JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Doctoral Thesis

Essays on Behavioral Finance and Corporate Governance

Submitted to: Justus Liebig University Giessen

Submitted on: July 26, 2022

Author: Darwin Semmler

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 **7**

List of Tables **9**

Acknowledgements **11**

General Introduction **13**

I What could possibly go wrong? Triggering misallocation **I-17**

I.1 Introduction I-19

I.2 Experimental setup and hypotheses I-22

I.3 Data I-29

I.4 Results I-31

I.5 Robustness checks I-38

I.6 Exploratory within-subject analyses I-39

I.7 Conclusion I-53

II What could possibly go wrong? Nudging and the Cuckoo Fallacy **II-57**

II.1 Introduction II-59

II.2 Experimental setup and hypotheses II-61

II.3 Data II-65

II.4 Results II-68

II.5 Robustness checks II-74

II.6 An additional experiment with independent rounds as a robustness check II-75

II.7 Conclusion II-83

III Addressing consumer misunderstanding in credit card debt repayment: Policy suggestions beyond the CARD Act III-85

III.1 Introduction III-87

III.2 Hypotheses development III-92

III.3 Experimental design and data III-97

III.4 Results III-100

III.5 Additional analyses and robustness checks III-108

III.6 Policy implication and discussion III-113

III.7 Conclusion III-120

IV Elemental Financial Decisions IV-123

IV.1 Introduction IV-125

IV.2 Theoretical background IV-130

IV.3 Experiment #1 . . . . . . IV-133

IV.3.1 General design IV-133

IV.3.2 Experimental Variables IV-136

IV.3.3 Results IV-139

IV.3.4 Robustness checks IV-147

IV.3.5 Discussion of Experiment 1 . . . . . . IV-148

IV.4 Experiment #2 . . . . . . IV-149

IV.4.1 General design IV-149

IV.4.2 Results IV-151

IV.4.3 Robustness checks IV-156

IV.5 Comparison of both experiments IV-157

IV.6 General Discussion and Conclusion IV-162

V Ethnic Diversity and the Glass Cliff - An examination of French CAC40 Boards V-167

V.1 Introduction V-169

V.2 Background and hypotheses development V-173

V.3 Data and method V-179

V.3.1 Data on appointment level V-179

V.3.2 Data on firm level V-183

V.3.3 Analyses V-187

V.4 Results V-191

V.5 Robustness checks V-197

V.6 Conclusion V-202

Appendix I (to Chapter I) 205

Appendix II (to Chapter II) 219

Appendix III (to Chapter III) 231

Appendix IV (to Chapter IV) 247

Appendix V (to Chapter V) 269

Bibliography 289

Affidavit 305

List of Figures

I.1 Comparison of choices between control and fallacy scenarios I-33

I.2 Proportion of optimal answers I-41

I.3 Visualization of the decision matrix I-45

I.4 Determination of cluster numbers I-50

II.1 Round development of misallocation and interaction plot between treatment and interest class II-70

II.2 Interaction plot between treatment and scenario type II-82

III.1 Development of misallocation over rounds III-109

IV.1 Experiment #1: Average misallocation and uncertainty IV-142

IV.2 Experiment #2: Average misallocation and uncertainty IV-153

V.1 Development of minority representation in boards V-181

V.2 Firm performance before and after appointments (split by appearance) . V-184

V.3 Firm performance before and after appointments (split by nationality) . V-185

List of Tables

I.1 Repayment options and scenarios I-28
I.2 Summary statistics of participants I-31
I.3 Logistic regression (fallacy-implicated option) I-35
I.4 Logistic regression (optimal option) I-36
I.5 Transition matrices I-42
I.6 Proportion of optimal or fallacy-implicated answers I-44
I.7 Correlation matrix of fallacy-implicated answers I-46
I.8 Behavior in the scenario pairs I-48
I.9 Description of cluster means I-51

II.1 Summary statistics of participants II-67
II.2 Misallocation split by treatment II-68
II.3 Misallocation split by round class II-73
II.4 Summary statistics of participants (additional experiment) II-78
II.5 Misallocation in additional experiment II-81

III.1 Descriptive statistics of participants III-101
III.2 Descriptive statistics of misallocation per treatment III-103
III.3 OLS of misallocation III-106

