p < 0.05$), whereas it is reduced by 7.8%-points ($p < 0.10$) among those who put greater weight on the intrinsic factor. As a comparison, the occurrence of crowding-in is 33.6% in the reference group. We therefore conclude that our results are by and large robust to using a within-person measure of motivation.

4.5. Conclusion

In this chapter we have examined the effect of monetary incentives on public-spirited behavior by analyzing a data set that contains the labor supply decisions of volunteer workers. We have shown that a large share of volunteers acts in line with standard economic theory, as the prevalence of crowding-out of labor supply due to financial incentives is relatively low. Furthermore, our results indicate that the impact of financial payments on labor supply depends crucially on the behavioral motives of the volunteers. In our setting, intrinsically motivated people are mostly unaffected by monetary incentives. They keep their work load unchanged, even in face of high monetary incentives. This is in contrast to people who are extrinsically motivated and who will respond even to low stake financial incentives by supplying more labor than before.

Our results suggest that the introduction of financial incentives leads to a change in the composition of persons who engage in prosocial behavior. In particular, the introduction of rewards may lead to a disproportionate crowding-in of extrinsically motivated

workers. From a public policy perspective it therefore seems important to decide whether the resulting sorting effects are desirable. Depending on the particular charitable activity, a policy maker might want to avoid having a high share of extrinsically motivated persons. If, on the other hand, the sorting effects are not problematic, our results have straightforward implications for the targeting of financial incentives to specific groups: we have shown that intrinsically motivated workers keep their work load unchanged, even if they are offered relatively large amounts of money. Thus, a cost-effective way to offer financial incentives would be achieved through a careful targeting towards more extrinsically motivated people.

APPENDIX A

Appendix to Chapter 1

TNS Infratest Sozialforschung
Landsberger Str. 338
80687 München
Tel.: 089 / 5600 - 1399

Persönlichkeit und Alltag

Unser Alltag wird von vielerlei Dingen bestimmt, über die die Wissenschaft wenig weiß. Zum Beispiel über unsere unterschiedlichen Temperamente, Interessen und Fähigkeiten. Die folgende Befragung ist deshalb etwas ungewöhnlich - aber für Sie persönlich auch besonders interessant. Neben einigen Fragen zu Ihrem Leben enthält das Interview auch kleine Tests. Am Ende können Sie bei einem Spiel auch Geld gewinnen. Mit etwas Glück bis zu 300 Euro. Auf jeden Fall erhalten Sie als Dankeschön ein Los der "Aktion Mensch"!

Listen-Nr. Lfd. Nr.

Bitte geben Sie hier Geschlecht, Geburtsjahr und Vorname der Zielperson ein:

Männlich ..... ☐
Weiblich ..... ☐

Geburtsjahr: 1 9 ☐ ☐

Vorname der Zielperson:

Bitte in Druckbuchstaben

Ihre Lebenssituation heute

1. Wie zufrieden sind Sie gegenwärtig mit den folgenden Bereichen Ihres Lebens?

☞ *Bitte kreuzen Sie für jeden Bereich auf der Skala einen Wert an:*
*Wenn Sie ganz und gar **unzufrieden** sind, den Wert "0",*
*wenn Sie ganz und gar **zufrieden** sind, den Wert "10".*
*Wenn Sie **teils zufrieden / teils unzufrieden** sind, einen Wert dazwischen.*

Wie zufrieden sind Sie ...	ganz und gar ***unzufrieden***										ganz und gar ***zufrieden***
	0	1	2	3	4	5	6	7	8	9	10
– mit Ihrer Gesundheit?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
(falls Sie erwerbstätig sind) – mit Ihrer Arbeit?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
(falls Sie im Haushalt tätig sind) – mit Ihrer Tätigkeit im Haushalt?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
– mit dem Einkommen Ihres Haushalts?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
– mit Ihrem persönlichen Einkommen?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
– mit Ihrer Wohnung?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
– mit Ihrer Freizeit?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
(falls Sie Kinder im Vorschulalter haben) – mit den vorhandenen Möglichkeiten der Kinderbetreuung?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
– mit Ihrem Lebensstandard insgesamt?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
– mit Ihrem Freundes- und Bekanntenkreis?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
– mit der Sicherheit in Ihrer Wohngegend?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

2. In welcher Stimmung sind Sie im Moment?

☞ *Antworten Sie bitte anhand der folgenden Skala, bei der "0"* ganz und gar ***schlecht,*** *"10"* ganz und gar ***gut*** *bedeutet.*

Meine Stimmung ist im Moment ...

☐–☐–☐–☐–☐–☐–☐–☐–☐–☐–☐

0 1 2 3 4 5 6 7 8 9 10

ganz und gar ***schlecht*** — ganz und gar ***gut***

3. Haben Sie die deutsche Staatsangehörigkeit?

Ja ..... ☐ Nein ..... ☐

4. Wo haben Sie vor der deutschen Wiedervereinigung, also vor 1989, gewohnt?

In der DDR (einschließlich Berlin-Ost) ..... ☐
In der Bundesrepublik (einschließlich Berlin-West) ..... ☐
In einem anderen Land ..... ☐

5. Haben Sie während der letzten 7 Tage irgendeine bezahlte Arbeit ausgeübt, auch wenn dies nur für eine Stunde oder für wenige Stunden war?

☞ *Bitte antworten Sie auch mit "Ja", wenn Sie in den letzten sieben Tagen normalerweise gearbeitet hätten, aber wegen Urlaub, Krankheit, schlechtem Wetter oder anderen Gründen zeitweise abwesend waren.*

Ja ..... ☐ Nein ..... ☐

6. Sind Sie zur Zeit im Mutterschutz oder in der gesetzlich geregelten "Elternzeit"?

Ja, Mutterschutz ..... ☐
Ja, Elternzeit ..... ☐ Nein ..... ☐

7. Sind Sie zur Zeit beim Arbeitsamt arbeitslos gemeldet?

Ja ..... ☐ ⇩ Nein ..... ☐ ➡ *Sie springen auf Frage **8!***

7a. Sind Sie bereits 12 Monate und länger arbeitslos?

Ja ..... ☐ Nein ..... ☐

8. Welchen Schulabschluss haben Sie?

Noch keinen Abschluss, gehe noch zur Schule ..... ☐

Schulabschluss in der Bundesrepublik Deutschland/ Westberlin:

Volksschul- / Hauptschulabschluss ..... ☐
Mittlere Reife, Realschulabschluss ..... ☐
Fachhochschulreife (Abschluss Fachoberschule) ..... ☐
Abitur (Hochschulreife) ..... ☐
Anderer Schulabschluss ..... ☐
Schule ohne Abschluss verlassen ..... ☐

In welchem Bundesland haben Sie Ihren Schulabschluss gemacht? ➔ ☐

Schulabschluss in der DDR:

Abschluss 8. Klasse ..... ☐
Abschluss 10. Klasse ..... ☐
Abitur (Hochschulreife) ..... ☐
Anderer Schulabschluss ..... ☐
Schule ohne Abschluss verlassen ..... ☐

Schulabschluss in einem anderen Land:

Dauer des Schulbesuchs ☐☐ Jahre,
und zwar:
Pflichtschule ohne Abschluss besucht ..... ☐
Abschluss einer Pflichtschule ..... ☐
Abschluss einer weiterführenden Schule ..... ☐

9. Können Sie sich an Ihr letztes Schulzeugnis erinnern? Welche Note hatten Sie in Ihrem letzten Zeugnis in den folgenden drei Fächern?

	Deutsch	Mathematik	Erste Fremdsprache
Sehr gut	☐	☐	☐
Gut	☐	☐	☐
Befriedigend	☐	☐	☐
Ausreichend	☐	☐	☐
Mangelhaft	☐	☐	☐
Ungenügend	☐	☐	☐
Hatte dieses Fach nicht	☐	☐	☐

10. Haben Sie eine Berufsausbildung oder ein Studium abgeschlossen?

Ja ..... ☐ ⇩ Nein ..... ☐ ➡ *Sie springen auf Frage **12!***

11. Was für ein Ausbildungs- oder Studienabschluss ist das?

☞ *Mehrfachnennungen möglich!*

- Gewerbliche oder landwirtschaftliche Lehre ..... ☐
- Kaufmännische oder sonstige Lehre ..... ☐
- Berufsfachschule, Handelsschule ..... ☐
- Schule des Gesundheitswesens ..... ☐
- Fachschule (z.B. Meister-, Technikerschule) ..... ☐
- Beamtenausbildung ..... ☐
- Fachhochschule, Ingenieurschule ..... ☐
- Universität, Hochschule ..... ☐
- Sonstiger Ausbildungsabschluss ..... ☐ ⇨ ☐

Bitte eintragen

12. In der folgenden Frage geht es darum, zu welchen Personen, Gruppen oder Institutionen Sie mehr oder weniger Vertrauen haben.

Wieviel Vertrauen haben Sie ...

☞ *In jede Zeile bitte **eine** Nennung!*

	Sehr viel Vertrauen	Ziemlich viel Vertrauen	Wenig Vertrauen	Überhaupt kein Vertrauen
– zur eigenen Familie	☐	☐	☐	☐
– zu Nachbarn	☐	☐	☐	☐
– zu Freunden	☐	☐	☐	☐
– zu Arbeitskollegen	☐	☐	☐	☐
– zu Fremden, denen Sie erstmals begegnen	☐	☐	☐	☐
– zu den Kirchen	☐	☐	☐	☐
– zu den Schulen und dem Bildungswesen	☐	☐	☐	☐
– zum Zeitungswesen	☐	☐	☐	☐
– zu den Gewerkschaften	☐	☐	☐	☐
– zur Polizei	☐	☐	☐	☐
– zum Bundestag	☐	☐	☐	☐
– zu den Behörden	☐	☐	☐	☐
– zur Europäischen Union	☐	☐	☐	☐
– zu den Gerichten	☐	☐	☐	☐
– zu großen Wirtschaftsunternehmen	☐	☐	☐	☐

13. Haben Sie jemals spontan von der Großzügigkeit einer Person profitiert, die Sie vorher nicht gekannt haben?

Ja ..... ☐ Nein ..... ☐

14. Wie oft kommt es vor, ...

☞ *In jede Zeile bitte **eine** Nennung!*

	Sehr oft	Oft	Manch-mal	Selten	Nie
– dass Sie Ihren Freunden persönliche Dinge leihen (z.B. CDs, Bücher, Auto, Fahrrad)?	☐	☐	☐	☐	☐
– dass Sie Ihren Freunden Geld leihen?	☐	☐	☐	☐	☐
– dass Sie die Tür zu Ihrer Wohnung unversperrt lassen?	☐	☐	☐	☐	☐
– dass Sie Ihr Wechselgeld beim Bezahlen nachzählen?	☐	☐	☐	☐	☐

15. Was würden Sie sagen: Wie viele enge Freunde haben Sie?

☐ enge Freunde

16. Glauben Sie, dass die meisten Leute ...

- Sie ausnützen würden, falls sie eine Möglichkeit dazu hätten ..... ☐
- oder versuchen würden, Ihnen gegenüber fair zu sein? ..... ☐

17. Würden Sie sagen, dass die Leute die meiste Zeit ...

- versuchen, hilfsbereit zu sein? ..... ☐
- oder nur ihre eigenen Interessen verfolgen? ..... ☐

18. Wie sehr stimmen Sie den folgenden Aussagen zu?

☞ *Kreuzen Sie bitte jeweils ein Kästchen an.*

	Stimme voll zu	Stimme eher zu	Lehne eher ab	Lehne voll ab
Im Allgemeinen kann man den Menschen vertrauen	☐	☐	☐	☐
Heutzutage kann man sich auf niemanden mehr verlassen	☐	☐	☐	☐
Wenn man mit Fremden zu tun hat, ist es besser, vorsichtig zu sein, bevor man ihnen vertraut	☐	☐	☐	☐
Wenn ich an die Zukunft denke, bin ich eigentlich zuversichtlich	☐	☐	☐	☐
Ich fühle mich oft einsam	☐	☐	☐	☐
Meine Arbeit macht mir eigentlich keine Freude	☐	☐	☐	☐
Die Verhältnisse sind so kompliziert geworden, dass ich mich ... fast nicht mehr zurecht finde	☐	☐	☐	☐

19. Wie schätzen Sie sich persönlich ein:
Sind Sie im allgemeinen ein risikobereiter Mensch oder versuchen Sie, Risiken zu vermeiden?

☞ *Bitte kreuzen Sie ein Kästchen auf der Skala an, wobei der Wert **0** bedeutet: "**gar nicht risikobereit**" und der Wert **10**: "**sehr risikobereit**". Mit den Werten dazwischen können Sie Ihre Einschätzung abstufen.*

Gar nicht risikobereit										Sehr risikobereit
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10

20. Stellen Sie sich vor, Sie nehmen an einer Quizshow teil, bei der Sie auch Geld gewinnen können. Dabei entscheiden Sie selbst in welcher Form: Sie können entweder einen festen Geldbetrag bekommen oder an einer Lotterie teilnehmen. Bei der Lotterie gewinnen Sie mit einer Chance von 50:50 entweder 200 Euro oder Sie bekommen nichts.

Bitte sehen Sie sich einmal Liste 20 an. Dort steht in jeder Zeile, wie Sie sich entscheiden können. Treffen Sie Ihre Entscheidung bitte zunächst in Zeile 1, dann in Zeile 2 usw.

Int: Liste 20 bitte mit ZP zusammen ansehen und bei Bedarf erläutern.

Wie würden Sie sich entscheiden ...

	A		B
in Zeile 1?	Fester Betrag 40 Euro ..... ☐	oder	Lotterielos ..... ☐
in Zeile 2?	Fester Betrag 70 Euro ..... ☐	oder	Lotterielos ..... ☐
in Zeile 3?	Fester Betrag 100 Euro ..... ☐	oder	Lotterielos ..... ☐
in Zeile 4?	Fester Betrag 130 Euro ..... ☐	oder	Lotterielos ..... ☐

ZP kann sich nicht entscheiden und verweigert die Antwort ..... ☐ ➡ *Sie springen auf Frage **22!***

21. Nun stellen Sie sich vor, in der Quizshow werden die Regeln etwas verändert. Wieder haben sie vorab die Wahl zwischen einem festen Geldbetrag und einer Lotterie. Bei der Lotterie können Sie mit einer Gewinnchance von 50:50 entweder 160 Euro gewinnen oder diesmal 40 Euro verlieren.

Bitte schauen Sie nun auf die Liste 21, dort steht in jeder Zeile wie Sie sich entscheiden können. Treffen Sie Ihre Entscheidung bitte zunächst in Zeile 1, dann in Zeile 2 usw.

Int: Liste 21 bitte mit ZP zusammen ansehen und bei Bedarf erläutern.

Wie würden Sie sich entscheiden ...

	A		B
in Zeile 1?	Fester Betrag 0 Euro ..... ☐	oder	Lotterielos ..... ☐
in Zeile 2?	Fester Betrag 30 Euro ..... ☐	oder	Lotterielos ..... ☐
in Zeile 3?	Fester Betrag 60 Euro ..... ☐	oder	Lotterielos ..... ☐
in Zeile 4?	Fester Betrag 90 Euro ..... ☐	oder	Lotterielos ..... ☐

ZP kann sich nicht entscheiden und verweigert die Antwort ..... ☐

Was für eine Persönlichkeit sind Sie?

22. Nun etwas ganz anderes. Unsere alltäglichen Handlungen werden davon beeinflusst, welche Grundüberzeugungen wir haben. Darüber ist in der Wissenschaft wenig bekannt.

Hier sind unterschiedliche Eigenschaften, die eine Person haben kann. Wahrscheinlich werden einige Eigenschaften auf Sie persönlich voll zutreffen und andere überhaupt nicht. Bei wieder anderen sind Sie vielleicht unentschieden.

