JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Doctoral Thesis

Five Essays on Experimental Finance

Submitted to: Justus Liebig University Giessen

Submitted on: November 7, 2023

Author: Florian Gärtner

Supervisors:

Prof. Dr. Christina E. Bannier

Chair of Banking & Finance, Justus Liebig University Giessen

Prof. Dr. Andreas Walter

Chair of Financial Services, Justus Liebig University Giessen

Contents

List of Figures

List of Tables

When I started this intellectual journey back in 2017, I was an outsider to the field of finance. While I have a bachelor's degree in economics, the bulk of my academic career so far had focused on sociology and behavioral sciences, and I had carefully circumvented anything close to finance, accounting, and all these other topics we in Germany sum up as the subject of "Betriebswirtschaftslehre". So when I applied to the newly formed research cluster, which would eventually be renamed into a research network called "Behavioral Finance and Accounting" at the JLU Gießen, I was not at all an obvious choice for one of the positions in that network. It took a leap of faith for the professors who created that cluster to hire someone like me. I want to begin my dissertation by thanking all of them.

First and foremost, I want to thank my supervisor Prof. Dr. Christina Bannier. The incredible degree to which she enabled me to follow my own ideas, research questions, and in particular methods, is something I can only hope for other Ph.D.-candidates to experience because I understand that this can be way different. She supported all my good work, be it in research or teaching, and intercepted the occasional bad idea before it could cause too much harm. Whenever I had questions or problems, I could rely on her advice, and her habit to enable simple, practical solutions.

I can only repeat the same gratitude and praise for my second supervisor, Prof. Dr. Andreas Walter. The thing that I am most grateful for might seem counterintuitive, but it is nevertheless true: Andreas is highly skeptical of experiments, and my work *only* consists of experiments. This made his advice and comments an invaluable source to shape my research designs and interpretations into way stronger versions than they would otherwise have been. If I relied less on his input than Christina's, this is attributed exclusively to the opportunities of the moment.

In addition, I want to thank the other professors who constituted the research network Behavioral and Social Finance and Accounting. Dr. Corinna Ewelt-Knauer, Prof. Dr. Peter Tillmann, Prof. Dr. Peter Winker, and Prof. Dr. Arnt Wöhrmann all shared helpful advice and guidance, in particular in the early stages of my doctorate. I am also deeply grateful for accepting an outsider such as me into the field.

From a practical point of view, the most important person for my dissertation was my colleague, co-author, and friend Dr. Darwin Semmler, with whom I worked for five years and wrote four out of the five dissertation chapters. Darwin is a mathematical genius (which he will deny, of course) with a natural gift for psychology, behavioral sciences, and programming, and I severely doubt that I would have been able to conduct most of my research without his help. I also want to thank my two other co-authors Alix Auzepy from BWL VI and Dr. Yannik Bofinger from our research network. Their knowledge and expertise about all things finance (way beyond "der kann ja eh nur Daten", Yannik) dramatically increased the quality of the projects we worked on together. Furthermore, I want to thank the fourth member of our research network, Benjamin Fiorelli, as well as Christina Stahlecker from the chair of social psychology, for a truckload of helpful feedback, additional ideas, and lengthy discussions. In addition, I want to thank my former and current colleagues from BWL VI, i.e. Karsten Bocks, Henry Flach, Dr. Thomas Heyden, Jan Reinschmidt, Dr. Björn Rock, and Dr. Carolin Schürg, for their support, helpful contributions and discussions in Brown Bags and other occasions. I also want to express my gratitude to all the other people from the department whom I discussed my work with, who gave feedback in seminars and otherwise, who participated in beta tests, or contributed in other ways. They are Prof. Dr. Petrit Ademi, Dr. Kim Heyden, Dr. Niklas Kreilkamp, Dr. Sebastian Krügel, Dr. Sascha Matanovic, Dr. Tobias Meyll, Karina Raith, Dr. Florian Röder, Philipp Schade, Dr. Maximilian Schmidt, Dr. Julia Schneider and Ferogh Zaman.

I cannot finish these acknowledgments without thanking three additional people: First our student assistant Julia Schwarz (hm, who did hire you again?) for her invaluable support in research and teaching - I'm looking forward to that Post-Doc job you promised me. I also want to thank Julia Körner and Annette Toalster, the former and current secretaries of the Chair of Banking & Finance. While I have heard some deeply disturbing horror stories involving academic bureaucracy, I never had to deal with such aspects of my job too much. I am *deeply* grateful for that, and I know I owe that to you and your tireless efforts to shield us researchers from all that tedious stuff.

Finally, I want to thank my family, in particular my parents Hans Gärtner and Anke Wenisch. As far as I know, I am the first in my family to ever aspire to achieve a Ph.D. Watching your child sailing into unknown waters is always a scary thing for parents, and I want to thank them for their trust and confidence in me and my success.

General Introduction

This dissertation's major topic is financial decision-making with a focus on nonprofessionals, and in particular on behavior that the decision-makers themselves might judge as sub-optimal or even come to regret. This approach, of course, nests this thesis within the subfield of Behavioral Finance. What sets it apart from other research in this field is the focus on very basic, or as my coauthor Darwin Semmler and I will call them in chapter IV, "elementary" financial decisions. Roughly defined, these are decisions that even most of the research on behavioral finance assumes to simply "work". As an example, consider the following examples, taken from chapter IV:

- Example 1: You can invest some money. Do you prefer to invest in a (safe) asset with 6% returns or in a (safe) asset with 12% returns, all else equal?
- Example 2: You need to borrow some money. Do you prefer to borrow for a 5% interest rate or a 10% rate, all else equal?

From a Finance point of view, these decisions are incredibly simple. Basically each model on financial decision-making, no matter if traditional or behavioral, predicts that investors maximize their returns by investing in the 12% asset, and borrowers minimize their interest payments by borrowing for 5% - in fact, I am not aware of any model that predicts a different outcome. Yet, as this thesis will document, people surprisingly often fail to find these optimal solutions in such and similar decisions. Starting from this empirical fact, I want to tackle three different questions.

The first question asks about where the problem starts. Instead of the common behavioral economics question "Where do people deviate from rationality?", I instead ask "Where can we expect people to be rational?" To investigate this question, I develop a series of experiments where participants have to make very basic financial decisions such as investing in two different assets or borrowing and repaying debts on two different credit cards, and test whether they differ from the rational benchmark. After all, it is reasonable to assume that the simpler the decision, there fewer biases I should find. In addition, biases in such simple decisions might spill over to more complex situations which might include risk, different time horizons, additional product features such as minimum repayments or performance-based fees, legal issues, etc. In general, I find that such biases are surprisingly common even in most of the simplest financial situations. In most experiments, the baseline misallocation, i.e. money invested in, borrowed from, or repaid to the dominated alternative, is around 20 to 25% and can go up to around 50%, essentially chance level. Most participants misallocate at least a little. The only domain that seems to work relatively well is investing, where misallocation is usually in the single percentage digits, but can go up to more than 20% as well, depending on the different experimental treatments.

Second, I want to gain a better understanding of the thought processes that lead to such misallocation. I use the psychological "heuristics and biases" tradition (Newell and Simon, 1972; Tversky and Kahneman, 1974) to explain financial decisions. Following this approach, people make decisions based on choice heuristics, i.e. simple rules of thumb. Such heuristics usually lead to good results, but if used under the wrong circumstances, they predictably fall short relative to the rational benchmark. In my experiments, I identify several such heuristics, and design situations where using them should lead to a predictable misallocation of financial means. I find that this approach indeed predicts financial decisions.

As a third question, I want to test methods to help people avoid costly misallocation. I test several nudging (Thaler and Sunstein, 2021) and boosting (Hertwig and Grüne-Yanoff, 2017) methods such as different framing, reminders, and information brochures. My experiments show that these interventions usually decrease misallocation, and are sometimes able to eradicate it more or less completely.

The first four out of the five chapters of this thesis are dedicated to these questions - however, in the order of how I finished the papers, which is not necessarily the order outlined above. The first three papers are a trilogy about credit card repayments, the fourth broadens the scope to investing and borrowing. While my coauthors and I published a summary of the key ideas from chapters 1 and 2 in an article called "What could possibly go wrong? Predictable misallocation in simple debt repayment experiments" in the Journal of Economic Behavior and Organization (Gärtner et al., 2023), each chapter is a complete stand-alone paper in its own right.

The first chapter, "What could possibly go wrong? Triggering misallocation", is coauthored by Darwin Semmler and Christina Bannier, and aims to identify reasons for non-optimal debt repayments. We investigate experimental situations where participants have debts on two credit card accounts and an income that they must use to repay a fraction of these debts. By manipulating the values of debts, interest rates and income, we can create certain situations (called "fallacy scenarios") where a certain repayment heuristic should lead to misallocated repayments, compared to a control scenario with almost the same values, but where the heuristic is either impossible to use, or would not lead to an error. As an example, consider the "complete repayment" heuristic (Amar et al., 2011). Participants following this heuristic repay the debts on one credit card completely, regardless of the card's interest rate. If this card is the one with the lower interest rate, this leads to misallocation. We designed a scenario where we endow the participants with just enough income to repay the low-interest card completely, while in the corresponding control treatment, we increase the debt on the low-interest card by $10, making it impossible to repay it completely. We test seven heuristics and find that for this particular heuristic, and three others, participant misallocate more of their income in the respective fallacy scenario than in the corresponding control scenarios. In an exploratory part of the chapter, we also show that only around a fourth of our subjects can reliably avoid misallocation in each scenario. These results imply that even for a very simple financial decision such as debt repayments, participants use heuristics, which might consequently lead them astray if the heuristics do not match the situation.

In contrast to the first chapter, the second chapter, "What could possibly go wrong? Nudging and the Cuckoo Fallacy", which is again coauthored by Darwin Semmler and Christina Bannier, narrows the focus down to one particular novel heuristic, which we investigate in detail. This heuristic is again best illustrated using an example: Consider that you have two credit cards. You have debts of $10.000 on Card 1, which charges 5% p.a., and debts of $1.000 on credit card 2, which charges 10% p.a. You can repay $200. How do you repay?

It is tempting, even for some professors and Ph.D. students in finance whom we encountered at conferences and workshops, to use all that money for the 5% card. After all, 5% of $10.000 is $500 worth of additional debts, while 10% of $1.000 is only worth $100, which in comparison seems negligible. However, using this simple heuristic, where people focus on the sum of new debts a credit card accumulates, is not optimal - neither in this example nor in general. It can be optimal, but only if the card with a higher interest rate accumulates more debt. This is often the case, of course, given that the interest rate is an important factor for debt growth. However, if the low-interest card accumulates more debts, as is the case in the example, this heuristic is not optimal. This might seem counterintuitive at first, but it becomes quite clear if the problem is reframed: No matter the balances, you can only reduce the debt growth per card by $\$200 \times \textit{interest rate}$. In this example, this is \$10 for the 5% card, and \$20 for the 10% card, and some number between these two extreme values for all possible repayment combinations in between. Whoever repays the low-interest first *because* it accumulates more new debts commits what we have termed the "Cuckoo Fallacy". Chapter II focuses on this fallacy. In a framing experiment, we either highlight or hide various information such as the interest rates, the amount of newly accumulated debts, or the amount of money saved given a certain repayment scheme, to steer the occurrence of the fallacy. In particular, we develop a nudge treatment, designed to reduce the misallocation which results from that fallacy. This treatment is successful. We also employ a sludge treatment, designed to worsen the misallocation, which seems to not have worked. These results imply that nudging can be successful in improving repayment decisions, while luckily suggesting that there is not much room for deliberately worsening them.

However, while understanding one particular fallacy in depth is important to comprehend human reasoning, and also helps to combat misallocation following from this fallacy, this approach would not be a very useful or practical solution to try to combat *each* of such fallacies, because that would require to shield every situation against each and every possible fallacies. For practical solutions, it is often sufficient to guide decision-makers to the optimal choice, without necessarily tackling all the details involved. In the spirit of research question 3, on how to minimize misallocation, chapter III called "Addressing consumer misunderstanding in credit card debt repayment: Policy suggestions beyond the CARD Act" (coauthored by Yannik Bofinger and Darwin

Semmler) tests such practical solutions in an experiment. We implement four different methods to reduce misallocation and conduct an experiment to test them. Two of these interventions are "general" interventions, which rely on the fact that there are generalities about each repayment decision, such as the fact that it is always optimal to repay the high-interest card first. The other two are "adapted" interventions, which are tailored to the specific situation, i.e. combination of interest rates, balances, and income. We find that all four interventions decrease misallocation, but the strongest effect comes from an adaptive software assistant which suggests the optimal repayment in each situation. Finally, the paper features a detailed discussion on some policy implications of our results, using the U.S. Credit Card Accountability Responsibility and Disclosure (CARD) Act of 2009 as a basis.

This third chapter finishes the trilogy on credit card repayments. However, credit card repayments are a relatively narrow range of financial decisions. I find misallocation, but does that mean that people have problems solving simple financial decisions in general, or are the problems limited to credit card repayments? And, in the spirit of research question 1, are debt repayments even the simplest financial decisions possible? To address such and similar concerns, chapter IV, which is joined work with Darwin Semmler and called "Elementary Financial Decisions", investigates investing and borrowing decisions. While algebraically identical to debt repayment decisions, decision-makers might perceive these two classes of decisions as simpler, because they do not necessarily involve balances, which removes one variable decision-makers might find important from the decision. Indeed, many of the explanations for misallocation in debt repayments championed and tested in the first three chapters use balances as an explanatory factor. The fourth chapter instead uses a novel concept called "cognitive uncertainty", recently introduced by Benjamin Enke and David Graeber (Enke and

Graeber, 2023), and tests whether this theory can shine a light on misallocation in decisions without balances. Unlike the canonical von Neumann-Morgenstern understanding of risk, where risk or uncertainty are features of the external world, i.e. the alternatives between which to choose from, cognitive uncertainty models "internal" uncertainty in the decision process of the decision-maker. The decision-maker might be aware that an optimal solution exists, but they are not necessarily certain how to achieve that, which in this model can result in a non-optimal action. We test this theory using experiments in which we vary cognitive uncertainty using three different independent variables. Participants in both experiments encounter decisions to invest or to borrow, they encounter positive or negative interest rates, and they encounter frames where we either present returns and interests as percentages or as absolute values. We show that investment decisions work reasonably well while borrowing decisions show some problems. According to our results, cognitive uncertainty might in part explain the misallocation stemming from these problems, but cannot fully account for them, and the results are in general not very robust. This implies that while cognitive uncertainty might play a role, the misallocation in these simplest financial decisions remains a mystery, which points to avenues for future research in this area.

The fifth and final chapter called "Looking beyond ESG preferences: The role of sustainable finance literacy in sustainable investing", coauthored by Alix Auzepy and Christina Bannier, is a bit of an outlier in my thesis, although it still shares a lot of common themes with the other chapters. The study is about financial decision-making, it is experimental, it investigates non-professionals, and it investigates non-optimal decision-making. However, it does not focus on very simple financial decisions. Quite the opposite, the chapter's central research question is derived from the complexity of a certain investment type. We investigate whether a certain form of financial literacy, called "sustainable finance literacy" (Filippini et al., 2023), influences sustainable investments. The basic research idea is straightforward: Actual investment decisions in the field are complex and complicated, and without sufficient knowledge, it is hard to translate one's own preferences into actions. In particular, to be able to invest sustainably, it is necessary to understand and to be able to identify sustainable financial products. We test this idea using an informational brochure designed to increase sustainable finance literacy. In an experiment, the treatment group is provided with the brochure, while the control group is not. Both groups have to make four investment decisions, where per decision they can decide to invest in one of three funds, some conventional and some sustainable. We find that the brochure increases the probability of investing sustainably, and, if given the choice between a "light green" and a "dark green" fund, brochure readers will more likely choose the dark green fund. We show that the brochure's effect is indeed in part mediated by sustainable finance literacy. However, these effects are moderated by sustainability preferences, and there is even some weak evidence that increasing sustainable finance literacy decreases sustainable investments for participants with very low sustainability preferences. This study shows that while preferences, on which most of the recent literature and regulation focuses, are important for sustainable investments, they must be accompanied by sufficient literacy.

Chapter I

What could possibly go wrong? Triggering misallocation

Coauthors:

Darwin Semmler

Christina E. Bannier

Relative share:

45%

Publication (jointly with chapter II):

Gärtner, Florian, Darwin Semmler, and Christina E. Bannier (2023), "What could possibly go wrong? Predictable misallocation in simple debt repayment experiments", Journal of Economic Behavior & Organization, 205, 28–43

Previous versions of this chapter have been presented at:

- 2nd Personal Finance Workshop 2019
- Jahrestagung des Vereins für Socialpolitik 2020 (conjoint version with chapter II)
- ASSA/AEA conference 2021, Poster Session (conjoint version with chapter II)

What could possibly go wrong? Triggering misallocation

Abstract

How do borrowers repay their debts? In a simple debt repayment experiment on Amazon Mechanical Turk, we elicit different repayment heuristics, i.e. predictable repayment rules used by our participants which can lead to specific deviations from debt minimizing repayments. We also show in which situations these heuristics can be triggered using supposedly irrelevant information. Furthermore, we identify four different clusters of participants based on their repayment decisions, which highlights the heterogeneity based on personal aspects.

Keywords: Household finance, credit cards, financial literacy, rationality, bias, cuckoo fallacy

JEL-Codes: D14 - D91 - G41 - G50

Funding: This work was financially supported by the "Frankfurter Institut für Risikomanagement und Regulierung" (FIRM). FIRM had no involvement in anything study-related.

Declarations of interest: none

I.1 Introduction

Over the past decades, several household finance puzzles, i.e. deviations from optimal behavior as deduced by rational choice, have been identified (Beshears et al., 2018; DellaVigna, 2009; Zinman, 2015). Recently, a specific credit card debt puzzle has received particular attention. It posits that, when endowed with several cards, a significant fraction of borrowers does not repay them in a debt minimizing way. Rather, two field studies from Mexico (Ponce et al., 2017) and the UK (Gathergood et al., 2019) show that, even after accounting for minimum repayments, around half of the repaid money is misallocated on cards with lower interest rates. More precisely, Gathergood et al. (2019) find for the UK that "[...] 85 percent of individuals should put 100 percent of their excess payments on the high-interest rate card but only 10 percent do so."

These results imply that a considerable number of people do not know how to repay debts optimally. This is puzzling because repaying credit cards is undoubtedly one of the simplest financial problems and the optimal strategy is straightforward: You fully repay the card with the highest interest rate first, then continue with the second most expensive card, etc. In order to better understand why people do not follow this strategy and how decisions can be improved, we develop an experiment. Participants hold two credit card accounts with different interest rates and negative balances, and are provided with an income to repay these debts. We run the experiments on the online platform Amazon Mechanical Turk (MTurk).

In the field studies preceding our work, individuals seem to rely on behavioral concepts such as mental accounting (Ponce et al., 2017; Thaler, 1985) or heuristics (Gathergood et al., 2019; Tversky and Kahneman, 1974) to reduce the complexity of their decision-making task. However, given that field studies are limited to the specific situation their subjects naturally find themselves in and which they endogenously created themselves, it is not clear whether such explanations remain to hold in more general environments. In order to understand whether and in which way different environments elicit different behavior, we design this experiment to identify situations that might lead to misallocation. We argue that individuals who do not know how to repay debts optimally use more information than just interest rates. In our analyses, we therefore consider additional information on income and account balances. These pieces of information are irrelevant for the optimal repayment strategy, but if a person does not know this, they might use them anyway. Manipulating the information environment in this way allows us to see how certain repayment decisions - both optimal and non-optimal - can be triggered.

To predict patterns of misallocation, we create several "scenarios", i.e. information environments, by changing the values of either interest rates, credit card balances, or income. Depending on the exact configuration of these values, we try to elicit the use of seven distinct heuristics, where some are taken from the literature, while others are novel. For each heuristic, we consider a pair of scenarios. In the "fallacy scenario", the heuristic should lead to a specific pattern of misallocation. In the corresponding control scenario, this misallocation should be weaker or be unable to occur at all. Pairs of scenarios only differ in one value,[1] all other variables remain constant. By comparing these seven pairs of scenarios, we are hence able to see whether different types of misallocation due to the use of certain heuristics can be reliably induced.

Indeed, we find evidence for 4 out of 7 predicted fallacies. Multivariate analyses show that these results are robust against controlling for person-specific characteristics such as gender and age. Financial literacy, however, measured via a sum index using six

[1]There is one exception where we change two values.

questions introduced by Lusardi and Mitchell (2011) and Lusardi and Tufano (2015), shows a nuanced relationship with misallocation: In general, financial literacy helps to find the optimal repayment solution. However, if participants with high financial literacy fail to choose the optimal solution, we find that they use the same heuristics as less financially literate subjects, and thus fall with the same frequency for the same fallacies.

In an exploratory within analysis using k-means-clustering, we find four distinct clusters of participants. Roughly a quarter of our participants belongs to a cluster with almost no misallocation throughout all decisions. Another quarter generally knows relatively well how to repay optimally, but is vulnerable to fallacies. A third cluster is also vulnerable, but from a much lower baseline. The final cluster chooses particularly bad, but is also vulnerable to fallacies. We also show that the choices of repayment heuristics do only mildly correlate with each other - knowing that a participant shows one particular fallacy only weakly predicts falling for another fallacy.

By studying how and why individuals make non-optimal debt repayment decisions, our work complements the literature on consumer finance puzzles (e.g. Agarwal et al. (2015); Gorbachev and Luengo-Prado (2019); Keys and Wang (2019); Stango and Zinman (2016)) and mental gaps (Handel and Schwartzstein, 2018) in general. More specifically, our work complements and enhances the findings on non-optimal credit card repayment from the field (Gathergood et al., 2019; Ponce et al., 2017). We chose to run an experiment since it is particularly helpful to broaden the field studies' results as it allows to exogenize decision parameters such as interest rates, card balances, or disposable income. This grants causal interpretations of changes in such parameters, as we employ in our experiment. Experiments further discard complications that may arise in the field: For instance, a person might organize their mental accounting system around their credit cards (Ponce et al. (2017) find evidence for that). Additionally, rational inattention (Sims, 2003) may lead a person to protect themselves against small print clauses that they suspect to exist.

The remainder of this chapter is structured as follows: Section I.2 describes the general experimental setup, section I.3 the data collection and I.4 presents the results. We show the robustness checks in section I.5 and describe an additional within-subject analysis in section I.6. Section I.7 discusses our results and concludes.

I.2 Experimental setup and hypotheses

In the experiment, subjects have to make fifteen different, independent decisions. For each decision we provide them with two credit card accounts and a checking account. To each credit card we assign a certain level of debt and a specific interest rate. On the checking account, participants have some disposable income that they can use to repay these debts. After the repayment decision is taken, the credit cards charge interests for one single time.

This experimental setting allows us to vary the values of five parameters: the two credit card balances, the two interest rates, and the income. We refer to a specific combination of these values as a "scenario". Comparing different scenarios then enables us to trace back the usage of certain repayment heuristics to these five values. We examine a comprehensive list of seven distinct repayment heuristics. Six of them stem from the literature or are natural variations of established heuristics, one is novel. For each heuristic, we develop a pair of scenarios in such a way that choosing a certain repayment heuristic either becomes more intuitive or less intuitive (or even impossible). We refer to the former scenarios as "fallacy scenarios", as we design the scenario such that the heuristic implies a non-optimal repayment decision, and to the latter as "control scenarios". We create a fallacy scenario by changing exactly one or (in one case) two values of its corresponding control scenario. All other values remain constant. Our objective is to establish whether the informational environment that we provide our subjects with can trigger certain repayment heuristics (and corresponding misallocations).

We investigate the following seven repayment heuristics and construct the corresponding pairs of scenarios:

1. Cuckoo Fallacy: Inspired by the results of a pretest, we consider a novel heuristic according to which borrowers repay most of their available money to the credit card which accumulates the highest amount of new debts. If this card is the cheaper one, which can happen if its starting balance is sufficiently larger than that of the expensive card, this heuristic induces misallocation (for an example, see footnote [2]). To the best of our knowledge, this fallacy has not been described before. We refer to it as the "Cuckoo Fallacy", as it mirrors behavior that is similar to parenting birds tending first to the largest fledgling in their nest, which may be a cuckoo chick. In the fallacy scenario, the low-interest rate card has a sufficiently high balance that it accumulates more new debts than the high-interest rate card. In the control scenario, the high-interest rate card accumulates more new debts.

2. Equalizing Balances: Participants might aim to simplify the decision problem for future decisions (regardless of whether this future really exists). Equalizing the account balances might serve that objective, because this reduces the information

[2]Consider a stylized example: You have debts of $4000 on a credit card account with a 3% interest rate, and $500 on a second account with a 5% rate. In the next period, the $4000 card will produce $120 of interest payments, i.e. "new debts", while the $500 card will accumulate only $25. So if you ignore the cheaper (3%-) card, its debt seems to "explode". Should you try to suppress this explosion? Rationally the answer is no, you should still repay the expensive card first, even though it accumulates less overall debt.

concerning the difference between balances to zero. We consequently construct a scenario where the income matches the difference in the account balances such that, if a subject uses the total income to repay the cheaper card, the balances of the two cards are equal. In the control scenario, the numbers do not match. While the income is larger than the balance difference, the experimental design does not allow equalizing balances, as we do not give our participants the options to do so (we explain how and why later in more detail, see also Table I.1a).

3. Complete Repayment: This heuristic replicates the concept of debt account aversion following Amar et al. (2011) which assumes that debtors prefer to reduce the number of open credits rather than their total amount of debt. This leads to a fallacy scenario where the available income matches exactly the balance of the cheaper card, and a control scenario where it is not possible to repay any card completely.

4. Balance Matching (Gathergood et al., 2019): This heuristic describes behavior where the share of repayments on the credit cards matches the share of the balances on each card, e.g. if 60% of the total outstanding debt is on one card, it receives 60% of the repayments. Gathergood et al. (2019) argue that this heuristic arises from balances as salient pieces of information and a general human tendency to show matching behavior in similar choice tasks, such as probability matching (Vulkan, 2000). Balance Matching should be easier to conduct if the income and the two account balances are immediately matchable (e.g., if the balances are simple multiples of the income), so we use this as a fallacy scenario. In the control scenario, the numbers do not match as smoothly.

5. Interest Matching: This heuristic is a natural extension of Balance Matching, as

the general argument for matching behavior should also apply to interest rates. According to this heuristic, subjects repay the available money in proportion to the interest rates, e.g., if one card charges 3% and the other 6%, 1/3 of the income is paid on the first card and 2/3 on the second. In the fallacy scenario, the income is therefore a multiple of the sum of the interest rates, so that matching on the individual rates can be easily done, while in the control scenario matching is not as easy.

6. 1/N Heuristic (Benartzi and Thaler, 2001): This repayment heuristic expresses the idea of naive diversification, i.e. repayment is split evenly on the credit cards. This heuristic is simple enough that it does not require any information. However, it implies that deviations from optimality become more severe the larger the spread between the cards' interest rates is. We use this argument to create a fallacy scenario with a small interest spread of 1 percentage point, and a control scenario with a large spread of 10 percentage points.

7. Equal Start: We are also interested in behavior that is triggered by a situation in which credit card balances are equal, as now the only distinguishing feature between the scenarios are the cards' interest rates. It should be noted that in this fallacy scenario, 1/N Heuristic and Balance Matching coincide to the same behavior. We design the control scenario such that the balances are not equal.

We finally investigate one further scenario that we denote as "Everything Equal", where the credit cards show equal balances and equal interest rates. Clearly, there is no optimal behavior anymore, and subjects should be indifferent between both cards. We use this scenario to measure behavior under indifference.

In order to determine unequivocally whether a subject succumbs to a certain fallacy in each of the scenarios, we offer only a limited set of repayment options. These options need to be symmetric to allow for interchangeability of credit cards, and identical in all scenarios to make them comparable. Additionally, the relations of values in the information sets and repayment options should be mathematically simple. To satisfy all these requirements, we offer five repayment options to our participants in all scenarios, as presented in Table I.1a. Apart from option 5, where the total income is repaid on the high-interest rate card, each repayment option implies a certain amount of misallocation. To incentivize our participants to minimize their overall debt, i.e. the misallocation, we offer a bonus at the end of the experiment that varies accordingly (details on the bonus design are provided in Section I.3). We use a fixed order for the repayment options in the experiment, but randomize the order of the credit cards[3]. If participants choose an option, they see the implicated new balances *before interests*, to minimize misallocation due to calculation errors.

Table I.1b summarizes the 15 scenarios and the corresponding details. In the experiment, we quasi-randomize the order of the scenarios by assigning one randomly chosen scenario of each heuristic to a random position from 1 to 7. The Everything Equal scenario is always the 8th scenario, and the remaining scenarios are assigned to a random position from 8 to 15. We break pure randomization for two reasons. First, we want to avoid that both scenarios of the same fallacy can be close together, because we suspect this to lead to a sharper contrast and thus more extreme behavior (more rational in the control scenarios, more misallocation in the fallacy scenarios), which would artificially boost our results. Second, we want the Everything Equal scenario to be right in the middle because we suspect that our scenarios might be too simple for many perfect re-

[3]The options in the experiment itself only refer to credit card 1 or credit card 2, and therefore depend on the random order of the credit cards. This means that in an actual scenario, either option 5 or option 1 can be optimal, depending on the random order of the credit cards. However, for this paper we need a standardized representation. Thus, we redefine the options regarding low- and high-interest credit card as shown in Table I.1a, such that option 5 consequently is the optimal option.

payers, in the sense that they would either not believe our experiment and give random answers just in case we might fool them, or simply fast-click without paying any attention, which could lead to more errors. Both effects potentially increase misallocation, which would bias the results towards higher misallocation. Placing the only scenario where other buttons than the two outer ones are at least in principle optimal right in the middle of the experiment might counter that problem somewhat.

As shown in the second column of Table I.1b, we expect each heuristic to trigger the use of one specific repayment option ("fallacy-implicated option"), which is denoted in parentheses. The remaining columns present the information on the income (checking account), the two card account balances and the two interest rates that define each scenario.

Note that this experiment is not a framing experiment, even though it is similar in spirit. In framing experiments (e.g. Tversky and Kahneman (1981)) the same information is delivered in different frames, e.g. by different wording or color schemes. Here we instead keep the frame constant, and change the information instead. But this information change is still irrelevant with respect to the optimal repayment strategy, just as changing frames is supposed to be irrelevant in classical framing experiments.

Our analyses focus on comparing the decisions of the participants in control and fallacy scenarios for each heuristic. The dependent variable in our analyses is therefore the repayment option that a participant chooses. In order to test whether our scenario design allows to predict the choice of the seven different heuristics and the corresponding repayment misallocation, we examine the following two hypotheses:

H1.1: The fallacy-implicated option is chosen more often in a fallacy scenario compared to the corresponding control scenario.

H1.2: The optimal option is chosen less often in a fallacy scenario compared to the corresponding control scenario.

Table I.1: Repayment options and scenarios

(a) Description of repayment options for all scenarios

Option no.	Notation	Description of payment	Implied bonus
1	All on low	All money → low-interest credit card	USD 0.00
2	2:1	$\frac{2}{3}$ → low-interest card, $\frac{1}{3}$ → high-interest credit card	USD 0.10
3	1:1	$\frac{1}{2}$ → low-interest card, $\frac{1}{2}$ → high-interest credit card	USD 0.15
4	1:2	$\frac{1}{3}$ → low-interest card, $\frac{2}{3}$ → high-interest credit card	USD 0.20
5	All on high	All money → high-interest credit card	USD 0.30

(b) Description of the scenarios[a]

Scenario No.	Triggered fallacy (Implicated option)	Checking account	Credit card 1	Credit card 2	Interest rate 1	Interest rate 2
1	Control	$120	$-1000	$-1400	13 %	15 %
2	Cuckoo Fallacy (1)	$120	$-1000	$-250	13 %	15 %
3	Control	$150	$-270	$-210	7 %	19 %
4	Equalize Balances (1)	$60	$-270	$-210	7 %	19 %
5	Control	$90	$-100	$-125	5 %	8 %
6	Complete Repayment (1)	$90	$-90	$-125	5 %	8 %
7	Control	$300	$-1400	$-1000	6 %	17 %
8	Balance Matching (2)	$300	$-2000	$-1000	6 %	17 %
9	Control	$600	$-1300	$-1700	1 %	11 %
10	1/N Heuristic (3)	$600	$-1300	$-1700	10 %	11 %
11	Control	$600	$-1100	$-1000	9 %	17 %
12	Interest Matching (4)	$600	$-1100	$-1000	10 %	20 %
13	Control	$540	$-1000	$-1700	4 %	5 %
14	Equal Start (3)	$540	$-1000	$-1000	4 %	5 %
15	Everything Equal	$60	$-100	$-100	6 %	6 %

[a] This table shows the values for the income on the checking account, for the credit card balances, and for the interest rates. Each double row contains a pair of control- and fallacy scenario. The number in parentheses denotes the option a subject would choose if they succumb to the concerning fallacy.

I.3 Data

We set up our experiment on the platform SoPHIE (Hendriks, 2012) and recruit the participants on Amazon's crowd-sourcing platform MTurk. Restricted to the US population, participants on MTurk (Turkers) are asked to solve individual Human Intelligence Tasks (HIT), which can then be approved or rejected by the requester of that HIT. Each of our experiments is one single HIT. We restrict participation to Turkers with at least 100 completed HITs to screen out throwaway accounts, bots and new Turkers, whom we expect to make more mistakes due to their unfamiliarity with MTurk. We require an approval rate on former HITs of at least 95%, a common threshold that was shown to ensure high data quality (Peer et al., 2014). No participant was allowed to take part in any of our experiments in this or any other chapter more than once.

The experimental design has three stages. We first explain the experiment to the participants. We then run the actual experimental stage, and finish with a post experiment questionnaire (PEQ). In stage 1, we ensure that the participants understand the rules of the experiment by running several comprehension tasks and two trial scenarios. Participants can also read the rules of the experiment during the experiment rounds anytime. To make sure that our subjects have a basic level of numeracy, we ask them in stage 1 to calculate the interest on a balance of $1000 with a 1% interest rate. Participants have to answer this question correctly to advance into the experimental stage. Here, participants go through all 15 scenarios in the quasi-randomized order as described above and make repayment decisions in each. We collect these decisions as our main data to be analyzed.

The PEQ includes questions on gender,[4] age, number of years of education and

[4]Unless stated otherwise, "female" is used as the reference category.

on financial literacy. To measure the latter comprehensively, we use the "Big Three" questions by Lusardi and Mitchell (2011) and three additional debt-related questions by Lusardi and Tufano (2015). We interpret the number of correct answers as a sum index measure of financial literacy. To exclude bots, we ask our participants to describe the strategies they have used in the experiment in an open question. Two different researchers analyze if the answers are meaningful for that question. Both agree that this is the case for all our subjects. We are therefore confident that our data does not contain any bot. We also include two attention check questions. In the first question, positioned in stage 1 right after the numeracy question, participants have to agree or disagree with the statement "All my friends are from outer space". Whoever agrees is screened out. The second question is included in the financial literacy questionnaire in stage 3. Subjects have to decide between choices we label "First answer" and "Second answer", where we ask them to select "Second answer". We screen out everyone who selects "First answer".

We pay a fixed participation fee of \$1, and a bonus of up to \$4.50, depending on the participant's decisions at the end of the experiment. A subject earns \$0.30 if they use the option to repay all their income on the high-interest card. For the other decisions we either pay \$0.20, \$0.15, \$0.10 or \$0 per decision, depending on the share of money repaid to the high-interest rate card (see also Table I.1a).[5] Thus, the maximum achievable payment in the experiment is \$5.50 ($= \$1 + 15 \cdot \$0.30$). On average, our participants earn \$4.45 in 22:29 minutes (participation fee already included), which implies an average hourly payment of around \$11.88. According to the literature these hourly payments are higher than the average payments on MTurk.[6] While we hence seem to overpay

[5] In the "everything" equal scenario we always pay \$0.30, as every choice is equally optimal.

[6] Hara et al. (2017) estimate the median wage on MTurk to be lower than \$2 and the mean wage slightly above \$3. Berg (2016) estimates an average hourly wage of around \$5.50.

our subjects relative to their expectations, our payments are comparable to common lab compensations, which is why we argue that the material incentives work.

468 MTurkers started our experiment, of which 343 finished it. Out of the 125 who did not finish the experiment, 89 dropped out before the basic numeracy question, 27 did not pass the basic numeracy question, and 9 dropped out within the experiment or the post experimental questionnaire. Out of the remaining 343 participants, 335 passed the attention tests; these form our eventual sample. The data was collected in January and February 2019. Table I.2 presents further summary statistics.

Table I.2: Summary statistics of participants

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Financial literacy	335	3.97	1.29	0	3	4	5	6
Age	335	35.86	10.49	20	28.5	33	41	72
Years of education	335	15.18	2.19	10	14	16	16	21
Experiment duration (min:sec)	335	22:29	10:08	05:55	15:04	20:26	27:00	59:08
Payoff (USD)	335	4.45	0.77	1.60	4.00	4.40	5.10	5.50
Gender info	Males: 195		Females: 140		Third gender: 0			

I.4 Results

Figure I.1 provides a first descriptive analysis of the data. It shows the distribution of the chosen repayment options in each of the fifteen scenarios (see also Table Appendix I.16). As can be seen, there is severe misallocation, i.e. money that was not repaid to the high-interest rate credit card (choice of options 1-4), in the fallacy scenarios, and some misallocation even in the control scenarios. This suggests that a large fraction of participants in our experiment does not know how to repay debt optimally. Furthermore, in the scenarios referring to the Cuckoo Fallacy, Complete Repayment, 1/N Heuristic and Equal Start, the fallacy-implicated option was indeed chosen noticeably more often.

Surprisingly, however, repayment choices in the control scenarios not only differ from their corresponding fallacy scenarios but also from each other. This can be seen as further indication that the information set has a strong influence on the repayment decision and also highlights the importance of using a specific control scenario for each fallacy. It is moreover interesting to note that in the Everything Equal scenario, around 82% of all subjects choose the equal split (option 3). As the values of account balances and interest rates are equal in this scenario, the actual repayment decisions neither matter for measuring misallocation nor for participants' bonus payments. Due to this presumed indifference, one might have expected a fairly equal distribution of decisions among the five repayment options. The strong observed focus on the 1:1 split (option 3) implies instead that naive diversification between multiple credits as proposed by the 1/N Heuristic might be the natural default repayment choice.

To test our hypotheses in a multivariate perspective, we split the data into seven different parts, where each part consists of the data from one fallacy scenario and its corresponding control scenario, and use two different dependent variables. The first dependent variable, $fallacy\ option_{i,j}$, is a dummy which takes the value 1 if and only if participant i selects the fallacy-implicated option as we show in Table I.1b in scenario j. The second dependent variable, $optimal_{i,j}$, is a dummy which takes the value 1 if and only if participant i chooses option 5 in scenario j. This leads to 14 different logistic regressions, two for each scenario pair.

Since we employ a within-subject design with 15 observations for each participant, we use a random intercept term u_i for subject i. As our control variables do not vary within one subject, a fixed effects regression is unable to estimate effects for any variable except for the scenarios. We follow the suggestion of Wooldridge (2010) to estimate a random effects model instead. In our reports, we omit the additional control variables;

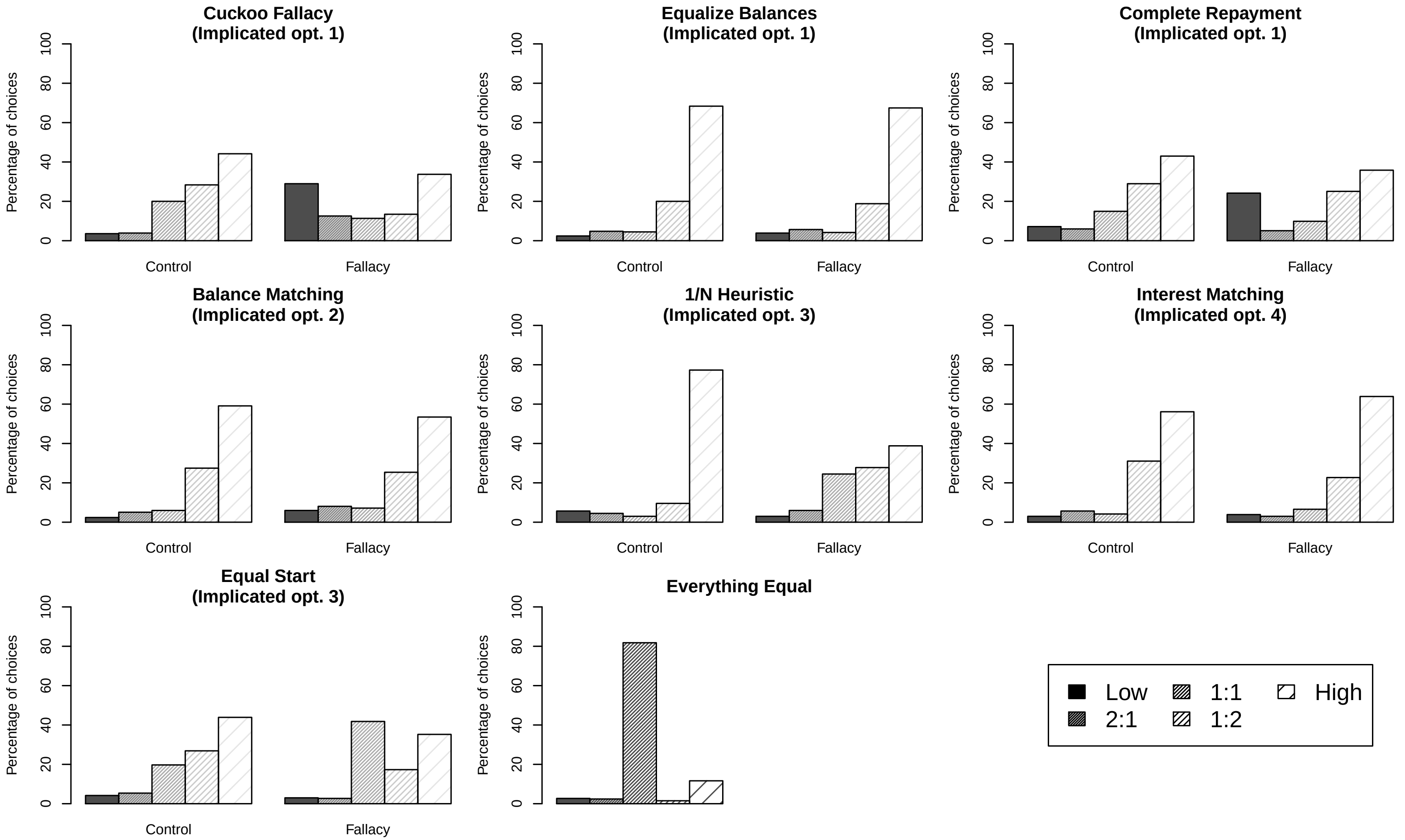


Figure I.1: Comparison of choices between control and fallacy scenarios. The columns show the percentage of choices for every repayment option. They are sorted by option no. 1 to 5. The leftmost column represents the number of participants choosing to repay all the money to the low interest credit card and the rightmost column represents the number of optimally repaying subjects. The Everything Equal scenario (third row, second column) does not have a control scenario, and since the interest rates are the same, the options do not imply any (non-)optimal repayments.

however, the full set of variables is reported in Appendix I (Tables Appendix I.14 and Appendix I.15). We use a 5% significance threshold in our regressions and apply a Bonferroni-Holm correction for the 28 coefficients we interpret in both tables combined. Since this correction drains test power, we report both the unadjusted and the adjusted p-values in the tables, but for our interpretation we rely only on the adjusted p-values.

We start by analyzing the behavior of our participants, and first ask whether participants use the fallacy implicated option more often in the fallacy scenarios. The variable "Fallacy scenario" in Table I.3 is a dummy variable with the value 1 if the respective scenario is the fallacy scenario. Its coefficient represents the difference in the probability of selecting the fallacy-implicated option relative to the control scenario. According to H1.1, we expect a significantly positive coefficient of this dummy variable. This is indeed what we observe for the Cuckoo Fallacy, Complete Repayment, 1/N Heuristic and Equal Start. Balance Matching does not survive the Bonferroni-Holm correction. We do not detect an effect in Equalize Balances. Interest Matching shows a significant effect that is opposite to what H1.1 prescribes.[7]

To test whether the fallacies draw subjects away from the optimal solution, we use the second set of regressions where $optimal_{i,j}$ is the dependent variable. If the information environment of the fallacy scenario leads subjects to select the optimal option less often (H1.2), we should find significantly negative coefficients of the "Fallacy scenario" dummy in these regressions. Table I.4 presents the results. Indeed, participants are significantly less likely to select the optimal option if they are in the fallacy scenarios of the Cuckoo Fallacy, 1/N Heuristic and Equal Start. Balance Matching again goes in the hypothesized direction but does not survive the Bonferroni-Holm correction. We

[7] Although we did not theorize the latter finding, an explanation could be that we have used too high-interest rates in the scenarios of this heuristics, so that other effects might have unduly influenced the decision-making.

Table I.3: Logistic regression model with random effects[a]

	Dependent variable: Choice of fallacy-implicated repayment option (1 = Chosen, 0 = Not chosen)						
	Cuckoo Fallacy	Equalize Balances	Complete Repayment	Balance Matching	1/N Heuristic	Interest Matching	Equal Start
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fallacy scenario	0.287***	0.023	0.175***	0.038	0.239***	-0.084*	0.216***
	(0.038)	(0.014)	(0.024)	(0.016)	(0.033)	(0.026)	(0.027)
	[0.000]	[0.096]	[0.000]	[0.020]	[0.000]	[0.001]	[0.000]
	[0.000]	[0.766]	[0.000]	[0.223]	[0.000]	[0.022]	[0.000]
Financial literacy	-0.007	-0.017	-0.013	-0.018	0.013	-0.011	-0.026
	(0.027)	(0.008)	(0.019)	(0.012)	(0.037)	(0.018)	(0.020)
	[0.785]	[0.031]	[0.473]	[0.109]	[0.723]	[0.534]	[0.208]
	[1.000]	[0.310]	[1.000]	[0.764]	[1.000]	[1.000]	[1.000]
Observations	670	670	670	670	670	670	670
Fall. scen. × Fin. lit.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: *$p<0.05$; **$p<0.01$; ***$p<0.001$ for the Holm-adjusted p-values

[a] Reported coefficients are margins. The seven models denote the seven scenario pairs, the differences of control- and fallacy scenario are denoted in the Fallacy scenario coefficients. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for 28 coefficients from two tables: The seven fallacy scenario coefficients for fallacy-implicated option as dependent variable, the seven fallacy scenario coefficients for optimal option as dependent variable, as well as the 14 financial literacy coefficients from both tables. Asterisks indicate significance after adjustment.

Table I.4: Logistic regression model with random effects[a]

	Dependent variable: Choice of optimal repayment option (1 = Chosen, 0 = Not chosen)						
	Cuckoo Fallacy	Equalize Balances	Complete Repayment	Balance Matching	1/N Heuristic	Interest Matching	Equal Start
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fallacy scenario	-0.103**	-0.010	-0.073	-0.057	-0.353***	0.079*	-0.086*
	(0.028)	(0.026)	(0.025)	(0.027)	(0.022)	(0.026)	(0.028)
	[0.000]	[0.695]	[0.004]	[0.032]	[0.000]	[0.002]	[0.002]
	[0.005]	[1.000]	[0.055]	[0.291]	[0.000]	[0.036]	[0.033]
Financial literacy	0.069*	0.083***	0.052	0.068*	0.070*	0.064*	0.058
	(0.021)	(0.018)	(0.019)	(0.021)	(0.020)	(0.020)	(0.020)
	[0.001]	[0.000]	[0.006]	[0.001]	[0.001]	[0.002]	[0.004]
	[0.016]	[0.000]	[0.073]	[0.021]	[0.011]	[0.031]	[0.056]
Observations	670	670	670	670	670	670	670
Fall. scen. × Fin. lit.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$ for the Holm-adjusted p-values

[a] Reported coefficients are margins. The seven models denote the seven scenario pairs, the differences of control- and fallacy scenario are denoted in the Fallacy scenario coefficients. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for 28 coefficients from two tables: The seven fallacy scenario coefficients for Fallacy-Implicated Option as dependent variable, the seven fallacy scenario coefficients for Optimal Option as dependent variable, as well as the 14 financial literacy coefficients from both tables. Asterisks indicate significance after adjustment.

also cannot establish the predicted effects for Equalize Balances, and Interest Matching is again significant in the opposite direction. The results from Table I.4 hence nicely complement those from the earlier analysis. The only difference is that Complete Repayment does not survive the Bonferroni-Holm correction in Table I.4. Taken together, the two tables provide support for our hypotheses: Choosing the information environment accordingly, we are able to trigger certain types of misallocation following from 4 out of a total of 7 hypothesized heuristics, and reduce the share of optimally choosing participants for 3 fallacy scenarios.

The effects of financial literacy are mixed, but display an interesting and distinct pattern. Financial literacy does not show significant main effects in either of the regressions of Table I.3 with $fallacy\ option_{i,j}$ as dependent variable, as the significance in Equalize Balances does not survive the Bonferroni-Holm correction. However, it has a significant, positive coefficient in each of the regressions of Table I.4 with $optimal_{i,j}$ as dependent variable, and five coefficients stay significant even after adjusting the p-values. This leads us to conclude that financial literacy plays an intricate role for debt repayment decisions: It helps to find the optimal solution, but if subjects with high financial literacy fail to make the optimal choice, they seem to use the same heuristics and thus fall for the same fallacies as financially less literate subjects. Moreover, this pattern does not seem to depend on the fallacy itself since none of the 14 interactions with the fallacy scenarios are significant (see Tables Appendix I.14 and Appendix I.15).

This finding inspires an additional analysis of what the financially illiterate do: If they select the optimal option less often, are they more likely to select the 1:1-split option "in the middle" to express a non-tendency to either of the options? To answer this question, we divide our sample into a group of participants with financial literacy below the median (three out of six correct answers at maximum) and a group of participants with financial literacy at and above the median (at least four correct answers). A binomial test reveals that the financially illiterate tend to choose the equal split option more often than the financially literate (17.3% of choices vs. 10.4% of choices, p-value $2.92 \cdot 10^{-11}$). However, note that a share of 17.3% of choices for option 3 is still lower than we would expect if choices of the five options were completely random (binomial test: p is different from 0.2 with a p-value of 0.008). Thus, while financially illiterate participants indeed choose the even split more often, we cannot assume that they simply choose one random standard option. Instead, they still use the provided information, but with less success.

I.5 Robustness checks

To render our results more robust, we consider several changes to the analyses. First, to ensure the robustness of the model specification, we run an LPM instead of logistic regressions (Tables Appendix I.17 and Appendix I.18). Second, we employ a multinomial regression analysis to consider all choice options simultaneously and get an overview how people switch among these options between control and fallacy scenarios (Table Appendix I.19). With a multinomial regression we depict the complete distribution of the chosen options in one analysis. Third, we take into account not only that participants switch options, but also consider to which extent they switch to a better or worse option. Thus, we can check if the results are robust against different strengths of effects of the fallacies. To do this we drop the within-subject design and consider only the choice differences in control and fallacy scenario for each participant. We run an OLS regression, where we use the difference of option choice, defined as the difference between the selected option number in the fallacy scenario and the selected option number in the corresponding control scenario, as dependent variable (Table Appendix I.20). Finally, we test the robustness of the main analyses of the Tables I.3 and I.4 by including the screened out subjects in the Tables Appendix I.21 and Appendix I.22. With 343 subjects we only have slightly more than the 335 subjects before. All robustness checks confirm the results we obtain with the logistic regressions above. The only exception lies in the multinomial regression where the Equal Start heuristic does not show any significant decline from the optimal option anymore. We conclude that we have robust evidence for the Cuckoo Fallacy, 1/N Heuristic and Complete Repayment and slightly less robust evidence for Equal Start.

Furthermore, a χ^2-Test shows no detectable dependency between the order of credit cards and the option choice (p = 0.2411). The same holds true if we use the order of scenarios instead (p = 0.361). We finally rule out learning effects via a two-sample binomial test of differences in choosing the optimal option between the first and the last displayed scenario (p = 0.2793).

I.6 Exploratory within-subject analyses

Our analysis of the hypothesized fallacies so far treated each individual's decisions independently. However, the within-subject behavior of the participants over the seven fallacies might be interesting in its own - and deliver further insights on decision-making processes. We therefore run additional exploratory analyses where we trace each participant's decisions throughout the experiment and compare decisions among participants and across scenarios. We start by counting the optimal answers of each participant and report the results in Figure I.2. 60 out of 335 participants (about 17.9%) always chose the optimal option 5, i.e. gave 14 optimal answers. On the other hand, 19 participants (about 5.7%) never chose option 5.

The transition matrices for each fallacy in Table I.5 confirm the results from a within-subjects perspective. Each cell in these matrices gives the proportion of participants that switch (or do not switch, in the cells on the main diagonal) from one option in the control scenarios to another in the fallacy scenario. Indeed, many participants switch between the five options comparing control and fallacy scenarios in all seven scenario pairs. However, the table also indicates that for the Cuckoo Fallacy, Complete Repayment, 1/N and Equal Start, more participants switch from any other option in the control scenario to the fallacy-implicated option in the fallacy scenario than the other way around.

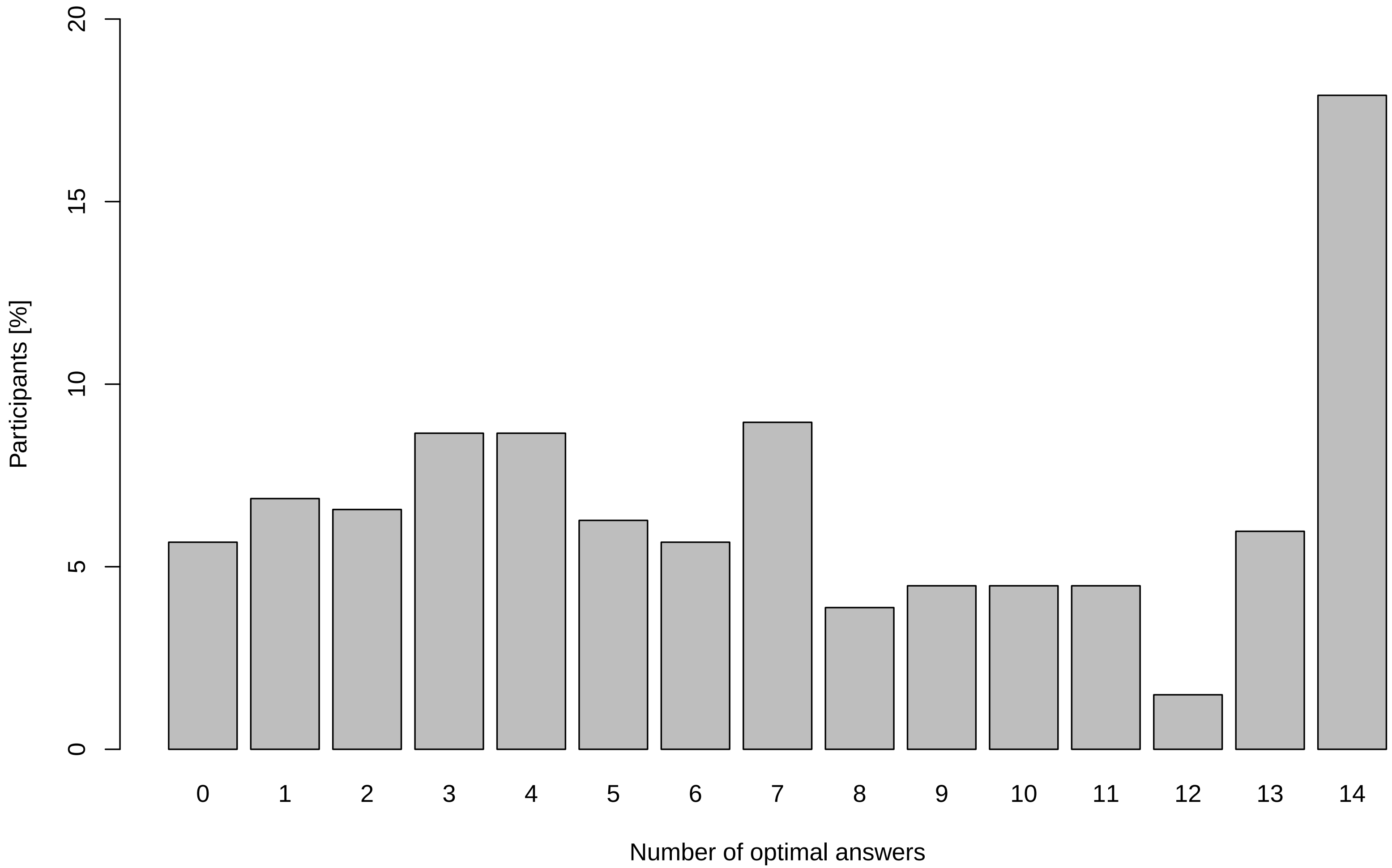


Figure I.2: The bars show which proportion of participants gave a certain amount of optimal answers (option 5) in the 14 scenarios (excluding the Everything Equal scenario).

Table I.5: Transition matrices between control and fallacy scenarios[a]

Fallacy	**Cuckoo Fallacy**					**Equalize Balances**				
Control	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
1	1.19%	0.00%	0.60%	0.00%	1.79%	1.19%	0.30%	0.00%	0.60%	0.30%
2	0.90%	0.30%	1.19%	0.30%	1.19%	0.90%	1.19%	0.90%	0.30%	1.49%
3	5.07%	3.28%	4.48%	4.18%	2.99%	0.00%	0.90%	1.19%	1.49%	0.90%
4	8.36%	6.87%	3.88%	6.57%	2.69%	0.90%	1.49%	1.49%	8.06%	56.42%
5	13.43%	2.09%	1.19%	2.39%	25.07%	0.90%	1.49%	1.49%	8.06%	56.42%
Fallacy	**Complete Repayment**					**Balance Matching**				
Control	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
1	5.07%	0.60%	0.60%	0.30%	0.60%	1.19%	0.30%	0.30%	0.30%	0.30%
2	2.09%	1.49%	0.30%	1.49%	0.60%	0.30%	2.39%	0.30%	1.19%	0.90%
3	3.28%	1.19%	3.88%	5.37%	1.19%	0.90%	0.60%	1.79%	2.09%	0.60%
4	4.18%	0.90%	4.48%	14.33%	5.07%	1.79%	2.99%	2.99%	11.94%	7.76%
5	9.55%	0.90%	0.60%	3.58%	28.36%	1.79%	1.79%	1.79%	9.85%	43.88%
Fallacy	**1/N Heuristic**					**Interest Matching**				
Control	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
1	1.19%	1.49%	1.79%	0.60%	0.60%	1.19%	0.60%	0.30%	0.90%	0.00%
2	0.00%	1.19%	1.19%	1.49%	0.60%	0.60%	0.60%	1.79%	1.19%	1.49%
3	0.00%	0.90%	1.19%	0.60%	0.30%	0.60%	0.60%	1.49%	0.30%	1.19%
4	0.30%	0.60%	3.28%	4.18%	1.19%	0.30%	0.90%	1.49%	14.93%	13.43%
5	1.49%	1.79%	17.01%	20.90%	36.12%	1.19%	0.30%	1.49%	5.37%	47.76%
Fallacy	**Equal Start**									
Control	**1**	**2**	**3**	**4**	**5**					
1	0.60%	0.30%	0.90%	0.60%	1.79%					
2	1.19%	0.30%	1.79%	1.79%	0.30%					
3	0.60%	1.19%	14.33%	1.79%	1.79%					
4	0.60%	0.30%	14.03%	6.87%	5.07%					
5	0.00%	0.60%	10.75%	6.27%	26.27%					

[a] This table shows the proportion of participants that switch from a certain option in the control scenario (rows) to a certain option in the fallacy scenario (columns) for all seven scenario pairs. Grey cells mark fallacy-implicated options and the participants switching to these option in the fallacy scenarios.

This impression is supported by Table I.6, where we calculate the proportion of optimal answers (panel (a)) and of fallacy-implicated answers (panel (b)) over all participants. We show the corresponding results for combinations of control (rows) and fallacy (column) scenarios. In panel (a), participants below the main diagonal (in grey) give more optimal answers in the control than in the fallacy scenarios. There are, for example, 18 participants (5.37%) who chose the optimal option twice in the control scenarios but only once in the fallacy scenarios, while there are only 9 people (2.69%) who show the opposite behavior. In line with our main findings, a Wilcoxon rank sum test of differences between optimal answers in control and fallacy scenarios reveals that participants indeed answered more optimally in the control scenarios (p-value: $2.2 \cdot 10^{-16}$). Panel (b) shows the corresponding proportions for the fallacy-implicated options instead. We expect more fallacy-implicated answers in the fallacy scenarios, i.e. higher proportions displayed above the main diagonal than below. This is confirmed by another Wilcoxon test (p-value: $2.2 \cdot 10^{-16}$).

A visualization of Table I.6 is given in Figure I.3, where each point represents one participant. We adapt the axes to the table, such that the x-axis denotes the count for the fallacy scenarios and the downwards directed y-axis denotes the count for the control scenarios.

Table I.6: Proportion of optimal or fallacy-implicated answers in control- and fallacy scenarios

(a) Proportion of optimal answers in control scenarios (rows) versus fallacy scenarios (columns)

Number of optimal answers								
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	5.67%	2.09%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
1	4.78%	3.58%	2.69%	0.60%	0.30%	0.00%	0.30%	0.00%
2	2.99%	5.37%	4.48%	1.49%	0.30%	0.00%	0.00%	0.00%
3	0.60%	3.58%	3.58%	2.99%	2.39%	0.60%	0.00%	0.00%
4	0.00%	0.60%	2.09%	3.88%	2.09%	1.19%	0.00%	0.00%
5	0.30%	0.30%	2.39%	0.90%	1.19%	1.79%	0.30%	0.00%
6	0.00%	0.00%	0.00%	2.09%	2.39%	1.49%	0.90%	1.49%
7	0.00%	0.30%	0.00%	0.30%	2.69%	0.60%	4.48%	17.91%
Wilcoxon rank sum test of differences in optimal answers control vs treatment: p-value $< 2.2 \cdot 10^{-16}$								

(b) Proportion of fallacy-implicated answers in control scenarios (rows) versus fallacy scenarios (columns)

Number of fallacy-implicated answers								
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	21.19%	14.33%	7.46%	4.78%	0.30%	0.00%	0.00%	0.00%
1	3.28%	8.06%	11.34%	9.55%	2.69%	0.00%	0.00%	0.00%
2	0.30%	2.69%	4.78%	5.97%	0.60%	0.00%	0.00%	0.00%
3	0.00%	0.30%	1.49%	0.00%	0.60%	0.00%	0.00%	0.00%
4	0.00%	0.00%	0.00%	0.30%	0.00%	0.00%	0.00%	0.00%
5	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
6	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
7	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Wilcoxon rank sum test of differences in fallacy-implicated answers control vs treatment: p-value $< 2.2 \cdot 10^{-16}$								

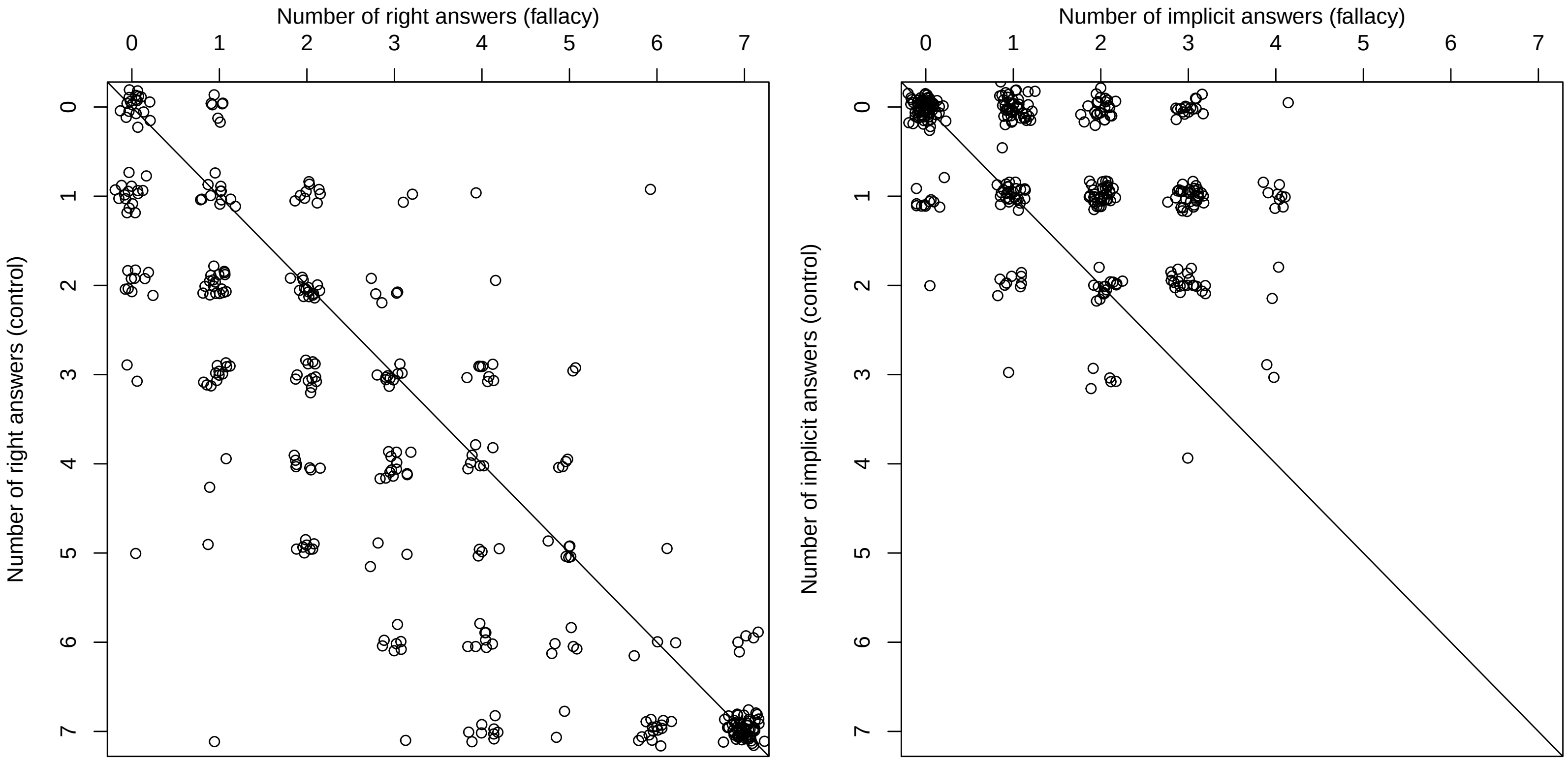


Figure I.3: This graph visualizes how many participants give certain answers in control (y-axis) vs fallacy scenarios (x-axis). The left graphic shows the number of optimal answers (option 5), the right graphic shows the number of fallacy-implicated answers.

We also investigate if the fallacies are correlated. Table I.7 shows the correlation matrix between choices of fallacy-implicated options in the fallacy scenarios, i.e. whether participants who choose a fallacy-implicated option for one particular heuristic tend to also choose the fallacy-implicated option in other fallacy scenarios. For most comparisons we cannot detect any significant dependencies between the fallacies. Only four correlations are significantly positive (between 1/N Heuristic, Interest Matching and Equal Start as well as between Cuckoo Fallacy and Balance Matching), but they are not particularly large (below 0.3). Thus, we cannot confirm clear linear dependencies between the fallacies.

Table I.7: Correlation matrix of fallacy-implicated answers in the fallacy scenarios

Correlation	**Cuckoo Fallacy**	**Equalize Balances**	**Complete Repayment**	**Balance Matching**	**1/N Heuristic**	**Interest Matching**	**Equal Start**
Cuckoo Fallacy	1	0.076	-0.068	0.125*	-0.042	-0.016	0.060
Equalize Balances	0.076	1	0.103	0.054	-0.042	0.002	0.049
Complete Repayment	-0.068	0.103	1	0.063	-0.013	-0.040	-0.012
Balance Matching	0.125*	0.054	0.063	1	-0.041	0.023	0.038
1/N Heuristic	-0.042	-0.042	-0.013	-0.041	1	0.156**	0.264***
Interest Matching	-0.016	0.002	-0.040	0.023	0.156**	1	0.220***
Equal Start	0.060	0.049	-0.012	0.038	0.264***	0.220***	1

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

In the next analysis we aim to identify groups of participants with similar answers to investigate whether the results are driven by a particular sub-population. We start by checking for each control scenario whether a participant selects the optimal option, the fallacy-implicated option, or any other non-optimal option combined. We then identify to which of these three possibilities the participant switches to (or stays) in the respective fallacy scenario. This allows us to identify nine distinct "types" of participants, e.g. participants who repay optimally in the control scenario and in the fallacy scenario (type 'optimal->optimal'), or participants who switch from optimal repayment in the control scenario to the fallacy-implicated option in the fallacy scenario (type ’optimal->implic’). In the next step, we count for each heuristic how many participants belong to a specific type and present the results as proportions in Table I.8. We compare the number of participants for the ’implic->optimal’ type and the ’optimal->implic’ type with binomial tests and report the p-values in the table. With the exception of Equalize Balances and Balance Matching, we detect large differences between the proportions of participants switching from optimal to fallacy-implicated option (row 4) and of participants that exhibit the reversed behavior (row 2) in all scenario pairs, which is in line with the differences we report in our main analyses (Interest Matching again shows the reversed sign, as it is the only pair where the value in row 4 is larger than in line 2).[8]

[8]Note that while we present switches in the direction from control to fallacy scenarios, the participants might also have answered the fallacy scenario first, depending on the random order of the scenarios.

Table I.8: Behavior in the scenario pairs[a]

Behavior	Cuckoo Fallacy	Equalize Balances	Complete Repayment	Balance Matching	1/N Heuristic	Interest Matching	Equal Start
implic->implic	1.19%	1.19%	5.07%	2.39%	1.19%	14.93%	14.33%
implic->optimal	1.79%	0.30%	0.60%	0.90%	0.30%	13.43%	1.79%
implic->other	0.60%	0.90%	1.49%	1.79%	1.49%	2.69%	3.58%
optimal->implic	13.43%	0.90%	9.55%	1.79%	17.01%	5.37%	10.75%
optimal->optimal	25.07%	56.42%	28.36%	43.88%	36.12%	47.76%	26.27%
optimal->other	5.67%	11.04%	5.07%	13.43%	24.18%	2.99%	6.87%
other->implic	14.33%	1.79%	9.55%	3.88%	6.27%	2.39%	16.72%
other->optimal	6.87%	10.75%	6.87%	8.66%	2.39%	2.69%	7.16%
other->other	31.04%	16.72%	33.43%	23.28%	11.04%	7.76%	12.54%
p-value	$\mathbf{3.1 \cdot 10^{-8}}$	0.616	$\mathbf{3.31 \cdot 10^{-7}}$	0.502	$\mathbf{4.15 \cdot 10^{-14}}$	$\mathbf{5.79 \cdot 10^{-4}}$	$\mathbf{3.80 \cdot 10^{-6}}$

[a] This table shows the proportion of participants exhibiting a certain behavior between control and fallacy scenario of a scenario pair. The behavior in the control scenario is denoted on the left before the '->', the behavior in the respective fallacy scenario is denoted on the right after the '->', where 'implic' means fallacy-implicated option, 'optimal' the optimal option (5) and 'other' every other option. The p-values refer to a binomial test of differences between the numbers of 'implic->optimal' and 'optimal->implic' participants for each scenario pair. We report significant p-values below 0.05 in bold to show that the numbers of participants switching from the optimal to the fallacy-implicated option differs from the participants switching the other way round.

In a final step, we use these nine types of participants to identify groups by employing a k-means cluster analysis. The cluster analysis helps us to identify a hypothetical "average" participant per group and use the information from the cluster to describe their typical decision-making more closely. We use the elbow criterion (Thorndike, 1953), Akaike's information criterion (Akaike, 1974) and the Bayesian information criterion (Schwarz, 1978) to determine the number of clusters. All three criteria are visualized in Figure I.4 for numbers of clusters between 1 and 30 (x-axis) and lead us to a choice of four clusters. To stabilize the clusters, we run the k-means algorithm 1000 times with different random starting values. We report the cluster centers of the four clusters in Table I.9. Each cluster center stands for the average participant in the respective cluster. The numbers in the cells denote in how many out of seven scenario pairs the average participant exhibits behavior of the respective type. For each column, we print the maximum number in bold as it drives the assignment to this cluster the most.

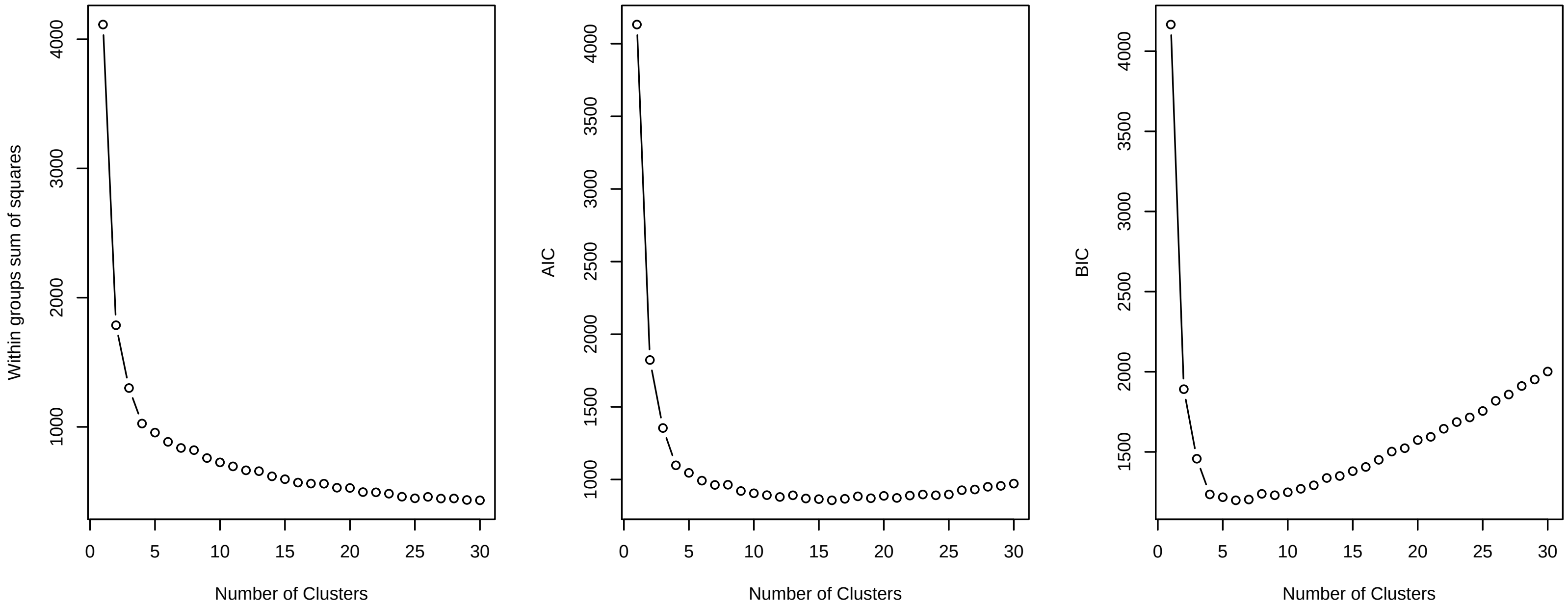


Figure I.4: This Figure shows three criteria to determine the number of clusters for a k-means clustering. The x-axes show the number of clusters between 0 and 30. The y-axes show the value of within groups sum of squares (elbow criterion, left figure), the AIC (Akaike's information criterion, middle figure) or the BIC (Bayesian information criterion, right Figure). Considering all three criteria we determined four as an appropriate number of clusters.

Table I.9: Description of cluster means[a]

Behavior	Cluster 1	Cluster 2	Cluster 3	Cluster 4
implic->implic	0.37	0.77	0.01	0.51
implic->optimal	0.21	0.16	0.02	0.37
implic->other	0.05	0.23	0.00	0.23
optimal->implic	0.82	0.35	0.23	0.93
optimal->optimal	**2.94**	0.20	**6.59**	0.53
optimal->other	1.05	0.48	0.05	1.18
other->implic	0.43	0.87	0.00	0.93
other->optimal	0.50	0.19	0.08	1.00
other->other	0.65	**3.76**	0.02	**1.31**
Cluster size	82	75	88	90
Within_SS	304.37	235.81	69.89	415.53
between_SS / total_SS	75.1 %			

[a] This table shows the cluster means of a k-means clustering with 1,000 random starting points. The columns show how many out of seven times a participants showed a specific behavior on average in each cluster. A number in bold stands for the maximum value in the respective cluster.

Analyzing the four clusters, the clearest assignment is to cluster 3 with 88 out of 335 participants. This cluster contains the optimally choosing participants. They choose the optimal option in the control scenarios in 6.77 (= $0.23+6.59+0.05$) out of seven control scenarios on average, and 6.69 times (= $0.02+6.59+0.08$) in the fallacy scenarios. They keep the optimal answer in 6.59 scenario pairs, and tend to correct the few errors they make. Out of the 0.13 times (= 0.01 + 0.02 + 0.00 + 0.00 + 0.08 + 0.02) they chose any non-optimal option in the control scenarios, they correct this error in 0.1 (= $0.02+0.08$) times in the fallacy scenario.

In contrast, the 82 participants assigned to cluster 1 seem to have a relatively good grasp on how to repay debts, but are vulnerable to fallacies. One average, they choose the optimal answer 4.81 times in the control scenarios, but only 3.65 times in the fallacy scenarios. They keep the optimal choice, provided they found it in the control scenarios, in only 2.94 times in the fallacy scenarios. Instead, they switch from the optimal to the fallacy-implicated answer in 0.82 times, and to any of the other three option in 1.05 times. They sometimes correct errors from the control scenario in the fallacy scenario (0.71 times in total), but these corrections do not offset the losses. On the other hand, they choose the fallacy-implicated option in only 0.63 control scenarios, but in 1.62 fallacy scenarios.

Cluster 4 (90 participants) seems to be similar to cluster 1, but with a much more erratic behavior, and starting from a lower level of optimality. Its participants choose the optimal option more often in the control scenarios than in the fallacy scenarios (2.64 to 1.9 times) too, and they show a vulnerability to getting distracted from the optimal option as well (0.93 times to the implicated option, 1.18 times to one of the other three). They also choose the fallacy-implicated option more often in the fallacy scenarios (2.37 times, vs. 1.11 times in the control scenarios). However, their erraticism also enables them to find the optimal decision in a fallacy scenario when they failed to do so in the control scenario relatively often (in around 1.37 of the 7 cases, compared to only 0.71 times for cluster 1).

The most striking feature of cluster 2 (75 participants) is that its participants rarely if ever find any optimal solution, be it in the control scenarios (1.03 times) or the fallacy scenarios (0.55 times). They are prone to fallacies and choose the implicated options in 1.99 fallacy scenarios but in only 1.16 control scenarios. Unlike cluster 4 however, they do not show the erraticism that helps them to correct errors (they switch from any of the four non-optimal options to the optimal option in only 0.35 fallacy scenarios).

It stands out that there is no specific cluster that shows switches to the fallacy-implicated option particularly often. While these switches do occur in the clusters 1, 2 and 4, and only cluster 3 seems to not be vulnerable for fallacies, the much more important features for the clusters seem to be optimality and consistency. This leads us to conclude that, while there are participants that choose optimally far more often than others, there is no particular group of participants who regularly choose fallacy-implicated options. At the same time, however, only a minority of participants seems to understand repayment problems well enough to resist fallacies. This supports our main results from I.4 where we have already shown that certain fallacies indeed lead to an increased number of participants choosing the fallacy-implicated option. The auxiliary results from this section do not allow us to pin this behavior to a distinct group of people, but the findings underline that a certain vulnerability to fallacies seems to be the norm rather than the exception.

I.7 Conclusion

Our experiment shows that the participants generally exhibit considerable amounts of misallocation and how different patterns of misallocation can be triggered by providing subjects with irrelevant pieces of information. Admittedly, our experimental design and interpretation of findings are based on the belief that participants have rational preferences for money but show mental gaps (Handel and Schwartzstein, 2018), which we use to trigger or manipulate different heuristics that lead to non-optimal decisions. We are, however, aware of alternative perspectives - and correspondingly deviating interpretations of our results - that deserve to be discussed.

First, one might argue that participants have rational preferences and do optimize, and that our findings are mere experimental artifacts. This implies that the utility function of our subjects comprises more arguments than just money. For instance, common problems of experimental studies such as experimenter demand effects, scrutiny effects or issues of stake size (Levitt and List, 2007; Zizzo, 2010) could overshadow preferences for money in our experiments. To counter this methodological objection, we will just briefly sketch three arguments and relegate a richer discussion to Appendix I. First, as our results are roughly in line with the literature from the field (Gathergood et al., 2019; Ponce et al., 2017) and earlier findings from the lab (Amar et al., 2011), this gives general support to our conclusions. Second, methodological standard objections do not easily explain the patterns we find in our data, but our heuristic approach does. And third, the evidence for methodological effects is mixed (e.g. Camerer (2015); Camerer and Hogarth (1999); Dhami (2016); Zizzo (2010)), and it is not clear why our experiments should suffer from methodological problems to such a degree that our results can be fully explained by them. For these reasons we do not think that our results are mere methodological artifacts.

A second alternative interpretation of our results could be that our participants have preferences that violate traditional assumptions of rationality, but still optimize given these non-standard preferences. This implies that in at least some parts of our experiments, subjects choose to misallocate because their preferences violate at least one rationality assumption. The most obvious of these possibly violated assumptions might be monotonicity - our participants choose to earn less than the maximum amount because they simply prefer to earn less. This, however, is not supported by our experimental evidence: If preferences are non-standard but at least stable between the scenarios, the different situational effects we find in our experiment should not matter, because this argument presupposes that our subjects know how to maximize the bonus, but deliberately choose not to do so. In our experiment however, the misallocation patterns change drastically, depending on the scenario.

Following up on this observation, one could argue that preferences for money are unstable. While this objection might be plausible in the long run, we do not believe that it can be applied to the short time horizon that our experiment covers. What is more, if we assumed that preferences change in every scenario, this would allow to explain any observation - without any scientific merit. For the same reason, we also do not entertain the final alternative perspective that participants neither have traditional rational preferences nor optimize.

This leaves us with the interpretation that our findings are more than experimental artifacts, and that they indeed reflect non-optimal decision-making. Taking our results seriously has several implications that we discuss briefly. One straightforward conclusion from our findings is that the misallocations observed by Gathergood et al. (2019) and Ponce et al. (2017) are not only caused by field aspects, but reflect deeper aspects of human decision-making. Differences in effect sizes to our results may nevertheless be explained by field effects and differences in the sample pools.

From a theoretical perspective, we argue that the results from both our study and from the broader literature which it complements (Amar et al., 2011; Gathergood et al., 2019; Handel and Schwartzstein, 2018; Ponce et al., 2017) show that it is important to incorporate systematic decision errors into models of financial decision-making (Köszegi and Rabin, 2008; Rabin, 2013). A large class of current (rational and behavioral) models uses some kind of obstacle in the utility function or side constraint in the budget restriction to explain deviations from simple rational choice predictions (Beshears et al., 2018). Their general argument is that people prefer to optimize, but the obstacle stops them from doing so. Examples for such obstacles are time, ego or other individuals.

Such "obstacle models" have shown empirical success (Beshears et al., 2018; Dhami, 2016), and since they are usually tweaked versions of standard economic models, they can rely on a rigorous, parsimonious and comprehensive theory of human behavior. Error-less rational choice also seems to be close to a standard of behavior most people prefer to achieve (Nielsen and Rehbeck, 2022). However, such models usually do not incorporate mental gaps or other errors. If obstacles and errors in decision-making are moreover correlated, such models might capture variance caused by errors and attribute it - wrongly - to the respective obstacle. To illustrate this argument, consider a recent paper by Enke and Graeber (2021). The authors use a distinct model of a mental gap, which they call "cognitive uncertainty" (see also Enke and Graeber (2023)). They develop and test a model where agents do not perfectly understand how to make decisions over time, and show that "(...) decisions associated with cognitive uncertainty *look like* they reflect very high impatience over short horizons. On the other hand, the inelasticity of observed choices with respect to the delay also means that cognitively uncertain decisions *look like* they reflect a lower degree of impatience over very long horizons" (emphasis in the original). This quote captures the essence of our argument - without accounting for errors in financial decisions, such as the ones we discovered, we run the risk of confusing mental gaps with, for example, non-standard preferences.

Our experimental setting can be easily expanded or adapted to financial investment decisions, or to test the relative strengths of repayment heuristics against each other. Furthermore, knowing or learning about fallacy-prone situations may be useful to create reminders for credit card debtors who find themselves in such a situation. All this could help broadening our understanding of basic financial decisions.

Chapter II

What could possibly go wrong? Nudging and the Cuckoo Fallacy

Coauthors:

Darwin Semmler

Christina E. Bannier

Relative share:

45%

Publication (jointly with chapter I):

Gärtner, Florian, Darwin Semmler, and Christina E. Bannier (2023), “What could possibly go wrong? Predictable misallocation in simple debt repayment experiments”, Journal of Economic Behavior & Organization, 205, 28–43

Previous versions of this chapter have been presented at:

- HypoVereinsbank Doktorandenseminar 2018
- Queen Mary University of London, Behavioural Finance Working Group Conference 2019
- Society for Experimental Finance Conference 2019
- Jahrestagung des Vereins für Socialpolitik 2020 (conjoint version with chapter I)
- ASSA/AEA Conference 2021, Poster Session (conjoint version with chapter I)

What could possibly go wrong? Nudging and the Cuckoo Fallacy

Abstract

We experimentally study a novel debt repayment heuristic, the "Cuckoo Fallacy", which is based on the amount of new debt rather than the interest rate. We show its existence, and demonstrate that simple framing can decrease repayment misallocation, nudging borrowers to more optimal behavior. Our results inform scholars and policy makers on how to improve household's financial decisions.

Keywords: Household finance, credit cards, financial literacy, rationality, bias, cuckoo fallacy

JEL-Codes: D14 - D91 - G41 - G50

Funding: This work was financially supported by the "Frankfurter Institut für Risikomanagement und Regulierung" (FIRM). FIRM had no involvement in anything study-related.

Declarations of interest: none

II.1 Introduction

Recent evidence shows that people often fail to repay their debts in an interest-minimizing and thus optimal way (see chapter I of this dissertation, but also Amar et al. (2011), Gathergood et al. (2019), Ozyılmaz and Zhang (2020) and Ponce et al. (2017)). This is important from a theoretical, but also from a practical point of view - if people make errors, one might want to help them avoiding misallocation. In principal, we see two broad ways to do that. First, one can educate people on how to find the correct solution, and give them guidance if they need it. In this case, the underlying assumption is that people do not know how to solve the repayment problem correctly, so education or guidance is needed. We adress this approach in chapter III. Second, one can help people avoid particular traps and fallacies. In this view, people generally know how to repay debts without misallocation, but some obstacle stops them, for example because the particular configuration of numbers in the repayment situation might trigger a wrong repayment heuristic - as we investigate in chapter I. But the latter problem might also include misguided attention to the wrong information pieces and therefore being influenced by framing effects (Tversky and Kahneman, 1981), which we investigate in the present chapter II. In particular, we explore ways to minimize certain decision errors stemming from the usage of one specific fallacy.

If we want to show that it is possible to circumvent a specific fallacy, we need to show the fallacy as a mechanism to misallocation, and that we can steer its occurrence. We focus on the "Cuckoo Fallacy", a novel fallacy that has proven to be particularly intriguing: Participants focus too strongly on the *amount of new debts* a card produces per round, rather than the *interest rate*. This triggers a repayment decision that is non-optimal if the low-interest rate card accumulates more new debt. We refer to this as

“Cuckoo Fallacy”, as it mirrors the behavior of parenting birds feeding the largest and most urgently pleading fledgling in their nest first, which might turn out to be a cuckoo.

Our experiment investigates the conditions for this fallacy with classical framing. We are particularly interested in whether we can frame the information environment - holding the values of all pieces of information constant - such that misallocation due to the fallacy is reduced. To do so, we use a simplified version of Amar et al. (2011)’s debt repayment game as a basis. The control group does not have any particular features intended to trigger or prevent the Cuckoo Fallacy. Additionally, we create two experimental treatments where we change the way we present the information about interest rates and balances. One treatment is supposed to protect from the fallacy (a "nudge treatment"), the other to increase misallocation (a "sludge treatment"). The sludge treatment tries to steer the attention to the possibly misleading amount of new debts a card will accumulate, thus triggering the Cuckoo Fallacy. The nudge treatment, in contrast, highlights the importance of the total money saved per round. We show that the nudge indeed decreases misallocation, but the sludge does not seem to work. These results are robust against control variables such as age, gender, and experience with credit cards. Interestingly, we find that financial literacy decreases misallocation, but does not interact with the treatments: Both treatments have similar effects regardless of how financially literate a subject is.

Just as the experiment in chapter I, our experiment contributes to the literature on consumer finance puzzles (Agarwal et al., 2015; Gorbachev and Luengo-Prado, 2019; Keys and Wang, 2019; Stango and Zinman, 2016), and non-optimal debt repayments in particular (Amar et al., 2011; Gathergood et al., 2019; Ozyılmaz and Zhang, 2020; Ponce et al., 2017). Additionally, we contribute to a strand of literature on improving financial decision-making. Two major approaches are financial education (Kaiser and

Menkhoff, 2020; Lusardi and Mitchell, 2014; Lusardi et al., 2020; Wagner and Walstad, 2019) and nudges (Thaler and Sunstein, 2021). We control for financial literacy and employ classical framing as a nudge to study its effects more closely. While nudging approaches have been successful in a variety of contexts (e.g. Benartzi and Thaler (1999); Blumenstock et al. (2018); Cai (2019); Choi et al. (2010); Frydman and Wang (2020); Gneezy and Potters (1997); Karlan et al. (2016)), evidence on the efficacy of framing is mixed (e.g. Beshears et al. (2017); Dimant et al. (2020)). Our work shows that nudging in an experimental setting is possible, which is an important first step for designing interventions in the field.

The remainder of this chapter is structured as follows: In section II.2 we describe the experimental setup, section II.3 shows the collection of the data and II.4 shows the results of the experiment. Section II.5 presents the robustness checks, section II.6 describes an additional experiment where we resolve the limiting factor of dependent experiment rounds, and section II.7 concludes.

II.2 Experimental setup and hypotheses

We provide our subjects with a fixed income on a checking account to repay debts on two credit cards that differ in their interest rates. The experiment lasts for ten dependent rounds, following Amar et al. (2011), where each new round starts with the card balances charging interest according to the last round's repayment decision. As a consequence, compound interests amplify the financial effects of misallocation, particularly from non-optimal repayment decisions in the early rounds. One credit card charges an interest rate of 3% per round, the other 5%. The checking account pays no interest. At the start of the experiment, both credit card accounts hold a negative balance of $2200, and the checking account an income of \$250. We let participants distribute their income freely on any account in each round.[1] Participants finalize their decision by actively finishing the current round and thereby entering the next round. Parallel to the interest being calculated and added to the card balances for the next round, the checking account is refilled with \$250 of income. After round 10, interests are calculated and added to the debt balances one final time. To rule out order effects, we assign the interest rates to the two credit cards randomly between the participants. For each participant, the order is stable. We pay \$1 as a show-up fee, and up to \$2 as a bonus.

For our dependent variable we define the *misallocation* of subject i (MA_i) as the percentage of available money that is *not* transferred to the high-interest rate credit card. Thus, MA_i is a value between 0 and 1. It is 0 if and only if participant i repays all the money to the high-interest rate card in all decisions rounds. We call this the "optimal behavior".[2] In order to translate the misallocation into a bonus payment, we employ the repayment efficiency, which is a percentage measure of how close the final debt balance comes to the minimum amount of debt (achieved by repaying optimally) relative to the maximum amount of debt (by not repaying at all).[3] If no money is left on the checking account, misallocation and repayment efficiency are linear dependent. However, if participants leave money on the checking account, they differ slightly, because

[1]There are no incentives to not repay debts, but technically it is possible to leave money on the checking account.

[2]We prove in Appendix II that this is a dominant strategy.

[3]Let *min* be the minimal possible amount of debts at the end of the experiment, which is the result when repaying optimally over all rounds (-\$2988.51), and *max* the maximal possible amount of debts, which is the result when nothing is ever repaid (-\$3790.20). Let *debt* denote the actual amount of debts a participant has at the end of the experiment, then the repayment efficiency is defined as $eff = 1 - \frac{debt - min}{max - min}$. We decided to use repayment efficiency to calculate bonuses for two reasons: First, we want to avoid having to explain misallocation to the subjects as it would already imply that there exists such a thing as "the right" credit card to repay. Secondly, repayment efficiency is directly bound to the overall goal of our participants to reduce the sum of the debt in the end, so it is just the logical monetary manifestation of our established incentives.

misallocation treats this as equally wrong as repaying on the low-interest card, while repayment efficiency differentiates these actions. The bonus is calculated as repayment efficiency multiplied by $2. As an example, consider a participant who repays optimally in each round. This leads to the minimal total final debt possible after the experiment, which is $2988.51. This participant has a repayment efficiency of 100%, and thus earns 100% of the bonus ($2). Now consider a participant who did not repay anything at all, not even on the low-interest card, and instead left everything on the checking account. They finish the experiment with the maximum possible amount of total debts, which is $3790.20. This person has a repayment efficiency of 0%, and earns no bonus at all.

The experiment consists of three treatments: A "nudge" treatment (Thaler and Sunstein, 2021) to decrease misallocation, a "sludge" treatment (Thaler, 2018) to increase it, and a control group as basic treatment. The nudge treatment tests if we can reduce misallocation. But while behavioral interventions usually intend to improve people's decisions, such interventions can backfire if they are designed poorly, might have unintended side effects (Medina, 2021), or can even be used to actively worsen decisions. We attempt to understand the potential magnitude of such problems by making use of the sludge treatment.

All treatments are based on the same set of information, but differ in the way this information is presented. They are designed as follows:

- The participants in the control group ("Basic treatment") see all three account balances and the two interest rates.

- The "ShowNewDebts" treatment is our sludge treatment. Instead of the interest rates in percent, we show the amount of new debts per card, given the chosen repayments so far. We also color the information on new debts in red in order to emphasize its importance (Bazley et al., 2021). We still present the interest rates, but in a less accessible manner: We show them once before the experiment rounds, and then hide them behind a button. Subjects can press this button at any time without any costs, but this should still work as a sludge.

- The "ShowSavedMoney" treatment is our nudge treatment where we try to decrease the misallocation by shifting the focus away from the new debts. Instead of displaying the balances of each credit card account, we only show the total debt, and hide the individual account balances during the rounds behind a button. This should make it harder to calculate or estimate the amount of new debts. We also present the interest rates as cents that can be saved in the next round for each dollar repaid in the current round, to shift attention to the interest rates.[4] Additionally we color the sum of the saved money in green.[5]

We use these treatments to test the following hypotheses:

H2.1: The misallocation in the ShowSavedMoney treatment is lower than in the Basic treatment.

H2.2: The misallocation in the ShowNewDebts treatment is higher than in the Basic treatment.

[4] For instance, instead of showing "3%" we write "For each dollar you repay on credit card 1, you will save 3 cents interest for the next round".

[5] This treatment is unusual as, compared to the basic treatment, we change several things at once. However, we are interested whether we can draw attention from the new debts at all, and not which particular change might be successful. Only for the latter question we would need to design several treatments to test all changes independently.

II.3 Data

We use the platform SoPHIE (Hendriks, 2012) to run the experiment and the crowdsourcing platform Amazon MTurk to recruit our participants. The subjects (Turkers) participate in our experiment via a Human Intelligence Task (HIT), where we can approve or reject their submission. We restrict the pool of Turkers to US Americans who have completed at least 100 HITs of which we require at least 95% to be approved as suggested by Peer et al. (2014). This ensures that the participants have substantial experience with the platform. 527 MTurkers started our experiment, of which 414 finished it. Out of the 113 who did not finish the experiment, 89 dropped out before the basic numeracy question, 36 did not pass the basic numeracy question, and 15 dropped out within the experiment or the post experimental questionnaire - 4 in each the Basic and the ShowNewDebts treatment, and 7 in the ShowSavedMoney treatment. Out of the remaining 414 participants, 404 passed both attention tests. These 404 participants form our sample. It should be noted that we recruit participants for the individual treatments in separate HITs on MTurk at the same time and using the same wording. As subjects cannot differentiate between the treatments, this should rule out selection effects. The data was collected in August and September 2018. On average our subjects earned \$2.80 in 19:01 minutes, implying an average hourly payment of around \$8.83.

We divide the experiment into three stages. In the first stage we explain the rules to the participants and use comprehension tasks to ensure participants read the rules properly. The subjects also have to calculate 1% of \$1000 to proceed to ensure basic understanding of interest rates. In the second stage - the experimental stage - participants have to make their decisions in 10 dependent experiment rounds. The last stage is the post experimental questionnaire (PEQ), where we collect demographics and other control variables. We ask for gender ("female" used as reference category), age, number of years of education, and measure financial literacy as number of right answer out of six questions. Three of them are the "Big Three" (Lusardi and Mitchell, 2011) and the other three are specifically about debt (Lusardi and Tufano, 2015). We employ an open question for the strategy used to exclude bots, rated by two different researchers, and furthermore include two attention checks, which we use to screen out everyone who does not answer correctly. Furthermore, we ask participants how many credit cards they own, and how many they additionally have access to (for instance via spouse) in order to measure credit card experience. We also ask them if they use credit cards at work, if they usually do not employ credit cards, or both together. "Credit card order" is a dummy variable indicating whether the more expensive card was the upper card on the experimental screen. Table II.1 presents the summary statistics - overall and for each treatment.

Table II.1: Summary statistics of participants

Overall statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Financial Literacy	404	3.73	1.38	0	3	4	5	6
Age	404	37.10	10.69	19	29	35	44	75
# Credit cards	386	2.65	2.68	0	1	2	4	20
# Additionally accessible credit cards	382	0.66	1.25	0	0	0	1	10
# Years of education	404	15.28	2.32	9	14	16	16	21
Experiment duration (min:sec)	404	19:01	08:35	04:57	13:05	16:56	23:20	58:08
Payoff (USD)	404	2.80	0.24	1.00	2.72	2.84	2.96	3.00
Gender info	Males: 213		Females: 190		Third gender: 1			
Basic treatment	**N**	**Mean**	**St. Dev.**	**Min**	**Pctl(25)**	**Median**	**Pctl(75)**	**Max**
Financial Literacy	131	3.75	1.37	0	3	4	5	6
Age	131	36.48	10.49	19	29	35	41.5	75
# Credit cards	126	3.01	3.51	0	1	2	4	20
# Additionally accessible credit cards	124	0.60	1.07	0	0	0	1	4
# Years of education	131	15.69	2.50	11	14	16	17	21
Experiment duration (min:sec)	131	17:42	08:13	05:53	12:41	15:16	21:08	50:57
Payoff (USD)	131	2.78	0.27	1.00	2.68	2.82	2.93	3.00
Gender info	Males: 72		Females: 58		Third gender: 1			
ShowNewDebts-treatment	**N**	**Mean**	**St. Dev.**	**Min**	**Pctl(25)**	**Median**	**Pctl(75)**	**Max**
Financial Literacy	135	3.67	1.44	0	3	4	5	6
Age	135	36.48	10.27	19	28	34	43.5	65
# Credit cards	126	2.13	1.92	0	1	2	3	10
# Additionally accessible credit cards	127	0.65	1.46	0	0	0	1	10
# Years of education	135	15.10	2.20	9	14	15	16	21
Experiment duration (min:sec)	135	18:47	08:18	04:57	13:02	16:59	22:46	52:03
Payoff (USD)	135	2.76	0.25	1.36	2.69	2.81	2.88	3.00
Gender info	Males: 73		Females: 62		Third gender: 0			
ShowSavedMoney-treatment	**N**	**Mean**	**St. Dev.**	**Min**	**Pctl(25)**	**Median**	**Pctl(75)**	**Max**
Financial Literacy	138	3.77	1.34	0	3	4	5	6
Age	138	38.28	11.25	22	29.2	36	45	70
# Credit cards	134	2.81	2.33	0	1	2	4	11
# Additionally accessible credit cards	131	0.73	1.20	0	0	0	1	8
# Years of education	138	15.06	2.20	9	13.2	15	16	21
Experiment duration (min:sec)	138	20:36	09:06	08:26	13:55	19:23	25:30	58:08
Payoff (USD)	138	2.85	0.20	1.21	2.77	2.88	3.00	3.00
Gender info	Males: 68		Females: 70		Third gender: 0			

II.4 Results

Table II.2 gives a first overview on the misallocation in the different treatments. Overall, around one quarter of the income is misallocated to the low-interest rate credit card on average, varying between 18.7% (ShowSavedMoney treatment) and 31.2% (ShowNewDebts treatment). The control group in the Basic treatment shows an average misallocation of 27.4%, which is above the nudge but below the sludge treatment. In addition, all treatments show variations in misallocation that cover the full interval between 0 and 1, implying that there are participants who consistently repay on the same card in all rounds. Finally, the ShowSavedMoney treatment is the only treatment in which more than one quarter (26.8%) of all the participants repay optimally over all rounds, while this share is only 11.1% in the ShowNewDebts treatment and 18.3% in the Basic treatment.

Table II.2: Misallocation, and the share of optimally repaying subjects in the treatments

Misallocation	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max	*Share of optimal-repaying subjects*
All data	404	0.257	0.202	0.000	0.080	0.295	0.378	1.000	*0.188*
Basic treatment	131	0.274	0.203	0.000	0.100	0.300	0.400	1.000	*0.183*
ShowNewDebts	135	0.312	0.202	0.000	0.230	0.314	0.388	1.000	*0.111*
ShowSavedMoney	138	0.187	0.180	0.000	0.000	0.147	0.335	1.000	*0.268*

As we measure misallocation for each participant in ten consecutive, dependent rounds, an analysis of temporal effects on this variable might be fruitful. Figure II.1 (left) shows the development of the average misallocation over the individual rounds of the experiment. As can be seen from the Figure, the average misallocation indeed increases, with an especially strong effect in the later rounds of the experiment. While this development shows in all treatments, it is comparably mild in the ShowSavedMoney treatment. The ShowNewDebts treatment, in contrast, in addition exhibits a particularly distinct zigzag pattern from round 6 on. This may be seen as a first indication of the Cuckoo Fallacy, as this fallacy implies that subjects switch their repayment to the cheaper credit card once this card starts to accumulate higher new debts. Since this requires that the balance on the high-interest rate card has to be reduced sufficiently relative to the low-interest card, this cannot happen before round 6 given our chosen specification values. Indeed, if a subject follows the "repay the card which produces more new debt"-heuristic perfectly, they would start by repaying the more expensive card first and then switch to the cheaper card in rounds 6, 8 and 9, indicating a zigzag pattern. These first observations hence lend credence to the assertion that the experiment is able to induce misallocation following from the Cuckoo Fallacy. At the same time, they imply that analyses of this misallocation need to account for the temporal dependence in the experiment.

In order to test the framing effect on the Cuckoo Fallacy conclusively, multivariate analyses hence have to take the dependency of rounds into account. We do so by controlling for the fact that this fallacy may not arise in all experimental rounds and measure the effect size of framing only in situations where the fallacy can occur. We therefore divide the treatment samples into rounds where either the expensive rate card produces more new debt (earmarked via the indicator variable $high\_int\_class = 1$) or where the cheap card does so ($high\_int\_class = 0$). The Cuckoo Fallacy is possible if and only if $high\_int\_class = 0$. We then include this dummy as our main explanatory variable of interest in a regression with MA_i as dependent variable. We report results from a minimal model which includes only the indicator variable $high\_int\_class$ as well as the treatment, a maximum model with all control variables, and the AIC-"optimal" model (Akaike, 1974). In each model, we adjust the p-values using the Bonferroni-

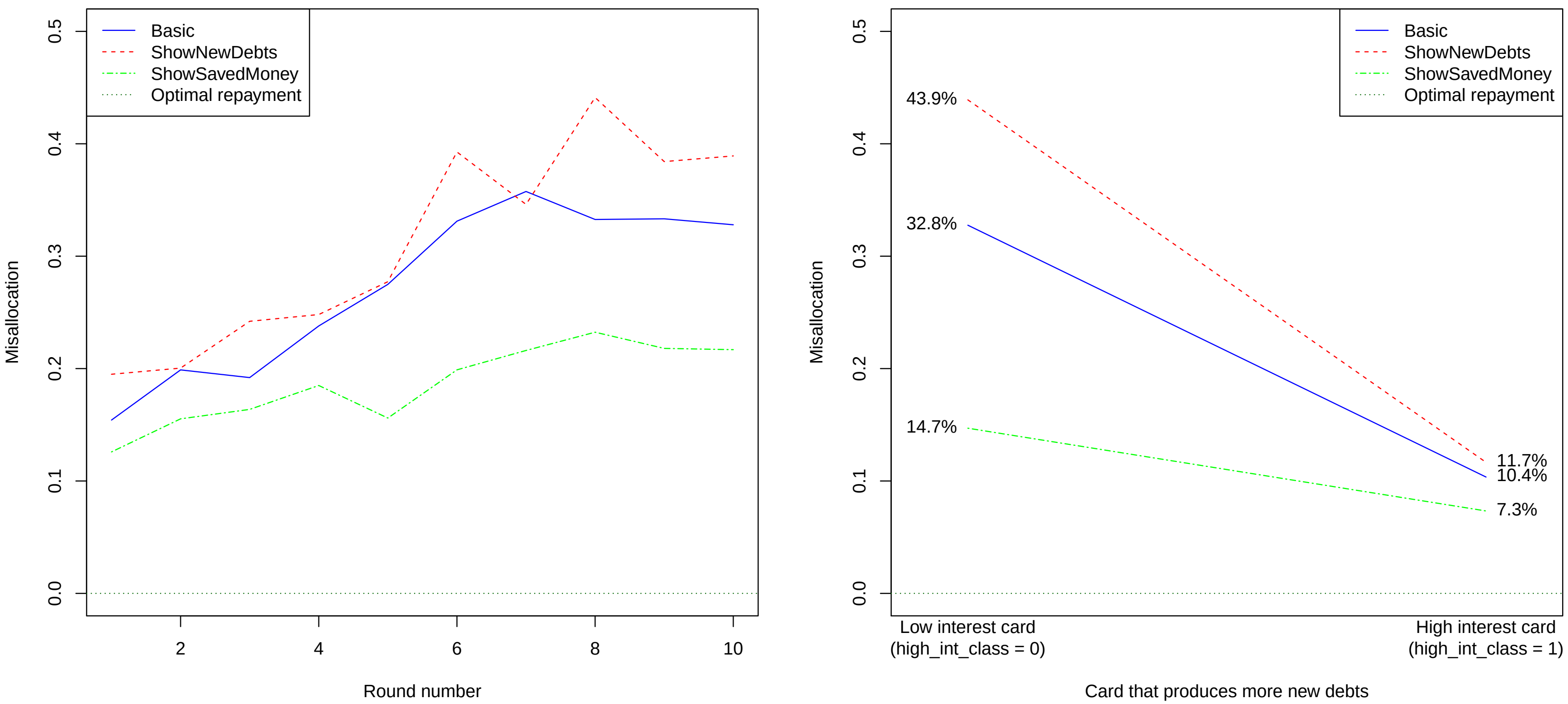


Figure II.1: The left Figure shows the development of the misallocation per treatment in the experiment rounds. The right Figure shows the interaction plot between treatment and interest class (i.e., which credit card accumulates more interest). ShowNewDebts (in red) is the sludge and ShowSavedMoney (in green) is the nudge treatment.

Holm method for all coefficients that we interpret (and report). Table II.3 and the right part of Figure II.1 show the results.[6]

As can be seen from Table II.3, the misallocation is significantly lower in rounds where the Cuckoo Fallacy is not possible, i.e. where *high_int_class* = 1. Also, the ShowSavedMoney treatment is associated with significantly lower misallocation compared to the Basic treatment (the omitted category in the regressions). Furthermore, its interaction with the *high_int_class* indicator variable shows a significantly positive coefficient in all models. Stated differently, in rounds where the Cuckoo Fallacy is possible, i.e. where *high_int_class* = 0, the nudge treatment significantly reduces the misallocation. The coefficient of the ShowNewDebts treatment variable, in contrast, is not significant in any model, as well as its interaction with the *high_int_class* indicator variable. Altogether, there is hence strong evidence that the ShowSavedMoney treatment indeed decreases the Cuckoo Fallacy (H2.1), but no evidence that the ShowNewDebts treatment increases it (H2.2).

However, we want to stress that under a different analysis, which would ignore multiple hypothesis testing *and* use one sided tests - this can be justified because H2.2. is a directed hypothesis - the ShowNewDebts coefficients in models 1 and 2 would be significant. We highlight this, because we view a sludge treatment as generally undesirable that should therefore be avoided. In such a situation, one would want to weigh up the advantages of the classical significance analysis, including a 5% significance level and multiple hypothesis testing, with the potential danger of committing a type II error by interpreting results too conservatively. The magnitude of the sludge treatment effect is still around 10 percentage points in all three models, and its statistical insignificance might come from a lack of power. If the effect exists, it has the potential to increase

[6]Also see Table Appendix II.24 for the full set of control variables.

misallocation quite strongly.

Figure II.1 (right) illustrates these findings nicely: Participants misallocate 43.9% of the available money in the ShowNewDebts treatment on average over all rounds where the Cuckoo Fallacy is possible, i.e. when the low-interest rate card produces more new debt than the high-interest rate card. In rounds where the Cuckoo Fallacy is not possible, only 11.7% are misallocated in this treatment. In the ShowSavedMoney treatment, in contrast, the average misallocation in rounds where the Cuckoo Fallacy is possible is 14.7%, whereas the average misallocation is 7.3% if it is not. We hence conclude that the Cuckoo Fallacy is weaker in the ShowSavedMoney treatment, even after controlling for the dependencies of rounds, and that we do not find an effect of the ShowNewDebts treatment. But the latter interpretation comes with the aforementioned caveat.

The regression results in Table II.3 also indicate that financial literacy is only weakly and not significantly negatively associated with misallocation. The complete table (Table Appendix II.24) also demonstrates that there is no significant interaction effect of financial literacy with any of the treatments. Apart from years of education, which is significantly negatively related with misallocation, there are no further significant effects of the other control variables, including the ones that approximate credit card experience.

Overall, our results hence support hypothesis H2.1, but not H2.2. The nudge is effective in manipulating misallocation, and the sludge is not - but the latter result should be interpreted carefully. In general, we conclude that framing is indeed relevant for credit card repayment and that it can be effectively employed to remedy the problem of misallocation, while it is harder to worsen misallocation.

Table II.3: Misallocation split by round class[a], OLS regression

	Dependent variable: Misallocation		
	Minimal	Akaike-optimal	Full model
	(1)	(2)	(3)
High_int_class	−0.224***	−0.224***	−0.219***
	(0.043)	(0.042)	(0.042)
	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]
ShowNewDebts	0.112	0.102	0.094
	(0.059)	(0.059)	(0.063)
	[0.057]	[0.080]	[0.134]
	[0.113]	[0.240]	[0.269]
ShowSavedMoney	−0.181***	−0.188***	−0.174**
	(0.046)	(0.0456)	(0.053)
	[0.000]	[0.000]	[0.001]
	[0.000]	[0.000]	[0.005]
High_int_class · ShowNewDebts	−0.098	−0.098	−0.098
	(0.062)	(0.061)	(0.063)
	[0.112]	[0.109]	[0.120]
	[0.113]	[0.240]	[0.360]
High_int_class · ShowSavedMoney	0.150**	0.150**	0.139*
	(0.049)	(0.048)	(0.049)
	[0.002]	[0.002]	[0.005]
	[0.006]	[0.007]	[0.019]
Financial literacy		−0.015	−0.023
		(0.009)	(0.016)
		[0.089]	[0.148]
		[0.240]	[0.360]
Constant	0.328***	0.523***	0.574***
	(0.040)	(0.083)	(0.095)
Observations	522	522	498
Interact. Fin.lit._treatments	No	No	Yes
Further control variables	No	only YOE[b]	Yes
R^2	0.230	0.245	0.246
Akaike Inf. Crit.	5.17	−0.83	17.95
F Statistic	21.606*** ($df = 5; 516$)	17.641*** ($df = 7; 514$)	7.251*** ($df = 17; 480$)

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$
Financial literacy is centralized at a value of 3.

[a] This table shows the misallocation when the Cuckoo Fallacy is possible (High_int_class = 0) vs. when it is not (High_int_class = 1). It also shows how it changes depending on the experimental treatments, and the interaction between these two variables. ShowNewDebts is the sludge and ShowSavedMoney is the nudge treatment. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for High_int_class, ShowNewDebts, ShowSavedMoney, High_int_class · ShowNewDebts, High_int_class · ShowSavedMoney and Financial literacy reported coefficients. Asterisks indicate significance after adjustment.

[b] Years of education

II.5 Robustness checks

The split of our data according to the indicator variable *high_int_class* might be seen as problematic because not all participants appear evenly in both groups. To test the robustness of our results, we therefore also run regressions without this explanatory variable. In a first robustness check, we hence pool the experiment rounds and run the earlier regression with the treatment indicator variables and controls (Table Appendix II.25). In a second analysis, we assume a linear relation between rounds and misallocation and include the round number as a further numeric variable to test the robustness against a possible learning effect during the ten rounds (Table Appendix II.26). In the latter regression we use random intercept terms, as we have multiple observations per participant. Both sets of results show that the ShowSavedMoney treatment exhibits significantly lower misallocation than the control group. The ShowNewDebts treatment, in contrast, has no significant effect. These additional findings support our earlier conclusion, that the sludge treatment is less effective than the nudge treatment, if it is effective at all. It should be noted, however, that weaker test power should be expected in these additional analyses, because the Cuckoo Fallacy can only occur in the later rounds, and hence in fewer cases, of the experiment.

We repeat the main analysis from Table II.3 with the screened out participants and report the results in Table Appendix II.27. This does not change our overall results.

We also replicate the basic treatment in a lab experiment (N=96) in an attempt to rule out that the observed misallocation itself is driven by the MTurk subject pool (see Appendix II for details). In the replication, we increase participants' incentives via a 4 Euro flat payment and pay up to 10 Euro as bonus. In the lab experiment, we also prohibit subjects to leave money on their checking account - a change which is conservative, as it renders some non-optimal behavior completely impossible. Surprisingly, the average misallocation in the lab is 5.4 percentage points higher (significant, p=0.0242) than on MTurk. We conclude from this observation that misallocation as a whole seems not to be driven by an MTurk effect.

II.6 An additional experiment with independent rounds as a robustness check

A potentially important limiting factor of this experiment design might be that subjects are fully comparable only at the beginning of the experiment. This is because repayment decisions in the earlier rounds determine whether the Cuckoo Fallacy becomes possible at all. More precisely, participants who repay sufficiently non-optimally from the very beginning often do not even get the chance to succumb to the Cuckoo Fallacy. This casts doubts on the internal validity in a very specific way - the differences between the treatments might be endogenous selection effects. However, real credit card repayments often *are* endogenous, since one important features of credit cards is that their debt is revolving, and credit card users often borrow from credit cards knowing perfectly fine that the height of their debts depends on their repayments in earlier time periods. If we ignore revolving as a feature, we not only lose external validity, but also internal validity in this regard. In particular, it could be the case that the dependency itself influences our participants' reactions to the nudge and the sludge. This is the reason why we focus heavily on the dependent rounds design, but to solve this dilemma, we run an additional experiment which uses ten independent rounds as a robustness check.

The setup of this altered experiment is as identical to the main experiment as possible. The values for income and interest rates stay at \$250, 3% and 5%, respectively. The first round also starts with \$2200 of debts on each card. The major change is that from round 2 on, participants go through a series of 9 independent decision problems where the values from the former rounds are not carried over. The decisions differ in the balances. In three of these rounds, the balance of the low-interest card is sufficiently high such that the Cuckoo Fallacy is possible (*CuckooPossible*). In three others, the lower interest card has a higher balance, but not high enough for the fallacy to occur ($card3p > card5p$). We implement these additional rounds to distinguish the effects the Cuckoo Fallacy from a simple "repay the higher balance card" heuristic. In the final three, the high-interest card has the higher balance, and the difference in the balances is of a similar magnitude as in the three Cuckoo Fallacy decision problems ($card5p > card3p$). Table II.5 shows the details. In the experiments we randomize the order of the 9 latter rounds, but to stay close to the main experiment, we only randomize the credit card order for the first round and then keep it the same for all other rounds. Additionally, while in the main experiment participants can see the results of their actions in the changes of the balances at the start of the next round, in this altered experiment we do not give them any information after the rounds. Instead we show them the total results after all ten rounds are finished. We implement this change to minimize any dependencies between the rounds, such as learning, as much as possible.

Outside of the experimental stage, we only make minor changes to the wording in the instructions and adapt the second comprehension task to the new instructions. We keep the three treatments and the post experimental questionnaire, and pay a \$1 show-up fee and up to \$2 as bonus, which is again calculated via the repayment efficiency.

805 MTurkers started our experiment, of which 660 finished it. Out of the 145 who did not finish the experiment, 37 dropped out before the basic numeracy question, 55

did not pass the basic numeracy question, and 53 dropped out within the experiment or the post experimental questionnaire - 17 in each the Basic and the ShowSavedMoney treatment, 19 in the ShowNewDebts treatment. Out of the remaining 660 participants, 496 passed both attention tests. This number already indicates that the data quality of this sample, which we collected in December 2021 and January 2022, is lower than the sample from the main experiment, which we collected more than three years earlier and where we only lost 10 participants to the attention checks. Other authors find such drops within the same time frame as well (Chmielewski and Kucker, 2020). The open anti-bot question in which we asked for the repayment strategies shows additional problems, which it did not in the main experiment, because there the few suspicious participants were already screened out due to the other data quality measures. In this additional experiment however, a large number of participants who passed the automatic screening process gave answers that did not fit the question ("i choose with my own perfection"), that were generic answers such as "good survey", "no" or "very interesting", or that clearly showed that the respective participant is not fluent in English. Others described not how they repaid in the experiment but how they generally think one uses credit cards (e.g. "Pay off your balance every month"), and some even copied the first sentences of some Google search results ("Two of the most popular strategies for paying off debt on your own are the snowball method and the avalanche method. Both methods require making the minimum monthly payments on all but one debt, which you put extra money towards" [remark: comment ends here]). Two raters went over the answers independently to mark them as "suspicious" based on these problems. For our main analysis we only use the data from participants who none of the raters marked as suspicious. These 291 participants form our main sample. We report summary statistics in Table II.4. As a robustness check we add the participants which only one rater found suspicious. We dropped everyone who both raters marked as suspicious.

Table II.4: Summary statistics of participants (additional experiment)

Overall statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Financial Literacy	291	3.54	1.38	0	3	4	5	6
Age	291	36.86	10.33	20	30	35	42	78
# Credit cards	268	2.27	1.82	0	1	2	3	12
# Additionally accessible credit cards	255	0.69	1.30	0	0	0	1	10
# Years of education	291	15.47	2.17	9	14	16	16	21
Experiment duration (min:sec)	291	19:13	13:08	4:11	10:59	14:37	22:22	100:58
Payoff (USD)	291	1.72	0.33	0.00	1.62	1.76	2.00	2.00
Gender info	Males: 168		Females: 121		Third gender: 2			
Basic treatment	**N**	**Mean**	**St. Dev.**	**Min**	**Pctl(25)**	**Median**	**Pctl(75)**	**Max**
Financial Literacy	95	3.46	1.48	0	2	4	5	6
Age	95	35.62	8.12	23	30	34	39	62
# Credit cards	85	1.96	1.73	0	1	2	2	12
# Additionally accessible credit cards	77	0.52	1.01	0	0	0	1	6
# Years of education	95	15.55	1.95	10	14	16	16	21
Experiment duration (min:sec)	95	18:44	11:40	4:11	10:58	14:32	22:16	66:46
Payoff (USD)	95	1.66	0.36	0.00	1.56	1.73	1.86	2.00
Gender info	Males: 57		Females: 37		Third gender: 1			
ShowNewDebts-treatment	**N**	**Mean**	**St. Dev.**	**Min**	**Pctl(25)**	**Median**	**Pctl(75)**	**Max**
Financial Literacy	94	3.66	1.29	1	3	4	5	6
Age	94	36.01	10.56	20	28.2	34	40.8	78
# Credit cards	87	2.56	2.02	0	1	2	3.5	10
# Additionally accessible credit cards	86	0.64	1.18	0	0	0	1	6
# Years of education	94	15.60	2.22	9	14	16	16	21
Experiment duration (min:sec)	94	17:50	13:36	4:59	10:42	13:47	20:37	100:58
Payoff (USD)	94	1.75	0.32	0.02	1.65	1.79	2.00	2.00
Gender info	Males: 55		Females: 39		Third gender: 0			
ShowSavedMoney-treatment	**N**	**Mean**	**St. Dev.**	**Min**	**Pctl(25)**	**Median**	**Pctl(75)**	**Max**
Financial Literacy	102	3.51	1.36	0	3	4	4.8	6
Age	102	38.78	11.69	21	30	37	45.5	77
# Credit cards	96	2.27	1.68	0	1	2	3	8
# Additionally accessible credit cards	92	0.89	1.57	0	0	0	1	10
# Years of education	102	15.28	2.32	10	14	16	16	21
Experiment duration (min:sec)	102	20:54	13:53	6:44	11:53	16:27	24:20	88:29
Payoff (USD)	102	1.75	0.31	0.02	1.64	1.77	2.00	2.00
Gender info	Males: 56		Females: 45		Third gender: 1			

Because the data quality is such an obvious confounder, we refrain from comparing the results from the main to the additional experiment and only focus on investigating our hypotheses. Unlike in experiment #1, we interpret both adjusted and unadjusted p-values, because we want to highlight that the combination of problematic data and the test-power-draining adjustment procedure increases the probability for a type II error considerably.

To investigate the effects of our treatments on the Cuckoo Fallacy, we mirror the comparison between *low_interest_card* and *high_interest_card* from the main experiment, but this time with the three types of rounds instead of two. We compare the three rounds where the fallacy is possible first with the three rounds where the low-interest card has a higher balance but the fallacy is not possible, and second with the three rounds where the high-interest card has the higher balance, and then interact these variables with the treatments. Figure II.2 shows that the misallocation decreases in all setups when the Cuckoo Fallacy is not possible compared to when it is possible. Table II.5 shows that this decrease is significant in the control group, which strongly suggests that the Cuckoo Fallacy has an additional effect beyond a simple balance effect. While we do not see any significant difference for the ShowNewDebts treatments, at least some models suggest after Bonferroni-Holm correction that the Cuckoo Fallacy is less of a problem in the ShowSavedMoney treatment (misallocation 16.7%) than in the control group (misallocation 23.9%). All models show significant unadjusted p-values, and three of these significances survive the multiple-hypothesis adjustment. The interactions between treatment and Scenario type show significantly weaker effects for the ShowSavedMoney treatment if we ignore adjusting, but with the Bonferroni-Holm correction, the significance in the full model survives the adjustment only when we screen out participants with at least one "suspicious" rating. Hence, the decrease when the Cuckoo Fallacy is possible compared to when the higher interest rate card has the higher debt balance might be a bit weaker in the ShowSavedMoney treatment, but since these effects often fail the multiple-hypothesis adjustment, we interpret this as weak evidence.

Figure II.2 visualizes these results: Misallocation is highest when the Cuckoo Fallacy is possible and lowest when the 5% card has a higher balance than the 3% card, and it is lower in the ShowSavedMoney treatment. Interestingly, in this version of the experiment the sludge treatment also has a lower misallocation, even if it is not significantly lower than in the control treatment. This might indicate that the potential for problems or abuse is not that high. To conclude, the results of the additional experiment show that the Cuckoo Fallacy is an important driver of misallocation, and there are clues that we can manipulate the presentation to make it less likely. However, the evidence for the latter claim is weaker than in the main experiment.

Table II.5: Misallocation in additional experiment, random effects regression[a]

	Dependent variable: Misallocation					
Model	Minimal	Minimal	Akaike-optimal	Akaike-optimal	Full model	Full model
Outscreening	Strict	Tolerant	Strict	Tolerant	Strict	Tolerant
	(1)	(2)	(3)	(4)	(5)	(6)
Treatments						
ShowNewDebts	−0.045	−0.035	−0.037	−0.025	−0.029	−0.035
	(0.030)	(0.030)	(0.028)	(0.028)	(0.029)	(0.027)
	[0.129]	[0.242]	[0.187]	[0.363]	[0.323]	[0.198]
	[0.517]	[0.968]	[0.747]	[1.000]	[0.970]	[0.793]
ShowSavedMoney	−0.071	-0.076^{*}	−0.069	−0.067	-0.072^{*}	-0.070^{*}
	(0.029)	(0.029)	(0.028)	(0.028)	(0.025)	(0.025)
	[0.013]	[0.008]	[0.014]	[0.015]	[0.004]	[0.004]
	[0.080]	[0.046]	[0.086]	[0.091]	[0.024]	[0.026]
Scenario types						
card3p>card5p	-0.059^{***}	-0.056^{**}	-0.059^{***}	-0.056^{**}	-0.072^{***}	-0.066^{**}
	(0.015)	(0.015)	(0.015)	(0.015)	(0.017)	(0.018)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.002]
card5p>card3p	-0.125^{***}	-0.115^{***}	-0.125^{***}	-0.115^{***}	-0.140^{***}	-0.133^{***}
	(0.018)	(0.017)	(0.018)	(0.017)	(0.020)	(0.018)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Interactions						
ShowNewDebts × card3p>card5p	−0.029	−0.019	−0.029	−0.019	−0.007	−0.009
	(0.022)	(0.022)	(0.023)	(0.022)	(0.024)	(0.024)
	[0.202]	[0.393]	[0.202]	[0.393]	[0.765]	[0.696]
	[0.605]	[1.000]	[0.747]	[1.000]	[0.970]	[1.000]
ShowNewDebts × card5p>card3p	0.010	0.013	0.010	0.013	0.024	0.022
	(0.026)	(0.024)	(0.026)	(0.024)	(0.027)	(0.025)
	[0.699]	[0.591]	[0.699]	[0.591]	[0.368]	[0.385]
	[1.000]	[1.000]	[1.000]	[1.000]	[0.970]	[1.000]
ShowSavedMoney × card3p>card5p	0.008	0.007	0.008	0.007	0.031	0.021
	(0.020)	(0.019)	(0.020)	(0.019)	(0.021)	(0.022)
	[0.679]	[0.722]	[0.679]	[0.722]	[0.134]	[0.342]
	[1.000]	[1.000]	[1.000]	[1.000]	[0.535]	[1.000]
ShowSavedMoney × card5p>card3p	0.053	0.047	0.053	0.047	0.070^{*}	0.054
	(0.025)	(0.024)	(0.025)	(0.024)	(0.026)	(0.025)
	[0.034]	[0.050]	[0.034]	[0.050]	[0.007]	[0.030]
	[0.170]	[0.251]	[0.170]	[0.251]	[0.036]	[0.150]
Financial literacy			-0.040^{***}	-0.048^{***}	-0.053^{**}	-0.060^{**}
			(0.007)	(0.007)	(0.015)	(0.015)
			[0.000]	[0.000]	[0.000]	[0.000]
			[0.000]	[0.000]	[0.002]	[0.001]
Constant	0.239^{***}	0.255^{***}	0.157^{*}	0.198^{**}	0.213^{*}	0.264^{**}
	(0.021)	(0.021)	(0.063)	(0.073)	(0.088)	(0.090)
Observations	2619	3015	2619	3015	2277	2493
Subjects	291	335	291	335	253	277
Interact. Fin.lit._treatments	No	No	No	No	Yes	Yes
Further control variables	No	No	only YOE[b]	only YOE[b]	Yes	Yes
R^2 overall	0.056	0.040	0.123	0.131	0.151	0.199

Note: $^{*}p$<0.05; $^{**}p$<0.01; $^{***}p$<0.001

Financial literacy is centralized at a value of 3.

[a] This table shows the misallocation when the Cuckoo Fallacy is possible (baseline) vs. when it is not, but the 3%-card still has the higher debt balance (card3p>card5p) vs. when it is not and the 5% card has the higher debt balance (card5p>card3p). Additionally the table shows the treatments and the interaction with these scenario types. For each set of control variables there is a more strict out-screening for subjects (at least one rater screens them out) and - for robustness - a more tolerant one (both raters have to screen them out). The Akaike-optimal models (3) and (4) have the same control variables as in the main analysis. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for all the reported coefficients, but not the control variables. Asterisks indicate significance after adjustment.

[b] Years of education

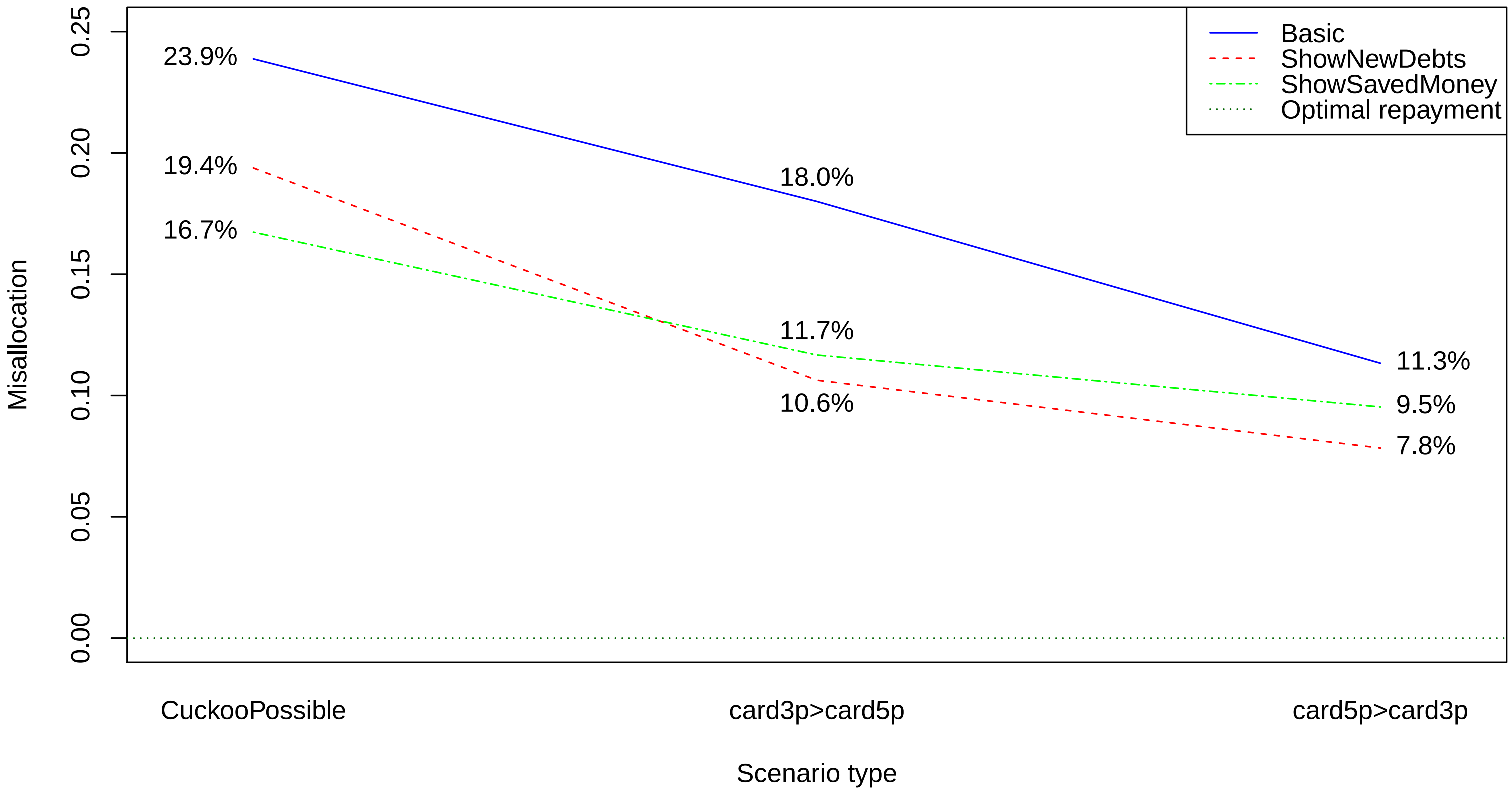


Figure II.2: This Figure shows the interaction plot between treatment and scenario type. That is we differ between scenarios whether the Cuckoo Fallacy was possible (CuckooPossible) or whether the 3% credit card produced more new debts than the 5% card, but not that much that the Cuckoo Fallacy was possible (card3p>card5p), or whether the 5% card produced more new debts than the 3% card (card5p>card3p). ShowNewDebts (in red) is the sludge and ShowSavedMoney (in green) is the nudge treatment.

II.7 Conclusion

In this paper, we conduct an experiment to study the behavior of the participants in credit card repayment. We find that only a small fraction of subjects repays optimally. Instead, a huge fraction of participants focuses on the card that produces more new debts, leading to a non-optimal split of repayments. Many of them repay the credit card first that produces a higher amount of debt in the next interest round - we call this deviation from optimality the Cuckoo Fallacy. The rounds in our experiment are interdependent to create a more realistic situation, but for the trade off of losing comparability between subjects. We tackle this problem by an additional experiment in which rounds are independent. The effects are weaker after adjusting, but we still find a considerable decline of misallocation in the ShowSavedMoney treatment. This shows that the Cuckoo Fallacy is a persistent effect beyond experimental design artifacts.

The experiment demonstrates that the Cuckoo Fallacy can be remedied by appropriate framing of information in the sense of a nudge. Highlighting and steering attention onto the information that shows how much money one can save in the next round with a given repayment allocation does increase optimal repayment behavior. However, increasing the salience of the new debt a credit card produces does not increase the misallocation, probably because an already large number of subjects seems to be worried about new debts in the basic treatment already.

Our work might be directly applied to design and test nudges to improve decisions in real-life situations. This could be particularly important to practitioners such as regulators or FinTechs striving to optimize their customers' financial decisions, probably with not too much room for backfiring or abuse. For instance, the way we present information in the ShowSavedMoney treatment could be used to avoid non-optimal behavior in credit card repayment decisions. Educating people about the existence of the fallacy and teaching them how to repay properly would be another huge step in that direction, which we follow up in chapter III. On the other side, an interesting question is to what degree financial institutions already try to sludge their credit card customers into worse repayment decisions today. Moreover, our experimental setting can be easily expanded or adapted to financial investment decisions.

A limiting factor of this study is the restriction to participants from the US, so in future works it would be interesting to find out, whether this results can also be obtained in countries where the use of credit cards is not part of everyday life. Furthermore, we limit our experiment to preferably homogeneous and comparable treatments in terms of the optimal behavior, meaning that there are no differences in the possibilities of the behavior one could show in the different treatments. That rules out changes in the number of credit cards and more realistic basic conditions like minimum payments, interest changes or overdrawing of an account. Future work should also look at different data sources than experimental data, preferably from a field experiment. This could help to improve our understanding of credit repayment, to educate debtors in an appropriate way and to find suitable regulatory rules for the credit market where necessary.

Chapter III

Addressing consumer misunderstanding in credit card debt repayment: Policy suggestions beyond the CARD Act

Coauthors:

Yannik Bofinger

Darwin Semmler

Relative share:

40%

A previous version of this chapter has been presented at:

- Society for Experimental Finance Conference 2021

Addressing consumer misunderstanding in credit card debt repayment: Policy suggestions beyond the CARD Act

Abstract

Recent studies find that people do not repay multiple credit cards in an interest minimizing way, which is usually interpreted as misallocation. We conduct an experiment on Amazon Mechanical Turk which tests four interventions to reduce this misallocation. We find that misallocation almost disappears when we provide participants with an assistant application which gives concrete repayment suggestions. Other interventions in the form of additional information, reminders and practice opportunities also help participants to reduce misallocation significantly, but not as strongly. Our results provide suggestions for policy makers on how to improve financial decisions in the context of debt repayment beyond the CARD Act.

Keywords: credit cards, financial decision-making, financial literacy, public policy, information disclosure

JEL-Codes: D14 - G41 - G51 - I18 - I22

Funding: This work was financially supported by the "Frankfurter Institut für Risikomanagement und Regulierung" (FIRM). FIRM had no involvement in anything study-related.

Declarations of interest: none

III.1 Introduction

When investigating financial decisions at the household level, economists repeatedly come across deviations from optimal behavior (Beshears et al., 2018; DellaVigna, 2009; Zinman, 2015). One recent instance for such a puzzle is that a large fraction of people does not repay debts on several credit cards with different interest rates in an interest minimizing way, a result which has been found both in empirical data and in experiments (Amar et al., 2011; Gathergood et al., 2019; Ponce et al., 2017). This is commonly interpreted as a failure in decision-making, because it is hard to believe that debtors would prefer to pay more than they need to without any apparent benefit. This misallocation adds to the rather common credit card debts, which amount to around \$800 billion in the U.S. alone (Federal Reserve Bank of New York, 2022). Large and further increasing amounts of credit card debts are paralleled by high finance costs for households, as the annual average percentage rate lies at 14.56%, according to the Federal Reserve Board (2022). These facts put further emphasis on the importance of tackling such misallocations and hence may be a goal for consumer financial policy.

In this paper, we test methods to decrease misallocation arising from non-optimal credit card repayment. This is particularly relevant as there exists ample evidence that despite the U.S. Credit Card Accountability Responsibility and Disclosure (CARD) Act of 2009 (H.R.627 — 111th Congress, 2009) which regulates financial firms with respect to their credit card offers and repayments, households are not able to optimally repay their debt. Even though Agarwal et al. (2014) and Jones et al. (2015) report a slightly positive effect of the CARD Act, Navarro-Martinez et al. (2011) and Salisbury (2014) also point to several negative consequences. Soll et al. (2013) even argue that additional policy interventions become necessary to improve the consumer understanding between debt reduction and monthly payments. These findings imply that despite the introduction of the CARD Act, certain inefficiencies in credit card debt repayment remain. Our study hence aims to fill this gap by designing and testing potential intervention scenarios which support consumers to understand the debt repayment process. To answer our research question to which extent policy intervention may help to avoid misallocation, we use an online experiment where we develop several financial interventions, and test on Amazon Mechanical Turk (MTurk) how successful they are in improving optimal repayment compared to a control treatment. In every treatment, we endow participants with debts on two credit card accounts and an income stream for ten rounds, which they can use to repay these debts. In the control group, we do not intervene with any help to solve this repayment problem. As we are particularly interested in the effect of certain types of intervention, we create four intervention treatments as proxy for possible policy interventions. These treatments are divided into two intervention groups: *General intervention treatments*, which are easy to implement and do not require any additional information about the credit, and *adapted intervention treatments* which are tailored for the credit situation, but require information the debtor voluntarily has to provide. We design two treatments for each of these groups. It is not our aim to distinguish effects within each group, but rather to suggest different practicable implementations for both approaches. The treatments in the general group are as follows: In the "pamphlet intervention", participants receive a three-page pamphlet to read, which explains the best strategy using text and graphics. In the "slider intervention", participants see a one-paged graphic including short explanations as well, but additionally they can practice repaying using an interactive slider which informs them how the debts change for a given repayment decision. Both the graphic and the slider are presented before the experimental stage starts. Although this slider is interactive, just as a pamphlet it is a mere information for the consumer and requires no information about the personal situation at all.

The two treatments for the adapted interventions work as follows: In the "reminder intervention", participants only see a short one-liner before the experiment which explains how they can repay optimally, but we also inform them that they receive a warning message whenever they deviate from the optimal strategy, which will be the main mechanic of this intervention. Once the participants finish a round with any misallocation, we inform them after this round about the misallocation, and again explain that they should use all their income to repay the more expensive card in future rounds. Finally, in the "assistant intervention" participants have a graphical tool - simulating the usage of an app - in every round that shows them which transfers the participants have to do in order to repay optimally. In case a participant misallocates some of the money, it also opens a popup with the information that the current transfers are not optimal and a calculation on how much money the participant can potentially save with the optimal allocation. Then the participants can revise or confirm their allocation. We announce the tool as an assistant to help finding the optimal repayment strategy before the trial rounds. Both of the latter two interventions vary in their presentation depending on whether a subject exhibits misallocation or not.

Our results show that in the control group without intervention, about 34.1% of the income is misallocated, while in each of the intervention groups misallocation is lower. The adapted interventions (6.6% misallocation) are stronger than the general interventions (11% misallocation). If we consider treatments alone, the strongest intervention is the assistant intervention, where misallocation drops to around 4%. In the other three interventions, roughly 10% of the income is misallocated. This implies that most of the difference between general and adapted interventions is driven by the assistant intervention. Financial literacy, measured as the sum of correct answers to six questions as introduced by Lusardi and Mitchell (2011) and Lusardi and Tufano (2015), has a strong negative effect on misallocation. Without intervention, financial literacy significantly improves the repayment decisions. Interactions between interventions and financial literacy show that the adapted interventions are strong enough to fully offset any advantages financial literacy brings. This implies that the provision of additional information and guidance on the optimal repayment process supports households to repay their credit card debt even though these households do not possess the relevant financial literacy.

Additional analyses reveal that participants tend to "unlearn" the optimal repayment strategy as misallocation increases over the experiment rounds. However, our interventions are able to reduce this increase and even completely offset the effect in the reminder intervention. This finding supports the conclusion that interventions need either to be permanent or to be renewed over time to increase optimal repayment decisions. Moreover, our results are robust to a battery of robustness checks. First, we include a measure of the credibility of our interventions to make sure that participants trust the provided information and guidance. Second, we repeat our analysis considering nonlinear data structure for our dependent variable. Finally, we also confirm the results when we only distinguish between people who employ optimal repayment every time and people repaying sub-optimal in at least one decision.

Our contribution to the literature is twofold. First, we add to the literature on misallocation in credit card repayments. Various studies have found significant deviations from optimal credit card repayment, both with field data (Gathergood et al., 2019; Ponce et al., 2017) and in experiments (Amar et al. (2011); Besharat et al. (2014); Ozyılmaz and Zhang (2020) as well as in chapters I and II). These studies document misallocation and try to explain it, using concepts such as heuristics, financial literacy, salience and framing. To the best of our knowledge, we are the first to design and test several general and adapted interventions to reduce misallocation with regard to the repayment decision of households. Moreover, we show that advantages from financial literacy are offset by the intervention, e.g. the additional provision of information and guidance to credit card customers.

Second, the literature on the effectiveness of the CARD Act is mixed. Jones et al. (2015) and Agarwal et al. (2014) report positive effects on household credit card repayments, whereas Navarro-Martinez et al. (2011), Salisbury (2014) as well as Hershfield and Roese (2015) point to remaining inefficiencies. Hence, we add to the literature by designing four potential interventions that are able to significantly improve repayment efficiency. Although the CARD Act already covers certain aspects of our interventions, e.g. credit card providers are obliged to distribute payments in excess of minimum payments to the highest interest credit card, our interventions nevertheless demonstrate the importance to provide households with information and guidance. This is due to the fact, that U.S. households hold on average 3.7 credit cards according to Foster et al. (2011) and also tend to use several credit card providers to benefit from extensive credit card rewards (Ching and Hayashi, 2010). Thus, even though the regulation prescribes the repayment process, households still face an individual decision with regards to different credit card providers. Furthermore, we give recommendations to policy makers for the implementation of these interventions as potential extensions to the current credit card regulation.

The remainder of this paper proceeds as follows: Section III.2 derives our hypotheses. Section III.3 outlines the experimental design and the data. Section III.4 illustrates our results and section III.5 provides additional analyses as well as robustness checks.

Section III.6 discusses our results under the consideration of the current U.S. credit card regulation and derives policy implications. Section III.7 concludes.

III.2 Hypotheses development

This study aims to shed light on the question on how misallocation in household debt repayment can be reduced. We tackle this question using a basic experimental design, which we modify to simulate different ways to intervene. This basic design resembles that of the other experimental studies on that subject (Amar et al., 2011; Ozyılmaz and Zhang, 2020) and the experiment in chapter II. In general, we endow participants with two credit card accounts, with debts on both. Participants also receive an income to repay these debts. After they finalized their repayment decision, the experiment continues with another round, where they start with the remaining debts, including the interests that are added between the rounds, and new income. This game is repeated for 10 rounds. In the control group, participants play this game without any intervention, while in the experimental intervention treatments ("interventions"), we use four different types of intervention.

Before we explain our interventions in detail, we can already set up the first hypothesis. We expect each intervention to lead to a reduction of misallocation, since even a weak intervention should still give a rough guideline, especially for those participants who do not have an idea how to place their repayments. Literature on the current credit card regulation (CARD Act) indeed shows that these interventions are able to improve the repayment decision (Agarwal et al., 2014; Jones et al., 2015). Furthermore, the identification of the problem alone might help participants to avoid mistakes. Thus, we formulate:

H1: Financial intervention lowers misallocation.

When policy makers or a financial advisor try to implement financial interventions for their customers, they have to think about how much data they can access from their customers in order to individualize the financial advice as much as possible. People might have different credits from different creditors and it may be either legally difficult to aggregate that data or inconvenient for a customer to give the precise details, provided that the customer is willing to share credit data at all. So it could be more practical to give general advice, especially because the advice in the problem we observe is universally valid: "Always repay the highest interest credit card completely before you touch any other credit". On the other hand, a general advice seems less suitable from the customer's point of view. The advice could be perceived as far too general to apply to individual cases. Tailored intervention might be more effective as it can be understood as a personalized nudge (Mills, 2022; Sunstein, 2012). While there is ample evidence that personalized nudges work in a variety of contexts (e.g. Bergman (2021); Castleman and Page (2015); Kraft and Rogers (2015); Page et al. (2020)), fewer studies investigated whether they outperform general interventions, but generally find that they do. Doss et al. (2019) show that a personalized texting-based program as educational intervention for kindergarten children is more effective than the analogous general programs. Other prior studies established the effectiveness of tailored advice in medical circumstances. Skinner et al. (1994) find that tailored mammography letters were read more carefully and patients were more likely to remember more information. Individually tailored advice also can lead to a behavioral change, as Kreuter and Strecher (1996) show with an increase in the effectiveness of risk health appraisals.

These considerations lead us to hypothesize that tailored advice might be more effective, but also not practicable in every situation. We want to do justice to both cases by a differentiation between two types of financial intervention: A general form of intervention that does not need to collect data of customers and is easier to employ (general intervention), and the individualized form of intervention that adapts to the customer's needs by collecting data (adapted intervention).

Before we construct the exact intervention treatments in each intervention group, we state the next hypothesis:

H2: Adapted interventions decrease misallocation stronger than general interventions.

For each of the two intervention styles we construct a group of two different treatments, four in total. Combined with the control group, we have 5 treatments in total. The two treatments per group are intended to experimentally test different types of practical implementations of financial interventions for banks, financial advisors and policy makers. Furthermore, this dual approach per intervention group stabilizes our results and weakens the influence of experimental artifacts. In addition to the comparison of the two intervention groups with the control group, we will also consider treatments individually in later analyses.

We first describe the two general interventions. The main idea for the first general intervention, the "pamphlet intervention", is to place information within a brochure that can be given to customers by a financial advisor or that is accessible in a bank. We develop such a brochure which we use as financial intervention (see Appendix III for the pamphlet we used). Participants are required to stay at least three minutes on the pamphlet's .pdf page, only then they can advance the experiment. A further incentive to use the three minutes for actually reading the pamphlet are three follow-up comprehension tasks. The tasks are announced on the pamphlet page and the pamphlet can also be downloaded again during the tasks. This procedure should strongly raise the probability that the participants deal with the pamphlet in depth.

In the second general intervention, the "slider intervention", we try to increase the learning effect of the pamphlet intervention by simplifying the detailed explanation in the pamphlet to a short text that just tells the participants to repay all available money to the highest interest rate credit card, and by showing the effect on the same graphic as at the end of the pamphlet (see Appendix III). This way, a participant can view the relevant information in a significantly less amount of time, which makes it easier to understand. To implement an adequate substitute of a reasoning leading to the correct solution, we provide the participants with an interactive repayment application. The participants can use a slider to view effects of a sample repayment for the next round as well as for the next five rounds by specifying a certain proportion of money they want to repay to the high-interest credit card. This way they can interactively experience the linear increase of the debts, the more the proportion shifts to the lower interest credit card, an idea based on Experiential Learning Theory (Kolb, 1984). Tang and Peter (2015) use this model to explain learning in financial contexts. A slider is also a visual representation of the repayment process. There is evidence that such visual tools can help in the context of financial decision-making (Killen et al., 2020; Lusardi et al., 2017). Finally, Kaufmann et al. (2013) find that using sliders in general can improve risk perceptions in financial decisions.

The financial intervention screens end with a comprehension task recapping the optimal repayment strategy. This intervention could be implemented within a website for financial advice or on an information screen of a banking app.

We continue with the two adapted interventions. The first adaptive intervention is the "reminder intervention". Reminders are a commonly used nudge, and a large literature shows that they can improve decision-making (for the seminal paper in the financial decision-making context see Karlan et al. (2016), for a recent paper see Medina (2021)). In our context, the reminder is a warning. If a participant finishes a round with any misallocation, no matter how small or large the fraction is, we show this warning in the after round screen. It reads as follows (assuming that the participant misallocated 250 out of 250 US-Dollar): "Warning: Your repayment was not optimal. You repaid only 0 out of 250 US-Dollar to the highest interest rate credit card. Therefore your efficiency was 0%. Try to repay all the available money to the highest interest rate credit card to minimize your overall debt." This text is highlighted with red color and in bold.

The second adaptive intervention is the "assistant intervention", which can be understood as a proof of concept for a "choice engine" (Thaler and Tucker, 2013). If we assume to have full access to a customer's credit data, financial intervention could be implemented via a system that can suggest to the customer how to repay in every situation in order to optimize debt payments. This could be done with a cell phone application. Nowadays FinTechs already offer multi-banking apps in which the management of accounts at other banks and credit institutions is possible. Third-party providers could also use digital interfaces of the banks or manual input of the customer to collect complete credit information and therefore be an advisor for individual situations. In our experimental setting, this means that we provide participants with an assistant interface during the decision situations, which gives information and guidance on what they need to do to repay their debts optimally.

The chapters I and II show that financial literacy is particularly relevant for credit card repayment decisions. The more financially educated a participant is, the smaller the misallocation in the experiments. Lusardi and Mitchell (2014) define financial literacy as "(...) peoples' ability to process economic information and make informed decisions about financial planning, wealth accumulation, debt, and pensions", and review a huge literature that shows connections between financial literacy and myriads of other variables. Hence, financial literacy can help participants to find the optimal allocation for themselves. However, financial interventions help less financial literate people to find a better way of allocating money, while more financially educated people already tend to know how to repay debts correctly. In other words, we think financial literacy and our different interventions are substitutes, which is why we argue that in our interventions the effect of financial literacy should be weaker. Technically speaking, we expect an interaction between each treatment and the measure of financial literacy. To be more specific, financial literacy should have the strongest impact in the control treatment, and a smaller one in all treatments.

H3: The effect of financial literacy on misallocation is lower in the intervention treatments compared to the control group.

III.3 Experimental design and data

We create our experiment using the Software Platform for Human Interaction Experiments (SoPHIE, Hendriks (2012)) and conduct it on Amazon's crowd-sourcing platform Mechanical Turk (MTurk) with U.S. residents[1]. The experimental design follows the experiment in chapter II. Participants have two credit card accounts, which both start with

[1]For a brief discussion of MTurk, see the section "Gathering Data on Amazon Mechanical Turk" in Appendix III.

$2200 of debts and charge 3% and 5% interest per round, respectively. In addition, participants have a checking account with an income of $250 per round, which participants must use to repay debts in that round.[2] The participants can freely choose how they want to allocate their income between both cards, but cannot leave any money on the checking account. The game repeats for ten dependent rounds. In each round, the remaining debt is carried over, respective interest payments are added, and the checking account gains additional $250. Participants are incentivized to minimize their overall debt, and have two trial rounds to familiarize themselves with the mechanics of the experiment. The participants receive $1 participation fee when finishing the experiment and up to $2 bonus payment depending on their overall debt in the end[3], ensuring that the main incentive was to minimize the overall debts.

The experiment starts with an explanation of the payment and the instructions. We ensure the understanding of the experiment with comprehension tasks and screen out every participant who does not pass our two attention tests[4]. To ensure basic numeracy, we ask participants to calculate the balance after one year if they had $1000 and earned 1% interest per year. The participants have the chance to test the mechanics of the experiment in two trial rounds. Then there is a brief digression with a financial intervention depending on the intervention (and none in the control group), and then 10 main experiment

[2]As people tend to repay any of the two accounts completely (Amar et al., 2011), we design the experiment in a way that such behavior is not possible which supports to focus on our research question. Given the interest rates and the income, starting with $2200 guarantees this.

[3]For *max*, *min* and *debt* as the maximal, the minimal and the actually achieved amount of debt in the experiment, the bonus calculates by $\$2 \cdot \frac{max-debt}{max-min}$. We explained this to the participants by the instruction that a smaller amount of total debt in the end leads to a higher bonus. We support our explanation by providing some examples.

[4]The first attention test is a mock question during the comprehension tasks which we formulate as if it was a question about credit card issuers, but then reveal that participants have to check the answer "other" and type a "h" in a free text field. We screen out all participants who did not do that. The second attention test is during the post-experimental questionnaire. The participants have to choose the 'second answer' from a choice of two answers to pass this test.

rounds. As described in section III.2, we perform four financial interventions. After the main rounds, the experiment concludes with a post-experimental questionnaire.

Furthermore, we ask in an open, text-based question if our participants were convinced by the proposed repayment strategy, as an additional safeguard against too inattentive participants and bots. Two raters independently analyzed the answers to check whether they are meaningful with respect to this question[5]. We use this screening process for an additional robustness check where we carefully exclude all the suspicious answers that go beyond the standard bot screening as described above. Only after that screen we present the results and the final payment to the participants. The questionnaire continues with questions to the number of credit cards (variable *# creditcards*)[6], the number of additional accessible credit cards (for example via friends, spouse, etc., variable *# credit access*) and two binary questions if credit cards are used at work (variable *Credit Card Usage at Work* and if the participant usually does not use credit cards, but generally knows how they work (variable *Unused but Knowledge*). Finally, we determine the financial literacy (variable *Fin. Literacy*) of the participants on a scale from 0 to 6 using the number of correct answers to six questions introduced by Lusardi and Mitchell (2011) and Lusardi and Tufano (2015), and finish the experiment with demographic questions.

The main dependent variable for our experiment is the average proportion of money a participant repays to the low-interest rate credit card in the ten experiment rounds. We call this measure the "misallocation". For each participant *i* the misallocation is a value between 0 and 1. It becomes 0 when participant *i* transfers all available money to the

[5]Since the control group has no proposed strategy, answers which indicate that participants were confused about the question itself are accepted as valid.

[6]One participant claims to have 51 credit cards. We therefore winsorize this variable at the one percent level to limit the influence of outliers in a robustness check. Our results remain qualitatively the same and are available from the authors upon request.

high-interest rate credit card, which implies optimal repayment behavior and thus no misallocation at all. Consequently, it becomes 1 if participant i transfers all available money to the low-interest rate credit card.

Table III.1 provides descriptive statistics of the experiment. The total number of participants is 660, the individual treatments consist of 125 to 139 participants. Approximately half of our participants are male and half female. The mean age is about 38 and the participants indicate 15.72 years of educations on average. More than half of them answered between 3 and 5 financial literacy questions correctly, on average 3.66 questions. Participants hold 2.47 credit cards on average, and in addition have further access to another 0.75 credit cards. This is roughly in line with findings from Foster et al. (2011), who argues that U.S. Americans hold 3.7 credit cards on average. Furthermore, the participants receive a total payoff of \$2.76 on average. Since the mean duration of a session was 18 minutes and 43 seconds[7], we paid an average hourly wage of \$8.85, which is far above the average payment for tasks on MTurk. Hara et al. (2017) quantifies the average hourly wage on MTurk slightly above \$3, Berg (2016) estimates approximately \$5.50. Thus, we are confident that the stake size is a large enough incentive for participants to minimize their debt. Finally, 235 of our participants answer the question if they do not use credit cards in general, but know how they work with yes, and about half of them uses credit cards at work.[8]

III.4 Results

We start our analysis by measuring the distribution of misallocation of the participants, both on the intervention group and on the treatment level. We find that we are able to

[7]Table Appendix III.29 shows summary statistics on the duration of all five treatments.

[8]For a description of summary statistics for each treatment separately, see Table Appendix III.30.