☞ *Antworten Sie bitte anhand der folgenden Skala.*
*Der Wert **1** bedeutet: **trifft überhaupt nicht zu.***
*Der Wert **7** bedeutet: **trifft voll zu**.*
Mit den Werten zwischen 1 und 7 können Sie Ihre Meinung abstufen.

Ich bin jemand, der ...	**Trifft überhaupt nicht zu** 1	2	3	4	5	6	**Trifft voll zu** 7
– gründlich arbeitet	☐	☐	☐	☐	☐	☐	☐
– kommunikativ, gesprächig ist	☐	☐	☐	☐	☐	☐	☐
– manchmal etwas grob zu anderen ist	☐	☐	☐	☐	☐	☐	☐
– originell ist, neue Ideen einbringt	☐	☐	☐	☐	☐	☐	☐
– sich oft Sorgen macht	☐	☐	☐	☐	☐	☐	☐
– verzeihen kann	☐	☐	☐	☐	☐	☐	☐
– eher faul ist	☐	☐	☐	☐	☐	☐	☐
– aus sich herausgehen kann, gesellig ist	☐	☐	☐	☐	☐	☐	☐
– künstlerische Erfahrungen schätzt	☐	☐	☐	☐	☐	☐	☐
– leicht nervös wird	☐	☐	☐	☐	☐	☐	☐
– Aufgaben wirksam und effizient erledigt	☐	☐	☐	☐	☐	☐	☐
– zurückhaltend ist	☐	☐	☐	☐	☐	☐	☐
– rücksichtsvoll und freundlich mit anderen umgeht	☐	☐	☐	☐	☐	☐	☐
– eine lebhafte Phantasie, Vorstellungen hat	☐	☐	☐	☐	☐	☐	☐
– entspannt ist, mit Stress gut umgehen kann	☐	☐	☐	☐	☐	☐	☐

23. In welchem Maße treffen die folgenden Aussagen auf Sie persönlich zu?

☞ *Antworten Sie bitte wieder anhand einer Skala. Der Wert **1** bedeutet: **trifft überhaupt nicht zu**, der Wert **7** bedeutet: **trifft voll zu**.*

	Trifft überhaupt nicht zu 1	2	3	4	5	6	Trifft voll zu 7
Wenn mir jemand einen Gefallen tut, bin ich bereit, dies zu erwidern	☐	☐	☐	☐	☐	☐	☐
Wenn mir schweres Unrecht zuteil wird, werde ich mich um jeden Preis bei der nächsten Gelegenheit dafür rächen	☐	☐	☐	☐	☐	☐	☐
Wenn mich jemand in eine schwierige Lage bringt, werde ich das Gleiche mit ihm machen	☐	☐	☐	☐	☐	☐	☐
Ich strenge mich besonders an, um jemandem zu helfen, der mir früher schon mal geholfen hat	☐	☐	☐	☐	☐	☐	☐
Wenn mich jemand beleidigt, werde ich mich ihm gegenüber auch beleidigend verhalten	☐	☐	☐	☐	☐	☐	☐
Ich bin bereit, Kosten auf mich zu nehmen, um jemandem zu helfen, der mir früher einmal geholfen hat	☐	☐	☐	☐	☐	☐	☐

24. Die folgenden Aussagen kennzeichnen verschiedene Einstellungen zum Leben und zur Zukunft. In welchem Maße stimmen Sie persönlich den einzelnen Aussagen zu?

☞ *Antworten Sie bitte wieder anhand einer Skala. Der Wert **1** bedeutet: **stimme überhaupt nicht zu**, der Wert **7** bedeutet: **stimme voll zu**.*

	Stimme überhaupt nicht zu 1	2	3	4	5	6	Stimme voll zu 7
Wie mein Leben verläuft, hängt von mir selbst ab	☐	☐	☐	☐	☐	☐	☐
Im Vergleich mit anderen habe ich nicht das erreicht, was ich verdient habe	☐	☐	☐	☐	☐	☐	☐
Was man im Leben erreicht, ist in erster Linie eine Frage von Schicksal oder Glück	☐	☐	☐	☐	☐	☐	☐
Wenn man sich sozial oder politisch engagiert, kann man die sozialen Verhältnisse beeinflussen	☐	☐	☐	☐	☐	☐	☐
Ich mache häufig die Erfahrung, dass andere über mein Leben bestimmen	☐	☐	☐	☐	☐	☐	☐
Erfolg muss man sich hart erarbeiten	☐	☐	☐	☐	☐	☐	☐
Wenn ich im Leben auf Schwierigkeiten stoße, zweifle ich oft an meinen Fähigkeiten	☐	☐	☐	☐	☐	☐	☐
Welche Möglichkeiten ich im Leben habe, wird von den sozialen Umständen bestimmt	☐	☐	☐	☐	☐	☐	☐
Wichtiger als alle Anstrengungen sind die Fähigkeiten die man mitbringt	☐	☐	☐	☐	☐	☐	☐
Ich habe wenig Kontrolle über die Dinge, die in meinem Leben passieren	☐	☐	☐	☐	☐	☐	☐
Ich bemühe mich immer, eine Geldreserve zu haben für den Fall, dass ich unerwartete Ausgaben tätigen muss	☐	☐	☐	☐	☐	☐	☐

25. Welche der folgenden Tätigkeiten üben Sie in Ihrer freien Zeit aus? Geben Sie bitte zu jeder Tätigkeit an, wie oft Sie das machen: jede Woche, jeden Monat, seltener oder nie?

	Jede Woche	Jeden Monat	Seltener	Nie
Besuch von kulturellen Veranstaltungen, z.B. Konzerten, Theater, Vorträgen	☐	☐	☐	☐
Kinobesuch, Besuch von Popkonzerten, Tanzveranstaltungen, Discos, Sportveranstaltungen	☐	☐	☐	☐
Aktiver Sport	☐	☐	☐	☐
Künstlerische und musische Tätigkeiten (Musizieren, Tanzen, Theater, Malen, Fotografieren)	☐	☐	☐	☐
Geselligkeit mit Freunden, Verwandten oder Nachbarn	☐	☐	☐	☐
Mithelfen, wenn bei Freunden, Verwandten oder Nachbarn etwas zu tun ist	☐	☐	☐	☐
Ehrenamtliche Tätigkeiten in Vereinen, Verbänden oder sozialen Diensten	☐	☐	☐	☐
Beteiligung in Bürgerinitiativen, in Parteien, in der Kommunalpolitik	☐	☐	☐	☐
Kirchgang, Besuch religiöser Veranstaltungen	☐	☐	☐	☐

26. Gehören Sie einer Kirche oder Religionsgemeinschaft an? Wenn ja, sind Sie ...

- katholisch ☐
- evangelisch ☐
- Mitglied einer anderen christlichen Religionsgemeinschaft ☐
- Mitglied einer islamischen Religionsgemeinschaft ☐
- Mitglied einer anderen Religionsgemeinschaft ☐ ⇨ ☐ *Bitte angeben*

Nein, konfessionslos ☐

27. Sind Sie Mitglied in einer der folgenden Organisationen bzw. Verbände?

	Ja	Nein
– in einer Gewerkschaft?	☐	☐
– in einem Berufsverband?	☐	☐
– im Betriebs- oder Personalrat der Firma oder Behörde, in der Sie arbeiten?	☐	☐
– in einer Gruppe oder Organisation, die sich für die Erhaltung und den Schutz von Umwelt und Natur einsetzt?	☐	☐
– in einem Verein oder einer sonstigen Organisation?	☐	☐

28. Benutzen Sie privat oder beruflich einen Computer?

Nein ☐ → *Sie springen auf Frage **36!***
Ja, privat ☐
Ja, beruflich ☐

29. Seit wann benutzen Sie einen Computer?

Seit dem Jahr ☐☐☐☐

30. Für welche der folgenden Tätigkeiten nutzen Sie den Computer?
Geben Sie bitte zu jeder Tätigkeit an, wie oft Sie das machen.

	täglich	mindestens einmal pro Woche	mindestens einmal pro Monat	seltener	nie
Texte oder Grafiken erstellen	☐	☐	☐	☐	☐
Spielen	☐	☐	☐	☐	☐
Lernen, Bildung	☐	☐	☐	☐	☐
Filme ansehen oder Musik hören	☐	☐	☐	☐	☐
Programmieren	☐	☐	☐	☐	☐

31. Benutzen Sie privat oder beruflich das Internet?

Nein ☐ → *Sie springen auf Frage **36!***
Ja, privat ☐
Ja, beruflich ☐

32. Seit wann benutzen Sie das Internet?

Seit dem Jahr ☐☐☐☐

33. Wie würden Sie Ihre persönliche Kompetenz im Umgang mit dem Internet einschätzen?

☞ *Antworten Sie bitte anhand der folgenden Skala: Der Wert **0** bedeutet:* ganz und gar ***unerfahren***.
*Der Wert **10** bedeutet:* ***sehr erfahren***. *Mit den Werten dazwischen können Sie abstufen.*

ganz und gar ***unerfahren*** — ***sehr erfahren***

0 1 2 3 4 5 6 7 8 9 10

34. Für welche der folgenden Tätigkeiten nutzen Sie das Internet?
Geben Sie bitte zu jeder Tätigkeit an, wie oft Sie das machen.

	täglich	mindestens einmal pro Woche	mindestens einmal pro Monat	seltener	nie
E-mails lesen und schreiben	☐	☐	☐	☐	☐
Surfen, Informationen suchen	☐	☐	☐	☐	☐
Chatten oder Foren nutzen	☐	☐	☐	☐	☐
Einkaufen, Bestellungen machen	☐	☐	☐	☐	☐
Bankgeschäfte, Kontoführung	☐	☐	☐	☐	☐
Webseiten gestalten	☐	☐	☐	☐	☐

35. Wenn Sie im Internet die nächste Zugverbindung nach Wien in Erfahrung bringen müssten, wie viele Minuten würden Sie dann für diese Suche im Internet etwa aufwenden müssen?

ca. ☐ Minuten Weiß nicht ☐

36. Üben Sie derzeit eine Erwerbstätigkeit aus?
Was trifft für Sie zu?

☞ *Rentner mit Arbeitsvertrag gelten hier als erwerbstätig.*

Voll erwerbstätig ☐
In Teilzeitbeschäftigung ☐
In betrieblicher Ausbildung / Lehre oder betrieblicher Umschulung ☐
Geringfügig oder unregelmäßig erwerbstätig ☐
→ *Sie springen auf Frage **41!***

In Altersteilzeit mit Arbeitszeit Null ☐ → *Sie springen auf Frage **60!***

Als Wehrpflichtiger beim Wehrdienst ☐
Zivildienstleistender / Freiwilliges soziales Jahr ☐
Nicht erwerbstätig ☐

37. Beabsichtigen Sie, in der Zukunft (wieder) eine Erwerbstätigkeit aufzunehmen?

Nein, ganz sicher nicht ☐ → *Sie springen auf Frage **63!***
Eher unwahrscheinlich ☐
Wahrscheinlich ☐
Ganz sicher ☐

38. Wann etwa wollen Sie Ihre Erwerbstätigkeit aufnehmen?

Möglichst sofort ☐
Innerhalb des kommenden Jahres ☐
In 2 bis 5 Jahren ☐
Erst später, in mehr als 5 Jahren ☐

39. Wenn man Ihnen <u>jetzt</u> eine passende Stelle anbieten würde, könnten Sie dann innerhalb der nächsten zwei Wochen anfangen zu arbeiten?

Ja ☐ Nein ☐

40. Haben Sie innerhalb der letzten vier Wochen aktiv nach einer Stelle gesucht?

Ja ☐ Nein ☐

*Sie springen auf Frage **63!***

41. Gehört der Betrieb, in dem Sie arbeiten, zum öffentlichen Dienst?

Ja ..... ☐ Nein ..... ☐

42. In welchem Wirtschaftszweig / welcher Branche / welchem Dienstleistungsbereich ist das Unternehmen bzw. die Einrichtung überwiegend tätig?

☞ *Bitte geben Sie die genaue Bezeichnung an, z.B. nicht "Industrie", sondern "Elektroindustrie"; nicht "Handel", sondern "Einzelhandel", nicht "öffentlicher Dienst", sondern "Krankenhaus".*

Bitte in Druckbuchstaben eintragen!

43. Wie viele Beschäftigte hat das Gesamtunternehmen etwa?

- unter 5 Beschäftigte ..... ☐
- 5 bis 10 Beschäftigte ..... ☐
- 11 bis unter 20 Beschäftigte ..... ☐
- 20 bis unter 100 Beschäftigte ..... ☐
- 100 bis unter 200 Beschäftigte ..... ☐
- 200 bis unter 2.000 Beschäftigte ..... ☐
- 2.000 und mehr Beschäftigte ..... ☐

Trifft nicht zu, da selbständig ohne weitere Beschäftigte ..... ☐

44. Seit wann sind Sie bei Ihrem jetzigen Arbeitgeber beschäftigt?

☞ *Falls Sie Selbständiger sind, geben Sie bitte an, seit wann Sie Ihre jetzige Tätigkeit ausüben.*

Seit ☐☐ Monat ☐☐☐☐ Jahr

45. Handelt es sich um ein Zeitarbeits- bzw. Leiharbeitsverhältnis?

Ja ..... ☐ Nein ..... ☐

46. Haben Sie einen unbefristeten Arbeitsvertrag oder ein befristetes Arbeitsverhältnis?

Unbefristeten Arbeitsvertrag ..... ☐ ➡ *Sie springen auf Frage **48!***

Befristetes Arbeitsverhältnis ..... ☐

Trifft nicht zu / habe keinen Arbeitsvertrag ..... ☐

47. Handelt es sich um eine ABM-Stelle oder eine Gemeinnützige Arbeit (1-Euro-Job)?

Ja, ABM-Stelle ..... ☐

Ja, Gemeinnützige Arbeit (1-Euro-Job) ..... ☐

Nein ..... ☐

48. In welcher beruflichen Stellung sind Sie derzeit beschäftigt?

☞ *Wenn Sie **mehr als eine** berufliche Tätigkeit ausüben, beantworten Sie die folgenden Fragen bitte nur für Ihre derzeitige berufliche **Haupt**tätigkeit.*

Arbeiter
(auch in der Landwirtschaft)

Ungelernte Arbeiter ..... ☐

Angelernte Arbeiter ..... ☐

Gelernte und Facharbeiter ..... ☐

Vorarbeiter, Kolonnenführer ..... ☐

Meister, Polier ..... ☐

Selbständige
(einschl. mithelfende Familienangehörige)

	Anzahl der Mitarbeiter: Keine	1 – 9	10 und mehr
Selbständige Landwirte	☐	☐	☐
Freie Berufe, selbständige Akademiker	☐	☐	☐
Sonstige Selbständige	☐	☐	☐

Mithelfende Familienangehörige ..... ☐

Auszubildende und Praktikanten

Auszubildende / gewerblich-technisch ..... ☐

Auszubildende / kaufmännisch ..... ☐

Volontäre, Praktikanten u.ä. ..... ☐

Angestellte

Industrie- und Werkmeister im Angestelltenverhältnis ..... ☐

Angestellte mit einfacher Tätigkeit

– ohne Ausbildungsabschluss ..... ☐

– mit Ausbildungsabschluss ..... ☐

Angestellte mit qualifizierter Tätigkeit (z.B. Sachbearbeiter, Buchhalter, technischer Zeichner) ..... ☐

Angestellte mit hochqualifizierter Tätigkeit oder Leitungsfunktion (z.B. wissenschaftlicher Mitarbeiter, Ingenieur, Abteilungsleiter) ..... ☐

Angestellte mit umfassenden Führungsaufgaben (z.B. Direktor, Geschäftsführer, Vorstand größerer Betriebe und Verbände) ..... ☐

Beamte
(einschl. Richter und Berufssoldaten)

Einfacher Dienst ..... ☐

Mittlerer Dienst ..... ☐

Gehobener Dienst ..... ☐

Höherer Dienst ..... ☐

49. Sind Sie anderen bei der Arbeit vorgesetzt, d.h. dass Personen nach Ihren Anweisungen arbeiten? Wenn ja, wie viele Personen arbeiten nach Ihren Anweisungen?