Table III.1: Descriptive statistics of participants

Statistic (continuous vars)	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Fin. Literacy	660	3.66	1.42	0	3	4	5	6
Age	660	38.03	12.55	18	29	34.5	46	82
Years of education	660	15.72	2.38	9	14	16	17	21
# credit cards	610	2.47	3.50	0	1	2	3	51
# credit access	604	0.75	1.33	0	0	0	1	12
Exp. duration (min:sec)	660	18:43	9:11	4:35	12:33	16:35	22:30	61:14
Total payoff	660	2.76	0.37	1.00	2.65	3.00	3.00	3.00
Statistic (count vars)	N	Type	Number		Type	Number		
Gender	660	Female:	304		Male:	356		
Unused but Knowledge	659	Yes:	235		No:	424		
Credit Card Usage at Work	659	Yes:	339		No:	320		

reproduce the behavior of misallocating money to the low-interest rate card as observed in chapter II in the control group. This is of major importance for our analyses, as we use the control group as reference in our regressions. However, as expected several differences between our interventions exist.

Table III.2 illustrates the summary statistics of misallocation split by treatment and intervention group, and additionally Figure Appendix III.8 represents a graphical representation via boxplots. The table shows that in the control treatment we measure an average misallocation of 34.1%. This value is distinctly higher than in the two intervention groups. With 6.6% misallocation on average, the adapted interventions are lower than the general interventions (10.9%). Considering the four interventions separately, the misallocation varies between 4.4% (assistant intervention) and 11.3% (pamphlet intervention). The slider intervention and the reminder intervention reveal misallocations of 10.5% and 8.8%, respectively. In all but the control treatment, at least one quarter of the participants ended the experiment without any misallocation, and in three of the financial interventions more than half of the participants do not show any misallocation at all. These results already indicate that financial interventions can lead to an improvement of misallocation and hence support households to efficiently repay their debt.

As a next step, we use the treatments and intervention groups as independent variables in a linear regression, with misallocation as dependent variable, and add financial literacy, as well as additional control variables[9]. For hypothesis H1 we compare the control group to all four intervention treatments. For hypothesis H2 we compare the two intervention groups with each other. For hypothesis H3 we separately let both the four treatments as well as the two treatment groups interact with financial literacy, to

[9]We perform these regressions without control variables as robustness check. The results remain qualitatively unchanged and are illustrated in Table Appendix III.31.

Table III.2: Descriptive statistics of misallocation per treatment[a]

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
All data	660	0.138	0.194	0.000	0.000	0.000	0.271	1.000
Control	132	0.341	0.221	0.000	0.234	0.345	0.426	1.000
General	262	0.109	0.173	0.000	0.000	0.000	0.200	1.000
Adapted	266	0.066	0.117	0.000	0.000	0.000	0.100	0.525
Pamphlet	125	0.113	0.195	0.000	0.000	0.000	0.140	1.000
Slider	137	0.105	0.150	0.000	0.000	0.000	0.210	0.684
Reminder	133	0.088	0.128	0.000	0.000	0.020	0.110	0.520
Assistant	133	0.044	0.100	0.000	0.000	0.000	0.000	0.525

[a] This table provides summary statistics of the misallocation values aggregated over all data as well as for the control group and interventions (pamphlet, slider, reminder, assistant and intervention groups). Detailed descriptions of the interventions can be found in section III.2. Misallocation measures the percentage of money, that was allocated to the low-interest rate credit card.

take into account that financial literacy might help to reduce misallocation to different extents in different treatments or groups. The results are shown in III.3.

The first column illustrates the differences in misallocation between the single treatments with the control group as reference. Each intervention shows a significantly lower average misallocation than the control group. The effects are economically large and vary between a decrease in misallocation of at least 20.9% (slider intervention) and maximum 27.1% (assistant intervention). In comparison to the control group, providing a pamphlet to the participants reduces misallocation by 22.2 %, while providing participants with reminders in case of inefficient repayment reduces misallocation by 22.3%. Consider the following numerical and hypothetical example that demonstrates the effect as well as the economic significance: a bank customer has debt of $1000 on each of two accounts with 5% and 3% interest rates, and repays $200 each month. Without any intervention, as in the control group, this implies a split of $132 on the 5% and $68 on the 3%-card (which corresponds to our mean misallocation of about 34% in the control treatment). With that split, the customer needs a total of \$2436.70 to repay their debt. In the assistant intervention, a reduction of misallocation by 27% lowers the amount of money needed to repay the debt to \$2395.99. Thus, our assistant leads to savings of \$40.71 in this example case. This is particularly important, as credit card debt in the U.S. tends to increase yearly and already amounts to \$800 billion in 2022 (Federal Reserve Bank of New York, 2022). These findings confirm our hypothesis (H1), as misallocation in every intervention is below the control group, which implies that interventions improve household decisions.

Columns (2) to (4) show the regression of the misallocation which compares the control group to the intervention groups (H2). All three columns show the same regression analysis, but with different reference groups in order to compare differences between the general and the adapted interventions. Since we have already shown that the single treatments reduce misallocation significantly, it is no surprise that this also applies for the grouped treatments. Column (2) shows that general interventions decrease the misallocation by 21.5 percentage points, while the adapted interventions decrease misallocation by 24.7 percentage points. The columns (3) and (4) show that this difference of 3.2 percentage points between these two groups is significant. This confirms hypothesis (H2). However, this reduction seems to be mostly driven by the assistant treatment, because column (1) shows that the reminder treatment reduces misallocation at about the same level as the general intervention treatments, while the assistant treatment shows a clearly stronger reduction.

As we expect financial literacy to have an impact on misallocation, we explicitly consider these effects in the following. As can be seen from Table III.3 in column (1), financial literacy decreases the mean misallocation in the control group by 6.5% per correctly answered question. This implies that financial literacy is indeed a relevant driver of misallocation in credit card debt repayment. Households with higher financial literacy are hence able to repay their debt more efficiently. To investigate (H3), i.e. whether the effects of financial literacy are weaker in the interventions compared to the control group, we need to consider column (1) and the interactions between financial literacy and the interventions. These interactions are significant in three out of the four interventions. The positive coefficients of the interaction terms indicate that the negative base effect of financial literacy becomes less negative in the interventions. Considering the reminder and the assistant intervention alone, these adapted interventions fully compensate any advantage of financial literacy. This effect is also shown in column (4) as the effect of financial literacy does not deviate from 0 significantly in the adapted interventions. However, (H3) does not apply for the pamphlet intervention, as its interaction term with financial literacy is not significant. This might occur because understanding the pamphlet might require some baseline financial literacy, and the pamphlet might not help as strongly for anyone below this threshold, while the other treatments work equally well for all levels of financial literacy. To test this claim, we calculate the coefficient for the pamphlet intervention again using only participants with a financial literacy score below the median value 4. Now the pamphlet intervention does not deviate in misallocation from the control group. The coefficient (<0.001) deviates significantly from the pamphlet coefficient (-0.222) from Table III.3 (p-value = 0.036 using the test in equation (4) of Paternoster et al. (1998) with reference to Clogg et al. (1995)). This is evidence that the pamphlet intervention is only helpful for financially more literate participants.

Therefore, the intervention might not work well enough, and a high value in financial literacy is still an advantage when solving the repayment problem. As a consequence, we can confirm (H3) for each intervention expect for the pamphlet treatment.

Table III.3: OLS regression of the misallocation with different reference categories[a]

	(1) Control	(2) Control	(3) General interventions	(4) Adapted interventions
		Dependent variable: Misallocation		
Treatment (group)				
Control	Reference	Reference	0.215***	0.247***
	(.)	(.)	(0.019)	(0.018)
	[.]	[.]	[0.000]	[0.000]
	[.]	[.]	[0.000]	[0.000]
Pamphlet	-0.222***			
	(0.021)			
	[0.000]			
	[0.000]			
Slider	-0.209***			
	(0.020)			
	[0.000]			
	[0.000]			
Reminder	-0.223***			
	(0.020)			
	[0.000]			
	[0.000]			
Assistant	-0.271***			
	(0.019)			
	[0.000]			
	[0.000]			
General		-0.215***	Reference	0.032**
		(0.019)	(.)	(0.011)
		[0.000]	[.]	[0.005]
		[0.000]	[.]	[0.009]
Adapted		-0.247***	-0.032*	Reference
		(0.018)	(0.011)	(.)
		[0.000]	[0.005]	[.]
		[0.000]	[0.013]	[.]
Financial literacy	-0.065***	-0.064***	-0.028***	-0.001
	(0.014)	(0.014)	(0.007)	(0.006)
	[0.000]	[0.000]	[0.000]	[0.853]
	[0.000]	[0.000]	[0.000]	[0.853]
Interactions between financial literacy and treatment (group)				
Control × FL	Reference	Reference	-0.037*	-0.063***
	(.)	(.)	(0.015)	(0.015)
	[.]	[.]	[0.018]	[0.000]
	[.]	[.]	[0.018]	[0.000]
Pamphlet × FL	0.027			
	(0.017)			
	[0.125]			
	[0.125]			
Slider × FL	0.044*			
	(0.017)			
	[0.010]			
	[0.020]			
Reminder × FL	0.058**			
	(0.017)			
	[0.001]			
	[0.002]			
Assistant × FL	0.068***			
	(0.015)			
	[0.000]			
	[0.000]			

continued on next page ...

... continued from previous page

	(1) Control	(2) Control	(3) General interventions	(4) Adapted interventions
		Dependent variable: Misallocation		
Interactions between financial literacy and treatment (group)				
General × FL		0.037*	Reference	-0.027**
		(0.015)	(.)	(0.009)
		[0.018]	[.]	[0.004]
		[0.018]	[.]	[0.009]
Adapted × FL		0.063***	0.027*	Reference
		(0.015)	(0.009)	(.)
		[0.000]	[0.004]	[.]
		[0.000]	[0.013]	[.]
Further control variables				
Age	-0.001*	-0.001*	-0.001*	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)
Years of education	0.000	0.000	0.000	0.000
	(0.003)	(0.003)	(0.003)	(0.003)
Male	0.005	0.003	0.003	0.003
	(0.013)	(0.012)	(0.012)	(0.012)
# credit cards	0.005	0.004	0.004	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
# credit access	0.002	0.003	0.003	0.003
	(0.005)	(0.005)	(0.005)	(0.005)
Cord	-0.002	0.000	0.000	0.000
	(0.012)	(0.012)	(0.012)	(0.012)
Unused but Knowledge	0.016	0.016	0.016	0.016
	(0.015)	(0.016)	(0.016)	(0.016)
Credit Card Usage at Work	0.027	0.025	0.025	0.025
	(0.014)	(0.014)	(0.014)	(0.014)
Constant	0.308***	0.310***	0.095	0.064
	(0.055)	(0.055)	(0.053)	(0.052)
Observations			595	
R^2	0.437	0.427	0.427	0.427
F-value	17.624	20.096	20.096	20.096

Note: *p<0.05; **p<0.01; ***p<0.001

[a] This table presents OLS regression results of the mean misallocation of participants. Misallocation serves as dependent variable in all regressions. The first column shows the control group in comparison to all intervention treatments. The other three columns show the same regression with either control, general or adapted intervention as base group. Further control variables are age, years of education, gender (reference: female), number of own credit cards, additional accessible credit cards, order of credit card presentation in the experiment (Cord=1 if 5%-card was second), and credit card dummies whether credit cards are generally not used, but known in principle (Unused but Knowledge) and whether they are used at work (Credit Card Usage at Work). Financial literacy (FL) is centralized at the median value 4. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values in model (1) are adjusted for Pamphlet, Slider, Reminder, Assistant, Pamphlet × FL, Slider × FL, Reminder × FL and Assistant × FL. The p-values in models (2-4) are adjusted for Control, General, Adapted, Control × FL, General × FL and Adapted × FL, depending on which four of the six variables are reported. Asterisks indicate significance after adjustment.

III.5 Additional analyses and robustness checks

Additional Analyses

Prior analyses in this study reveal an overall reduction of misallocation for our four interventions as compared to the control group. We now analyze the development of misallocation over the ten experiment rounds. Figure III.1 shows a graphical representation in order to investigate if misallocation is stable, or whether there is a - possibly treatment-dependent - fluctuation between rounds, which would suggest a dependency on account balances and previous decisions. Therefore, the x-axis delineates the round number and the y-axis illustrates the mean value of misallocation. As can be seen from Figure III.1, all interventions exhibit a small increasing tendency of misallocation over the ten experiment rounds. This effect is particularly strong in the control group.

Table III.4 illustrates OLS regressions showing the average increase of the misallocation per round for all interventions. The average increase per experiment round is significant at 2.7% in the control group and declines to 0.8% in the pamphlet intervention and 1.5% in the slider intervention. With regards to the reminder and assistant intervention, the average increase per round reduces to 0.4%, although the effect is not significant in the reminder intervention. Thus, in all but the reminder intervention the misallocation significantly increases on average per round. The reminder hence seems to significantly support the decision-making of households towards the efficient solution. People seem to "unlearn" the optimal repayment behavior, maybe because if one repays money to the highest interest rate card, the debt on the other card increases every round and therefore appears more urgent after a few rounds. This is in line with chapter II, where we observe a similar increase and refer to the non-optimal assumption of increasing urgency of the lower interest card as "Cuckoo Fallacy". However,

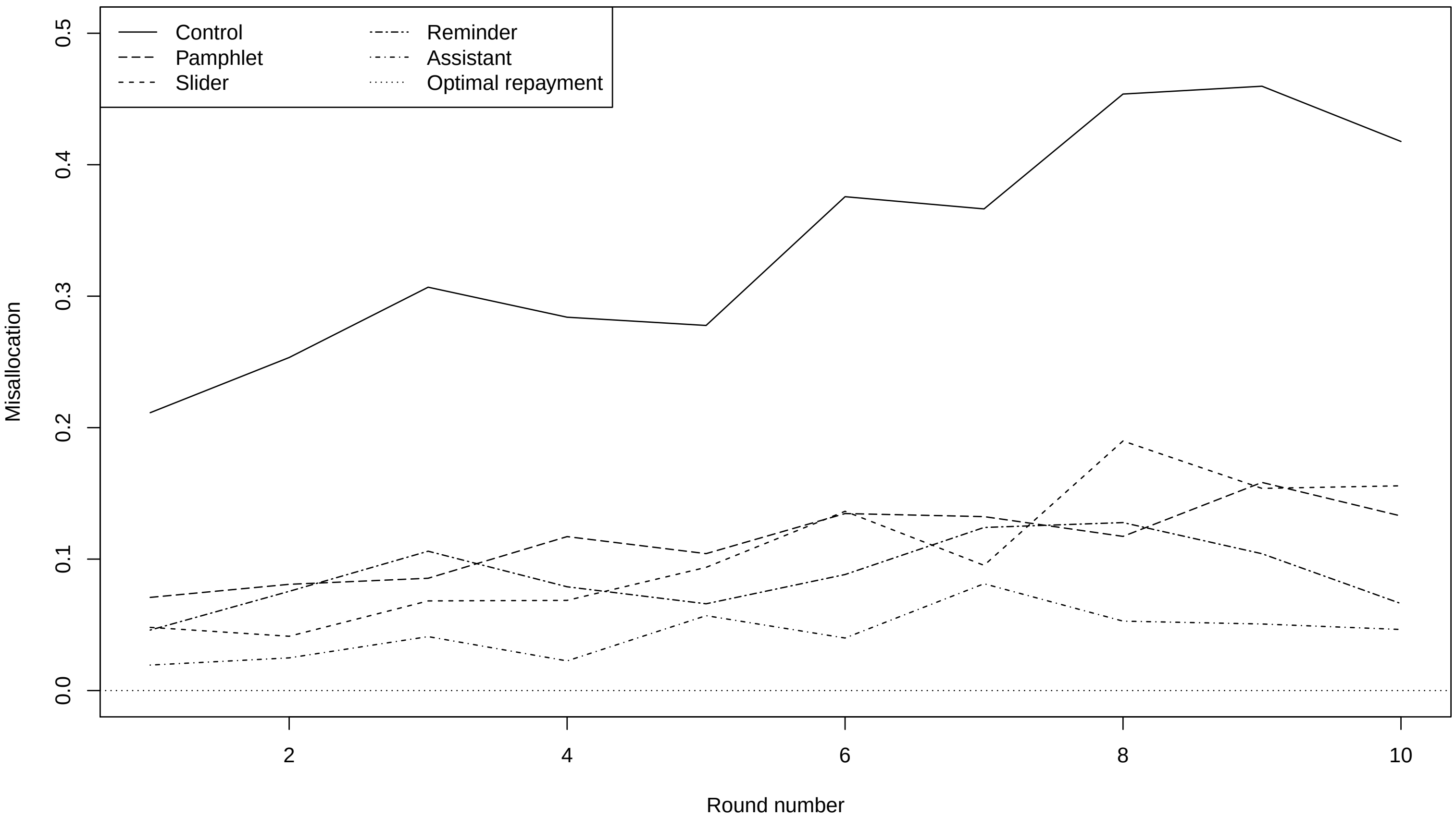


Figure III.1: This figure shows the development of the average misallocation over ten experiment rounds split by treatment. The round numbers are on the x-axis, the mean misallocation on the y-axis. The different values for each treatment are shown as different line types. The classification of the line types can be found in the legend.

the interventions seem to clarify this issue slightly, such that the effect of increasing misallocation per round is much weaker in our interventions compared to the control group.

Table III.4: Average increase of misallocation per experiment round split by treatment[a]

Treatment	Control	Pamphlet	Slider	Reminder	Assistant
Increase of misallocation per round	0.027***	0.008**	0.015***	0.004	0.004*
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Constant	0.195***	0.069***	0.022*	0.067***	0.022**
	(0.020)	(0.015)	(0.011)	(0.013)	(0.008)
Observations	1320	1250	1370	1330	1330
R^2	0.043	0.007	0.028	0.002	0.004

Note: *p<0.05; **p<0.01; ***p<0.001

[a] This table provides OLS regressions of misallocation per experiment round and by intervention. Misallocation serves as dependent variable in all regressions. The table shows the slope of an OLS regression which corresponds to the average increase of mean misallocation per round. Standard errors are robust and reported in parentheses.

Robustness Checks

To underline the robustness of our findings, we perform various robustness checks. First, we drop the 13 participants which gave a suspicious or non-fitting answer to the open anti-bot question after the experimental stage, evaluated by two independent raters. The results remain qualitatively unchanged, except for the interaction between financial literacy and the slider intervention (with the control group as base), which becomes insignificant. Thus we cannot establish that the slider intervention reduces the effect of financial literacy compared to the control group. Table Appendix III.32 illustrates the results.

Second, we use fractional regressions as illustrated in Table Appendix III.33 to account for the fact that the misallocation variable is a number between zero and one. This check is essentially important, considering the strong decrease of misallocation in the financial interventions. There are many participants with an overall misallocation at or near the minimal value zero, therefore the assumption of normally distributed residuals in an OLS regression might not be valid. Table Appendix III.33 confirms that all interventions significantly reduce misallocation overall and hence underlines the robustness of all our base effects. The consideration of the interactions between treatments and financial literacy reveals that only the coefficient for the assistant intervention remains significant. Thus, the established effect of financial literacy decreases only in the assistant intervention robustly.

Third, we use an alternative measure for optimal repayment instead of misallocation: In an attempt to more conservatively evaluate the effects of a financial intervention, we only consider an intervention as successful if it was able to completely nullify misallocation. Only then, the participants fully followed the strategy suggested by the financial intervention. So we replace misallocation with the dummy variable *Optimal repayment*, which takes the value 1 if and only if a participant has zero misallocation, and 0 otherwise[10]. We apply a logistic regression model as illustrated in Table Appendix III.34. In general, we can draw the same conclusions as in the main analysis: Financial interventions that reduce misallocation also significantly increase the probability for a participant to repay optimally. However, we now additionally see a difference between the pamphlet and the reminder intervention; in the pamphlet intervention the chance of repaying optimally is significantly higher than in the reminder intervention (p-value =

[10]Note that for this dependent variable, positive regression coefficients imply that more participants repay optimally.

0.008), which is surprising since the pamphlet intervention is weaker in terms of reducing misallocation. This can be explained with a higher variance of misallocation in the pamphlet intervention (sd = 0.195) than in the reminder intervention (sd = 0.128, F-test: p-value for ratio = 1: $< 2.8 \cdot 10^{-6}$). Thus, there are more participants with optimal repayments, but also more participants with higher misallocation in the pamphlet intervention. This increased variance could be caused by varying degrees of attention paid to the pamphlet. Some participants may have read it to the end carefully, others may have not. This is not a weakness in the experiment design, but rather shows a problem of the "pamphlet" approach itself. People have to actively deal with the content of the pamphlet in order to learn lessons for their actions. As in our main analysis, the effects of financial literacy lose significance in the adapted interventions. However, we do not measure any difference in the control group compared to the interventions regarding financial literacy anymore. In other words, while our interventions help to replace financial literacy as a reducing factor with respect to misallocation, financial literacy remains important for the understanding of the optimal repayment.

At last and in order to ensure the reliability of our results, participants have to trust and believe in our interventions and experiment design and quality. Hence, a difference in the quality level of our interventions can be one potential confounder of the experiment results. If so, it is not the type of intervention, but our concrete implementation that leads to different values of misallocation. If, say, our pamphlet is too poorly written, participants may not perceive it as convincing and thus ignore its advice, even though a better pamphlet would be useful. To tackle this issue, we employ the variable *Credibility*. For that variable, we ask participants to rate the statement "I was convinced that the strategy proposed by [the intervention] would give me the highest bonus" on a Likert scale from -2 ("I totally disagree") to 2 ("I totally agree")[11]. This enables us to

[11] We cannot ask that question in the control treatment since there is no intervention, so we only include

check for differences in the perceived intervention quality, using a regression where we add the credibility variable (*Credibility*) and interact it with the interventions. In case a confounding effect exists, these interaction terms would be significant. Table Appendix III.35 in illustrates the results if we include credibility as well as its interactions with the single interventions. Higher credibility of the experiment participants indeed goes along with a significant reduction in misallocation. This implies that participants who believe the intervention to be credible achieve better repayments. Second, all interactions between credibility and the interventions are insignificant. Thus, we do not find evidence that the influence of the credibility systematically varies between different interventions. This allows us to conclude that the experiment credibility is given among our participants and equal among interventions, thus it does not confound the results.

III.6 Policy implication and discussion

Prior evidence shows that households repay their credit card debt sub-optimally (Gathergood et al., 2019; Amar et al., 2011). This behavior facilitates in a repayment choice that is not interest minimizing and hence comes at additional interest costs for households. Moreover, our findings indicate that the provision of additional information and guidance can improve repayment decisions. [12]

In order to address inefficient repayment decisions, the United States initiated the CARD Act of 2009. With a variety of disclosure obligations and restrictions, the CARD Act seeks to protect consumers and improve transparency. More specifically, the imple-

it in the interventions.

[12]We want to emphasize that our policy implications have to be considered against the background of proportionality and appropriateness with regards to public interventions. This paper intends to provide scenarios that might improve household welfare and stability in case of public interventions. Hence, it is far beyond the scope of our paper to discuss the adequacy of potential regulatory measures.

mentation restricts and defines caps on various fees or increases in fees and requires lenders to provide credit card users with early notice of risen charges or other changes in terms and conditions and of their right to cancel in such an event (Agarwal et al., 2014). Moreover, the CARD Act mandates the publication of information on the credit card billing statement, such as penalty interest rates, payment due dates, late fees and payoff times. The details regarding the payoff times are divided into two scenarios, one when only minimum payments are made and another one in which the debts are settled within three years. Both scenarios show the respective monthly payments and durations required to pay off the debts and the respective accrued total interest charges (Jones et al., 2015). This information is presented as a minimum payment warning to all credit card users not to pay only the required minimum monthly payments, and aims to enhance cardholders' repayment behavior (Navarro-Martinez et al., 2011). Additionally, the regulator tries to tackle inefficient repayment structures, as credit card issuers have to allocate amounts in excess of the minimum payment to the highest interest rate card. These regulatory requirements, however, only apply if a customer holds several cards from the same issuer.

Jones et al. (2015) observe the repayment history of credit card debts before and after the CARD Act. They find a significant impact of the additional disclosures on how participants repay their credit card bills. Credit card users who paid attention to this new information tend to repay higher amounts of debt monthly after the modification, especially the probability of a full settlement increases (Jones et al., 2015). Agarwal et al. (2014) find similar results. They note a slight but significant influence of the CARD Act on consumer's behavior to repay their credit card debt. Furthermore, they determine a decline of credit costs due to the implementation, while other costs and the total lending remain stable (Agarwal et al., 2014).

On the other hand, Navarro-Martinez et al. (2011) show negative effects of disclosures such as the minimum payment information on repayment behavior. Even additional information as required by the CARD Act does not significantly change the outcome. Rather they note that cardholders tend to reduce their repayments based on details about future interest costs. Furthermore, similar investigations reveal stronger reactions of credit card holders when alternative repayment structures were given, such as the three-years plan. The tested people tend to orientate themselves by the alternative to the minimum payment, whereby some people increased their payments and others who were willing to pay back higher amounts first reduced their payments after the new information (Salisbury, 2014). This result is confirmed by Hershfield and Roese (2015). Besides their finding of declining repayments in cases where people would have been willing to make higher payments than the three-year amount, they show evidence that credit card users are less inclined to repay their debts in full when a second payoff scenario is presented. As a solution, they specify a range between 0 and the full settlement in addition to the dual payoff scenario, with the indication that any amount within the range can be paid. Additionally, they show participants the amount of their total balance directly before the payment. Both interventions weaken the previously mentioned effect and prevent that information such as the minimum payment or the three-year payment amount serves as an anchor (Hershfield and Roese, 2015).

Although these regulations aim to protect consumers, the current literature allows to draw the conclusion that households still lack financial knowledge, inducing them to fall for certain repayment fallacies. This is in line with surveys that examine financial knowledge in the context of debt in general (Lusardi and Mitchell, 2011). Moore (2003) even shows in a Washington-State residents survey that people face issues understanding interest compounding as well as terms and conditions of loans. Similar to Lusardi and

Mitchell (2011) and Moore (2003), an analysis of Soll et al. (2013) also shows massive mathematical comprehension problems of the link between credit card debt and monthly repayments. Soll et al. (2013) determine that especially people with a lower numeracy miscalculate this situation. Even though the introduction of the CARD Act mitigates this problem, they point out that the mandatory information on payoff times is still misunderstood by many users. Based on their findings, they even further recommend policy interventions that help to improve credit card holders understanding between payments and debt elimination. As the previous literature states that inefficiencies might remain despite the introduction of the CARD Act, our study designs and analyses four interventions that might further help to reduce non interest optimizing repayment of debts by providing further background information on the functioning of debt repayment for customers. We even distinguish between general and adapted interventions to illustrate the differences in the effects on debt repayment.

Recent EU-wide financial regulations (e.g. the Markets in Financial Instruments Directive II (MIFID II)) enhance investor protection, in particular with regard to investment vehicles, by forcing financial institutions to provide further information to customers. As a consequence, higher transparency through the provision of information supports the functioning of capital markets and hence protects investors. In a similar vein and as far as debt instruments (e.g. loans, mortgages, credit card debt) are concerned, we provide evidence that providing additional information can - besides the credit card regulation that is already in place - improve the debt repayment behavior of households. This is particularly relevant, as it shows that financial regulators can improve households' debt repayments decisions and hence the financial position and stability, respectively. Moreover, we can even trace out the economic significance of our four interventions and hence aim to provide guidance to financial regulators. In the following, we discuss our four interventions by contrasting them with the current regulation and by providing opportunities for improvements to the regulator under the consideration of the economic significance of potential interventions.

First, we give participants additional information on optimal credit card repayment by providing the pamphlet (sheet including relevant information to the repayment process, see Appendix III). This is similar to the payoff time disclosures by the CARD Act shown in two scenarios: the payoff time when minimum payments are made compared to higher repayments to settle within three years, and which are intended to encourage the consumers to make higher repayments. However, the difference is that information required by the CARD Act is only provided on the credit card statement and therefore after the transaction. The advantage of our intervention is the provision of information before the participants decide on how to allocate the available sum. Thus, participants have the opportunity to take the given information directly into account in their decision. Furthermore, the pamphlet intervention provides an explanation, whereas in contrast the CARD Act disclosures only shows facts, which may not be comprehensible for the participant. Moreover, both interventions aim to improve repayment behavior, however the CARD Act targets the absolute amount of repayment, while our intervention refers to its distribution. Furthermore, the CARD Act requires any repayment which exceeds the minimum payment to be allocated by the lender to the credit card with the highest interest rate first, and the credit card with the lower interest rates in each case will not be serviced until the first card has been paid in full. This is almost equal to our request to the participants to pay the card with the highest charges first. However, in our intervention, the decision of distribution is not with the lender but with the consumer. This is particularly relevant as most people hold more than one credit card[13]. Gathergood

[13]In our sample we find that on average participants hold 2.5 credit cards as illustrated in Table III.1

et al. (2019) report that according to Trans Union data from 2015 71.5% of credit hard holders have two or more credit cards. Foster et al. (2011) even shows that U.S. Americans hold 3.7 credit cards on average. What is even more, these credit cards are often provided by different issuers, which emphasizes the relevance for consumers to understand how to repay efficiently, as credit card issuer do not possess information of other issuers. The tested people should understand for themselves what the best allocation is. Salisbury (2014) notices a deterioration of the payment behavior associated with a poor understanding of the financial context. This problem should be improved in the pamphlet intervention, as we provide participants with background information on the repayment process reducing the necessity of financial literacy for this problem.

The second intervention simplifies the long explanation of the first one to a short description of the optimal repayment decision, i.e. paying back the highest interest credit card. Soll et al. (2013) point out in their study that despite the disclosures of the CARD Act, individuals still have problems comprehending the calculations and understanding the published content. In addition, in the slider intervention participants are provided with an application where they can use a slider to experience the effects of their repayment decision more directly. Although this intervention significantly improves repayment behavior as compared to the control group, it does not outperform the pamphlet. The advantage from financial literacy seems to be weaker in the slider intervention as compared to the control group. In the pamphlet intervention, a high value in financial literacy remains an advantage when solving the repayment problem.

Third, the reminder intervention has the intention to intervene in case participants sub-optimally allocate money to both credit cards. As money that is paid back in excess of the minimum payment has to be repaid to the highest interest rate credit card as

and have additional access to another 0.75 credit cards.

per regulatory requirement, a comprehensive analysis over several credit card providers becomes crucial since the regulation then no longer applies. The implementation of a reminder significantly depends on aggregated data availability and hence might be implemented by a FinTech that collects information on household debt over several debt institutions. One possible solution are multi-banking apps, where households can voluntarily aggregate their banking data. As our results indicate, such a reminder significantly reduces misallocation. Without a reminder, people tend to repay less efficient as the increasing average misallocation in subsequent experiment rounds of the control group shows. Our reminder intervention is far beyond what is currently enforced by the regulator. As of the current CARD Act regulation, consumers are reminded that minimum repayment only comes at higher costs than the repayment of larger amounts, and still no information is provided regarding the optimal allocation of repayments between two credit cards. However, as our results show that households vary their repayment over time and hence unlearn the correct repayment process, a regular reminder that takes into account the repayment per month over aggregated debt data can significantly reduce misallocation.

Fourth, the assistant intervention represents the strongest of the four interventions. Again, this intervention requires data among a variety of credit card providers to find a perfect solution. It provides both a short one-liner and a variant of a reminder, as in the reminder intervention, and an assistant which informs about the optimal distribution and warns about non-optimal distributions before the payment is confirmed. Financial literacy does not play a role anymore, which implies that these strong interventions both in the reminder and in the assistant intervention allow all households independent of the prevailing level of financial literacy to reduce the misallocation. Even though the implementation of the pamphlet intervention allows households to repay their debt more efficiently, policy support should especially focus on adapted financial interventions, as all households, independent of their financial literacy, benefit. Since these two interventions require access to client related data from the respective institution, policy makers might pave the way and require financial institutions to provide an interface for multi-banking apps to gather the data if requested by the customer.

III.7 Conclusion

In this study, we focus on non-optimal household credit card repayment, which is a comparably new problem in the literature (Gathergood et al., 2020; Amar et al., 2011; Ponce et al., 2017). When people are faced with multiple credits, they do not use all their available money to repay the credit with the highest interest rate. In order to find practical methods to reduce such misallocation, we use an experimental setting in which participants are required to allocate a certain amount of money to two credits with different interest rates in 10 subsequent experiment rounds. We develop and test four treatments employing different financial interventions with a variation in the categories general vs. adapted interventions. Two interventions feature generalized interventions, the other two feature interventions adapted to the individual situation.

We find that misallocation almost vanishes when we provide participants with an assistant that tells them which credit has to be repaid first. This finding could be the basis for an app that helps people to organize all their credits and accounts. We also test less invasive financial interventions that need less personal information on the credit situation, such as providing a pamphlet or a program that tells participants the outcome of user-defined money splits exemplary before the experiment rounds. Although not as effective as an assistant app, all other interventions strongly reduce misallocation to a comparably degree. Furthermore we find that financial literacy of participants helps to reduce misallocation, but seems to become less effective in the adaptive financial interventions.

Our findings bear implications for policy makers, as - despite the CARD Act of 2009 - financial interventions can improve the repayment behavior of households. First, we explicitly educate participants using an information brochure that describes the optimal repayment process thoroughly. Even though the CARD Act provides additional facts on the repayment decision taken by individuals, it does not financially educate people. Furthermore, in case households possess credit cards from several credit card issuers, it becomes necessary to understand the repayment problem, as CARD Act requirements (e.g. issuers are obliged to distribute amounts in excess of minimum payments to the highest interest rate card) do not work in case of several credit card issuers. Additionally, the slider intervention provides an application which enables households to learn the repayment process. Policy makers should therefore consider to prescribe financial institutions the provision of additional information and guidance to reduce misallocation and significantly improve household welfare and stability. Second, the reminder intervention as well as the assistant intervention rely on the availability of data. We argue that financial services providers (e.g. FinTechs, banks, financial advisors) that aggregate data (e.g. voluntarily provided by households to a multi-banking app) are able to implement these interventions and hence significantly improve debt repayment. As a consequence, policy makers might consider to instruct financial institutions to provide an interface for data exchanges to multi-banking apps.

Even though we intensively analyze the relationship between misallocation and financial intervention, our study might be subject to certain limitations. While we find some evidence that the interventions might need reinforcement, our study cannot make any substantial statements about long-term learning effects. Furthermore, we only examine the interventions with regard to their differences in the misallocation, but not, for example, in the time it takes to learn the message, as would be necessary for a cost-benefit analysis of financial intervention. The interventions can be understood as different starting ideas for developing more practical and workable systems in real-life applications. While it is well beyond the scope of this paper, transferring our ideas could be tested using a field experiment in future research.

Chapter IV

Elementary Financial Decisions

Coauthor:

Darwin Semmler

Relative share:

50%

A previous version of this chapter has been presented at:

- Society for Experimental Finance Conference 2022

Elementary Financial Decisions

Abstract

We investigate elementary financial decisions such as "You can invest some money. Do you prefer to invest in a (safe) asset with 5% returns or in a (safe) asset with 10% returns, all else equal and no additional strings attached?" Such decisions are fundamental for all financial decisions, yet they have not been investigated experimentally. Using four different independent variables, we find that participants on average misallocate between around 3% to around 51% of the available money. Investment works far better than borrowing, while negative interest rate induce higher misallocation. A change in framing and reducing the options to a binary choice do not decrease misallocation. These effects might partly be driven by cognitive uncertainty, which is a particular form of confusion.

Keywords: Household finance, investing, experimental finance, elementary financial decisions

JEL-Codes: D14 - D91 - G41 - G51

IV.1 Introduction

Why do people fail to make optimal financial decisions? Researchers have amassed a huge mountain of evidence that they actually do (e.g. Beshears et al. (2018); DellaVigna (2009); Zinman (2015)), but the reason remains an open question. Theories to explain this phenomenon usually employ complex dimensions, such as uncertainty in prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), or time in (quasi-)hyperbolic discounting models (for an overview, see Cohen et al. (2020)). But to the best of our knowledge, no one ever checked if people "get the basics right". To illustrate what we mean by that, consider two examples:

- Example 1: You can invest some money. Do you prefer to invest in a (safe) asset with 6% returns or in a (safe) asset with 12% returns, all else equal, and no additional strings attached?
- Example 2: You need to borrow some money. Do you prefer to borrow for a 5% interest rate or a 10% rate, all else equal, and no additional strings attached?

Both examples offer a dominant alternative and abstract away from any complication and thus are very elementary. If we assume that people prefer more money over less money, they have a simple solution - invest in the 12% asset, borrow for 5%. However, when we experimentally investigate such *elementary financial decisions* similar to the examples, we find that our participants invest 7.8% of the money in the low-return asset and borrow 22.7% from the high-interest credit. When we vary such questions using four independent variables, the misallocation ranges from around 3% to around 51%. These results are puzzling, precisely because all these decisions are so simple. They are also important because most real financial decisions are more complicated since they are composed of such elementary financial decisions. If people do not consistently behave optimally in the elementary financial decisions, this non-optimality might spill over to the more complex decisions as well. Thus the goal of this paper is to investigate such elementary financial decisions. We document deviations from optimal behavior, and shed some light on when they happen, and why they happen.

We run two very similar, pre-registered (Gärtner and Semmler (2022) or see Appendix IV) experiments with three independent variables each. In both experiments, participants make 16 different financial decisions where they have an income to invest in two assets, or must cover expenditures by borrowing from two credits. Every decision problem has an optimal option for a participant who has a rational, monotonous preference for money because the assets or credits always differ in their interest rates. We then observe which fraction of the "financial means" (the money our participants decide about) is misallocated, i.e. either invested in the low-return asset or taken from the high-interest credit. The difference between both experiments is whether the financial means are freely divisible. In experiment #1, the financial means are divisible and participants can freely distribute them over both alternatives. In experiment #2, we force participants into binary choices, i.e. the whole sum must be invested in one asset, or borrowed from one credit.

The three independent variables both experiments share are motivated by the idea of "cognitive uncertainty" (Enke and Graeber, 2023). We assume that our participants have monotonous preferences for money, and we use Enke and Graeber (2023)'s model of cognitive uncertainty, which they understand as "subjectively perceived uncertainty about what the optimal action is", as our theoretical framework to explain non-optimal decision-making. In this model, people solve problems with an (possibly subjective) optimal solution, but might not find this optimal solution, for example, because they do not know how to make sense of the provided information. People are aware of this cognitive noise, which creates cognitive uncertainty. One core result of Enke and Graeber (2023) is that this uncertainty leads to a "shrunken action", i.e. an action that is dampened to a prior. The higher the cognitive uncertainty, the more dampened the reaction is. We argue that in our experiments, the profit-maximizing solution is what participants would want to implement if they experienced zero cognitive uncertainty, but cognitive uncertainty dampens their reaction towards an even split, which we assume is the ignorance prior if both assets or credits are perceived as equally likely to be the profit-maximizing solution. Shrinking to this prior creates misallocation.

We exogenously manipulate cognitive uncertainty using three within-subject variables. We use simple framing with which we intend to decrease cognitive uncertainty, by reporting either the interest rates of the alternatives or the already calculated payments expressed as sums of money - we believe the latter to be simpler. In a second treatment, we use negative interest rates to increase cognitive uncertainty, because we believe that our participants are not really familiar with them. Third, we argue that borrowing induces more cognitive uncertainty than investing.

We find that cognitive uncertainty increases misallocation, but only under divisible money. We argue that divisibility is required to properly translate cognitive uncertainty into behavior by splitting the money. In a binary decision, any doubts participants might have about the perceived optimal solution cannot be expressed properly. Beyond that, we find clear effects for borrowing. Participants report less cognitive certainty for borrowing, and also misallocate substantially more financial means compared to investment decisions, up to around 20 percentage points in difference. Negative interest rates also increase cognitive uncertainty and misallocation, but their effect on misallocation is way stronger for borrowing than for investment decisions. Forcing participants into binary decisions decreases cognitive uncertainty, but not misallocation. The percentage frame increases misallocation in the investment decisions significantly in the simplest model under divisible money, but this result is not robust, so we conclude the frame did not work for investing. For borrowing, the percentage frame has an effect, but contrary to our hypothesis, percentages actually help participants decrease their misallocation.

Our paper contributes to the literature in several ways. First, we simply document that systematic misallocation occurs even on the most elementary levels of financial decision-making. This is in line with recent literature on credit card repayments (Amar et al., 2011; Gathergood et al., 2019; Ozyılmaz and Zhang, 2020; Ponce et al., 2017) and borrowing (Agarwal et al., 2015), but our experiment is even simpler than the decisions analyzed in these papers. Credit card repayment is not quite as an elementary decision as investing or borrowing, because debt repayment decisions necessarily need to include negative balances (i.e. you need to have debts to repay them). All these papers use balances as an explanation in some way. Ozyılmaz and Zhang (2020) show experimentally that balances influence repayment decisions roughly as strongly as the interest rates, Gathergood et al. (2019) find that people use a balance matching heuristic in the field, and Amar et al. (2011) - as well as our experiment in chapter I - find additional heuristics and fallacies which rely on specific combinations of balances, income, and interest rates. We do not model any balances in our experiments, yet we still find misallocation, which indicates a need to explore explanations beyond balances, such as cognitive uncertainty. Agarwal et al. (2015) show for borrowing that in a field experiment where participants could decide between a credit card with an annual fee and a lower APR and one with no annual fee but a higher APR, around 40% choose the suboptimal card. However, this decision is again more complex than ours, because we only model the APR in our experiments.

Second, on a broader scale, our results suggest that explanations that use additional dimensions to explain non-optimal behavior cannot explain all the variance in non-optimal behavior, because they usually at least assume some kind of monotonicity. This fits nicely with the results of Dembo et al. (2021) who find a similar pattern in experiments with situations of uncertainty. Their experiments show that while participants do violate the relatively high-level assumption of independence from irrelevant alternatives (which common modern theories such as rank dependent utility ((Quiggin, 1982), Quiggin (1993)) or cumulative prospect theory (Tversky and Kahneman, 1992) give up), they far more often violate lower level assumptions such as ordering or first-order stochastic dominance (which these modern theories still assume as well). We do not argue to abandon such higher-level theories, but to complement them with theories about something like confusion.

Third, we discuss if one of these supplemental theories can be Enke and Graeber (2023)'s model of cognitive uncertainty, as this model leads to non-optimal behavior even if agents have monotonous preferences. We use the model to successfully predict misallocation by exogenously varying cognitive uncertainty. However, we also show that this model needs to be complemented by other explanations because it is not strong enough to explain all the differences between our treatments, nor is it particularly robust. In particular, we show that an alternative explanation might be "account confusion": Participants understand that the optimal solution is to use the financial means for one account only, but might be confused about which of the accounts is the correct one. This goes against cognitive uncertainty, which would predict that participants distribute their financial means over both accounts, based on the shrinking argument. In investment decisions, we find that shrinking is more important than account confusion, but for borrowing decisions it is the opposite. This suggests that for these decisions, mental gaps

(Handel and Schwartzstein, 2018) might be more important than cognitive uncertainty, i.e. people are not only confused about whether their preferred solution is actually the optimal one, but have problems finding the optimal solution in the first place.

The remainder of this paper is structured as follows: In section IV.2 we develop the theoretical background of our experiment. Section IV.3 describes the general design, variables, hypotheses, results, and robustness checks of experiment #1, section IV.4 those of experiment #2. In Section IV.5 we compare the results of both experiments. Section IV.6 discusses the results and concludes.

IV.2 Theoretical background

Our definition of an "elementary financial decision" for this paper contains three aspects:

- The decision-maker needs to decide about *something*, in our case about "financial means" - that is, some money or a money-equivalent about which they decide, including income, wealth, expenditures, or debt in any form.
- The decision-maker has exactly two alternatives to choose from, which differ in only one dimension, where a dimension refers to one property expressed through one variable. Having less than two alternatives constitutes no decision. Having more than two alternatives can be decomposed into sequences of choices with two alternatives, thus it is not the most elementary decision.
- The alternatives need to be presented with as few dimensions as possible. This aspect is important for three reasons. First, people have to evaluate if a dimension is important, and the fewer dimensions there are, the simpler this process is. This argument holds true even if the dimensions are (supposedly) irrelevant to the

actual decision problem, or have identical values in both alternatives. Second, after the relevant dimensions are acknowledged, this aspect minimizes the minimal number of required comparisons to see if the alternatives differ on a dimension, and if they differ, how. Third, each extra dimension might constitute an interaction effect with another dimension, even if this other dimension does not differ between the alternatives. The fewer possible interactions, the more elementary the decision.

Additionally, we need an auxiliary assumption about the alternatives to distinguish optimal from non-optimal behavior. Alternatives in financial decision problems usually differ in dimensions such as returns, uncertainty, liquidity, maturity, and so on. For simplicity, we focus on returns. While all these dimensions are preference-based, which makes it hard to observe non-optimal behavior, it is common to assume monotonicity for the preference for money. We assume that people invest to make as much money as possible, and prefer to pay as little for credit as possible. This additional assumption enables us to conceptualize *misallocation* of money, which we define as the share of financial means put into a dominated alternative. This is in line with the interpretation in other recent papers which focus on non-optimal borrowing (Agarwal et al., 2015) and debt repayments (Amar et al., 2011; Gathergood et al., 2019; Ozyılmaz and Zhang, 2020; Ponce et al., 2017). Focusing on returns is also probably the most generous setting for our null hypothesis, which is rational choice, i.e., zero misallocation. However, in the field, it is rare that such an elementary financial decision as we understand it exists, if any at all. This is why we use an experimental approach.

We conceptualize elementary financial decisions as situations of cognitive uncertainty (Enke and Graeber, 2023). Enke and Graeber (2023) define cognitive uncertainty as "subjectively perceived uncertainty about what the optimal action is". Unlike the canonical concept of uncertainty, which understands uncertainty as random outcomes of lotteries, cognitive uncertainty can occur in situations of perfect objective certainty (i.e. a choice between only degenerated lotteries). In Enke and Graeber (2023)'s model, people solve problems with an (possibly subjective) optimal decision p, but only have noisy access to that p. People have a prior p^d about p, which Enke and Graeber (2023) assume to be non-informative, and then receive a noisy signal $s = p + e$, where e is an error term indicating cognitive noise. People are aware of this cognitive noise, which creates cognitive uncertainty. Their optimal action depends on a weighted linear combination of the prior p^d and the signal s. The respective weights depend on cognitive uncertainty. Enke and Graeber (2023) show that this setup leads to a reaction that is dampened to the prior, and the higher the cognitive uncertainty, the more dampened the reaction is.

We argue that this is the situation in our experiments. Here, participants have to make several decisions where they either decide about financial means, concretely in which of two assets to invest a sum of money, or from which of two credits to borrow to cover some expenditures. The alternatives differ in returns or interests. These are paid or charged with certainty, which creates choices between two degenerated lotteries. In our case, p is the share of the financial means dedicated to the dominating alternative (i.e. the high-return asset or the low-interest credit) and always equals 1. Yet participants experience uncertainty because they do not fully understand that p should always equal 1. Before observing the interest rates, participants are indifferent between both alternatives. Once they observe the difference in the interest rate, they understand that this difference favors one alternative over the other, but they do not necessarily know which exact action should follow from that understanding. Applying the model by Enke and Graeber (2023) implies that our participants' reactions are biased towards the uninformative prior $p^d = 0.5$, which results in misallocation. This leads us to our first hypothesis:

Hypothesis 1 (H1): The higher the cognitive uncertainty, the higher the misallocation.

To investigate H1, we need variation in cognitive uncertainty. However, cognitive uncertainty is a state of mind, and we are not aware of any methods to manipulate a state of mind in a direct and controlled manner. Instead, we follow Enke and Graeber (2023)'s approach and manipulate cognitive uncertainty indirectly. For example, in one of their experiments, participants have to make risky decisions and the authors compare behavior when the alternatives are compounded with behavior when the alternatives are not compounded. They show that compounding increases cognitive uncertainty and that cognitive uncertainty influences behavior in the respective experiment. For a causal interpretation, they assume that cognitive uncertainty is the only causal pathway between compounding and the respective dependent variable. We follow that example by using three different independent variables to exogenously vary cognitive uncertainty - a difference in framing, negative interest rates, and income valence (investing vs. borrowing).

IV.3 Experiment #1

IV.3.1 General design

We preregistered the general idea, hypotheses, variables, outscreen processes, N, and the analyses for both experiments (Gärtner and Semmler (2022), or see Appendix IV).

Experiment #1 starts with the experimental stage, which consists of 19 financial decision problems. In each decision, participants have financial means, which is either some amount of money to invest in one of two assets, or a deficit to cover by borrowing from one of two credits. Participants can distribute the financial means freely over both alternatives. The first three decision problems are unincentivized trials. In these trials, participants can test the mechanics of the experiment. While the first trial only features assets with positive returns, we confront the participants with negative returns in the other trials and - exclusive to the third trial - with credits to borrow from. We use browser message boxes in the second and third trials to remind the participants to pay attention to which variant of the decisions they are dealing with. We excluded the three trials from the data analysis.

For the 16 remaining decisions, we pay participants a bonus between 0 to 20 pence per decision. The bonus scales linearly with the share of financial means put into the optimal alternative, i.e. the high-interest asset or low-interest credit. We only show the total sum of bonuses a participant earned after they made the last decision, without any performance feedback within the experimental stage. In each decision, we vary the financial means and the returns/interests. Participants type the sum of the means they want to use for each alternative into a text field. They have to invest or borrow the full sum. To make the utilization of the text field approach easier, we interactively show participants the remaining amount of money to distribute in real-time.

Before we confront participants with the experiment decisions, we start the experiment with instructions, in which we explain the rules and incentivization. Since the experiment makes use of Javascript, we exclude participants who disabled Javascript in their browser right from the start. We use three comprehension tasks to ensure the understanding of the incentivization and experiment rules. Furthermore, we ensure a basic understanding of percentages by requiring participants to calculate 1% of 1000. The participants have to correctly answer this question as well as the comprehension tasks in order to proceed.

After the experimental stage ends, we ask participants to briefly describe the strategy they used in their last decision problem in an open question. We do not analyze this question, but instead use it to screen out people who gave nonsensical answers to this questions. Two raters independently analyzed whether the answers matched the question, no matter what was actually answered. We screened our every participant where both raters agreed that the answer was nonsensical.

A post-experiment questionnaire follows the experimental decisions, where we measure experience with assets and credits, financial literacy, preference for numerical information, numeracy, consumer confidence, risk affinity, and basic demographics in this order. We start with measuring the experience of the participants with credits or assets by asking them if they have credit card debts, and how many investment and borrowing transactions they usually execute per year. We measure financial literacy by counting the correct answers of the Big3-questions from Lusardi and Mitchell (2011) as well as three questions especially tailored for debt literacy by Lusardi and Tufano (2015). The measures for preference for numerical information, numeracy, and consumer confidence are all taken from Fernandes et al. (2014). Preference for numerical information is measured as the mean of eight questions on a 6-point Likert scale between 1=strongly disagree and 6=strongly agree. We measure numeracy (from study 2 of Fernandes et al. (2014)) as the number of correct answers out of eleven questions mainly covering calculations about percentages. Consumer confidence is calculated as the mean of five questions on a Likert scale between 1 and 6. Finally, we measure risk affinity as suggested by Falk et al. (2023) by letting the participants make decisions along a decision tree. Participants have to make five hypothetical choices between a sure payment and a lottery with a 50 percent chance of payment. The lottery stays the same in all choices, while the sure payment varies depending on the decisions participants make. The final measure for risk affinity varies between 0 and 31, where 31 is the maximum risk affinity.

Furthermore, we include three attention checks in the post-experimental questionnaire, but *not* in the experimental stage. We reject participants who fail at least two of these checks. For our analyses, we additionally exclude participants who failed any attention check. The experiment closes with demographic questions (gender, age, and years of education) and lets participants comment on the experiment.

We run the experiment using the experimental software SoPHIE (Hendriks, 2012) and recruit our participants from the online crowd-sourcing platform Prolific (For a discussion of Prolific, see Palan and Schitter (2018)). Following our preregistration plan, we recruit 240 participants. We restrict our sample to US participants who claim to be fluent in English to avoid language problems and enforce an equal gender split. We pay a show-up fee of £2.50 and a bonus of up to £3.20.

IV.3.2 Experimental variables

We measure misallocation, our dependent variable, as the share of financial means dedicated to the dominated option, i.e. either invested in the low return asset or borrowed from the high-interest credit.

To measure cognitive uncertainty we ask our participants how certain they are that their solution maximizes their payoff in this decision, which they indicate with a percentage scale slider. This follows the approach from Enke and Graeber (2021), except for that these authors did not use a slider with 1% steps, but a horizontal list with 5% steps.

The first experimental variable we investigate is the context of investment or borrowing; we call this variable the income valence. We vary the income valence for two reasons. First, we believe that borrowing induces higher cognitive uncertainty, and second to not accidentally miss patterns that may be different for different valences. At latest since the seminal work of Kahneman and Tversky (1979), which among other concepts introduced the idea of loss aversion, economists acknowledge that the valence of a decision problem can have an influence on decisions. In a paper very close to ours, Ozyılmaz and Zhang (2020) for example find that in an experiment that compares debt repayment and investing decisions, their participants misallocate less in the investment decisions. However, while most theories explain differences between gains and losses on a preference base, we argue that borrowing also increases cognitive uncertainty. People are more familiar with positive numbers, which increase in their absolute value as the number itself increases. This concept is inversed with negative numbers: A greater absolute value results in a smaller number, which effectively inverses the measure of misallocation compared to absolute values. We argue that participants struggle with this additional notion. Therefore, the cognitive uncertainty - and subsequently the misallocation - in the negative income valence (borrowing) should increase compared to the positive income valence (investment).

Our second independent variable is the sign of the interest rates, which is supposed to increase cognitive uncertainty. We hypothesize that participants have more problems understanding negative interest rates than positive interest rates, which creates different levels of cognitive uncertainty. We argue that in financial contexts people expect returns to increase investment, and interest rates to increase a credit sum, but negative interest rates decrease investments and debts instead. We assume that this mismatch with expectations induces cognitive uncertainty. Additionally, nominal negative interest rates are very rare in the field, such that we can expect participants to be less familiar with them, which should increase cognitive uncertainty as well.

Our final independent variable is a framing intervention which we expect to decrease misallocation. Consider the following two decision problems:

- "You have to invest a sum of £200. The returns of one asset are 4%. The returns of the other asset are 12%. Which asset do you prefer to invest in?"

vs.

- "You have to invest a sum of £200. The returns of one asset are £10, if all money is invested there. The returns of the other asset are £30, if all money is invested there. Which asset do you prefer to invest in?"

Both decision problems are almost identical, they only differ insofar as the first problem presents the returns as percentages and the second as the actual amount of pound sterling. We argue that it is easier for people to deal with concrete terms such as pound sterling (see e.g. Hoffrage et al. (2000); Gigerenzer et al. (2007), while percentages may be more confusing because participants do not understand that when comparing different percentages of the same base, the comparison is just as simple. The concrete terms are also the results from the calculation that the percentages may induce participants to make. Thus percentages should increase cognitive uncertainty and misallocation.

We summarize our hypotheses with respect to the indirect manipulation of cognitive uncertainty:

Hypothesis 2a (H2a): Cognitive uncertainty is higher in the negative income valence treatments (borrowing).

Hypothesis 2b (H2b): Misallocation is higher in the negative income valence treatments *(borrowing).*

Hypothesis 3a (H3a): Cognitive uncertainty is higher in the negative interest rates treatments.

Hypothesis 3b (H3b): Misallocation is higher in the negative interest rates treatments.

Hypothesis 4a (H4a): Cognitive uncertainty is higher in the percentage treatments.

Hypothesis 4b (H4b): Misallocation is higher in the percentage treatments.

IV.3.3 Results

We conducted the experiment in February 2022. Due to our out-screening procedures, we had to recruit two additional participants in April 2022. A total of 302 participants started the study. 32 participants quit before finishing the experiment, further 5 participants dropped out due to time-out. From the remaining 265 participants, 22 participants did not pass all attention checks and a further 3 participants were rated as potential bots by two raters. Our final data set consists of 240 participants, thereof 88 males, 111 females, 6 people of a third gender, and 35 persons who denied information about their gender. The average participant in our data set is about 40 years old, with 16 years of education. The study took a mean duration of 26.5 minutes with an average payment of £4.98 (including the participation fee of £2.5[1]). The average hourly wage was around £14.02, which is in line with usual experimental payments. Table IV.1 shows the summary statistics.

We start the analysis with an overview of misallocation in general. In Table IV.2 we report the average misallocation of participants in different treatments. For a graphical

[1]£3 for the two participants recruited in April due to an increase in the minimal hourly wage on Prolific.

Table IV.1: Summary statistics of experiment #1

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Uncertainty	240	21.21	25.33	0	0	11.5	34.2	100
Age	239	39.97	13.89	18.00	29.00	37.00	50.00	77.00
Years of education	239	16.05	2.60	10.00	14.00	16.00	17.50	23.00
Fin. literacy	240	3.80	1.27	0	3	4	5	6
Numeracy	240	9.52	1.42	3	9	10	10	11
Cons. Confidence	240	3.60	1.28	1.00	2.80	3.80	4.60	6.00
Pref. num. info	240	4.51	1.00	1.38	3.75	4.69	5.28	6.00
Risk seek	240	9.07	5.11	1	5	9	12	32
# of yearly invest transactions	206	14.54	27.36	0.00	0.00	2.00	15.00	150.00
# of yearly credit transactions	202	313.43	4,223.34	0.00	0.00	0.00	1.00	60,000.00
Duration total (min:sec)	240	26:29	17:09	9:20	16:53	21:53	32:34	202:26
Duration pre exp	240	6:45	9:09	0:46	2:56	4:04	7:17	94:09
Duration exp	240	9:03	7:24	2:27	5:39	7:15	10:32	98:55
Duration PEQ	240	10:41	5:36	2:47	7:02	9:25	13:04	44:37
Payoff (USD)	240	4.98	0.58	3.71	4.50	4.98	5.50	5.70
Gender info	Males: 88		Females: 111		Third gender: 6		NA: 35	
Credit card debt info	Has debt: 111			Does not have debt: 126			NA: 3	

representation of misallocation and uncertainty in general as well as in different treatments see Figures Appendix IV.9 and Appendix IV.10. The misallocation varies in the full range between 0% and 100% in each treatment, i.e. there are always participants investing or borrowing perfectly optimal, but also perfectly non-optimal. It stands out that while in most of the treatment variations (except in the borrowing treatments with negative interest rates) more than half of the decisions do not exhibit any misallocation at all, the average misallocation greatly differs. We find far more misallocation in the borrowing treatments. The average misallocation also increases for negative interest rates. This effect is even stronger for borrowing decisions, where the average misallocation almost reaches random level. However, there seems to be no consistent effect of the percentage frame. It slightly decreases misallocation in the investment treatments, but increases it in the borrowing treatments.

Table IV.2: Misallocation statistics of experiment #1

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Investing						
Pos. int. rates & No percentages	0.0%	0.0%	0.0%	7.0%	0.0%	100.0%
Pos. int. rates & Percentages	0.0%	0.0%	0.0%	8.5%	0.0%	100.0%
Neg. int. rates & No percentages	0.0%	0.0%	0.0%	12.3%	1.2%	100.0%
Neg. int. rates & Percentages	0.0%	0.0%	0.0%	13.6%	9.3%	100.0%
Borrowing						
Pos. int. rates & No percentages	0.0%	0.0%	0.0%	24.9%	50.0%	100.0%
Pos. int. rates & Percentages	0.0%	0.0%	0.0%	19.5%	18.7%	100.0%
Neg. int. rates & No percentages	0.0%	0.0%	52.8%	50.0%	100.0%	100.0%
Neg. int. rates & Percentages	0.0%	0.0%	37.5%	45.0%	100.0%	100.0%

Figure IV.1 shows barplots for both misallocation and cognitive uncertainty (shorter relabeled as "Uncertainty"). Although uncertainty generally takes on low values around 20% and does not vary in the same magnitude as the misallocation, it varies jointly with misallocation.

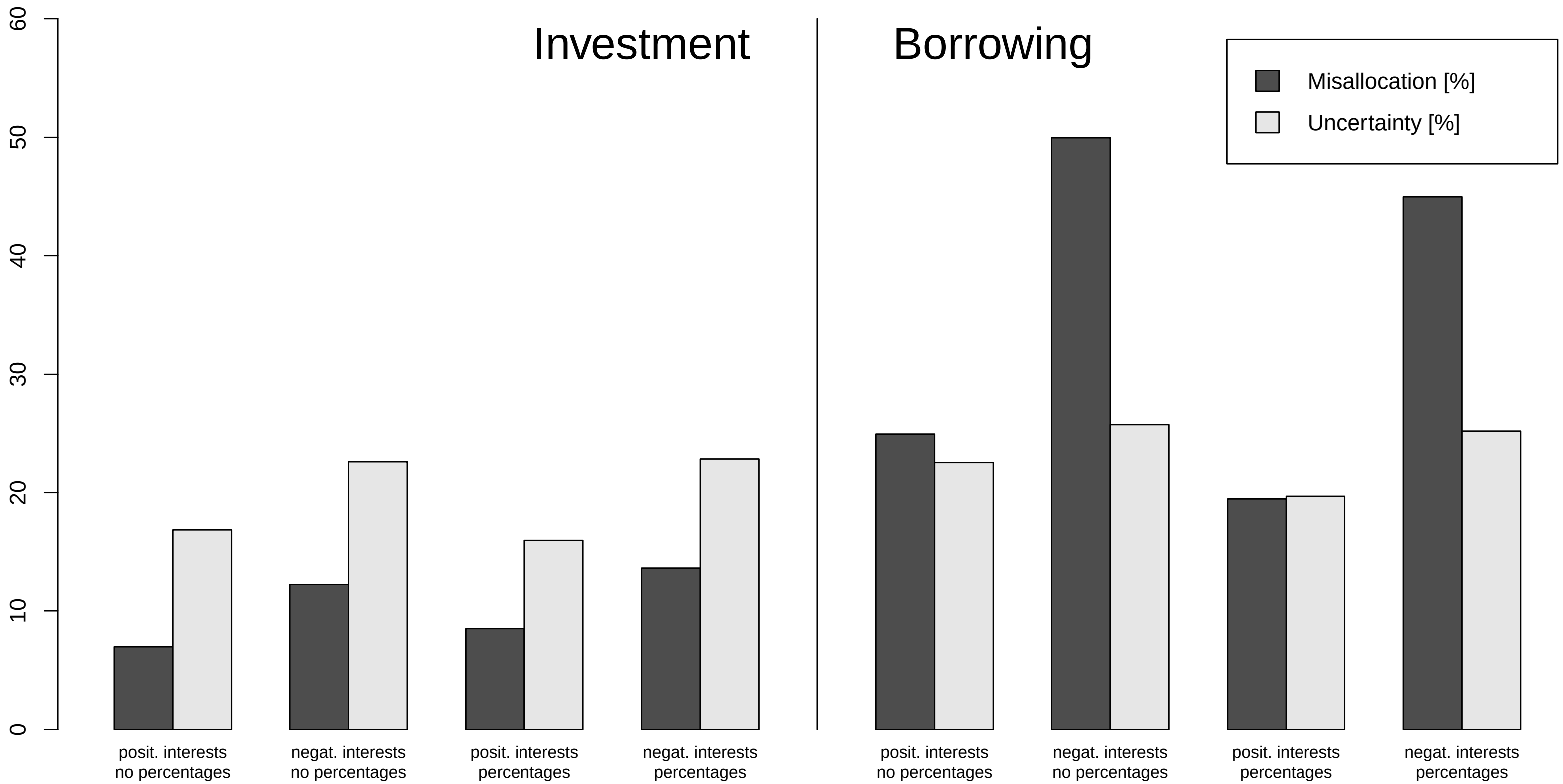


Figure IV.1: The Figure shows barplots of average percentage points in misallocation and uncertainty split by the 8 treatments. The barplots on the left side correspond to investment in assets, the one on the right correspond to borrowing.

In the next step, we test our hypotheses with regression models. Since we have a within-design, we employ random effects regressions (i.e. with a random intercept term for every participant), where the "round", i.e. the randomized position of a certain decision problem from 1 to 16, constitutes the time dimension. We use random effects since a fixed effects regression would not be able to estimate the effect sizes of the constant control variables as their influences are completely captured by the participant-wise intercept terms (Wooldridge, 2010). We test our hypotheses with a two-fold regression analysis. In the first step, we run a manipulation check, i.e. we investigate the influence of our treatments on cognitive uncertainty using three different sets of independent variables and control variables. We then regress misallocation to the same variables and add cognitive uncertainty as an additional regressor. For an experiment, it is a bit unusual to use cognitive uncertainty directly as an independent variable instead of only using the treatment variables as the main independent variables[2]. We do so because the analysis will show that the treatments themselves have other effects well beyond those from cognitive uncertainty, so interpreting the treatments as the effects of cognitive uncertainty would overestimate its effect dramatically. Table IV.3 shows the results for both dependent variables. For the sake of brevity, we do not display the individual control variables but include the complete regression table as Table Appendix IV.36.

Columns (1), (3) and (5) describe the models with uncertainty as the dependent variable, measured on a scale between 0 and 100. Columns (2), (4) and (6) model the influences on misallocation, also measured on a scale between 0 and 100. We adjust p-values using the Holm-Bonferroni method. We adjust within the models for each hypothesis, which are reflected in the main effects of borrowing, negative interest rates, percentage frame and uncertainty (i.e., three adjustments for models with uncertainty

[2]See chapter V for a detailed discussion of that point.

Table IV.3: Random effects regression of experiment #1[a]

	DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	5.669***	17.329***	5.674***	17.325***	6.915***	13.702***
	(0.939)	(2.225)	(0.937)	(2.213)	(1.150)	(2.524)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.735***	4.652***	5.750***	4.673***	6.082***	3.441
	(0.877)	(1.383)	(0.879)	(1.382)	(0.930)	(1.556)
	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.027]
	[0.000]	[0.002]	[0.000]	[0.002]	[0.000]	[0.081]
Percentage frame	-0.885	1.630*	-0.451	-0.653	0.981	-2.819
	(0.638)	(0.781)	(0.968)	(1.737)	(0.902)	(2.083)
	[0.166]	[0.037]	[0.641]	[0.707]	[0.276]	[0.176]
	[0.166]	[0.037]	[0.641]	[0.707]	[0.276]	[0.176]
Uncertainty		0.111*		0.112*		0.084
		(0.040)		(0.040)		(0.049)
		[0.006]		[0.005]		[0.088]
		[0.012]		[0.011]		[0.175]
Borrowing × Negative int. rates	-2.542*	20.034***	-2.579*	19.995***	-2.550	21.162***
	(1.263)	(3.450)	(1.269)	(3.436)	(1.544)	(4.116)
Borrowing × Percentage frame	-1.952	-6.777***	-1.938	-6.793***	-2.519*	-6.888**
	(1.051)	(1.706)	(1.047)	(1.704)	(1.246)	(2.096)
Negative int. rates × Percentage frame	1.121	-0.271	1.132	-0.339	0.466	0.121
	(0.967)	(1.583)	(0.966)	(1.586)	(0.944)	(1.924)
Triple interaction	1.173	0.464	1.193	0.540	1.305	-0.307
	(1.621)	(3.157)	(1.620)	(3.146)	(2.046)	(3.838)
Round			0.112*	0.026	0.169*	-0.004
			(0.057)	(0.107)	(0.067)	(0.126)
Right 2nd			0.065	0.439	0.102	-0.510
			(0.418)	(0.999)	(0.479)	(1.128)
Starkness			0.002	-0.033	0.005	-0.048
			(0.011)	(0.023)	(0.011)	(0.028)
Starkness × Percentage frame			-0.008	0.043	-0.022	0.060
			(0.014)	(0.032)	(0.015)	(0.039)
Constant	16.858***	5.099***	15.753***	6.354***	78.036***	62.887***
	(1.423)	(1.082)	(1.607)	(1.896)	(16.789)	(12.154)
Observations	3840	3840	3840	3840	2624	2624
# participants	240	240	240	240	164	164
Individual control variables	No	No	No	No	Yes	Yes

Mediation analysis of uncertainty mediating misallocation - Sobel test

	(1)–(2)	(3)–(4)	(5)–(6)
Borrowing	0.629*	0.635*	0.581
	(0.254)	(0.254)	(0.358)
	[0.013]	[0.012]	[0.105]
	[0.036]	[0.034]	[0.306]
Negative int. rates	0.636*	0.644*	0.511
	(0.253)	(0.254)	(0.312)
	[0.012]	[0.011]	[0.102]
	[0.036]	[0.034]	[0.102]
Percentage frame	-0.098	-0.051	0.082
	(0.083)	(0.117)	(0.100)
	[0.239]	[0.665]	[0.411]
	[0.239]	[0.665]	[0.411]

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the regression results for uncertainty and misallocation under divisible money, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates and percentage frame, as well as uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

as the dependent variable, and four adjustments for models with misallocation as the dependent variable). The first two columns describe the minimal model, which only captures the influence of the three varying variables for the treatments as well as all interactions. Models (3) and (4) add all experiment-specific control variables. These are first the *Round*, i.e. the position of a decision, to capture potential learning effects. Second, we include the dummy variable *Right 2nd*, which takes on the value of 1 if and only if the optimal alternative is shown as the second option in the second line of the experimental screen, in case participants generally prefer the first option. Third, the variable *starkness* is equal to the difference in returns/interests of the two alternatives, either in percentages or in absolute values.[3] We furthermore include an interaction term between starkness and the *percentage frame*-variable, because percentage spreads - even on roughly the same scale - might not be comparable to differences in absolute money values. Models (5) and (6) add participant-specific control variables. We also include interaction terms between all treatments, but since we do not hypothesize any, we refrain from deep interpretations and simply highlight significant effects in an exploratory spirit. Because of that spirit, we do not adjust their p-values.

We analyze the results in the order of our hypotheses, starting with the influence of uncertainty on misallocation. In the models describing misallocation without including individual control variables, we see a significant positive effect of uncertainty on the misallocation. That is, the more uncertain participants are, the more misallocation they expose in their decisions. This effect vanishes in the model which includes all variables, however, we also lose roughly a third of our observations, so this might be a test power problem. We interpret these results as weak evidence for hypothesis H1, that cognitive

[3]In the case of percentages we multiply the difference by 10 to keep the starkness on a comparable scale with the starkness of the absolute values.

uncertainty increases misallocation.

Borrowing leads to more uncertainty and misallocation. In all models, the average misallocation increases by more than 13% when participants have to borrow instead of investing in assets. This is completely in line with the hypotheses H2a and H2b. Additionally, in all models, we detect a strong positive effect of negative interest rates on uncertainty. This is true for misallocation as well, even controlled for uncertainty, except in the full model (6), which gets insignificant when adjusting p-values. These results strongly indicate the correctness of hypotheses H3a and H3b.

While model (2) shows a positive effect of percentages instead of absolute values on misallocation, this effect is not stable in the models including control variables. We will later reanalyze this model with an additional independent variable in Table IV.10, and in this later model, the effect does not survive p-value adjustment. Because of that, and since there is also no effect on uncertainty, we decide against confirming any of the hypotheses H4.

Exploring the interaction terms, we find that debts combined with interests strengthen the effect on misallocation by another 20 percentage points, even though in two out of three models, uncertainty actually decreases in these decisions - this goes against H1. These results are reflected in Table IV.2 and Figure Appendix IV.10, where these treatment variations accumulate the highest misallocation of nearly 50% on average. It is also notable that we measure a strong negative interaction between *borrowing* and *percentage frame* on misallocation suggesting that for borrowing, percentages indeed help avoid misallocation, which is the opposite of H4b.

We finally turn to the question of whether cognitive uncertainty is a mediator on the path to misallocation. We follow the approach of Baron and Kenny (1986b) and run a Sobel test (Sobel, 1982) to test for each treatment variation if the effect on misallocation is mediated by cognitive uncertainty. We run the test for each set of control variables. The lower part of Table IV.3 shows the results. We detect a significant mediating effect of uncertainty for the borrowing treatment and negative interest rates, but not for the percentage frame. The latter result is unsurprising as we also do not measure any effect of the percentage frame on uncertainty at all. However, if we include all control variables, the mediating effect of uncertainty for the other treatment variables loses significance. Therefore, we interpret this as weak evidence for mediation of the effects of valence and misallocation via uncertainty. Cognitive uncertainty plays a certain role when determining misallocation in financial decisions, but it is far from explaining non-optimal decisions completely, and the size of its effect is dwarfed by the other channels through which the treatments have effects on misallocation - the coefficients of the Sobel test are around 1 to 1.5 orders of magnitude lower than the treatment effects of borrowing and negative interest rates.

IV.3.4 Robustness checks

We employ several additional checks to ensure robustness in our results. First, we investigate a different notion of misallocation by comparing decisions that have no misallocation to decisions that exhibit any form of misallocation. Therefore, we change the dependent variable to a new dummy variable *misallodummy* that takes on the value 1 if a participant in the observed round misallocates any of the available money, and 0 only if the misallocation was exactly zero. In 1379 out of 3840 decisions (about 35.9%) we detect misallocation greater than zero. Table Appendix IV.37 shows the results. We do not display the regressions with uncertainty as a dependent variable here as they are identical to the ones displayed in Table IV.3. In this analysis, the main effect of borrowing remains, while the negative interest rates lose their significance. However, the Sobel test shows that the mediation still remains significant in all three models. We interpret this as a complete mediation of the effects of negative interest rates via cognitive uncertainty, which is perfectly in line with our hypothesis. Borrowing also shows an indirect effect via uncertainty but has an additional direct effect which we did not hypothesize. The percentage frame does not work in any instance. The interactions between borrowing and negative int. rates as well as between borrowing and the percentage frame are still significant.

In the next robustness check we repeat the main analysis but now include subjects that we originally screened out because they failed the attention tests or the raters interpreted their answers to the open question after the experimental stage as nonsensical. Table Appendix IV.38 shows our results. The interpretations for the hypotheses still hold. For concerns regarding participants who took too long or were too quick to complete the experiment, we created a subset of our data that excludes participants who were below the 2.5% (corresponding to 11 minutes and 17 seconds) or above the 97.5% quantile (corresponding to 59 minutes and 59 seconds) in the duration. Overall the results and interpretations for our hypotheses as shown in Table Appendix IV.39 remain the same as in the main analysis.

We also apply several hypothesis tests to check our model assumptions. All the methodological variables in Table IV.3 are insignificant, except for the round, which increases uncertainty. However, a χ^2-test detects no significant influence of the order of the rounds on uncertainty (p=0.8951) or misallocation (p=0.6834). Also, the order of the assets or credits - that is which asset or credit was presented first - does not matter significantly for uncertainty (p=0.7039) or misallocation (p=0.3400). Furthermore, we check for a potential learning effect for the round and run a paired sample t-test for differences of the first decision to the last - the sixteenth - decision. Again there is no significant difference for uncertainty (p=0.2531) and misallocation (p=0.4928). If the round actually increases cognitive uncertainty, this effect is very mild.

IV.3.5 Additional analyses about cognitive uncertainty

In this subsection, we employ additional analyses to investigate the effects of cognitive uncertainty on misallocation and to test the model of cognitive uncertainty more rigorously. For transparency, we want to highlight that these analyses are not part of the original research agenda for this paper, but interesting ideas we developed after the experiment. We neither preregistered them, nor did we optimize our experiment for them, and some are motivated by the results.

Lack of test power due to missing values

Considering Table IV.3 again, we find that uncertainty loses significance once we include all control variables, which sparks doubts about its actual explanatory power because cognitive uncertainty might simply proxy the effects of the other cognitive financial variables such as financial literacy, numeracy or experience. However, in the full model, we lose 1216 observations which correspond to 76 participants - almost one-third of our data - due to missing values on some control variables. So the question arises whether cognitive uncertainty has real explanatory power that simply got undermined due to low power, or whether it merely is an alternative measure for other "finance-cognitive" variables.

The summary statistics in Table IV.1 show that three variables are responsible for the loss of observations: the two transaction variables which we jointly use to proxy financial experience, and the gender. Therefore we run additional regressions of misallocation and uncertainty where we omit different constellations of these variables. Table IV.4 shows the results.

The models (1) and (2) denote the regressions of uncertainty and misallocation without the experience variables. In models (3) and (4) we omit gender and in models (5) and (6) we omit both gender and the experience variables. First, in models (1), (3) and (5) we see that cognitive uncertainty correlates with financial literacy, numeracy, and consumer confidence, and also the number of yearly transactions in model (3). This suggests that cognitive uncertainty could indeed proxy for these other variables. With respect to misallocation, cognitive uncertainty only remains significant (and has a mediating effect for borrowing and negative interest rates) when we exclude gender only. It is not significant when we also exclude financial experience, or when we only exclude experience. These results suggest that on the one hand, we cannot confirm that the lack of significance when including all control variables is a mere power problem. In model (6) we only lose 4 out of 240 participants and still uncertainty remains insignificant. On the other hand, it stands out that uncertainty is significant in the analysis where we include the transaction variables and only exclude gender. If cognitive uncertainty really was only a proxy for all these cognitive financial variables and itself without merit, we should have found no significance in model (3), but oddly enough this is the only configuration where cognitive uncertainty has a significant influence on misallocation. This strengthens the case for cognitive uncertainty. Its insignificance in the models where we include gender could indicate that cognitive uncertainty measures some particular gender effects, but then we should also see a significant coefficient of cognitive uncertainty in model 6, which we do not. Overall the results are mixed enough to neither accept nor discard cognitive uncertainty based on test power concerns.

Table IV.4: Additional random effects regression without low observation variables[a]

	DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	6.206***	16.003***	6.263***	15.034***	5.759***	17.506***
	(1.049)	(2.343)	(1.026)	(2.383)	(0.953)	(2.223)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.958***	4.545**	5.933***	3.648*	5.836***	5.066**
	(0.942)	(1.495)	(0.893)	(1.399)	(0.892)	(1.405)
	[0.000]	[0.002]	[0.000]	[0.009]	[0.000]	[0.000]
	[0.000]	[0.007]	[0.000]	[0.027]	[0.000]	[0.001]
Percentage frame	0.214	-0.797	0.605	-2.369	-0.445	-0.951
	(1.048)	(1.889)	(0.822)	(1.899)	(0.986)	(1.730)
	[0.839]	[0.673]	[0.462]	[0.212]	[0.652]	[0.583]
	[0.839]	[0.922]	[0.462]	[0.212]	[0.652]	[0.583]
Uncertainty		0.031		0.105*		0.060
		(0.042)		(0.046)		(0.042)
		[0.461]		[0.022]		[0.150]
		[0.922]		[0.043]		[0.300]
Borrowing × Negative int. rates	-2.629	19.124***	-2.516	22.201***	-2.604*	19.134***
	(1.404)	(3.656)	(1.417)	(3.731)	(1.290)	(3.464)
Borrowing × Percentage frame	-2.368*	-7.907***	-1.873	-5.613**	-1.950	-7.010***
	(1.162)	(1.876)	(1.129)	(1.934)	(1.065)	(1.736)
Negative int. rates × Percentage frame	0.500	-1.087	0.941	0.381	1.165	-0.289
	(1.063)	(1.688)	(0.902)	(1.801)	(0.982)	(1.610)
Triple interaction	1.469	2.535	1.187	-1.794	1.195	0.384
	(1.831)	(3.345)	(1.843)	(3.573)	(1.648)	(3.196)
Round	0.154*	-0.005	0.122*	0.027	0.113	0.025
	(0.063)	(0.116)	(0.062)	(0.117)	(0.058)	(0.108)
Right 2nd	0.118	-0.178	0.140	-0.401	0.067	0.237
	(0.469)	(1.029)	(0.428)	(1.042)	(0.424)	(0.971)
Starkness	0.002	-0.047	0.006	-0.034	0.002	-0.036
	(0.012)	(0.025)	(0.010)	(0.026)	(0.011)	(0.023)
Starkness × Percentage frame	-0.013	0.047	-0.021	0.054	-0.009	0.049
	(0.016)	(0.033)	(0.014)	(0.036)	(0.015)	(0.031)
Age	-0.024	0.057	-0.015	0.017	-0.058	-0.001
	(0.100)	(0.084)	(0.085)	(0.077)	(0.080)	(0.076)
Female	-1.184	-5.170*				
	(2.986)	(2.388)				
Third gender	8.257	-2.687				
	(9.277)	(6.293)				
Has credit card debts	-1.737	-3.153	-3.552	-1.048	-3.464	-1.462
	(2.680)	(2.269)	(2.965)	(2.220)	(2.478)	(2.090)
# of yearly credit transactions			-0.000***	0.000***		
			(0.000)	(0.000)		
# of yearly investment transactions			-0.032	0.016		
			(0.040)	(0.042)		
Risk seek	-0.044	0.264	0.053	-0.154	0.146	0.287
	(0.255)	(0.242)	(0.290)	(0.232)	(0.217)	(0.213)
Years of education	0.778	-0.506	0.494	-0.959*	0.583	-0.605
	(0.513)	(0.497)	(0.589)	(0.437)	(0.491)	(0.443)
Financial Literacy	-4.696***	-5.145***	-4.424**	-2.664**	-4.202***	-4.247***
	(1.263)	(0.989)	(1.458)	(0.982)	(1.143)	(0.916)
Numeracy	-3.080**	-1.968*	-1.664	-2.173**	-2.596*	-2.064**
	(1.124)	(0.775)	(1.359)	(0.765)	(1.091)	(0.707)
Cons. Confidence	-4.887***	1.015	-4.311**	0.835	-4.615***	1.960*
	(1.223)	(1.035)	(1.322)	(1.096)	(1.125)	(0.948)
Pref. num. info.	-3.618	-3.209**	-3.007	-2.268	-3.289	-3.187**
	(1.914)	(1.237)	(2.158)	(1.184)	(1.834)	(1.137)
Constant	86.787***	66.722***	70.208***	60.284***	81.311***	58.579***
	(13.629)	(11.353)	(14.767)	(10.788)	(11.715)	(10.081)
Observations	3248	3248	3072	3072	3776	3776
# participants	203	203	192	192	236	236
Mediation analysis of uncertainty mediating misallocation - Sobel test						
Borrowing	0.194		0.657		0.345	
	(0.269)		(0.309)		(0.249)	
	[0.471]		[0.034]		[0.167]	
	[1.000]		[0.095]		[0.494]	
Negative int. rates	0.186		0.623		0.349	
	(0.258)		(0.290)		(0.251)	
	[0.469]		[0.032]		[0.165]	
	[1.000]		[0.095]		[0.494]	
Percentage frame	0.007		0.063		-0.027	
	(0.056)		(0.098)		(0.074)	
	[0.905]		[0.517]		[0.720]	
	[1.000]		[0.517]		[0.720]	

Note: * $p<0.05$; ** $p<0.01$; *** $p<0.001$

[a] This table shows the regression results for uncertainty and misallocation, each with three different models without certain variables with many missing values: The models (1) and (2) include all control variables except the numbers of yearly credit or investment transactions; the models (3) and (4) omit the gender variables; and the models (5) and (6) omit both sets of variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for Borrowing, Negative Interest Rates and Percentage frame, as well as Uncertainty, if applicable. Asterisks indicate significance after adjustment.

Fixed effects regression

The effects of cognitive uncertainty might be obstructed due to correlations with financial experience. Since we measure uncertainty multiple times per participant, our experiment allows us to differentiate between intra- and inter-individual effects of cognitive uncertainty. The intra-individual variation of cognitive uncertainty cannot be explained by financial experience as we can reasonably assume that it remains constant within the same participant in the short time window of our experiment.

Therefore we employ a fixed effects regression, where the total inter-individual variation is captured by the fixed effect term of each participant. Although we have no theory about the structure and relation of an intra- and an inter-individual part of uncertainty, we can at least investigate a part of uncertainty that is completely separated from financial experience. We report the results in Table IV.5.

Cognitive uncertainty does not show any significant effects regardless of the set of control variables we use. Consequently, we cannot detect any influences of intra-individual cognitive uncertainty on misallocation. If cognitive uncertainty plays a role in generating misallocation - which is not clear considering the non-robust results in our other analyses - it seems to generate its effects between different subjects, while the level of cognitive uncertainty within a participant seems to be irrelevant. However, as our experiment does not allow us to disentangle inter-individual effects of cognitive uncertainty from effects due to a lack of financial experience, we can again neither accept nor discard the concept. Future research could either find ways to separate cognitive uncertainty from financial experience or find practical instruments to proxy cognitive uncertainty and better catch its explanatory power.

Table IV.5: Additional random effects regression, fixed effects[a]

	DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	5.669***	17.857***	5.673***	17.864***	6.914***	14.020***
	(0.939)	(2.220)	(0.937)	(2.209)	(1.148)	(2.518)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.735***	5.186**	5.749***	5.219**	6.081***	3.720
	(0.877)	(1.378)	(0.879)	(1.376)	(0.927)	(1.546)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.017]
	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.052]
Percentage frame	-0.885	1.547	-0.457	-0.472	0.970	-2.571
	(0.638)	(0.785)	(0.968)	(1.735)	(0.896)	(2.079)
	[0.167]	[0.050]	[0.637]	[0.786]	[0.281]	[0.218]
	[0.167]	[0.100]	[0.637]	[1.000]	[0.281]	[0.436]
Uncertainty		0.018		0.017		0.038
		(0.051)		(0.051)		(0.064)
		[0.727]		[0.733]		[0.552]
		[0.727]		[1.000]		[0.552]
Borrowing × Negative int. rates	-2.542*	19.797***	-2.577*	19.748***	-2.548	21.045***
	(1.263)	(3.454)	(1.269)	(3.440)	(1.541)	(4.107)
Borrowing × Percentage frame	-1.952	-6.958***	-1.937	-6.976***	-2.518*	-6.998**
	(1.051)	(1.700)	(1.047)	(1.699)	(1.243)	(2.089)
Negative int. rates × Percentage frame	1.121	-0.167	1.133	-0.221	0.467	0.153
	(0.967)	(1.579)	(0.966)	(1.583)	(0.942)	(1.927)
Triple interaction	1.173	0.574	1.193	0.651	1.304	-0.257
	(1.621)	(3.155)	(1.620)	(3.146)	(2.041)	(3.833)
Round			0.112*	0.036	0.169*	0.004
			(0.057)	(0.107)	(0.067)	(0.126)
Right 2nd			0.087	0.415	0.123	-0.550
			(0.417)	(0.996)	(0.477)	(1.132)
Starkness			0.002	-0.033	0.005	-0.046
			(0.011)	(0.023)	(0.011)	(0.028)
Starkness × Percentage frame			-0.008	0.038	-0.022	0.055
			(0.014)	(0.032)	(0.015)	(0.039)
Constant	16.858***	6.669***	15.737***	7.903***	12.968***	7.301**
	(0.580)	(1.387)	(1.056)	(2.094)	(1.164)	(2.326)
Observations	3840	3840	3840	3840	2624	2624
# participants	240	240	240	240	164	164
Indiv. control variables	No	No	No	No	Yes	Yes
Mediation analysis of uncertainty mediating misallocation - Sobel test						
Borrowing	0.101		0.099		0.262	
	(0.294)		(0.295)		(0.448)	
	[0.731]		[0.737]		[0.559]	
	[1.000]		[1.000]		[1.000]	
Negative int. rates	0.102		0.100		0.231	
	(0.296)		(0.298)		(0.393)	
	[0.730]		[0.736]		[0.558]	
	[1.000]		[1.000]		[1.000]	
Percentage frame	-0.016		-0.008		0.037	
	(0.057)		(0.057)		(0.091)	
	[0.781]		[0.889]		[0.685]	
	[1.000]		[1.000]		[1.000]	

Note: * $p<0.05$; ** $p<0.01$; *** $p<0.001$

[a] This table shows the regression results for uncertainty and misallocation under divisible money, each with three different models without certain variables with many missing values: The models (1) and (2) include all control variables except the numbers of yearly credit or investment transactions; the models (3) and (4) omit the gender variables; and the models (5) and (6) omit both sets of variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for Borrowing, Negative Interest Rates and Percentage frame, as well as Uncertainty, if applicable. Asterisks indicate significance after adjustment.

Does behavior shrink?

While the theory of cognitive uncertainty predicts misallocation, it does not predict any arbitrary pattern of misallocation. Rather, with growing cognitive uncertainty, behavior should shrink to a prior p^d. This prior reflects a decision without knowing the interest rates. We did not measure any priors in our experiment, but we argue that the most plausible average ignorance prior is $p^d_{average} = 0.5$ (a fifty-fifty distribution). Participants might not have the same ignorance priors, but at least on average, it is plausible, because any $p^d_{average} \neq 0.5$ requires a systematic mechanism that enables the participants to identify the optimal account without any knowledge of the interest rates. We cannot think of any, except for some experimental features such as the account order, for which we already control both in the design and the statistical analysis. If this assumption is true, the model of cognitive uncertainty implies that the behavior shrinks closer to 0.5 the more cognitively uncertain our participants are.

Figure IV.2 shows the histograms for misallocation in each combination from each combination investment vs. borrowing and positive vs. negative interest rates (We ignore the percentage frame for brevity, as it is insignificant in most analyses). It is striking that we do not see that much shrinking in the data between the treatments, but rather a complete confusion of accounts. We first investigate shrinking in this section, and then account confusion in the following section.

We test shrinking more formally in two different ways. For both approaches, we construct a new variable *prior distance*, which measures the absolute distance from the actual decision to our assumed $p^d_{average} = 0.5$. It is simply calculated by $abs(misallocation - 50)$. This variable reaches between 0 and 50, where 0 means that a participant distributes the financial means equally, and 50 means that they distribute all financial means in one account (no matter the "right" or the "wrong" account). We then regress this variable on cognitive uncertainty in a univariate model. We use this first approach to test whether cognitive uncertainty holds at all. We find that one unit of uncertainty significantly decreases the prior distance by 0.191 on average (p-value $< 2 \cdot 10^{-16}$). Thus, we indeed measure a shrinkage to our assumed prior of $p^d_{average} = 0.5$ with growing cognitive uncertainty.

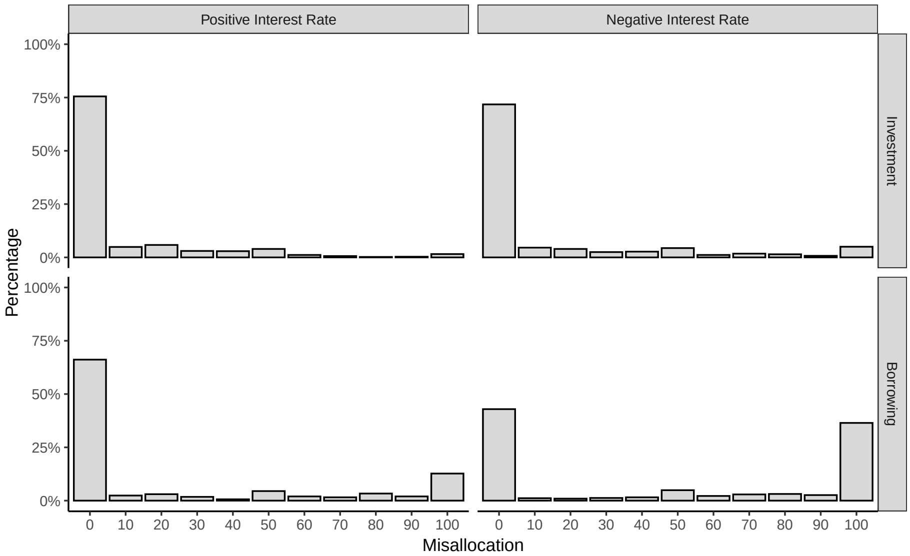


Figure IV.2: These histograms show the distribution of misallocation (x-axis) split by Investment vs. borrowing and positive vs. negative interest rates. The y-axis denotes the percentage of the 960 decisions in each treatment. Shrinking would imply an increase in the center for the treatments with more misallocation, for which we do not find much evidence.

The second approach uses the experimental treatments. We know from Table IV.3 that borrowing and negative interest rates increase cognitive uncertainty. This means that for the conditions "Borrowing with positive interest rates", "Investment with negative interest rates" and "Borrowing with negative interest rates", the distribution of the prior distance should be closer to 0 than for "Investment with positive interest rates".

Figure IV.3 shows the prior distance distribution for each of these treatments.

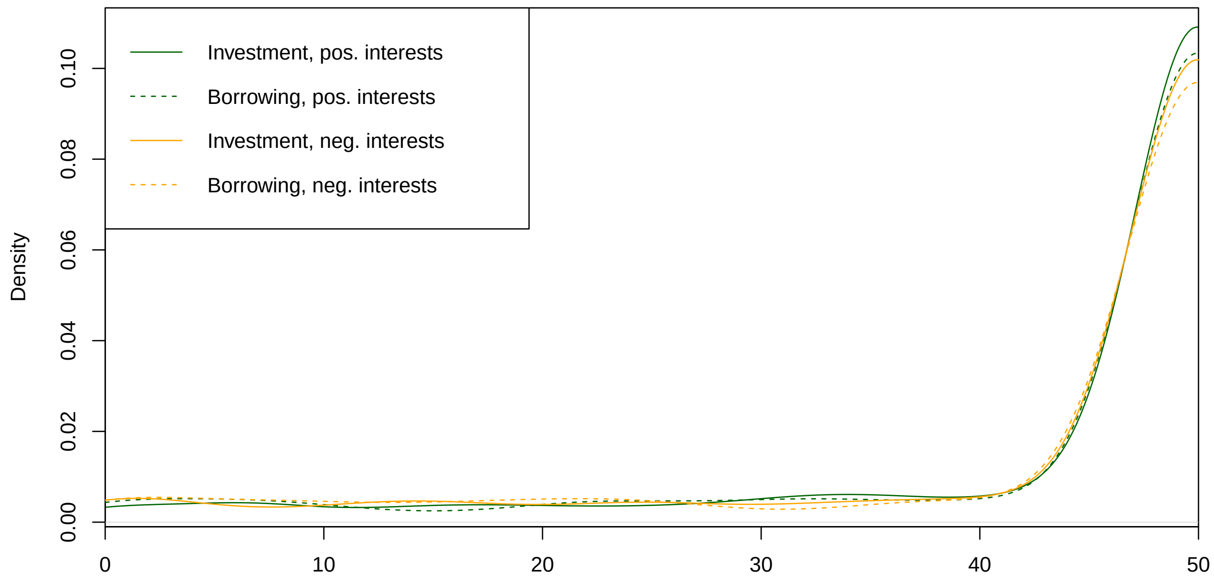


Figure IV.3: The Figure shows the estimated density functions for the distributions of the prior distance. The solid lines correspond to the investment treatments, the dashed lines correspond to the borrowing treatments. Green color indicates positive interest rates, orange color indicates negative interest rates.

The four empirical distributions all seem to be almost identical. The small differences at 50 indicate a tendency of fewer allocations to exactly one account for borrowing compared to investment and for negative interest rates compared to positive ones. This is in line with our finding of higher cognitive uncertainty in these situations. However, Kolmogorov-Smirnov tests of all pairwise constellations show no difference between these four distributions. Even the most dissimilar distributions - investment with positive interest rates and borrowing with negative interest rates - are not significantly different from each other (p-value = 0.3454). This indicates that the differences in misallocation between these four treatments are only negligibly explainable by shrinking. In summary, we find evidence for shrinking, but not necessarily as predicted. We would have expected more differences between the experimental treatments.

Account confusion

In a final additional analysis, we investigate an alternative reason for misallocation suggested by Figure IV.2: Participants might allocate all financial means to the non-optimal account. Such a behavior might also stem from some kind of uncertainty. For example, a subject might know that one of the accounts must be the financially optimal one, but variations in the valence of balances and interests impede the comprehension of which one the right account is. Such a subject would report more uncertainty and we are also likely to measure more misallocation, because of a higher probability of choosing the sub-optimal account. In this case, there exists a form of uncertainty that manifests in higher misallocation, but it is not the kind of uncertainty postulated by Enke and Graeber (2023) in the sense of a dampening to the prior. Instead, it seems more in line with the concept of "mental gaps" (Handel and Schwartzstein, 2018), which posits that people do not model the world correctly, which implies that they might not be able to identify the optimal solution in the first place. Therefore, we employ a dummy variable 100% *misallocation* with a value of 1 if the misallocation in one round is 100, and 0 otherwise. We run a logistic regression with 100% *misallocation* as the dependent variable and use the three sets of independent variables from the main analyses. We report the results in Table IV.6.

Table IV.6: Random effects logistic regression[a]

	DIVISIBLE		
	Dep. var.: 100% misallocation (= 1 if misallocation equals 100, = 0 otherwise)		
	(1)	(2)	(3)
Borrowing	3.877**	3.867**	3.293*
	(1.060)	(1.059)	(1.089)
	[0.000]	[0.000]	[0.003]
	[0.001]	[0.001]	[0.01]
Negative int. rates	2.504*	2.495*	2.274
	(1.068)	(1.065)	(1.091)
	[0.019]	[0.019]	[0.037]
	[0.038]	[0.038]	[0.074]
Percentage frame	1.135	0.786	-0.553
	(0.967)	(0.988)	(0.796)
	[0.240]	[0.426]	[0.487]
	[0.240]	[0.426]	[0.487]
Uncertainty	-0.014*	-0.014*	-0.019*
	(0.005)	(0.005)	(0.007)
	[0.004]	[0.004]	[0.004]
	[0.012]	[0.012]	[0.012]
Borrowing × Negative int. rates	-0.426	-0.420	-0.092
	(1.142)	(1.141)	(1.206)
Borrowing × Percentage frame	-1.507	-1.476	-0.721
	(0.979)	(0.981)	(0.797)
Negative int. rates × Percentage frame	-0.798	-0.772	0.168
	(1.014)	(1.016)	(0.834)
Triple interaction	0.869	0.854	0.115
	(1.037)	(1.039)	(0.908)
Round		0.015	0.002
		(0.014)	(0.017)
Right 2nd		0.091	0.021
		(0.129)	(0.155)
Starkness		-0.001	-0.003
		(0.003)	(0.004)
Starkness× Percentage frame		0.006	0.010
		(0.004)	(0.006)
Constant	-6.594***	-6.724***	-3.491
	(1.041)	(1.029)	(2.153)
Observations	3840	3840	2624
# participants	240	240	164
Individual control variables	No	No	Yes
Mediation analysis of uncertainty mediating misallocation - Sobel test			
Borrowing	-0.079*	-0.079*	-0.133*
	(0.031)	(0.031)	(0.052)
	[0.010]	[0.010]	[0.010]
	[0.027]	[0.027]	[0.027]
Negative int. rates	-0.080*	-0.080*	-0.117*
	(0.030)	(0.031)	(0.045)
	[0.009]	[0.009]	[0.009]
	[0.027]	[0.027]	[0.027]
Percentage frame	0.012	0.006	-0.019
	(0.010)	(0.014)	(0.020)
	[0.233]	[0.663]	[0.333]
	[0.233]	[0.663]	[0.333]

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the regression results for misallocation (as dummy variable) under divisible money with three different models: The simple model (1) which includes only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the model (2) which includes some technical aspects of the experiment; and the complete model (3) with all control variables. We omit regressions for uncertainty as they are identical to the main regressions. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for Borrowing, Negative Interest Rates and Percentage frame, as well as Uncertainty, if applicable. Asterisks indicate significance after adjustment.

The table shows that in all models there are significantly more rounds where participants choose the non-optimal account when borrowing or dealing with negative interest rates. This shows that some of the higher misallocation in these treatments can be explained by people confusing the two accounts. However, we cannot confirm account confusion in the percentage frame - which was expected as we also have no effect in the main analyses - and furthermore cannot detect any interactions between treatments. It is noteworthy that cognitive uncertainty has a negative effect. The more cognitively confused participants are, the less likely they are to misallocate all their financial means. This is again consistent with shrinking.

Furthermore, we run a "horse race" between account confusion and cognitive uncertainty to investigate which idea explains misallocation better. For each treatment, we select the decisions that exhibit misallocation and calculate whether they are closer to the prediction of the cognitive uncertainty model or closer to 100% misallocation, which corresponds to a confusion of accounts. To calculate a prediction of misallocation induced by cognitive uncertainty, we assume that the optimal decision for each participant is zero misallocation and that the non-informative prior always is an equal split between both accounts ($p = 0$ and $p^d = 0.5$). Furthermore, we assume that we can linearly translate cognitive uncertainty in a decision between p and p^d. Therefore, the predicted misallocation is always half the measured cognitive uncertainty.

Figure IV.4 shows the results. While in the investment treatments, the cognitive uncertainty model explains misallocation better than account confusion, it is the other way around in the borrowing treatments. Account confusion also seems to play a more prominent role under negative interest rates, but not enough to overshadow cognitive uncertainty in investment decisions. This analysis shows that the relative importance of both concepts depends on the valences.

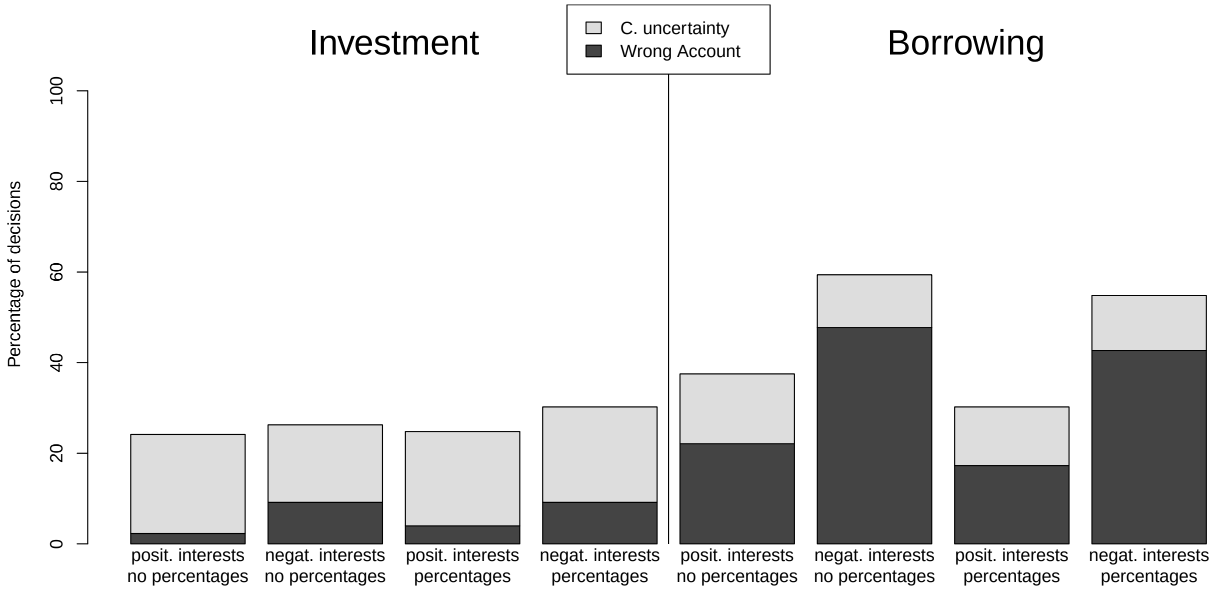


Figure IV.4: This barplot shows for each treatment the percentage of decisions that are either closer to the prediction of the cognitive uncertainty model or closer to 100% misallocation (account confusion). We only take decisions into account which exhibit misallocation at all.

Overall, all these additional analyses show that cognitive uncertainty might play an important role, but that it also has several problems. We find evidence of shrinking, and that it can explain misallocation in investment decisions relatively well. However, its effects do not survive every additional analysis and are overshadowed by account confusion in borrowing decisions. This in itself is an interesting insight, as cognitive uncertainty seems to be a relevant, but weaker driver of misallocation than we thought, but future research could nevertheless find ways to better disentangle intra- and inter-individual effects of cognitive uncertainty and better separate its effects from other valence-induced effects.

IV.3.6 Discussion of Experiment 1

In experiment #1, we allow our participants to freely distribute their financial means over both alternatives. Since in our experiment, financial means are basically money, this is only natural because divisibility is one of the fundamental properties of money. However, divisibility technically violates the "only two alternatives" condition we used to define "elementary financial decisions". With divisible financial means, a decision-maker has not only two alternatives, but a+1 options, where a is the amount of means represented in the smallest currency unit[4]. In this sense, experiment #1 is not the simplest analysis of our research question, because that would require only two options to choose from. However, one core attribute of money is its divisibility, so a binary choice would lose external validity in that regard. Additionally, this might also cost internal validity as well - if we take away a core element of financial decisions, do we still investigate financial decisions, or merely decisions that look like financial decisions, but really are not? We believe the simplest way to solve is tension is to run both experiments, so for conceptual clarity, experiment #2 investigates binary decisions where participants can only choose from two options.

IV.4 Experiment #2

IV.4.1 General design

With respect to the general design, variables, definitions and hypotheses, experiment #2 is as identical to experiment #1 as possible. The major difference is that participants

[4]For example, if you want to invest $100 and have to decide between two assets P and Q, you can invest 0 cents in P and 10,000 in Q, or 1 cent in P and 9,999 in Q, etc, up to 10,000 cents in P and 0 in Q, which gives you 10,001 alternatives to choose from.

cannot freely distribute their financial means over both alternatives, but have to choose exactly one option, which they use their entire financial means for. As a particular detail, we again use a text field as input. Technically this is an unnecessarily complicated format for a binary choice, but it enables a better comparison with experiment #1. Once participants type in a number, the other text field becomes closed and greyed (which can be undone by deleting the number). This ensures that a non-splitting decision is not substantially easier to apply than a splitting decision.

Divisibility of money might influence cognitive uncertainty via two possible channels: First, divisibility might increase cognitive uncertainty directly because it increases the number of effective options to choose from. If this mechanism exists, we should find that cognitive uncertainty is higher in the divisibility treatments. Second, divisibility allows to express cognitive uncertainty much better. If participants in experiment #2 are biased to their priors, but still lean towards one alternative, the binary nature requires them to choose that alternative. Under divisible money in experiment #1, they can express their bias, which should result in allocations that are less extreme, which in turn implies higher misallocation[5]. In the extreme case, this effect completely offsets any effect of cognitive uncertainty.

To summarize the hypotheses:

Hypothesis 5a (H5a): Cognitive uncertainty is higher in experiment #1 than in experiment #2.

Hypothesis 5b (H5b): Misallocation is higher in experiment #1 than in experiment #2.

Hypothesis 5c (H5c): The effect of cognitive uncertainty on misallocation is stronger in

[5]For example, if you think that you should invest 85% of your money in asset A, you misallocate 15% in experiment #1. In experiment #2, you cannot split the money, and - assuming you invest based on your tendency - instead invest 100% in asset A, this behavior results in 0% misallocation.

experiment #1.

Note that we will investigate hypotheses 5a and 5b as well as the interaction term cognitive uncertainty $\times$ experiment later in section IV.5, where we compare the results of both experiments. In this section, we instead investigate the hypotheses 1 to 4b.

We recruited 240 participants for experiment #2 as well, using the same exclusion criteria as in experiment #1.

IV.4.2 Results

We conducted the experiment in February 2022. We recorded 301 participants starting our study. We lost 28 participants who returned the study, and 2 participants due to time-out. We rejected another 2 participants because they failed multiple attention checks. Out of the remaining 269 approved participants, we removed a further 26 participants for failing at least one attention check. Another 3 participants were rated as potential bots. We remain with 240 participants in our final data set, 109 males, 96 females, 2 people of a third gender, and 33 persons who denied information about their gender. With a mean age of 38.6 years and roughly 16 years of education, the sample in our second study is comparable to the sample in experiment #1. The study took a mean duration of around 26 minutes. The average payment was £5.10 (including the participation fee of £2.5). The average hourly wage was £14.22, slightly higher than in the first experiment. We report the full summary statistics for these participants in Table IV.7.

We again start by exploring the average misallocation in different treatments. The results are shown in Table IV.8 and Figure IV.5. We show further depictions of uncertainty and misallocation in general and in different treatments in Figures Appendix IV.11

Table IV.7: Summary statistics of experiment #2

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Uncertainty	240	15.36	21.37	0	0	4.5	24.2	97
Age	240	38.57	12.10	18	30	36.5	46	76
Years of education	240	15.99	2.48	7	15	16	17	23
Fin. literacy	240	3.77	1.31	0	3	4	5	6
Numeracy	240	9.55	1.50	2	9	10	10	11
Cons. Confidence	240	3.63	1.37	1.00	2.60	3.80	4.60	6.00
Pref. num. info	240	4.51	1.01	1.00	3.75	4.62	5.28	6.00
Risk seek	240	9.38	5.16	1	5	10	13	25
# of yearly invest transactions	202	18.06	43.45	0.00	0.00	2.00	12.00	300.00
# of yearly credit transactions	207	79.50	1,044.53	0.00	0.00	0.00	1.00	15,000.00
Duration total (min:sec)	240	26:02	12:58	9:19	17:28	23:11	30:04	91:05
Duration pre exp	240	6:48	8:01	0:41	2:53	4:31	7:19	85:14
Duration exp	240	8:32	5:44	2:39	5:25	7:08	9:29	47:42
Duration PEQ	240	10:42	5:28	2:33	6:58	9:25	13:04	39:03
Payoff (USD)	240	5.10	0.58	3.50	4.70	5.30	5.70	5.70
Gender info	Males: 109		Females: 96		Third gender: 2		NA: 33	
Credit card debt info	Has debt: 114			Does not have debt: 123			NA: 3	

and Appendix IV.12. The results, especially the differences between the treatments, are almost identical to experiment #1, but misallocation is slightly lower. Borrowing leads to more misallocation than investing, ranging from 17.3% to 48.5%, while the values for the latter only vary between 3.3% and 8.1%. Negative interest rates seem to increase misallocation, while the percentages again show mixed results.

For a more detailed investigation of whether the results from experiment #1 hold, we replicate the random effects models. Table IV.9 shows the results. We find very similar patterns, with two important exceptions: The first exception is uncertainty no longer staying significant in the models (2) and (4). Together with the significant results from experiment #1, this provides evidence for H5c. The second exception is that negative interest rates are also no longer significant for misallocation in the investment treatments. Borrowing shows all the predicted patterns, which confirms H2a and H2b, but the percentage frame shows no main effects. We also find a similar interaction term pattern as in experiment #1. Misallocation is significantly higher for borrowing with negative interest rates, and the percentage frame decreases misallocation for borrowing decisions.

When running the Sobel test in a mediation analysis of the treatments on misallocation via cognitive uncertainty in none of the models we detect a significant mediation. This is consistent with the idea that since we force participants to choose exactly one of two options, they are not able to express cognitive uncertainty. So for the same reason we measure no main effect of cognitive uncertainty on misallocation, we consequently also cannot measure any mediating effect. This result can also explain why we do not find any significant effect of negative interests on misallocation anymore: The indirect channel via uncertainty is closed, and unlike the income valance, negative interest rates only have a very weak direct effect on misallocation, if at all.

Table IV.8: Misallocation statistics of experiment #2

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Investing						
Pos. int. rates & No percentages	0.0%	0.0%	0.0%	3.3%	0.0%	100.0%
Pos. int. rates & Percentages	0.0%	0.0%	0.0%	3.3%	0.0%	100.0%
Neg. int. rates & No percentages	0.0%	0.0%	0.0%	5.8%	0.0%	100.0%
Neg. int. rates & Percentages	0.0%	0.0%	0.0%	8.1%	0.0%	100.0%
Borrowing						
Pos. int. rates & No percentages	0.0%	0.0%	0.0%	22.9%	0.0%	100.0%
Pos. int. rates & Percentages	0.0%	0.0%	0.0%	17.3%	0.0%	100.0%
Neg. int. rates & No percentages	0.0%	0.0%	0.0%	48.5%	100.0%	100.0%
Neg. int. rates & Percentages	0.0%	0.0%	0.0%	41.2%	100.0%	100.0%

Table IV.9: Random effects regression of experiment #2[a]

	NOT DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	4.352***	19.519***	4.372***	19.506***	4.703***	16.805***
	(0.737)	(2.440)	(0.738)	(2.448)	(0.852)	(2.726)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	6.306***	2.407	6.307***	2.460	5.344***	3.020
	(0.886)	(1.489)	(0.892)	(1.484)	(0.901)	(1.488)
	[0.000]	[0.106]	[0.000]	[0.097]	[0.000]	[0.042]
	[0.000]	[0.318]	[0.000]	[0.292]	[0.000]	[0.127]
Percentage frame	0.179	-0.003	0.130	1.426	0.837	1.791
	(0.410)	(0.935)	(0.738)	(2.094)	(0.828)	(2.150)
	[0.662]	[0.998]	[0.860]	[0.496]	[0.312]	[0.405]
	[0.662]	[1.000]	[0.860]	[0.992]	[0.312]	[0.810]
Uncertainty		0.015		0.015		0.015
		(0.045)		(0.046)		(0.062)
		[0.747]		[0.748]		[0.811]
		[1.000]		[0.992]		[0.811]
Borrowing × Negative int. rates	-1.431	23.146***	-1.476	23.186***	-1.992	22.319***
	(1.080)	(3.443)	(1.087)	(3.442)	(1.327)	(3.988)
Borrowing × Percentage frame	-1.202	-5.607**	-1.227	-5.673**	-2.174**	-5.678*
	(0.755)	(2.028)	(0.756)	(2.028)	(0.837)	(2.527)
Negative int. rates × Percentage frame	0.394	2.286	0.380	2.138	0.428	0.799
	(0.692)	(1.778)	(0.692)	(1.748)	(0.780)	(1.693)
Triple interaction	1.494	-3.980	1.543	-3.777	3.185*	-5.066
	(1.196)	(3.465)	(1.200)	(3.455)	(1.241)	(4.125)
Round			0.051	-0.114	0.035	-0.107
			(0.048)	(0.124)	(0.057)	(0.148)
Right 2nd			0.047	1.219	0.067	0.756
			(0.392)	(1.127)	(0.460)	(1.355)
Starkness			-0.005	-0.005	0.008	0.001
			(0.010)	(0.030)	(0.012)	(0.035)
Starkness × Percentage frame			0.001	-0.024	-0.016	-0.026
			(0.012)	(0.035)	(0.014)	(0.039)
Constant	12.229***	3.154***	12.021***	3.706	48.677**	60.187***
	(1.127)	(0.952)	(1.314)	(2.147)	(16.466)	(17.789)
Observations	3840	3840	3840	3840	2656	2656
# participants	240	240	240	240	166	166
Individual control variables	No	No	No	No	Yes	Yes
Mediation analysis of uncertainty mediating misallocation - Sobel test						
Borrowing	0.064		0.064		0.070	
	(0.201)		(0.202)		(0.296)	
	[0.750]		[0.752]		[0.815]	
	[1.000]		[1.000]		[1.000]	
Negative int. rates	0.093		0.092		0.079	
	(0.290)		(0.290)		(0.336)	
	[0.749]		[0.751]		[0.814]	
	[1.000]		[1.000]		[1.000]	
Percentage frame	0.003		0.002		0.012	
	(0.021)		(0.036)		(0.074)	
	[0.901]		[0.958]		[0.867]	
	[1.000]		[1.000]		[1.000]	

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the regression results for uncertainty and misallocation under indivisible money, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. The reference group for gender is male. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates and percentage frame, as well as uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

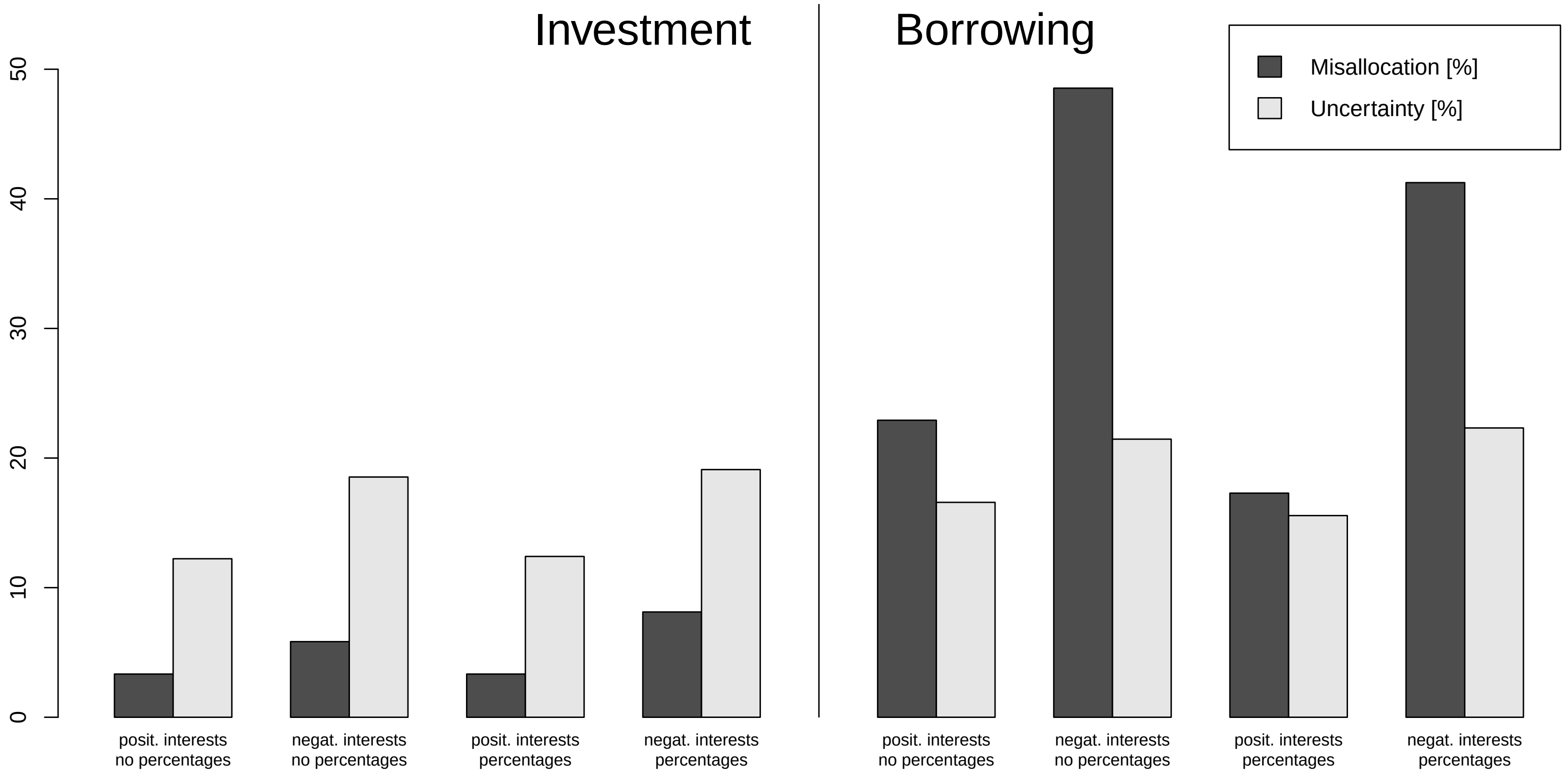


Figure IV.5: The Figure shows barplots of average percentage points in misallocation and uncertainty split by the 8 treatments. The barplots on the left side correspond to investment in assets, the one on the right correspond to borrowing.

IV.4.3 Robustness checks

For experiment #2 we run the same robustness checks as for the first experiment, except for the first check which used a dummy variable for misallocation. Since the misallocation in experiment #2 is measured binary by design, an additional check for this case is not necessary. Thus, we start with a repetition of the main analysis which includes screened-out subjects due to failed attention tests or bot-like answers in the open question. We show the results in Table Appendix IV.41. All coefficients keep their significance, but negative interest rates now are significant in model (6). This might indicate a very weak direct main effect of negative interest rates, but given that we have to include out-screened participants, this is very weak evidence. We also run a regression in which we screened out the 5% of participants with extreme experiment time, that is all participants who took less than 10 minutes and 53 seconds or more than 62 minutes and 38 seconds to complete the experiment. We display the results in Table Appendix IV.42. None of the significances from the main analysis change.

For the check of the experimental technicalities, we first want to highlight that unlike in experiment #1, the coefficient for round is insignificant. We use a χ^2-test to detect possible influences of the order of the rounds but do not find any connections to uncertainty (p=0.9298) or misallocation (p=0.8882). Furthermore, the χ^2-test for the order of the assets or credits does not detect significant influences on uncertainty (p=0.7055) or misallocation (p=0.6770). In contrast to experiment #1 the paired sample t-test to compare round 1 and round 16 shows a significant increase of uncertainty (p=0.0305, in round 16 approx. 2.9 units higher than in round 1), but not on misallocation (p=0.4072). So the participants show no learning effects over the round with respect to misallocation, but there is some very weak evidence that they become less certain in their decisions the longer the experiment lasts.

IV.5 Comparison of both experiments

We compare the results of the experiments #1 and #2 to investigate the hypotheses H5a and H5b. We pool the data of both experiments and add a variable *NotDivisible*, which equals 1 for experiment #2, and 0 for experiment #1. Furthermore, we include interaction terms between the treatment dummy variables and the experiment variable, to account for possible differences of influences of the treatments between the two experiments. Table IV.10 shows these results.

The *NotDivisible* coefficient indicates that uncertainty indeed decreases if money is not divisible, however not in the complete model. We interpret this as weak evidence for H5a. However, misallocation is not significantly lower in each model, so we cannot confirm H5b. For H5c, the situation is more complex. Strictly speaking, the interaction effect NotDivisible $\times$ Uncertainty is insignificant, so we cannot confirm H5c. However, recall that in Table IV.9 from the former section there is no significant effect of uncertainty on misallocation in any of the models. If this result here in Table IV.10 would be best interpreted as a true null result, this would imply significant effects of uncertainty on misallocation in Table IV.9 - after all, if the effect of uncertainty on misallocation is significant in the divisibility treatment, and the difference between divisibility and non-divisibility is truly non-existent, one would expect that misallocation has a significant effect in the non-divisibility treatment as well. But the coefficient there is insignificant, which means that we also do not have strong evidence for a true null effect. Additionally, when returning to Table IV.10, recall that the main effect of Uncertainty is the effect of Uncertainty on Misallocation in the Divisibility treatment. Note that this coefficient and the interaction term NotDivisible $\times$ Uncertainty almost cancel each other to 0. So even if the interaction term is insignificant, we still think it is plausible to conclude that

H5c is confirmed - the interaction term is insignificant, because its effect is too small to turn significant, given our test power, not because the insignificance reflects a true null result. However, this is weak evidence.

The interactions between *NotDivisible* and the other independent variables are all insignificant, so we assume that the results are roughly identical for all treatments.

We finally repeat the robustness checks we used for the single experiments, namely including all screened-out subjects in Table Appendix IV.44 and excluding participants who took less than 11 minutes and 2 seconds (2.5%-quantile) or more than 60 minutes and 55 seconds (97.5%-quantile) in Table Appendix IV.45, but do not detect considerable deviations from the main results.

In the next step, we take the mean of uncertainty and misallocation over all participants to average out individual fluctuations and obtain an overall difference between both experiments. As we remain with only one observation per participant, this renders the within-treatment variables (Borrowing, negative int. rates, percentage frame, and their interactions) irrelevant, as well as the experiment round specific variables (Round, Right 2nd, and Starkness). Thus, we only have to regress the influence on uncertainty and misallocation of the variables NotDivisible, Uncertainty, and their interaction, and additionally the models including the participant-specific control variables. We do this with OLS regression models and show the results in Table IV.11. The positive influence of uncertainty on misallocation in experiment #1 stays significant and does not vary significantly between both experiments, just like in Table IV.10. Furthermore, misallocation does not vary significantly between both experiments in the averaged data set.

As a side mark, an interesting question to analyze over both experiments is the correlation between uncertainty and the time taken for an experiment round. Although we did not state an official hypothesis it is reasonable to assume that participants who take longer for their decisions are less certain. We test this assumption with a simple OLS regression of uncertainty as the dependent variable and the duration of each experiment round as the independent variable. The resulting influence is significant and confirms the suspicion: For each second taken for an experiment round the uncertainty of a participant increases by a value of around 0.039 units (on the scale between 0 and 100, $p = 1.93 \cdot 10^{-5}$). The effect seems small, but given the fact that the standard derivation of the duration of one experiment round is around 30.9 seconds (with a mean of 21.37 seconds), this leads to a notable fluctuation on the uncertainty scale.

Table IV.10: Comparison of experiments: Random effects regressions[a]

	COMPARISON NOT DIVISIBLE - DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
NotDivisible	-4.629*	-1.795	-4.619*	-1.877	-3.396	-2.589
	(1.814)	(1.441)	(1.814)	(1.432)	(1.858)	(1.473)
	[0.011]	[0.213]	[0.011]	[0.190]	[0.068]	[0.079]
	[0.021]	[0.213]	[0.022]	[0.380]	[0.135]	[0.284]
Borrowing	5.669***	17.304***	5.676***	17.253***	6.909***	13.623***
	(0.938)	(2.223)	(0.938)	(2.216)	(1.146)	(2.505)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.735***	4.627**	5.751***	4.589**	6.066***	3.336
	(0.876)	(1.382)	(0.877)	(1.385)	(0.930)	(1.556)
	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.032]
	[0.000]	[0.004]	[0.000]	[0.004]	[0.000]	[0.160]
Percentage frame	-0.885	1.634	-0.702	1.132	0.808	-0.667
	(0.638)	(0.780)	(0.795)	(1.415)	(0.769)	(1.600)
	[0.165]	[0.036]	[0.377]	[0.424]	[0.294]	[0.677]
	[0.165]	[0.109]	[0.377]	[0.424]	[0.294]	[0.733]
Uncertainty		0.115*		0.116*		0.080
		(0.040)		(0.040)		(0.044)
		[0.004]		[0.004]		[0.071]
		[0.016]		[0.015]		[0.284]
NotDivisible × Uncertainty		-0.107		-0.107		-0.062
		(0.061)		(0.061)		(0.068)
		[0.079]		[0.077]		[0.367]
		[0.160]		[0.231]		[0.733]
NotDivisible × Borrowing	-1.317	2.242	-1.307	2.327	-2.206	3.211
	(1.192)	(3.300)	(1.196)	(3.297)	(1.430)	(3.662)
NotDivisible × Negative int. rates	0.571	-2.180	0.554	-2.105	-0.740	-0.419
	(1.246)	(2.030)	(1.251)	(2.027)	(1.295)	(2.150)
Borrowing × Negative int. rates	-2.542*	20.045***	-2.574*	20.090***	-2.526	21.323***
	(1.262)	(3.447)	(1.265)	(3.438)	(1.539)	(4.113)
NotDivisible × Borrowing × Negative int. rates	1.110	3.092	1.090	3.014	0.510	0.971
	(1.660)	(4.870)	(1.662)	(4.863)	(2.028)	(5.722)
NotDivisible × Percentage frame	1.065	-1.635	1.076	-1.513	0.194	0.102
	(0.758)	(1.217)	(0.757)	(1.205)	(0.759)	(1.183)
Borrowing × Percentage frame	-1.952	-6.768***	-1.944	-6.756***	-2.520*	-6.768**
	(1.050)	(1.705)	(1.048)	(1.701)	(1.244)	(2.089)
NotDivisible × Borrowing × Percentage frame	0.750	1.153	0.721	1.068	0.325	1.047
	(1.293)	(2.648)	(1.293)	(2.643)	(1.499)	(3.260)
Negative int. rates × Percentage frame	1.121	-0.276	1.126	-0.257	0.451	0.325
	(0.966)	(1.582)	(0.965)	(1.581)	(0.944)	(1.911)
NotDivisible × Negative int. rates × Percentage frame	-0.727	2.565	-0.748	2.458	-0.026	0.576
	(1.188)	(2.379)	(1.189)	(2.363)	(1.225)	(2.552)
Triple interaction	1.173	0.459	1.193	0.476	1.283	-0.503
	(1.619)	(3.154)	(1.619)	(3.143)	(2.040)	(3.833)
NotDivisible × Triple interaction	0.321	-4.430	0.356	-4.302	1.957	-4.626
	(2.012)	(4.683)	(2.016)	(4.661)	(2.388)	(5.613)
Constant	16.858***	5.025***	16.207***	5.968***	65.702***	63.204***
	(1.422)	(1.075)	(1.509)	(1.629)	(12.531)	(10.088)
Observations	7680	7680	7680	7680	5280	5280
# participants	480	480	480	480	330	333
Further experimental control variables	No	No	Yes	Yes	Yes	Yes
Further individual control variables	No	No	No	No	Yes	Yes
Note:						$^{*}p<0.05$;$^{**}p<0.01$;$^{***}p<0.001$

[a] This table shows the regression results for uncertainty and misallocation where we compare divisibility with non-divisibility, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates, percentage frame and NotDivisible, as well as uncertainty and NotDivisible × Uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

Table IV.11: Comparison of experiments: OLS regressions of participant average values[a]

	COMPARISON NOT DIVISIBLE - DIVISIBLE			
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)
NotDivisible	-4.146*	-2.262	-4.326*	-1.866
	(1.915)	(2.001)	(2.022)	(1.920)
	[0.031]	[0.259]	[0.033]	[0.332]
	[0.031]	[0.518]	[0.033]	[0.663]
Uncertainty		0.279***		0.153**
		(0.042)		(0.048)
		[0.000]		[0.002]
		[0.000]		[0.005]
NotDivisible × Uncertainty		-0.020		0.008
		(0.071)		(0.067)
		[0.784]		[0.902]
		[0.784]		[0.902]
Age			-0.025	0.066
			(0.086)	(0.064)
			[0.769]	[0.303]
Female			-0.644	-5.040**
			(2.539)	(1.722)
			[0.800]	[0.004]
Third gender			6.150	2.031
			(9.428)	(5.552)
			[0.515]	[0.715]
Has credit card debts			-0.962	-1.154
			(2.241)	(1.565)
			[0.668]	[0.461]
# of yearly credit transactions			-0.000***	0.000***
			(0.000)	(0.000)
			[0.000]	[0.000]
# of yearly investment transactions			-0.012	0.020
			(0.021)	(0.020)
			[0.556]	[0.329]
Risk seek			0.057	-0.335
			(0.224)	(0.171)
			[0.799]	[0.051]
Years of education			0.714	-0.214
			(0.476)	(0.382)
			[0.134]	[0.576]
Financial literacy			-3.544**	-3.758***
			(1.249)	(0.806)
			[0.005]	[0.000]
Numeracy			-1.965	-2.614**
			(1.098)	(0.795)
			[0.074]	[0.001]
Cons. Confidence			-4.039***	0.990
			(1.008)	(0.768)
			[0.000]	[0.199]
Pref. num. info.			-3.262*	-1.941*
			(1.543)	(0.914)
			[0.035]	[0.034]
Constant	21.421***	16.612***	72.252***	67.096***
	(1.412)	(1.452)	(12.769)	(10.076)
	[0.000]	[0.000]	[0.000]	[0.000]
Observations	480	480	330	330

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the OLS regression results for uncertainty and misallocation where we take the average of uncertainty and misallocation for each participant. This renders the dummy treatment variables and the experimental control variables irrelevant. We compare divisibility with non-divisibility, each with two different models: The simple models (1) and (2) which include only the NotDivisible dummy (and in model (2) the interaction with uncertainty and its main effect); and the complete models (3) and (4) with all control variables. The reference group for gender is male, p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values in models (2) and (4) are adjusted for NotDivisible, Uncertainty and NotDivisible × Uncertainty. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

IV.6 General discussion and conclusion

Our study shows that people seem to have some problems in solving the two easiest and most elementary decisions that we could think of, namely investing and borrowing. These results are usually driven by a minority - the median misallocation is 0% in all but two treatment combinations - but are predictable, relatively stable, and often quite strong. For investment, the deviations from optimality seem moderate, since even in the least favorable condition the misallocation averages no more than around 14%. However, for borrowing, the misallocation is around 15 to 20 percentage points higher, especially if combined with negative interest rates where it can reach values around 50% - basically random level. These effects are in part explainable by cognitive uncertainty, at least if it is possible to translate this uncertainty into actual behavior under divisible money. So in general, our predictions were reasonable.

However, there are some exceptions. While cognitive uncertainty seems to play a role, it can explain no more than around 12% points of the maximal misallocation in experiment #1 (comparing the estimates for 0 and 100% uncertainty, respectively), many treatment effects stay significant even after controlling for it, and the mediation analyses show that the other effects of the treatments beyond cognitive uncertainty are qualitatively way more relevant. Therefore, the model of cognitive uncertainty alone seems to miss important aspects. Misallocation under negative interest rates in particular offers something like "familiarity" as a natural additional explanation: People rarely if ever choose between *nominally* negative interest rates, even if the *real* interest rates might be negative. If people are not familiar with converting nominal into real terms, this might influence their behavior via another type of uncertainty which is different from cognitive uncertainty and also from the other similar variables for which we control, such as experience, financial literacy, education, and so on. This effect might be particularly strong for borrowing because investing in assets that turn out to have negative returns ex-post is common, so the concept of negative interest rates in the investing context might be familiar, while credits basically never have nominally negative interest rates, even ex-post.

A second explanation might point to the intuition behind prospect theory (Tversky and Kahneman, 1992). It is striking that misallocation increases for both the negative income valence and negative interest rates, and is maximized if both are combined - borrowing covers expenditures that participants might interpret as losses, and negative interest rates shrink the pie. However, there are some caveats to this interpretation. Prospect theory is a theory about the behavior under classical outcome uncertainty, which we do not model in our experiments, so core aspects such as compressed probability weighting or reversed risk preferences for losses cannot apply. The parts that can be adapted to situations under certainty are concerned with non-standard preferences, but note that the condition where the losses in the form of interests for credits are the highest is borrowing with *positive* interest rates, a condition which does not induce the highest misallocation. This suggests that developing preferences that can explain a maximum misallocation probably has to include some very arbitrary assumptions. However, the general intuition that losses might be more troubling than gains might apply to cognitive uncertainty or the behavior under confusion in general.

The percentage frame did not work as we expected. Indeed when we explore the interaction terms we find weak evidence that it might actually help to avoid misallocation when borrowing. This might indicate that we failed with our design choices. An alternative explanation is that the usefulness of percentages depends on the context. While there is evidence that they are more confusing in the context of probabilities than natural frequencies (Gigerenzer et al., 2007), they are the standard measure of returns or interests in the context of finance, so people might be familiar enough with them.

If we make money indivisible in our experiments, misallocation shrinks and the relationship between cognitive uncertainty and misallocation vanishes. This result suggests that divisibility as one of the core characteristics of money causes misallocation because it allows people to translate their uncertainty into behavior. We can interpret this as a hidden cost of using money. Usually, divisibility of money is seen as a desirable quality, because it allows a smoother expression of preferences and production costs which leads to more mutually beneficial trades[6]. We believe that this effect dominates the misallocation that stems from the possibility of expressing cognitive uncertainty in general. But there might be special cases, such as our experiment #2, where this is reversed. A more general hypothesis following this argument states that we should observe less optimal decisions for any variable that allows us to express uncertainty compared to a variable that does not, as long as the smoothing effect is not too strong.

We finish by drawing some additional conclusions for future research. First, we want to highlight that we do not dismiss the importance of cognitive uncertainty, or more generally, confusion. We still find that cognitive uncertainty seems to matter for elementary financial decisions, even though its effects are not that large and not that robust. However, our experiments model the simplest financial decisions imaginable. It is reasonable to assume that other important dimensions, such as time or risk, might be influenced stronger by cognitive uncertainty, and more complex decisions even more strongly. Second, our results suggest that in times of negative interest rates, the average

[6]Consider for example a situation where a seller has production costs of $4.40 and a buyer has a willingness to buy of $4,70. In this case, a mutually beneficial deal is possible at a price anywhere between these values. If we were to restrict prices to steps of one dollar, this deal would not be realized, because $4 is a too small incentive for the seller to produce, while $5 is too expensive for the buyer.

decision quality should decrease. And third, there might be other relatively elementary decisions, such as selling assets, where we might find misallocation, and it should also be fruitful to investigate who exactly misallocates, and why. Research that generalizes from investment decisions might underestimate behavioral phenomena because investing might turn out to be the one family of decisions where mistakes are relatively rare.

Chapter V

Looking beyond ESG preferences: The role of sustainable finance literacy in sustainable investing

Coauthors:

Alix Auzepy

Christina E. Bannier

Relative share:

50%

Looking beyond ESG preferences: The role of sustainable finance literacy in sustainable investing

Abstract

This paper investigates how knowledge of regulations, norms, and standards for financial products with sustainable characteristics, a concept called sustainable finance literacy, affects sustainable investment behavior. Using a large, pre-registered experiment, we propose a method for measuring sustainable finance literacy and its impact on investment decisions. We find that an increase in sustainable finance literacy leads to a 4 to 5% increase in the probability of investing sustainably. This effect holds particularly for individuals with high sustainability preferences, while for individuals with low sustainability preferences, we find some evidence that sustainable finance literacy might even reduce sustainable investments. We also find that individuals with high sustainability preferences discriminate more strongly between light green and dark green financial products, given a higher level of sustainable finance literacy. Our results underscore the role of knowledge in shaping sustainable investment decisions, highlighting the importance of factors beyond sustainability preferences.

Keywords: Sustainable finance literacy, sustainable investments, behavioral finance, SFDR, MIFID

JEL-Codes: G11, G18, G53

Funding: This work was financially supported by the "Stiftung für die Wissenschaft". Stiftung für die Wissenschaft had no involvement in anything study-related.

V.1 Introduction

The number of sustainable investment products, i.e., products considering environmental, social and governance (ESG) factors in portfolio selection and management, has risen substantially in recent years (GSIA, 2021).[1] However, according to a study by the German Institute for Retirement Provision, there exists "widespread uncertainty about how retail investors understand and evaluate sustainable investments and how these affect their investment decisions in detail" (DIA, 2020).

While retail investors often express a clear preference for sustainable choices, sustainable finance products currently account for only a small portion of their portfolios (DIA, 2020). In surveys, investors frequently cite a lack of product transparency and insufficient knowledge as barriers to sustainable investing (DIA, 2020; Dumas and Louche, 2015; Friede et al., 2015; Gutsche and Zwergel, 2020; Phillips and Johnson, 2019). In fact, retail investors wanting to invest sustainably are often faced with complex and at times intransparent information. As a result, making sustainable investment decisions typically involves additional layers of information complexity that prevent these investors from being able to align their investment choices with their stated sustainability preferences (Filippini et al., 2023; Paetzold and Busch, 2014; Anderson and Robinson, 2021).

Financial decisions in general are largely influenced by specific knowledge and experience (Lusardi and Mitchell, 2014). Consequently, it is reasonable to infer that sustainable financial decisions require not only financial knowledge, but also a solid understanding of the sustainability criteria applied to corresponding financial products. For

[1] In this paper, the terms "sustainable investments", "sustainable funds", "sustainable financial products", "ESG funds", "ESG financial products" etc. are all used synonymously. In addition, we use terms such as "sustainable investing", "ESG investing", "sustainable investment decisions" and "sustainable investment behavior" synonymously.

example, a person who has a comprehensive understanding of ESG is more likely to respond to ESG information and make investment decisions based on such information than someone who has never heard of the acronym. Following Filippini et al. (2023), we refer to this concept as "sustainable finance literacy", which these authors define as the "knowledge of regulations, norms, and standards about financial products that have sustainable characteristics". In this paper, we explore the effects of sustainable finance literacy on sustainable investing. Motivated by the importance of preferences for investment decisions (Bauer et al., 2021; Brodback et al., 2019; Hong and Kostovetsky, 2012; Riedl and Smeets, 2017; von Wallis and Klein, 2014), we also examine the relationship between sustainability preferences and sustainable finance literacy, and their role in shaping sustainable investment behavior.

To investigate the causal effect of sustainable finance literacy on ESG investing, we run a pre-registered experiment with a large sample of German participants. We randomly assign our participants either to a treatment group, which receives a brochure with simple information on key aspects of sustainable investing, or to a control group, which does not receive any information. The information in the brochure focuses on three key dimensions: ESG criteria, sustainable investment strategies, and regulation of sustainable investments in the European Union (EU), and is thus consistent with Filippini et al. (2023)'s definition of sustainable finance literacy.

In the context of our experiment, participants in both treatments, i.e., the brochure treatment and the control treatment, have to make four investment decisions. For each round of investment decision, they have to choose one out of three funds from a given selection of sustainable and conventional funds. Investment decisions are incentivized by a bonus mechanism that leads to potential real payoffs for the participants. The funds to choose from are actual financial products. We present these funds using the web interface of a large direct bank offering retail banking products and services.[2] In addition to a variety of financial information, this web interface also provides sustainability information for each fund. Finally, we include several questions aimed at measuring not only sustainable finance literacy, but also other important factors influencing sustainable investment decisions, including financial literacy, economic and sustainability preferences, environmental literacy, and perceived impact (Lusardi and Mitchell, 2014; Falk et al., 2023; Anderson and Robinson, 2021; Heeb et al., 2022).

German retail investors are particularly well-suited to study the effects of sustainable finance literacy on investment decisions as they are directly impacted by the Sustainable Finance Disclosure Regulation (SFDR). This regulation, which applies to all financial firms that market their financial products in the European Union (EU), classifies financial products such as mutual funds and exchange-traded funds (ETFs) according to the extent to which ESG objectives are pursued and promoted. This categorization is reviewed and enforced by the German Federal Financial Supervisory Authority (BaFin). One key objective of this regulation is to increase transparency of sustainable investment products and provide investors with additional ESG-related information. In our experiment, this regulation allows us to verifiably differentiate between conventional and ESG funds, which is necessary for the construction of our dependent variables.

Our analysis delivers the following key findings: First, we provide evidence that sustainable finance literacy plays a key role in shaping investment choices. Providing some basic information about ESG criteria and portfolio selection strategies can have a substantial effect on individuals' knowledge of sustainable products and, consequently, on their probability of engaging in such investments. As a starting point, we analyze the effect of our brochure treatment using three different model classes (simple, medium,

[2]We use the interface of ING-DiBa AG, which is part of the Dutch ING Group.

complex), which incorporate different sets of control variables. We find that the total effect of the treatment is an increase in the probability of choosing a sustainable fund of around 9%. We examine this result further using a causal mediation analysis and find that 4 to 5% of this increase can be directly attributed to an increase in sustainable finance literacy. Similarly, we find that the brochure treatment increases the share of participants who claim to use ESG criteria in their investment decisions, for which around 12-14% can be directly attributed to an increase in sustainable finance literacy.

Second, we investigate the relationship between sustainability preferences and sustainable financial literacy and how both influence sustainable investment behavior. We show that sustainable financial literacy must be coupled with at least a moderate level of sustainability-oriented preferences to positively influence ESG investments. In the absence of moderate sustainability preferences, any additional increase in sustainable finance literacy is at minimum irrelevant, and we find some weak evidence that it might even reduce sustainable investments.

Finally, we show that for the participants who choose sustainable over conventional funds, an increase in sustainable finance knowledge increases the probability of investing in the more sustainable fund out of two ESG funds. In particular, we find that participants who possess the required knowledge to distinguish "dark green" (SFDR Article 9) funds from "light green" (SFDR Article 8) funds have an around 12% increased probability to choose the dark green over the light green fund. This increase is again moderated by the level of ESG preferences. In other words, without at least a moderate level of sustainability preferences, sustainable investment behavior is not influenced by knowledge.

We take several measures to ensure the robustness of our findings. During the experiment, we control for the possibility of an experimenter demand effect (EDE). EDE refers to a phenomenon in experiments in which the subjects form beliefs about the experimental objectives and adapt their actions in the direction most congruent to such objectives (Zizzo, 2009). Therefore, we divide the treatment group into two different random subgroups: "High EDE" and "Low EDE". Each subgroup gets to read different statements about our expectations with regard to their investment behavior, i.e., "we expect that participants in the experiment who read these instructions will be less (more) likely to invest in sustainable funds than they normally would" (de Quidt et al., 2018). In contrast, the control group does not get any particular statement. We run several tests and provide evidence that our results are not driven by an EDE. Furthermore, we perform a battery of robustness checks and show that our baseline results are robust to alternative model specifications.

Our paper complements the nascent literature on sustainable finance literacy. So far, only one recent study by Filippini et al. (2023) has investigated the relationship between sustainable finance literacy and financial decisions. Using survey data from Switzerland, their analysis shows that sustainable finance literacy is relatively rare, but it nevertheless has an important influence on whether people own sustainable finance products. We build on this study in three different ways: First, we conduct an experiment in which participants not only indicate whether they own sustainable assets, but also have to make active investment decisions from a given selection of conventional and sustainable funds. Second, we are the first to provide causal evidence for the effect of sustainable finance literacy on investment behavior by exogenously increasing the level of sustainable finance literacy of one group of participants (treatment) relative to another group (control). Third, we extend the set of questions proposed by Filippini et al. (2023) to measure sustainable finance literacy by formulating nine questions focusing on general (ESG criteria, sustainable investment strategies) and local (EU regulation) issues.

Our study also contributes to the growing stream of literature on the determinants of sustainable investing. Previous studies show that investors value sustainability and respond to corresponding information when making investment decisions (Hartzmark and Sussman, 2019; Ceccarelli et al., 2019). Several studies focus on institutional investors (e.g., pension funds) and discuss how such investors should align their investment practices with their clients' preferences (Bauer et al., 2021). However, the findings of such studies are not directly applicable to retail investors, who form a distinct subset of non-professional investors with distinct characteristics, motivations, and constraints.

Among prior studies that take the viewpoint of retail investors, Anderson and Robinson (2021) analyze Swedish households and find that households with stronger pro-environmental values do not necessarily hold greener portfolios. Briere and Ramelli (2021) observe that the offering of responsible investment options increases the propensity of left-wing and pro-social individual investors to invest in equity products due to a better alignment with their own personal values. Finally, Heeb et al. (2022) investigate the investment behavior of experienced private investors. They find that investors are willing to pay for sustainable investments, but that this willingness does not increase with the additional impact generated by such investments. Except for the aforementioned study by Filippini et al. (2023), these studies do not include an indicator of the knowledge of retail investors about sustainable finance products as an explanatory variable. Yet, as our results show, knowledge about sustainable financial products has a causal impact on financial decisions.

Another contribution of our paper is to provide and complement a comprehensive set of survey questions gathered from the existing literature on the determinants of ESG investments. Appendix V.6 provides an overview of these questions and shows which thematic modules we complement. While several studies have examined different determinants of socially responsible investments (SRI) individually, our work attempts to systematically analyze all of these dimensions in the context of one single experiment. This allows us to truly narrow down the specific effect that sustainable finance literacy plays in this context. More precisely, our study contains questions on risk, trust and time preferences developed by Falk et al. (2023), on financial literacy by Lusardi and Mitchell (2014), on financial experience by Gutsche and Zwergel (2020) and Anderson and Robinson (2021), on sustainable finance literacy by Filippini et al. (2023), on perceived impact by Heeb et al. (2022), on financial expectations with respect to ESG financial products by Riedl and Smeets (2017) and Bauer et al. (2021), and on environmental literacy by Anderson and Robinson (2021); Geiger and Holzhauer (2020); Zwickle and Jones (2018). All these factors provide complementary information and together contribute to a comprehensive understanding of sustainable investment decisions.

A final contribution is our proposed design of choice environment, which conveys a high degree of external validity to our experiment. There are some experiments studying sustainable investments in the laboratory (e.g., Barreda-Tarrazona et al. (2011), Bassen et al. (2018), Gutsche and Ziegler (2019), Heeb et al. (2022)), but these studies usually employ imaginary funds using stylized financial and sustainable features, which are designed for the purpose of the study. In contrast, our participants decide between funds that actually exist: We use screenshots of a real web interface of a large direct bank, and each fund conveys the information exactly as it is presented in the field. We also link bonus payments to the actual performance of these funds, including a time window of approximately half a year between the investment decision and the disbursement. While other experiments in prior literature have some of these features, to the best of our knowledge, our study is the only one that incorporates all of them. Thus, it simulates the actual decision with one of the highest degrees of credibility yet achieved in a laboratory or online experiment.

A key implication of our findings is that fostering sustainable choices and a "green" transformation goes beyond merely understanding investors' ESG preferences. In recent years, regulatory authorities have actively sought to channel capital flows towards green assets, for example by increasing disclosure and transparency of investment funds' ESG strategies. Such initiatives have been shown to have an impact on institutional investors (see e.g., Scherer and Hasaj (2023)). However, when considering retail investors, the success of such initiatives is dependent on investors' capacity to not only understand their preferences, but also actively translate these preferences into appropriate investment decisions, an aspect that cannot be assumed as a given. In terms of practical implications, our paper therefore highlights the need for educational initiatives and information campaigns on sustainable investments.

Understanding what kind of knowledge and preferences lead individuals to invest in certain ways is important not only to academics but also to investment professionals who invest on behalf of individuals. This is particularly true in the EU, where the revised Markets in Financial Instruments Directive (MiFID) II now mandates investment professionals to gather information about clients' sustainability preferences and integrate such preferences into the investment process. It is therefore becoming increasingly important for institutional investors to understand the sustainability preferences of their clients. At the same time, such sustainability preferences are likely to be influenced by these clients' knowledge and understanding of sustainable investment products, highlighting the importance of understanding the linkages between sustainable finance literacy, sustainability preferences and investment behavior among retail investors.

The rest of this paper is organized as follows. Section V.2 presents the theoretical background of our paper. Section V.3 describes the study design and provides descriptive statistics. Section V.4 presents the empirical results. Section V.5 discusses the limitations and implications of our paper, and Section V.6 concludes.

V.2 Literature background and brochure development

Sustainable investing is an investment approach that considers environmental, social and governance (ESG) criteria in portfolio selection and management (GSIA, 2021). Much of the recent literature explains the demand for sustainable investing as taste-based. Several theoretical models incorporate types of agents who derive utility from investing sustainably (Pástor et al., 2021; Pedersen et al., 2021; Oehmke and Opp, 2020), and a large body of empirical literature finds evidence for this family of explanations (e.g. Riedl and Smeets (2017); Hartzmark and Sussman (2019); Barber et al. (2021); Bauer et al. (2021); Bofinger et al. (2022); Heeb et al. (2022)). However, this literature usually focuses on establishing a link between preferences and demand, while rarely investigating how this link is mediated, or under which conditions this relationship holds. In particular, while the number of sustainable investment products has grown rapidly in recent years, the literature on whether and how such products are understood and perceived by retail investors remains limited. Compared to institutional investors, retail investors often have fewer resources and less expertise at their disposal. Therefore, it is crucial to shed light on how retail investors engage with such products, given their inherent complexity.

There is little literature that investigates the role of literacy in the context of sustainable investments, which seems surprising given the important role of financial literacy in financial decision-making as a whole (Lusardi and Mitchell, 2014). Aristei and Gallo (2021) and Bethlendi et al. (2022) investigate the influence of financial literacy on sustainable investing and find a positive relationship. Bethlendi et al. (2022) find a similar result for green, or environmental, literacy. In contrast, Anderson and Robinson (2021) and Filippini et al. (2023) do not find any influence of environmental or sustainability literacy.

However, from a theoretical point of view, it is not quite clear why financial literacy or environmental literacy, i.e., knowledge about concepts such as inflation, compound interest, the influence of carbon dioxide on the earth's climate, or the natural habitat of polar bears (see Anderson and Robinson (2021)), should influence the tendency to invest sustainably, other than via a correlation with some other aspects of sustainable investing, such as preferences or specific knowledge. This is why Filippini et al. (2023) develop the concept of sustainable finance literacy, which is tailored to this specific knowledge, and defined as the "knowledge of regulations, norms, and standards about financial products that have sustainable characteristics." The authors find that this special knowledge, while in general not widespread among individuals, nonetheless predicts the probability of sustainable investing in an observational study in Switzerland. To measure sustainable finance literacy, the authors develop a set of eight questions that cover several topics, including the definition of ESG, rules and certifications of ESG products, the difference between sustainability characteristics and ecology, and the difference between sustainable investing and impact investing.[3]

Our paper aims to provide evidence that there is a causal relationship between sustainable finance literacy and investment behavior. To this end, and in contrast to Filip-

[3] As their study focuses on Swiss investors, several of their questions are framed to fit the Swiss context.

pini et al. (2023), we use an experiment. Our treatment consists of a short educational brochure with key information for retail investors based on the definition developed by Filippini et al. (2023). Specifically, the brochure is organized around three central dimensions: (1) the definition and components of ESG criteria, (2) the various investment strategies incorporating these criteria, and (3) the EU regulation governing such investments. A copy of the brochure can be found in Appendix V.6.[4] Several examples cited in the brochure also follow a book written by the Stiftung Warentest, a foundation originally established by the German Bundestag with the aim of giving guidance to consumers by providing impartial and objective information (Stiftung Warentest, 2021).

Following Filippini et al. (2023)'s definition, which emphasizes the importance of norms and standards, the first part of our brochure explains the acronym "ESG" and addresses the specific components that fall under each of the three pillars, which together contribute to the assessment of the sustainability "profile" of a company or stock. The second part of our brochure highlights how such ESG criteria are applied in various investment strategies. Examples of these strategies include "negative screens," which deliberately avoid investing in certain stocks that do not meet pre-defined criteria. Moreover, the brochure elaborates on alternative strategies such as the "best-in-class" screening. Unlike negative screening, this approach seeks to invest in companies that are industry leaders in sustainability, irrespective of whether the industry itself is inherently "green" (Gutsche and Ziegler, 2019). Consequently, the second part of the brochure also helps explain why some investment portfolios classified as sustainable may still include stocks of companies in industries that are not necessarily inherently environmentally-friendly. This is important because a lack of knowledge about the investment strategies

[4]Please note that the brochure in the appendix is an English translation. The original document used in the experiment was in German.

underlying sustainable investments often leads to misconceptions about stock selection in these portfolios.

The third part of the brochure is dedicated to the regulation of such investment products in the EU. As highlighted by Filippini et al. (2023)'s definition of sustainable finance literacy, an understanding of regulations is crucial in the context of sustainable investing. This is particularly the case in the EU where the SFDR, which took effect in 2021, imposes a set of mandatory disclosure requirements on asset managers and other financial market participants.[5] An important aspect of this regulation is the classification of investment products according to three different categories: Article 6, Article 8 and Article 9 financial products (European Parliament and Council of the European Union, 2019). Each of these three categories comes with its own disclosure requirements, resulting in more ESG-related information for retail investors.

Article 6 funds do not have sustainable investment as their objective. Nevertheless, the incorporation of sustainability risks into investment decision-making and the impact of sustainability risks on the fund's returns must be described in the fund's pre-contractual disclosures (European Parliament and Council of the European Union, 2019). When a fund manager does not consider sustainability risks in the decision-making process, the disclosure should explain why, under the principle of "comply or explain". In contrast, Article 8 products (also referred to as "light green" funds) promote investments with environmental or social characteristics, or a combination of those characteristics, provided that the companies in which the investments are made follow good governance practices (European Parliament and Council of the European Union, 2019). While sustainable investment is not the primary objective of Article 8 products, it re-

[5]This regulation also applies to all US financial firms that market their financial products in the EU. Thus, US companies that sell to EU-based clients or are domiciled in the EU must also adhere to SFDR requirements for each fund.

mains an aspect of the investment process. Finally, Article 9 products (also referred to as "dark green" funds) have a sustainable investment as their objective. In this context, a sustainable investment means an investment in an economic activity that contributes to an environmental objective, as measured, for example, by key resource efficiency indicators and greenhouse gas emissions, or an investment in an economic activity that contributes to a social objective, such as tackling inequality (European Parliament and Council of the European Union, 2019). Such products must also comply with the "do no significant harm" principle by demonstrating that they do not in any way significantly harm any other important sustainability objectives. Furthermore, the investee companies also have to follow good governance practices with respect to management structures, employee relations, remuneration and tax compliance.

Taken together, the SFDR regulation results overall in an increase in ESG information available to retail investors, especially regarding the characteristics and strategies applied by these financial products.

In addition to our brochure, we create nine survey questions that aim to test and measure sustainable finance literacy. We divide these questions into two distinct categories, which we refer to as "global" and "local". Within the "global" category, we include five questions related to ESG considerations, as well as investment strategies that hold relevance across the globe. For instance, the incorporation of ESG criteria and the application of positive or negative screens for sustainable investments are practices embraced by investment firms worldwide, and are therefore not limited to only the EU. However, since regulations and norms about financial products differ between regulatory contexts, we argue, in line with Filippini et al. (2023) that sustainable finance literacy cannot be measured by relying only on questions that measure aspects that are identical across jurisdictions. Thus, we add questions centered around the regulatory context of the EU's

SFDR. Specifically, we include in the "local" category four questions on issues related to the SFDR's Articles 6, 8 and 9, which are specific to the EU context.

We hypothesize that increasing sustainable finance literacy has two effects: First, we expect that increasing sustainable finance literacy increases the probability of investing sustainably (H1), as such knowledge allows to translate sustainable preferences into action. Second, based on the aforementioned literature on ESG preferences (Riedl and Smeets, 2017; Hartzmark and Sussman, 2019; Barber et al., 2021; Bauer et al., 2021; Bofinger et al., 2022; Heeb et al., 2022), we only expect this effect if combined with a sufficiently high level of sustainable preferences (H2). If participants do not have any sustainable preferences, the family of preference explanations for sustainable investing predicts no effect of sustainable finance literacy.

V.3 Study design

V.3.1 Brochure treatment and experimenter demand effect

We address the question of the relationship between sustainable finance literacy and investment behavior using a preregistered experimental procedure.[6] The experiment was conducted in June 2023 with a sample of 1,000 participants recruited from the Prolific platform. To ensure the relevance and contextual validity of the results, the experiment was carried out specifically in German and targeted German residents within the Prolific platform.