Ja ..... ☐ und zwar ☐ Personen

Nein ..... ☐

50. Wie viele Wochenstunden beträgt Ihre vereinbarte Arbeitszeit ohne Überstunden?

☐☐ , ☐ Stunden pro Woche Keine festgelegte Arbeitszeit ..... ☐

51. Und wieviel beträgt im Durchschnitt Ihre tatsächliche Arbeitszeit pro Woche einschließlich eventueller Überstunden?

☐☐ , ☐ Stunden pro Woche

Wenn unter 30 Stunden: ⇩

Wenn 30 Stunden und mehr ➡ *Sie springen auf Frage **54!***

52. Ist das eine geringfügige Beschäftigung nach der 400- bzw. 800-Euro-Regelung (Mini- bzw. Midi-Job)?

Ja, Mini-Job (bis 400 Euro) ..... ☐

Ja, Midi-Job (401 bis 800 Euro) ..... ☐

Nein ..... ☐

53. Ist das eine Teilzeitbeschäftigung nach dem Gesetz zur Elternteilzeit (Bundeserziehungsgeldgesetz)?

Ja ..... ☐ Nein ..... ☐

54. Bitte geben Sie an, in welchem Maße Sie den folgenden Aussagen zustimmen.

	stimme gar nicht zu	stimme eher nicht zu	stimme eher zu	stimme voll zu
Beim Arbeiten komme ich leicht in Zeitdruck	☐	☐	☐	☐
Es passiert mir oft, dass ich schon beim Aufwachen an Arbeitsprobleme denke	☐	☐	☐	☐
Wenn ich nach Hause komme, fällt mir das Abschalten von der Arbeit sehr leicht	☐	☐	☐	☐
Diejenigen, die mir am nächsten stehen sagen, dass ich mich für meinen Beruf zu sehr aufopfere	☐	☐	☐	☐
Die Arbeit lässt mich selten los, das geht mir abends im Kopf rum	☐	☐	☐	☐
Wenn ich etwas verschiebe, was ich eigentlich heute tun müsste, kann ich nachts nicht schlafen	☐	☐	☐	☐

55a. Ich lese Ihnen Aussagen über mögliche berufliche Belastungen an Ihrem derzeitigen Arbeitsplatz vor. Bitte geben Sie an, ob der jeweilige Punkt bei Ihnen zutrifft und, falls ja, wie stark Sie das belastet.

	Zutreffend?		Und wie stark belastet Sie das?			
	Nein	Ja	Gar nicht	Mäßig	Stark	Sehr stark
Aufgrund des hohen Arbeitsaufkommens besteht häufig großer Zeitdruck	☐	☐ ⇨	☐	☐	☐	☐
Bei meiner Arbeit werde ich häufig unterbrochen und gestört	☐	☐ ⇨	☐	☐	☐	☐
Im Laufe der letzten beiden Jahre ist meine Arbeit immer mehr geworden	☐	☐ ⇨	☐	☐	☐	☐
Die Aufstiegschancen in meinem Betrieb sind schlecht	☐	☐ ⇨	☐	☐	☐	☐
Ich erfahre - oder erwarte - eine Verschlechterung meiner Arbeitssituation	☐	☐ ⇨	☐	☐	☐	☐
Mein eigener Arbeitsplatz ist gefährdet	☐	☐ ⇨	☐	☐	☐	☐

55b. Und wie ist es mit den folgenden Punkten?

	Zutreffend?		Und wie stark belastet Sie das?			
	Ja	Nein	Gar nicht	Mäßig	Stark	Sehr stark
Ich erhalte von meinen Vorgesetzen die Anerkennung, die ich verdiene	☐	☐ ⇨	☐	☐	☐	☐
Wenn ich an all die erbrachten Leistungen und Anstrengungen denke, halte ich die erfahrene Anerkennung für angemessen	☐	☐ ⇨	☐	☐	☐	☐
Wenn ich an all die erbrachten Leistungen und Anstrengungen denke, halte ich meine persönlichenen Chancen des beruflichen Fortkommens für angemessen	☐	☐ ⇨	☐	☐	☐	☐
Wenn ich an all die erbrachten Leistungen denke, halte ich mein Gehalt/meinen Lohn für angemessen	☐	☐ ⇨	☐	☐	☐	☐

56. Wenn Sie den Umfang Ihrer Arbeitszeit selbst wählen könnten und dabei berücksichtigen, dass sich Ihr Verdienst entsprechend der Arbeitszeit ändern würde:

Wie viele Stunden in der Woche würden Sie dann am liebsten arbeiten?

☐☐ , ☐ Stunden pro Woche

57. Kommt es vor, dass Sie Überstunden leisten?

Ja ..... ☐ ⇩

Nein ..... ☐
Trifft nicht zu, da Selbständiger ..... ☐

*Sie springen auf Frage **59!***

58. Wie war das im letzten Monat: Haben Sie da Überstunden geleistet? Wenn ja, wieviele?

Ja ..... ☐ ⇨ und zwar: ☐☐ Stunden, davon bezahlt: ..... ☐☐ Stunden

Bitte "00" eintragen wenn unbezahlt!

Nein ..... ☐

59. Die Arbeitsverdienste sind ja nicht in allen Regionen Deutschlands gleich. Daher die Frage: Liegt Ihre Arbeitsstätte . . .

. . . in einem der alten Bundesländer ..... ☐
. . . in einem der neuen Bundesländer ..... ☐
. . . im Westteil Berlins ..... ☐
. . . im Ostteil Berlins ..... ☐
. . . oder im Ausland? ..... ☐

60. Wie hoch war Ihr Arbeitsverdienst im letzten Monat?

☞ *Wenn Sie im letzten Monat Sonderzahlungen hatten, z.B. Urlaubsgeld oder Nachzahlungen, rechnen Sie diese bitte **nicht** mit. Entgelt für Überstunden rechnen Sie dagegen mit.*

☞ *Falls Sie selbständig sind: Bitte schätzen Sie Ihren monatlichen Gewinn vor und nach Steuer.*

Bitte geben Sie nach Möglichkeit beides an:

- **den Bruttoverdienst, das heißt Lohn oder Gehalt vor Abzug der Steuern und Sozialversicherung**
- **und den Nettoverdienst, das heißt den Betrag nach Abzug von Steuern und Beiträgen zur Renten-, Arbeitslosen- und Krankenversicherung.**

Der Arbeitsverdienst betrug: brutto ☐ EURO

netto ☐ EURO

61. Beeinhaltete Ihr Einkommen des letzten Jahres Sonderzahlungen oder Zuschläge der folgenden Art?

☞ *Bitte alles Zutreffende ankreuzen!*

Schicht-/Spätarbeits-/Wochenendzuschläge ..... ☐
Überstundenzuschläge ..... ☐
Erschwerniszulagen ..... ☐
Funktionszulagen ..... ☐
Persönliche Zulagen ..... ☐
13. Monatsgehalt ..... ☐
14. Monatsgehalt ..... ☐
Weihnachtsgeld ..... ☐
Urlaubsgeld ..... ☐
Jahresprämie ..... ☐
Gruppenprämie ..... ☐
Gewinnbeteiligung ..... ☐
verbilligt überlassene Aktienoptionen ..... ☐
Fahrtgeld ..... ☐
Trinkgelder ..... ☐
Sonstiges ..... ☐ ⇨ ☐
Nichts davon ..... ☐

Bitte eintragen

62. Haben Sie innerhalb des letzten Jahres von Ihrem Arbeitsgeber andere Zusatzleistungen zum Gehalt erhalten?

☞ *Bitte alles Zutreffende ankreuzen!*

Firmenwagen für private Nutzung ..... ☐
Spesen, die über die reine Kostenerstattung hinausgehen ..... ☐
verbilligte Wohnung/Haus ..... ☐
Zahlungen für Telefonkosten ..... ☐
Betriebskindergarten oder Beiträge für Kinderbetreuung ..... ☐
Betriebskantine oder Essenszuschuss ..... ☐
Zuschüsse für öffentlichen Nahverkehr (Job-Ticket) ..... ☐
Computer für zu Hause/Laptop ..... ☐
Handy für den privaten Gebrauch ..... ☐
Günstige Kredite oder Hypotheken ..... ☐
Deputate (Sachleistungen, Naturalien) ..... ☐
Sonstiges ..... ☐ ⇨ ☐
Nichts davon ..... ☐

Bitte eintragen

63. Sind Sie in den letzten 12 Monaten aus einer beruflichen Tätigkeit bzw. einer Stelle, die Sie vorher gehabt haben, ausgeschieden?

Ja ..... ☐ ⇩ Nein ..... ☐ ➡ *Sie springen auf Frage **65!***

64. Auf welche Weise wurde diese Beschäftigung beendet?

☞ *Bitte nur **einen** Punkt ankreuzen!*

Wegen Betriebsstillegung / Auflösung der Dienststelle ..... ☐
Durch eigene Kündigung ..... ☐
Durch Kündigung von seiten des Arbeitgebers ..... ☐
Durch Auflösungsvertrag / im Einvernehmen ..... ☐
Befristete Beschäftigung oder Ausbildung war beendet ..... ☐
Erreichen der Altersgrenze / Rente / Pension ..... ☐
Beurlaubung ..... ☐
Aufgabe der selbständigen Tätigkeit / des Geschäfts ..... ☐

Gesundheit und Krankheit

65. Wie würden Sie Ihren gegenwärtigen Gesundheitszustand beschreiben?

Sehr gut ..... ☐
Gut ..... ☐
Zufriedenstellend ..... ☐
Weniger gut ..... ☐
Schlecht ..... ☐

66. Wenn Sie Treppen steigen müssen, also mehrere Stockwerke zu Fuß hochgehen: Beeinträchtigt Sie dabei Ihr Gesundheitszustand stark, ein wenig oder gar nicht?

Stark ..... ☐
Ein wenig ..... ☐
Gar nicht ..... ☐

67. Und wie ist das mit anderen anstrengenden Tätigkeiten im Alltag, wo man z.B. etwas Schweres heben muss oder Beweglichkeit braucht: Beeinträchtigt Sie dabei Ihr Gesundheitszustand stark, ein wenig oder gar nicht?

Stark ..... ☐
Ein wenig ..... ☐
Gar nicht ..... ☐

67a. Wie ist das mit Ihrem Schlaf? Schlafen Sie nachts in der Regel ...

Sehr gut ..... ☐
Gut ..... ☐
Zufriedenstellend ..... ☐
Weniger gut ..... ☐
Schlecht ..... ☐

68. Bitte denken Sie einmal an die letzten vier Wochen. Wie oft kam es in dieser Zeit vor, . . .

	Immer	Oft	Manch-mal	Fast nie	Nie
● dass Sie sich gehetzt oder unter Zeitdruck fühlten?	☐	☐	☐	☐	☐
● dass Sie sich niedergeschlagen und trübsinnig fühlten?	☐	☐	☐	☐	☐
● dass Sie sich ruhig und ausgeglichen fühlten?	☐	☐	☐	☐	☐
● dass Sie jede Menge Energie verspürten?	☐	☐	☐	☐	☐
● dass Sie starke körperliche Schmerzen hatten?	☐	☐	☐	☐	☐
● dass Sie wegen gesundheitlicher Probleme körperlicher Art in Ihrer Arbeit oder Ihren alltäglichen Beschäftigungen					
– weniger geschafft haben als Sie eigentlich wollten?	☐	☐	☐	☐	☐
– in der Art Ihrer Tätigkeiten eingeschränkt waren?	☐	☐	☐	☐	☐
● dass Sie wegen seelischer oder emotionaler Probleme in Ihrer Arbeit oder Ihren alltäglichen Beschäftigungen					
– weniger geschafft haben als Sie eigentlich wollten?	☐	☐	☐	☐	☐
– Ihre Arbeit oder Tätigkeit weniger sorgfältig als sonst gemacht haben?	☐	☐	☐	☐	☐
● dass Sie wegen gesundheitlicher oder seelischer Probleme in Ihren sozialen Kontakten, z.B. mit Freunden, Bekannten oder Verwandten, eingeschränkt waren?	☐	☐	☐	☐	☐

69. Wie oft betreiben Sie aktiv Sport, Fitness oder Gymnastik?

Regelmäßig ..... ☐
Gelegentlich ..... ☐
Nie ..... ☐

70. Inwieweit achten Sie auf gesundheitsbewusste Ernährung?

Sehr stark ..... ☐
Stark ..... ☐
Ein wenig ..... ☐
Gar nicht ..... ☐

71. Wie häufig trinken Sie die folgenden alkoholischen Getränke?

	regelmäßig	ab und zu	selten	nie
Bier	☐	☐	☐	☐
Wein, Sekt	☐	☐	☐	☐
Spirituosen (Schnaps, Weinbrand etc.)	☐	☐	☐	☐
Mischgetränke (Alkopops, Cocktails etc.)	☐	☐	☐	☐

72. Rauchen Sie gegenwärtig, seien es Zigaretten, Pfeifen oder Zigarren?

Nein ..... ☐
Ja ..... ☐ ⇨ **Wieviel Zigaretten, Pfeifen oder Zigarren rauchen Sie pro Tag?**
☞ *Geben Sie bitte den täglichen Durchschnitt der letzten Woche an.*

Zigaretten ..... ☐☐ pro Tag
Pfeifen ..... ☐☐ pro Tag
Zigarren/Zigarillos ..... ☐☐ pro Tag

73. Was ist Ihre Körpergröße in cm?

☞ *Wenn Sie es nicht genau wissen, schätzen Sie bitte.*

☐☐☐ cm

74. Wieviel Kilogramm wiegen Sie gegenwärtig?

☞ *Wenn Sie es nicht genau wissen, schätzen Sie bitte.*

☐☐☐ kg

75. Sind Sie nach amtlicher Feststellung erwerbsgemindert oder schwerbehindert?

Ja ..... ☐ ⇨ **Wie hoch ist Ihre Erwerbsminderung oder Schwerbehinderung nach der letzten Feststellung?** ..... ☐☐☐ %

Und welcher Art ist die Behinderung? ☐
Bitte eintragen!

Nein ..... ☐

76. Was glauben Sie:
Wie wahrscheinlich ist es, dass Sie 75 Jahre alt oder älter werden?

☞ *Antworten Sie bitte anhand der folgenden Skala, bei der "0" absolut **unwahrscheinlich,** "10" höchst **wahrscheinlich** bedeutet.*

☐–☐–☐–☐–☐–☐–☐–☐–☐–☐–☐

0 1 2 3 4 5 6 7 8 9 10
absolut ***unwahrscheinlich*** — *höchst* ***wahrscheinlich***

77. Wir möchten nun mit Ihnen einen sogenannten "Greifkrafttest" durchführen. Mit diesem Test wird gemessen, welche Kraft ein Mensch beim Greifen ausüben kann. Das kann zur Beurteilung der körperlichen Verfassung wichtig sein. Zu Beginn des Tests haben wir zwei Fragen.

Sind Sie von Natur aus Rechts- oder Linkshänder?

Rechtshänder ..... ☐
Linkshänder ..... ☐

78. Mit welcher Hand schreiben Sie tatsächlich?

Rechte Hand ..... ☐
Linke Hand ..... ☐

79. Jetzt folgt der eigentliche Test.

Der Test ist nicht schädlich und kann in jedem Alter durchgeführt werden, außer bestimmte medizinische Gründe sprechen dagegen, wie zum Beispiel:
- **Schwellungen**
- **Entzündungen**
- **Schmerzen**
- **Operationen oder Verwundungen in den letzten sechs Monaten**

Ist nur eine Hand betroffen, führen Sie bitte nur die zwei Messungen mit der beschwerdefreien Hand durch.