The treatment in our experimental setting is a brochure containing information on ESG criteria, sustainable investment strategies and the SFDR regulation. Participants

[6] The experiment was preregistered with the American Economic Association (AEA). For preregistration details, see Auzepy et al. (2023), https://doi.org/10.1257/rct.11325-2.0.

are randomly assigned to either the treatment group, which gets to read this brochure, or the control group, which does not get to see the brochure. To ensure that the treatment group actually reads the brochure, the participants are required to remain on the appropriate page for at least three minutes before moving on to the next page of the experiment. The actual average time spent reading the brochure was 328 seconds, i.e., approximately 5.5 minutes. Our experimental procedure is displayed in Figure V.1.

To increase the internal validity of our experiment, we control for the possibility of experimenter demand effects (EDE) and socially desirable responses in several ways. An experimenter demand effect refers to a phenomenon in experimental research in which survey participants unintentionally modify their behavior or responses based on cues and expectations they perceive from the experimenter or the experimental setting (Zizzo, 2009). To account for the possibility that participants may change their behavior to conform to what they believe we expect in the study, we divide our treatment group into three subgroups. This is represented by the black triangle in Figure V.1.

The first pool is presented with a short introductory text prior to the investment decisions. The text reads as follows: "We expect that participants in the experiment who read these instructions will be less likely to invest in sustainable funds than they normally would." We refer to this subsample as the "Low EDE" treatment. The second pool receives the following sentence: "We expect that participants in the experiment who read these instructions will be more likely to invest in sustainable funds than they normally would." We call this subsample the "High EDE treatment". Finally, the last pool did not get to read any of these sentences.

By communicating these expectations, we aim to induce experimenter demand effects[7]. We test the presence of an EDE in several model specifications and robustness

[7]This procedure follows de Quidt et al. (2018). These instructions are not deceiving. In other words,

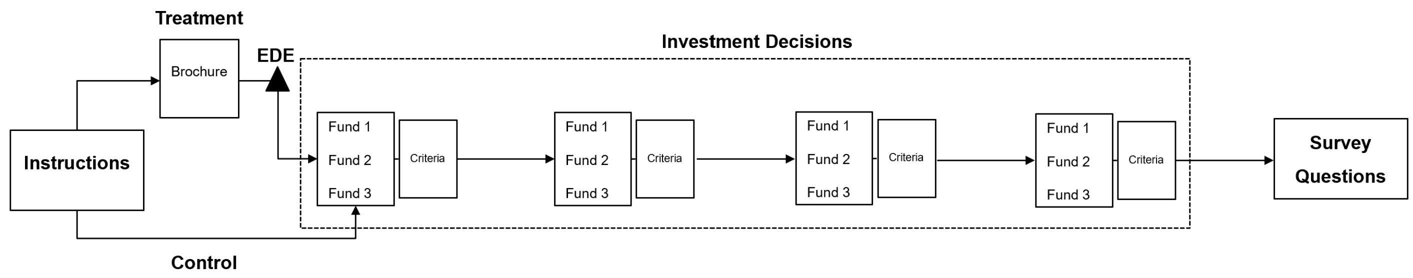


Figure V.1: This figure provides an illustration of the various steps of the experimental procedure we use in our experiment. Participants are randomly assigned either to the treatment group with the brochure or the control group without the brochure. The experiment contains four rounds of investment decisions. In each round, participants have to choose one of three funds and indicate which criteria were relevant to their investment decision. We test for the presence of an experiment demand effect (EDE).

checks. In addition, we perform a mediation analysis, as described in Section V.4.2, in order to isolate the effects of a change in sustainable finance literacy from other potential effects induced by the brochure, such as experimenter demand effects. This allows us to precisely disentangle the impact of the brochure on sustainable investment behavior via the sustainable finance literacy channel.

V.3.2 Investment decisions

Our experiment includes four rounds of investment decisions. In each round, the participants have to choose one out of three different funds or ETFs. These funds are real funds from actual asset management companies commonly known to German retail investors. Appendix V.6 provides an overview of the different funds used in our experiment.

The information provided for each fund is taken directly from an online account of ING, Germany's largest direct bank, and therefore reflects the information that a retail investor would typically access online. The details presented for each fund follow a standard format and include the fund's provider, its name, a performance chart showing the fund's performance over the past year, the issuing company, Morningstar's risk rating, whether it distributes or reinvests gains, the currency used, the fund's size, costs and ongoing charges, its major holdings, its exposure to different countries and industries. Additionally, the fund information includes the fund's SFDR article[8] and a textual description of the fund's investment strategy. Finally, participants have the option to download the fund's full fact sheet.

Based on the information provided, in each round the participants have to choose

based on an experimenter demand argument, we *truly* expect participants whom we tell to invest more sustainably to actually invest more sustainably, and vice versa for the Low EDE group.

[8] At the time of the experiment, out of the largest German online investing platforms, the ING platform was one of the few that explicitly displayed the fund's SFDR article. This was another reason why we chose this particular platform for our experiment.

a single fund to invest 200 Euro. In the first three decisions, participants are given a choice between two conventional investments and one sustainable investment offered by the same asset management company per decision round (taken from the asset management firms of the three banking groups with large customer bases in Germany). The funds are not explicitly labeled as sustainable or conventional, but with sufficient knowledge, it is possible to infer this from the information provided (e.g. the fund's name, the SFRD article, and the description of the investment strategy). In the final round, participants have the choice between two sustainable investment options (SFDR Article 8 and Article 9) and one conventional investment option (SFDR Article 6). To ensure that the sustainable investments are objectively more sustainable than the conventional funds, we reviewed their Morningstar Sustainability Rating, Carbon Risk Score, and share of fossil fuel companies prior to selecting them. Furthermore, we made sure that the sustainable funds that we selected did not exhibit strikingly more favorable risk-return profiles or cost attributes compared to the conventional funds in order to maintain fairly comparable sets of funds for each round of investment decisions.

We use these investment decisions as one of our dependent variables in several model specifications, measuring the likelihood that participants choose sustainable over conventional investments. A description of this dependent variable is provided in Appendix V.6.

After each investment round, participants are asked to indicate which of the displayed pieces of information about the funds played a role in their investment decision. The participants can select specific aspects from a list of pre-defined criteria that we provide, or write additional criteria in a text box. Our pre-defined list of criteria is based on standard information available for all funds. For example, we ask the participants whether they considered the fund's provider, the fund's name, its risk and return profile, past performance, size, top holdings, country exposure, industry exposure, and costs. We also ask whether sustainability-related information played a role in their decision. We use this information as a second dependent variable, which measures the conscious part in the decision-making process. Specifically, we measure the number of cases in which sustainability information was reported as one of the criteria for the investment decision after each investment round.

V.3.3 Incentives

In the initial phase, participants receive comprehensive instructions about the experiment, including information about their compensation. The compensation per participant is 4.50 pounds (about 5.20 euros). To increase data quality, compensation is only paid if participants answer two attention questions correctly. These attention questions are easy to identify, and we have provided clear instructions on how to answer them. If a participant answers both of these questions differently from the instructions, we reject the submission.[9]

In order to increase the chances of measuring actual investment behavior, we follow Heeb et al. (2022) as well as Bauer et al. (2022) and include a bonus payment, which every participant is also informed about before making the first investment decision. The bonus payment takes the form of a lottery. For 20 participants, we implement one randomly selected investment choice each. After half a year, we pay out the value of this investment to the selected participants. Since the payout is affected by both gains and losses, this makes the investment decisions more realistic and increases the stakes of the experiment.

[9]This happened in only 5 cases. These participants do not count for our goal of 1,000 participants and are excluded from each analysis.

The bonus calculation utilizes a simplified net return, representing the raw return earned by the fund minus the fund's ongoing costs for six months. For simplicity, other cost factors, such as performance fees and sales charges, are disregarded. For instance, if a selected fund achieved a 10.5% return by December 1, 2023, with ongoing expenses of 0.5%, the net return is 10%, resulting in a bonus payment of 220 euros. Conversely, in the case of a loss of 9.3% and running costs of 0.7%, the net return would be -10%, leading to a 180 euro bonus. In our experiment, the maximum payout is capped at 300 euros. Additionally, a floor is established, guaranteeing a minimum payout of 100 euros even if the investment's value is lower on the cut-off date.

V.3.4 Survey questions and control variables

To complete our dataset, we collect additional data on the survey participants. First, we collect standard demographic data on age, gender, years of education, and household income. Due to the linkages between political views and sustainability preferences found in previous literature, we ask the participants for which party they would vote in a hypothetical upcoming general election. Anderson and Robinson (2021) measure pro-environmental attitudes using Green Party voting records. Briere and Ramelli (2021) report that responsible stock funds provide incentives for left-leaning individuals to increase their stock market participation given that such funds are more in line with their personal values.

Second, we collect data on other variables that are also likely to influence investment decisions. Specifically, we collect data on individual preferences, financial literacy, financial experience, environmental literacy, sustainable finance literacy, perceived impact, and expectations regarding sustainable investment products. We made the deliberate choice to rely on a large set of questions that have already been used and validated in previous literature. All questions discussed in this section are grouped into thematic modules summarized in Appendix V.6. In addition, this appendix also provides a detailed description of each variable derived from these questions.

Regarding individual preferences, we focus on two types of preferences: economic preferences related to risk, time, trust and altruism, and sustainability preferences. We measure economic preferences using the experimentally validated survey module introduced by Falk et al. (2023) and previously employed in related literature (see e.g., Heeb et al. (2022)). In total, we use five questions to determine how risk-averse the participants are, how much they discount time by preferring present rewards to future ones, and how willing they are to trust and share with others. Each of these questions is on a 10-point scale. To further elicit intrinsic social preferences, the preference module uses the responder behavior in an ultimatum game.

In order to measure sustainability preferences, we use questions that are political in nature and involve implicit individual cost-benefit trade-offs. To this end, we select seven statements from the so-called Wahl-O-Mat, a publicly accessible online tool of the German Federal Agency for Civic Education ("Bundeszentrale für politische Bildung") that contains political statements from various political parties and is intended to help citizens understand how political parties align with their own preferences on various issues. We select a set of statements intended to measure environmental and social preferences in a German context. Participants can indicate how much they agree on a 5-point scale with statements about climate neutrality, the planned phase-out of coal-fired power generation, combustion engines, subsidies for organic farming, expansion of rail transportation, mandatory photovoltaic systems for new housing, and an increase in the minimum wage. Since the answers are not labeled with numbers, we code them as ranging from 0 to 4 for convenience. From these seven questions, we calculate the average and refer to this variable as the "ESG Pref Score".

To assess the financial literacy of the participants, we use the standard test developed by Lusardi and Mitchell (2014). Specifically, we employ their three core questions (often referred to as the "Big-3"), which assess the knowledge of interest rates, inflation, and portfolio diversification. Each question can be answered correctly or incorrectly. Following the literature (see e.g., Filippini et al. (2023)), we construct a financial literacy indicator by summing the correct answers given by the participants to each of the three questions. In addition, we ask them about their agreement with the statement developed by Riedl and Smeets (2017), "I often talk with other people about investments" to measure signaling effects. Furthermore, we try to capture the extent to which participants are financially active by measuring self-assessed investment experience. We also ask whether they make financial decisions for themselves or whether someone else does (Gutsche and Zwergel, 2020). Finally, to measure financial self-monitoring, we collect information on how often participants check their investment portfolio and in which financial products (e.g., stocks, savings accounts) they are or were invested (Anderson and Robinson, 2021; Gutsche and Zwergel, 2020).

As shown by Anderson and Robinson (2021) and Filippini et al. (2023), it is also important to account for the environmental literacy of the participants, as it differs from both sustainability preferences and financial literacy. Thus, we ask five questions designed to capture households' knowledge about climate change and the environmental costs of different consumption choices. To this end, we begin with a question on the definition of sustainable development and sustainable forestry, which was developed by Zwickle and Jones (2018) and adopted by Filippini et al. (2023). We also add a question on energy use related to heating or cooling homes, proposed by Anderson and Robinson (2021). Finally, we add questions about carbon footprints (Geiger and Holzhauer, 2020) and the rise in global temperatures. Each question has several answers, out of which only one is right. We sum up the number of correct answers.

We follow Riedl and Smeets (2017) and elicit return expectations and risk perceptions regarding sustainable investment products compared to conventional investment products. We ask the participants how they assess the returns of sustainable investments compared to conventional investments on a scale that ranges from "much lower" and "somewhat lower", to "same", "somewhat higher" and "much higher". We then ask the same question about the risk of sustainable investments compared to conventional investments.

In addition to risk and return expectations, we account for the perceived impact of certain investment decisions. As shown in previous literature (Heeb et al., 2022), positive emotions derived from choosing sustainable investments are also an important driver of sustainable investing. To capture the extent to which participants perceive their investments as making a meaningful contribution to addressing societal challenges, we ask them after each of the four investment decisions to rate their investment in terms of perceived impact on a scale from 0 ("no contribution") to 5 ("very positive contribution"). In a separate question, we ask the participants which of the following dimensions are important to them, in general, when investing: returns, risk, environment, social, and governance. The participants can provide a response ranging from "not important" to "very important" for each dimension.

Finally, we take into account the perceived skepticism towards sustainable investments and ask the participants whether they think that "sustainable financial products are just greenwashing". Respondents can give an answer ranging from "strongly disagree" to "strongly agree".

V.4 Empirical specification and results

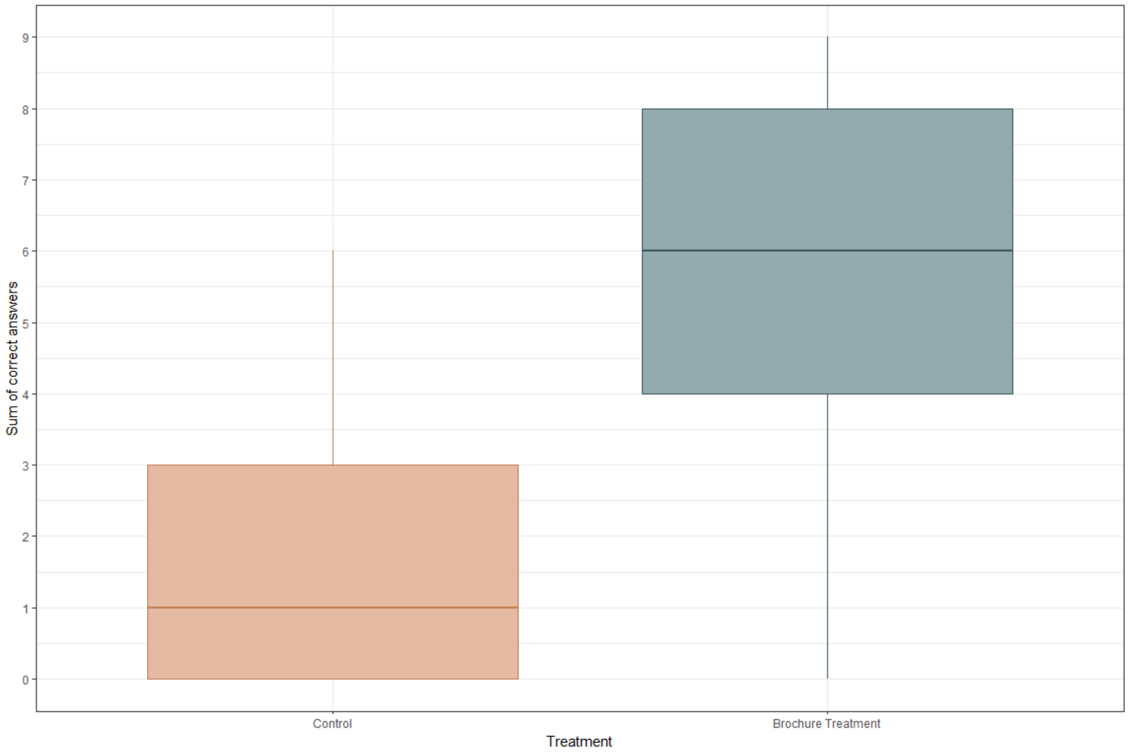
V.4.1 Does the brochure treatment increase sustainable finance literacy?

We start by analyzing whether the brochure is indeed successful in increasing sustainable finance literacy. As highlighted in Section V.2, and in contrast to financial literacy, there is no established procedure to measure sustainable finance literacy so far. As a result, we adopt Filippini et al. (2023)'s definition of sustainable finance literacy and develop a set of nine questions that address general ESG considerations as well as more specific considerations that relate primarily to the EU's SFDR regulation. An overview of these questions can be found in Appendix V.6.

For the treatment to be effective, we expect the treatment group, which gets to read the brochure, to answer significantly more questions correctly than the control group, which does not get to see the brochure. In order to test this, we use the nine questions referred to above and add up the number of correct answers per participant in a sum index. Figure V.2 shows that the median in the treatment group answers on average 6 out of 9 sustainable finance literacy questions correctly. In contrast, the median in the control group answers only 1 out of 9 questions correctly.

To determine whether the difference between the two groups is also statistically significant, we further investigate our results in an untabulated OLS regression analysis where we regress the sum index on the brochure treatment variable. The coefficient of the brochure treatment corresponds to 4 more correct answers, and is significant at the 1% level. In addition, the explanatory power of the brochure is high: The R^2 of this simple regression is 0.38. We conclude that the treatment effect is both statistically and economically significant and substantially increases sustainable finance literacy.

Figure V.2: This figure shows the results of the treatment and control groups in relation to our 9 questions on sustainable finance literacy. The correct answers are added to form a sum index, where 9 means that all 9 questions were answered correctly and 0 means that none of the questions was answered correctly.



Our analysis so far serves to show that the brochure treatment is effective and increases sustainable finance literacy in a significant way. This allows us to use the brochure treatment indicator as the main independent variable in the rest of our model specifications. To understand why this choice is most appropriate and why we should not resort to employing the sum index of sustainable finance literacy instead, we need to consider two distinct causes of heterogeneity in our participant group. First, participants are randomly assigned to either the treatment or the control group. As a result, the randomization determines for which participants we increase sustainable finance literacy. This creates a source of variation between the brochure treatment group and the control group which is typically referred to as "between variation". The second source of variation is the one that the participants *naturally* show, i.e., the differences that our participants display before they enter the experiment. This variation exists within each treatment group and is therefore referred to as the "within variation".

In our experiment, due to the randomization, the treatment and control groups are identical in expectation, i.e., the "within variation" is identical for both groups. What differs between the groups is the "between variation" induced by the treatment. As a result, by using the brochure treatment as the main independent variable in our analyses, we only use the "between variation" in sustainable finance literacy to explain differences in investment behavior. If we were to use the measured sustainable finance literacy instead, we would employ the total variation, which includes the within variation that we cannot control.

It should be noted that we also do not use sustainable finance literacy as a control variable in our analyses. We hypothesize that the treatment variable explains variation in sustainable investment behavior because it increases sustainable finance literacy, and sustainable finance literacy in turn leads to an increase in the probability of investing sustainably. Technically, this means that sustainable finance literacy is a mediator on the causal path from the treatment variable to the investment behavior variable. Thus, if we were to use sustainable finance literacy as a control variable in our model specifications, the brochure treatment variable would no longer capture the "between variation" in sustainable finance literacy. Instead, it would only capture all differences between the control and treatment groups except for the differences in sustainable finance literacy. This approach would therefore not test any of our hypotheses. As a result, we do not use the sum index of sustainable finance literacy as an independent variable or control variable in the following model specifications, but instead employ the brochure treatment variable.

V.4.2 Effects of sustainable finance literacy and preferences on investment decisions

Sample and descriptive statistics

Tables V.1, V.2 and V.3 present descriptive statistics for our sample, categorized by measurement scale (nominal plus ordinal; metric plus Likert >=10 scale points; and Likert =5 scale points, respectively). In this section, we discuss the key descriptive statistics extracted from these three tables.

Out of the survey participants who indicated their gender, 601 individuals identify as male, 378 as female, and 16 as non-binary. The median age of the respondents is 28 years, and their education level is 16 years, which is slightly higher than a high school diploma but lower than a fully completed bachelor's degree. As the experiment replicates investment decisions made online, using screenshots from a web interface of a large direct bank, our sample aligns with a younger demographic that is more likely to favor digital investment options over traditional banking advice.

The "Frequency Portfolio Checks" variable in Table V.1 indicates that the individuals in our sample exhibit a diverse range of financial monitoring behaviors. The majority of respondents engage in weekly portfolio checks (431), indicating a frequent and active interest in their financial situation. This is also in line with the "Talks often about Investments" variable in Table V.3, where 339 respondents selected "rather agree", indicating an inclination towards engaging in frequent discussions on the topic.

Table V.1: This table reports the demographics and sample proportions of several key survey questions. The total sample consists of 1000 survey participants. Total N contains all responses minus refused responses

Gender	Total N = 995
Female	378
Male	601
Non-binary	16
Frequency Portfolio Checks	**Total N = 999**
Weekly	431
Monthly	221
Several times per year, but less frequently than monthly	97
Once a year	11
More rarely	21
Never	5
Only when I create a deposit, or change it	9
I don’t have a financial deposit	204
Financial Decision Maker	**Total N = 998**
I decide for myself and/or my household alone	624
I do not decide but someone else does (e.g., partner, parents)	60
I decide together with my partner	314
Monthly Net Income	**Total N = 943**
Less than 500€	43
500€ to less than 1000€	115
1000€ to less than 2000€	154
2000€ to less than 3000€	212
3000€ to less than 4000€	153
4000€ to less than 5000€	119
5000€ to less than 6000€	65
6000€ to less than 7000€	48
7000€ or more	34
Party Preference	**Total N = 950**
CDU/CSU	60
SPD	123
Green Party	307
FDP	116
The Left	98
AfD	26
Other Party	100
Would not vote	57
I am not eligible to vote because I do not have German citizenship	63

Table V.2: This table reports the summary statistics for metric variables, aggregated indices, and Likert scales with >= 10 scale points. The total sample consists of 1000 survey participants. Total N contains all responses minus refused responses.

Variable	N	Min	Max	Median	Mean	Std. Dev.
Financial Literacy	998	0	3	3.00	2.67	0.63
Environmental Literacy	996	0	5	3.00	3.30	0.89
Perceived Impact	989	0	5	2.75	2.68	1.01
ESG Pref Score	993	0	4	3.00	2.91	0.72
Questionnaire Time (in sec.)	1000	251	4857	1247	1358	587
Data Usable	996	2	10	10.00	9.49	1.09
Age	988	18	72	28.00	29.86	8.44
Years of Education	996	8	23	16.00	14.70	2.97
Risk Preference	999	0	10	5.00	4.94	2.19
Time Preference	998	0	10	7.00	6.94	1.96
Trust	999	0	10	5.00	4.83	2.43
Social Preferences	997	0	10	7.00	6.58	2.03
Social Preferences, costly	998	0	10	6.00	5.70	2.24
Minimal Acceptance in UG	997	0	100	50.00	41.33	15.10
Financial Experience	1000	1	22	6.00	9.32	6.64

Table V.3: This table reports the number of answers for Likert scales with 5 scale points. The total sample consists of 1000 survey participants. Total N contains all responses minus refused responses.

Variable	Total N	much lower	somewhat lower	same	somewhat higher	much higher	I don't know
Expected Return	996	37	529	238	141	24	27
Expected Risk	997	24	191	386	324	51	21
Variable	Total N	strongly disagree	rather disagree	neither agree nor disagree	rather agree	strongly agree	
Greenwashing	999	72	303	410	177	37	
Talks often about Inv.	999	192	339	235	190	43	
Variable	Total N	not important	slightly important	moderately important	important	very important	
Importance Returns	998	10	31	148	441	368	
Importance Risk	982	1	67	221	420	273	
Importance E	982	94	224	301	279	94	
Importance S	969	95	199	302	278	95	
Importance G	997	63	189	298	299	148	

A substantial number prefers monthly checks (221), reflecting a somewhat less intensive approach. In contrast, a smaller portion of respondents opt for more infrequent checks, with 97 individuals doing so several times per year but less frequently than monthly, and only 11 respondents checking once a year. Moreover, a minimal number never engage in portfolio checks (5), and 204 respondents mentioned not having an investment portfolio.

The majority of respondents (624) make financial decisions independently ("Financial Decision Maker") or in conjunction with their partner (314). On the other hand, 60 respondents do not make financial decisions themselves but delegate this responsibility to someone else. The "Monthly Net Income" variable reports the income distribution among the respondents. Notably, the largest group of respondents falls into the income category of 2000€ to less than 3000€, comprising 212 individuals. In addition, the second largest group of respondents belongs to the adjacent income groups, with 154 individuals earning between 1000€ and less than 2000€ per month, and 153 individuals earning between 3000€ to less than 4000€. Lastly, the "Party Preference" variable provides insights into the political preferences of the respondents and the political diversity within the surveyed population. The data reveals a range of political affiliations, with the Green Party being the most popular choice (307), followed by the SPD (123), FDP (116), and The Left (98). Smaller numbers of respondents align with CDU/CSU (60), AfD (26), or other parties (100).

Table V.3 presents the summary statistics for several control variables with a Likert scale of five points. The table presents responses related to "Return expectations" and "Risk expectations" of ESG financial products as compared to conventional products. In the case of return expectations, the majority of participants (529) rated it as "somewhat lower", followed by 238 respondents who felt the returns were "the same". On the other hand, for risk expectations, a substantial portion (386) indicated that the risk was "the same", while 324 participants felt it was "somewhat higher". Regarding greenwashing behind ESG financial products, a notable number (410) chose the "neither agree nor disagree" option, while 303 respondents "rather disagreed" indicating that a majority of respondents do not necessarily associate ESG products with greenwashing.

Additionally, the table highlights respondents' perceptions of the importance of various factors, including returns, risk and ESG considerations when making investment decisions. Notably, for "Importance of Returns," a majority found it "important" (441) or even "very important" (368), indicating a strong emphasis on financial returns. In contrast, the "Importance of Risk" is somewhat weaker, with 420 participants saying risk is rather "important" and 273 participants considering it a "very important" dimension.

Interestingly, participants exhibited a more diverse range of opinions when assessing the importance of environmental, social, and governance factors. While 279 individuals indicated that environmental factors were "important", 301 considered them to be "moderately important", and 224 respondents felt they were only "slightly important". A similar pattern emerges with regard to the importance of social factors. Of the 969 respondents, 278 individuals rated social factors as "important", 302 considered them "moderately important" and 199 respondents found social factors to be only "slightly important". The importance of governance factors also drew varied responses. Interestingly, a large number of respondents (148) regarded governance factors as "very important". Relatively speaking, more respondents seemed to rate governance factors as "very important" compared to environmental and social factors. Furthermore, 299 participants rated governance factors as "important" and 298 as "moderately important".

Model specifications

Our empirical strategy proceeds in two steps. In the first step, we investigate whether a higher level of sustainable finance literacy, as indicated via the brochure treatment, leads to a higher probability of investing in a sustainable fund (H1), using two different dependent variables, both binary. Specifically, we estimate the following equations, using a logistic regression:

$$\text{Dependent Variable}_{i,p} = \beta_1 \text{Brochure Treatment}_p + \beta_2 \text{controls}_{i,p} + \alpha_p + \epsilon_{i,p} \tag{V.1}$$

where $Dependent\ Variable_{i,p}$ is either the indicator variable $Chose\ ESG_{i,p}$, which is equal to 1 if the decision i of participant p is to invest in a sustainable fund and 0 otherwise, or $Used\ Criterion_{i,p}$, an indicator variable which equals 1 if the participant p indicated the use of an ESG criterion in decision i and 0 otherwise. $Brochure\ Treatment_p$ is an indicator variable that equals 1 if the participant p received the brochure treatment, and 0 otherwise. In our regression results, the coefficient β_1 represents the variable of interest as it captures the effect of the brochure treatment on investment decisions. α_p is a random intercept for each participant p, which accounts for the fact that the decisions are clustered at the participant level, and $\epsilon_{i,p}$ is the error term. Furthermore, $controls_{i,p}$ is an optional vector of additional control variables, depending on the complexity of the model.

For each dependent variable, we run three model types with different degrees of complexity: In the *simple* model type, we do not include any control variable at all. This model type measures the net effect of the brochure itself on the dependent variables. In the *complex* model type, we include all control variables as outlined in Section V.3.4. In an experiment, the main role of control variables, aside from reducing standard errors by controlling for potential randomization failures, is to account for alternative mediators, i.e., other causal channels by which the brochure treatment might influence sustainable investing, other than through sustainable finance literacy. The advantage of the complex model type is therefore to deliver the most precise effect of sustainable finance we can measure, given all of our control variables. However, the complex model type appears to often overfit the data, as indicated by singularity problems (Bates et al., 2015, 2018). A common solution for that problem is to develop a reduced model (Matuschek et al., 2017).

Table V.4: This table shows the correlation coefficients of each variable with the treatment variable, sorted by p-value

Variable	Correlation	P-Value
Pearson correlation coefficients of "Treatment" with numeric and nominal variables		
Perceived Impact	0.23	0.000***
Importance S	0.07	0.028*
Data usable	-0.07	0.038*
Trust	0.06	0.042*
Party Preference NA	0.06	0.072
Greenwashing	-0.05	0.129
Social Preferences	0.04	0.154
Importance E	0.04	0.163
Gender Non Binary	0.04	0.187
Importance Risk	-0.04	0.194
Importance Returns	-0.04	0.226
Gender Female	0.03	0.291
Party not eligible	-0.03	0.291
Party Greens	0.03	0.301
Time Preference	0.03	0.319
Party FDP	-0.03	0.375
Ln UG min. demand	-0.03	0.379
Importance G	-0.03	0.403
Party none	0.03	0.406
Party AfD	-0.03	0.416
ESG Pref Score	0.03	0.419
Risk Preference	0.02	0.514
Ln Interview Time	0.02	0.529
Talks often about Inv.	-0.02	0.578
Party The Left	0.02	0.629
Financial Literacy	0.01	0.673
Party other	0.01	0.709
Gender NA	0.01	0.724
Ln Age	0.01	0.73
Financial Experience	-0.01	0.752
Social Preferences, costly	0.01	0.851
Years of education	-0.00	0.949
Environmental Literacy	-0.00	0.96
Party SPD	-0.00	0.987
Spearman correlation coefficients of treatment with ordinal variables		
Return expectations of ESG Funds	-0.03	0.284
Monthly Net Income	-0.02	0.44
Depot Check Count	-0.02	0.563
Risk expectations of ESG Funds	-0.00	0.993
Note:		$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

We account for this with the *medium* model type, where we strive for a balance between controlling for the most important potential alternative mediators, while also keeping the model as simple as possible. Thus, this model only includes control variables that significantly correlate with the treatment variable, as shown in Table V.4. These variables are "perceived impact", "importance S", "data usable" and "trust". The medium model type is simple enough not to cause overfitting, at the cost of potentially overlooking more complex mediations.

In the second step, we examine the various effects of sustainable finance knowledge combined with sustainability preferences on investment behavior (H2). We argue that the effect of the brochure depends on the level of ESG preferences. We mirror the analysis for H1, but now include an interaction term between the treatment and the ESG pref score. Specifically, we estimate the following model based on a logistic regression:

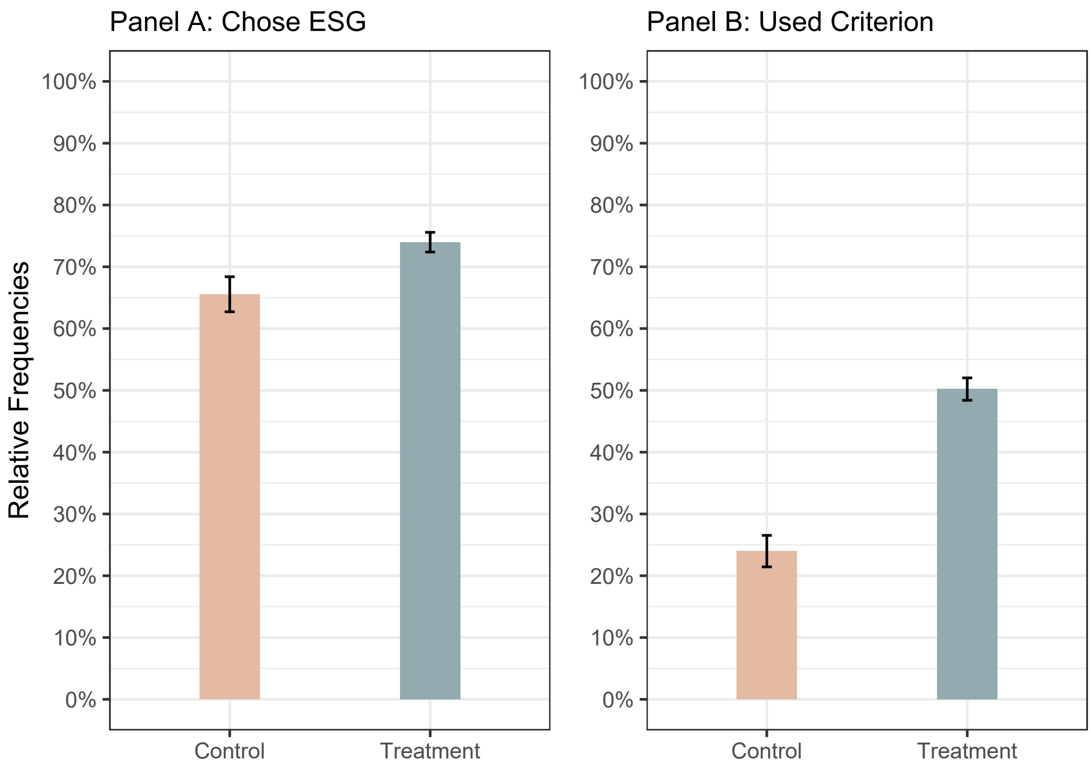
$$\text{Dependent Variable}_{i,p} = \beta_1 \text{Brochure Treatment}_p * \beta_2 \text{ESG\ \ Pref Score}_p + \beta_3 \text{controls}_{i,p} + \alpha_p + \epsilon_{i,p} \tag{V.2}$$

All the variables and model types (simple, medium, complex) stay the same, and in addition the ESG Pref Score$_p$ is the average answer from participant p for the seven ESG preference questions. These questions are five-point Likert scales, but the labels for the points do not include any numeric values. Thus, we scale the variable as a number between 0 and 4, which conveniently gives the coefficient β_1 for the brochure treatment in the regression model a meaningful interpretation: It is the effect of the brochure for the participants with the lowest sustainability preferences.

Does an increase in sustainable finance literacy lead to an increase in the probability of investing sustainably (H1)?

As a first step, we hypothesize that the brochure treatment leads to a higher probability of investing in sustainable funds and to base investment decisions on ESG-related information. Figure V.3 illustrates the results, showing bar plots for both dependent variables, split by treatment condition.

Figure V.3: This figure illustrates the results of H1, showing bar plots for both dependent variables ("Chose ESG" and "Used Criterion"), split by treatment condition. Panel A shows the relative frequencies of sustainable investment decisions for the control and treatment groups. Panel B reports the relative frequencies of participants in both the control and treatment groups who reported taking ESG criteria into account in their investment decisions.



Panel A shows the relative frequencies of sustainable investment decisions for the control and treatment groups. As can be seen, sustainable investment decisions, i.e., the choice of a sustainable fund in a specific investment round from the available fund selection, account for about 65% of the total number of investment decisions made by the control group. In contrast, sustainable investment decisions account for approximately

74% of total investment decisions made by the treatment group, which represents an increase of 9 percentage points compared to the control group.

In Panel B we show the relative frequencies of participants in both the control and treatment groups who reported taking ESG criteria into account in their investment decisions. In particular, in the control group, ESG criteria played a role in about 25% of investment decisions. In stark contrast, the brochure treatment group had a significantly higher usage rate, with ESG criteria used in about 50% of their investment decisions. Thus, the stated use of an ESG criterion roughly doubles from the control treatment to the brochure treatment.

Comparing Panel A and Panel B also indicates that participants often pick sustainable funds even though they do not explicitly state using ESG criteria. This is particularly true for the control group. The control group is less likely to show a conscious tendency to select sustainable funds based on ESG criteria. In contrast, the brochure group appears to make more decisions in favor of sustainable investments and tends to base its decisions more consciously on corresponding ESG information. Overall, this indicates that the control group relies less on ESG information than the brochure group.

The regression models confirm these results. Table V.5 reports the logit coefficients and margins (average marginal effects, i.e., the average effect of the brochure, given that the effect of the brochure for a given decision is nonlinear and also depends on the control variables) of six regressions with a random intercept. We present the results for the dependent variables "Chose ESG" in columns (1) to (3), and for "Used Criterion" in columns (4) to (6).

Table V.5: This table presents the results for Hypothesis 1, and reports the logits and margins of mixed models regressions with a random intercept on the decision level, using two different dependent variables as the measure for ESG Investment decision. "Chose ESG" is a Dummy that captures whether a sustainable fund was chosen or not. This measure does not differentiate between Article 8 funds ("light green") and Article 9 funds ("dark green"). "Used Criterion" is a dummy that captures whether a participant reported to have used any ESG criterion for their decision. "Simple" models do not include any control variables. "Medium" models only include control variables that significantly correlate with the Brochure Treatment variable. "Complex" models include all control variables.

	Chose ESG, simple	Chose ESG, medium	Chose ESG, complex	Used Criterion, simple	Used Criterion, medium	Used Criterion, complex
	(1)	(2)	(3)	(4)	(5)	(6)
Logits						
Brochure Treatment	0.480***	0.263*	0.248*	2.134***	1.511***	1.436***
	(0.106)	(0.104)	(0.106)	(0.217)	(0.194)	(0.189)
Intercept	0.757***	−1.143*	−2.719*	−2.116***	−8.125***	−11.995***
	(0.090)	(0.458)	(1.289)	(0.191)	(0.837)	(2.228)
Margins						
Brochure Treatment	0.093***	0.049*	0.043*	0.306***	0.214***	0.198***
	(0.021)	(0.020)	(0.019)	(0.028)	(0.026)	(0.025)
Controls	none	correlated	all	none	correlated	all
N	3993	3934	3438	4000	3936	3440
R^2 marg.	0.01	0.06	0.19	0.10	0.28	0.43
R^2 cond.	0.21	0.21	0.24	0.65	0.66	0.67
Note:						*p<0.05; **p<0.01; ***p<0.001

The regression results in columns (1) to (3) are both statistically and economically significant. In our simple model without control variables, the margins imply that the brochure treatment leads to an increase in the probability of choosing a sustainable over a conventional fund by around 9%. The coefficient for Brochure Treatment is significant at the 0.1% level. The results also hold for the medium model, with selected control variables, and the complex model, with all control variables. The brochure treatment variable loads positively on choosing an ESG fund, and is in both models significant at the 5% level, with an effect size of around 4 to 5% in both models. Hence, while the brochure treatment does not seem to be the main driver of sustainable investment behavior, it nevertheless represents an important factor to consider for investment decisions.

In columns (4) to (6) we explore the extent to which the brochure treatment leads to the use of ESG information more consciously in investment decisions. The average marginal effects vary from around 31% in the simple model to around 20% in the medium and complex model. All effects are significant at the 0.1% level. Again, these effects are substantially larger than the actual behavior effects.

To summarize the results for H1, we find that the brochure has a positive effect on both sustainable investment and on taking ESG criteria into account for financial decisions. We explore these results further using a mediation analysis in the following section.

Mediation Analysis: What is the effect of the brochure on investment behavior through sustainable finance literacy?

The results so far show that the brochure affects sustainable investments. However, the brochure's effect sizes drop substantially for both our dependent variables once we include control variables. This indicates that the brochure's effects are not only driven by an increase in sustainable literacy. We control for all observed variables, but the brochure may have some other unobserved effects beyond merely increasing sustainable finance literacy.

As an illustration, the brochure could trigger mental associations related to ESG, including prior knowledge, attitudes or expectations. Thus, it is possible that the brochure induces a so-called "priming" effect, i.e., it could simply increase the level of attention paid to ESG criteria among participants in the treatment group, without a similar priming effect in the control group. Consequently, this increased attention to ESG might also lead to a higher tendency to engage in sustainable investing – which is not directly caused by an increased sustainable finance literacy. A similar argument could be made for the EDE, where the mere display of the brochure could be indicative of our research hypothesis, motivating participants in the treatment group to invest more sustainably, while the control group has no additional motivation to do so. These arguments imply that the margins of the four medium and complex models of Table V.5 should best be interpreted as an upper limit for the isolated effects of sustainable finance literacy. At worst, sustainable finance literacy could have no effect at all.

To exclude this hypothesis and measure the effect of the brochure on sustainable investment behavior *only* via sustainable finance literacy, we conduct a causal mediation analysis following the approach of Baron and Kenny (1986a), Imai et al. (2011) and Acharya et al. (2016), and recently used in research related to financial literacy (see e.g., Carpena and Zia (2020)) and ESG (e.g., Zhou et al. (2022)).[10] This approach is based on the idea that the total effect of an independent variable is composed of several channels, i.e. the causal chain between the independent and the dependent variable incorporates some intermediary variables, which are called mediator variables. The conceptually simplest way to decompose the total effect is to divide it into two sub-effects. These are the "indirect" effect, which quantifies the extent to which a treatment influences an outcome through a specific mediating variable of interest, in our case sustainable finance literacy, and the "direct" effect, which is the aggregate of any other possible mediator, including unobserved variables. The indirect effect is usually operationalized as the average causal mediation effect (ACME). Thus, we focus in this analysis on the ACME of sustainable finance literacy.

Table V.6 shows the results, for both the nonlinear models from table V.5 and for linear models which we include as a robustness check, for both model types (medium and complex) and for both dependent variables.[11] For the "Chose ESG" variable, each

[10]We use the R package "mediation" (Tingley et al., 2014).

[11]We conduct the mediation analysis for the medium and complex models since we already know from the analysis of H1 that the effect from the simple model drops to roughly half after controlling for other

total effect of the brochure is in line with the results from the earlier analyses, with estimates ranging from 4.3% to 5.6%. The ACME of sustainable finance literacy in the complex models is around 4 to 5%, which also confirms our initial results.

Table V.6: This table shows the estimations for total effects and average causal mediation effects (ACME) for both dependent variables with sustainable finance literacy as the mediator, varied by which control variables and which regression formulas are used. Complex models include all control variables, medium models include Perceived Impact, Importance S, Trust, and Use Data. Linear models use mixed effects linear models on both stages. Nonlinear models use Poisson regressions for the mediator and logit regressions for the dependent variables.

Dependent variable	Model type	Total effect	ACME sufili
Chose ESG	medium nonlinear	0.056***	0.101***
Chose ESG	medium linear	0.046*	0.097***
Chose ESG	complex noinlinear	0.047**	0.049***
Chose ESG	complex linear	0.043*	0.044***
Used Criterion	medium nonlinear	0.217***	0.207***
Used Criterion	medium linear	0.188***	0.179***
Used Criterion	complex nonlinear	0.197***	0.143***
Used Criterion	complex linear	0.177***	0.121***

For the medium models, the analysis actually suggests that the effect of sustainable finance literacy is larger, at around 10%. This implies that the net effect of all alternative mediators combined would actually be negative, meaning that we underestimate the effect of sustainable finance literacy. We do not follow this interpretation since the medium models in the table might not incorporate some relevant effects while the complex models do. However, this actually provides more evidence that the effect size of the brochure via sustainable finance literacy is around 4 to 5% for the "Chose ESG" variable, and that alternative uncontrolled mediators such as increased attention or an experimenter demand effect cannot explain this finding away.

Next, we turn to the "Used Criterion" variable. The four models estimate the total ef-

variables, indicating alternative mediators.

fects to be around 17.7% to 21.7%, which is in line with the earlier results. The ACME estimates in the medium models are around the same size as the total effects, which implies that the brochure's effect after controlling for observables is purely driven by sustainable finance literacy. The complex models, however, suggest that the ACME of sustainable finance literacy is smaller than the total effect. For these models, the ACMEs vary around 12.1% to 14.3%, which is around two-thirds of the total effects of the respective models. To err on the conservative side, we again champion the interpretation from the complex models. It suggests that while the brochure's effect on the decision criteria is in part due to an increase in sustainable finance literacy, other mediators, such as priming, play a role as well. The channel through sustainable finance literacy still seems to be the most relevant, as it accounts for roughly two-thirds of the total effect both in the linear and nonlinear models.

In sum, the results from the mediation analysis provide additional support for our initial results discussed in Section V.4.2 and confirm that the effect size of the brochure via sustainable finance literacy is around 4 to 5% for actual behavior. However, we do see that the brochure has some unmeasured influences beyond sustainable finance literacy regarding the conscious usage of ESG criteria.

Does an increase in sustainable finance literacy, combined with high ESG preferences, lead to an increase in the probability of investing sustainably (H2)?

Does the brochure work for all the participants in the same way, and what role do ESG preferences play in this context? A person who is knowledgeable about ESG and sustainable investing but has no strong environmental and/or social preferences could make a conscious decision *not* to invest in sustainable finance products. Conversely, a person with strong environmental and/or social preferences but insufficient knowledge of sustainable investing may have difficulty effectively translating those preferences into actionable investment decisions. As a next step, we therefore explore the role of both sustainability preferences and sustainable finance literacy in shaping investment decisions. In other words, we focus on ESG preferences and interact such preferences with the brochure treatment.

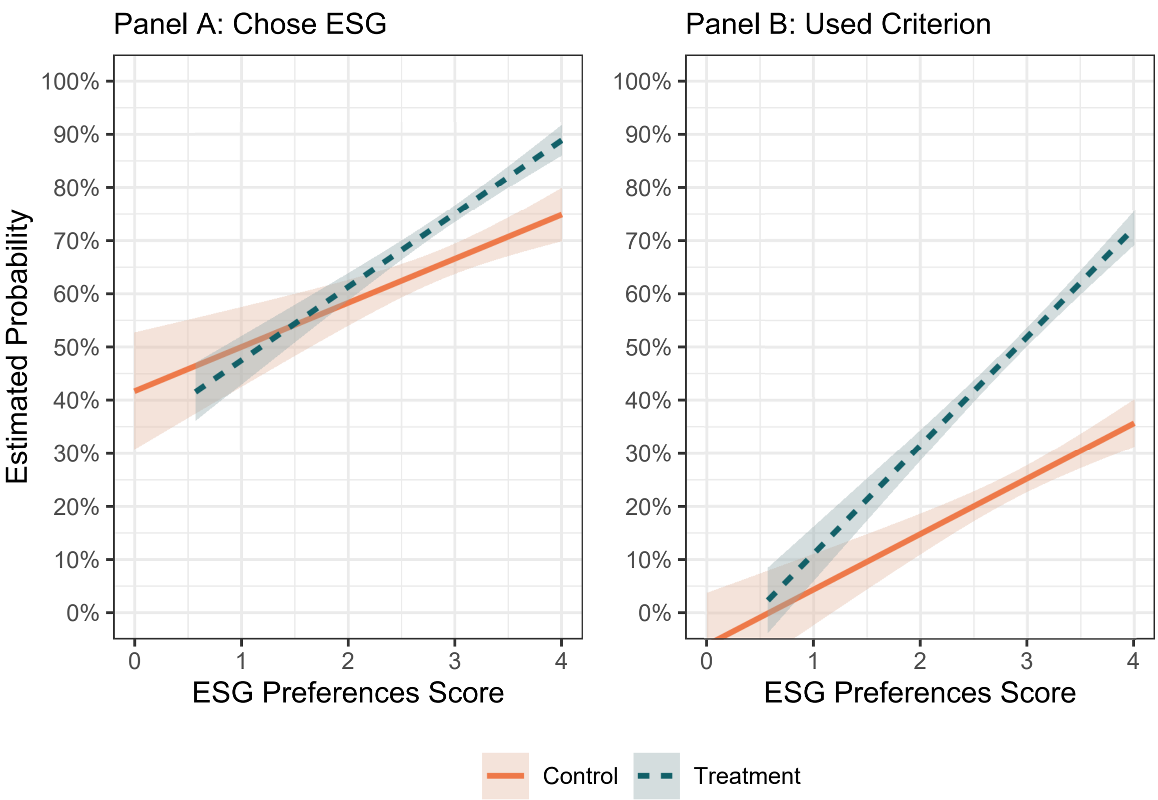
Figure V.4 is an interaction plot that illustrates the results. It plots the relationship between ESG preferences and the two dependent variables "Chose ESG" and "Used Criterion" for each of the two experimental groups. Our second hypothesis (H2) implies that for these relationships, the slope for the brochure treatment should be steeper than for the control group. We find this result for both dependent variables. Panel A shows that the probability of investing sustainably increases with ESG preferences. The slope is steeper for the brochure treatment group compared to the control group, and both groups start to differ significantly as the ESG preferences score increases. Panel B shows that this pattern also holds for the incorporation of an ESG criterion into the decision-making process, and the differences become significant at a slightly lower level of ESG preferences.

Table V.7 corroborates these findings, primarily for the dependent variable "Chose ESG". The table reports the logits[12] for each of the six models, which again are combinations of the two different dependent variables and the three model types. Columns (1) to (3) show the results for the actual behavior as a dependent variable. In each of the models, the interaction term is significant on the 1% level, with the predicted sign. However, for the "Used Criterion" variable, in columns (4) to (6), we find the predicted signs, but only the coefficient in the complex model is significant at the 5% level. We

[12]Unlike for H1, we do not report margins (AMEs) for H2 because we are interested in the interaction term. In this case, margins cannot be determined (Williams, 2012). We can, however, use the logits to infer the interaction term's statistical significance.

attribute this to the relatively large standard errors, which are approximately twice as large as for the "Chose ESG" variable. Thus, the "Used Criterion" variable appears to be noisier.

Figure V.4: This figure is an interaction plot illustrating the results of H2. It plots the relationships between ESG preferences and the two dependent variables "Chose ESG" and "Used Criterion" for the control and treatment groups. Panel A shows the relationship between the probability of investing sustainably and ESG preferences. Panel B plots the relationship between the incorporation of ESG criteria into investment decisions and ESG preferences.



Interestingly, we find in columns (1) to (3) a negative effect of the treatment variable, which is even significant at the 5% level in the medium model in column (2). Since we deliberately mapped the ESG preferences score on a scale from 0 to 4, this coefficient represents the behavior of the participants with the lowest ESG preferences. Therefore, it might even be argued that, for individuals with low ESG preferences, the brochure reduces sustainable investments. This seems reasonable because if sufficient sustainable finance literacy makes it possible to identify sustainable funds, this very literacy combined with low preferences might help such individuals to actively avoid sustainable funds. In addition, it could also be that individuals who have a negative view of ESG issues in general may have a negative reaction to the brochure treatment. This could be related not only to anti-ESG sentiment, but also to the perception of sustainable financial products as a form of greenwashing.

However, further research would be needed to underpin this finding. Admittedly, we do not find this pattern in the case of the other dependent variable, which should be the case if the decision to avoid sustainable funds were a conscious one. Furthermore, the results are based on relatively few observations. Only 125 participants in both groups combined have an ESG preference score of 2 or less, and only 1 participant has a value of 0. Nevertheless, these results might point to a more nuanced understanding of sustainable finance literacy to be explored in later studies.

In sum, we conclude that we find strong evidence for H2 for the behavior and weaker, more mixed evidence for the conscious use of ESG criteria. Specifically, a higher level of sustainable finance literacy, combined with high ESG preferences, leads to a higher probability of choosing an ESG fund. This suggests that sustainable finance literacy helps individuals to better align their preferences with their investment decisions. However, the effectiveness of the brochure seems limited among individuals who have low ESG preferences and could potentially have unintended negative effects in some cases.

Table V.7: This table presents the results for Hypothesis 2, and reports the logits of mixed models regressions with a random intercept on the decision level, using two different dependent variables as the measure for ESG Investment decision. "Chose ESG" is a Dummy that captures whether a sustainable fund was chosen or not. This measure does not differentiate between Article 8 funds ("light green") and Article 9 funds ("dark green"). "Used Criterion" is a dummy that captures whether a participant reported to have used any ESG criterion for their decision. "Simple" models do not include any control variables. "Medium" models only include control variables that correlate with the Brochure Treatment variable. "Complex" models include all control variables. The relevant variable is the interaction term. The ESG Pref Score is not centered.

	Chose ESG, simple	Chose ESG, medium	Chose ESG, complex	Used Criterion, simple	Used Criterion, medium	Used Criterion, complex
	(1)	(2)	(3)	(4)	(5)	(6)
Brochure Treatment	−0.644	−0.827*	−0.743	0.760	0.095	−0.209
	(0.396)	(0.387)	(0.392)	(0.860)	(0.826)	(0.797)
ESG Pref Score	0.412***	0.299**	−0.032	1.135***	0.688**	−0.030
	(0.110)	(0.110)	(0.120)	(0.244)	(0.238)	(0.244)
Brochure * ESG Pref Score	0.387**	0.385**	0.350**	0.452	0.487	0.550*
	(0.135)	(0.132)	(0.133)	(0.283)	(0.270)	(0.260)
Intercept	−0.437	−1.450**	−2.108	−5.374***	−9.139***	−10.960***
	(0.323)	(0.534)	(1.310)	(0.756)	(1.060)	(2.308)
Controls	none	correlated	all	none	correlated	all
N	3969	3910	3438	3972	3912	3440
R^2 marg.	0.07	0.10	0.19	0.20	0.33	0.43
R^2 cond.	0.22	0.22	0.25	0.66	0.67	0.67

Note: *p<0.05; **p<0.01; ***p<0.001

V.4.3 Robustness checks

We conduct an array of additional tests to check the robustness of our baseline results. First, we estimate the nonlinear models with a probit link function ("probits") instead of logits. Second, we also run a linear model ("LPM"), which we not only use as a robustness check, but also as a second method to estimate effect sizes. Third, we restrict the sample to participants who gave a score of at least 5 out of 10 to the statement "I have given my answers and made my decisions carefully and to the best of my knowledge, and therefore think that my data should be used for the study" ("Use data 5"). We also run a robustness check excluding the fastest and slowest 2.5% of the participants ("Time 95%"). Finally, we check for the presence of an EDE, as described in Section V.3.1.

Table V.8 shows the results of the robustness checks for H1. In the probit models, the effect of the brochure treatment remains consistent with the baseline results. The treatment has a positive and statistically significant effect on both "Chose ESG" and "Used Criterion" across all levels of model complexity (simple, medium, complex) in columns (1) to (6). The LPM model results also show a consistent positive effect of the brochure treatment on ESG investment decisions: The treatment is statistically significant and positively associated with "Chose ESG" and "Used Criterion" across all columns, and the coefficients have a similar, but slightly smaller size as the margins in Table V.5. When restricting the sample to participants who gave a score of at least 5 out of 10 for their data use statement, the positive effect of the brochure treatment on ESG decisions remains robust. Similarly, excluding the fastest and slowest 2.5% of participants from the sample does not substantially alter the results.

Table V.9 reports the results of the robustness checks for H2, where we interact the brochure treatment with the sustainability preferences. The results from the probit and LPM models generally confirm the direction of effects observed in Table V.7, although there are differences in the magnitude and statistical significance of some coefficients. Specifically, for the probit models in columns (1) to (3), the interaction term retains its significance ($p < 0.01$) and remains consistent with the main results. For the LPM, the coefficients are generally smaller but remain statistically significant, except in column (3). The results also hold in the robustness checks "Use data 5" and "Time 95%". Thus, all four types of robustness checks (probits, LPM, Use data 5 and Time 95%) provide strong and consistent support for the results of H2 with "Chose ESG" as the dependent variable.

Table V.8: This table summarises the margins for each robustness check for Hypothesis 1. We estimate the nonlinear models with a probit link function ("Probits"). We also run a linear model ("LPM"). We restrict the sample to participants who gave a score of at least 5 out of 10 to the statement "I have given my answers and made my decisions carefully and to the best of my knowledge, and therefore think that my data should be used for the study" ("Use data 5"). We also run a robustness check excluding the fastest and slowest 2.5% of the participants ("Time 95%"). Finally, we check for the presence of an experimenter demand effect (EDE). "Chose ESG" is a dummy that captures whether a sustainable fund was chosen or not. "Used Criterion" is a dummy that captures whether a participant reported to have used any ESG criterion for their decision. "Simple" models do not include any control variables. "Medium" models only include control variables that correlate with the Brochure Treatment variable. "Complex" models include all control variables.

	Dependent Variable, Model:					
Robustness Check, Variables	Chose ESG, simple	Chose ESG, medium	Chose ESG, complex	Used Criterion, simple	Used Criterion, medium	Used Criterion, complex
	(1)	(2)	(3)	(4)	(5)	(6)
Probits						
Brochure Treatment	0.094***	0.051*	0.046*	0.307***	0.216***	0.198***
	(0.021)	(0.020)	(0.019)	(0.027)	(0.026)	(0.024)
LPM						
Brochure Treatment	0.084***	0.045*	0.042*	0.262***	0.190***	0.181***
	(0.019)	(0.019)	(0.019)	(0.026)	(0.025)	(0.025)
Use data 5						
Brochure Treatment	0.095***	0.052*	0.046*	0.314***	0.220***	0.202***
	(0.021)	(0.020)	(0.019)	(0.028)	(0.026)	(0.025)
Time 95%						
Brochure Treatment	0.100***	0.057**	0.054**	0.309***	0.210***	0.200***
	(0.022)	(0.020)	(0.019)	(0.029)	(0.027)	(0.026)
EDE						
High EDE-Treatment	−0.017	−0.025	−0.012	0.054	0.046	0.062*
	(0.025)	(0.024)	(0.023)	(0.037)	(0.034)	(0.032)
Low EDE-Treatment	−0.036	−0.022	−0.010	−0.033	0.000	0.021
	(0.026)	(0.024)	(0.022)	(0.037)	(0.034)	(0.032)
Note:						*p<0.05; **p<0.01; ***p<0.001

Turning to columns (4) to (6) with "Used Criterion" as the dependent variable, the results are more contrasted. The coefficient of the interaction term keeps its significance ($p<0.05$) in the complex model in column (6) for the probit models, and even becomes strongly significant ($p<0.001$) across all three columns in the LPM models. The results hold consistently in the "Time 95%" robustness check, but not entirely when considering the "Use data 5" check where the statistical coefficient remains in the complex model in Column (3) and disappears in the others. Overall, the results of the four robustness checks confirm the main results and even provide some evidence that the main analyses for H2 with "Used Criterion" as a dependent variable might underestimate the significance of the interaction term. We find significant results for models (4) and (5) in two out of four robustness checks. To err on the conservative side, however, we conclude that the evidence for H2 with "Used criterion" as the dependent variable is weaker than that for the "Chose ESG".

In both Table V.8 and Table V.9, we find very little and inconclusive evidence for the presence of an EDE. Most coefficients are insignificant, and they often have the wrong sign. For example, the "High EDE" coefficients in Table V.8 should be positive because a stronger EDE should increase the probability to invest sustainably. Instead, the coefficients for the "Chose ESG" variable are all negative. Furthermore, we find that the interaction term in Table V.9 is significantly larger for the "Low EDE" group compared to the treatment group without any EDE manipulation. If there was an EDE, this coefficient should actually be smaller. The only result that speaks in favor of an EDE is the significant coefficient for the complex "Used Criterion" model in Table V.8. Therefore, we conclude that there is very little and weak evidence for the presence of an EDE affecting our baseline results.

Table V.9: This table summarises the margins for each robustness check for Hypothesis 2. We estimate the nonlinear models with a probit link function ("Probits"). We also run a linear model ("LPM"). We restrict the sample to participants who gave a score of at least 5 out of 10 to the statement "I have given my answers and made my decisions carefully and to the best of my knowledge, and therefore think that my data should be used for the study" ("Use data 5"). We also run a robustness check excluding the fastest and slowest 2.5% of the participants ("Time 95%"). Finally, we check for the presence of an experimenter demand effect (EDE). "Chose ESG" is a dummy that captures whether a sustainable fund was chosen or not. "Used Criterion" is a dummy that captures whether a participant reported to have used any ESG criterion for their decision. "Simple" models do not include any control variables. "Medium" models only include control variables that correlate with the Brochure Treatment variable. "Complex" models include all control variables.

Robustness Check, Variables	*Dependent Variable, Model:*					
	Chose ESG, simple	Chose ESG, medium	Chose ESG, complex	Used Criterion, simple	Used Criterion, medium	Used Criterion, complex
	(1)	(2)	(3)	(4)	(5)	(6)
Probits						
Brochure Treatment	−0.381	−0.497*	−0.446	0.561	0.072	−0.146
	(0.236)	(0.230)	(0.231)	(0.493)	(0.470)	(0.456)
Brochure * ESG Pref Score	0.230**	0.232**	0.211**	0.229	0.282	0.331*
	(0.080)	(0.078)	(0.078)	(0.162)	(0.154)	(0.149)
LPM						
Brochure-Treatment	−0.080	−0.110	−0.090	−0.033	−0.119	−0.129
	(0.073)	(0.072)	(0.073)	(0.100)	(0.093)	(0.095)
Brochure * ESG Pref Score	0.055*	0.054*	0.045	0.100**	0.108***	0.107***
	(0.025)	(0.024)	(0.024)	(0.034)	(0.031)	(0.031)
Use data 5						
Brochure Treatment	−0.639	−0.496*	−0.717	0.919	0.219	−0.134
	(0.398)	(0.230)	(0.395)	(0.861)	(0.832)	(0.799)
Brochure * ESG Pref Score	0.390**	0.232**	0.345*	0.410	0.456	0.534*
	(0.136)	(0.078)	(0.134)	(0.283)	(0.272)	(0.260)
Time 95%						
Brochure Treatment	−0.751	−0.858*	−0.648	0.400	−0.121	−0.421
	(0.415)	(0.406)	(0.409)	(0.887)	(0.849)	(0.816)
Brochure * ESG Pref Score	0.442**	0.414**	0.338*	0.574*	0.549*	0.623*
	(0.142)	(0.138)	(0.139)	(0.292)	(0.278)	(0.267)
EDE						
High EDE-Treatment	−0.189	−0.321	−0.951	−0.423	−0.355	−0.420
	(0.591)	(0.573)	(0.605)	(1.067)	(0.578)	(1.038)
Low EDE-Treatment	−0.953	−0.789	−1.242*	−2.227	−1.183	−1.296
	(0.620)	(0.603)	(0.614)	(1.138)	(0.618)	(1.078)
High EDE * ESG Pref Score	0.058	0.082	0.302	0.300	0.197	0.293
	(0.204)	(0.198)	(0.206)	(0.358)	(0.194)	(0.343)
Low EDE * ESG Pref Score	0.267	0.228	0.410*	0.677	0.394	0.488
	(0.211)	(0.206)	(0.208)	(0.375)	(0.204)	(0.352)
Note:						*p<0.05; **p<0.01; ***p<0.001

V.4.4 Additional analysis: Does sustainable finance literacy lead to differentiation between light green and dark green funds?

While little is known to date about how retail investors understand and are influenced by ESG information, there is clear empirical evidence that institutional investors respond to ESG information, particularly SFDR labels of funds (see e.g. Hartzmark and Sussman (2019); Becker et al. (2022); Scherer and Hasaj (2023)). Therefore, we examine whether sustainable finance literacy in combination with high sustainability preferences leads participants to differentiate between light green (SFDR Article 8) and dark green (SFDR Article 9) funds. Specifically, we hypothesize that individual investors with high sustainability preferences will deliberately invest in funds that explicitly pursue environmental or social objectives that are aligned with their preferences if they are able to identify information that allows them to recognize such funds.

To analyze this question, we focus on the fourth round of investment decisions, which includes all three types of SFDR funds: a dark green fund, a light green fund, and a conventional fund. We restrict the sample to the participants who chose one of the two sustainable funds. As the dependent variable, we use a dummy variable which indicates whether the light green or the dark green fund was chosen. As independent variables, we use the treatment variable, and additionally, we focus on one particular question from the sustainable finance literacy module: "Sufili local 4". This question specifically tests knowledge about financial products classified as SFDR Article 9.

If sustainable finance literacy influences the choice between article 8 and article 9 financial products, it should be the specific knowledge about the SFDR's article 9 in particular that causes this choice, but not necessarily other aspects of sustainable finance literacy. Therefore, we expect this question to have the strongest influence on the decision between the two fund classes, while the other questions should not be as influential. We again expect an interaction with ESG preferences, for the same reasons as in the main analyses. Thus, we mirror our analyses from the main results, and label the different models as "H1" and "H2". We only compute models with the full set of control variables.

Table V.10 provides support for both hypotheses. This table reports the results of logistic regressions as logits, for both hypotheses as well as for both independent variables. The brochure treatment increases the probability of choosing the article 9 fund over the article 8 fund (column 1), but this increase depends on the ESG preferences (column 2). Specific knowledge about the SFDR's article 9 also increases the probability of choosing a corresponding fund over an article 8 fund (column 3), again moderated by ESG preferences (column 4). In both columns (2) and (4) the coefficients of the interaction terms taking ESG preferences into account become larger than the individual effects of the brochure and the Sufili local 4-question.

Calculating the margins, as reported in Table V.11, shows that the brochure treatment increases the probability of choosing the article 9 fund over the article 8 one by around 9.4%, and knowing the correct answer to the Sufili local 4-question by around 12.2%. The interaction terms have the predicted signs.

We conclude that sustainable finance literacy not only increases the probability of investing sustainably at all, but also increases the probability of choosing the more sustainable option out of several sustainable alternatives. Since the other sustainable finance literacy questions do not increase the probability of choosing a dark green fund, it is reasonable to conclude that this effect is driven by specific knowledge about the SFDR's article 9 funds.

Table V.10: This table presents the results for the additional analyses of whether participants with higher sustainable finance literacy prefer article 9 funds ("dark green") over article 8 funds ("light green"), and reports the logits of logistic regressions. The sample is limited to decision 4, and only includes the decisions for any of the two sustainable funds. The dependent variable in each model is a dummy variable which equals 1 if participants chose the article 9 fund, and 0 if they chose the article 8 fund. The models differ in whether the main explanatory variable is the treatment or a question that specifically measures knowledge about article 9 funds ("Sufili local 4"), and whether this variable is interacted with the ESG Preferences Score. Each model includes all control variables. The ESG Pref Score is not centered.

	Chose dark green over light green, H1	Chose dark green over light green, H2	Chose dark green over light green, H1	Chose dark green over light green, H2
	(1)	(2)	(3)	(4)
Brochure Treatment	0.427*	−1.193		
	(0.197)	(0.819)		
Brochure * ESG Pref Score		0.547*		
		(0.269)		
Sufili local 4			0.567*	−1.654*
			(0.234)	(0.796)
Sufili local 4 * ESG Pref Score				0.764**
				(0.262)
ESG Pref Score	−0.023	−0.417	−0.032	−0.461*
	(0.163)	(0.253)	(0.166)	(0.223)
Sufili global 1			0.147	0.176
			(0.240)	(0.242)
Sufili global 2			0.028	0.036
			(0.241)	(0.242)
Sufili global 3			−0.227	−0.235
			(0.272)	(0.274)
Sufili global 4			−0.179	−0.202
			(0.208)	(0.210)
Sufili global 5			0.176	0.144
			(0.205)	(0.207)
Sufili local 1			0.002	−0.075
			(0.204)	(0.208)
Sufili local 2			−0.274	−0.253
			(0.231)	(0.232)
Sufili local 3			0.075	0.073
			(0.203)	(0.204)
Intercept	−1.059	0.144	−0.888	0.367
	(2.396)	(2.477)	(2.470)	(2.516)
Controls	all	all	all	all
N	714	714	709	709
Pseudo R^2	0.11	0.12	0.12	0.13

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table V.11: This table presents the results for the additional analyses of whether participants with higher sustainable finance literacy prefer article 9 funds ("dark green") over article 8 funds ("light green"), and reports the margins of logistic regressions. The sample is limited to decision 4, and only includes the decisions for any of the two sustainable funds. The dependent variable in each model is a dummy variable which equals 1 if participants chose the article 9 fund, and 0 if they chose the article 8 fund. The models differ whether the main explanatory variable is the treatment or a question that specifically measures knowledge about article 9 funds ("Sufili local 4"). Each model includes all control variables. The ESG Pref Score is not centered.

	Chose dark green over light green, H1	Chose dark green over light green, H1
Brochure Treatment	0.094*	
	(0.043)	
Sufili local 4		0.122*
		(0.050)
ESG Pref Score	−0.005	−0.007
	(0.036)	(0.036)
Sufili global 1		0.032
		(0.052)
Sufili global 2		0.006
		(0.052)
Sufili global 3		−0.049
		(0.059)
Sufili global 4		−0.039
		(0.045)
Sufili global 5		0.038
		(0.044)
Sufili local 1		0.000
		(0.044)
Sufili local 2		−0.059
		(0.049)
Sufili local 3		0.016
		(0.044)
Controls	all	all
Note:		*p<0.05; **p<0.01; ***p<0.001

At first glance, this result might seem to contradict the main findings of (Heeb et al., 2022), who show in their experiment that participants do not differentiate between different degrees of sustainable impact. Although article 9 funds are not necessarily impact funds and therefore may not necessarily achieve more impact compared to article 8 funds (Chesney and Lambillon, 2023), we would like to highlight a key difference between their experimental design and ours. While the treatments differ in how much of an impact the ESG fund has ("low impact" versus "high impact"), participants in Heeb et al. (2022)'s main experiment do not explicitly decide between two ESG funds directly. They either have to choose between a high-impact fund and a conventional fund, or a low-impact fund and a conventional fund.

In contrast, the participants in our experiment can decide between one conventional and two different ESG funds. A vast literature on preference construction and preference reversals emphasizes such contrasts in the choice environment as deciding factors (see Dhami (2016) and Lic (2006) for an overview). In our experiment, participants are able to compare both ESG funds, which may allow them to construct their preferences differently as the differences between these ESG funds become more apparent. A decision with only one ESG fund would not allow for that.

In an additional analysis, Heeb et al. (2022) actually find this "comparability" effect of an additional ESG fund as well. In their experiment, the participants decide between a conventional fund without any positive environmental impact, a second fund with a positive but relatively moderate impact, and a third fund with a relatively large positive impact. Heeb et al. (2022) conclude from their analyses that "the joint evaluation demonstrates that comparability creates some sensitivity to impact" (see p. 1765), which is consistent with our results.

Interestingly, for participants with very low levels of ESG preferences, we again find some evidence that sustainable finance literacy can decrease sustainable investments. In the H2 models in columns (2) and (4), the main effects are negative, and even significantly negative ($p<0.05$) for the Sufili local 4-question.

V.4.5 Additional analysis: Contrasting evidence on the determinants of ESG investing

In this section, we discuss further factors that influence ESG investment decisions by contrasting our results with the existing literature. The objective of this analysis is not to test established findings with our data,[13] but rather to analyze whether our dataset contains some of these established results in order to provide further evidence for the validity of our data. To this end, we use a correlation matrix, as shown in Table V.12, which reports the correlation coefficients between the "Chose ESG" variable and each of the control variables used in this study for which prior literature exists. We discuss whether these correlations are consistent with the literature in terms of coefficient sign and statistical significance. It is important to note, however, that there is no established consensus in the literature for several of the variables discussed below.

Age In our sample, age has a negative, but statistically insignificant correlation with the likelihood to invest in ESG funds. This is broadly in line with the literature, which usually finds that younger individuals invest more sustainably. Bauer et al. (2021), Bauer et al. (2022), Brodback et al. (2019), Giglio et al. (2023) and Gutsche and Zwergel (2020) observe a significant negative age effect. Bauer and Smeets (2015), Filippini et al. (2023) and Riedl and Smeets (2017) find a negative effect, but their results contain model specifications where age is not significant. In the context of a Swedish pension scheme, Anderson and Robinson (2021) derive a positive and significant correlation when the default investment option cannot be interpreted as having sustainable characteristics, and a negative and significant correlation when it can.

[13] See Hünermund and Beyers (2022) for why this would not be a feasible endeavor.

Table V.12: This table shows the correlation coefficients of each control variable for which there exists literature with the "Chose ESG" variable, sorted by p-value

Variable	Correlation	P-Value
Pearson correlation coefficients of "Chose ESG" with numeric and nominal control variables		
Perceived Impact	0.18	0.000***
Party Greens	0.12	0.000***
Greenwashing	-0.10	0.000***
Social Preferences	0.09	0.000***
Risk Preferences	-0.09	0.000***
Party FDP	-0.08	0.000***
Talks often about Inv.	-0.08	0.000***
Gender Female	0.07	0.000***
Party AfD	-0.06	0.000***
Financial Literacy	0.05	0.001**
Environmental Literacy	0.05	0.001**
Trust	0.05	0.002**
Social Preferences, costly	0.04	0.008**
Party The Left	0.04	0.015*
Time Preferences	0.03	0.084
Ln Age	-0.02	0.172
Party SPD	-0.01	0.374
Financial Experience	0.00	0.777
Years of education	0.00	0.786
Spearman correlation coefficients of "Chose ESG" with ordinal control variables		
Expected Risk of ESG Funds	-0.07	0.000***
Monthly Net Income	-0.04	0.034*
Expected Performance of ESG Funds	0.00	0.957
Note:		*p<0.05; **p<0.01; ***p<0.001

Gender While we find that gender predicts sustainable investments significantly, the evidence in the literature is rather mixed. Giglio et al. (2023) show that although women agree more with the fact that ESG investments are "the right thing to do" than men, they often do not translate these preferences into action. Bauer et al. (2021) analyze a Dutch pension fund that bases its sustainable investment policies on the decision of its members and find that women are more in favor of a sustainable investment policy. Gutsche and Zwergel (2020) and Bauer et al. (2022) observe that women invest significantly more in sustainable funds. Other studies (e.g., Anderson and Robinson (2021); Brodback et al. (2019); Filippini et al. (2023); Riedl and Smeets (2017) do not find a significant gender difference, although the coefficient sign is usually in favor of women investing more sustainably.

Education We find that the correlation between education and the probability of investing sustainably is zero. This finding is also common in the literature. Some studies do not find any results for education (e.g., Bauer et al. (2021); Gutsche and Zwergel (2020)). For those who derive significant results, the effects are mixed. Filippini et al. (2023) present six models, out of which four have a positive relationship. Bauer and Smeets (2015) and Riedl and Smeets (2017) both find that university education does not increase the absolute amount of sustainable investments, but on the contrary decreases their share in the portfolio. Finally, Bauer et al. (2022) find a positive relationship, but operationalize education as having a Ph.D., which limits generalizability.

Income and wealth While we observe a negative relationship between income and sustainable investments, the literature on income and wealth, as another measure of financial well-being, is very mixed. Anderson and Robinson (2021) exemplify this in their study, having significant correlations with income in both directions, depending on the model. Bauer and Smeets (2015) find no relationship for income, but a negative relationship for wealth. Bauer et al. (2021) and Filippini et al. (2023) also find no relationship with income, but Filippini et al. (2023) reports a positive correlation between wealth and sustainable investment. Giglio et al. (2023) show that the share of sustainable investments increases with wealth, while Riedl and Smeets (2017) show the opposite. Finally, Brodback et al. (2019) and Gutsche and Zwergel (2020) find positive associations between income and ESG investing, while Bauer et al. (2022) find a negative relationship.

Financial literacy and environmental literacy We find positive associations for both financial literacy and environmental literacy. The literature on these relationships is contrasted. Filippini et al. (2023) find no significant correlation for both these variables with sustainable investments. Similarly, Anderson and Robinson (2021) find no clear relationship for environmental literacy, and in the case of financial literacy, the relationship is in some models even significantly negative. In contrast, Aristei and Gallo (2021) find a positive relationship between financial literacy and sustainable investments, and Bethlendi et al. (2022) report a positive association between environmental literacy and ESG investing.

Return expectations In our data the correlation between expected returns of ESG investments and the likelihood to invest in ESG funds is zero. In a similar way, Heeb et al. (2022) find that neither investors' risk expectations nor their return expectations correlate significantly with their willingness to pay for sustainable investments. Bauer et al. (2021) report that individuals favor sustainable investments independent of return expectations. Specifically, they find that the majority of respondents in their experiment chooses to expand sustainable investing at their pension fund, even those who have negative return expectations or are uncertain about what to expect in terms of returns. Furthermore, Anderson and Robinson (2021) note that a green pro-social value orientation is strongly related to the willingness to pay higher fees for environmentally sustainable funds. Overall, this combined evidence suggests that return expectations are not the primary determinant of ESG investment decisions. Finally, Giglio et al. (2023) conduct a survey of retail investors and report considerable heterogeneity among these investors in their ESG return expectations and motivations for ESG investing, with 25% of respondents saying they are primarily motivated by ethical considerations and only 7% by return expectations.

Risk perception We find an inverse and statistically significant relationship between the perceived risk of ESG investments and the probability of investing in ESG funds. This seems reasonable, as we expect survey respondents to invest less in sustainable funds if they perceive them to be riskier – even more so as returns do not appear to be their primary motive. In contrast, Gutsche and Zwergel (2020) do not find any relationship between perceived higher risks of sustainable funds and investments in sustainable funds. Bauer and Smeets (2015) investigate the risk perceptions of retail investors with regard to SRI funds and find that these investors do not expect higher risks from such investments compared to conventional funds. Rather, they expect such investments to have both higher returns and lower risk, indicating that investors might have a poor understanding of the relation between risk and return on securities, or that they are overconfident about sustainable investments.

Political preferences We find that Green Party and Left party voters are more likely to invest sustainably than CDU/CSU voters (i.e., the reference category in Table V.12), while voters from the pro-business party FDP and the far-right party AfD are less likely to do so. In general, these findings, in particular the effects observed for the Green Party voters, are in line with the literature. Based on an experiment with German households, Gutsche and Zwergel (2020) show that participants with an ecological political identification invest more sustainably in a stated choice experiment. Briere and Ramelli (2021) demonstrate that French individuals living in regions with a high share of Green Party voters invest more sustainably. For the U.S., Giglio et al. (2023) find that there is a higher ESG participation by retail investors resident in predominately Democratic areas compared to Republican ones. Even Anderson and Robinson (2021), who find no relationship between green attitudes and sustainable investment, still provide evidence for a strong association between voting in favor of the Swedish Green Party and ESG investing.

Social preferences and trust We show a positive and statistically significant relationship between social preferences (for both "social preferences" and "social preferences, costly") and probability to invest in ESG funds. Riedl and Smeets (2017) find that social preferences are key to investing in sustainable funds at all, but they do not explain how much wealth is allocated to these funds. Bauer et al. (2021) show that social preferences rather than financial beliefs drive the choice for more sustainability. Our results also show a positive correlation between trust and ESG investing. This is in line with Gutsche and Zwergel (2020), but not perfectly aligned with Filippini et al. (2023), who do find positive correlations, but except for one model they are insignificant.

Time and risk preferences We find a positive, but statistically insignificant correlation between time preferences and the likelihood to invest in ESG funds. This is broadly consistent with Bauer et al. (2022) who report that individuals with a longer-term perspective are more significantly likely to invest in a sustainable funds. Turning to risk preferences, we find that such preferences negatively predict sustainable investments. This is in line with Bauer and Smeets (2015), but not fully consistent with Filippini et al. (2023) who find a weakly positive relationship between risk preferences and ownership of sustainable financial products. Likewise, Riedl and Smeets (2017) find a positive correlation between risk tolerance and the amount invested in sustainable equity funds, but no significant impact on the probability to invest in a sustainable manner.