Int: Sie haben sich bereits mit der Anleitung und dem Gerät vertraut gemacht.

Hier zur Erinnerung die wichtigsten Punkte.

1. Den Pfeil des Gerätes per Hand auf Null stellen.
2. Den Griff des Gerätes auf die Handgröße des Befragten anpassen.
3. Störenden Schmuck an den Händen, ablegen.
4. Die Zielperson soll möglichst stehen.
5. Der Arm der Zielperson muss am Körper anliegen und darf bei der Messung nicht angehoben werden.
6. Der Unterarm muss im rechten Winkel gehalten werden.
7. Die Zielperson soll ein paar Sekunden lang, so fest wie Sie kann, zudrücken.
8. Den erreichten Wert, auf ein halbes Kilo genau eingeben.
9. Zwei Werte für jede Hand ermitteln (abwechselnd zwischen rechts und links).
10. Abweichungen von der Testanordnung in der folgenden Frage F080 vermerken.

Zuerst:		**Dann, falls möglich:**
Erste Messung (rechts-links)		**Zweite Messung (rechts-links)**
Rechts ..... ☐☐, ☐ kg	nicht möglich ..... ☐	Rechts ..... ☐☐, ☐ kg
Links ..... ☐☐, ☐ kg	nicht möglich ..... ☐	Links ..... ☐☐, ☐ kg

Greifkrafttest verweigert ..... ☐

80. Gab es unvermeidliche Abweichungen vom vorgesehenen Ablauf?

Ja ..... ☐ und zwar: ☐
Nein ..... ☐

81. Auf dieser Liste stehen Personen, die für Sie in irgendeiner Weise bedeutsam sein können. Wie ist es bei Ihnen, wenn es um folgende Dinge geht?

☞ *Nennen Sie bitte jeweils bis zu drei Personen von der Liste.*

a) Mit wem teilen Sie persönliche Gedanken und Gefühle oder sprechen über Dinge, die Sie nicht jedem erzählen würden?

Kennziffer: ☐ ☐ ☐

Mit niemand .. ☐

b) Wer unterstützt Sie in Ihrem beruflichen Fortkommen oder Ihrer Ausbildung und hilft Ihnen, dass Sie vorankommen?

Kennziffer: ☐ ☐ ☐

Niemand ..... ☐

c) Nur einmal hypothetisch gefragt: Wie wäre es bei einer langfristigen Pflegebedürftigkeit, z.B. nach einem schweren Unfall:
Wen würden Sie um Hilfe bitten?

Kennziffer: ☐ ☐ ☐

Niemand ..... ☐

d) Mit wem haben Sie gelegentlich Streit oder Konflikte, die belastend für Sie sind?

Kennziffer: ☐ ☐ ☐

Mit niemand .. ☐

e) Wer kann Ihnen auch mal unangenehme Wahrheiten sagen?

Kennziffer: ☐ ☐ ☐

Niemand ..... ☐

Kennziffer ⇩

Ehe, Partnerschaft

01 (Ehe-)Partner, (Ehe-)Partnerin
02 ehemalige/r Ehepartner(in)/Partner(in)

Familie

03 Mutter
04 Vater
05 Stiefmutter oder Pflegemutter
06 Stiefvater oder Pflegevater

07 Schwiegermutter
08 Schwiegervater

09 Tochter
10 Sohn

11 Schwester
12 Bruder

13 Großmutter
14 Großvater
15 Enkel

16 Tante/Nichte
17 Onkel/Neffe
18 andere weibliche Verwandte
19 andere männliche Verwandte

Andere Personen

20 Kollegen von der Arbeit
21 Vorgesetzte im Betrieb
22 aus Schule/Ausbildung/Studium
23 aus der Nachbarschaft
24 aus Verein oder Freizeitaktivitäten
25 bezahlte Helfer, ambulanter Dienst, Sozialhelfer
26 Andere

82. Nun eine Frage zu Ihrem Bekannten- und Freundeskreis:
Denken Sie bitte an <u>drei Personen</u>, mit denen Sie näher befreundet sind und mit denen Sie sich häufig treffen.
Es kann sich dabei sowohl um Verwandte als auch um Nicht-Verwandte handeln, nur nicht um Personen, die mit Ihnen hier im selben Haushalt wohnen.

Sagen Sie uns bitte zur ersten, zur zweiten und zur dritten Person:

		erste Person	zweite Person	dritte Person
a) Ist er oder sie mit Ihnen verwandt?	ja	☐	☐	☐
	nein	☐	☐	☐
b) Ist er oder sie	männlich	☐	☐	☐
	weiblich	☐	☐	☐
d) Wie alt ist er oder sie? Wenn Sie es nicht genau wissen, bitte schätzen	Jahre	☐☐	☐☐	☐☐
e) Welche Nationalität oder Herkunft hat er oder sie?				
Ist aus den alten Bundesländern		☐	☐	☐
Ist aus den neuen Bundesländern		☐	☐	☐
Ist aus einem anderen Land oder Ausländer		☐ ⇨	☐ ⇨	☐ ⇨
Kommen Sie selbst aus demselben Land?	ja	☐	☐	☐
	nein	☐	☐	☐
f) Ist er oder sie?				
- in Vollzeit erwerbstätig		☐	☐	☐
- in Teilzeit erwerbstätig		☐	☐	☐
- arbeitslos gemeldet		☐	☐	☐
- in Schule/Studium/Ausbildung		☐	☐	☐
- in Rente/Pension		☐	☐	☐
- Sonstiges		☐	☐	☐
g) Welchen höchsten Bildungsabschluss hat er oder sie?				
Keinen Abschluss		☐	☐	☐
Volks-/Hauptschulabschluss (DDR: 8. Klasse)		☐	☐	☐
Mittlere Reife, Realschulabschluss (DDR: 10. Klasse)		☐	☐	☐
Abitur/Hochschulreife (DDR: EOS)/Fachhochschulreife		☐	☐	☐
Weiß nicht		☐	☐	☐

83. Und nun zu Ihrem engeren wie weiteren Familienkreis.

Welche der folgenden Verwandten haben Sie? Für diese Personen geben Sie bitte die Anzahl an und zusätzlich, ob sie in Ihrem Haushalt leben bzw. wie weit entfernt sie wohnen.

☞ *Wenn es mehrere Personen der jeweiligen Kategorie gibt, geben Sie den Ort nur für die nächstwohnende Person an.*

	Ja vorhanden		Kennziffern 0 - 6 lt. Liste
(Ehe-) Partner	☐ ⇨		lebt wo: ☐
ehemaliger (Ehe-) Partner	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Mutter	☐ ⇨		lebt wo: ☐
Vater	☐ ⇨		lebt wo: ☐
Stiefmutter oder Pflegemutter	☐ ⇨		lebt wo: ☐
Stiefvater oder Pflegevater	☐ ⇨		lebt wo: ☐
Tochter / Töchter	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Sohn / Söhne	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Schwester / Schwestern (auch Halbgeschwister)	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Bruder / Brüder (auch Halbgeschwister)	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Großmutter	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Großvater	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Enkel	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Tanten/Onkel	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Nichten/Neffen	☐ ⇨	Anzahl: ☐	lebt wo: ☐
Sonstige Verwandte, mit denen Sie näheren Kontakt haben	☐ ⇨	Anzahl: ☐	lebt wo: ☐

Kennziffer ⇩	Wo lebt die Person?
0	im gleichen Haushalt
1	im gleichen Haus
2	in der Nachbarschaft
3	im gleichen Ort, aber mehr als 15 Minuten Fußweg
4	in einem anderen Ort, aber innerhalb 1 Stunde erreichbart
5	weiter entfernt (aber in Deutschland)
6	im Ausland

84. Verschiedenen Menschen sind verschiedene Dinge wichtig.
Wie wichtig sind für Sie persönlich heute die folgenden Dinge?

☞ *Antworten Sie bitte anhand der folgenden Skala:*
*Der Wert **1** bedeutet: **ganz und gar unwichtig**.*
*Der Wert **7** bedeutet: **sehr wichtig**.*
Mit den Werten dazwischen können Sie abstufen.

	ganz und gar unwichtig 1	2	3	4	5	6	sehr wichtig 7
Sich etwas leisten können	☐	☐	☐	☐	☐	☐	☐
Für andere da sein	☐	☐	☐	☐	☐	☐	☐
Sich selbst verwirklichen	☐	☐	☐	☐	☐	☐	☐
Erfolg im Beruf haben	☐	☐	☐	☐	☐	☐	☐
Ein eigenes Haus haben	☐	☐	☐	☐	☐	☐	☐
Eine glückliche Ehe/Partnerschaft haben	☐	☐	☐	☐	☐	☐	☐
Kinder haben	☐	☐	☐	☐	☐	☐	☐
Sich politisch, gesellschaftlich einsetzen	☐	☐	☐	☐	☐	☐	☐
Die Welt sehen, viele Reisen machen	☐	☐	☐	☐	☐	☐	☐

85. In welchem Maße treffen die folgenden Aussagen auf Sie persönlich zu?

☞ *Antworten Sie bitte wieder anhand einer Skala.*
*Der Wert **1** bedeutet: **trifft überhaupt nicht zu**,*
*der Wert **7** bedeutet: **trifft voll zu**.*

	Trifft überhaupt nicht zu 1	2	3	4	5	6	Trifft voll zu 7
Ich habe immer jemanden in erreichbarer Nähe, zu dem ich mit alltäglichen Problemen kommen kann	☐	☐	☐	☐	☐	☐	☐
Mir fehlt ein wirklich guter Freund/eine gute Freundin	☐	☐	☐	☐	☐	☐	☐
Manchmal habe ich das Gefühl, dass mich große Leere erfasst	☐	☐	☐	☐	☐	☐	☐
Es gibt ausreichend viele Menschen auf die ich zählen kann, wenn ich Schwierigkeiten habe	☐	☐	☐	☐	☐	☐	☐
Ich vermisse die Geselligkeit mit anderen Menschen	☐	☐	☐	☐	☐	☐	☐
Ich finde, der Kreis meiner Freunde und Bekannten ist zu klein	☐	☐	☐	☐	☐	☐	☐
Es gibt viele Personen, denen ich voll vertrauen kann	☐	☐	☐	☐	☐	☐	☐
Ich habe genügend Personen, denen ich mich sehr verbunden fühle	☐	☐	☐	☐	☐	☐	☐
Ich vermisse einen häufigeren Kontakt mit meinen Familienangehörigen und Verwandten	☐	☐	☐	☐	☐	☐	☐
Oft fühle ich mich im Stich gelassen	☐	☐	☐	☐	☐	☐	☐
Wenn ich meine Freunde brauche, kann ich jederzeit auf sie zählen	☐	☐	☐	☐	☐	☐	☐

86. Wie wichtig ist es für Sie in den folgenden Lebensbereichen, in die Zukunft zu planen?

☞ *Antworten Sie bitte wieder anhand einer Skala.*
*Der Wert **1** bedeutet: **ganz und gar unwichtig,***
*der Wert **7** bedeutet: **sehr wichtig.***
Mit den Werten dazwischen können Sie abstufen.

	ganz und gar unwichtig 1	2	3	4	5	6	sehr wichtig 7
Gesundheit	☐	☐	☐	☐	☐	☐	☐
Familie	☐	☐	☐	☐	☐	☐	☐
Beruf	☐	☐	☐	☐	☐	☐	☐
Finanzielle Fragen	☐	☐	☐	☐	☐	☐	☐

87. Über seine Zukunft kann man sich mehr oder weniger Gedanken machen: Wie oft denken Sie bei Ihren Entscheidungen ein Jahr voraus, wie oft 5 Jahre und wie oft 10 Jahre voraus?

	ein Jahr voraus	5 Jahre voraus	10 Jahre voraus
Sehr oft	☐	☐	☐
Oft	☐	☐	☐
Manchmal	☐	☐	☐
Selten	☐	☐	☐
Nie	☐	☐	☐

88. Menschen machen sich manchmal Vorsätze, Dinge zu ändern, die sie an sich selbst nicht mögen. Wenn Sie sich solche Vorsätze machen, wie häufig brechen Sie diese dann wieder?

☞ *Antworten Sie bitte wieder anhand einer Skala.*
*Der Wert **1** bedeutet: **niemals,***
*der Wert **7** bedeutet: **sehr oft.***
Mit den Werten dazwischen können Sie abstufen.

	niemals 1	2	3	4	5	6	sehr oft 7	trifft nicht zu
Im Allgemeinen	☐	☐	☐	☐	☐	☐	☐	☐
Ernährung	☐	☐	☐	☐	☐	☐	☐	☐
Fitness	☐	☐	☐	☐	☐	☐	☐	☐
Geld ausgeben	☐	☐	☐	☐	☐	☐	☐	☐
Rauchen	☐	☐	☐	☐	☐	☐	☐	☐
Alkoholkonsum	☐	☐	☐	☐	☐	☐	☐	☐

89. Was erwarten Sie in Bezug auf Ihre finanzielle Situation in der Zukunft? Wird sie besser, etwa gleich oder eher schlechter sein als heute?

Und zwar ...	besser	etwa gleich	eher schlechter
in sechs Monaten	☐	☐	☐
in einem Jahr	☐	☐	☐
in zwei Jahren	☐	☐	☐

90. Wie sicher oder unsicher ist es, dass die Entwicklung Ihrer finanziellen Situation so eintreten wird?

☞ *Antworten Sie bitte anhand der folgenden Skala,*
bei der "0" absolut **unsicher,** *"10"* absolut ***sicher*** *bedeutet.*

Und zwar ...	absolut unsicher 0	1	2	3	4	5	6	7	8	9	absolut sicher 10
in sechs Monaten	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
in einem Jahr	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
in zwei Jahren	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

91. Ist Ihr Konto gegenwärtig überzogen?

Ja ☐

Nein ☐ Ich habe kein Konto ☐ ➡ *Sie springen auf Frage **93!***

92. An wie vielen Tagen im Jahr ist Ihr laufendes Konto überzogen?

Niemals ☐
1 - 30 Tage ☐
31 - 90 Tage ☐
Mehr als 90 Tage ☐

93. Wenn Sie plötzlich in eine unvorhergesehene Situation geraten würden, und Sie innerhalb von zwei Wochen etwa 1.000 Euro bezahlen müßten, könnten Sie das schaffen?

Ja ☐ ⇩ Nein ☐ ➡ *Sie springen auf Frage **95!***

94. Und wie würden Sie das Geld beschaffen?

Vom eigenen Bankkonto ☐
Darlehen von einem Haushaltsmitglied ☐
Darlehen von Verwandten / Freunden ☐
Bankkredit ☐
Sonstiges ☐

und zwar:

95. Wenn Sie einmal überschlägig das gesamte Vermögen dieses Haushalts zusammenrechnen (Geld- und Sachvermögen einschließlich des selbstgenutzen Wohneigentums), wie hoch schätzen Sie den Gesamtwert?
Geben Sie bitte die zutreffende Größenordnung an.

☞ *Eventuell aufgenommene Hypotheken, Darlehen oder Kredite sind abzuziehen.*

- Keinerlei Vermögen ..... ☐
- unter 5.000 Euro ..... ☐
- 5.000 bis unter 20.000 Euro ..... ☐
- 20.000 bis unter 50.000 Euro ..... ☐
- 50.000 bis unter 100.000 Euro ..... ☐
- 100.000 bis unter 200.000 Euro ..... ☐
- 200.000 bis unter 500.000 Euro ..... ☐
- 500.000 bis unter 1.000.000 Euro ..... ☐
- 1.000.000 Euro und mehr ..... ☐
- Weiß nicht ..... ☐

96. K-Test 1(Tiere nennen)

97. Nehmen Sie an, Sie werfen eine Münze, die gleichmäßig auf die eine oder die andere Seite fällt. Nach acht Würfen beobachten Sie folgendes Ergebnis:

Zahl - Zahl - Zahl - Kopf - Zahl - Kopf - Kopf - Kopf

Wie hoch ist die Wahrscheinlichkeit, ausgedrückt in Prozent, dass der nächste Wurf "Zahl" ist?

☐☐☐ % Weiß nicht ..... ☐

98. Stellen Sie sich vor, Sie sind im Urlaub in einer Gegend, in der meistens schönes Wetter ist und Sie fragen sich, wie wohl morgen das Wetter sein wird. Angenommen an Ihrem Urlaubsort scheint im Durchschnitt an 90 von 100 Tagen die Sonne und an 10 von 100 Tagen regnet es. Für morgen lautet der Wetterbericht der Wettervorhersage, dass es regnen wird. Die Wettervorhersage sagt das Wetter für den nächsten Tag im Durchschnitt an 80 von 100 Tagen richtig voraus.
Was glauben Sie: Wie hoch ist die Wahrscheinlichkeit, ausgedrückt in Prozent, dass es morgen regnet?

☐☐☐ % Weiß nicht ..... ☐

APPENDIX B

Appendix to Chapter 2

Table B.1: Probability Judgment - Detailed Descriptive Statistics

	[0%,20%]	(20%,40%]	(40%,60%]	(60%,80%)	80%	(80%,100%]	Don't know	Total
All	205 (20.7%)	58 (5.9%)	149 (15.1%)	56 (5.7%)	239 (24.2%)	152 (15.4%)	129 (13.1%)	988 (100%)
Female	117 (22.0%)	31 (5.8%)	71 (13.3%)	29 (5.5%)	116 (21.8%)	85 (16.0%)	83 (15.6%)	532 (100%)
Male	88 (19.3%)	27 (5.9%)	78 (17.1%)	27 (5.9%)	123 (27.0%)	67 (14.7%)	46 (10.1%)	456 (100%)
Age < 50	107 (19.3%)	27 (4.9%)	82 (14.8%)	35 (6.3%)	158 (28.5%)	98 (17.7%)	47 (8.5%)	554 (100%)
Age ≥ 50	98 (22.6%)	31 (7.1%)	67 (15.4%)	21 (4.8%)	81 (18.7%)	54 (12.4%)	82 (18.9%)	434 (100%)
Word fluency measure > 0	114 (20.2%)	26 (4.6%)	80 (14.2%)	31 (5.5%)	151 (26.7%)	94 (16.6%)	69 (12.2%)	565 (100.0%)
Word fluency measure ≤ 0	91 (21.5%)	32 (7.6%)	69 (16.3%)	25 (5.9%)	88 (20.8%)	58 (13.7%)	60 (14.2%)	423 (100.0%)
Perceptual speed measure > 0	112 (19.8%)	24 (4.2%)	76 (13.5%)	40 (7.1%)	148 (26.2%)	96 (17.0%)	69 (12.2%)	565 (100.0%)
Perceptual speed measure ≤ 0	93 (22.0%)	34 (8.0%)	73 (17.3%)	16 (3.8%)	91 (21.5%)	56 (13.2%)	60 (14.2%)	423 (100.0%)
Years of schooling > 10	68 (21.0%)	16 (4.9%)	37 (11.4%)	23 (7.1%)	102 (31.5%)	51 (15.7%)	27 (8.3%)	324 (100%)
Years of schooling ≤ 10	137 (20.6%)	42 (6.3%)	112 (16.9%)	33 (5.0%)	137 (20.6%)	101 (15.2%)	102 (15.4%)	664 (100%)
University degree	30 (25.9%)	3 (2.6%)	11 (9.5%)	9 (7.8%)	44 (37.9%)	12 (10.3%)	7 (6.0%)	116 (100%)
No univ. degree	175 (20.1%)	55 (6.3%)	138 (15.8%)	47 (5.4%)	195 (22.4%)	140 (16.1%)	122 (14.0%)	872 (100%)

Table B.1: Probability Judgment - Detailed Descriptive Statistics. Bayesian answer is located in (20%,40%], base rate neglect leads to 80%, pure base rate is in [0%,20%]. Cognitive ability measures are standardized to have mean 0 and standard deviation 1.

Table B.2: Detailed Multinomial Analysis (I)

Dependent variable: answer in %, base category is x=80% (base rate neglect)

	[0%,20%]	(20%,40%]	(40%,60%]	(60%,80%)	(80%,100%]	Don't know
Word fluency measure	-0.219**	-0.304	-0.110	-0.301*	-0.042	-0.482***
	[0.109]	[0.193]	[0.120]	[0.167]	[0.116]	[0.149]
1 if female	0.441**	0.294	0.033	0.146	0.285	0.667**
	[0.210]	[0.351]	[0.230]	[0.326]	[0.229]	[0.270]
Age	0.010	0.008	0.006	-0.013	-0.006	0.025***
	[0.006]	[0.010]	[0.007]	[0.010]	[0.007]	[0.008]
Years of schooling	-0.036	-0.388***	-0.199***	0.130	-0.089	-0.196**
	[0.058]	[0.109]	[0.068]	[0.084]	[0.066]	[0.080]
Constant	-0.498	1.733	1.266	-2.378**	0.553	-0.524
	[0.734]	[1.231]	[0.827]	[1.075]	[0.814]	[0.986]
N. Obs.	802					
Prob χ^2	0.000					
Pseudo-R^2	0.032					

Table B.2: Detailed Multinomial Analysis (I). Multinomial logit estimates, reference category is base rate neglect ($x = 80\%$). Word fluency measure is standardized to have mean 0 and standard deviation 1. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

Table B.3: Detailed Multinomial Analysis (II)

Dependent variable: answer in %, base category is x=80% (base rate neglect)

	[0%,20%]	(20%,40%]	(40%,60%]	(60%,80%)	(80%,100%]	Don't know
Perceptual speed measure	-0.189	-0.424**	-0.277**	-0.130	-0.083	-0.313**
	[0.124]	[0.194]	[0.135]	[0.185]	[0.134]	[0.160]
1 if female	0.326	0.229	-0.049	0.206	0.233	0.706***
	[0.204]	[0.328]	[0.226]	[0.309]	[0.220]	[0.271]
Age	0.008	0.003	-0.000	-0.008	-0.008	0.019**
	[0.007]	[0.011]	[0.008]	[0.011]	[0.008]	[0.009]
Years of schooling	-0.062	-0.396***	-0.202***	0.070	-0.089	-0.241***
	[0.056]	[0.101]	[0.066]	[0.081]	[0.062]	[0.081]
Constant	-0.058	2.145*	1.611**	-1.909*	0.731	0.149
	[0.720]	[1.155]	[0.805]	[1.052]	[0.779]	[1.000]
N. Obs.	850					
Prob χ^2	0.000					
Pseudo-R^2	0.028					

Table B.3: Detailed Multinomial Analysis (II). Multinomial logit estimates, reference category is base rate neglect ($x = 80\%$). Perceptual speed measure is standardized to have mean 0 and standard deviation 1. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

Table B.4: Detailed Multinomial Analysis (III)

Dependent variable: answer in %, base category is x=80% (base rate neglect)

	[0%,20%]	(20%,40%]	(40%,60%]	(60%,80%)	(80%,100%]	Don't know
Word fluency measure	-0.244**	-0.488***	-0.135	-0.248	-0.043	-0.457***
	[0.108]	[0.177]	[0.116]	[0.164]	[0.113]	[0.144]
1 if female	0.435**	0.243	0.008	0.069	0.285	0.559**
	[0.210]	[0.323]	[0.224]	[0.318]	[0.224]	[0.261]
Age	0.011*	0.004	0.011*	-0.014	-0.003	0.030***
	[0.006]	[0.009]	[0.006]	[0.010]	[0.007]	[0.007]
1 if university degree	-0.226	-1.070*	-1.248***	0.210	-0.938**	-1.361***
	[0.285]	[0.630]	[0.411]	[0.423]	[0.379]	[0.503]
Constant	-0.882***	-1.713***	-0.860***	-0.934**	-0.400	-2.537***
	[0.315]	[0.471]	[0.335]	[0.453]	[0.329]	[0.422]
N. Obs.	843					
Prob χ^2	0.000					
Pseudo-R^2	0.029					

Table B.4: Detailed Multinomial Analysis (III). Multinomial logit estimates, reference category is base rate neglect ($x = 80\%$). Word fluency test measures cognitive pragmatics. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

Table B.5: Detailed Multinomial Analysis (IV)

Dependent variable: answer in %, base category is x=80% (base rate neglect)

	[0%,20%]	(20%,40%]	(40%,60%]	(60%,80%)	(80%,100%]	Don't know
Perceptual speed measure	-0.189	-0.616***	-0.251**	-0.046	-0.013	-0.312**
	[0.119]	[0.170]	[0.128]	[0.179]	[0.129]	[0.151]
1 if female	0.347*	0.170	-0.058	0.162	0.259	0.599**
	[0.200]	[0.305]	[0.220]	[0.301]	[0.216]	[0.260]
Age	0.009	-0.003	0.006	-0.007	-0.003	0.025***
	[0.007]	[0.010]	[0.007]	[0.010]	[0.007]	[0.008]
1 if university degree	-0.325	-1.307**	-1.214***	-0.144	-1.018***	-1.478***
	[0.271]	[0.624]	[0.389]	[0.410]	[0.360]	[0.497]
Constant	-0.753**	-1.273***	-0.594	-1.171**	-0.330	-2.330***
	[0.342]	[0.485]	[0.363]	[0.502]	[0.360]	[0.457]
N. Obs.	894					
Prob χ^2	0.000					
Pseudo-R^2	0.027					

Table B.5: Detailed Multinomial Analysis (IV). Multinomial logit estimates. Reference category is base rate neglect ($x = 80\%$). Perceptual speed test measures cognitive mechanics. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

APPENDIX C

Appendix to Chapter 3

C.1. Instructions for Laboratory Experiment

Welcome to today's decision experiment.

To start, please read these instructions carefully. At the end of the instructions you will find some example questions. The experiment starts as soon as all participants have answered these questions correctly.

Please note that it is not allowed to communicate with other participants of the experiment from now on. If this should happen, the experiment loses its scientific value and we have to stop the experiment. If you have any questions, please hold your hand out of the cubicle; we will then come to you.

The experiment consists of two parts. They will be called **work phase** and **shopping phase**. During the work phase you have the possibility to earn talers. You can then use these talers for shopping during the shopping phase. The value your purchases have for you will be denoted in points during the experiment. Directly after the experiment, the points you achieved will be summed up and paid in cash to you according to an exchange rate of

1 point = 0.01 Euro

In addition, you receive **2.50 Euro** for having showed up on time. The 2.50 Euro will be paid after the experiment independently of your decisions and **additionally** to the amount you earn during the experiment.

Work phase

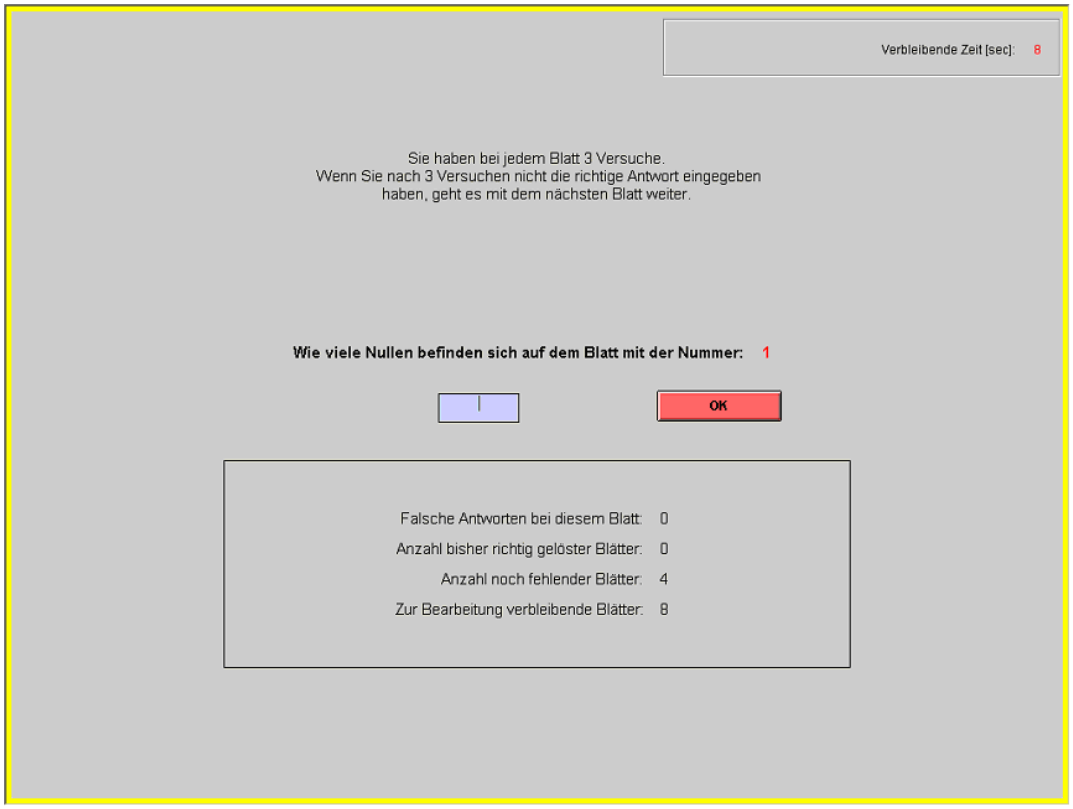
During the **work phase** you have the opportunity to earn 100 talers. The work consists of counting the number of zeros in tables filled with zeros and ones. Below, you see an example table with 3 rows and 8 columns. The tables used in the experiment are larger, they contain 10 rows and 30 columns.

Example of work phase

1	**1**	**1**	**0**	**0**	**0**	**1**	**0**
1	**0**	**1**	**0**	**1**	**1**	**0**	**1**
1	**0**	**0**	**0**	**1**	**0**	**1**	**1**

You earn the 100 talers if you succeed in finding the correct number of zeros in four tables within 15 minutes. If you do not succeed in finding the correct number of zeros in four tables you earn 10 talers instead.

Work phase screen



During the work phase, you will receive eight sheets with zeros and ones. Please begin on sheet 1 and count the number of zeros on this sheet. Enter the number of zeros in the input box in the middle of the computer screen. After entering the number click on the OK-button. If you entered the correct number, you may continue with sheet 2. If you entered a number that is higher by 1 or lower by 1 than the correct number, your number will also be rated as correct. If you enter a number that deviates by more than plus/minus 1 from the correct number, your input will be rated as false. You then have another two tries to enter the correct number for this sheet. Thus, you have three tries in total for each sheet. In the top-right hand corner of the screen, you can see the remaining time in seconds. The time starts at 900 seconds = 15 minutes and counts backwards.

Please note: the **red number** above the OK-button indicates the number of the current sheet. If you enter three times a wrong number for a sheet, the counter for the current sheet changes to the next sheet. If this occurs, please put the current sheet aside and start the next one.

You have a total of eight sheets at your disposal. As soon as you found the correct number of zeros on four sheets, the task is completed successfully and you receive 100 talers. **You then have finished the work phase.** If you do not succeed in completing the task within 15 minutes, you earn 10 talers instead.

Please note: Experience shows that is helpful to mark the 50th, 100th... counted zero. If you miscount in this case you do not have to start all over again but you can continue from the last marked zero.

Shopping phase

The **shopping phase** starts as soon as it has been determined for every participant if he or she completed the task of the work phase successfully. You will make **two** shopping decisions. Your credit balance is split equally between the two decisions. If you completed the task of the work phase successfully you have $100/2 = 50$ talers at your disposal per purchasing decision, otherwise you have $10/2 = 5$ talers.