Financial experience The correlation between financial experience and likelihood to invest in ESG products is zero in our data. This suggests that having financial experience doesn't make one more or less likely to invest in ESG funds. Anderson and Robinson (2021) find that households that exhibit strong pro-environmental behaviors and beliefs are financially disengaged and generally uninterested in financial matters. In addition, Kaustia and Torstila (2011) show that left-wing investors are less inclined to invest in stocks because of their general aversion toward financial markets. Briere and Ramelli (2021) report that the offering of sustainable investment options increases the willingness of investors, including those with a strong pro-social orientation, to participate in financial markets due to a better value alignment. Thus, if these individuals do invest, their investments are likely more driven by environmental and social considerations than by financial expertise.

Talks often about investments Riedl and Smeets (2017) use this variable as a proxy for social signaling and report that investors who talk more often about their investments are more likely to invest in a socially responsible way. In our experiment, we find the opposite result. However, as Riedl and Smeets (2017) note, this variable is likely not a pure measure of social signaling. Instead, it may also be associated with other determinants of sustainable investing, such as financial experience or financial engagement. In particular, we expect that individuals who are more financially engaged are also more likely to talk about their investments with their peers. The observed negative relationship between the variable "talks often about investments" and the probability of investing in ESG funds seems to confirm this and might once again suggest that individuals with strong sustainability preferences are more disengaged from financial decisions.

Perceived impact We find a positive and statistically significant relationship between perceived impact and the likelihood to invest in ESG funds. This is consistent with studies using similar variables. Riedl and Smeets (2017) report that people who perceive ESG funds as having a positive impact on society have a higher likelihood of holding ESG equity. Brodback et al. (2019) find such a positive relationship in a survey among retail investors on the relative importance of social responsibility. In addition, Heeb et al. (2022), find no effect of the actual impact in their main experiments but some effect using the comparability treatments, which might point to the positive influence of perceived impact on sustainable investments.

In sum, the correlations found in our dataset are broadly in line with the patterns and relationships observed in the literature. This validates our dataset and strengthens its external validity, and consequently our main results.

V.5 Discussion & implications

In this paper, we show that knowledge about sustainable investments combined with at least moderate sustainability preferences can significantly influence sustainable investment behavior. In particular, we show that an educational brochure designed to increase sustainable finance literacy increases the probability of investing in an ESG-labelled fund by around 9%, of which around 4 to 5% are causally mediated by an increase in sustainable finance literacy. We also show that the brochure leads participants to consciously use more ESG criteria in their investment decisions by around 20%. Around two-thirds of this effect size can be attributed to sustainable finance literacy. However, the brochure does not work for participants with low sustainability preferences, and there is some weak evidence that it might even reduce their willingness to invest sustainably. For participants who decided to invest sustainably, the brochure also increases the probability of investing in "dark green" article 9 funds, compared to "light green" article 8 funds, by around 12%.

Our study entails some limitations, which warrant careful consideration but also may highlight the potential for further research. One of the primary limitations of our study is that it relies on a sample of German participants recruited through Prolific, which may not fully represent the diversity of characteristics and preferences of retail investors worldwide. Future research could therefore extend our findings to additional geographical and cultural backgrounds, preferably by conducting field experiments in an actual financial context. This approach would provide a more complete understanding of the generalizability of our findings across different investor demographics.

Furthermore, in keeping with Filippini et al. (2023)'s definition of sustainable finance literacy, a portion of our brochure is focused on the regulatory context of the EU, and in particular, the SFDR. Future research could explore the adaptation of educational materials to other regulatory contexts. This approach would provide further insights into the definition of sustainable finance literacy and the effectiveness of tailored educational content in promoting sustainable investment behavior.

In addition, our study focuses on the immediate effects of providing an educational brochure on participants' investment decisions. While this is a faithful implementation of our research question, for practical purposes, it would be desirable to gain further insights into the longer-term impacts of sustainable finance literacy. Tracking participants' investment decisions over time could shed light on whether such a brochure treatment is effective in the long run and has a lasting impact on investment behavior.

By using screenshots of funds offered by a bank through an online financial account, the selection environment of our experiment has high external validity, but it is also accompanied by a simplified and standardized presentation of information. This selection may not be universally applicable, as some retail investors may face a broader range of ESG information in other investment situations. As Anderson and Robinson (2021) point out, one of the biggest challenges in making sustainable investment decisions is overcoming the informational hurdles associated with such decisions. In other words, in an environment with many different sources of information that are presented in a non-standardized way, the task of choosing a sustainable fund that matches one's sustainability preferences might be more difficult than in our environment. In addition, even the task of searching for reliable and trustworthy information can be daunting, which might deter investors from investing sustainably in the first place. Both these arguments suggest a larger role of sustainable finance literacy in the field, which we cannot capture in our experiment. Future research might delve into that aspect.

Our results, combined with this argument, also provide important implications for the way ESG-related information should be displayed and communicated to retail investors. The fund classification emerging from the SFDR lowers information hurdles substantially. That such a simplification of the choice environment is useful might best be illustrated by the results of the additional analyses, where we show that the specific knowledge about the SFDR's article 9 increases the probability of choosing such funds. It is unlikely that we would find such results if the requirements behind article 9 were not condensed into such a concise format – the caveat being, of course, that such requirements are implemented and enforced faithfully, and that this classification does not degrade into a tool for greenwashing.

Our findings also have important implications for policymakers. Policymakers in the EU are increasingly emphasizing, for example through MiFID II, the need for financial professionals to assess and take into account their clients' sustainability preferences when making investment recommendations. While this focus on quantifying sustainability preferences is justified, our research highlights a potential pitfall of exclusively concentrating on preferences without considering individuals' knowledge of sustainable investments. In particular, our findings show that it is equally important to assess retail investors' knowledge of sustainable investments. Policymakers should therefore strive to strike a balance between measuring ESG preferences and promoting sustainable finance literacy. This entails ensuring that retail investors have access to educational resources and information that allow them to make informed decisions.

V.6 Conclusion

We present evidence that investors' sustainable investment behavior is not only driven by their ESG preferences, but also by their knowledge of sustainable financial products. We arrive at this result using a pre-registered experiment based on a large sample of German participants recruited through the Prolific platform. Our findings have important implications for our understanding of how to model and predict investors' sustainable investment behavior. Moreover, our research bears implications for policymakers seeking to integrate sustainability considerations into the financial system and to steer capital flows towards sustainable investments.

Appendix I (to Chapter I)

Methodological discussion

While evidence on methodological "standard objections" tends to be mixed and effects are often quite small (Camerer (2015); Camerer and Hogarth (1999); Dhami (2016); Zizzo (2010)), stake size might be a reasonable objection for our case, since high incentives can induce higher effort and may thereby improve performance in decision-making and problem solving (Camerer and Hogarth, 1999). After all, repaying all the money on the low-interest rate card reduces bonus payments by only 83 US-cents in our MTurk-based experiment in chapter II. However, due to the change that no money can be left on the checking account, in the lab replication this difference is €10, and we find even more misallocation (see Appendix II). The fact that stake size has rather the opposite effect in our study should hence be seen as an argument against this objection.

Another problem with such standard methodological objections is that they do not imply any predictable patterns in the data. We, in contrast, find distinct patterns. In order to explain these as artifacts would hence require specific methodological objections that allow to predict the Cuckoo Fallacy, the Complete Repayment fallacy, Equal Start effects *and* the 1/N strategy. A second example for strong structural effects that are hard to reconcile with methodological objections are the strong differences in the distributions of the chosen options in the experiment, say between the Everything Equal and the control scenario for the 1/N Heuristic. Subjects that are not responding to our incentives or instructions should not behave that differently between scenarios as they do. In the same vein, we find that the effects of financial literacy on misallocation are stable and negative, which raises the question which objection allows to explain a constant relationship between financial literacy and the vulnerability to experimental effects. The argument that "subjects find the experiment too easy to be true", for instance, predicts a positive correlation between financial literacy and misallocation, since the more literate a person is, the easier the task should appear - but we find the opposite. Relationships between financial literacy and experimenter demand effects, scrutiny or other typical experimental artifacts are also not obvious, so again the question remains which methodological error, or combination of errors, could cause the observed patterns.

One specific experimenter demand effect might exist for some of the fallacies in this experiment, however, only if we assume that participants anticipate our hypothesis that fallacy scenarios cause misallocation. For fallacies such as the Cuckoo Fallacy or Complete Repayment, it might be easier in the fallacy scenario to predict which button we "want" our subjects to click than in the control scenario, which might explain why they use it more often. We tried to tackle this problem by ruling out that a fallacy scenario and its control scenario can occur directly after each other to muddle the contrasts somewhat, but ultimately we cannot exclude this explanation. However, we are not sure about the direction of that effect, because we have severe doubts that our participants systematically anticipate our hypothesis, and even if they do, that they are motivated to follow "our demands". Consider what kind of a situation this experimenter demand effect assumes: A subject correctly identifies a fallacy scenario as a trap, and then believes we "want" them to step into the trap, so they do. But we believe it is just as likely

- if not more so - that in such a situation a participant thinks we "want" them to identify the trap and avoid it - which would lead to the exact opposite behavior. And to the degree that participant interpret setting up traps as a negative behavior by us, reciprocity might change their motive from "helping" us to "showing" us, which again predicts the opposite effect. In combination with standard objections against experimenter demand effects such as that they go against the incentive structure and that they are less severe in online experiments than in a lab experiment where we as researchers are physically present, we do not think this explanation works for the scenarios.

A final methodological objection that we want to raise is the strong effect size in our data. Set against some parts of the earlier literature, the number of subjects that make at least one non-optimal choice in our experiment is with around 82% very high (see I.6). Keys et al. (2016), for instance, find that around 20% of US households who could refinance mortgages more cheaply did not, even though this task is clearly more complex than our experiments. Agarwal et al. (2015) show that in a natural experiment where consumers could acquire a credit card, roughly 40% chose the higher interest rate card. In Keys and Wang (2019), only up to 20% of credit card owners are influenced by anchoring due to minimum repayments. Our results appear less outlandish, however, when compared to the literature most closely related to our work. In experiment #1 of Amar et al. (2011) the misallocation is 41% or 49%, depending on the treatment, and virtually no participant finished their 25 rounds game without any misallocation. The two field studies that resemble our work most closely have similar results: While the share of misallocating people is not directly reported in Ponce et al. (2017), Gathergood et al. (2019) indicate that "85 percent of individuals should put 100 percent of their excess payments on the high-interest rate card but only 10 percent do so". And the results in the field are even stronger than in our data if we refer to misallocation itself. In both

Gathergood et al. (2019) and Ponce et al. (2017), the average observed misallocation is around 50%, while in our data - determined by lost bonus payments in the scenarios 1 to 14 - it is around 25% (on average \$ 1.05 loss of bonus per participant on a maximum bonus of \$ 4.20 in 14 scenarios, see chapter I, especially Table I.2). In fact, we should have expected the effect to be smallest in a pure, but potentially immeasurable state, modest in an experimental setting with some methodological problems or a design to provoke misallocation, and highest in the field, where most distractions and a selection effect with respect to the use of credit cards exist. Altogether, we hence admit that some or a combination of methodological objections might influence the effect size, or even specify interaction effects, but we believe that they cannot negate the existence of the reported effects.

Additional tables and figures

Table Appendix I.13: Duration statistics for the experimental (in minutes)

Duration statistic (min:sec)	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Total	335	22:28	10:08	05:54	15:02	20:25	26:59	59:06
Instructions	335	10:56	07:28	01:39	06:20	08:55	12:34	52:03
Experimental stages	335	06:11	04:02	02:02	03:50	05:02	07:17	38:55
Post exp. questionnaire	335	05:20	03:19	01:15	03:23	04:39	06:19	36:33

Table Appendix I.14: Logistic regression model with random effects[a]

	Dependent variable: Choice of fallacy-implicated repayment option (1 = Chosen, 0 = Not chosen)						
	Cuckoo Fallacy	Equalize Balances	Complete Repayment	Balance Matching	1/N Heuristic	Interest Matching	Equal Start
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fallacy scenario	0.287***	0.023	0.175***	0.038	0.239***	-0.084*	0.216***
	(0.038)	(0.014)	(0.024)	(0.016)	(0.033)	(0.026)	(0.027)
	[0.000]	[0.096]	[0.000]	[0.020]	[0.000]	[0.001]	[0.000]
	[0.000]	[0.766]	[0.000]	[0.223]	[0.000]	[0.022]	[0.000]
Financial literacy	-0.007	-0.017	-0.013	-0.018	0.013	-0.011	-0.026
	(0.027)	(0.008)	(0.019)	(0.012)	(0.037)	(0.018)	(0.020)
	[0.785]	[0.031]	[0.473]	[0.109]	[0.723]	[0.534]	[0.208]
	[1.000]	[0.310]	[1.000]	[0.764]	[1.000]	[1.000]	[1.000]
Fall. scen. × Fin. lit.	0.019	0.008	0.019	0.019	-0.033	-0.009	0.004
	(0.030)	(0.007)	(0.019)	(0.013)	(0.037)	(0.021)	(0.024)
Age	-0.004**	-0.001	-0.004**	-0.003*	0.002	-0.001	0.001
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Dummy: Male	-0.016	-0.024	-0.030	-0.030	-0.004	-0.011	-0.021
	(0.028)	(0.018)	(0.032)	(0.023)	(0.027)	(0.043)	(0.040)
Years of education (yoe)	-0.007	-0.004	-0.010	-0.003	-0.006	0.002	-0.030***
	(0.006)	(0.004)	(0.007)	(0.005)	(0.006)	(0.010)	(0.009)
Observations	670	670	670	670	670	670	670

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$ for the Holm-adjusted p-values

[a] Reported coefficients are margins. The seven models denote the seven scenario pairs, the differences of control- and fallacy scenario are denoted in the Fallacy scenario coefficients. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for 28 coefficients from two tables: The seven fallacy scenario coefficients for Fallacy-Implicated Option as dependent variable, the seven fallacy scenario coefficients for Optimal Option as dependent variable, as well as the 14 financial literacy coefficients from both tables. Asterisks indicate significance after adjustment.

Table Appendix I.15: Logistic regression model with random effects[a]

	Dependent variable: Choice of optimal repayment option (1 = Chosen, 0 = Not chosen)						
	Cuckoo Fallacy	Equalize Balances	Complete Repayment	Balance Matching	1/N Heuristic	Interest Matching	Equal Start
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fallacy scenario	-0.103**	-0.010	-0.073	-0.057	-0.353***	0.079*	-0.086*
	(0.028)	(0.026)	(0.025)	(0.027)	(0.022)	(0.026)	(0.028)
	[0.000]	[0.695]	[0.004]	[0.032]	[0.000]	[0.002]	[0.002]
	[0.005]	[1.000]	[0.055]	[0.291]	[0.000]	[0.036]	[0.033]
Financial literacy	0.069*	0.083***	0.052	0.068*	0.070*	0.064*	0.058
	(0.021)	(0.018)	(0.019)	(0.021)	(0.020)	(0.020)	(0.020)
	[0.001]	[0.000]	[0.006]	[0.001]	[0.001]	[0.002]	[0.004]
	[0.016]	[0.000]	[0.073]	[0.021]	[0.011]	[0.031]	[0.056]
Fall. scen. × Fin. lit.	-0.014	-0.008	0.024	-0.011	0.000	0.020	-0.002
	(0.026)	(0.020)	(0.025)	(0.022)	(0.023)	(0.022)	(0.024)
Age	-0.000	0.005*	0.004	0.005*	0.001	0.004	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Dummy: Male	0.088	0.042	0.089	0.031	0.030	0.050	0.093*
	(0.045)	(0.044)	(0.047)	(0.048)	(0.042)	(0.047)	(0.047)
Years of education (yoe)	0.001	0.007	0.008	0.009	0.006	0.002	0.014
	(0.011)	(0.010)	(0.011)	(0.011)	(0.009)	(0.011)	(0.010)
Observations	670	670	670	670	670	670	670

Note: $^{*}p$<0.05; $^{**}p$<0.01; $^{***}p$<0.001 for the Holm-adjusted p-values

[a] Reported coefficients are margins. The seven models denote the seven scenario pairs, the differences of control- and fallacy scenario are denoted in the Fallacy scenario coefficients. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for 28 coefficients from two tables: The seven fallacy scenario coefficients for Fallacy-Implicated Option as dependent variable, the seven fallacy scenario coefficients for Optimal Option as dependent variable, as well as the 14 financial literacy coefficients from both tables. Asterisks indicate significance after adjustment.

Table Appendix I.16: Number of choices for each repayment option in each scenario.

	All on low	2:1	1:1	1:2	All on high	∅ MA
Scenario 01: Cuckoo Fallacy, Control	12	13	67	95	148	0.26
Scenario 02: Cuckoo Fallacy, Treatment	97	42	38	45	113	0.47
Scenario 03: Equalize Balances, Control	8	16	15	67	229	0.14
Scenario 04: Equalize Balances, Treatment	13	19	14	63	226	0.16
Scenario 05: Complete Repayment, Control	24	20	50	97	144	0.28
Scenario 06: Complete Repayment, Treatment	81	17	33	84	120	0.41
Scenario 07: Balance Matching, Control	8	17	20	92	198	0.18
Scenario 08: Balance Matching, Treatment	20	27	24	85	179	0.23
Scenario 09: 1/N Heuristic, Control	19	15	10	32	259	0.13
Scenario 10: 1/N Heuristic, Treatment	10	20	82	93	130	0.28
Scenario 11: Interest Matching, Control	10	19	14	104	188	0.19
Scenario 12: Interest Matching, Treatment	13	10	22	76	214	0.17
Scenario 13: Equal Start, Control	14	18	66	90	147	0.27
Scenario 14: Equal Start, Treatment	10	9	140	58	118	0.31
Scenario 15: Everything Equal	9	8	274	5	39	-

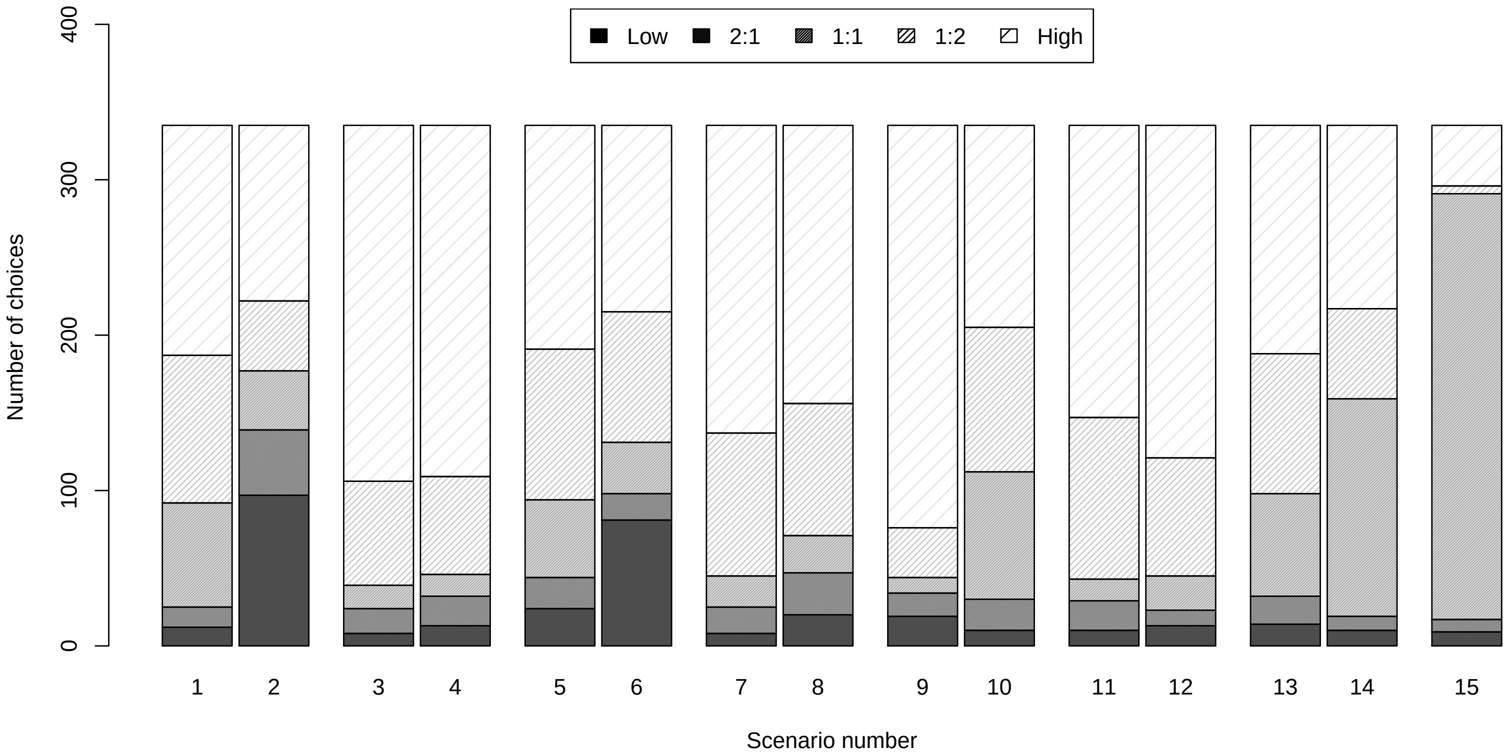


Figure Appendix I.5: Relative proportion of choices in the scenarios

Table Appendix I.17: OLS regression model with random effects, with control variables[a]

	Dependent variable: Choice of fallacy-implicated repayment option (1 = Chosen, 0 = Not chosen)						
	Cuckoo Fallacy (All low)	Equalize Balances (All low)	Complete repayment (All low)	Balance Matching (2:1)	1/N Heuristic (1:1)	Interest Matching (1:2)	Equal Start (1:1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fallacy scenario	0.254***	0.015	0.170***	0.030	0.214***	-0.084*	0.221***
	(0.027)	(0.011)	(0.023)	(0.016)	(0.025)	(0.026)	(0.029)
	[0.000]	[0.159]	[0.000]	[0.051]	[0.000]	[0.001]	[0.000]
	[0.000]	[1.000]	[0.000]	[0.508]	[0.000]	[0.025]	[0.000]
Financial literacy	0.003	-0.017	-0.003	-0.018	0.003	-0.012	-0.021
	(0.010)	(0.011)	(0.014)	(0.014)	(0.011)	(0.020)	(0.018)
	[0.732]	[0.101]	[0.837]	[0.196]	[0.807]	[0.550]	[0.237]
	[1.000]	[0.913]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Constant	0.292*	0.127*	0.378**	0.186*	0.046	0.327	0.607***
	(0.119)	(0.063)	(0.122)	(0.086)	(0.102)	(0.174)	(0.142)
Observations	670	670	670	670	670	670	670
Fall. scen. × Fin. lit.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.131	0.031	0.073	0.022	0.110	0.012	0.085

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$ for the Holm-adjusted p-values

[a] The seven models denote the seven scenario pairs. In each model, we compare a fallacy scenario with its respective control scenario. The differences between these two scenarios are denoted in the Fallacy scenario coefficients. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for 28 coefficients from two tables: The seven fallacy scenario coefficients for Fallacy-Implicated Option as dependent variable, the seven fallacy scenario coefficients for Optimal Option as dependent variable, as well as the 14 financial literacy coefficients from both tables. Asterisks indicate significance after adjustment.

Table Appendix I.18: OLS regression model with random effects, with control variables[a]

	Dependent variable: Choice of optimal repayment option (1 = Chosen, 0 = Not chosen)						
	Cuckoo Fallacy	Equalize Balances	Complete repayment	Balance Matching	1/N Heuristic	Interest Matching	Equal Start
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fallacy scenario	-0.105**	-0.009	-0.071	-0.057	-0.385***	0.078*	-0.087*
	(0.028)	(0.026)	(0.026)	(0.027)	(0.030)	(0.027)	(0.028)
	[0.000]	[0.727]	[0.005]	[0.035]	[0.000]	[0.003]	[0.002]
	[0.005]	[1.000]	[0.069]	[0.388]	[0.000]	[0.048]	[0.030]
Financial literacy	0.071*	0.089***	0.054	0.069*	0.069*	0.068*	0.059*
	(0.020)	(0.020)	(0.019)	(0.022)	(0.021)	(0.021)	(0.020)
	[0.001]	[0.000]	[0.006]	[0.001]	[0.001]	[0.001]	[0.003]
	[0.011]	[0.000]	[0.069]	[0.025]	[0.021]	[0.027]	[0.044]
Constant	0.373*	0.380*	0.109	0.268	0.651***	0.363	0.090
	(0.180)	(0.183)	(0.192)	(0.188)	(0.156)	(0.194)	(0.182)
Observations	670	670	670	670	670	670	670
Fall. scen. × Fin. lit.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.052	0.084	0.061	0.054	0.194	0.067	0.054

Note: $^{*}p$<0.05;** p<0.01;*** p<0.001 for the Holm-adjusted p-values

[a] The seven models denote the seven scenario pairs. In each model, we compare a fallacy scenario with its respective control scenario. The differences between these two scenarios are denoted in the Fallacy scenario coefficients. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for 28 coefficients from two tables: The seven fallacy scenario coefficients for Fallacy-Implicated Option as dependent variable, the seven fallacy scenario coefficients for Optimal Option as dependent variable, as well as the 14 financial literacy coefficients from both tables. Asterisks indicate significance after adjustment.

Table Appendix I.19: Multinomial Regression analysis[a]

	Dependent variable: Chosen option						
	Cuckoo Fallacy	Equalize Balances	Complete Repayment	Balance Matching	1/N Heuristic	Interest Matching	Equal Start
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fallacy scenario							
Option 1	0.260***	0.024	0.182***	0.036	-0.014	0.008	-0.014
	(0.036)	(0.018)	(0.031)	(0.021)	(0.014)	(0.017)	(0.015)
	[0.000]	[0.178]	[0.000]	[-]	[-]	[-]	[-]
	[0.000]	[1.000]	[0.000]	[-]	[-]	[-]	[-]
Option 2	0.130***	0.013	-0.011	0.038	-0.010	-0.019	-0.024
	(0.035)	(0.020)	(0.023)	(0.022)	(0.018)	(0.018)	(0.019)
	[-]	[-]	[-]	[0.078]	[-]	[-]	[-]
	[-]	[-]	[-]	[0.861]	[-]	[-]	[-]
Option 3	-0.116***	-0.010	-0.054*	0.013	0.225***	0.014	0.217***
	(0.027)	(0.018)	(0.026)	(0.020)	(0.036)	(0.019)	(0.032)
	[-]	[-]	[-]	[-]	[0.000]	[-]	[0.000]
	[-]	[-]	[-]	[-]	[0.000]	[-]	[0.000]
Option 4	-0.148***	-0.014	-0.037	-0.027	0.150***	-0.082	-0.094**
	(0.028)	(0.031)	(0.034)	(0.034)	(0.030)	(0.034)	(0.031)
	[-]	[-]	[-]	[-]	[-]	[0.017]	[-]
	[-]	[-]	[-]	[-]	[-]	[0.272]	[-]
Option 5	-0.125**	-0.014	-0.081	-0.061	-0.351***	0.079	-0.086
	(0.035)	(0.036)	(0.036)	(0.038)	(0.027)	(0.037)	(0.036)
	[0.000]	[0.704]	[0.026]	[0.106]	[0.000]	[0.033]	[0.017]
	[0.008]	[1.000]	[0.370]	[0.848]	[0.000]	[0.427]	[0.285]
Financial literacy							
Option 1	0.001	-0.017	-0.017	-0.020*	-0.021**	-0.013	-0.000
	(0.029)	(0.008)	(0.022)	(0.010)	(0.007)	(0.008)	(0.008)
	[0.978]	[0.037]	[0.450]	[-]	[-]	[-]	[-]
	[1.000]	[0.449]	[1.000]	[-]	[-]	[-]	[-]
Option 2	-0.062***	-0.024**	-0.039***	-0.019	-0.022*	-0.020*	-0.024***
	(0.016)	(0.009)	(0.009)	(0.013)	(0.010)	(0.008)	(0.007)
	[-]	[-]	[-]	[0.138]	[-]	[-]	[-]
	[-]	[-]	[-]	[0.969]	[-]	[-]	[-]
Option 3	-0.023	-0.019*	-0.017	-0.021*	0.022	-0.011	-0.034
	(0.014)	(0.007)	(0.013)	(0.010)	(0.044)	(0.008)	(0.021)
	[-]	[-]	[-]	[-]	[0.619]	[-]	[0.105]
	[-]	[-]	[-]	[-]	[1.000]	[-]	[0.941]
Option 4	0.002	-0.023	0.023	-0.012	-0.026	-0.017	0.001
	(0.016)	(0.017)	(0.019)	(0.019)	(0.024)	(0.018)	(0.017)
	[-]	[-]	[-]	[-]	[-]	[0.342]	[-]
	[-]	[-]	[-]	[-]	[-]	[1.000]	[-]
Option 5	0.082*	0.083***	0.050	0.072*	0.047	0.062*	0.057
	(0.025)	(0.019)	(0.021)	(0.022)	(0.029)	(0.021)	(0.021)
	[0.001]	[0.000]	[0.020]	[0.001]	[0.104]	[0.003]	[0.006]
	[0.018]	[0.000]	[0.300]	[0.018]	[1.000]	[0.048]	[0.113]
Observations	670	670	670	670	670	670	670
Fall. scen. × Fin.Lit	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: $^{*}p$<0.05;** p<0.01;*** p<0.001 for the Holm-adjusted p-values

[a] Reported coefficients are margins and denote the estimated differences in the probability that a certain option is chosen (rows) for all seven scenarios (columns). The first block "Fallacy scenario" shows the average differences in percentage of chosen options when switching from the control to the corresponding fallacy scenario. The second block "Financial literacy" shows how each correctly answered financial literacy question changes the probability to choose a certain option. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values for the variables we interpret in brackets. The p-values are adjusted to include all the 28 coefficients for which we present adjusted p-values. Asterisks indicate significance after adjustment.

Table Appendix I.20: OLS of the changes between control and fallacy scenario[a]

	Dependent variable: Option change between control and fallacy scenario						
	Cuckoo Fallacy (All low)	Equalize Balances (All low)	Complete Repayment (All low)	Balance Matching (2:1)	1/N Heuristic (1:1)	Interest Matching (1:2)	Equal Start (1:1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	-0.952***	-0.069	-0.513***	-0.236***	-0.549***	0.081	-0.218**
	(0.102)	(0.054)	(0.086)	(0.061)	(0.066)	(0.053)	(0.064)
	[0.000]	[0.207]	[0.000]	[0.000]	[0.000]	[0.126]	[0.001]
	[0.000]	[0.252]	[0.000]	[0.001]	[0.000]	[0.252]	[0.002]
Observations	335	335	335	335	335	335	335
Further control variables	No	No	No	No	No	No	No
R^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$ for the Holm-adjusted p-values

[a] The constant estimates the mean number of options a participant changes between the two scenario types, negative values implicate a change away from the optimal option. The seven models denote the seven scenario pairs. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for all 7 coefficients. Asterisks indicate significance after adjustment.

Table Appendix I.21: Logistic regression model with random effects (including screened out participants)[a]

	Dependent variable: Choice of fallacy-implicated repayment option (1 = Chosen, 0 = Not chosen)							
	Cuckoo Fallacy	Cuckoo Fallacy	Equalize Balances	Complete Repayment	Balance Matching	1/N Heuristic	Interest Matching	Equal Start
	(1.1)	(1.2)	(2)	(3)	(4)	(5)	(6)	(7)
fallacy scenario	0.279	0.281***	0.026	0.171***	0.038	0.224***	-0.080*	0.220***
	(2.494)	(0.037)	(0.015)	(0.024)	(0.017)	(0.032)	(0.026)	(0.026)
	[0.911]	[0.000]	[0.077]	[0.000]	[0.025]	[0.000]	[0.002]	[0.000]
	[see caption]	[0.000]	[0.561]	[0.000]	[0.221]	[0.000]	[0.025]	[0.000]
Financial literacy	-0.017	-0.006	-0.019*	-0.012	-0.020	-0.020	-0.007	-0.023
	(0.142)	(0.026)	(0.008)	(0.018)	(0.011)	(0.028)	(0.018)	(0.020)
	[0.902]	[0.824]	[0.014]	[0.493]	[0.070]	[0.474]	[0.690]	[0.257]
	[see caption]	[1.000]	[0.157]	[1.000]	[0.561]	[1.000]	[1.000]	[1.000]
Observations	686	684	686	686	686	686	686	686
Fall. scen. × Fin. lit.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$ for the Holm-adjusted p-values

[a] Reported coefficients are margins. The seven models denote the seven scenario pairs, the differences of control- and fallacy scenario are denoted in the Fallacy scenario coefficients. The first column shows a model where we believe there was a technical error in the algorithm that prevented the calculations of the standard errors for the Cuckoo Fallacy from succeeding properly. Only when we include a particular person *and* use robust standard errors *and* include the age a a control variable *and* report the margins, we get a standard error of 2.494. If any of these conditions is not fulfilled, the standard error decreases by a factor of around 65. The particular participant shows no anomalies (e.g., she is 28 years old). We do not think this is a "legit" standard error but a problem of the margin calculation (else we would see the problem in the logit calculation as well, but we do not), so we solved the problem by screening out the problematic participant. The results are in column (1.2). For completeness, we still report the erroneous calculation in column (1.1), but do ignore this model for the Bonferroni-Holm correction. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for 28 coefficients from two tables: The seven fallacy scenario coefficients for Fallacy-Implicated Option as dependent variable, the seven fallacy scenario coefficients for Optimal Option as dependent variable, as well as the 14 financial literacy coefficients from both tables. Asterisks indicate significance after adjustment.

Table Appendix I.22: Logistic regression model with random effects (including screened out participants)[a]

	Dependent variable: Choice of optimal repayment option (1 = Chosen, 0 = Not chosen)						
	Cuckoo Fallacy	Equalize Balances	Complete Repayment	Balance Matching	1/N Heuristic	Interest Matching	Equal Start
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
fallacy scenario	-0.103**	-0.013	-0.068	-0.059	-0.347***	0.079*	-0.090*
	(0.027)	(0.026)	(0.025)	(0.026)	(0.022)	(0.026)	(0.027)
	[0.000]	[0.628]	[0.008]	[0.024]	[0.000]	[0.002]	[0.001]
	[0.004]	[1.000]	[0.091]	[0.240]	[0.000]	[0.030]	[0.017]
Financial literacy	0.073**	0.083***	0.060*	0.072**	0.076**	0.063*	0.065*
	(0.020)	(0.018)	(0.018)	(0.020)	(0.020)	(0.020)	(0.020)
	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.002]	[0.001]
	[0.005]	[0.000]	[0.016]	[0.007]	[0.003]	[0.026]	[0.017]
Observations	670	670	670	670	670	670	670
Fall. scen. × Fin. lit.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: *p<0.05; **p<0.01; ***p<0.001

[a] Reported coefficients are margins. The seven models denote the seven scenario pairs, the differences of control- and fallacy scenario are denoted in the Fallacy scenario coefficients. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for 28 coefficients from two tables: The seven fallacy scenario coefficients for Fallacy-Implicated Option as dependent variable, the seven fallacy scenario coefficients for Optimal Option as dependent variable, as well as the 14 financial literacy coefficients from both tables. Asterisks indicate significance after adjustment.

Appendix II (to Chapter II)

Explanation of numbers

We set the starting debts on each credit card to $2200 and the starting income on the checking account to $250. One of the credit cards has an interest rate of 3% per round and the other one of 5% per round. In every round the checking account is refilled with $250. We choose these particular values for account levels and interest rates because they fulfill several conditions:

1. Both credit cards start with the same amount of debts, so the balances do not "favor" any card in the first round.

2. It is not possible to repay one of the cards completely in ten rounds. For our research questions we are only interested in situations where subjects actually have to make a choice between two cards. Therefore, ruling out this possibility ensures that we can evaluate every round of each subject.

3. The total new debt on both cards in each round does not exceed the income in the checking account, so subjects would not get the impression of "pointless repayments" due to runaway debts.[14]

[14] If a subject does not repay anything at all, then their total new debts do exceed the checking account deposits in rounds 9 and 10. But if subjects already did not repay anything in the 8 rounds before, the numbers in rounds 9 and 10 could not have retrospectively induced such feelings of fatalism anyway.

Additional tables and figures

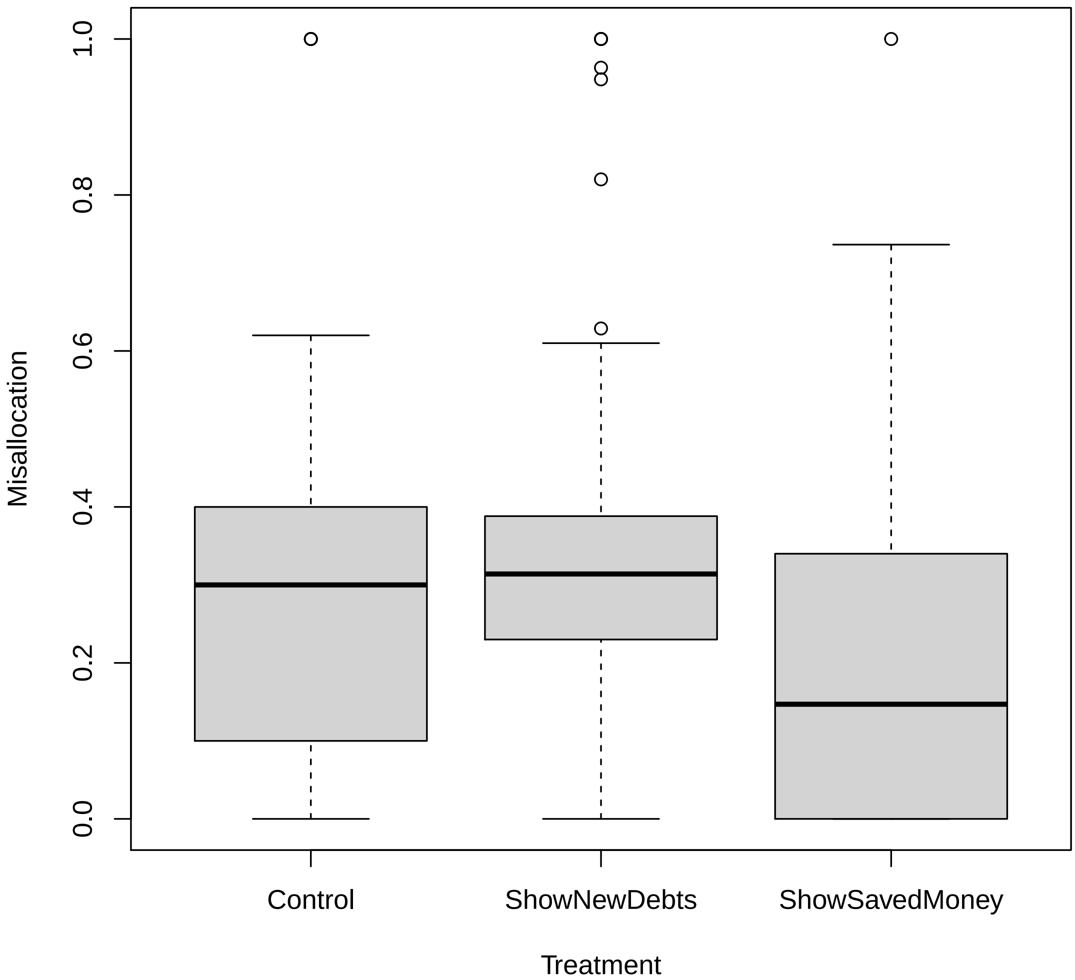


Figure Appendix II.6: Boxplots of misallocation split by treatment. ShowNewDebts is the sludge and ShowSavedMoney is the nudge treatment.

Table Appendix II.23: Duration statistics for the experimental (in minutes)

Duration statistic (min:sec)	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Basic treatment								
Total	131	17:42	08:13	05:53	12:41	15:16	21:08	50:57
Instructions	131	07:38	05:12	02:18	04:17	05:51	09:14	30:18
Experimental stages	131	04:57	03:16	00:45	03:03	04:08	05:38	23:59
Post exp. questionnaire	131	05:07	02:43	01:22	03:23	04:14	05:53	21:13
ShowNewDebts treatment								
Total	135	18:47	08:18	04:57	13:02	16:59	22:46	52:03
Instructions	135	08:19	05:55	01:49	04:35	06:25	10:04	34:41
Experimental stages	135	05:22	02:47	01:41	03:29	04:40	06:08	20:48
Post exp. questionnaire	135	05:07	02:39	01:21	03:20	04:30	06:17	15:29
ShowSavedMoney treatment								
Total	138	20:36	09:06	08:26	13:55	19:23	25:30	58:08
Instructions	138	08:54	06:28	02:41	05:00	07:08	10:11	48:36
Experimental stages	138	05:51	03:00	01:53	03:44	05:14	06:47	17:27
Post exp. questionnaire	138	05:51	03:20	01:01	03:38	05:03	07:14	21:55

Table Appendix II.24: Misallocation split by round class[a]

	Dependent variable: Misallocation		
	Minimal	Akaike-optimal	Full model
	(1)	(2)	(3)
High_int_class	−0.224***	−0.224***	−0.219***
	(0.043)	(0.042)	(0.042)
	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]
ShowNewDebts	0.112	0.102	0.094
	(0.059)	(0.059)	(0.063)
	[0.057]	[0.080]	[0.134]
	[0.113]	[0.240]	[0.269]
ShowSavedMoney	−0.181***	−0.188***	−0.174**
	(0.046)	(0.0456)	(0.053)
	[0.000]	[0.000]	[0.001]
	[0.000]	[0.000]	[0.005]
High_int_class · ShowNewDebts	−0.098	−0.098	−0.098
	(0.062)	(0.061)	(0.063)
	[0.112]	[0.109]	[0.120]
	[0.113]	[0.240]	[0.360]
High_int_class · ShowSavedMoney	0.150**	0.150**	0.139*
	(0.049)	(0.048)	(0.049)
	[0.002]	[0.002]	[0.005]
	[0.006]	[0.007]	[0.019]
Financial literacy		−0.015	−0.023
		(0.009)	(0.016)
		[0.089]	[0.148]
		[0.240]	[0.360]
Years of education (yoe)		−0.011*	−0.012*
		(0.004)	(0.005)
Credit card order (desc.)			−0.024
			(0.022)
Dummy: Male			−0.003
			(0.022)
Age			−0.001
			(0.001)
No. credit card access			−0.007
			(0.009)
No. own credit cards			−0.0004
			(0.005)
Dummy: Use credit card at work			0.007
			(0.030)
Dummy: At work, but usually don't use			−0.035
			(0.066)
Dummy: Usually do not use credit cards			0.007
			(0.032)
ShowNewDebts · Financial literacy			0.011
			(0.023)
ShowSavedMoney · Financial literacy			0.001
			(0.020)
Constant	0.328***	0.523***	0.574***
	(0.040)	(0.083)	(0.095)
Observations	522	522	498
R^2	0.230	0.245	0.246
Akaike Inf. Crit.	5.17	−0.83	17.95

Note: $^{*}p$<0.05; $^{**}p$<0.01; $^{***}p$<0.001 for the Holm-adjusted p-values
Financial literacy is centralized at a value of 3.

[a] High_int_class = 1 if the high interest rate credit card produces more debt in the observed round, High_int_class = 0 otherwise. ShowNewDebts is the sludge and ShowSavedMoney is the nudge treatment. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. P-values adjusted for High_int_class, ShowNewDebts, ShowSavedMoney, High_int_class · ShowNewDebts, High_int_class · ShowSavedMoney and Financial literacy reported coefficients. Asterisks indicate significance after adjustment.

Table Appendix II.25: Comparison of the three treatments via OLS regression, with misallocation as dependent variable.[a]

	Dependent variable: Misallocation		
	Minimal	Akaike-optimal	Full model
	(1)	(2)	(3)
ShowNewDebts	0.038	0.030	0.031
	(0.025)	(0.024)	(0.030)
	[0.131]	[0.212]	[0.303]
	[0.131]	[0.212]	[0.303]
ShowSavedMoney	-0.087^{***}	-0.091^{***}	-0.080^{*}
	(0.023)	(0.022)	(0.029)
	[0.000]	[0.000]	[0.006]
	[0.000]	[0.000]	[0.011]
Financial literacy		-0.046^{***}	-0.044^{**}
		(0.008)	(0.014)
		[0.000]	[0.002]
		[0.000]	[0.006]
Constant	0.274^{***}	0.421^{***}	0.409^{***}
	(0.018)	(0.067)	(0.077)
Observations	404	404	379
Interact. Fin.lit._treatments	No	No	Yes
Further control variables	No	only YOE[b]	Yes
R^2	0.068	0.183	0.189
Akaike Inf. Crit.	−168.51	−217.46	−213.4
Note:	$^{*}p$<0.05; $^{**}p$<0.01; $^{***}p$<0.001 for the Holm-adjusted p-values		

[a] Model 1 includes dummy variables for the respective treatments, model 3 includes all control variables and interactions, model 2 is the AIC-optimal model. ShowNewDebts is the sludge and ShowSavedMoney is the nudge treatment. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for all the reported coefficients, but not the control variables. Asterisks indicate significance after adjustment.
All models show that the misallocation is smaller in the ShowSavedMoney treatment. An increase of one unit in the financial literacy sum index leads to an average decrease in the misallocation by more than 4% in every treatment, so pre-knowledge seems to have an effect on the overall misallocation.

[b] Years of education

Table Appendix II.26: Linear regression of misallocation with a random intercept term for each round[a]

	Dependent variable: Misallocation		
	Minimal	Akaike-optimal	Full model
	(1)	(2)	(3)
Round	0.020***	0.020***	0.020***
	(0.002)	(0.002)	(0.002)
	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]
ShowNewDebts	0.038	0.030	0.031
	(0.024)	(0.023)	(0.026)
	[0.117]	[0.187]	[0.242]
	[0.117]	[0.187]	[0.242]
ShowSavedMoney	−0.087***	−0.091***	−0.080**
	(0.024)	(0.023)	(0.027)
	[0.000]	[0.000]	[0.003]
	[0.000]	[0.000]	[0.006]
Financial literacy		−0.046***	−0.044**
		(0.007)	(0.013)
		[0.000]	[0.001]
		[0.000]	[0.002]
Constant	0.166***	0.313***	0.298***
	(0.019)	(0.065)	(0.079)
Observations	4,040	4,040	3,790
Interact. Fin.lit._treatments	No	No	Yes
Further control variables	No	only YOE[b]	Yes
Akaike Inf. Crit.	1,746.204	1,715.055	1,667.742

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$ for the Holm-adjusted p-values
Financial literacy was centralized at a value of 3.

[a] Model (1) is without further control variables, model (3) contains all control variables and model (2) contains only the variables that are optimal according to Akaike's information criterion from the main analysis in table II.3 (note that the full model has a lower AIC in this particular analysis). ShowNewDebts is the sludge and ShowSavedMoney is the nudge treatment. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for all the reported coefficients, but not the control variables. Asterisks indicate significance after adjustment.

[b] Years of education

Table Appendix II.27: Misallocation split by round class (including screened out participants)[a], OLS regression

	Dependent variable: Misallocation		
	Minimal	Akaike-optimal	Full model
	(1)	(2)	(3)
High_int_class	−0.226***	−0.226***	−0.222***
	(0.042)	(0.041)	(0.042)
	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]
ShowNewDebts	0.110	0.101	0.092
	(0.058)	(0.058)	(0.063)
	[0.060]	[0.084]	[0.142]
	[0.119]	[0.252]	[0.384]
ShowSavedMoney	−0.185***	−0.193***	−0.180**
	(0.045)	(0.045)	(0.052)
	[0.000]	[0.000]	[0.001]
	[0.000]	[0.000]	[0.003]
High_int_class · ShowNewDebts	−0.096	−0.096	−0.095
	(0.062)	(0.061)	(0.063)
	[0.119]	[0.116]	[0.128]
	[0.119]	[0.252]	[0.384]
High_int_class · ShowSavedMoney	0.156**	0.156**	0.145*
	(0.048)	(0.048)	(0.049)
	[0.001]	[0.001]	[0.003]
	[0.004]	[0.004]	[0.012]
Financial literacy		−0.015	−0.023
		(0.009)	(0.016)
		[0.089]	[0.151]
		[0.252]	[0.384]
Years of education (yoe)		−0.011*	−0.011*
		(0.004)	(0.005)
Credit card order (desc.)			−0.024
			(0.022)
Dummy: Male			−0.005
			(0.022)
Age			−0.001
			(0.001)
No. credit card access			−0.006
			(0.009)
No. own credit cards			−0.001
			(0.005)
Dummy: Use credit card at work			0.008
			(0.030)
Dummy: At work, but usually don't use			−0.035
			(0.066)
Dummy: Usually do not use credit cards			0.005
			(0.031)
ShowNewDebts · Financial literacy			0.011
			(0.023)
ShowSavedMoney · Financial literacy			0.002
			(0.020)
Constant	0.329***	0.520***	0.572***
	(0.040)	(0.081)	(0.092)
Observations	528	528	504
R^2	0.232	0.246	0.247
Akaike Inf. Crit.	0.4	−5.32	13.71
Note:	$^{*}p$<0.05;** p<0.01;*** p<0.001 for the Holm-adjusted p-values Financial literacy is centralized at a value of 3.		

[a] High_int_class = 1 if the high interest rate credit card produces more debt in the observed round, High_int_class = 0 otherwise. ShowNewDebts is the sludge and ShowSavedMoney is the nudge treatment. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values for the variables we interpret in brackets. The p-values are adjusted for High_int_class, ShowNewDebts, ShowSavedMoney, High_int_class · ShowNewDebts, High_int_class · ShowSavedMoney and Financial literacy, but not the control variables. Asterisks indicate significance after adjustment.

MTurk legitimization

We replicate the Basic treatment in the experimental econ laboratory of the University of Heidelberg in July 2019 (n=96). The experiment was translated in German and adapted to the lab. Overall we find more misallocation than in the MTurk experiment, although we had higher financial incentives (participation fee: Euro 4, bonus: up to Euro 10). We present the results in Table Appendix II.28 and Figure Appendix II.7. The lab either shows significantly higher misallocation compared to MTurk (in Models (1) and (2)) or no significant difference (in model (3)). Thus, we can conclude that participants on MTurk at least do not perform worse in the sense of higher misallocation than lab participants, despite of clearly lower stakes. This legitimizes the usage of MTurk for this experiment.

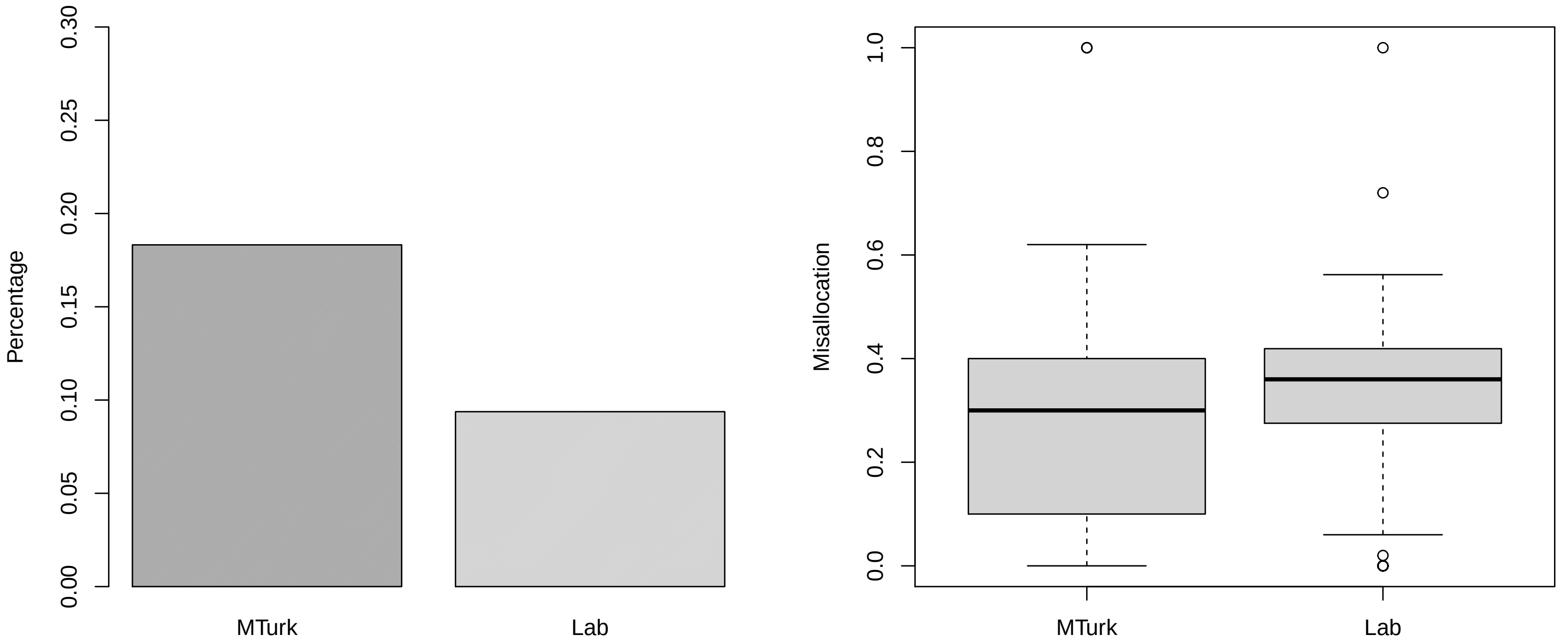


Figure Appendix II.7: The bars in the left Figure show the proportion of subjects without any misallocation in all the experiment rounds. 18.3% of the subjects in MTurk and 9.4% of the subjects in the lab do not exhibit misallocation. The difference is not significant (p=0.0526). The right Figure shows the boxplots of misallocation of all participants on average. The average misallocation in the lab is significantly higher than on MTurk (p = 0.0242).

Table Appendix II.28: Comparison Lab and MTurk via OLS regression, with misallocation as dependent variable.[a]

	Dependent variable: Misallocation		
	Minimal	Akaike-optimal	Full model
	(1)	(2)	(3)
Dummy Laboratory	0.057*	0.060**	0.049
	(0.024)	(0.022)	(0.027)
	[0.018]	[0.007]	[0.065]
	[0.018]	[0.007]	[0.065]
Credit card order (desc.)		0.067**	0.067**
		(0.023)	(0.023)
		[0.004]	[0.004]
		[0.008]	[0.007]
Financial literacy		−0.048***	−0.049***
		(0.010)	(0.010)
		[0.000]	[0.000]
		[0.000]	[0.000]
Constant	0.274***	0.421***	0.451***
	(0.018)	(0.041)	(0.069)
Observations	227	227	227
Further control variables	No	No	Yes
R^2	0.022	0.167	0.170
Akaike Inf. Crit.	−115.79	−148.06	−143.03

Note: $^{*}p$<0.05;** p<0.01;*** p<0.001 for the Holm-adjusted p-values

[a] Model (1) includes only a dummy for the lab, model (3) includes all control variables, model (2) is the AIC-optimal model. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for all the reported coefficients, but not the control variables. Asterisks indicate significance after adjustment.

Theorem

The following theorem proves that repaying the high-interest rate credit card is indeed the debt minimizing way of credit card repayment:

Let there be two credit cards x and y with start balances $b_x, b_y \in \mathbb{R}$ and interest rates $i_x > i_y > 0$, such that x is the high-interest rate credit card. Let there be n repayment rounds following the procedure as proposed in Section II.2. Furthermore, let $a > 0$ be the money available every round on the checking account and let $r_k \in [0, 1]$ for $1 \leq k \leq n$ be the share of money that is repaid on the high-interest rate credit card x. Consequently $1 - r_k$ is the share of money that is repaid on y. Then the overall debt after n rounds are minimized if and only if $r_k = 1 \; \forall \; 1 \leq k \leq n$.

Proof: The overall balance f is the sum of the balances of the credit cards x and y after n rounds depending on the choices of r_k for $1 \leq k \leq n$. Thus, f is a function

$$[0,1]^n \to \mathbb{R} : \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \mapsto b_x i_x^n + \sum_{k=1}^{n} \left(a r_k i_x^{n-k+1} \right) + b_y i_y^n + \sum_{k=1}^{n} \left(a(1 - r_k) i_y^{n-k+1} \right)$$

Therefore,

$$f \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = b_x i_x^n + b_y i_y^n + a \cdot \sum_{k=1}^{n} \left(r_k i_x^{n-k+1} + (1 - r_k) i_y^{n-k+1} \right)$$

and

$$Df = \begin{pmatrix} \frac{\partial f}{\partial r_1} \\ \vdots \\ \frac{\partial f}{\partial r_n} \end{pmatrix} = a \cdot \underbrace{\begin{pmatrix} i_x^n - i_y^n \\ \vdots \\ i_x^1 - i_y^1 \end{pmatrix}}_{>0,\text{ because of } i_x > i_y}.$$

The derivative of f is constant positive, therefore f is strictly increasing in all its components r_k. Thus, f takes on the absolute maximum if and only if $r_k = 1 \; \forall \; 1 \leq k \leq n$. The absolute maximum of the account balances corresponds to a minimum of the debt. Note that this proof also applies for positive account balances, meaning that you maximize your money by investing in an asset with the highest interest rate. □

Appendix III (to Chapter III)

Gathering Data on Amazon Mechanical Turk

Amazon Mechanical Turk (MTurk) is a crowd-sourcing platform on which paid workers work on different tasks by various requesters. Requesters post a HIT ("Human Intelligence Task"), which is usually a series of tasks (e.g. a treatment of our experiment from start to finish is one single HIT), and Turkers decide to join that HIT. Requesters can approve or reject the work of a Turker after the Turker has finished the HIT.

Crowd-sourcing platforms are a relatively new way to conduct experiments, but are becoming more and more common as a convenient sample in the social sciences. As with any convenient sample, their usage is controversial. Sceptics raise concerns about a lack of attention by the Turkers, control problems, too experienced subjects and low external validity (e.g. Chandler et al. (2014, 2015); Ford (2017)). Since most of those concerns are relatively easy to study, Turkers are an extensively and thoroughly examined sample population. Recent papers that discuss potential issues include Chandler and Shapiro (2016), Goodman and Paolacci (2017), Hauser et al. (2019) and Miller et al. (2017). The findings in general seem to imply that the data quality of Turkers is somewhat worse than that of actual representative samples, but outperforms that of common convenient samples such as undergraduates. Turkers from the US seem to resemble the general US population relatively well (Huff and Tingley, 2015; McCredie and Morey, 2018; Paolacci et al., 2010; Snowberg and Yariv, 2021), and especially better than student samples (Snowberg and Yariv (2021); Roulin (2015), however see Krupnikov and Levine (2014)). They produce data of relatively high quality (Kees et al., 2017), with similar noise levels as representative studies (Snowberg and Yariv, 2021), and seem to be more attentive than students (Hauser and Schwarz, 2016; Ramsey et al., 2016). Replications of classical studies of psychology, political sciences and economics are usually successful (e.g. Amir et al. (2012); Berinsky et al. (2012); Coppock (2019); Crump et al. (2013); Horton et al. (2011); Mullinix et al. (2015); Wolfson and Bartkus (2013)), though not every result is replicable – which is not too surprising given the recent replication problems in social sciences and economics (e.g. Camerer et al. (2018); Open Science Collaboration (2015)). However, data quality seems to fall once non-native English speakers are included (e.g. Goodman et al. (2013)), which is why we restrict our sample to the US population. We also require Turkers to have finished at least 100 other HITs with an approval rate of at least 95% (as recommended by Peer et al. (2014), and guarantee that no worker can accept more than one of our HITs by filtering the Turker's IDs. Based on these arguments and the additional checks described in the main text of this paper, we argue that our data is of high quality and performs at least as well as any data we could acquire by using common lab samples.

Additional tables and figures

Table Appendix III.29: Experiment duration in minutes split by treatment[a]

Treatment (min : sec)	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Control	04 : 35	11 : 10	15 : 17	16 : 41	20 : 03	57 : 17
Pamphlet	09 : 31	15 : 13	20 : 38	22 : 35	27 : 05	54 : 12
Slider	06 : 28	12 : 49	17 : 16	19 : 28	23 : 32	61 : 14
Reminder	05 : 41	10 : 55	15 : 02	17 : 05	19 : 56	47 : 31
Assistant	05 : 06	11 : 48	15 : 50	17 : 58	21 : 38	58 : 56

[a] Summary statistics of the duration of the experiment for each treatment. The time is denoted in the format minutes : seconds.

Table Appendix III.30: Descriptive statistics of participants split by treatment

Control (continuous vars)	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Fin. Literacy	132	3.38	1.57	0	2	4	5	6
Age	132	36.97	12.43	19	28	33	44.2	76
Years of education	132	15.75	2.54	9	14	16	17	21
# credit cards	122	2.56	4.70	0	1	2	3	50
# credit access	124	0.81	1.42	0	0	0	1	9
Total payoff	132	2.41	0.46	1.00	2.20	2.42	2.70	3.00
Statistic (count vars)	N	Type	Number		Type	Number		
Gender	132	Female:	68		Male:	64		
Unused but Knowledge	131	Yes:	42		No:	89		
Credit Card Usage at Work	131	Yes:	71		No:	60		

Pamphlet (continuous vars)	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Fin. Literacy	125	3.75	1.31	0	3	4	5	6
Age	125	38.34	12.91	18	28	34	47	82
Years of education	125	15.82	2.25	9	14	16	17	21
# credit cards	120	2.42	2.20	0	1	2	3	13
# credit access	116	0.72	1.10	0	0	0	1	5
Credibility	125	1.30	0.94	-2	1	2	2	2
Total payoff	125	2.80	0.38	1.00	2.76	3.00	3.00	3.00
Statistic (count vars)	N	Type	Number		Type	Number		
Gender	125	Female:	54		Male:	71		
Unused but Knowledge	125	Yes:	37		No:	88		
Credit Card Usage at Work	125	Yes:	76		No:	49		

Slider (continuous vars)	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Fin. Literacy	137	3.82	1.44	0	3	4	5	6
Age	137	38.56	12.24	18	29	37	47	70
Years of education	137	15.63	2.51	9	14	16	17	21
# credit cards	127	2.72	4.87	0	1	2	3	51
# credit access	128	0.84	1.45	0	0	0	1	12
Credibility	137	1.20	1.02	-2	1	2	2	2
Total payoff	137	2.84	0.26	1.51	2.68	3.00	3.00	3.00
Statistic (count vars)	N	Type	Number		Type	Number		
Gender	137	Female:	71		Male:	66		
Unused but Knowledge	137	Yes:	51		No:	86		
Credit Card Usage at Work	137	Yes:	68		No:	69		

Reminder (continuous vars)	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Fin. Literacy	133	3.74	1.35	0	3	4	5	6
Age	133	38.22	12.90	18	30	35	46	79
Years of education	133	15.70	2.22	9	14	16	16	21
# credit cards	120	2.16	1.79	0	1	2	3	10
# credit access	120	0.74	1.51	0	0	0	1	10
Credibility	133	0.87	1.33	-2	0	1	2	2
Total payoff	133	2.84	0.24	1.89	2.74	2.97	3.00	3.00
Statistic (count vars)	N	Type	Number		Type	Number		
Gender	133	Female:	57		Male:	76		
Unused but Knowledge	133	Yes:	54		No:	79		
Credit Card Usage at Work	133	Yes:	57		No:	76		

Assistant (continuous vars)	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Fin. Literacy	133	3.59	1.41	0	3	4	5	6
Age	133	38.04	12.40	18	29	34	47	76
Years of education	133	15.72	2.38	10	14	16	17	21
# credit cards	121	2.47	2.61	0	1	2	3	20
# credit access	116	0.63	1.11	0	0	0	1	5
Credibility	133	0.91	1.20	-2	0	1	2	2
Total payoff	133	2.93	0.18	1.83	3.00	3.00	3.00	3.00
Statistic (count vars)	N	Type	Number		Type	Number		
Gender	133	Female:	54		Male:	79		
Unused but Knowledge	133	Yes:	51		No:	82		
Credit Card Usage at Work	133	Yes:	67		No:	66		

Table Appendix III.31: OLS of the misallocation with different reference categories, w/o control variables[a]

	(1) Control	(2) Control	(3) General interventions	(4) Adapted interventions
	Dependent variable: Misallocation			
Treatment (group)				
Control	Reference	Reference	0.232***	0.275***
	(.)	(.)	(0.022)	(0.021)
	[.]	[.]	[0.000]	[0.000]
	[.]	[.]	[0.000]	[0.000]
Pamphlet	-0.227***			
	(0.026)			
	[0.000]			
	[0.000]			
Slider	-0.236***			
	(0.023)			
	[0.000]			
	[0.000]			
Reminder	-0.252***			
	(0.022)			
	[0.000]			
	[0.000]			
Assistant	-0.297***			
	(0.021)			
	[0.000]			
	[0.000]			
General		-0.232***	Reference	0.043***
		(0.022)	(.)	(0.013)
		[0.000]	[.]	[0.001]
		[0.000]	[.]	[0.000]
Adapted		-0.275***	-0.043***	Reference
		(0.021)	(0.013)	(.)
		[0.000]	[0.001]	[.]
		[0.000]	[0.001]	[.]
Constant	0.341***	0.341***	0.109***	0.066***
	(0.019)	(0.019)	(0.011)	(0.007)
Observations	660	660	660	660
Further controls	No	No	No	No
R-sqr	0.288	0.283	0.283	0.283
F-value	49.999	89.979	89.979	89.979

Note: *p<0.05; **p<0.01; ***p<0.001

[a] This table presents OLS regression results of the mean misallocation of participants. Misallocation serves as dependent variable in all regressions. The first column shows the control group in comparison to all intervention treatments. The other three columns show the same regression with either control, general or adapted intervention as base group. Treatment is the only regressor variable, there are no control variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for all variables per model. Asterisks indicate significance after adjustment.

Table Appendix III.32: OLS regression of the misallocation for all treatments with different reference categories (subjects with suspicious or non-fitting answers to the open question screened out)[a]

	(1) Control	(2) Control	(3) General interventions	(4) Adapted interventions
		Dependent variable: Misallocation		
Treatment (group)				
Control	Reference	Reference	0.212***	0.242***
	(.)	(.)	(0.019)	(0.018)
	[.]	[.]	[0.000]	[0.000]
	[.]	[.]	[0.000]	[0.000]
Pamphlet	-0.223***			
	(0.021)			
	[0.000]			
	[0.000]			
Slider	-0.203***			
	(0.020)			
	[0.000]			
	[0.000]			
Reminder	-0.219***			
	(0.020)			
	[0.000]			
	[0.000]			
Assistant	-0.266***			
	(0.019)			
	[0.000]			
	[0.000]			
General		-0.212***	Reference	0.030**
		(0.019)	(.)	(0.011)
		[0.000]	[.]	[0.008]
		[0.000]	[.]	[0.008]
Adapted		-0.242***	-0.030*	Reference
		(0.018)	(0.011)	(.)
		[0.000]	[0.008]	[.]
		[0.000]	[0.016]	[.]
Financial literacy	-0.055***	-0.055***	-0.028***	-0.003
	(0.014)	(0.014)	(0.007)	(0.006)
Interactions between financial literacy and treatment (group)				
Control × FL	Reference	Reference	-0.026	-0.052**
	(.)	(.)	(0.016)	(0.015)
	[.]	[.]	[0.099]	[0.001]
	[.]	[.]	[0.099]	[0.002]
Pamphlet × FL	0.021			
	(0.018)			
	[0.239]			
	[0.2394]			
Slider × FL	0.030			
	(0.017)			
	[0.082]			
	[0.165]			
Reminder × FL	0.048*			
	(0.018)			
	[0.007]			
	[0.020]			
Assistant × FL	0.057***			
	(0.015)			
	[0.000]			
	[0.000]			

continued on next page ...

... continued from previous page

	(1) Control	(2) Control	(3) General interventions	(4) Adapted interventions
		Dependent variable: Misallocation		
Interactions between financial literacy and treatment (group)				
General × FL		0.026	0.000	-0.026*
		(0.016)	(.)	(0.010)
		[0.099]	[.]	[0.007]
		[0.099]	[.]	[0.014]
Adapted × FL		0.052*	0.026*	0.000
		(0.015)	(0.010)	(.)
		[0.001]	[0.007]	[.]
		[0.001]	[0.022]	[.]
Further control variables				
Age	-0.001*	-0.001*	-0.001*	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)
Years of education	0.002	0.002	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)
Male	0.007	0.006	0.006	0.006
	(0.012)	(0.012)	(0.012)	(0.012)
# credit cards	0.003	0.003	0.003	0.003
	(0.004)	(0.004)	(0.004)	(0.004)
# credit access	0.000	0.001	0.001	0.001
	(0.005)	(0.005)	(0.005)	(0.005)
Cord	-0.006	-0.004	-0.004	-0.004
	(0.012)	(0.012)	(0.012)	(0.012)
Unused but Knowledge	0.015	0.015	0.015	0.015
	(0.016)	(0.016)	(0.016)	(0.016)
Credit Card Usage at Work	0.025	0.023	0.023	0.023
	(0.014)	(0.014)	(0.014)	(0.014)
Constant	0.288***	0.289***	0.077	0.047
	(0.054)	(0.054)	(0.051)	(0.050)
Observations	582	582	582	582
R-sqr	0.400	0.391	0.391	0.391
F-value	15.988	18.126	18.126	18.126

Note: *p<0.05; **p<0.01; ***p<0.001

[a] This table presents OLS regression results of the mean misallocation of participants, but we screen out 13 participants with suspicious or unfitting answers to the open question. Misallocation serves as dependent variable in all regressions. The first column shows the control group in comparison to all intervention treatments. The other three columns show the same regression with either control, general or adapted intervention as base group. Further control variables are age, years of education, gender (reference: female), number of own credit cards, additional accessible credit cards, order of credit card presentation in the experiment (Cord=1 if 5%-card was second), and credit card dummies whether credit cards are generally not used, but known in principle (Unused but Knowledge) and whether they are used at work (Credit Card Usage at Work). Financial literacy (FL) is centralized at the median value 4. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values in model (1) are adjusted for Pamphlet, Slider, Reminder, Assistant, Pamphlet × FL, Slider × FL, Reminder × FL and Assistant × FL. The p-values in models (2-4) are adjusted for Control, General, Adapted, Control × FL, General × FL and Adapted × FL, depending on which four of the six variables are reported. Asterisks indicate significance after adjustment.

Table Appendix III.33: Fractional regression of the misallocation with different reference categories[a]

	(1) Control	(2) Control	(3) General interventions	(4) Adapted interventions
		Dependent variable: Misallocation		
Treatment (group)				
Control	Reference	Reference	0.846***	1.051***
	(.)	(.)	(0.075)	(0.078)
	[.]	[.]	[0.000]	[0.000]
	[.]	[.]	[0.000]	[0.000]
Pamphlet	-0.919***			
	(0.101)			
	[0.000]			
	[0.000]			
Slider	-0.795***			
	(0.088)			
	[0.000]			
	[0.000]			
Reminder	-0.873***			
	(0.088)			
	[0.000]			
	[0.000]			
Assistant	-1.301***			
	(0.116)			
	[0.000]			
	[0.000]			
General		-0.846***	Reference	0.205*
		(0.075)	(.)	(0.083)
		[0.000]	[.]	[0.013]
		[0.000]	[.]	[0.027]
Adapted		-1.051***	-0.205*	Reference
		(0.078)	(0.083)	(.)
		[0.000]	[0.013]	[.]
		[0.000]	[0.040]	[.]
Financial literacy	-0.175***	-0.174***	-0.152***	-0.012
	(0.037)	(0.037)	(0.037)	(0.046)
Interactions between financial literacy and treatment (group)				
Control × FL	Reference	Reference	-0.021	-0.162**
	(.)	(.)	(0.052)	(0.059)
	[.]	[.]	[0.678]	[0.006]
	[.]	[.]	[0.678]	[0.018]
Pamphlet × FL	-0.049			
	(0.065)			
	[0.451]			
	[0.557]			
Slider × FL	0.065			
	(0.060)			
	[0.278]			
	[0.557]			
Reminder × FL	0.133			
	(0.069)			
	[0.054]			
	[0.163]			
Assistant × FL	0.208*			
	(0.073)			
	[0.004]			
	[0.017]			

continued on next page ...

... continued from previous page

	(1) Control	(2) Control	(3) General interventions	(4) Adapted interventions
		Dependent variable: Misallocation		
Interactions between financial literacy and treatment (group)				
General × FL		0.021	Reference	-0.140*
		(0.052)	(.)	(0.059)
		[0.678]	[.]	[0.018]
		[0.678]	[.]	[0.027]
Adapted × FL		0.162*	0.140*	Reference
		(0.059)	(0.059)	(.)
		[0.006]	[0.018]	[.]
		[0.012]	[0.040]	[.]
Further control variables				
Age	-0.006*	-0.006*	-0.006*	-0.006*
	(0.002)	(0.002)	(0.002)	(0.002)
Years of education	0.005	0.003	0.003	0.003
	(0.015)	(0.015)	(0.015)	(0.015)
Male	0.023	0.013	0.013	0.013
	(0.067)	(0.067)	(0.067)	(0.067)
# credit cards	0.015	0.014	0.014	0.014
	(0.009)	(0.010)	(0.010)	(0.010)
# credit access	0.008	0.013	0.013	0.013
	(0.025)	(0.024)	(0.024)	(0.024)
Cord	-0.023	-0.015	-0.015	-0.015
	(0.065)	(0.065)	(0.065)	(0.065)
Unused but Knowledge	0.098	0.104	0.104	0.104
	(0.094)	(0.095)	(0.095)	(0.095)
Credit Card Usage at Work	0.168	0.156	0.156	0.156
	(0.087)	(0.087)	(0.087)	(0.087)
Constant	-0.537	-0.516	-1.362***	-1.567***
	(0.274)	(0.271)	(0.276)	(0.272)
Observations			595	

Note: *p<0.05; **p<0.01; ***p<0.001

[a] This table presents fractional regressions of the mean misallocation of participants. Misallocation serves as dependent variable in all regressions. The first column shows the control group in comparison to all intervention treatments. The other three columns show the same regression with either control, general or adapted intervention as base group. Further control variables are age, years of education, gender (reference: female), number of own credit cards, additional accessible credit cards, order of credit card presentation in the experiment (Cord=1 if 5%-card was second), and credit card dummies whether credit cards are generally not used, but known in principle (Unused but Knowledge) and whether they are used at work (Credit Card Usage at Work). Financial literacy (FL) is centralized at the median value 4. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values in model (1) are adjusted for Pamphlet, Slider, Reminder, Assistant, Pamphlet × FL, Slider × FL, Reminder × FL and Assistant × FL. The p-values in models (2-4) are adjusted for Control, General, Adapted, Control × FL, General × FL and Adapted × FL, depending on which four of the six variables are reported. Asterisks indicate significance after adjustment.

Table Appendix III.34: Logistic regression of optimal repaying subjects with different reference categories[a]

	(1) Control	(2) Control	(3) General interventions	(4) Adapted interventions
		Dependent variable: Misallocation		
Treatment (group)				
Control	Reference	Reference	-2.244***	-2.382***
	(.)	(.)	(0.307)	(0.308)
	[.]	[.]	[0.000]	[0.000]
	[.]	[.]	[0.000]	[0.000]
Pamphlet	2.438***			
	(0.347)			
	[0.000]			
	[0.000]			
Slider	2.087***			
	(0.336)			
	[0.000]			
	[0.000]			
Reminder	1.681***			
	(0.331)			
	[0.000]			
	[0.000]			
Assistant	3.286***			
	(0.374)			
	[0.000]			
	[0.000]			
General		2.244***	Reference	-0.138
		(0.307)	(.)	(0.201)
		[0.000]	[.]	[0.491]
		[0.000]	[.]	[0.491]
Adapted		2.382***	0.138	Reference
		(0.308)	(0.201)	(.)
		[0.000]	[0.491]	[.]
		[0.000]	[0.982]	[.]
Financial literacy	0.451*	0.437*	0.490***	0.118
	(0.223)	(0.221)	(0.117)	(0.104)
Interactions between financial literacy and treatment (group)				
Control × FL	Reference	Reference	-0.053	0.319
	(.)	(.)	(0.247)	(0.242)
	[.]	[.]	[0.831]	[0.187]
	[.]	[.]	[0.982]	[0.374]
Pamphlet × FL	0.114			
	(0.284)			
	[0.688]			
	[1.000]			
Slider × FL	0.011			
	(0.267)			
	[0.968]			
	[1.000]			
Reminder × FL	-0.237			
	(0.268)			
	[0.375]			
	[1.000]			
Assistant × FL	-0.371			
	(0.268)			
	[0.167]			
	[0.667]			

continued on next page ...

... continued from previous page

	(1) Control	(2) Control	(3) General interventions	(4) Adapted interventions
		Dependent variable: Misallocation		
Interactions between financial literacy and treatment (group)				
General × FL		0.053	Reference	0.372*
		(0.247)	(.)	(0.154)
		[0.831]	[.]	[0.015]
		[0.831]	[.]	[0.046]
Adapted × FL		-0.319	-0.372*	Reference
		(0.242)	(0.154)	(.)
		[0.187]	[0.015]	[.]
		[0.374]	[0.046]	[.]
Further control variables				
Age	0.014	0.013	0.013	0.013
	(0.008)	(0.008)	(0.008)	(0.008)
Years of education	-0.031	-0.026	-0.026	-0.026
	(0.043)	(0.041)	(0.041)	(0.041)
Male	-0.312	-0.220	-0.220	-0.220
	(0.206)	(0.198)	(0.198)	(0.198)
# credit cards	0.016	0.018	0.018	0.018
	(0.030)	(0.031)	(0.031)	(0.031)
# credit access	-0.090	-0.103	-0.103	-0.103
	(0.078)	(0.071)	(0.071)	(0.071)
Cord	0.099	0.040	0.040	0.040
	(0.194)	(0.186)	(0.186)	(0.186)
Unused but Knowledge	-0.451	-0.427	-0.427	-0.427
	(0.288)	(0.275)	(0.275)	(0.275)
Credit Card Usage at Work	-0.435	-0.329	-0.329	-0.329
	(0.269)	(0.255)	(0.255)	(0.255)
Constant	-1.243	-1.371	0.873	1.011
	(0.809)	(0.789)	(0.761)	(0.771)
Observations			595	

Note: *p<0.05; **p<0.01; ***p<0.001

[a] This table presents logistic regressions of a dummy variable whether participants repay optimally (=1 if mean misallocation is zero, =0 otherwise). Misallocation serves as dependent variable in all regressions. The first column shows the control group in comparison to all intervention treatments. The other three columns show the same regression with either control, general or adapted intervention as base group. Further control variables are age, years of education, gender (reference: female), number of own credit cards, additional accessible credit cards, order of credit card presentation in the experiment (Cord=1 if 5%-card was second), and credit card dummies whether credit cards are generally not used, but known in principle (Unused but Knowledge) and whether they are used at work (Credit Card Usage at Work). Financial literacy (FL) is centralized at the median value 4. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values in model (1) are adjusted for Pamphlet, Slider, Reminder, Assistant, Pamphlet × FL, Slider × FL, Reminder × FL and Assistant × FL. The p-values in models (2-4) are adjusted for Control, General, Adapted, Control × FL, General × FL and Adapted × FL, depending on which four of the six variables are reported. Asterisks indicate significance after adjustment.