During the shopping phase you can spend your money on two things that will be called **housing** and **clothing**. You decide which amount of housing and clothing you want to buy. Expenses for housing denote the rent of the apartment.

The value housing and clothing have for you are expressed in points that are exchanged into Euro at the end of the experiment and paid out to you. How valuable a specific amount of housing or clothing is for you is denoted in two tables during the experiment. Below you see an example. In this example numbers of points and prices take on **different values** than in the experiment. The sole purpose of this example is to help you become familiar with the procedure of the purchasing decision.

Example of shopping phase

Housing	
Units	Points
0	0
1	6
2	11
3	15
4	18
5	20

Clothing	
Units	Points
0	0
1	16
2	24
3	27
4	29
5	30

Your credit balance
20 talers

Prices per unit
Housing: 4 talers
Clothing: 3 talers

In the left column of each table, the different amounts that are offered for sale are presented. The right column indicates how many points you get for the purchase of the corresponding amount. You can read from the table "Housing" that in this example 0 units of housing have a value of 0 points for you, 1 unit of housing has a value of 6 points, 2 units 11 points, and so on.

Your credit balance for the purchase is indicated in the top-right panel; in this example 20 talers. In the bottom-right panel you find the prices (in talers) for housing and clothing; prices are per unit. The **prices** for housing and clothing are **different**. The table "Prices per unit" shows that in this example a unit of housing costs 4 talers while clothing costs 3 talers per unit.

In the purchasing decision, you decide how many units of housing and how many units of clothing you want to buy. You can choose freely how many units to buy as long as the total price does not exceed your credit balance.

The **total price of your purchase** is calculated as follows:

Total price of purchase =	(units of housing × price per unit of housing)
	+ (units of clothing × price per unit of clothing)

As soon as you have decided how many units of housing and how many units of clothing to buy, it is determined how many points you will get for this decision. If you do not spend your entire credit balance, **the talers not spent are forfeited**. Additionally, talers from the first purchasing decision cannot be kept for the second purchasing decision.

The **total number of points** is calculated as follows:

Total number of points =	points for purchased units of housing
	+ points for purchased units of clothing

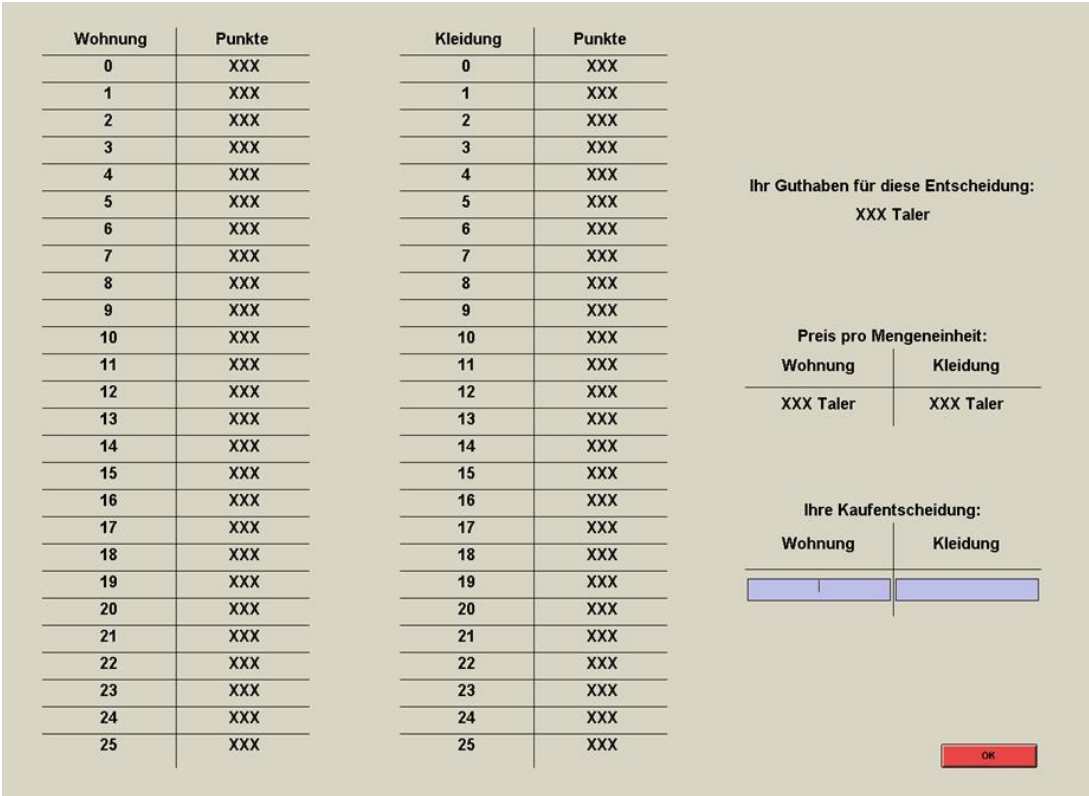
Example of a purchase

In the example mentioned above, you have a credit balance of 20 talers. Imagine you wanted to buy 3 units of housing and 2 units of clothing. Then you have to pay [(3 × price per unit of housing) + (2 × price per unit of clothing)] talers, i.e., $12 + 6 = 18$ talers. This purchase is possible with your credit balance.
In the tables, you find the number of points you get for this purchase. You get **15 points** for 3 units of housing and **24 points** for 2 units of clothing. Your purchase would thus earn you $15 + 24 = 39$ **points**

Please note: It is only possible to buy **one** amount of each good. For example, if you want to buy altogether 4 units of clothing, the point value that is noted next to the number 4 (29 points) matters for you. You cannot buy first one unit of clothing and then another 3 units of clothing, for example.

On the computer, you make your decisions on the input screen of the shopping phase. Below you see a screen shot of this input screen. The screen contains all information that you need for your decision: tables for the point values of housing and clothing, your credit balance and the prices per unit. The actual point values and prices used in the experiment have been replaced with "XXX".

Shopping phase screen



Wohnung	Punkte
0	XXX
1	XXX
2	XXX
3	XXX
4	XXX
5	XXX
6	XXX
7	XXX
8	XXX
9	XXX
10	XXX
11	XXX
12	XXX
13	XXX
14	XXX
15	XXX
16	XXX
17	XXX
18	XXX
19	XXX
20	XXX
21	XXX
22	XXX
23	XXX
24	XXX
25	XXX

Kleidung	Punkte
0	XXX
1	XXX
2	XXX
3	XXX
4	XXX
5	XXX
6	XXX
7	XXX
8	XXX
9	XXX
10	XXX
11	XXX
12	XXX
13	XXX
14	XXX
15	XXX
16	XXX
17	XXX
18	XXX
19	XXX
20	XXX
21	XXX
22	XXX
23	XXX
24	XXX
25	XXX

In the bottom-right hand corner of the screen, you can see two input fields. After having decided how many units of housing and of clothing to buy you enter your decision in these two fields and confirm your choice by clicking on the OK-button. **After having clicked on the OK-button you cannot change your decision anymore.** Your decision will be shown again on the screen. Please write your decision on the decision sheet that was handed out with these instructions. If you click on the OK-button although you would spend more talers than you have at your disposal, an error message is displayed and you have the possibility to correct your decision.

If you have any questions please hold your hand out of the cubicle; we will then come to you.

When all participants have answered the example questions correctly, the experiment starts with the working phase. When all participants have finished the working phase, you

will be presented again short instructions for the first purchasing decision on the computer screen. Also for the second purchasing decision, the screen will show short instructions. As soon as all participants have taken the second purchasing decision the computer screen shows a questionnaire. After the questionnaire, the experiment is over.

Please answer the example questions handed out with these instructions before the experiment starts.

On-screen Instructions

Before the Working Phase. The working phase is about to start now. If you succeed in counting the correct number of zeros on four sheets within 15 minutes, you have completed the task successfully and you get 100 talers. If you do not succeed in completing the task successfully you get 10 talers instead.

Please click on the OK-button to start the working phase.

Before the First Purchasing Decision. You completed the task successfully. Your credit balance per purchasing decision is thus 50 talers.

In the following shopping phase you will make **two** purchasing decisions.

You decide how many units of housing and how many units of clothing to buy. You can read from the tables on the screen how many points you will get for your decision. If you do not spend all your credit balance, the talers not spent will be forfeited.

Before the Second Purchasing Decision. Label treatment

For the second purchasing decision, you get a **housing subsidy** of **30 talers** in addition to your credit balance of 50 talers. You can spend the housing subsidy **only on housing**.

If the amount you spend on housing is lower than the amount of the housing subsidy, i.e., lower than 30 talers, the part of the subsidy that is not spent is **forfeited**.

The **housing subsidy** is the **only difference** compared to the first purchasing decision. All prices and point values remain the same.

Please note: When entering your purchasing decision, please report the **total** number of units you buy, no matter whether you paid them out of your **own credit balance** or out of the **housing subsidy**.

Cash treatment

For the second purchasing decision, you get a **subsidy** of **30 talers** in addition to your credit balance of 50 talers. You can spend the subsidy on housing, on clothing or on both.

If you do not spend the whole subsidy, the part of it that is not spent is **forfeited**.

The **subsidy** is the **only difference** compared to the first purchasing decision. All prices and point values remain the same.

Please note: When entering your purchasing decision, please report the **total** number of units you buy, no matter whether you paid them out of your **own credit balance** or out of the **subsidy**.

Moral Obligation Scenarios

Student-Support Scenario.

> Mr Smith is a first-year Biology student who wants to apply for Bafög.[1] When he reads up on Bafög he notices that he has to specify the income of his parents and additionally his own wealth. He recently received part of his bequest, amounting to 32 000 Euro. If he declares this amount his application will be rejected. He decides to not declare the bequest in his application in order to receive Bafög anyway.

What do you think about the behavior of Mr Smith? 1: "Not justified at all", 6: "Fully justified"

Unemployment-Benefit Scenario.

> Ms Newman has finished her studies of Law and is looking for a job. She has already found one but this position is only available in three months. She knows that she is eligible for unemployment benefit. She could easily bridge the time until the job starts since she has savings of 10 000 Euro. Additionally, her parents support her with 800 Euro per month until the new job starts. Ms Newman decides to claim unemployment benefit in addition, amounting to 300 Euro per month.

What do you think about the behavior of Ms Newman? 1: "Not justified at all", 6: "Fully justified"

C.2. Wording of Vignette Survey

Hello!

We are conducting a survey as part of our doctoral research. Please imagine to be in the following situation and answer the two questions.

Suppose you have finished your studies and are about to start postgraduate studies in a new city. The university provides you with a monthly scholarship of 900 Euro. This is your only income. You have decided to live on your own. You find two apartments between which you can choose. Both are in the same building and are of similar standard. Apartment A has 25 sqm and the rent is 250 Euro per month. Apartment B has 37 sqm

[1] "Bafög" is the student support provided by the state in Germany. The amount depends on own income, own wealth and parents' income.

and the rent is 410 Euro per month. Which apartment would you rather choose in the financial situation described above?

☐ Apartment A ☐ Apartment B

[Cash treatment]: Now suppose that the city administration provides you with a monthly scholarship of 200 Euro, in addition to your university scholarship of 900 Euro. Which apartment would you rather choose in this new situation?

[Label treatment]: Now suppose that the city administration provides you with a monthly housing subsidy of 200 Euro, in addition to your university scholarship of 900 Euro. This housing subsidy has to be spent exclusively on rent expenditures. Which apartment would you rather choose in this new situation?

☐ Apartment A ☐ Apartment B

Field of Study: .....

Age: .....

☐ Female ☐ Male

Thank you!

APPENDIX D

Appendix to Chapter 4

Fragebogen zur Motivation ehrenamtlicher Arbeit

Sehr geehrte Senior Expertinnen und Experten,

zunächst noch einmal vielen Dank für Ihre Bereitschaft, an dieser Umfrage zur Motivation ehrenamtlicher Arbeit teilzunehmen.

Dieser Fragebogen ist in drei Teile gegliedert. Der erste Teil beschäftigt sich mit Angaben zu Ihrer Person und Ihren Ansichten. Im zweiten Teil möchten wir gerne einige Fragen zu Ihrer Tätigkeit als Senior Experte an sich stellen. Im dritten Teil geht es schließlich um verschiedene Faktoren, die die Motivation ehrenamtlicher Arbeit beeinflussen können.

Die Beantwortung aller Fragen nimmt etwa 20 Minuten in Anspruch. Wir möchten Sie an dieser Stelle noch einmal darauf hinweisen, dass all Ihre Angaben selbstverständlich anonym behandelt werden.

Bitte senden Sie den Fragebogen mit Hilfe des beiliegenden Rückumschlags bis zum 15.12.2004 an das IZA zurück.

Und nun viel Spaß beim Beantworten der Fragen!

Teil 1: Angaben zu Ihrer Person

Bitte kreuzen Sie bei den Fragen des ersten Teils die jeweils zutreffende Antwortalternative an oder schreiben Sie die entsprechende Angabe in das dafür vorhergesehene Feld.

1. Ihr Geschlecht und Geburtsjahr:

Männlich ☐
Weiblich ☐

Geburtsjahr: ☐

2. Wie ist Ihr Familienstand?

Verheiratet, mit dem Ehepartner zusammenlebend ☐
Verheiratet, dauernd getrennt lebend ☐
Ledig, in einer festen Partnerschaft lebend ☐
Ledig, nicht in einer festen Partnerschaft lebend ☐
Geschieden ☐
Verwitwet ☐

3. Über welchen Schulabschluss verfügen Sie?

Hauptschulabschluss ☐
Realschulabschluss ☐
Fachhochschulreife ☐
Abitur bzw. Hochschulreife ☐
Sonstiger Schulabschluss ☐

4. Welche Ausbildung haben Sie?

Lehre ☐
Berufsfachschule, Handelsschule, Schule des Gesundheitswesens ☐
Fachschule, z.B. Meister-, Technikerabschluss ☐
Beamtenausbildung ☐
Fachhochschule, Berufsakademie ☐
Universitäts-, Hochschulabschluss ☐
sonstiger Abschluss ☐

5. Wie alt waren Sie zum Zeitpunkt Ihrer Pensionierung?
Falls Sie noch arbeiten geben Sie bitte an, wie alt Sie zum Zeitpunkt Ihrer Registrierung als Senior Experte waren.

☐

6. In welcher beruflichen Stellung waren Sie vor Ihrer Pensionierung zuletzt beschäftigt (bzw. sind Sie beschäftigt, falls Sie noch erwerbstätig sind)?
*Wenn Sie mehr als **eine** berufliche Tätigkeit ausüben/ausübten, beantworten Sie diese Frage bitte nur für Ihre berufliche **Haupt**tätigkeit*

Arbeiter
(auch in der Landwirtschaft)

- Ungelernte Arbeiter ☐
- Angelernte Arbeiter ☐
- Gelernte und Facharbeiter ☐
- Vorarbeiter, Kolonnenführer ☐
- Meister, Polier ☐

Angestellte

- Industrie- und Werkmeister im Angestelltenverhältnis ☐
- Angestellte mit einfacher Tätigkeit
  - ohne Ausbildungsabschluss ☐
  - mit Ausbildungsabschluss ☐
- Angestellte mit qualifizierter Tätigkeit (z.B. Sachbearbeiter, Buchhalter, techn. Zeichner) ☐
- Angestellte mit hochqualifizierter Tätigkeit oder Leitungsfunktion (z.B. wissenschaftliche Mitarbeiter, Ingeneure, Abteilungsleiter) ☐
- Angestellte mit umfassenden Führungsaufgaben (z.B. Direktor, Geschäftsführer, Vorstand größerer Betriebe und Verbände) ☐

Selbständige
(einschl. mithelfende Familienangehörige)

	Anzahl Mitarbeiter 0	1-9	≥10
Selbständige Landwirte	☐	☐	☐
Freie Berufe, selbständige Akademiker	☐	☐	☐
Sonstige Selbständige	☐	☐	☐
Mithelfende Familienangehörige	☐	☐	☐

Beamten / Staatsverwaltung
(einschl. Richter und Berufssoldaten)

- Einfacher Dienst ☐
- Mittlerer Dienst ☐
- Gehobener Dienst ☐
- Höherer Dienst ☐

7. Wenn man alle Einkünfte zusammenzählt, wie hoch ist das monatliche Einkommen <u>Ihres Haushaltes</u> derzeitig?
Bitte geben Sie den monatlichen Nettobetrag an, also nach Abzug von Steuern und Sozialabgaben. Einkünfte Ihres Ehepartners oder im Haushalt lebender Kinder sowie regelmäßige Zahlungen wie Renten, Wohngeld, Kindergeld, Unterhaltszahlungen etc. rechnen Sie bitte dazu.
Falls nicht genau bekannt, bitte schätzen Sie den monatlichen Betrag.

☐

8. Wenn man alle Einkünfte zusammenzählt, wie hoch ist Ihr monatliches <u>individuelles</u> Nettoeinkommen derzeitig?
Falls nicht genau bekannt, bitte schätzen Sie den monatlichen Betrag.

☐

9. Waren Sie früher bereits beruflich im Ausland tätig?

Ja ☐ Nein ☐

10. Sind Sie zur Zeit beim Arbeitsamt arbeitslos gemeldet?

Ja ☐ Nein ☐

11. Falls Sie noch arbeiten, nehmen Sie zur Zeit eine der gesetzlichen Regelungen zur Altersteilzeit wahr?
Bitte lassen Sie diese Frage einfach unbeantwortet, falls Sie nicht mehr arbeiten.

Ja ☐ → **Und zwar** a) reduzierte Arbeitszeit ☐
b) reduziertes Gehalt bei gleich bleibender Arbeitszeit ☐
Nein ☐

12. Verschiedenen Menschen sind verschiedene Dinge wichtig. Sind für Sie persönlich die folgenden Dinge
- **sehr wichtig**
- **wichtig**
- **weniger wichtig**
- oder **nicht wichtig?**

	Sehr Wichtig	Wichtig	Weniger Wichtig	Nicht Wichtig
Sich etwas leisten können	☐	☐	☐	☐
Für andere da sein	☐	☐	☐	☐
Sich selbst verwirklichen	☐	☐	☐	☐
Erfolg im Beruf haben	☐	☐	☐	☐
Ein eigenes Haus haben	☐	☐	☐	☐
Eine glückliche Ehe/Partnerschaft haben	☐	☐	☐	☐
Kinder haben	☐	☐	☐	☐
Sich politisch, gesellschaftlich einzusetzen	☐	☐	☐	☐
Die Welt sehen, viele Reisen machen	☐	☐	☐	☐
Anerkennung finden	☐	☐	☐	☐

13. Wie viele Kinder haben Sie?

- 0 ☐
- 1 bis 2 ☐
- 3 bis 4 ☐
- 5 bis 6 ☐
- 7 und mehr ☐

14. Wie alt sind Ihre Kinder? (mehrere Antworten sind möglich)
Bitte springen Sie auf Frage 18, falls Sie keine Kinder haben.

- ☐ 0 bis 15 Jahre
- ☐ 16 bis 30 Jahre
- ☐ 31 bis 45 Jahre
- ☐ 46 und mehr Jahre

15. Wie weit wohnt das am nächsten lebende Kind von Ihnen entfernt?

- ☐ Im gleichen Haus
- ☐ Unter 5 km entfernt
- ☐ 5 bis unter 20 km entfernt
- ☐ 20 bis unter 50 km entfernt
- ☐ 50 bis unter 100 km entfernt
- ☐ 100 km und mehr entfernt

16. Wie würden Sie das Verhältnis zu Ihren Kindern beschreiben?

- ☐ Sehr gut
- ☐ Gut
- ☐ Mittelmäßig
- ☐ Weniger gut
- ☐ Schlecht

17. Haben Sie Enkelkinder?

Ja ☐ Nein ☐

Wenn ja, wie viele? ☐

18. Wie viele Stunden pro Woche verbringen Sie durchschnittlich mit Ihrer Familie (gegenseitiges Besuchen, gemeinsame Aktivitäten, etc.)?

- ☐ Unter 5 Stunden
- ☐ 5 bis unter 10 Stunden
- ☐ 10 bis unter 15 Stunden
- ☐ 15 Stunden und mehr

19. Wie eng ist der Kontakt zu Ihren Nachbarn bei Ihnen im Haus oder Wohngebiet?

- ☐ Sehr eng
- ☐ Eng
- ☐ Mittel
- ☐ Nur flüchtig
- ☐ Gar kein Kontakt

20. Wie würden Sie Ihren derzeitigen Gesundheitszustand beschreiben?

- ☐ Sehr gut
- ☐ Gut
- ☐ Zufriedenstellend
- ☐ Weniger gut
- ☐ Schlecht

21. Haben Sie in den letzten 3 Monaten Ärzte aufgesucht? Wenn ja, bitte geben Sie an wie häufig.

Zahl der Arztbesuche in den letzten 3 Monaten ☐

Keinen Arzt in Anspruch genommen ☐

22. Wie schätzen Sie sich persönlich ein:
Sind Sie im Allgemeinen ein risikobereiter Mensch oder versuchen Sie, Risiken zu vermeiden?
Bitte kreuzen Sie ein Kästchen auf der Skala an, wobei der Wert 0 bedeutet: „gar nicht risikobereit" und der Wert 10. „sehr risikobereit".
Mit den Werten dazwischen können Sie Ihre Einschätzung abstufen.

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10
Gar nicht risikobereit										Sehr risikobereit

23. Noch eine Frage zur Risikobereitschaft.
Überlegen Sie bitte, was Sie in folgender Situation tun würden: Stellen Sie sich vor, dass Sie in einer Lotterie 100.000 Euro gewinnen. Unmittelbar nach Erhalt des Gewinns bekommen Sie von einer angesehenen Bank ein Angebot für eine Geldanlage, die folgendes beinhaltet:
Es gibt die Chance, das Geld innerhalb von zwei Jahren zu verdoppeln.
Es gibt aber auch ein gleich hohes Risiko, die Hälfte des eingesetzten Geldes zu verlieren.
Sie können das Geld ganz oder teilweise in folgender Weise anlegen oder das Angebot ablehnen. Welchen Teil des Lotteriegewinnes würden Sie für die einerseits riskante, andererseits Gewinn versprechende Geldanlage einsetzen?

- ☐ Den ganzen Betrag von 100.000 €
- ☐ Den Betrag von 60.000 €
- ☐ Den Betrag von 40.000 €
- ☐ Den Betrag von 20.000 €
- ☐ Überhaupt nichts, würde das Angebot ablehnen

24. Wie ist es mit den folgenden Gebieten – machen Sie sich da Sorgen?

	Große Sorgen	Einige Sorgen	Keine Sorgen
Um die allgemeine wirtschaftliche Situation?	☐	☐	☐
um die eigene wirtschaftliche Situation?	☐	☐	☐
um Ihre Gesundheit?	☐	☐	☐
Um den Schutz der Umwelt?	☐	☐	☐
Um die Erhaltung des Friedens?	☐	☐	☐
Über steigende Kriminalität in Deutschland?	☐	☐	☐
Über die Folgen der EU-Erweiterung?	☐	☐	☐
Über die Zuwanderung nach Deutschland?	☐	☐	☐
Über Ausländerfeindlichkeit in Deutschland?	☐	☐	☐

Oder was sonst macht Ihnen Sorgen? (Bitte angeben)

25. Wie wurden Sie auf den SES aufmerksam?

- ☐ Durch die Presse (Zeitungen/Fernsehen)
- ☐ Durch andere Experten aus dem Bekanntenkreis / der Familie
- ☐ Durch das Arbeitsamt
- ☐ Durch meinen alten Arbeitgeber

Sonstiges (Bitte angeben)

Teil 2: Ihre Tätigkeit als Senior Experte

Wir möchten gerne zunächst einmal wissen, inwieweit die folgenden Aussagen zur Tätigkeit als Senior Experte auf Sie zutreffen. Bitte kreuzen Sie bei diesen Aussagen jeweils ein Kästchen auf der Skala an, wobei der Wert 1 bedeutet „trifft überhaupt nicht auf mich zu“ und der Wert 5 „trifft völlig auf mich zu“.
Mit den Werten dazwischen können Sie Ihre Einschätzung abstufen.

26. Ich denke, es ist moralisch richtig und wichtig, sich um das Wohl seiner Mitmenschen zu kümmern und habe mich daher entschieden, ehrenamtlich tätig zu werden.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Trifft nicht zu — Trifft völlig zu

27. Ich erzähle gerne, auch in der Presse oder im Rahmen der Öffentlichkeitsarbeit des SES von meiner Arbeit als Senior Experte.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Trifft nicht zu — Trifft völlig zu

28. Ich möchte nach meiner Pensionierung nicht einfach einen Schlussstrich unter mein Berufsleben ziehen, sondern mich weiterhin mit meiner alten Tätigkeit beschäftigen.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Trifft nicht zu — Trifft völlig zu

29. Es ist sehr wichtig und motivierend zu sehen, dass meine Arbeit als Senior Experte vor Ort auch tatsächlich Früchte trägt und bestehende Probleme zumindest in Teilen gelöst wurden.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Trifft nicht zu — Trifft völlig zu

30. Ich mag meinen alten Beruf sehr und schätze daher die Möglichkeit, als Senior Experte weiterhin in meinem Beruf tätig sein zu können.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Trifft nicht zu — Trifft völlig zu

31. Ich finde die Möglichkeit, im Rahmen eines SES-Einsatzes in die verschiedensten Länder reisen zu können, sehr reizvoll.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Trifft nicht zu — Trifft völlig zu

32. Ich möchte gerne weiterhin beruflich aktiv sein. Ich sehe die Tätigkeit als Senior Experte als Möglichkeit, nicht den Anschluss an neue Entwicklungen in meinem Beruf zu verlieren und nützliche Kontakte knüpfen zu können.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Trifft nicht zu — Trifft völlig zu

33. Meine lebenslange Berufserfahrung soll nach meiner Pensionierung nicht plötzlich unwichtig werden. Ich freue mich, als Senior Experte die Möglichkeit zu haben, mein Wissen sinnvoll weiterzugeben.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Trifft nicht zu — Trifft völlig zu

34. Ich fühle mich eigentlich noch zu aktiv für den Ruhestand. Die Tätigkeit als Senior Experte bietet mir die Möglichkeit, meine Zeit sinnvoll zu nutzen.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Trifft nicht zu — Trifft völlig zu

Die folgende Art von Fragen kennen Sie schon aus dem ersten Teil. Bitte kreuzen Sie bei jeder Frage die jeweils zutreffende Antwort an.

35. Wie ist die Meinung Ihres Umfeldes (Familie, Freunde oder andere Ihnen nahe stehende Personen) von Ihrer Tätigkeit als Senior Experte?

- Völlige Zustimmung ☐
- Eher Zustimmung ☐
- Die Meinungen gehen auseinander ☐
- Eher Ablehnung ☐
- Ablehnung ☐

36. Wenn Sie sich mit Freunden oder Verwandten unterhalten, wie oft sprechen Sie dann über Ihre Tätigkeit als Senior Experte?

- Sehr häufig ☐
- Häufig ☐
- Gelegentlich ☐
- Manchmal ☐
- Kaum ☐

37. Haben Sie in den letzten 5 Jahren vor Ihrer Registrierung als Senior Experte eine andere ehrenamtliche Tätigkeit ausgeübt?

Ja ☐ Nein ☐

38. Stand bzw. steht es für Sie zur Diskussion, auch eine ehrenamtliche Tätigkeit in Ihrem Wohnort oder nahe Ihrem Wohnort auszuüben (beispielsweise im kirchlichen oder kulturellen Bereich)?

- Ich übe parallel eine solche Tätigkeit aus ☐
- Ich habe darüber nachgedacht, mich aber dagegen entschieden ☐
- Es gibt kein solches Angebot in meinem Wohnort ☐
- Eine solche Tätigkeit stand bzw. steht für mich nicht zur Diskussion ☐

39. Gab es zum Zeitpunkt Ihrer Registrierung als Senior Experte jemanden in Ihrem Bekanntenkreis, der schon für den SES tätig war?

Ja ☐ Nein ☐

40. Bitte rechen Sie einmal die Zeit, die Sie in den vergangenen 12 Monaten auf Einsätze sowie deren Vor- und Nachbereitung verwendet haben, auf einen wöchentlichen Durchschnittswert herunter. Wie viele Stunden pro Woche sind Sie dann durchschnittlich in den letzten 12 Monaten für den SES tätig gewesen?

- Unter 10 Stunden ☐
- 10 bis unter 20 Stunden ☐
- 20 bis unter 30 Stunden ☐
- 30 Stunden und mehr ☐

41. Wie viel Geld haben Sie im Jahr 2003 für gemeinnützige Zwecke gespendet?

- unter 50 € ☐
- 50 bis unter 200 € ☐
- 200 € oder mehr ☐

42. Es wurde schon des Öfteren in Zeitungs- oder auch Fernsehberichten über die Arbeit von Senior Experten berichtet. Fänden Sie es gut, wenn auch von einem Ihrer Einsätze in dieser Form berichtet wird?

- Ja ☐
- Nein ☐
- Es spielt keine Rolle für mich ☐
- Es wurde schon in den Medien über meine Arbeit berichtet ☐

Teil 3: Mögliche Einflüsse auf die Motivation

Bitte kreuzen Sie bei den folgenden, hypothetischen Fragen wieder die jeweils zutreffende Antwort an.

In seiner Arbeit ist der SES auf Ihre Bereitschaft ehrenamtlich zu arbeiten angewiesen. Stellen Sie sich nun einmal vor, es könnten die finanziellen Mittel zur Verfügung gestellt werden, um Sie für Ihre Arbeit zu bezahlen.

43. Wie würde Ihre Bereitschaft Einsätze durchzuführen sich im Gegensatz zu jetzt verändern, wenn Ihnen pro Einsatzwoche 100 € bezahlt würden?

- Ich würde in jedem Fall mehr Einsätze durchführen wollen ☐
- Ich würde eventuell mehr Einsätze durchführen wollen ☐
- Meine Bereitschaft wäre unverändert ☐
- Meine Bereitschaft würde eher sinken ☐
- Ich denke nicht, dass ich unter diesen Umständen SES-Einsätze durchführen würde ☐

44. Wie würde Ihre Bereitschaft Einsätze durchzuführen sich im Gegensatz zu jetzt verändern, wenn Ihnen pro Einsatzwoche 500 € bezahlt würden?

- Ich würde in jedem Fall mehr Einsätze durchführen wollen ☐
- Ich würde eventuell mehr Einsätze durchführen wollen ☐
- Meine Bereitschaft wäre unverändert ☐
- Meine Bereitschaft würde eher sinken ☐
- Ich denke nicht, dass ich unter diesen Umständen SES-Einsätze durchführen würde ☐

45. Wie würde Ihre Bereitschaft Einsätze durchzuführen sich im Gegensatz zu jetzt verändern, wenn Ihnen pro Einsatzwoche 1000 € bezahlt würden?

- Ich würde in jedem Fall mehr Einsätze durchführen wollen ☐
- Ich würde eventuell mehr Einsätze durchführen wollen ☐
- Meine Bereitschaft wäre unverändert ☐
- Meine Bereitschaft würde eher sinken ☐
- Ich denke nicht, dass ich unter diesen Umständen SES-Einsätze durchführen würde ☐

46. Wie würde Ihre Einsatzbereitschaft sich im Gegensatz zu jetzt verändern, wenn der Erfolg Ihrer Einsätze stärker durch die Zentrale kontrolliert würde?