Table Appendix III.35: OLS of the misallocation with different reference categories (incl. Credibility)[a]

	(1) Pamphlet	(2) Slider	(3) Reminder	(4) Assistant	(5) General interventions	(6) Adapted interventions
			Dependent variable: Misallocation			
Treatment (group)						
Pamphlet	Reference	0.015	0.042	0.095**		
	(.)	(0.033)	(0.030)	(0.030)		
	[.]	[0.652]	[0.169]	[0.002]		
	[.]	[0.652]	[0.338]	[0.003]		
Slider	-0.015	Reference	0.027	0.080**		
	(0.033)	(.)	(0.023)	(0.023)		
	[0.652]	[.]	[0.250]	[0.001]		
	[0.652]	[.]	[0.338]	[0.002]		
Reminder	-0.042	-0.027	Reference	0.054**		
	(0.030)	(0.023)	(.)	(0.020)		
	[0.169]	[0.250]	[.]	[0.007]		
	[0.338]	[0.499]	[.]	[0.007]		
Assistant	-0.095**	-0.080**	-0.054*	Reference		
	(0.030)	(0.023)	(0.020)	(.)		
	[0.002]	[0.001]	[0.007]	[.]		
	[0.005]	[0.002]	[0.020]	[.]		
General					Reference	0.060**
					(.)	(0.018)
Adapted					-0.060**	Reference
					(0.018)	(.)
Credibility	-0.059***	-0.042***	-0.044***	-0.038***	-0.050***	-0.041***
	(0.014)	(0.012)	(0.008)	(0.009)	(0.009)	(0.006)
Financial literacy	-0.035***	-0.020	-0.015	-0.003	-0.026***	-0.008
	(0.010)	(0.010)	(0.009)	(0.004)	(0.008)	(0.005)
Interactions between Credibility and treatment (group)						
Pamphlet × Credibility	Reference	-0.017	-0.016	-0.021		
	(.)	(0.018)	(0.016)	(0.017)		
Slider × Credibility	0.017	Reference	0.002	-0.004		
	(0.018)	(.)	(0.014)	(0.015)		
Reminder × Credibility	0.016	-0.002	Reference	-0.006		
	(0.016)	(0.014)	(.)	(0.012)		
Assistant × Credibility	0.021	0.004	0.006	Reference		
	(0.017)	(0.015)	(0.012)	(.)		
General × Credibility					Reference	-0.009
					(.)	(0.011)
Adapted × Credibility					0.009	Reference
					(0.011)	(.)
Further control variables						
Age	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years of education	0.001	0.001	0.001	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Male	0.004	0.004	0.004	0.004	0.002	0.002
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
# credit cards	0.005	0.005	0.005	0.005	0.005	0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
# credit access	-0.003	-0.003	-0.003	-0.003	-0.002	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Cord	-0.006	-0.006	-0.006	-0.006	-0.004	-0.004
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Unused but Knowledge	0.015	0.015	0.015	0.015	0.015	0.015
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Credit Card Usage at Work	0.025*	0.025*	0.025*	0.025*	0.024	0.024
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Constant	0.137*	0.122*	0.096*	0.042	0.130*	0.071
	(0.057)	(0.052)	(0.048)	(0.050)	(0.052)	(0.047)
Observations	476	476	476	476	476	476
Interactions FL and treatment (group)	Yes	Yes	Yes	Yes	Yes	Yes
R-sqr	0.261	0.261	0.261	0.261	0.235	0.235
F-value	9.842	9.842	9.842	9.842	10.438	10.438

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table presents OLS regression results of the mean misallocation of participants. Misallocation serves as dependent variable in all regressions. In order to analyze the additional variable Credibility (a Likert scale from -2 to 2 on how convinced participants were by our suggested strategy) we exclude the control group from this table, since it cannot be measured there. The columns (1)-(4) show the very same regression, but with a different treatment group as reference. The columns (5) and (6) show the regressions for the general and the adapted intervention group. Further control variables are age, years of education, gender (reference: female), number of own credit cards, additional accessible credit cards, order of credit card presentation in the experiment (Cord=1 if 5%-card was second), and credit card dummies whether credit cards are generally not used, but known in principle (Unused but Knowledge) and whether they are used at work (Credit Card Usage at Work). Financial literacy is centralized at the median value 4. All regressions also include interaction terms between financial literacy and treatment (group), but for brevity we do not report these terms. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values in models (1-4) are adjusted for Pamphlet, Slider, Reminder and Assistant, depending on which three of the four coefficients are reported. In models 5 and 6, adjustment is not needed. Asterisks indicate significance after adjustment.

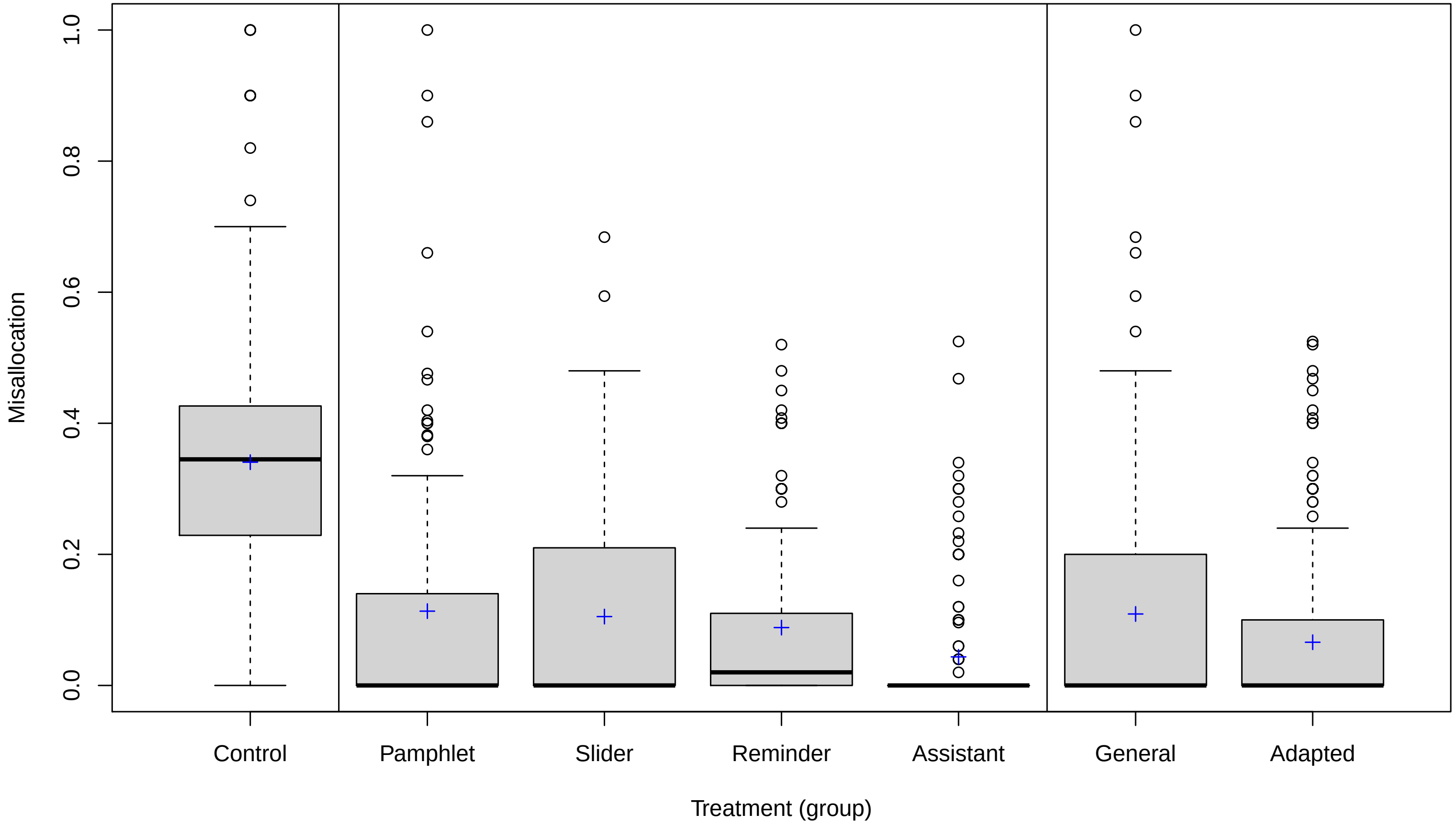


Figure Appendix III.8: This figure shows boxplots of the mean misallocation of all participants split by treatment and intervention group. The thick black line within each boxplot denotes the median, the blue "+"-symbol indicates the mean. The maximum length of the whiskers is 1.5 times the interquartile range. All participants with a mean misallocation outside the range of the whiskers are shown as separate points.

Pamphlet

Efficient credit card repayment

Imagine you have debts on several credit cards, and for each card, you have to pay a different interest rate. How should you repay the debt, if you want to save as much interest payments as possible? This is how you do it:

First, you should cover all the required minimum payments on each card and after that, you should try to stay within the limits of your credit card contract. When this has been done and you still have some money left, the optimal way to repay the money is to **always settle the debts on the credit card with the <u>highest interest rate</u> first** before you even touch any other cards. Only if you have fully repaid the debt on that card, you should start settling the debts on the card with the second highest interest rate, while still ignoring the debts on all the other cards. This is how you proceed with each credit card until the one with the lowest interest rate is left.

<u>Why is that?</u>

You can think of the interest rate as the amount of cents that you have to pay per dollar in the next period. If the interest rate of a credit card is 10% annually, you have to pay 10 cents per dollar every year. Another way of looking at this is to bear in mind that for every dollar you repay you save 10 cents per year, because the repayment will no longer be subject to the next interest payment. So basically, the interest rate of a credit card tells you how much money you can save by repaying your debts on that card – the higher the interest rate is, the more you save when settling the debts.

<u>An example:</u>

Let us assume you have two credit cards. The monthly interest rate of the first credit card is 5%, while the rate of the other is only 1%. For each dollar on the first credit card that is repaid, the interest payment falls by 5 cent per month. However, for the second credit card the interest savings per $ 1 that is repaid will be 1 cent only. Therefore, the interest savings will be highest if and only if you fully settle the debts on the 5% credit card first before starting to repay the debts 1% card.

Many people do not consider this rule when repaying debts, and thus they pay more than they have to. Below are two typical examples of how people handle their debts, even though they lose money in the process:

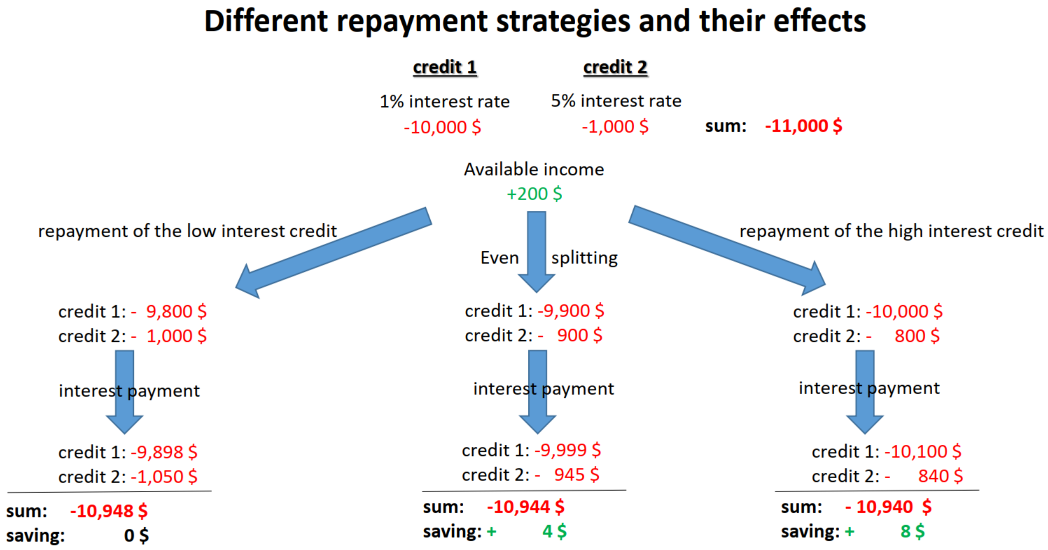
<u>Costly repayment 1: Inclusion of debt levels</u>

Many people tend to take into account the amount of debt on a credit card, especially if the amounts vary considerably. So if there is a debt of 1000$ on one card, and only 100$ on the other, people will first settle the amount of 1000$ although the 100$-card may be more expensive in terms of interest rates. This may be a reasonable decision if one has exhausted the limit on a credit card and would have to pay extra money or would otherwise get into trouble for crossing the limit. Nevertheless, many people think that the amount of debt matters even if they are well within their limits, and as a result, they pay more interest than they have to.

Costly repayment 2: Repay several cards simultaneously

Many people settle the debts on several credit cards simultaneously although the interest rates are different. Again, this procedure generates less savings than a full repayment of the debt on the card with the highest interest rate would.

These above-described two examples are illustrated in the graph below. We start with two credit cards. The cheaper card has a 1% interest rate per month, the other card has an interest rate of 5%. We start with 10,000$ of debts on the cheaper card and with 1,000 on the 5% card. In the next step we repay 200$ of our debts. We now present three possible ways to repay the debt, each showing the debt that is left after the payment.

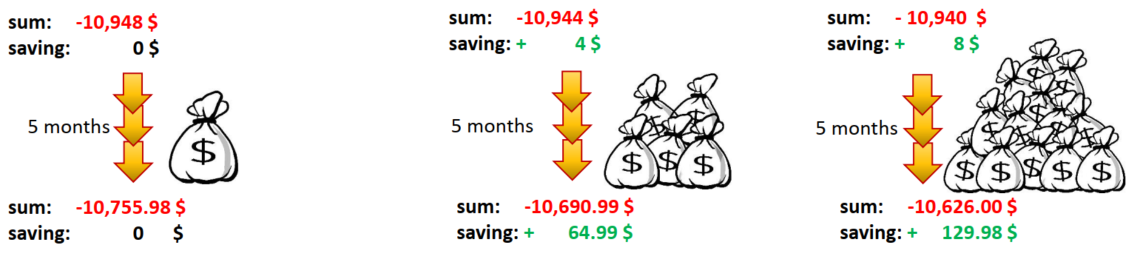


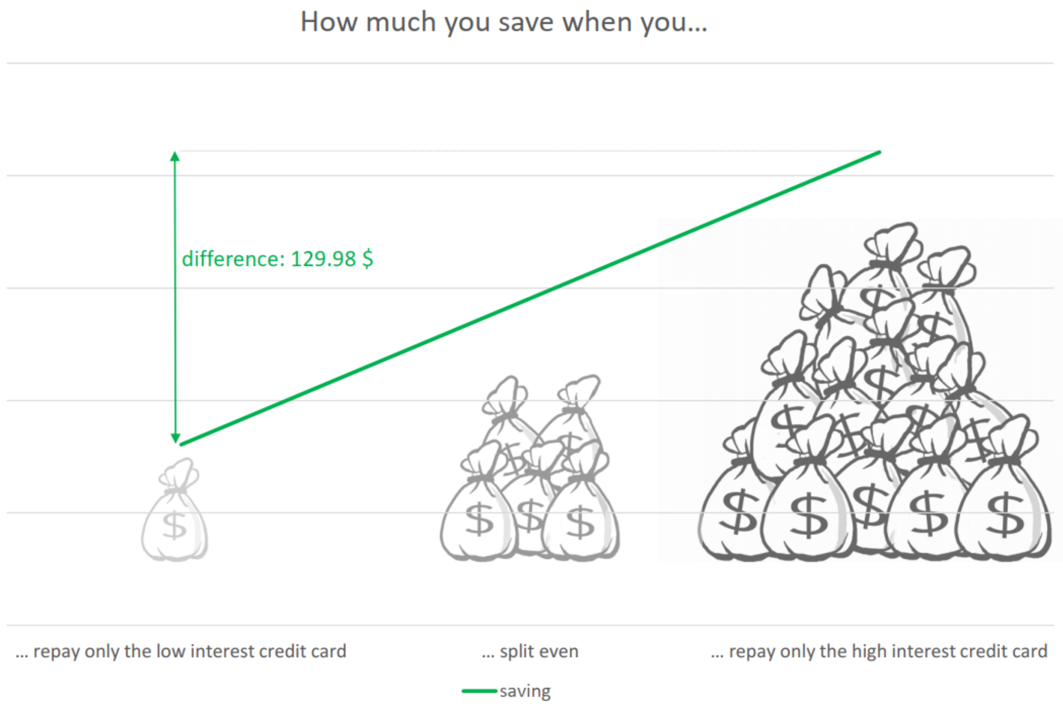
The example on the left shows what happens if we repay the highest debt first: we use the full amount of 200$ to repay part of the debt on the 10,000$-card and we leave the 1,000$-card untouched. If we calculate with the interest, 10,948$ are left after that month.

The example on the right shows what happens if we repay the debt on the card with the high interest rate first: we use the full amount of 200$ to repay the debt on the 5%-card, and ignore the debt on the 1%-card. After calculating the interests, we are left with 10,940$. This is 8$ less than what we would obtain in the example on the left.

The example in the middle shows what happens if we split the money into equal parts and repay 100$ of the debt on each card. The month after that we are left with debts of 10,944$. This is exactly the middle of the other two examples and instead of 8$ we only save 4$.

These differences do not seem very significant at first, but consider what happens after five months: if we choose the left example and repay 200$ every month for five ongoing months we will still be left with 10,755.98$ of debts. If we choose the right example instead and pay the debt on the card with the highest interest rate first, we save almost 130$ in interests. This is due to the compound interest we now no longer have to pay.





<u>Conclusion:</u>

When repaying debts, make sure that you make all the minimum repayments first and that you stay within your limits. After doing so only look at the interest rate (the APR) and repay the debt on the cards starting with the one with the lowest interest rate and ending with the one with the highest interest rate. Ignore any urge to split the amount of repayment or to take into account the balances, and you will have more money left. This payment advice also holds true for any other type of credit or loan such as mortgages, student debts or car loans.

Appendix IV (to Chapter IV)

Additional tables and figures

Experiment #1

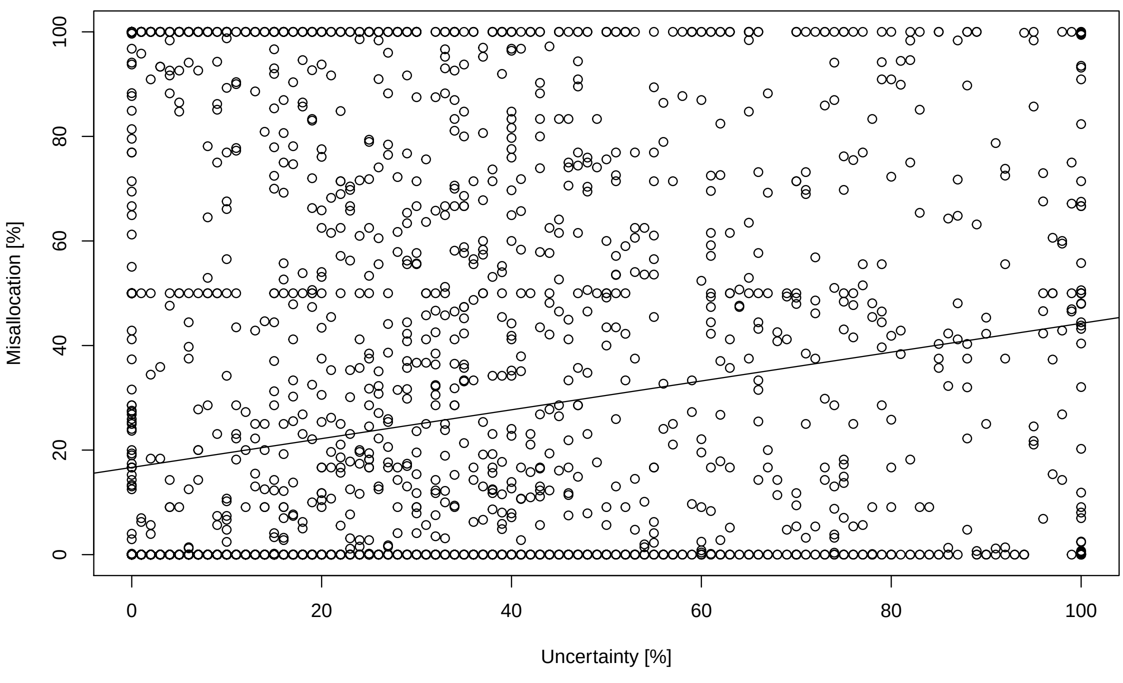


Figure Appendix IV.9: This scatter plot shows the percentage points of uncertainty (x-axis) in relation to misallocation (y-axis). The additional OLS-regression line shows that uncertainty and misallocation correlate positively.

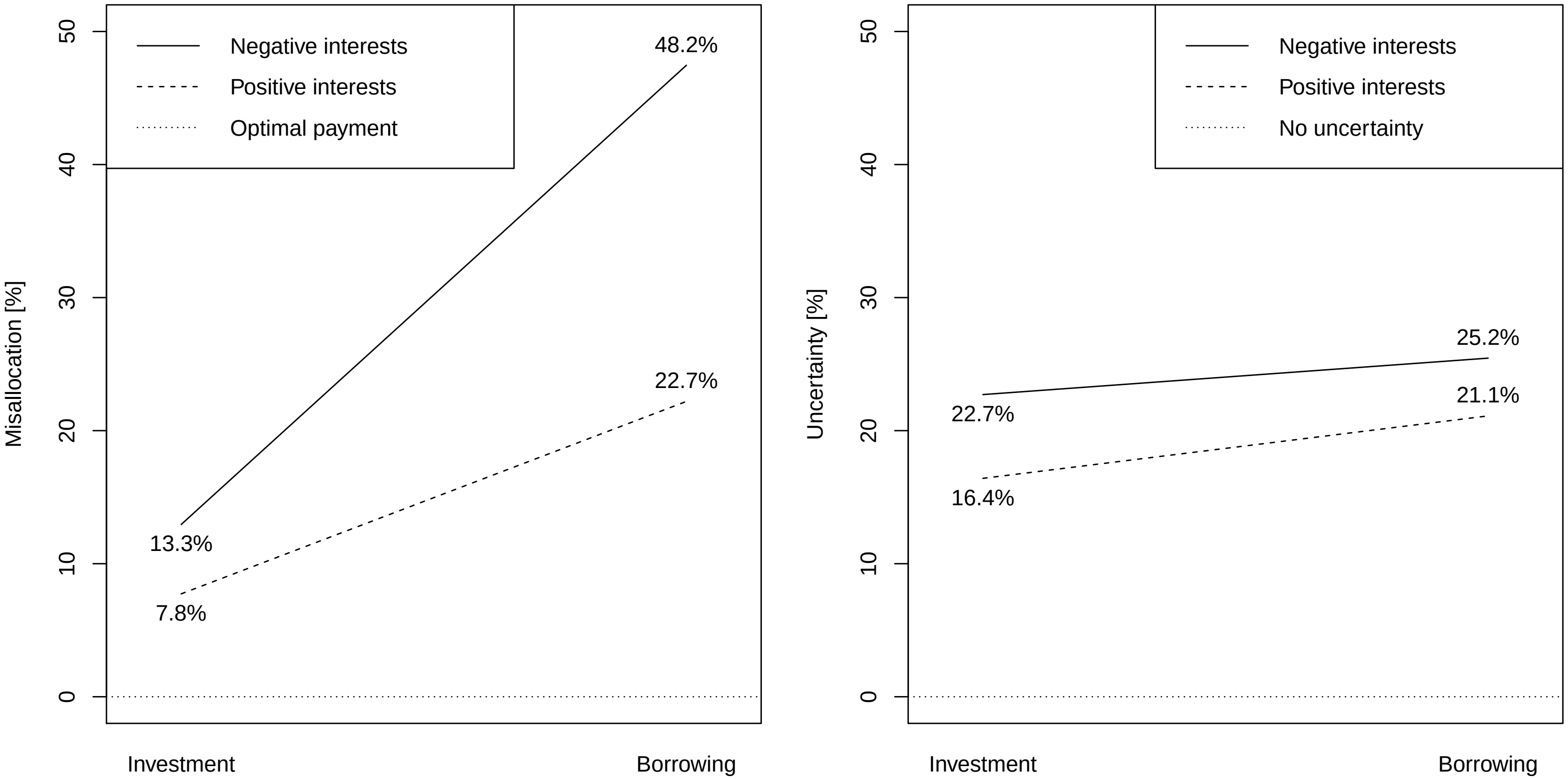


Figure Appendix IV.10: These interaction plots show the average values for misallocation (left graphic) and uncertainty (right graphic) in experiment #1 split by treatment. We differentiate by the borrowing variable (x-axis in both graphics) and the negative int. rates variable (line types in both graphics). For a better overview we do not include the percentage frame dummy and instead average out its effects, as they are not significant in the regression models.

Table Appendix IV.36: Random effects regression showing all used variables[a]

	DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	5.669***	17.329***	5.674***	17.325***	6.915***	13.702***
	(0.939)	(2.225)	(0.937)	(2.213)	(1.150)	(2.524)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.735***	4.652***	5.750***	4.673***	6.082***	3.441
	(0.877)	(1.383)	(0.879)	(1.382)	(0.930)	(1.556)
	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.027]
	[0.000]	[0.002]	[0.000]	[0.002]	[0.000]	[0.081]
Percentage frame	-0.885	1.630*	-0.451	-0.653	0.981	-2.819
	(0.638)	(0.781)	(0.968)	(1.737)	(0.902)	(2.083)
	[0.166]	[0.037]	[0.641]	[0.707]	[0.276]	[0.176]
	[0.166]	[0.037]	[0.641]	[0.707]	[0.276]	[0.176]
Uncertainty		0.111*		0.112*		0.084
		(0.040)		(0.040)		(0.049)
		[0.006]		[0.005]		[0.088]
		[0.012]		[0.011]		[0.175]
Borrowing × Negative int. rates	-2.542*	20.034***	-2.579*	19.995***	-2.550	21.162***
	(1.263)	(3.450)	(1.269)	(3.436)	(1.544)	(4.116)
Borrowing × Percentage frame	-1.952	-6.777***	-1.938	-6.793***	-2.519*	-6.888**
	(1.051)	(1.706)	(1.047)	(1.704)	(1.246)	(2.096)
Negative int. rates × Percentage frame	1.121	-0.271	1.132	-0.339	0.466	0.121
	(0.967)	(1.583)	(0.966)	(1.586)	(0.944)	(1.924)
Triple interaction	1.173	0.464	1.193	0.540	1.305	-0.307
	(1.621)	(3.157)	(1.620)	(3.146)	(2.046)	(3.838)
Round			0.112*	0.026	0.169*	-0.004
			(0.057)	(0.107)	(0.067)	(0.126)
Right 2nd			0.065	0.439	0.102	-0.510
			(0.418)	(0.999)	(0.479)	(1.128)
Starkness			0.002	-0.033	0.005	-0.048
			(0.011)	(0.023)	(0.011)	(0.028)
Starkness × Percentage frame			-0.008	0.043	-0.022	0.060
			(0.014)	(0.032)	(0.015)	(0.039)
Age					0.031	0.079
					(0.108)	(0.087)
Female					-3.157	-5.161*
					(3.392)	(2.422)
Third gender					13.385	2.518
					(11.071)	(7.579)
Has credit card debts					-1.150	-2.586
					(3.152)	(2.389)
# of yearly credit transactions					-0.000***	0.000***
					(0.000)	(0.000)
# of yearly investment transactions					-0.043	0.023
					(0.042)	(0.045)
Risk seek					-0.326	-0.272
					(0.364)	(0.287)
Years of education					0.960	-0.612
					(0.597)	(0.489)
Financial Literacy					-4.805**	-3.651***
					(1.751)	(1.071)
Numeracy					-2.548	-1.937*
					(1.403)	(0.881)
Cons. Confidence					-4.857***	-0.095
					(1.467)	(1.161)
Pref. num. info.					-3.417	-2.176
					(2.216)	(1.250)
Constant	16.858***	5.099***	15.753***	6.354***	78.036***	62.887***
	(1.423)	(1.082)	(1.607)	(1.896)	(16.789)	(12.154)
Observations	3840	3840	3840	3840	2624	2624
# participants	240	240	240	240	164	164
Mediation analysis of uncertainty mediating misallocation - Sobel test						
Borrowing	0.629*		0.635*		0.581	
	(0.254)		(0.254)		(0.358)	
	[0.013]		[0.012]		[0.105]	
	[0.036]		[0.034]		[0.306]	
Negative int. rates	0.636*		0.644*		0.511	
	(0.253)		(0.254)		(0.312)	
	[0.012]		[0.011]		[0.102]	
	[0.036]		[0.034]		[0.102]	
Percentage frame	-0.098		-0.051		0.082	
	(0.083)		(0.117)		(0.100)	
	[0.239]		[0.665]		[0.411]	
	[0.239]		[0.665]		[0.411]	

Note: * p<0.05;** p<0.01;*** p<0.001

[a] This table shows the regression results for uncertainty and misallocation under divisible money, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates and percentage frame, as well as uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

Table Appendix IV.37: Random effects logistic regression[a]

	DIVISIBLE		
	Dep. var.: dummymisallo (= 1 if misallo is greater than 0, = 0 otherwise)		
	(1)	(2)	(3)
Borrowing	1.284***	1.297***	1.072**
	(0.261)	(0.260)	(0.360)
	[0.000]	[0.000]	[0.003]
	[0.000]	[0.000]	[0.009]
Negative int. rates	0.048	0.048	0.170
	(0.266)	(0.267)	(0.339)
	[0.856]	[0.859]	[0.616]
	[0.937]	[1.000]	[0.616]
Percentage frame	0.145	-0.050	-0.413
	(0.201)	(0.254)	(0.347)
	[0.468]	[0.844]	[0.234]
	[0.937]	[1.000]	[0.468]
Uncertainty	0.038***	0.038***	0.039***
	(0.006)	(0.006)	(0.008)
	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]
Borrowing × Negative int. rates	1.752***	1.741***	1.614**
	(0.412)	(0.411)	(0.511)
Borrowing × Percentage frame	-0.761**	-0.780**	-0.679*
	(0.253)	(0.254)	(0.344)
Negative int. rates × Percentage frame	0.278	0.275	0.311
	(0.267)	(0.267)	(0.340)
Triple interaction	-0.005	0.019	0.071
	(0.373)	(0.372)	(0.461)
Round		-0.015	-0.013
		(0.014)	(0.017)
Right 2nd		0.035	0.006
		(0.108)	(0.142)
Starkness		-0.005	-0.005
		(0.003)	(0.004)
Starkness × Percentage frame		0.004	0.005
		(0.004)	(0.005)
Constant	-3.086***	-2.718***	6.324*
	(0.292)	(0.350)	(2.457)
Observations	3840	3840	2624
# participants	240	240	164
Individual control variables	No	No	Yes
Mediation analysis of uncertainty mediating misallocation - Sobel test			
Borrowing	0.215***	0.216***	0.269***
	(0.049)	(0.049)	(0.070)
	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]
Negative int. rates	0.218***	0.218***	0.237***
	(0.047)	(0.047)	(0.060)
	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]
Percentage frame	-0.034	-0.017	0.038
	(0.025)	(0.037)	(0.037)
	[0.180]	[0.646]	[0.296]
	[0.180]	[0.646]	[0.296]

Note: $^{*}p<0.05$;$^{**}p<0.01$;$^{***}p<0.001$

[a] This table shows the regression results for misallocation (as dummy variable) under divisible money with three different models: The simple model (1) which includes only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the model (2) which includes some technical aspects of the experiment; and the complete model (3) with all control variables. We omit regressions for uncertainty as they are identical to the main regressions. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates and percentage frame, as well as uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

Table Appendix IV.38: Random effects regression including screened out subjects[a]

	DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	5.491***	17.944***	5.476***	17.961***	6.587***	14.407***
	(0.870)	(2.128)	(0.864)	(2.114)	(1.056)	(2.395)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.519***	4.963***	5.523***	4.975***	5.854***	3.906*
	(0.837)	(1.343)	(0.840)	(1.337)	(0.852)	(1.545)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.011]
	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.034]
Percentage frame	-0.915	2.153*	-0.532	0.883	0.899	-0.788
	(0.597)	(0.835)	(0.952)	(1.796)	(0.953)	(2.154)
	[0.126]	[0.010]	[0.576]	[0.623]	[0.345]	[0.714]
	[0.126]	[0.010]	[0.576]	[0.623]	[0.345]	[0.714]
Uncertainty		0.105*		0.106**		0.090
		(0.038)		(0.038)		(0.047)
		[0.005]		[0.005]		[0.054]
		[0.011]		[0.009]		[0.108]
Borrowing × Negative int. rates	-2.394*	19.905***	-2.391*	19.852***	-2.303	20.925***
	(1.192)	(3.346)	(1.194)	(3.331)	(1.430)	(3.981)
Borrowing × Percentage frame	-1.600	-7.339***	-1.555	-7.363***	-2.017	-7.172***
	(1.002)	(1.714)	(0.997)	(1.703)	(1.193)	(2.088)
Negative int. rates × Percentage frame	1.130	0.984	1.150	0.958	0.656	1.157
	(0.902)	(1.551)	(0.901)	(1.556)	(0.863)	(1.888)
Triple interaction	0.945	-0.797	0.918	-0.728	1.066	-2.071
	(1.524)	(3.048)	(1.519)	(3.037)	(1.883)	(3.715)
Round			0.113*	0.019	0.170**	0.005
			(0.055)	(0.104)	(0.063)	(0.123)
Right 2nd			0.105	0.350	0.115	-0.698
			(0.394)	(0.967)	(0.448)	(1.073)
Starkness			0.006	-0.024	0.011	-0.029
			(0.011)	(0.023)	(0.011)	(0.028)
Starkness × Percentage frame			-0.008	0.024	-0.022	0.037
			(0.014)	(0.031)	(0.015)	(0.038)
Constant	16.962***	5.538***	15.645***	6.439***	69.735***	64.213***
	(1.346)	(1.044)	(1.541)	(1.903)	(14.636)	(11.410)
Observations	4240	4240	4240	4240	2928	2928
# participants	265	265	265	265	183	183
Individual control variables	No	No	No	No	Yes	Yes

Mediation analysis of uncertainty mediating misallocation - Sobel test

	(1)–(2)	(3)–(4)	(5)–(6)
Borrowing	0.579*	0.582*	0.595
	(0.229)	(0.229)	(0.327)
	[0.012]	[0.011]	[0.069]
	[0.033]	[0.031]	[0.198]
Negative int. rates	0.582*	0.587*	0.528
	(0.229)	(0.229)	(0.287)
	[0.011]	[0.010]	[0.066]
	[0.033]	[0.031]	[0.198]
Percentage frame	-0.096	-0.057	0.081
	(0.075)	(0.109)	(0.106)
	[0.200]	[0.605]	[0.442]
	[0.200]	[0.605	[0.442]

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the regression results for uncertainty and misallocation under divisible money, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates and percentage frame, as well as uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

Table Appendix IV.39: Random effects regression excluding subjects in the lower and upper 2.5% quantile of experiment duration[a]

	DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	5.741***	17.255***	5.741***	17.247***	6.984***	13.886***
	(0.984)	(2.269)	(0.982)	(2.260)	(1.212)	(2.595)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.969***	4.412**	5.978***	4.434**	6.359***	3.304
	(0.913)	(1.425)	(0.914)	(1.424)	(0.968)	(1.611)
	[0.000]	[0.002]	[0.000]	[0.002]	[0.000]	[0.040]
	[0.000]	[0.006]	[0.000]	[0.006]	[0.000]	[0.121]
Percentage frame	-0.827	1.702*	-0.515	-0.618	0.993	-2.855
	(0.666)	(0.820)	(1.018)	(1.797)	(0.955)	(2.172)
	[0.215]	[0.038]	[0.613]	[0.731]	[0.298]	[0.189]
	[0.215]	[0.038]	[0.613]	[0.731]	[0.298]	[0.189]
Uncertainty		0.122**		0.123**		0.089
		(0.041)		(0.041)		(0.051)
		[0.003]		[0.003]		[0.078]
		[0.006]		[0.006]		[0.156]
Borrowing × Negative int. rates	-2.522	19.985***	-2.543	19.948***	-2.518	20.669***
	(1.320)	(3.557)	(1.324)	(3.544)	(1.621)	(4.233)
Borrowing × Percentage frame	-2.039	-6.982***	-2.028	-6.990***	-2.593*	-7.277***
	(1.103)	(1.762)	(1.099)	(1.763)	(1.314)	(2.178)
Negative int. rates × Percentage frame	1.099	-0.332	1.097	-0.399	0.449	0.051
	(1.012)	(1.631)	(1.013)	(1.632)	(0.992)	(1.963)
Triple interaction	1.081	0.496	1.094	0.569	1.295	-0.159
	(1.700)	(3.223)	(1.697)	(3.212)	(2.160)	(3.883)
Round			0.113	0.043	0.167*	0.020
			(0.059)	(0.111)	(0.071)	(0.130)
Right 2nd			0.092	0.424	0.154	-0.585
			(0.439)	(1.021)	(0.504)	(1.165)
Starkness			0.002	-0.032	0.005	-0.047
			(0.012)	(0.024)	(0.012)	(0.029)
Starkness × Percentage frame			-0.006	0.044	-0.021	0.062
			(0.015)	(0.032)	(0.016)	(0.040)
Constant	17.083***	4.910***	15.997***	5.981**	84.075***	62.205***
	(1.469)	(1.111)	(1.674)	(1.953)	(16.762)	(12.627)
Observations	3648	3648	3648	3648	2480	2480
# participants	228	228	228	228	155	155
Individual control variables	No	No	No	No	Yes	Yes
Mediation analysis of uncertainty mediating misallocation - Sobel test						
Borrowing	0.699*		0.704*		0.622	
	(0.268)		(0.268)		(0.374)	
	[0.009]		[0.008]		[0.096]	
	[0.022]		[0.021]		[0.276]	
Negative int. rates	0.727*		0.733*		0.567	
	(0.272)		(0.272)		(0.337)	
	[0.007]		[0.007]		[0.092]	
	[0.011]		[0.021]		[0.276]	
Percentage frame	-0.101		-0.063		0.089	
	(0.092)		(0.133)		(0.110)	
	[0.274]		[0.635]		[0.421]	
	[0.274]		[0.635]		[0.421]	

Note: $^{*}p$<0.05; $^{**}p$<0.01; $^{***}p$<0.001

[a] This table shows the regression results for uncertainty and misallocation under divisible money, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates and percentage frame, as well as uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

Experiment #2

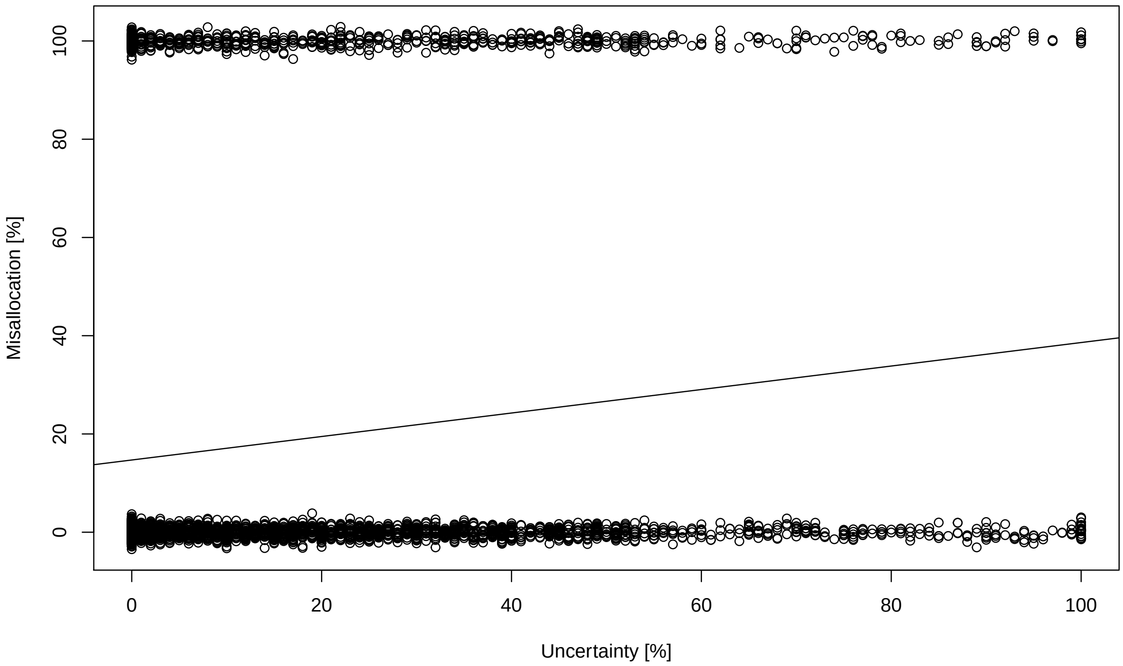


Figure Appendix IV.11: This scatter plot shows the percentage points of uncertainty (x-axis) in relation to misallocation (y-axis). For a better overview the points are jittered randomly in y-direction (although they can only take on the value 0 or 100). The additional OLS-regression line shows that uncertainty and misallocation correlate positively.

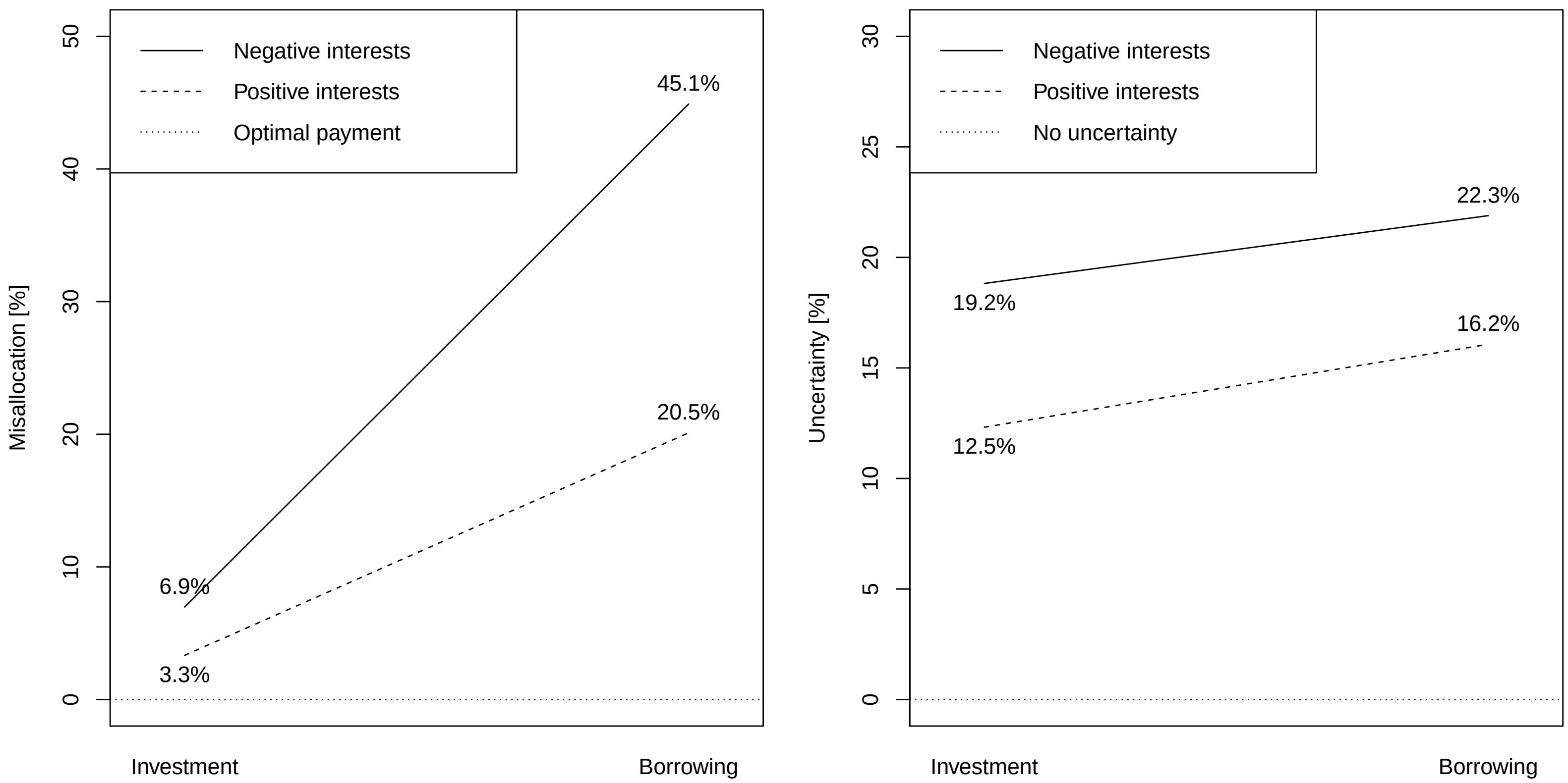


Figure Appendix IV.12: These interaction plots show the average values for misallocation (left graphic) and uncertainty (right graphic) in experiment #2 split by treatment. We differentiate by the borrowing variable (x-axis in both graphics) and the negative int. rates variable (line types in both graphics). For a better overview we do not include the percentage frame dummy and instead average out its effects, as they are not significant in the regression models.

Table Appendix IV.40: Random effects regression showing all used variables[a]

	NOT DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	4.352***	19.519***	4.372***	19.506***	4.703***	16.805***
	(0.737)	(2.440)	(0.738)	(2.448)	(0.852)	(2.726)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	6.306***	2.407	6.307***	2.460	5.344***	3.020
	(0.886)	(1.489)	(0.892)	(1.484)	(0.901)	(1.488)
	[0.000]	[0.106]	[0.000]	[0.097]	[0.000]	[0.042]
	[0.000]	[0.318]	[0.000]	[0.292]	[0.000]	[0.127]
Percentage frame	0.179	-0.003	0.130	1.426	0.837	1.791
	(0.410)	(0.935)	(0.738)	(2.094)	(0.828)	(2.150)
	[0.662]	[0.998]	[0.860]	[0.496]	[0.312]	[0.405]
	[0.662]	[1.000]	[0.860]	[0.992]	[0.312]	[0.810]
Uncertainty		0.015		0.015		0.015
		(0.045)		(0.046)		(0.062)
		[0.747]		[0.748]		[0.811]
		[1.000]		[0.992]		[0.811]
Borrowing × Negative int. rates	-1.431	23.146***	-1.476	23.186***	-1.992	22.319***
	(1.080)	(3.443)	(1.087)	(3.442)	(1.327)	(3.988)
Borrowing × Percentage frame	-1.202	-5.607**	-1.227	-5.673**	-2.174**	-5.678*
	(0.755)	(2.028)	(0.756)	(2.028)	(0.837)	(2.527)
Negative int. rates × Percentage frame	0.394	2.286	0.380	2.138	0.428	0.799
	(0.692)	(1.778)	(0.692)	(1.748)	(0.780)	(1.693)
Triple interaction	1.494	-3.980	1.543	-3.777	3.185*	-5.066
	(1.196)	(3.465)	(1.200)	(3.455)	(1.241)	(4.125)
Round			0.051	-0.114	0.035	-0.107
			(0.048)	(0.124)	(0.057)	(0.148)
Right 2nd			0.047	1.219	0.067	0.756
			(0.392)	(1.127)	(0.460)	(1.355)
Starkness			-0.005	-0.005	0.008	0.001
			(0.010)	(0.030)	(0.012)	(0.035)
Starkness × Percentage frame			0.001	-0.024	-0.016	-0.026
			(0.012)	(0.035)	(0.014)	(0.039)
Age					-0.076	0.095
					(0.137)	(0.090)
Female					1.087	-5.280*
					(3.533)	(2.491)
Third gender					4.615	-3.606
					(13.137)	(6.617)
Has credit card debts					-0.330	-0.119
					(2.668)	(2.209)
# of yearly credit transactions					0.033***	0.010
					(0.008)	(0.006)
# of yearly investment transactions					-0.002	0.015
					(0.023)	(0.021)
Risk seek					0.312	-0.387*
					(0.263)	(0.194)
Years of education					0.556	0.357
					(0.693)	(0.531)
Financial Literacy					-2.167	-4.859***
					(1.762)	(1.190)
Numeracy					-1.295	-3.597**
					(1.403)	(1.376)
Cons. Confidence					-3.532**	1.354
					(1.329)	(0.933)
Pref. num. info.					-3.086	-2.706*
					(1.788)	(1.250)
Constant	12.229***	3.154***	12.021***	3.706	48.677**	60.187***
	(1.127)	(0.952)	(1.314)	(2.147)	(16.466)	(17.789)
Observations	3840	3840	3840	3840	2656	2656
# participants	240	240	240	240	166	166
Mediation analysis of uncertainty mediating misallocation - Sobel test						
Borrowing	0.064		0.064		0.070	
	(0.201)		(0.202)		(0.296)	
	[0.750]		[0.752]		[0.815]	
	[1.000]		[1.000]		[1.000]	
Negative int. rates	0.093		0.092		0.079	
	(0.290)		(0.290)		(0.336)	
	[0.749]		[0.751]		[0.814]	
	[1.000]		[1.000]		[1.000]	
Percentage frame	0.003		0.002		0.012	
	(0.021)		(0.036)		(0.074)	
	[0.901]		[0.958]		[0.867]	
	[1.000]		[1.000]		[1.000]	

Note: $^{*}p$<0.05; $^{**}p$<0.01; $^{***}p$<0.001

[a] This table shows the regression results for uncertainty and misallocation under indivisible money, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. The reference group for gender is male. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates and percentage frame, as well as uncertainty, if applicable. For a definition of the variables, see the glossary in Appendix IV.

Table Appendix IV.41: Random effects regression including screened out subjects[a]

	NOT DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	4.287***	18.935***	4.303***	18.898***	4.453***	16.944***
	(0.687)	(2.315)	(0.687)	(2.320)	(0.773)	(2.605)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	6.567***	2.564	6.558***	2.621	5.799***	3.833*
	(0.856)	(1.466)	(0.858)	(1.460)	(0.881)	(1.536)
	[0.000]	[0.080]	[0.000]	[0.073]	[0.000]	[0.013]
	[0.000]	[0.241]	[0.000]	[0.218]	[0.000]	[0.038]
Percentage frame	0.124	-0.560	0.241	0.816	1.034	0.961
	(0.386)	(0.928)	(0.684)	(2.019)	(0.758)	(2.095)
	[0.748]	[0.546]	[0.724]	[0.686]	[0.172]	[0.646]
	[0.748]	[0.886]	[0.724]	[0.900]	[0.172]	[1.000]
Uncertainty		0.032		0.032		0.034
		(0.042)		(0.042)		(0.060)
		[0.443]		[0.448]		[0.573]
		[0.886]		[0.900]		[01.000]
Borrowing × Negative int. rates	-1.281	22.819***	-1.313	22.881***	-1.697	20.588***
	(0.996)	(3.495)	(0.999)	(3.488)	(1.187)	(4.121)
Borrowing × Percentage frame	-1.276	-4.588*	-1.295	-4.608*	-2.131**	-5.144*
	(0.724)	(2.009)	(0.724)	(2.016)	(0.760)	(2.417)
Negative int. rates × Percentage frame	0.502	3.132	0.494	3.011	0.238	1.499
	(0.658)	(1.759)	(0.656)	(1.738)	(0.720)	(1.691)
Triple interaction	0.881	-3.918	0.928	-3.799	2.182	-4.777
	(1.135)	(3.441)	(1.139)	(3.435)	(1.158)	(4.081)
Round			0.045	-0.146	0.039	-0.131
			(0.046)	(0.117)	(0.053)	(0.139)
Right 2nd			-0.090	1.606	-0.058	1.149
			(0.372)	(1.184)	(0.427)	(1.429)
Starkness			-0.006	0.001	0.009	-0.001
			(0.010)	(0.028)	(0.011)	(0.033)
Starkness × Percentage frame			-0.002	-0.023	-0.017	-0.017
			(0.012)	(0.034)	(0.013)	(0.038)
Constant	12.920***	3.839***	12.901***	4.153*	61.473***	62.735***
	(1.080)	(1.007)	(1.300)	(2.029)	(12.545)	(13.590)
Observations	4320	4320	4320	4320	3040	3040
# participants	270	270	270	270	190	190
Individual control variables	No	No	No	No	Yes	Yes

Mediation analysis of uncertainty mediating misallocation - Sobel test	(1)–(2)	(3)–(4)	(5)–(6)
Borrowing	0.139	0.139	0.151
	(0.185)	(0.186)	(0.273)
	[0.452]	[0.456]	[0.580]
	[1.000]	[1.000]	[1.000]
Negative int. rates	0.213	0.212	0.197
	(0.282)	(0.282)	(0.354)
	[0.449]	[0.454]	[0.578]
	[1.000]	[1.000]	[1.000]
Percentage frame	0.004	0.008	0.035
	(0.021)	(0.038)	(0.081)
	[0.850]	[0.837]	[0.666]
	[1.000]	[1.000]	1.000]

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the regression results for uncertainty and misallocation under indivisible money, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. The reference group for gender is male. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates and percentage frame, as well as uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

Table Appendix IV.42: Random effects regression excluding subjects in the lower and upper 2.5% quantile of experiment duration[a]

	NOT DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing	4.326***	19.516***	4.355***	19.450***	4.887***	16.606***
	(0.761)	(2.514)	(0.762)	(2.524)	(0.894)	(2.808)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	6.485***	2.512	6.496***	2.561	5.549***	2.775
	(0.930)	(1.537)	(0.936)	(1.529)	(0.942)	(1.537)
	[0.000]	[0.102]	[0.000]	[0.094]	[0.000]	[0.071]
	[0.000]	[0.307]	[0.000]	[0.282]	[0.000]	[0.213]
Percentage frame	0.053	-0.221	-0.028	1.853	0.942	1.893
	(0.408)	(0.964)	(0.758)	(2.082)	(0.862)	(2.122)
	[0.897]	[0.818]	[0.971]	[0.374]	[0.274]	[0.372]
	[0.897]	[1.000]	[0.971]	[0.747]	[0.274]	[0.744]
Uncertainty		0.020		0.020		0.029
		(0.047)		(0.047)		(0.066)
		[0.668]		[0.677]		[0.658]
		[1.000]		[0.747]		[0.747]
Borrowing × Negative int. rates	-1.344	22.935***	-1.397	23.046***	-2.035	22.252***
	(1.112)	(3.539)	(1.119)	(3.535)	(1.364)	(4.146)
Borrowing × Percentage frame	-1.167	-5.263*	-1.214	-5.242*	-2.346**	-5.879*
	(0.764)	(2.072)	(0.766)	(2.070)	(0.869)	(2.624)
Negative int. rates × Percentage frame	0.430	2.194	0.399	2.092	0.313	0.542
	(0.702)	(1.842)	(0.703)	(1.813)	(0.786)	(1.760)
Triple interaction	1.233	-3.769	1.315	-3.666	3.149*	-3.847
	(1.189)	(3.561)	(1.196)	(3.555)	(1.259)	(4.280)
Round			0.037	-0.153	0.021	-0.109
			(0.048)	(0.128)	(0.057)	(0.155)
Right 2nd			0.177	1.158	0.170	0.855
			(0.400)	(1.165)	(0.468)	(1.397)
Starkness			-0.007	0.007	0.003	0.007
			(0.010)	(0.031)	(0.012)	(0.036)
Starkness × Percentage frame			0.002	-0.037	-0.016	-0.028
			(0.013)	(0.035)	(0.015)	(0.039)
Constant	11.982***	3.062**	11.936***	3.399	47.783**	66.062***
	(1.110)	(0.981)	(1.319)	(2.219)	(15.873)	(17.660)
Observations	3632	3632	3632	3632	2528	2528
# participants	227	227	227	227	158	158
Individual control variables	No	No	No	No	Yes	Yes
Mediation analysis of uncertainty mediating misallocation - Sobel test						
Borrowing	0.087		0.086		0.142	
	(0.207)		(0.209)		(0.326)	
	[0.674]		[0.682]		[0.664]	
	[1.000]		[1.000]		[1.000]	
Negative int. rates	0.131		0.128		0.161	
	(0.309)		(0.310)		(0.370)	
	[0.672]		[0.680]		[0.663]	
	[1.000]		[1.000]		[1.000]	
Percentage frame	0.001		-0.001		0.027	
	(0.021)		(0.039)		(0.087)	
	[0.960]		[0.989]		[0.754]	
	[1.000]		[1.000]		[1.000]	

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the regression results for uncertainty and misallocation under indivisible money, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates and percentage frame, as well as uncertainty, if applicable. For a definition of the variables, see the glossary in Appendix IV.

Overall analysis

Table Appendix IV.43: Random effects regressions showing all used variables[a]

	COMPARISON NOT DIVISIBLE - DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
NotDivisible	-4.629*	-1.795	-4.619*	-1.877	-3.396	-2.589
	(1.814)	(1.441)	(1.814)	(1.432)	(1.858)	(1.473)
	[0.011]	[0.213]	[0.011]	[0.190]	[0.068]	[0.079]
	[0.021]	[0.213]	[0.022]	[0.380]	[0.135]	[0.284]
Borrowing	5.669***	17.304***	5.676***	17.253***	6.909***	13.623***
	(0.938)	(2.223)	(0.938)	(2.216)	(1.146)	(2.505)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.735***	4.627**	5.751***	4.589**	6.066***	3.336
	(0.876)	(1.382)	(0.877)	(1.385)	(0.930)	(1.556)
	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.032]
	[0.000]	[0.004]	[0.000]	[0.004]	[0.000]	[0.160]
Percentage frame	-0.885	1.634	-0.702	1.132	0.808	-0.667
	(0.638)	(0.780)	(0.795)	(1.415)	(0.769)	(1.600)
	[0.165]	[0.036]	[0.377]	[0.424]	[0.294]	[0.677]
	[0.165]	[0.109]	[0.377]	[0.424]	[0.294]	[0.733]
Uncertainty		0.115*		0.116*		0.080
		(0.040)		(0.040)		(0.044)
		[0.004]		[0.004]		[0.071]
		[0.016]		[0.015]		[0.284]
NotDivisible × Uncertainty		-0.107		-0.107		-0.062
		(0.061)		(0.061)		(0.068)
		[0.079]		[0.077]		[0.367]
		[0.160]		[0.231]		[0.733]
NotDivisible × Borrowing	-1.317	2.242	-1.307	2.327	-2.206	3.211
	(1.192)	(3.300)	(1.196)	(3.297)	(1.430)	(3.662)
NotDivisible × Negative int. rates	0.571	-2.180	0.554	-2.105	-0.740	-0.419
	(1.246)	(2.030)	(1.251)	(2.027)	(1.295)	(2.150)
Borrowing × Negative int. rates	-2.542*	20.045***	-2.574*	20.090***	-2.526	21.323***
	(1.262)	(3.447)	(1.265)	(3.438)	(1.539)	(4.113)
NotDivisible × Borrowing × Ne. int. rates	1.110	3.092	1.090	3.014	0.510	0.971
	(1.660)	(4.870)	(1.662)	(4.863)	(2.028)	(5.722)
NotDivisible × Percentage frame	1.065	-1.635	1.076	-1.513	0.194	0.102
	(0.758)	(1.217)	(0.757)	(1.205)	(0.759)	(1.183)
Borrowing × Percentage frame	-1.952	-6.768***	-1.944	-6.756***	-2.520*	-6.768**
	(1.050)	(1.705)	(1.048)	(1.701)	(1.244)	(2.089)
NotDivisible × Borrowing × Perc. frame	0.750	1.153	0.721	1.068	0.325	1.047
	(1.293)	(2.648)	(1.293)	(2.643)	(1.499)	(3.260)
Negative int. rates × Percentage frame	1.121	-0.276	1.126	-0.257	0.451	0.325
	(0.966)	(1.582)	(0.965)	(1.581)	(0.944)	(1.911)
NotDivisible × Neg. int. rates × Perc. frame	-0.727	2.565	-0.748	2.458	-0.026	0.576
	(1.188)	(2.379)	(1.189)	(2.363)	(1.225)	(2.552)
Triple interaction	1.173	0.459	1.193	0.476	1.283	-0.503
	(1.619)	(3.154)	(1.619)	(3.143)	(2.040)	(3.833)
NotDivisible × Triple interaction	0.321	-4.430	0.356	-4.302	1.957	-4.626
	(2.012)	(4.683)	(2.016)	(4.661)	(2.388)	(5.613)

continued on next page ...

... continued from previous page

	COMPARISON NOT DIVISIBLE - DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
Round			0.082*	-0.044	0.102*	-0.054
			(0.037)	(0.082)	(0.044)	(0.097)
Right 2nd			0.049	0.832	0.075	0.143
			(0.285)	(0.752)	(0.331)	(0.882)
Starkness			-0.001	-0.018	0.006	-0.023
			(0.007)	(0.019)	(0.008)	(0.022)
Starkness × Percentage frame			-0.004	0.009	-0.019	0.017
			(0.010)	(0.024)	(0.010)	(0.028)
Age					-0.025	0.060
					(0.085)	(0.063)
Female					-0.651	-5.003**
					(2.495)	(1.705)
Third gender					6.090	2.718
					(9.267)	(5.470)
Has credit card debts					-0.971	-1.255
					(2.203)	(1.550)
# of yearly credit transactions					-0.000***	0.000***
					(0.000)	(0.000)
# of yearly investment transactions					-0.012	0.018
					(0.020)	(0.020)
Risk seek					0.058	-0.320
					(0.220)	(0.170)
Years of education					0.714	-0.150
					(0.468)	(0.370)
Financial Literacy					-3.545**	-4.120***
					(1.227)	(0.790)
Numeracy					-1.969	-2.799***
					(1.079)	(0.776)
Cons. Confidence					-4.040***	0.570
					(0.990)	(0.725)
Pref. num. info.					-3.268*	-2.258*
					(1.515)	(0.894)
Constant	16.858***	5.025***	16.207***	5.968***	65.702***	63.204***
	(1.422)	(1.075)	(1.509)	(1.629)	(12.531)	(10.088)
Observations	7680	7680	7680	7680	5280	5280
# participants	480	480	480	480	330	333

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the regression results for uncertainty and misallocation where we compare divisibility with non-divisibility, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. The reference group for gender is male. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates, percentage frame and NotDivisible, as well as uncertainty and NotDivisible × Uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

Table Appendix IV.44: Random effects regressions including screened out subjects[a]

	COMPARISON NOT DIVISIBLE - DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
NotDivisible	-4.042*	-1.546	-4.027*	-1.616	-3.197	-2.604
	(1.724)	(1.450)	(1.725)	(1.441)	(1.736)	(1.449)
	[0.019]	[0.286]	[0.020]	[0.262]	[0.066]	[0.072]
	[0.038]	[0.286]	[0.039]	[0.419]	[0.131]	[0.217]
Borrowing	5.491***	17.919***	5.495***	17.860***	6.593***	14.278***
	(0.869)	(2.127)	(0.867)	(2.118)	(1.054)	(2.379)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.519**	4.938***	5.533**	4.875***	5.846***	3.762
	(0.836)	(1.342)	(0.837)	(1.342)	(0.851)	(1.547)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.015]
	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.075]
Percentage frame	-0.915	2.157*	-0.675	2.141	0.761	0.536
	(0.597)	(0.834)	(0.767)	(1.450)	(0.775)	(1.687)
	[0.125]	[0.010]	[0.378]	[0.140]	[0.326]	[0.751]
	[0.125]	[0.029]	[0.378]	[0.419]	[0.326]	[0.800]
Uncertainty		0.110*		0.111*		0.091
		(0.037)		(0.037)		(0.042)
		[0.003]		[0.003]		[0.032]
		[0.013]		[0.013]		[0.128]
NotDivisible × Uncertainty		-0.083		-0.084		-0.055
		(0.057)		(0.057)		(0.065)
		[0.143]		[0.140]		[0.399]
		[0.290]		[0.419]		[0.800]
NotDivisible × Borrowing	-1.204	1.041	-1.200	1.100	-2.146	2.689
	(1.107)	(3.142)	(1.108)	(3.138)	(1.311)	(3.497)
NotDivisible × Negative int. rates	1.048	-2.336	1.026	-2.246	-0.058	-0.020
	(1.196)	(1.987)	(1.200)	(1.979)	(1.228)	(2.173)
Borrowing × Negative int. rates	-2.394*	19.916***	-2.411*	19.964***	-2.304	21.094***
	(1.191)	(3.344)	(1.192)	(3.333)	(1.425)	(3.981)
NotDivisible × Borrowing × Negative int. rates	1.113	2.896	1.090	2.852	0.590	-0.515
	(1.552)	(4.835)	(1.552)	(4.823)	(1.853)	(5.726)
NotDivisible × Percentage frame	1.039	-2.716*	1.053	-2.601*	0.408	-1.022
	(0.710)	(1.247)	(0.709)	(1.235)	(0.712)	(1.248)
Borrowing × Percentage frame	-1.600	-7.332***	-1.579	-7.313***	-2.036	-7.033***
	(1.002)	(1.713)	(0.999)	(1.703)	(1.195)	(2.089)
NotDivisible × Borrowing × Percentage frame	0.324	2.736	0.289	2.682	-0.113	1.845
	(1.235)	(2.639)	(1.233)	(2.632)	(1.416)	(3.177)
Negative int. rates × Percentage frame	1.130	0.979	1.143	1.018	0.637	1.343
	(0.901)	(1.550)	(0.901)	(1.548)	(0.863)	(1.871)
NotDivisible × Negative int. rates × Percentage frame	-0.628	2.156	-0.650	2.052	-0.407	0.236
	(1.116)	(2.343)	(1.116)	(2.330)	(1.124)	(2.514)
Triple interaction	0.945	-0.802	0.939	-0.788	1.071	-2.254
	(1.522)	(3.046)	(1.521)	(3.031)	(1.881)	(3.704)
NotDivisible × Triple interaction	-0.064	-3.111	-0.011	-3.029	1.153	> -2.550
	(1.899)	(4.593)	(1.898)	(4.572)	(2.207)	(5.494)
Constant	16.962***	5.460***	16.305***	6.103***	68.808***	65.573***
	(1.344)	(1.037)	(1.438)	(1.600)	(9.784)	(8.873)
Observations	8560	8560	8560	8560	5968	5968
# participants	535	535	535	535	373	373
Further experimental control variables	No	No	Yes	Yes	Yes	Yes
Further individual control variables	No	No	No	No	Yes	Yes

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the regression results for uncertainty and misallocation where we compare divisibility with non-divisibility, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates, percentage frame and NotDivisible, as well as uncertainty and NotDivisible × Uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

Table Appendix IV.45: Random effects regressions excluding subjects in the lower and upper 2.5% quantile of experiment duration[a]

	COMPARISON NOT DIVISIBLE - DIVISIBLE					
Dependent variable	Uncertainty	Misallocation	Uncertainty	Misallocation	Uncertainty	Misallocation
	(1)	(2)	(3)	(4)	(5)	(6)
NotDivisible	-5.057*	-1.824	-5.062*	-1.891	-4.432*	-2.898
	(1.846)	(1.497)	(1.848)	(1.487)	(1.880)	(1.527)
	[0.006]	[0.223]	[0.006]	[0.203]	[0.018]	[0.058]
	[0.012]	[0.223]	[0.012]	[0.407]	[0.037]	[0.216]
Borrowing	5.747***	17.399***	5.747***	17.352***	6.978***	14.167***
	(0.971)	(2.272)	(0.971)	(2.266)	(1.194)	(2.610)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Negative int. rates	5.929***	4.327*	5.937***	4.292*	6.249***	3.101
	(0.901)	(1.406)	(0.902)	(1.409)	(0.958)	(1.595)
	[0.000]	[0.002]	[0.000]	[0.002]	[0.000]	[0.052]
	[0.000]	[0.010]	[0.000]	[0.010]	[0.000]	[0.212]
Percentage frame	-0.794	1.652	-0.691	1.582	0.891	-0.367
	(0.658)	(0.809)	(0.829)	(1.446)	(0.805)	(1.644)
	[0.227]	[0.041]	[0.405]	[0.274]	[0.268]	[0.823]
	[0.227]	[0.136]	[0.405]	[0.407]	[0.268]	[0.846]
Uncertainty		0.123*		0.124*		0.089
		(0.040)		(0.040)		(0.044)
		[0.002]		[0.002]		[0.043]
		[0.010]		[0.010]		[0.216]
NotDivisible × Uncertainty		-0.114		-0.114		-0.057
		(0.062)		(0.062)		(0.071)
		[0.068]		[0.066]		[0.423]
		[0.136]		[0.197]		[0.846]
NotDivisible × Borrowing	-1.376	2.425	-1.350	2.490	-2.045	2.716
	(1.239)	(3.408)	(1.244)	(3.408)	(1.499)	(3.813)
NotDivisible × Negative int. rates	0.641	-1.711	0.644	-1.638	-0.638	-0.384
	(1.302)	(2.098)	(1.308)	(2.092)	(1.351)	(2.224)
Borrowing × Negative int. rates	-2.615*	19.839***	-2.631*	19.878***	-2.534	20.711***
	(1.304)	(3.510)	(1.307)	(3.503)	(1.597)	(4.180)
NotDivisible × Borrowing × Negative int. rates	1.233	3.388	1.181	3.336	0.401	1.766
	(1.723)	(5.012)	(1.725)	(5.005)	(2.107)	(5.913)
NotDivisible × Percentage frame	0.850	-1.876	0.877	-1.774	0.184	0.135
	(0.777)	(1.267)	(0.775)	(1.253)	(0.785)	(1.243)
Borrowing × Percentage frame	-2.054	-7.081***	-2.053	-7.066***	-2.617*	-7.340***
	(1.088)	(1.749)	(1.086)	(1.745)	(1.298)	(2.155)
NotDivisible × Borrowing × Percentage frame	0.878	1.735	0.833	1.675	0.217	1.298
	(1.335)	(2.731)	(1.334)	(2.726)	(1.566)	(3.399)
Negative int. rates × Percentage frame	1.115	-0.065	1.113	-0.038	0.429	0.686
	(0.999)	(1.630)	(0.999)	(1.628)	(0.978)	(1.969)
NotDivisible × Negative int. rates × Percentage frame	-0.686	2.293	-0.717	2.199	-0.129	-0.050
	(1.226)	(2.477)	(1.228)	(2.459)	(1.261)	(2.651)
Triple interaction	1.152	-0.085	1.171	-0.072	1.221	-1.185
	(1.680)	(3.204)	(1.679)	(3.192)	(2.130)	(3.875)
NotDivisible × Triple interaction	0.105	-4.168	0.171	-4.071	2.038	-3.358
	(2.066)	(4.801)	(2.070)	(4.781)	(2.480)	(5.755)
Constant	17.188***	5.057***	16.637***	5.920***	64.754***	65.907***
	(1.468)	(1.111)	(1.557)	(1.688)	(12.823)	(10.088)
Observations	7280	7280	7280	7280	5008	5008
# participants	455	455	455	455	313	313
Further experimental control variables	No	No	Yes	Yes	Yes	Yes
Further individual control variables	No	No	No	No	Yes	Yes

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a] This table shows the regression results for uncertainty and misallocation where we compare divisibility with non-divisibility, each with three different models: The simple models (1) and (2) which include only the treatment variables as dummies, as well as their interactions and uncertainty for the misallocation model; the models (3) and (4) which include some technical aspects of the experiment; and the complete models (5) and (6) with all control variables. Robust standard errors in parentheses, unadjusted p-values and Bonferroni-Holm adjusted p-values in brackets. The p-values are adjusted for borrowing, negative interest rates, percentage frame and NotDivisible, as well as uncertainty and NotDivisible × Uncertainty, if applicable. Asterisks indicate significance after adjustment. For a definition of the variables, see the glossary in Appendix IV.

Glossary

Variable	Description
# of yearly credit transactions	Gives the number of borrowing transactions a participant reported to execute typically per year (individual control variable).
# of yearly investment transactions	Gives the number of investment transactions a participant reported to execute typically per year (individual control variable).
Age	Measures the age of a participant in years (individual control variable).
Borrowing	Dummy variable equal to one in decisions where participants have to take debts from two credits (within-subject varying)
Cons. Confidence	A participant's consumer confidence measured by five questions on a Likert scale from one to six (individual control variable).
Duration exp	Duration of the 19 experiment rounds (also including the three trial rounds).
Duration PEQ	Duration of the post experimental questionnaire after the experiment rounds, including the measuring of all individual control variables.
Duration pre exp	Duration of all proceedings before the experiment rounds, including reading the instructions and completing the comprehension tasks.

Duration total	Total duration of the experiment.
Female	Dummy variable that equals one if a participant is female. Gives the differences to the reference level "male" (individual control variable).
Financial literacy	Measure for a participant's financial literacy as number of correctly answered questions out of six questions (individual control variable).
Has credit card debts	Dummy variable that equals one if a participant reported to have credit card debts (individual control variable).
Negative int. rates	Dummy variable equal to one in decisions where interest rates (or absolute interests) are negative (within-subject varying).
NotDivisible	Dummy variable equal to one in decisions from experiment #2 and zero in decisions from experiment #1. Used in comparison of both experiments (between-subject varying).
Numeracy	Measure of a participant's numeracy as number of correctly answered questions out of 11 questions (individual control variable).

Misallocation	Measure of the percentage of money that a participant does not allocate to the financially optimal asset or credit. We multiplied the values with 100 to keep the measure on the same scale as uncertainty (between 0 and 100).
Percentage frame	Dummy variable equal to one in decisions where interests are displayed as percentages instead of absolute values (within-subject varying).
Pref. num. info	A participants preference for numerical information measured by eight questions on a Likert scale from one to six (individual control variable).
Right 2nd	Dummy variable that is equal to one when the asset or credit presented secondly (that is under the first asset/credit) is the asset/credit a participant has to transfer money to to avoid misallocation. This variable captures potential order effects in the presentation of the assets/credits (experimental control variable).
Risk seek	Measure of risk affinity of a participant on a scale between 0 (risk averse) to 31 (risk affine) (individual control variable).
Round	Number of the decision round for one participant. Captures potential learning effects during the experiment (experimental control variable).

Starkness	Measures the spread of the values between the two assets or credits presented in one experimental round. In case of "Percentage frame = 1" we multiply this value with 10 to keep it approximately on the same scale as when we present absolute values instead (experimental control variable).
Third gender	Dummy variable that equals one if a participant feels affiliated to a third gender. Gives the differences to the reference level "male" (individual control variable).
Triple interaction	Short term for the triple interaction term between the within-subject varying variables borrowing, negative int. rates and percentage frame.
Uncertainty	Measures cognitive uncertainty on a scale between 0 (completely certain) and 100 (completely uncertain) in each decision. The participants self-report this value when we ask them how certain they are about their decision.
Years of education	Number of years of education of a participant, reported in full-time equivalents and includes compulsory years of schooling (individual control variable).

Appendix V (to Chapter V)

Description of variables & survey questions

(This page intentionally left blank for formatting purposes)

Table Appendix V.47: Overview of survey questions. This table contains a brief overview of all key questions grouped into thematic modules. The letters and numbers in parentheses correspond to their respective survey IDs. Sources are provided for relevant variables.

Variable	Survey ID	Source
Module: Demographics		
Age	(N1)	
Gender	(N2)	
Party	(N3)	Anderson and Robinson (2021), Bauer et al. (2021)
Years of education	(O1)	
Income	(O2)	
Module: Sustainability preferences		
Wahl-o-Mat questions	(Q1)	Wahl-o-Mat
Module: Environmental Literacy		
Question on Sustainable Development	(R1)	Filippini et al. (2023)
Question on heating	(R2)	Anderson and Robinson (2021)
Question on CO2 Footprint	(R3)	Geiger and Holzhauer (2020)
Question on temperature	(R4)	Own question
Question on forestry	(R5)	Zwickle and Jones (2018)
Module: Risk, time, social preferences and trust		
Risk preferences	(P1)	Falk et al. (2023)
Time preferences	(P2)	Falk et al. (2023)
Trust	(P3)	Falk et al. (2023)
Social preferences	(P4)	Falk et al. (2023)
Social preferences, costly	(P5)	Falk et al. (2023)
UG Mininimal Demand	(P6)	Falk et al. (2023)
Module: Financial experience		
Financial Decision Maker	(S1)	Gutsche and Zwergel (2020)
Checks Portfolio	(S2)	modified version from Anderson and Robinson (2021)
Talks often about investments	(S3)	Riedl and Smeets (2017)
Current investments	(T1)	Gutsche and Zwergel (2020)
Past investments	(T2)	Gutsche and Zwergel (2020)
Module: Financial expectations of ESG products		
Return expectations	(U1)	Riedl and Smeets (2017), Bauer et al. (2021)
Risk perception	(U2)	Riedl and Smeets (2017), Bauer et al. (2021)
Dimension importance	(U4)	Own question
Financial scepticism	(U5)	Own question
Module: Financial literacy		
Interest rates	(V1)	Lusardi and Mitchell (2014)
Inflation	(V2)	Lusardi and Mitchell (2014)
Diversification	(V3)	Lusardi and Mitchell (2014)
Module: Sustainable finance literacy - Global		
Question ESG	(W1)	Filippini et al. (2023)
Question Sustainable Financial Product	(W2)	Filippini et al. (2023)
Question ESG Components	(W3)	Filippini et al. (2023)
Question Exclusion-based investing	(W4)	Own question
Question Best-in-class approach	(W5)	Own question
Module: Sustainable finance literacy - Local		
Question SFDR Article 6	(X1)	Own question
Question SFDR Article 8	(X2)	Own question
Question SFDR Article 8 social	(X3)	Own question
Question Question Article 9	(X4)	Own question
Module: Perceived impact		
Positive contribution	(F2), (H2), (J2), (L2)	Heeb et al. (2022)

Description of variables In the following, we provide a description of all key variables used in our analyses, as well as the original questionnaire format. Note that the questionnaire was originally conducted in German.

Main Independent Variables

- *Brochure Treatment*. This dummy variable equals 1 if the participant was given the brochure to read and 0 otherwise.