- Ich würde in jedem Fall mehr Einsätze durchführen wollen ☐
- Ich würde eventuell mehr Einsätze durchführen wollen ☐
- Meine Bereitschaft wäre unverändert ☐
- Meine Bereitschaft würde eher sinken ☐
- Ich denke nicht, dass ich unter diesen Umständen SES-Einsätze durchführen würde ☐

Zum Abschluss dieses Fragebogens möchten wir Sie gerne noch nach Ihrer Zufriedenheit mit Ihrem Leben insgesamt fragen. Antworten Sie bitte anhand der folgenden Skala, bei der 0 „ganz und gar unzufrieden“ und 10 „ganz und gar zufrieden“ bedeutet

47. Wie zufrieden sind Sie gegenwärtig, alles in allem, mit Ihrem Leben?

☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10
ganz und gar unzufrieden										Ganz und gar zufrieden

Vielen herzlichen Dank für Ihre Unterstützung!

Falls Sie noch Anmerkungen haben, so können Sie diese hier aufschreiben.

Bibliography

Abeler, J., and F. Marklein (2008): "Fungibility, Labels, and Consumption," *IZA Discussion Paper No. 3500*.

Ariely, D., A. Bracha, and S. Meier (2007): "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially," *IZA Discussion Paper No. 2968*.

Arkes, H. R., C. A. Joyner, M. V. Pezzo, J. Nash, K. Siegel-Jacobs, and E. Stone (1994): "The psychology of windfall gains," *Organizational behavior and human decision processes*, 59(3), 331–347.

Arrow, K. J. (1972): "Gifts and Exchanges," *Philosophy and Public Affairs*, 1(4), 343–362.

Ayton, P., and I. Fischer (2004): "The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness?," *Memory & Cognition*, 32(8), 1369–1378.

Baltes, P. B., U. Lindenberger, and U. M. Staudinger (2006): "Life span theory in developmental psychology," in *Handbook of child psychology: Vol. 1. Theoretical models of human development (6th ed.)*, ed. by W. Damon, and R. M. Lerner, pp. 569–664. Wiley, New York.

Bar-Hillel, M. (1980): "The Base-Rate Fallacy in Probability Judgments," *Acta Psychologica*, 44(3), 211–233.

Barberis, N., and M. Huang (2001): "Mental Accounting, Loss Aversion, and Individual Stock Returns," *Journal of Finance*, 56(4), 1247–1292.

Barberis, N., M. Huang, and T. Santos (2001): "Prospect Theory and Asset Prices," *Quarterly Journal of Economics*, 116(1), 1–53.

Barberis, N., M. Huang, and R. Thaler (2006): "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing," *American Economic Review*, 96(4), 1069–1090.

Benartzi, S., and R. H. Thaler (1995): "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110(1), 73–92.

Benjamin, D. J., S. A. Brown, and J. M. Shapiro (2006): "Who is 'Behavioral'? Cognitive Ability and Anomalous Preferences," *Harvard University Working Paper*.

Bernheim, B. D., and A. Rangel (2007): "Behavioral Public Economics: Welfare and Policy Analysis with Non-Standard Decision-Makers," in *Behavioral Economics and its Applications*, ed. by P. Diamond, and H. Vartiainen. Princeton University Press.

Bertrand, M., S. Mullainathan, and E. Shafir (2004): "A Behavioral-Economics View of Poverty," 94, 419–423.

Blow, L., I. Walker, and Y. Zhu (2004): "Who Benefits from Child Benefit?," *Working Paper, University of Warwick*.

Bénabou, R., and J. Tirole (2003): "Intrinsic and Extrinsic Motivation," *Review of Economic Studies*, 70(3), 489–520.

——— (2006): "Incentives and Prosocial Behavior," *American Economic Review*, 96(5), 1652–1678.

Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue, and M. Rabin (2003): "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism'," *University of Pennsylvania Law Review*, 151(3), 1211–1254.

Carpenter, J., and C. Myers (2007): "Why Volunteer? Evidence on the Role of Altruism, Reputation, and Incentives," *IZA Discussion Paper No. 3021*.

Casari, M., J. C. Ham, and J. H. Kagel (forthcoming): "Selection Bias, Demographic Effects and Ability Effects in Common Value Auction Experiments," *American Economic Review*.

Casscells, W., A. Schoenberger, and T. Graboys (1978): "Interpretation by Physicians of Clinical Laboratory Results.," *New England Journal of Medicine*, 299(18), 999–1001.

Cattell, R. B. (1963): "Theory of Fluid and Crystallized Intelligence: A Critical Experiment," *Journal of Educational Psychology*, 54(1), 1–22.

Charness, G., and D. Levin (2005): "When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect," *American Economic Review*, 95, 1300–1309.

Cherry, T. L., P. Frykblom, and J. F. Shogren (2002): "Hardnose the Dictator," *American Economic Review*, 92(4), 1218–1221.

Clotfelter, C. T., and P. J. Cook (1993): "The 'Gambler's Fallacy' in Lottery Play," *Management Science*, 39, 1521–1525.

Conlisk, J. (1996): "Why Bounded Rationality?," *Journal of Economic Literature*, 34, 669–700.

Cosmides, L., and J. Tooby (1996): "Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions From the Literature on Judgment Under Uncertainty," *Cognition*, 58(1), 1–73.

Croson, R., and J. Sundali (2005): "The Gambler's Fallacy and the Hot Hand: Empirical Data from Casinos," *Journal of Risk and Uncertainty*, 30, 195–209.

Deci, E. L. (1971): "Effects of externally mediated rewards on intrinsic motivation," *Journal of Personality and Social Psychology*, 18(1), 105–115.

Deci, E. L., R. Koestner, and R. M. Ryan (1999): "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation," *Psychological Bulletin*, 125(6), 627–68.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2007): "Are Risk Aversion and Impatience Related to Cognitive Ability?," *IZA Discussion Paper No. 2735*.

Eddy, D. M. (1982): "Probabilistic Reasoning in Clinical Medicine: Problems and Opportunities," in *Judgment Under Uncertainty: Heuristics and Biases*, ed. by D. Kahneman, P. Slovic, and A. Tversky, pp. 249–267. Cambridge University Press, Cambridge.

Ellingsen, T., and M. Johannesson (2007): "Pride and Prejudice: The Human Side of Incentive Theory," *Working Paper, Stockholm School of Economics*.

Epley, N., D. Mak, and L. C. Idson (2006): "Bonus or Rebate?: The Impact of Income Framing on Spending and Saving," *Journal of Behavioral Decision Making*, 19, 213–227.

Fack, G. (2006): "Are housing benefits an effective way to redistribute income? Evidence from a natural experiment in France," *Labour Economics*, 13(6), 747–771.

Falk, A., D. Huffman, and U. Sunde (2006a): "Do I Have What it Takes? Equilibrium Search with Type Uncertainty and Non-Participation," *IZA Discussion Paper No. 2531*.

——— (2006b): "Self-Confidence and Search," *IZA Discussion Paper No. 2525*.

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10(2), 171–178.

Fowler, F. J. (2002): *Survey Research Methods*. Sage Publications, London.

Frederick, S. (2005): "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19(4), 25–42.

Freeman, R. B. (1997): "Working for Nothing: The Supply of Volunteer Labor," *Journal of Labor Economics*, 15(1), S140–S166.

Frey, B. S., and R. Jegen (2001): "Motivation Crowding Theory: A Survey of Empirical Evidence," *Journal of Economic Surveys*, 15(5), 589–611.

Gächter, S., and A. Riedl (2005): "Moral Property Rights in Bargaining with Infeasible Claims," *Management Science*, 51(2), 249–263.

Gibbons, S., and A. Manning (2006): "The Incidence of UK Housing Benefit: Evidence from the 1990s Reforms," *Journal of Public Economics*, 90(4), 799–822.

Gigerenzer, G. (2003): *Calculated Risks: How to Know When Numbers Deceive You*. Simon & Schuster, New York.

Gigerenzer, G., and U. Hoffrage (1995): "How to Improve Bayesian Reasoning without Instructions: Frequency Formats," *Psychological Review*, 102(2), 684–704.

Gilovich, T., R. Vallone, and A. Tversky (1985): "The Hot Hand in Basketball: On the Misperception of Random Sequences," *Cognitive Psychology*, 17, 295–314.

Gneezy, U., and J. Potters (1997): "An Experiment on Risk Taking and Evaluation Periods," *Quarterly Journal of Economics*, 112(2), 631–645.

Gneezy, U., and A. Rustichini (2000): "Pay Enough or Don't Pay at All," *Quarterly Journal of Economics*, 115(3), 791–810.

Grether, D. M. (1980): "Bayes Rule as a Descriptive Model: The Representativeness Heuristic," *Quarterly Journal of Economics*, 95(3), 537–557.

Heath, C., and J. Soll (1996): "Mental Budgeting and Consumer Decisions," *Journal of Consumer Research*, 23(1), 40–52.

Hoffrage, U., S. Lindsey, R. Hertwig, and G. Gigerenzer (2000): "Communicating Statistical Information," *Science*, 290(5500), 2261–2262.

Hogarth, R. M. (2005): "The challenge of representative design in psychology and economics," *Journal of Economic Methodology*, 12(2), 253–263.

Isserstedt, W., E. Middendorff, S. Weber, K. Schnitzer, and A. Wolter (2004): "Die wirtschaftliche und soziale Lage der Studierenden in der Bundesrepublik Deutschland 2003," *Report for the Federal Ministry of Education and Research*, http://www.bmbf.de/pub/040714_ soz17hauptbericht.pdf.

Kahneman, D., and D. Lovallo (1993): "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking," *Management Science*, 39(1), 17–31.

Kahneman, D., and A. Tversky (1972): "On Prediction and Judgment," *ORI Research monograph*, 12.

Koehler, J. J. (1996): "The Base Rate Fallacy Reconsidered: Descriptive, Normative and Methodological Challenges," *Behavioral and Brain Sciences*, 19(1), 1–53.

Kooreman, P. (2000): "The Labeling Effect of a Child Benefit System," *American Economic Review*, 90(3), 571–583.

Laferrère, A., and D. Le Blanc (2004): "How do housing allowances affect rents? An empirical analysis of the French case," *Journal of Housing Economics*, 13(1), 36–67.

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112(2), 443–477.

Lang, F. R. (2005): "Erfassung des kognitiven Leistungspotenzials und der 'Big Five' mit Computer-Assisted-Personal-Interviewing (CAPI): Zur Reliabilität und Validität zweier ultrakurzer Tests und des BFI-S," *DIW Berlin, Research Notes, 9/2005*.

Laplace, P. (1820): *Philosophical Essays on Probabilities*. Dover, New York, translated by F.W. Truscott and F. L. Emory, 1951.

Lazear, E. P. (2000): "Performance Pay and Productivity," *American Economic Review*, 90(5), 1346–1361.

Linardi, S., and M. A. McConnell (2008): "Volunteering and Image Concerns," *Caltech Social Science Working Paper No. 1282*.

Lindenberger, U., and P. B. Baltes (1997): "Intellectual Functioning in Old and Very Old Age: Cross-Sectional Results From the Berlin Aging Study," *Psychology and Aging*, 12(3), 410–432.

Mellström, C., and M. Johannesson (2005): "Crowding Out in Blood Donation: Was Titmuss Right?," *Working Paper No. 180, Göteborg University, Department of Economics*.

Modigliani, F., and R. Brumberg (1954): "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data," in *Post-Keynesian Economics*, ed. by K. K. Kurihara. Rutgers University Press, New Brunswick, NJ.

Moffitt, R. (1989): "Estimating the Value of an In-Kind Transfer: The Case of Food Stamps," *Econometrica*, 57, 385–409.

Munro, A. (2005): "The Economics of the Winter Fuel Allowance," *mimeo*.

O'Curry, S. (1997): "Income Source Effects," *DePaul University, Discussion Paper*.

Odean, T. (1998): "Are investors reluctant to realize their losses?," *Journal of Finance*, 53(5), 1775–1798.

Oechssler, J., A. Roider, and P. W. Schmitz (2008): "Cognitive Abilities and Behavioral Biases," *IZA Discussion Paper No. 3481*.

Prelec, D., and G. Loewenstein (1998): "The Red and the Black: Mental Accounting of Savings and Debts," *Marketing Science*, 17(1), 4–28.

Prendergast, C. (1999): "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37(1), 7–63.

Rabin, M. (2002): "Inference by Believers in the Law of Small Numbers," *Quarterly Journal of Economics*, 117(3), 775–816.

Rabin, M., and G. Weizsäcker (2007): "Narrow bracketing and dominated choices," *IZA Discussion Paper*.

Read, D., G. Loewenstein, and M. Rabin (1999): "Choice Bracketing," *Journal of Risk and Uncertainty*, 19:1–3, 171–197.

Rockenbach, B. (2004): "The Behavioral Relevance of Mental Accounting for the Pricing of Financial Options," *Journal of Economic Behavior & Organization*, 53, 513–527.

Ryan, R., and E. Deci (2000): "Intrinsic and extrinsic motivations: Classic definitions and new directions," *Contemporary Educational Psychology*, 25(1), 54–67.

Seabright, P. (2004): "Continous Preferences Can Cause Discontinous Choices: an Application to the Impact of Incentives on Altruism," *IDEI Working Paper 257, Toulouse University*.

SES (2005): "Annual Report 2005," *Senior Expert Service, Bonn*, http://www.ses-bonn.de.

Shefrin, H., and R. Thaler (1988): "The Behavioral Life-Cycle Hypothesis," *Economic Inquiry*, 26, 609–643.

Singer, W. (1995): "Development and plasticity of cortical processing architectures," *Science*, 270, 758–764.

Smith, V., and J. Walker (1993): "Monetary Rewards and Decision Costs," *Economic Inquiry*, 31, 245–261.

Stango, V., and D. Zinman (2007): "How a Cognitive Bias Shapes Competition: Evidence from Consumer Credit Markets," *Working Paper, Dartmouth College*.

Stanovich, K. E., and R. F. West (1998): "Individual Differences in Rational Thought," *Journal of Experimental Psychology: General*, 127(2), 161–188.

——— (2000): "Individual Differences in Reasoning: Implications for the rationality Debate?," *Behavioral and Brain Sciences*, 23, 645–726.

Susin, S. (2002): "Rent vouchers and the price of low-income housing," *Journal of Public Economics*, 83(1), 109–152.

Terrell, D. (1994): "A Test of the Gambler's Fallacy: Evidence from Pari-Mutuel Games," *Journal of Risk and Uncertainty*, 8, 309–317.

Thaler, R. (1985): "Mental Accounting and Consumer Choice," *Marketing Science*, 4, 199–214.

——— (1999): "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12, 183–206.

Thaler, R., A. Tversky, D. Kahneman, and A. Schwartz (1997): "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test," *Quarterly Journal of Economics*, 112(2), 647–661.

Thaler, R. H., and C. R. Sunstein (2003): "Libertarian Paternalism," *American Economic Review*, 93(2), 175–179.

Titmuss, R. M. (1970): *The Gift Relationship: from Human Blood to Social Policy*. London: Allen and Unwin.

Tversky, A., and D. Kahneman (1971): "Belief in the Law of Small Numbers," *Psychological Bulletin*, 2, 105–110.

——— (1981): "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 453–458.

——— (1982): "Evidential Impact of Base Rates," in *Judgment Under Uncertainty: Heuristics and Biases*, ed. by D. Kahneman, P. Slovic, and A. Tversky, pp. 153–160. Cambridge University Press, Cambridge.

Viscusi, W. K., and C. J. O'Connor (1984): "Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers?," *American Economic Review*, 74(5), 942–956.

White, R. J. (2006): "Format Matters in the Mental Accounting of Funds: The Case of Gift Cards and Cash Gifts," *University of Waterloo Discussion Paper*.