- *ESG Pref Score*. The average over the seven Wahl-o-Mat questions (Q1). Item 3 is reversed. The questions do not have any numbers associated with the answers. For a convenient interpretation of the lowest category, we use a scale from 0 ("strongly disagree") to 4 ("strongly agree").

Main Dependent Variables - Decision Level

- *Chose ESG*. Dummy variable based on the fund decisions of questions (E2), (G2), (I2), (K2). Equals 1 if the participant chooses an article 8 or 9 fund.

- *Used Criterion*. Dummy variable based on the investment criteria of questions (F1), (H1), (J1), (L1). Equals 1 if the participant indicates that sustainability information played a role in the investment decision.

Demographics

- *Age*. The natural log of the participants' self-stated age (N1).

- *Gender.* Answers to the gender question (N2) are split into dummy variables, with "male" as the reference category.

- *Gender Female*. This dummy variable equals 1 if the participant chooses Female from among the options Female, Male, Non-Binary and Prefer not to say, and 0 if not.
- *Gender Non Binary*. This dummy variable equals 1 if the participant chooses Non Binary from among the options Female, Male, Non-Binary and Prefer not to say, and 0 if not.

• *Party Preference.* Answers to the party preference question (N3) are split into dummy variables, with CDU/CSU as the reference category.

- *Party SPD*. This dummy variable equals 1 if the participant chooses SPD.
- *Party Greens*. This dummy variable equals 1 if the participant chooses Bündnis 90 / Die Grünen
- *Party FDP*. This dummy variable equals 1 if the participant chooses FDP
- *Party The Left*. This dummy variable equals 1 if the participant chooses Die Linke
- *Party AfD*. This dummy variable equals 1 if the participant chooses AfD
- *Party Other*. This dummy variable equals 1 if the participant chooses "Other party"
- *Party None*. This dummy variable equals 1 if the participant chooses "Would not vote"
- *Party not eligible*. This dummy variable equals 1 if the participant chooses "I am not eligible"

• *Years of education*. Self-reported highest degree (O1), which we translate into the implied years of education the participant has completed: No degree (yet) = 8 years, elementary school = 9 years, secondary school or Realschule or completed apprenticeship = 11 years, advanced technical college entrance qualification = 12 years, Abitur or "erweiterte Oberschule" with completion of 12th grade (university entrance qualification) = 13 years, bachelor's degree = 16 years, master's degree or equivalent = 18 years, doctorate or postdoctoral qualification = 23 years.

- *Income*. Self-reported net monthly household income class (O2). The options start with "less than EUR 500" and "500 to less than EUR 1000", then move in steps of EUR 1000 until EUR 7000, and the final class is "EUR 7000 or more".

Risk Preferences, Time Preferences, Trust and Altruism

- *Risk Preference*. Answer to the question "How do you see yourself: Are you a person who is generally willing to take risks, or do you try to avoid taking risks" (P1), on a 10-point scale (0 = "Completely unwilling to take risks"; 10 = "Very willing to take risks") according to the experimentally validated survey module of Falk et al. (2023).

- *Time Preference*. Answer to the question "In comparison to others, are you a person who is generally willing to give up something today in order to benefit from that in the future or are you not willing to do so?" (P2) on a 10-point scale (0 = "Completely unwilling"; 10 = "Very willing to do so"), following Falk et al. (2023).

- *Trust*. Answer to the question "How well does the following statement describe you as a person? As long as I am not convinced otherwise, I assume that people have only the best intentions." (P3) on a 10-point scale (0 = "Does not describe me at me"; 10 = "Describes me very well"), following Falk et al. (2023).

- *Social Preferences*. Answer to the question “How do you assess your willingness to share with others without expecting anything in return when it comes to charity?” (P4) on a 10-point scale (0 = “Completely unwilling to share”; 10 = “Very willing to share”), following Falk et al. (2023).

- *Social Preferences, costly*. Answer to the question "How do you see yourself: Are you a person who is generally willing to punish unfair behavior even if this is costly?" (P5) on a 10-point scale (0 = "not willing at all"; 10 = "Very willing to punish unfair behavior"), following Falk et al. (2023).

- *UG minimal demand*. The natural log of the minimum demand as player 2 in a hypothetical Ultimatum Game with EUR 100 to share (P6), following Falk et al. (2023).

Financial Experience and Signaling

- *Financial Decision Maker*. Dummy variables based on the answers to the item "Please indicate which of the following statements applies to you personally when it comes to financial matters, e.g. investments." (S1)
  - *Decides finances with partner*. Dummy variable, equals 1 if a participant gave that answer.
  - *Does not decide about own finances*. Dummy variable, equals 1 if a participant gave that answer.

- *Checks Portfolio*. Answers to the "Checks Portfolio" question (S2) are split into dummy variables, with "weekly" as the reference category, following Anderson and Robinson (2021).

  - *Checks Portfolio 12 times/year*. This dummy variable equals 1 if the participant responds "monthly".
  - *Checks Portfolio 2-11 times/year*. This dummy variable equals 1 if the participant responds "several times per year but less frequently than monthly".
  - *Checks Portfolio 1 time/year*. This dummy variable equals 1 if the participant responds "once a year".
  - *Checks Portfolio <1 time/year*. This dummy variable equals 1 if the participant responds "more rarely".
  - *Checks Portfolio never*. This dummy variable equals 1 if the participant responds "never".
  - *Checks Portfolio only when opening/changing*. This dummy variable equals 1 if the participant responds "only when I create a account or change it".
  - *Has no Portfolio*. This dummy variable equals 1 if the participant responds "I don't have an investment account".

  For the correlation analyses, we recode this variable into an ordinal variable, where we join "Has no Portfolio" with "never" and "Checks Portfolio only when opening/changing" with "Checks Portfolio <1 time/year".

- *Talks often about inv.*. Likert scale response to the statement "I often talk about investment with others" (1 fully disagree, . . . , 5 strongly agree) (S3), slightly modifying Riedl and Smeets (2017).

- *Financial experience.*. Counts the numbers of current and past types of investments a participant has or had invested in (T1 and T2), following Gutsche and Zwergel (2020).

Literacy

- *Financial literacy*. The sum of correct answers to the "Big Three" questions (V1-V3), following Lusardi and Mitchell (2014).

- *Environmental literacy*. The sum of correct answers to the five questions from the module "Environmental Literacy" (R1-R5) in Table Appendix V.47. The questions follow Filippini et al. (2023); Anderson and Robinson (2021); Geiger and Holzhauer (2020); Zwickle and Jones (2018).

- *Sustainable finance literacy*. This variable is the sum of correct answers to the five questions from the modules "Sustainable finance literacy - Global" (W1-W5) and "Sustainable finance literacy - Local" (X1-X4) in Table Appendix V.47. Three of these questions follow Filippini et al. (2023).

Financial Expectations and Perceived Importance

- *ESG return expectations / ESG risk expectations*. Answers to the questions about return expectations and risk expectations of ESG funds (U1 and U2) are split into dummy variables with "much lower" as the reference category in both cases, following Riedl and Smeets (2017)

  - *ESG return expectations / ESG risk expectations somewhat lower*. This dummy variable equals 1 if the participant answers "somewhat lower".
  - *ESG return expectations / ESG risk expectations similar*. This dummy variable equals 1 if the participant answers "similar".
  - *ESG return expectations / ESG risk expectations somewhat higher*. This dummy variable equals 1 if the participant answers "somewhat higher".

  - *ESG return expectations / ESG risk expectations much higher*. This dummy variable equals 1 if the participant answers "much higher".
  - *ESG return expectations / ESG risk expectations does not know*. This dummy variable equals 1 if the participant answers "I do not know".

  For the correlation analyses, we recode these variables into ordinal variables, where we code "much lower" as 1, "somewhat lower" as 2, "similar" as 3, "somewhat higher" as 4, and "much higher" as 5.

- *Perceived Importance*. Likert scale response to the statement "How important are the following dimensions to you when investing?" (U4), on a scale from 1 (not important) to 5 (very important).
  - *Importance Returns*.
  - *Importance Risk*.
  - *Importance E*.
  - *Importance S*.
  - *Importance G*.
- *Financial Scepticism*. The average over the 3 questions from module U5, on a scale of 1 to 5, where 1 is "strongly disagree" and 5 is "strongly agree". Item 2 is reversed.
- *Perceived Impact*. The average of the 4 0-5 Likert scale questions about the perceived impact of the respective investment from modules F2, H2, J2 and L2.

Experimenter Demand Effect

- *Low EDE*. Based on statement (D2). A random subsample of the brochure group gets to read the following statement: "We expect that participants in the experiment who read these instructions will be less likely to invest in sustainable funds than they normally would.", but does not see statement (D3).

- *High EDE*. Based on statement (D3). A random subsample of the brochure group gets to read the following statement: "We expect that participants in the experiment who read these instructions will be more likely to invest in sustainable funds than they normally would.", but does not see statement (D2).

Note that for the reference group for this variable, we only show (D1).

Technical

- *Interview time*. The variable represents the natural log of the time spent responding to the survey, excluding the time spent reading the brochure.

- *Use data*. The variable represents the self-stated assessment to the statement "I have given and made my answers and decisions carefully, and to the best of my knowledge, and therefore think that my data should be used for the study" (Y1).

Note that for the two complex linear models in section V.4.2 we had to z-standardize several variables since the downdated VtV matrix of the regression in the first stage of the analysis was not positive definite. This affects the variables Environmental Literacy, Education, Financial Literacy, Perceived impact, Use Data, Risk Preferences, Time Preferences, Social Preferences, Social Preferences, costly, and Financial Experience.

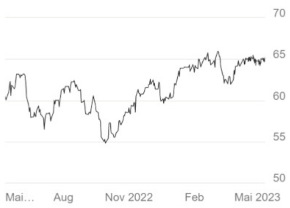
width=!,height=!,pages=-, pagecommand=

Used screenshots of the funds & brochure

UniNachhaltig Aktien Europa A

Fonds | 988475 | LU0090707612

65,50 EUR



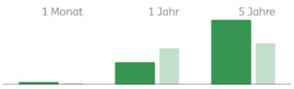
Basisinformationen

Fondskategorie	Aktienfonds
Emittent	Union Investment Luxembourg S.A.
Morningstar-Rating	★★★★☆
Nachhaltigkeitsbezug (SFDR)	Artikel 8
Ausschüttungsart	Ausschüttend
Fondswährung	EUR
Fondsvermögen	1,66 Mrd. EUR (Groß)
Sparplan möglich	Nein

Kosten

Laufende Kosten p.a.	1,55 %
Fondstransaktionskosten p.a.	0,05 %
Performancegebühr	Nein

Performance im Vergleich



	1 Monat	1 Jahr	5 Jahre
UniNachhaltig Aktien Europa A	+1,24 %	+12,05 %	+35,03 %
EURO STOXX 50	+0,72 %	+19,60 %	+22,25 %

Zusammensetzung

Anlageklassen

Aktien	96,8 %
Sonstige	3,2 %

Positionen

Astrazeneca (ADRs)	3,7 %
ASML NY	3,3 %
Nordisk (ADRs)	3,2 %
Siemens	3,1 %

Länder

Frankreich	23,6 %
Deutschland	22,5 %
Großbritannien	14,0 %
Niederlande	6,9 %

Branchen

Finanzwesen	19,2 %
Gesundheitswesen	14,3 %
Industrie	12,3 %
IT	10,2 %

Strategie

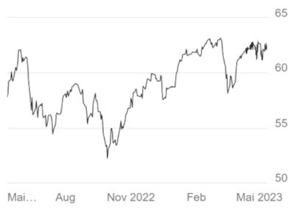
Das Fondsvermögen wird überwiegend in europäische Aktien investiert. Daneben kann das Fondsvermögen in fest und variabel verzinslichen Wertpapieren, in Bankguthaben und/oder Geldmarktinstrumenten sowie bis zu 10 Prozent in Zielfonds angelegt werden. Der Fonds berücksichtigt bei der Auswahl der Emittenten ethische, soziale und ökologische Kriterien. Derivate können zu Investitions- und Absicherungszwecken eingesetzt werden.

Dow Jones Euro Stoxx 50misst die Wertentwicklung der 50 meistgehandelten Aktien der EUR-Zone.

UniValueFonds: Europa A

Fonds | 630948 | LU0126314995

62,639 EUR



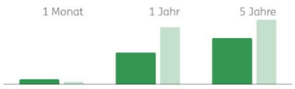
Basisinformationen

Fondskategorie	Aktienfonds
Emittent	Union Investment Luxembourg S.A.
Morningstar-Rating	★★★★☆
Nachhaltigkeitsbezug (SFDR)	Artikel 6
Ausschüttungsart	Ausschüttend
Fondswährung	EUR
Fondsvermögen	158,63 Mio. EUR (Klein)
Sparplan möglich	Nein

Kosten

Laufende Kosten p.a.	1,52 %
Fondstransaktionskosten p.a.	0,05 %
Performancegebühr	Nein

Performance im Vergleich



	1 Monat	1 Jahr	5 Jahre
UniValueFonds: Europa A	+1,71 %	+10,87 %	+15,93 %
EURO STOXX 50	+0,72 %	+19,60 %	+22,25 %

Zusammensetzung

Anlageklassen

Aktien	97,3 %
Sonstige	2,7 %

Positionen

Novartis	4,8 %
Siemens	4,4 %
TotalEnergies	4,4 %
Shell	4,2 %

Länder

Deutschland	23,0 %
Frankreich	18,4 %
Großbritannien	16,2 %
Niederlande	8,3 %

Branchen

Finanzwesen	24,0 %
Energie	10,5 %
Gesundheitswesen	10,4 %
Versorgungsbetriebe	10,0 %

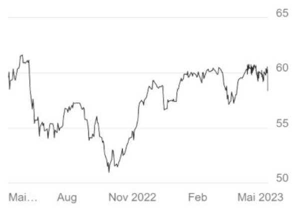
Strategie

Der Fonds investiert in ausgewählte europäische Aktien mit nachhaltiger Wertorientierung (Substanzwerte). Die Unternehmen müssen einem klar definierten Anforderungskatalog entsprechen. Wichtige Kriterien für die gezielte Einzeltitelauswahl (Stock Picking) sind dabei die Unternehmensbewertung, der reale Gegenwert und die tendenziell geringe Schwankungsintensität der Aktien. Der Fonds wird hinsichtlich Länder, Branchen- und Unternehmensauswahl sowie Timing und Investitionsgrad aktiv gemanagt.

UniDividendenAss -net- A

Fonds | A0B821 | LU0186860663

59,109 EUR



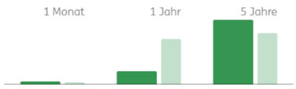
Basisinformationen

Fondskategorie	Aktienfonds
Emittent	Union Investment Luxembourg S.A.
Morningstar-Rating	★★★☆☆
Nachhaltigkeitsbezug (SFDR)	Artikel 6
Ausschüttungsart	Ausschüttend
Fondswährung	EUR
Fondsvermögen	1,18 Mrd. EUR (Groß)
Sparplan möglich	Nein

Kosten

Laufende Kosten p.a.	1,87 %
Fondstransaktionskosten p.a.	0,10 %
Performancegebühr	Ja, siehe Verkaufsunterlagen

Performance im Vergleich



	1 Monat	1 Jahr	5 Jahre
UniDividendenAss -net- A	+1,08 %	+5,51 %	+28,07 %
EURO STOXX 50	+0,72 %	+19,60 %	+22,25 %

Zusammensetzung

Anlageklassen

Aktien	95,2 %
Sonstige	4,8 %

Positionen

Unilever	5,8 %
AXA	5,6 %
Rio Tinto PLC	4,9 %
National Grid PLC	4,1 %

Länder

Großbritannien	27,7 %
Frankreich	14,8 %
Deutschland	14,7 %
Schweiz	11,9 %

Branchen

Roh-, Hilfs- & Betriebsstoffe	17,4 %
Finanzwesen	17,0 %
Basiskonsumgüter	14,2 %
Versorgungsbetriebe	11,9 %

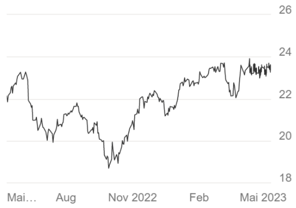
Strategie

Der Fonds investiert in aus gewählten dividendenstarke Unternehmen mit Gewinnstabilität. Der Anlageschwerpunkt liegt dabei auf den europäischen Aktienmärkten. Darüber hinaus ist eine Beimischung von internationalen Dividendentiteln möglich.

Deka EURO iSTOXX ex Fin Dividend+ UCITS ETF

ETF | ETFL48 | DE000ETFL482

23,555 EUR



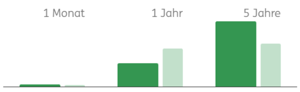
Basisinformationen

Fondskategorie	Aktienfonds
Emittent	Deka Investment GmbH
Morningstar-Rating	★★★★☆
Nachhaltigkeitsbezug (SFDR)	Artikel 6
Ausschüttungsart	Ausschüttend
Nachbildung	Vollständig
Fondswährung	EUR
Fondsvermögen	286,62 Mio. EUR (Mittel)
Sparplan möglich	Nein

Kosten

Laufende Kosten p.a.	0,30 %
Performancegebühr	Nein

Performance im Vergleich



	1 Monat	1 Jahr	5 Jahre
Deka EURO iSTOXX ex Fin Dividend+ UCITS ETF	+1,19 %	+11,99 %	+33,60 %
EURO STOXX 50	+0,72 %	+19,60 %	+22,25 %

Zusammensetzung

Anlageklassen

Aktien	98,4 %
Sonstige	1,6 %

Positionen

OCI	6,1 %
ENGIE	4,0 %
Enel	3,1 %
Telefonica	3,0 %

Länder

Deutschland	21,6 %
Frankreich	20,7 %
Niederlande	16,6 %
Spanien	11,2 %

Branchen

Versorger	23,1 %
Chemie	14,8 %
Autoindustrie	11,9 %

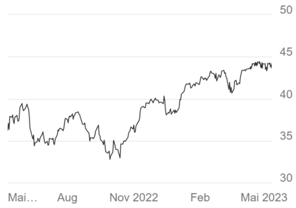
Strategie

Der Deka EURO iSTOXX ex Fin Dividend+ UCITS ETF ist ein UCITS IV konformer Publikumsfonds, der die Performance des EURO iSTOXX ex Financials High Dividend 50 Preis Index (Preisindex) nachbildet. Der Index umfasst die 50 Unternehmen des EURO STOXX® ex Financials mit der höchsten Dividendenrendite aus der Eurozone. Die Bestandteile des Index werden anhand der Dividende entsprechend gewichtet.

Deka EURO STOXX 50® UCITS ETF

ETF | ETFL02 | DE000ETFL029

44,155 EUR



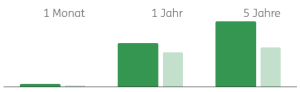
Basisinformationen

Fondskategorie	Aktienfonds
Emittent	Deka Investment GmbH
Morningstar-Rating	★★★★☆
Nachhaltigkeitsbezug (SFDR)	Artikel 6
Ausschüttungsart	Ausschüttend
Nachbildung	Vollständig
Fondswährung	EUR
Fondsvermögen	829,52 Mio. EUR (Groß)
Sparplan möglich	Nein

Kosten

Laufende Kosten p.a.	0,15 %
Performancegebühr	Nein

Performance im Vergleich



	1 Monat	1 Jahr	5 Jahre
Deka EURO STOXX 50® UCITS ETF	+1,51 %	+24,60 %	+37,13 %
EURO STOXX 50	+0,72 %	+19,60 %	+22,25 %

Zusammensetzung

Anlageklassen

Aktien	99,6 %

Positionen

ASML	7,4 %
LVMH	7,2 %
TotalEnergies	4,6 %
SAP	4,0 %

Länder

Frankreich	40,1 %
Deutschland	26,4 %
Niederlande	16,2 %
Spanien	6,3 %

Branchen

Konsumgüter	15,0 %
Technologie	14,3 %
Industrie	13,2 %
Banken	10,1 %

Strategie

Der Fonds bildet die Performance des Dow Jones EURO STOXX 50 Index (Preisindex) nach. Der Index umfasst das Segment der Blue Chips innerhalb der Aktienmärkte aus den EU-Staaten der Eurozone. Er besteht aus 50 Werten, die mindestens an einer nationalen Börse, oft aber an mehreren Europäischen Handelsplätzen, gehandelt werden können.

Der Euro Stoxx 50 misst die Wertentwicklung der 50 bedeutendsten europäischen Aktien.

Deka EURO STOXX 50® ESG Filtered UCITS ETF

ETF | ETFL46 | DE000ETFL466

97,32 EUR

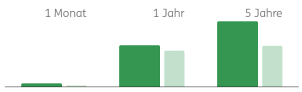
Basisinformationen

Fondskategorie	Aktienfonds
Emittent	Deka Investment GmbH
Morningstar-Rating	★★★★★
Nachhaltigkeitsbezug (SFDR)	Artikel 8
Ausschüttungsart	Thesaurierend
Nachbildung	Vollständig
Fondswährung	EUR
Fondsvermögen	115,31 Mio. EUR (Klein)
Sparplan möglich	Nein

Kosten

Laufende Kosten p.a.	0,15 %
Performancegebühr	Nein

Performance im Vergleich



	1 Monat	1 Jahr	5 Jahre
Deka EURO STOXX 50® ESG Filtered UCITS ETF	+1,64 %	+22,41 %	+35,41 %
EURO STOXX 50	+0,72 %	+19,60 %	+22,25 %

Zusammensetzung

Anlageklassen

Aktien	99,5 %

Positionen

ASML	9,5 %
LVMH	9,3 %
SAP	5,2 %
Sanofi	4,7 %

Länder

Frankreich	38,2 %
Deutschland	23,6 %
Niederlande	18,5 %
Spanien	6,2 %

Branchen

Konsumgüter	19,3 %
Technologie	18,4 %
Banken	13,0 %
Industrie	9,5 %

Strategie

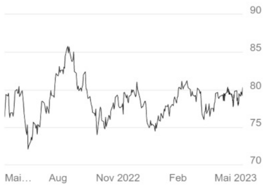
Der Deka EURO STOXX 50® ESG Filtered UCITS ETF ist ein passiv gemanagter börsengehandelter Indexfonds (Exchange Traded Fund, ETF). Ziel des Fondsmanagements ist die exakte Abbildung der Wertentwicklung des zu Grunde liegenden Index. Der Fonds bildet den EURO STOXX 50® ESG Filtered nach. Der EURO STOXX 50® ESG Filtered umfasst Aktien der 50 größten, nachhaltigen Unternehmen mit Sitz in der Eurozone. Für die Auswahl der Indexkonstituenten werden Unternehmen auf Grundlage von umweltbezogenen, sozialen oder die Unternehmensführung betreffende Kriterien (ESG-Kriterien) bewertet.

Der EURO STOXX 50® ESG umfasst Aktien der 50 größten, nachhaltigen Unternehmen mit Sitz in der Eurozone. Für die Auswahl der Indexkonstituenten werden Unternehmen auf Grundlage von umweltbezogenen, sozialen oder die Unternehmensführung betreffende Kriterien (ESG-Kriterien) bewertet.

Xtrackers MSCI World UCITS ETF 1C

ETF | A1XB5U | IE00BJ0KDQ92

80,154 EUR



Basisinformationen

Fondskategorie	Aktienfonds
Emittent	DWS Investment S.A.
Morningstar-Rating	★★★★☆
Nachhaltigkeitsbezug (SFDR)	Artikel 6
Ausschüttungsart	Thesaurierend
Nachbildung	Optimiert
Fondswährung	USD
Fondsvermögen	9,72 Mrd. EUR (Groß)

Kosten

Laufende Kosten p.a.	0,22 %
Performancegebühr	Nein

Performance im Vergleich

	1 Monat	1 Jahr	5 Jahre
Xtrackers MSCI World UCITS ETF 1C	+1,11 %	+7,25 %	+45,64 %
MSCI World	+0,37 %	+2,28 %	+44,12 %

Zusammensetzung

Anlageklassen

Aktien	99,5 %

Positionen

Apple	4,9 %
Microsoft	4,0 %
Amazon	1,8 %

Länder

USA	64,8 %
Japan	6,1 %
Großbritannien	4,7 %
Kanada	3,5 %

Branchen

Informationstechnologie	19,8 %
Finanzwesen	15,0 %
Gesundheitswesen	13,5 %
Industrie	10,8 %

Strategie

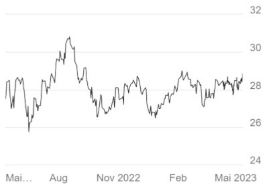
Das Anlageziel des Fonds besteht darin, die Wertentwicklung des Basiswerts, des MSCI Total Return Net World Index abzubilden. Hierzu investiert der Fonds in erster Linie in übertragbare Wertpapiere und setzt zudem derivative Techniken ein.

Der MSCI World TR Net gibt die Wertentwicklung von 23 entwickelten Aktienmärkten wieder (Total Return = Kurse + Dividenden).

Xtrackers MSCI World ESG UCITS ETF 1C

ETF | A2AQST | IE00BZ02LR44

28,815 EUR



Basisinformationen

Fondskategorie	Aktienfonds
Emittent	DWS Investment S.A.
Morningstar-Rating	★★★★★
Nachhaltigkeitsbezug (SFDR)	Artikel 8 (Hellgrün)
Ausschüttungsart	Thesaurierend
Nachbildung	Vollständig
Fondswährung	USD
Fondsvermögen	3,32 Mrd. EUR (Groß)

Kosten

Laufende Kosten p.a.	0,20 %
Fondstransaktionskosten p.a.	0,10 %
Performancegebühr	Nein

Performance im Vergleich

	1 Monat	1 Jahr	5 Jahre
Xtrackers MSCI World ESG UCITS ETF 1C	+1,39 %	+7,10 %	+50,43 %
MSCI World	+0,37 %	+2,28 %	+44,12 %

Zusammensetzung

Anlageklassen

Aktien	99,9 %

Positionen

Microsoft	8,7 %
NVIDIA	2,7 %
Alphabet A	2,6 %
Alphabet C	2,4 %

Länder

USA	66,5 %
Japan	6,1 %
Großbritannien	5,2 %
Frankreich	3,0 %

Branchen

Informationstechnologie	22,1 %
Finanzwesen	17,1 %
Gesundheitswesen	16,1 %
Industrie	10,9 %

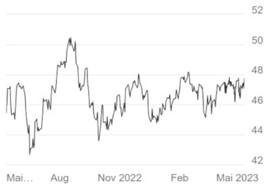
Strategie

Zur Erreichung des Anlageziels versucht der Fonds, den Index vor Gebühren und Aufwendungen nachzubilden, indem er alle (oder in Ausnahmefällen eine wesentliche Anzahl der) Aktien im Index im gleichen Verhältnis wie der Index erwirbt, wie von Gesellschaften der DWS bestimmt. Der Fonds kann Techniken und Instrumente für das Risikomanagement, zur Kostenreduzierung und zur Ergebnisverbesserung einsetzen. Diese Techniken und Instrumente können den Einsatz von Finanzkontrakten (Derivategeschäften) umfassen.

Xtrackers MSCI World Quality UCITS ETF 1C

ETF | A1103D | IE00BL25JL35

47,69 EUR



Basisinformationen

Fondskategorie	Aktienfonds
Emittent	DWS Investment S.A.
Morningstar-Rating	★★★★☆
Nachhaltigkeitsbezug (SFDR)	Artikel 6 (Grau)
Ausschüttungsart	Thesaurierend
Nachbildung	Optimiert
Fondswährung	USD
Fondsvermögen	1,01 Mrd. EUR (Groß)

Kosten

Laufende Kosten p.a.	0,26 %
Fondstransaktionskosten p.a.	0,10 %
Performancegebühr	Nein

Performance im Vergleich

	1 Monat	1 Jahr	5 Jahre
Xtrackers MSCI World Quality UCITS ETF 1C	+1,00 %	+7,18 %	+48,32 %
MSCI World	+0,37 %	+2,28 %	+44,12 %

Zusammensetzung

Anlageklassen

Aktien	99,9 %

Positionen

Home Depot	4,3 %
Microsoft	4,0 %
Apple	3,6 %
Meta	3,1 %

Länder

USA	67,6 %
Schweiz	6,5 %
Großbritannien	4,6 %
Dänemark	2,9 %

Branchen

Informationstechnologie	18,5 %
Finanzwesen	15,9 %
Gesundheitswesen	14,0 %
Industrie	10,7 %

Strategie

Die Bestandteile des DB Equity Quality Factor Index (der Index) werden aus den Aktien von Unternehmen mit hoher und mittlerer Marktkapitalisierung aus rund 31 Industrieländern ausgewählt. Die Aktien werden mithilfe einer qualitätsbasierten Strategie ausgewählt, die die Qualität der Erträge von Unternehmen analysiert. Diese Strategie basiert auf der Annahme, dass zu bestimmten Zeiten Unternehmen mit qualitativ hochwertigen Erträgen eine bessere und Unternehmen mit qualitativ schlechteren Erträgen eine schlechtere Wertentwicklung aufweisen als der Aktienmarkt insgesamt. Die Strategie verwendet eine regelbasierte Formel zur Analyse der Aktien, die Bestandteil des MSCI World Index sind, und berechnet für jede Aktie einen Qualitätsscore, der sich aus der Gesamtkapitalrendite und aufgelaufenen Beträgen ergibt. Die Indexbestandteile werden aus den im MSCI World Index vertretenen Aktien ausgewählt, mit dem Ziel, diejenigen Aktien überzugewichten, die einen höheren Qualitätsscore aufweisen, und diejenigen mit einem niedrigeren Qualitätsscore unterzugewichten. Die Gesamtkapitalrendite setzt das Nettoergebnis nach Steuern vor Dividenden und Zinsaufwendungen ins Verhältnis zum Gesamtkapital und den kurzfristigen Verbindlichkeiten des Vorjahres sowie den langfristigen Verbindlichkeiten des laufenden Jahres. Aufgelaufene Beträge messen die Wertentwicklung während eines Zeitraums, wobei Aufwendungen und damit verbundene Umsätze gegenübergestellt werden. Die zugrunde liegenden Bestandteile notieren in verschiedenen Währungen. Der Index wird auf Basis der Netto-Gesamtrendite (Total Return Net) berechnet, alle Dividenden und Ausschüttungen der Unternehmen werden nach Steuern wieder in den Aktien angelegt. Der Index wird vierteljährlich neu gewichtet.

JPMorgan Funds - Global Natural Resources A (acc) - EUR

Fonds | A0DPLL | LU0208853274

20,005 EUR

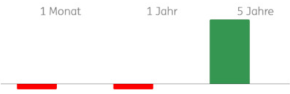
Basisinformationen

Fondskategorie	Aktienfonds
Emittent	JPMorgan Asset Management (Europe) S.à r.l.
Morningstar-Rating	★★★★☆
Nachhaltigkeitsbezug (SFDR)	Artikel 6
Ausschüttungsart	Thesaurierend
Fondswährung	EUR
Fondsvermögen	1,06 Mrd. EUR (Groß)

Kosten

Laufende Kosten p.a.	1,73 %
Performancegebühr	Nein
Ausgabeaufschlag	2,50 %

Performance im Vergleich



	1 Monat	1 Jahr	5 Jahre
JPMorgan Funds - Global Natural Resources A (acc) - EUR	-3,53 %	-3,35 %	+40,85 %

Zusammensetzung

Anlageklassen

Aktien	100,1 %

Positionen

Chevron	7,0 %
Rio Tinto PLC	6,3 %
Shell	5,7 %
Freeport McMoRan	5,4 %

Länder

USA	43,1 %
Großbritannien	15,4 %
Australien	12,8 %
Kanada	12,2 %

Branchen

Materials	50,0 %
Energy	49,1 %

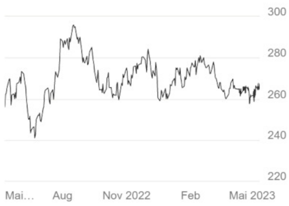
Strategie

Anlageziel des Fonds ist langfristiges Kapitalwachstum. Hierzu investiert der Fonds vorwiegend in Aktien von Unternehmen, die mit der Suche nach und Förderung von Rohstoffen, der Raffinierung, Verarbeitung und Vermarktung der Rohstoffe und deren Sekundärprodukten weltweit befasst sind.

Swisscanto (LU) Equity Fund Sustainable Global Water AT

Fonds | A0MSPX | LU0302976872

267,669 EUR



Basisinformationen

Fondskategorie	Aktienfonds
Emittent	Swisscanto Asset Management International S.A., Luxemburg
Morningstar-Rating	★★★☆☆
Nachhaltigkeitsbezug (SFDR)	Artikel 9 (Dunkelgrün)
Ausschüttungsart	Thesaurierend
Fondswährung	EUR
Fondsvermögen	390,98 Mio. EUR (Mittel)
Sparplan möglich	Nein

Kosten

Laufende Kosten p.a.	1,85 %
Performancegebühr	Nein
Ausgabeaufschlag	2,50 %

Performance im Vergleich

	1 Monat	1 Jahr	5 Jahre
Swisscanto (LU) Equity Fund Sustainable Global Water AT	+1,35 %	+0,63 %	+51,81 %

Zusammensetzung

Anlageklassen

Aktien	97,2 %
Sonstige	2,8 %

Positionen

Stantec	4,3 %
Thermo Fisher Sc Rg	4,0 %
Valmont Ind	3,8 %
Evoqua Wtr Tech Rg	3,6 %

Länder

USA	63,2 %
Kanada	6,5 %
Japan	5,0 %
Großbritannien	4,5 %

Branchen

Industrie	51,4 %
Versorger	16,9 %
IT	12,0 %
Gesundheitswesen	11,3 %

Strategie

Anlageziel ist langfristiges Kapitalwachstum, verbunden mit angemessenem Ertrag. Der Fonds investiert weltweit vorwiegend in Aktien und andere Beteiligungspapiere von Unternehmen, die Technologien, Produkte oder Dienstleistungen mit Bezug zur Wertschöpfungskette des Wassers anbieten. Avisiert werden insbesondere Unternehmen aus dem Bereich der Wasserversorgung, Wassertechnologien, Wasseraufbereitung, Wasserdienstleistungen, Wasserreinigung und Wasserrecycling.

M&G (Lux) IF 1 - Global Listed Infrastructure - EUR A Acc

Fonds | A2DXT8 | LU1665237704

15,286 EUR

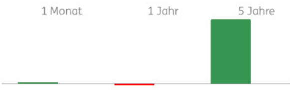
Basisinformationen

Fondskategorie	Aktienfonds
Emittent	M&G Luxembourg S.A.
Morningstar-Rating	★★★★☆
Nachhaltigkeitsbezug (SFDR)	Artikel 8 (Hellgrün)
Ausschüttungsart	Thesaurierend
Fondswährung	EUR
Fondsvermögen	2,66 Mrd. EUR (Groß)
Sparplan möglich	Nein

Kosten

Laufende Kosten p.a.	2,14 %
Performancegebühr	Nein
Ausgabeaufschlag	2,00 %

Performance im Vergleich



	1 Monat	1 Jahr	5 Jahre
M&G (Lux) IF 1 - Global Listed Infrastructure - EUR A Acc	+0,95 %	-1,58 %	+56,95 %

Zusammensetzung

Anlageklassen

Aktien	98,9 %

Positionen

Enel Societa per Azioni	4,0 %
Segro	3,7 %
Crown Castle Intl.	3,5 %
E.ON	3,5 %

Länder

USA	39,5 %
Kanada	14,7 %
Großbritannien	12,5 %
Italien	9,5 %

Branchen

Versorger	34,0 %
Kommunikation	17,4 %
Transportwesen	14,0 %
Soziale Infrastruktur	

Strategie

Mindestens 80 % des Fonds sind in Aktien von Infrastrukturgesellschaften, Investmentgesellschaften und Immobilieninvestmentgesellschaften beliebiger Größe mit Sitz in beliebigen Ländern, einschließlich Schwellenmärkten*, investiert. Der Fonds hält in der Regel Aktien von weniger als 50 Unternehmen. Die Mindestallokation von 80 % kann Stammaktien, Vorzugsaktien und Wandelanleihen umfassen. Zu den Infrastrukturgesellschaften gehören Unternehmen in den folgenden Sektoren: Versorgung, Energie, Transport, Gesundheit, Ausbildung, Sicherheit, Kommunikation und Transaktionen. Unternehmen, die mehr als 30 % ihres Umsatzes aus Strom von Kohle- und Kernkraftwerken erzielen, sind vom Anlageuniversum ausgeschlossen, ebenso Branchen wie Tabak, Alkohol, Unterhaltung für Erwachsene, Glücksspiele und Waffen. Die Prinzipien des United Nations Global Compact zu Menschenrechten, Arbeits- und Umweltrechten sowie der Korruptionsbekämpfung werden bei der Analyse der Unternehmen ebenfalls berücksichtigt.

OBJECTIVE OF THIS BROCHURE

Dear Participants,

Sustainability is on everyone's lips these days. But what does it have to do with our financial investments? In this brochure, you will learn what sustainability means when it comes to investing. In particular, you will learn about the criteria and investment strategies that are at play in the composition of sustainable funds and ETFs. Most importantly, it will help you make more informed investment decisions. Enjoy the reading!

Best regards,

Christina E. Bannier   Alix Auzepy   Florian Gärtner

WHO WE ARE

Justus-Liebig-Universität Gießen
Chair of Banking & Finance (BWL VI)
Prof. Dr. Christina E. Bannier
Licher Straße 62
D-35394 Gießen

Contact person: Florian Gärtner, M.A.
Florian.gaertner@wirtschaft.uni-giessen.de

General Knowledge: Sustainable Investments

ESG – WHAT IS IT?

Anyone interested in sustainable investing will repeatedly come across the so-called ESG criteria. The acronym stands for "**Environmental, Social and Governance**".

These three generic terms have become widely accepted by professional investors, such as pension funds or asset managers, as **important criteria for classifying sustainable investments**. ESG refers to three key areas of responsibility that can be assessed in order to determine the sustainability performance of a company.

FOR A COMPANY TO BE CONSIDERED A SUSTAINABLE COMPANY ON THE FINANCIAL MARKETS, IT MUST PERFORM WELL IN ONE OF THESE AREAS - OR PREFERABLY ALL THREE.

However, there is **no universally accepted definition of sustainability**: A "green finish" does not always mean that a financial product is truly sustainable. Rather, one must carefully examine the product and its composition, for example, the individual companies in a fund. In general, a company can be considered sustainable on the financial markets if it performs well in one of the ESG areas - or preferably in all three.

Sustainability features may vary significantly from one financial product to another. Investors should therefore be aware of the **specific investment criteria and strategy applied before investing in a fund and ETF.**

THREE CENTRAL CRITERIA

Environmental

This area includes companies that are, for example, characterized by their adoption of environmentally friendly production processes, efficient utilization of raw materials, waste reduction efforts, and a focus on minimizing greenhouse gas emissions. Additionally, companies that make significant investments in renewable energy sources and technologies geared towards combating climate change tend to achieve high scores in this crucial category.

Social

This area includes companies that set high standards for the rights and well-being of their employees. This includes maintaining strict policies against child labor and avoiding any form of discrimination based on gender, ethnicity or other minority characteristics. Companies that score high in this area provide fair compensation to their employees. They also actively engage with their suppliers to ensure compliance with the necessary standards and promote a responsible and inclusive working environment throughout their supply chain.

Governance

Companies with good corporate governance reject corruption and ensure independence and transparency in their decision-making. They place a high priority on compliance with legal requirements and create a supportive environment for whistleblowers. The sustainability strategy is integrated into their management and undergoes regular review by the company's supervisory bodies.

SUSTAINABLE FUNDS AND ETFS

How do sustainable funds or ETFs actually select their stocks? What benchmarks and criteria do they use, and what methods do they employ?

A common strategy of sustainable funds is to **exclude harmful companies and business practices**. Depending on how sustainability is understood, different exclusion criteria may come into play. Here are some examples:

- **Fossil energy:** This includes various business activities related to coal and oil. Special processes such as fracking or the use of oil sands may also be excluded.
- **Nuclear power:** This often includes not only the operation of nuclear power plants, but also the production of nuclear components and uranium mining.
- **Armaments and weapons:** This category includes, for example, manufacturers of controversial weapons and suppliers of critical components.
- **Disregard for human and labor rights:** Companies may be excluded if trade union rights are not respected, if children are used as labor, or if forced labor is used.
- **Business ethics:** This refers to unethical or criminal business practices such as corruption, tax evasion and money laundering. Companies with serious violations in these areas may be excluded.

A COMMON STRATEGY OF FUND COMPANIES IS THE EXCLUSION OF HARMFUL COMPANIES AND BUSINESS PRACTICES.

Certain **thresholds** are often used when applying exclusion criteria. This means, for example, that a company can generate up to 5% of its revenue from fossil fuels without violating the specific exclusion criteria applied by the fund.

In addition to the use of exclusion criteria (as "negative" criteria), "**positive" criteria** are also used in the composition of sustainable funds. The positive criteria are based on three additional strategies, among others: ESG integration, best in class and theme selection.

THREE ADDITIONAL STRATEGIES

ESG integration

ESG integration is an investment strategy that entails including companies in a fund's composition based on specific minimum thresholds related to financially significant ESG factors. This approach involves overweighting stocks of companies with robust ESG practices and underweighting those with poor ESG performance. To evaluate these thresholds, ESG ratings or similar scores are commonly used.

Best-in-class

The best-in-class strategy focuses on including companies in a portfolio that are industry leaders in specific sustainability areas. This means selecting companies that excel in environmental, social, and governance practices, even if they operate in traditionally unsustainable sectors, as long as they are at the forefront of positive change in their industry. For example, investing in an energy company that operates coal-fired power plants might be an option if it is also has a strong commitment to renewable energy.

Theme selection

Thematic investing is an investment strategy where funds select companies based on specific themes or problem areas they focus on, such as renewable energies, water, health, and other relevant sectors. By targeting companies aligned with these themes, the funds aim to capitalize on emerging trends and innovations within these areas.

REGULATION IN THE E.U.

The multitude of terms and strategies surrounding sustainable investing can be confusing for investors. However, the European Union has taken a step towards providing clarity and transparency with the introduction of the **Sustainable Finance Disclosure Regulation (SFDR)** in 2021.

This regulation mandates asset managers, banks, and fund companies to categorize financial products like funds and ETFs into one of the **three following categories**:

- **Article 6 funds:** These funds take into consideration ESG factors and sustainability risks in their investment decisions. However, they do not explicitly promote or market their specific environmental or social features.
- **Article 8 funds:** In contrast to Article 6 funds, Article 8 funds not only consider and integrate ESG factors and sustainability risks into their investment strategies but also actively advertise their environmental or social features.
- **Article 9 funds:** Article 9 funds are specifically designed to pursue sustainable investment objectives with measurable impacts.

The SFDR facilitates investor decision-making by requiring funds to **document** their sustainability objectives and policies in fund brochures and on their websites.

THREE CENTRAL ARTICLES

Article 6 funds: Traditional financial products

Classification under Article 6 signifies that ESG considerations are not the primary focus of the investment strategy. As part of the disclosure requirements, Article 6 funds must describe how sustainability risks could potentially influence their investment policy and impact their financial position. If sustainability risks are not considered relevant, they must provide a clear explanation for this decision.

Article 8 funds: Light green products

Classification under Article 8 indicates that the funds not only consider, but also actively promote, sustainability features in its investment policy, alongside financial objectives. These products are required to disclose how they incorporate environmental and/or social features, while also ensuring good corporate governance practices.

Article 9 funds: Dark green products

Classification under Article 9 indicates that the funds have a clearly defined sustainable investment objective. They must explicitly state an environmental, social or similar objective. These funds also seek to reduce negative impacts on other environmental or social considerations. In addition, they are must transparently disclose how they ensure the achievement of their sustainable investment objective.

KEY TAKEAWAYS

- Sustainable funds are funds that use ESG criteria in their stock selection. ESG stands for the dimensions: **Environmental, Social & Governance**. These criteria are critical to evaluating a company's sustainability performance.
- Although these criteria are commonly used, there is **no binding or universally accepted definition of sustainability**, which means that sustainability characteristics can vary greatly from one financial product to another. Differences may arise due to how the individual dimensions are measured and weighted.
- In the context of sustainable funds, the selection criteria and sustainability characteristics for the inclusion of companies in the fund are determined by the **fund's strategy**. It is common for the fund strategy to have an **emphasis on certain ESG criteria** to the **detriment of other sustainability aspects**.
- It is therefore **not necessary** for a company to have an equally good performance in all three dimensions - environment, social and governance - in order to be considered sustainable by the financial markets. A single dimension may be sufficient.
- In practice, there are **various approaches** to how sustainable funds make their selection decisions based on ESG criteria. These include exclusion procedures, ESG integration, best-in-class approach and theme selection.
- The EU Sustainable Finance Disclosure Regulation (SFDR) requires asset managers and fund companies to classify their financial products into sustainability categories known as **Article 6 (traditional financial products), Article 8 (light green financial products)** and **Article 9 (dark green financial products).** The aim is to create transparency about the actual extent to which ESG criteria are taken into account in the investment strategy and the pursuit of sustainability goals.

Bibliography

Acharya, Avidit, Matthew Blackwell, and Maya Sen (2016), “Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects”, *American Political Science Review*, 110 (3), 512–529.

Agarwal, Sumit, Souphala Chomsisengphet, Chunlin Liu, and Nicholas S. Souleles (2015), “Do Consumers Choose the Right Credit Contracts?”, *Review of Corporate Finance Studies*, 4 (2), 239–257.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel (2014), “Regulating Consumer Financial Products: Evidence from Credit Cards”, *The Quarterly Journal of Economics*, 130 (1), 111–164.

Akaike, Hirotugu (1974), “A new look at the statistical model identification”, *IEEE Transactions on Automatic Control*, 19 (6), 716–723.

Amar, Moty, Dan Ariely, Shahar Ayal, Cynthia E. Cryder, and Scott I. Rick (2011), “Winning the Battle but Losing the War: The Psychology of Debt Management”, *Journal of Marketing Research*, 48, 38–50.

Amir, Ofra, David G. Rand, and Ya’akov Kobi Gal (2012), “Economic Games on the Internet: The Effect of $1 Stakes”, *PLoS ONE*, 7 (2).

Anderson, Anders and David T. Robinson (2021), “Financial Literacy in the Age of Green Investment”, Working Paper.

Aristei, David and Manuela Gallo (2021), “Financial Knowledge, Confidence, and Sustainable Financial Behavior”, *Sustainability*, 13 (19), 10926.

Auzepy, Alix, Christina Bannier, and Florian Gärtner (2023), “"Sustainable Finance Literacy and Sustainable Investment Behavior."”, AEA RCT Registry. July 31. https://doi.org/10.1257/rct.11325-2.0.

Barber, Brad M., Adair Morse, and Ayako Yasuda (2021), “Impact investing”, *Journal of Financial Economics*, 139 (1), 162–185.

Baron, Reuben M. and David A. Kenny (1986a), “The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations.”, *Journal of Personality and Social Psychology*, 51 (6), 1173–1182.

——— (1986b), “The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations”, *Journal of Personality and Social Psychology*, 51 (6), 1173–1182.

Barreda-Tarrazona, Iván, Juan Carlos Matállín-Sàez, and Maria Rosario Balaguer-Franch (2011), “Measuring Investors’ Socially Responsible Preferences in Mutual Funds”, *Journal of Business Ethics*, 103 (2), 305–330.

Bassen, Alexander, Katrin Gödker, Florian Lüdeke-Freund, and Josua Oll (2018), “Climate Information in Retail Investors’ Decision-Making: Evidence From a Choice Experiment”, *Organization & Environment*, 32 (1), 62–82.

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth, and Harald Baayen (2018), “Parsimonious Mixed Models”, Working Paper.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker (2015), “Fitting Linear Mixed-Effects Models Using lme4”, *Journal of Statistical Software*, 67 (1), 1–48.

Bauer, Rob, Marco Ceccarelli, Katrin Gödker, and Paul Smeets (2022), “Measuring sustainable preferences of pension members A methodological proposition and a case study of a UK pension fund”,Technical report.

Bauer, Rob, Tobias Ruof, and Paul Smeets (2021), “Get Real! Individuals Prefer More Sustainable Investments”, *The Review of Financial Studies*, 34 (8), 3976–4043.

Bauer, Rob and Paul Smeets (2015), “Social identification and investment decisions”, *Journal of Economic Behavior & Organization*, 117, 121–134.

Bazley, William J., Henrik Cronqvist, and Milica Mormann (2021), “Visual Finance: The Pervasive Effects of Red on Investor Behavior”, *Management Science*, 67 (9), 5616–5641.

Becker, Martin G., Fabio Martin, and Andreas Walter (2022), “The power of ESG transparency: The effect of the new SFDR sustainability labels on mutual funds and individual investors”, *Finance Research Letters*, 47, 102708.

Benartzi, Shlomo and Richard H. Thaler (1999), “Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments”, *Management Science*, 45 (3), 364–381.

——— (2001), “Naive Diversification Strategies in Defined Contribution Saving Plans”, *American Economic Review*, 91 (1), 79–98.

Berg, Janine (2016), “Income Security in the On-demand Economy: Findings and Policy Lessons from a Survey of Crowdworkers”, Conditions of Work and Employment Series No. 74.

Bergman, Peter (2021), “Parent-Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment”, *Journal of Political Economy*, 129 (1), 286–322.

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz (2012), “Evaluating Online Labor Markets for Experimental Research: Amazon.com’s Mechanical Turk”, *Political Analysis*, 20 (3), 351–368.

Besharat, Ali, François A. Carrillat, and Daniel M. Ladik (2014), “When Motivation is against Debtors’ Best Interest: The Illusion of Goal Progress in Credit Card Debt Repayment”, *Journal of Public Policy & Marketing*, 33 (2), 143–158.

Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian (2017), “Does Aggregated Returns Disclosure Increase Portfolio Risk Taking?”, *Review of Financial Studies*, 30 (6), 1971–2005.

——— (2018), “Behavioral Household Finance”, in B. Douglas Bernheim, David Laibson, Stefano DellaVigna ed. *Handbook of Behavioral Economics*, Chap. 3, 177–276: Elsevier.

Bethlendi, András, László Nagy, and András Póra (2022), “Green finance: the neglected consumer demand”, *Journal of Sustainable Finance & Investment*, 0 (0), 1–19.

Blumenstock, Robin, Michael Callen, and Tarek Ghani (2018), “Why Do Defaults Affect Behavior? Experimental Evidence from Afghanistan”, *American Economic Review*, 108 (10), 2868–2901.

Bofinger, Yannik, Kim J. Heyden, and Björn Rock (2022), “Corporate social responsibility and market efficiency: Evidence from ESG and misvaluation measures”, *Journal of Banking & Finance*, 134, 106322.

Briere, Marie and Stefano Ramelli (2021), “Responsible Investing and Stock Allocation”, *SSRN Electronic Journal*.

Brodback, Daniel, Nadja Guenster, and David Mezger (2019), “Altruism and egoism in investment decisions”, *Review of Financial Economics*, 37 (1), 118–148.

Cai, Cynthia Weiyi (2019), “Nudging the financial market? A review of the nudge theory”, *Accounting & Finance*, Early View, 1–25.

Camerer, Colin F., Anna Dreber, Felix Holzmeister et al. (2018), “Evaluating the Replicability of Social Science Experiments in Nature and Science between 2010 and 2015”, *Nature Human Behaviour*, 2 (9), 637–644.

Camerer, Colin F. and Robin M. Hogarth (1999), “The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework”, *Journal of Risk and Uncertainty*, 19 (1-3), 7–42.

Camerer, Colin Farrell (2015), “The Promise and Success of Lab-Field Generalizability in Experimental Economics: A Critical Reply to Levitt and List”, in Fréchette, Guillaume and Andrew Schotter eds. *Handbook of Experimental Economic Methodology*, Chap. 14, 249–296: Oxford University Press.

Carpena, Fenella and Bilal Zia (2020), “The causal mechanism of financial education: Evidence from mediation analysis”, *Journal of Economic Behavior &amp Organization*, 177, 143–184.

Castleman, Benjamin L. and Lindsay C. Page (2015), “Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates?”, *Journal of Economic Behavior & Organization*, 115, 144–160.

Ceccarelli, Marco, Stefano Ramelli, and Alexander F. Wagner (2019), “When Investors Call for Climate Responsibility, How Do Mutual Funds Respond?”, *SSRN Electronic Journal*.

Chandler, Jesse, Pam Mueller, and Gabriele Paolacci (2014), “Nonnaïveté among Amazon Mechanical Turk Workers: Consequences and Solutions for Behavioral Researchers”, *Behavior Research Methods*, 46 (1), 112–130.

Chandler, Jesse, Gabriele Paolacci, Eyal Peer, Pam Mueller, and Kate A. Ratliff (2015), “Using Nonnaive Participants Can Reduce Effect Sizes”, *Psychological Science*, 26 (7), 1131–1139.

Chandler, Jesse and Danielle Shapiro (2016), “Conducting Clinical Research Using Crowdsourced Convenience Samples”, *Annual Review of Clinical Psychology*, 12, 53–81.

Chesney, Marc and Drien-Paul Lambillon (2023), “How green is ‘dark green’? An analysis of SFDR Article 9 funds”, Working Paper.

Ching, Andrew T. and Fumiko Hayashi (2010), “Payment card rewards programs and consumer payment choice”, *Journal of Banking & Finance*, 34 (8), 1773–1787.

Chmielewski, Michael and Sarah C. Kucker (2020), “An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results”, *Social Psychological and Personality Science*, 11 (4), 1–10.

Choi, James J., David Laibson, and Brigitte C. Madrian (2010), “Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds”, *Review of Financial Studies*, 23 (4), 1405–1432.

Clogg, Clifford C, Eva Petkova, and Adamantios Haritou (1995), “Statistical Methods for Comparing Regression Coefficients Between Models”, *American Journal of Sociology*, 100 (5), 1261–1293.

Cohen, Jonathan, Keith Marzilli Ericson, David Laibson, and John Myles White (2020), “Measuring Time Preferences”, *Journal of Economic Literature*, 58 (2), 299–347.

Coppock, Alexander (2019), “Generalizing from Survey Experiments Conducted on Mechanical Turk: A Replication Approach”, *Political Science Research and Methods*, 9 (3), 613–628.

Crump, Matthew J. C., John V. McDonnell, and Todd M. Gureckis (2013), “Evaluating Amazon’s Mechanical Turk as a Tool for Experimental Behavioral Research”, *PLoS ONE*, 8 (3).

DellaVigna, Stefano (2009), “Psychology and Economics: Evidence from the Field”, *Journal of Economic Literature*, 47 (2), 315–372.

Dembo, Aluma, Shachar Kariv, Matthew Polisson, and John K.-H. Quah (2021), “Ever Since Allais”, Working Paper.

Dhami, Sanjit (2016), *The Foundations of Behavioral Economic Analysis*: Oxford University Press.

DIA (2020), “Wie halten es die Anleger mit der Nachhaltigkeit?”,Technical report, Deutsches Institut für Altersvorsorge.

Dimant, Eugen, Gerben A. van Kleef, and Shaul Shalvi (2020), “Requiem for a Nudge: Framing effects in nudging honesty”, *Journal of Economic Behavior & Organization*, 172, 247–266.

Doss, Christopher J., Erin M. Fahle, Susanna Loeb, and Benjamin N. York (2019), “More Than Just a Nudge: Supporting Kindergarten Parents with Differentiated and Personalized Text-Messages”, *Journal of Human Resources*, 54 (3), 567–603.

Dumas, Christel and Céline Louche (2015), “Collective Beliefs on Responsible Investment”, *Business & Society*, 55 (3), 427–457.

Enke, Benjamin and Thomas Graeber (2021), “Cognitive Uncertainty in Intertemporal Choice”, Working Paper.

——— (2023), “Cognitive Uncertainty”, *The Quarterly Journal of Economics*, 138 (4), 2021–2067.

European Parliament and Council of the European Union (2019), “REGULATION (EU) 2019/2088 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL”, `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32019R2088`.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde (2023), “The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences”, *Management Science*, 69 (4), 1935–1950.

Federal Reserve Bank of New York (2022), “Quarterly Report on Household Debt and Credit”, 2nd Quarter, May 2022.

Federal Reserve Board (2022), “Federal Reserve Statistical Release on Consumer Credit”, March 2022.

Fernandes, Daniel, John G. Lynch, and Richard G. Netemeyer (2014), “Financial Literacy, Financial Education, and Downstream Financial Behaviors”, *Management Science*, 60 (8), 1861–1883.

Filippini, Masimo, Markus Leippold, and Tobias Wekhof (2023), “Sustainable Finance Literacy and the Determinants of Sustainable Investing”, Swiss Finance Institute Research Paper Series No. 22-02.

Ford, John B. (2017), “Amazon’s Mechanical Turk: A Comment”, *Journal of Advertising*, 46 (1), 156–158.

Foster, Kevin, Erik Meijer, Scott D. Schuh, and Michael A. Zabek (2011), “The 2009 Survey of Consumer Payment Choice”, Public Policy Discussion Paper, Boston Reserve Bank.

Friede, Gunnar, Timo Busch, and Alexander Bassen (2015), “ESG and financial performance: aggregated evidence from more than 2000 empirical studies”, *Journal of Sustainable Finance & Investment*, 5 (4), 210–233.

Frydman, Cary and Baolian Wang (2020), “The Impact of Salience on Investor Behavior: Evidence from a Natural Experiment”, *Journal of Finance*, 75 (1), 229–276.

Gathergood, John, David Hirshleifer, David Leake, Hiroaki Sakaguchi, and Neil Stewart (2020), “Naïve *Buying* Diversification and Narrow Framing by Individual Investors”, NBER Working Paper.

Gathergood, John, Neale Mahoney, Neil Stewart, and Jörg Weber (2019), “How Do Individuals Repay Their Debt? The Balance-Matching Heuristic”, *American Economic Review*, 109 (3), 844–875.

Geiger, Sonja and Brigitte Holzhauer (2020), “Weiterentwicklung einer Skala zur Messung von zentralen Kenngrößen des Umweltbewusstseins”,Technical report, Umweltbundesamt.

Gigerenzer, Gerd, Wolfgang Gaissmaier, Elke Kurz-Milcke, Lisa M. Schwartz, and Steven Woloshin (2007), “Helping Doctors and Patients Make Sense of Health Statistics”, *Psychological Science in the Public Interest*, 8 (2), 53–96.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, Zhenhao Tan, Stephen Utkus, and Xiao Xu (2023), “Four Facts About ESG Beliefs and Investor Portfolios”,Technical report.

Gneezy, Uri and Jan Potters (1997), “An Experiment on Risk Taking and Evalutation Periods”, *Quarterly Journal of Economics*, 112 (2), 631–645.

Goodman, Joseph E. and Gabriele Paolacci (2017), “Crowdsourcing Consumer Research”, *Journal of Consumer Research*, 44 (1), 196–210.

Goodman, Joseph K., Cynthia E. Cryder, and Amar Cheema (2013), “Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples”, *Journal of Behavioral Decision Making*, 26, 213–224.

Gorbachev, Olga and María José Luengo-Prado (2019), “The Credit Card Debt Puzzle: The Role of Preferences, Credit Access Risk, and Financial Literacy”, *Review of Economics and Statistics*, 101 (2), 294–309.

GSIA (2021), “Global Sustainable Investment Review 2020”,Technical report, Global Sustainable Investment Alliance.

Gutsche, Gunnar and Andreas Ziegler (2019), “Which private investors are willing to pay for sustainable investments? Empirical evidence from stated choice experiments”, *Journal of Banking & Finance*, 102, 193–214.

Gutsche, Gunnar and Bernhard Zwergel (2020), “Investment Barriers and Labeling Schemes for Socially Responsible Investments”, *Schmalenbach Business Review*, 72 (2), 111–157.

Gärtner, Florian and Darwin Semmler (2022), *Elemental Financial Decisions - Preregistration*: AEA RCT Registry.

Gärtner, Florian, Darwin Semmler, and Christina E. Bannier (2023), "What could possibly go wrong? Predictable misallocation in simple debt repayment experiments", *Journal of Economic Behavior & Organization*, 205, 28–43.

Handel, Benjamin and Joshua Schwartzstein (2018), "Frictions or Mental Gaps: What's Behind the Information We (Don't) Use and When Do We Care?", *Journal of Economic Perspectives*, 32 (1), 155–78.

Hara, Kotaro, Abi Adams, Kristy Milland, Saiph Savage, Chris Callison-Burch, and Jeffrey P. Bigham (2017), "A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk", Working Paper.

Hartzmark, Samuel M. and Abigail B. Sussman (2019), "Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows", *The Journal of Finance*, 74 (6), 2789–2837.

Hauser, David J., Gabriele Paolacci, and Jesse Chandler (2019), "Common Concerns with MTurk as a Participant Pool: Evidence and Solutions", in Kardes, Frank R., Paul M. Herr, and Norbert Schwarz eds. *Handbook of Research Methods in Consumer Psychology*, Chap. 17: Taylor & Francis, London.

Hauser, David J. and Norbert Schwarz (2016), "Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants", *Behavior Research Methods*, 48 (1), 400–407.

Heeb, Florian, Julian F Kölbel, Falko Paetzold, and Stefan Zeisberger (2022), "Do Investors Care about Impact?", *The Review of Financial Studies*, hhac066.

Hendriks, Achim (2012), "SoPHIE - Software Platform for Human Interaction Experiments", Working Paper.

Hershfield, Hal E. and Neal J. Roese (2015), "Dual payoff scenario warnings on credit card statements elicit suboptimal payoff decisions", *Journal of Consumer Psychology*, 25 (1), 15–27.

Hertwig, Ralph and Till Grüne-Yanoff (2017), "Nudging and Boosting: Steering or Empowering Good Decisions", *Perspectives on Psychological Science*, 12 (6), 973–986, PMID: 28792862.

Hoffrage, Ulrich, Samuel Lindsey, Ralph Hertwig, and Gerd Gigerenzer (2000), "Communicating Statistical Information", *Science*, 290 (5500), 2261–2262.

Hong, Harrison and Leonard Kostovetsky (2012), “Red and blue investing: Values and finance”, *Journal of Financial Economics*, 103 (1), 1–19.

Horton, John J., David G. Rand, and Richard J. Zeckhauser (2011), “The online laboratory: conducting experiments in a real labor market”, *Experimental Economics*, 14 (3), 399–425.

H.R.627 — 111th Congress (2009), “Credit Card Accountability Responsibility and Disclosure Act of 2009”, `https://www.congress.gov/bill/111th-congress/house-bill/627`.

Huff, Connor and Dustin Tingley (2015), ““Who are these people?” Evaluating the demographic characteristics and political preferences of MTurk survey respondents”, *Research & Politics*, 2 (3), 1–12.

Hünermund, Paul and Louw Beyers (2022), “On the Nuisance of Control Variables in Regression Analysis”, Working Paper.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto (2011), “Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies”, *American Political Science Review*, 105 (4), 765–789.

Jones, Lauren E., Cäzilia Loibl, and Sharon Tennyson (2015), “Effects of informational nudges on consumer debt repayment behaviors”, *Journal of Economic Psychology*, 51, 16–33.

Kahneman, Daniel and Amos Tversky (1979), “Prospect Theory: An Analysis of Decision under Risk”, *Econometrica*, 47 (2), 263–292.

Kaiser, Tim and Lukas Menkhoff (2020), “Financial education in schools: A meta-analysis of experimental studies”, *Economics of Education Review*, 78.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman (2016), “Getting to the Top of Mind: How Reminders Increase Saving”, *Management Science*, 62 (12), 3393–3411.

Kaufmann, Christine, Martin Weber, and Emily Haisley (2013), “The Role of Experience Sampling and Graphical Displays on One’s Investment Risk Appetite”, *Management Science*, 59 (2), 323–340.

Kaustia, Markku and Sami Torstila (2011), “Stock market aversion? Political preferences and stock market participation”, *Journal of Financial Economics*, 100 (1), 98–112.

Kees, Jeremy, Christopher Berry, Scot Burton, and Kim Sheehan (2017), "An Analysis of Data Quality: Professional Panels, Student Subject Pools, and Amazon's Mechanical Turk", *Journal of Advertising*, 46 (1), 141–155.

Keys, Benjamin J., Devin G. Pope, and Jaren C. Pope (2016), "Failure to Refinance", *Journal of Financial Economics*, 122 (3), 482–499.

Keys, Benjamin J. and Jialan Wang (2019), "Minimum Payments and Debt Paydown in Consumer Credit Cards", *Journal of Financial Economics*, 131 (3), 528–548.

Killen, Catherine P., Joana Geraldi, and Alexaner Kock (2020), "The role of decision makers' use of visualizations in project portfolio decision making", *International Journal of Project Management*, 38, 267–277.

Kolb, David A. (1984), *Experiential Learning: Experience as the Source of Learning and Development*: Englewood Cliffs, NJ: Prentice Hall.

Köszegi, Botond and Matthew Rabin (2008), "Revealed Mistakes and Revealed Preferences", in Caplin, Andrew and Andrew Schotter eds. *The Foundations of Positive and Normative Economics: A Handbook*, Chap. 8, 193–209: Oxford University Press.

Kraft, Matthew A. and Todd Rogers (2015), "The underutilized potential of teacher-to-parent communication: Evidence from a field experiment", *Economics of Education Review*, 47, 49–63.

Kreuter, Matthew W. and Victor J. Strecher (1996), "Do tailored behavior change messages enhance the effectiveness of health risk appraisal? Results from a randomized trial", *Health Education Research*, 11 (1), 97–105.

Krupnikov, Yanna and Adam Seth Levine (2014), "Cross-Sample Comparisons and External Validity", *Journal of Experimental Political Science*, 1 (1), 59–80.

Levitt, Steven D. and John A. List (2007), "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?", *Journal of Economic Perspectives*, 21 (2), 153–174.

(2006), *The Construction of Preference*: Cambridge University Press.

Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S. Mitchell (2020), "Assessing the impact of financial education programs: A quantitative model", *Economics of Education Review*, 78.

Lusardi, Annamaria and Olivia S. Mitchell (2011), "Financial literacy around the world: an overview", *Journal of Pension Economics and Finance*, 10 (4), 497–508.

——— (2014), "The Economic Importance of Financial Literacy: Theory and Evidence", *Journal of Economic Literature*, 52 (1), 5–44.

Lusardi, Annamaria, Anya Savikhin Samek, Arie Kapteyn, Lewis Glinert, Angela Hung, and Aileen Heinberg (2017), "Visual tools and narratives: new ways to improve financial literacy", *Journal of Pension Economics Finance*, 16 (Special Issue 3: Financial Knowledge and Key Retirement Outcomes), 297–323.

Lusardi, Annamaria and Peter Tufano (2015), "Debt literacy, financial experiences, and overindebtedness", *Journal of Pension Economics & Finance*, 14 (4), 332–368.

Matuschek, Hannes, Reinhold Kliegl, Shravan Vasishth, Harald Baayen, and Douglas Bates (2017), "Balancing Type I error and power in linear mixed models", *Journal of Memory and Language*, 94, 305–315.

McCredie, Morgan N. and Leslie C. Morey (2018), "Who Are the Turkers? A Characterization of MTurk Workers Using the Personality Assessment Inventory", *Assessment*, 26 (5), 759–766.

Medina, Paolina C. (2021), "Side Effects of Nudging: Evidence from a Randomized Intervention in the Credit Card Market", *Review of Financial Studies*, 34 (5), 2580–2607.

Miller, Joshua D., Michael Crowe, Brandon Weiss, Jessica L. Maples-Keller, and Donald R. Lynam (2017), "Using Online, Crowdsourcing Platforms for Data Collection in Personality Disorder Research: The Example of Amazon's Mechanical Turk", *Personality Disorders: Theory, Research, and Treatment*, 8 (1), 26–34.

Mills, Stuart (2022), "Personalized nudging", *Behavioral Public Policy*, 6 (1), 150–159.

Moore, Danna L. (2003), "Survey of financial literacy in Washington State: Knowledge, behavior, attitudes, and experiences", SESRC Technical Report 03-39, Social and Economic Sciences Research Center, Washington State University.

Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese (2015), "The Generalizability of Survey Experiments", *Journal of Experimental Political Science*, 2 (2), 109–138.

Navarro-Martinez, Daniel, Linda C. Salisbury, Katherine N. Lemon, Neil Stewart, William J. Matthews, and Adam J.L. Harris (2011), "Minimum Required Payment and Supplemental Information Disclosure Effects on Consumer Debt Repayment Decisions", *Journal of Marketing Research*, 48 (SPL), 60–77.

Newell, Allen and Herbert Simon (1972), *Human Problem Solving*: Prentice-Hall.

Nielsen, Kirby and John Rehbeck (2022), "When Choices are Mistakes", *American Economic Review*, 112 (7), 2237–2268.

Oehmke, Martin and Marcus Opp (2020), "A Theory of Socially Responsible Investment", Working Paper.

Open Science Collaboration (2015), "Estimating the Reproducibility of Psychological Science", *Science*, 349 (6251), 943–951.

Ozyılmaz, Hakan and Guangli Zhang (2020), "The Debt Payment Puzzle: An Experimental Investigation", Working Paper.

Paetzold, Falko and Timo Busch (2014), "Unleashing the Powerful Few", *Organization & Environment*, 27 (4), 347–367.

Page, Lindsay C., Benjamin L. Castleman, and Katharine Meyer (2020), "Customized Nudging to Improve FAFSA Completion and Income Verification", *Educational Evaluation and Policy Analysis*, 42 (1), 3–21.

Palan, Stefan and Christian Schitter (2018), "Prolific.ac—A subject pool for online experiments", *Journal of Behavioral and Experimental Finance*, 17, 22–27.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis (2010), "Running Experiments on Amazon Mechanical Turk", *Judgment and Decision Making*, 5 (5), 411–419.

Pástor, L'uboš, Robert F. Stambaugh, and Lucian A. Taylor (2021), "Sustainable investing in equilibrium", *Journal of Financial Economics*, 142 (2), 550–571.

Paternoster, Ray, Robert Brame, Paul Mazerolle, and Alex Piquero (1998), "Using the Correct Statistical Test for Equality of Regression Coefficients", *Criminology*, 36, 859 – 866.

Pedersen, Lasse Heje, Shaun Fitzgibbons, and Lukasz Pomorski (2021), "Responsible investing: The ESG-efficient frontier", *Journal of Financial Economics*, 142 (2), 572–597.

Peer, Eyal, Joachim Vosgerau, and Alessandro Acquisti (2014), "Reputation as a sufficient condition for data quality on Amazon Mechanical Turk", *Behavior Research Methods*, 46 (4), 1023–1031.

Phillips, Susan D. and Bernadette Johnson (2019), "Inching to Impact: The Demand Side of Social Impact Investing", *Journal of Business Ethics*, 168 (3), 615–629.

Ponce, Alejandro, Enrique Seira, and Guillermo Zamarripa (2017), "Borrowing on the Wrong Credit Card? Evidence from Mexico", *American Economic Review*, 107 (4), 1335–1361.

de Quidt, Jonathan, Johannes Haushofer, and Christopher Roth (2018), "Measuring and Bounding Experimenter Demand", *American Economic Review*, 108 (11), 3266–3302.

Quiggin, John (1982), "A theory of anticipated utility", *Journal of Economic Behavior & Organization*, 3 (4), 323–343.

——— (1993), *Generalized Expected Utility Theory: The Rank-Dependent Model*: Dordrecht: Kluwer.

Rabin, Matthew (2013), "Incorporating Limited Rationality into Economics", *Journal of Economic Literature*, 51 (2), 528–543.

Ramsey, Sarah R., Kristen L. Thompson, Melissa McKenzie, and Alan Rosenbaum (2016), "Psychological research in the internet age: The quality of web-based data", *Computers in Human Behavior*, 58, 354–360.

Riedl, Arno and Paul Smeets (2017), "Why Do Investors Hold Socially Responsible Mutual Funds?", *The Journal of Finance*, 72 (6), 2505–2550.

Roulin, Nicolas (2015), "Don't Throw the Baby Out With the Bathwater: Comparing Data Quality of Crowdsourcing, Online Panels, and Student Samples", *Industrial and Organizational Psychology*, 8 (2), 190–196.

Salisbury, Linda C. (2014), "Minimum Payment Warnings and Information Disclosure Effects on Consumer Debt Repayment Decisions", *Journal of Public Policy & Marketing*, 33 (1), 49–64.

Scherer, Bernd and Milot Hasaj (2023), "Greenlabelling: How valuable is the SFDR Art 9 label?", *Journal of Asset Management*.

Schwarz, Gideon (1978), "Estimating the Dimension of a Model", *The Annals of Statistics*, 6 (2), 461–464.

Sims, Christopher A. (2003), "Implications of rational inattention", *Journal of Monetary Economics*, 50 (3), 665–690.

Skinner, Celette S., Victor J. Strecher, and Harm Hospers (1994), "Physicians' recommendations for mammography: do tailored messages make a difference?", *American Journal of Public Health*, 84 (1), 43–49.

Snowberg, Erik and Leeat Yariv (2021), "Testing the Waters: Behavior across Participant Pools", *American Economic Review*, 111 (2), 687–719.

Sobel, Michael E. (1982), "Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models", *Sociological Methodology*, 13, 290–312.

Soll, Jack B., Ralph L. Keeney, and Richard P. Larrick (2013), "Consumer Misunderstanding of Credit Card Use, Payments, and Debt: Causes and Solutions", *Journal of Public Policy & Marketing*, 32 (1), 66–81.

Stango, Victor and Jonathan Zinman (2016), "Borrowing High versus Borrowing Higher: Price Dispersion and Shopping Behavior in the U.S. Credit Card Market", *Review of Financial Studies*, 29 (4), 979–1006.

Stiftung Warentest (2021), "Nachhaltig Geld anlegen: ökologisch, sozial und ethisch investieren",: Stiftung Warentest.

Sunstein, Cass R. (2012), "Impersonal Default Rules vs. Active Choices vs. Personalized Default Rules: A Triptych", Working Paper.

Tang, Ning and Paula C. Peter (2015), "Financial knowledge acquisition among the young: The role of financial education, financial experience, and parents' financial experience", *Financial Services Review*, 24 (2), 119–137.

Thaler, Richard H. (1985), "Mental Accounting and Consumer Choice", *Marketing Science*, 4 (3), 199–214.

——— (2018), "Nudge, not sludge", *Science*, 361 (6401), 431–431.

Thaler, Richard H. and Cass R. Sunstein (2021), *Nudge: The Final Edition*: Penguin.

Thaler, Richard H. and Will Tucker (2013), "Smarter Information, Smarter Consumers", *Harvard Business Review*, 91 (1–2), 44–54.

Thorndike, Robert L. (1953), "Who belongs in the family?", *Psychometrika*, 18, 267–276.

Tingley, Dustin, Yamamoto Teppei, Kentaro Hirose, Luke Keele, and Kosuke Imai (2014), "mediation: R Package for Causal Mediation Analysis", *Journal of Statistical Software*, 59 (5), 1–38.

Tversky, Amos and Daniel Kahneman (1974), "Judgment under Uncertainty: Heuristics and Biases", *Science*, 186 (4157), 1124–1131.

——— (1981), "The Framing of Decisions and the Psychology of Choice", *Science*, 211 (4481), 453–458.

——— (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty", *Journal of Risk and Uncertainty*, 5 (4), 297–323.

Vulkan, Nir (2000), "An Economist's Perspective on Probability Matching", *Journal of Economic Surveys*, 14 (1), 101–118.

Wagner, Jamie and William B. Walstad (2019), "The Effects of Financial Education on Short-Term and Long-Term Financial Behaviors", *Journal of Consumer Affairs*, 53 (1), 234–259.

von Wallis, Miriam and Christian Klein (2014), "Ethical requirement and financial interest: a literature review on socially responsible investing", *Business Research*, 8 (1), 61–98.

Williams, R. (2012), "Using the margins command to estimate and interpret adjusted predictions and marginal effects", *Stata Journal*, 12 (2), 308–331, `https://www.stata-journal.com/article.html?article=st0260`.

Wolfson, Shael N. and James R. Bartkus (2013), "An Assessment of Experiments run on Amazon's Mechanical Turk", *Mustang Journal of Business and Ethics*, 5, 119–129.

Wooldridge, Jeffrey M. (2010), *Econometric Analysis of Cross Section and Panel Data*: The MIT Press.

Zhou, Guangyou, Lian Liu, and Sumei Luo (2022), "Sustainable development, ESG performance and company market value: Mediating effect of financial performance", *Business Strategy and the Environment*, 31 (7), 3371–3387.

Zinman, Jonathan (2015), "Household Debt: Facts, Puzzles, Theories, and Policies", *Annual Review of Economics*, 7, 251–276.

Zizzo, Daniel J. (2010), "Experimenter demand effects in economic experiments", *Experimental Economics*, 13 (1), 75–98.

Zizzo, Daniel John (2009), "Experimenter demand effects in economic experiments", *Experimental Economics*, 13 (1), 75–98.

Zwickle, Adam and Keith Jones (2018), "Sustainability Knowledge and Attitudes - Assessing Latent Constructs", in Filho, Walter Leal, Robert W. Marans, and Dr. John Callewaert eds. *Handbook of Sustainability and Social Science Research*, 435–451: Springer International Publishing.

Affidavit

Ich erkläre hiermit, dass ich die vorgelegten und nachfolgend aufgelisteten Aufsätze selbstständig und nur mit den Hilfen angefertigt habe, die im jeweiligen Aufsatz angegeben oder zusätzlich in der nachfolgenden Liste aufgeführt sind. In der Zusammenarbeit mit den angeführten Koautoren war ich mindestens anteilig beteiligt. Bei den von mir durchgeführten und in den Aufsätzen erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis niedergelegt sind, eingehalten.

________________________ 07.11.2023

Signature author Date

Submitted Papers:

1. Gärtner, Florian, Darwin Semmler and Christina E. Bannier (2022), "What could possibly go wrong? Predictable misallocation in simple debt repayment experiments". Journal of Economic Behavior & Organization, 205, 28–43. (Chapters I and II).

    (a) Gärtner, Florian, Darwin Semmler and Christina E. Bannier (2019), "Identifying situations and behavior leading to non-optimal credit card repayment". Working Paper (Chapter I).

    (b) Gärtner, Florian, Darwin Semmler and Christina E. Bannier (2019), "Does credit card repayment behavior depend on the presentation of interest payments? The cuckoo fallacy". Working Paper (Chapter II).

2. Bofinger, Yannik, Florian Gärtner and Darwin Semmler (2021), "Addressing consumer misunderstanding in credit card debt repayment: Policy suggestions beyond the CARD Act". Working Paper (Chapter III).

3. Gärtner, Florian and Darwin Semmler (2022), "Elementary Financial Decisions". Working Paper (Chapter IV)

4. Auzepy, Alix, Christina E. Bannier and Florian Gärtner (2023), "Looking beyond ESG preferences: The role of sustainable finance literacy in sustainable investing". Working Paper (Chapter V)