III.4 Increase of misallocation per round III-110

IV.1 Summary statistics of experiment #1 . . . . IV-140

IV.2 Misallocation statistics of experiment #1 . . . . IV-141

IV.3 Random effects regression of experiment #1 . . . . IV-144

IV.4 Summary statistics of experiment #2 . . . . IV-152

IV.5 Misallocation statistics of experiment #2 . . . . IV-154

IV.6 Random effects regression of experiment #2 . . . . IV-155

IV.7 Comparison of experiments IV-160

IV.8 Comparison of experiments (participant average values) IV-161

V.1 Summary statistics of appointments V-180

V.2 Summary statistics of average firm variables from 2002 to 2018 . . . . V-186

V.3 t-tests for propensity score matching of non-whites V-189

V.4 t-tests for propensity score matching of non-Frenchs V-190

V.5 Firm level logistic regression V-192

V.6 Firm performance panel regression for non-white V-195

V.7 Firm performance panel regression for non-French V-196

In the beginning of this dissertation I would like to thank all the people who made this work possible, first and foremost my supervisor Prof. Dr. Christina Bannier. I am deeply grateful for her constant support and confidence in my work. She has managed to create a free and pleasant working atmosphere that left room for my own ideas and creativity, as well as always being there for questions, having an open ear for problems and giving very helpful advice. Such a supportive working relationship is by no means a matter of course and I deeply appreciate that.

I also would like to thank my second supervisor Prof. Dr. Andreas Walter. In addition to his supervision and advice, his lecture on behavioral finance sparked my interest in decision-making experiments and inspired my colleague and me to investigate deviations from rational choice.

Furthermore, besides Prof. Bannier and Prof. Walter, I would like to thank the other four professors leading our research network Behavioral and Social Finance and Accounting, namely Prof. Dr. Corinna Ewelt-Knauer, Prof. Dr. Peter Tillmann, Prof. Dr. Peter Winker and Prof. Dr. Arnt Wöhrmann. They all not only contributed to my dissertation with helpful advice, but also made this doctorate possible for me - as a non-specialist career changer - by creating this research network and welcomed me to Faculty 02 in a pleasant atmosphere.

Next, I would like to thank my colleague and main coauthor Florian Gärtner, who shared an office with me for the past five years and was always there for interesting and fruitful discussions. He complemented my mathematical with his sociological experience and knowledge in the field of behavioral finance. Also I would like to thank my other coauthors Yannik Bofinger, Benjamin Fiorelli (both also in the BSFA research network) and Jan-Niklas Reinschmidt from BWL VI for the great collaboration and giving me the opportunity to explore the field of finance beyond experiments. Furthermore, I thank my other colleagues Alix Auzepy, Karsten Bocks, Björn Rock and Dr. Thomas Heyden from BWL VI for their constant support and insightful discussions in several Brown Bags. My gratitude also goes to several other colleagues for the many fascinating and fruitful conversations: Petrit Ademi, Dr. Kim Heyden, Dr. Niklas Kreilkamp, Dr. Sascha Matanovic, Philipp Schade, Dr. Maximilian Schmidt, Christina Stahlecker and Ferogh Zaman.

I would also like to express my gratitude to our student assistants Anja Heller, Julia Schwarz, Jan Speier and Luisa Trebing, who have relieved me of a lot of work with teaching and research assignments. And I also acknowledge Julia Körner and Annette Toalster, the former and current secretaries in the Chair of Banking & Finance for their extraordinary administration of everything organizational in the university, often far beyond their actual job.

Lastly I thank my beloved family, both my parents Günther and Dietlinde Semmler as well as my brother Dr. Diego Semmler, for their outstanding support.

General Introduction

This doctoral thesis consists of five papers, of which four deal with experiments in household finance to investigate deviations from rational choice, one deals with diversity in corporate boards. Two of the household finance puzzles are in a conjoint version published at the *Journal of Economic Behavior & Organization* (Gärtner et al. (2023), https://doi.org/10.1016/j.jebo.2022.10.032).

The first four papers in this thesis, coauthored by Florian Gärtner, Christina Bannier and Yannik Bofinger, evaluate the question, how consumers deal with multiple accounts, mainly credits. The research questions we investigate in these papers all sparked from a specific household finance puzzle discovered by Ponce et al. (2017) and Gathergood et al. (2020) with data from the field: Consumers allocate around half of their money they use to repay credit card debts *not* on the credit card with the highest interest rate, even after accounting for contractual constraints as minimum payments. The reason for such deviations remains ambiguous. While Ponce et al. (2017) focus on the concept of mental accounting (Thaler, 1985), Gathergood et al. (2020) see an explanation in heuristics (Tversky and Kahneman, 1974). The main goal of our studies is to investigate such heuristics in this deviation from optimal behavior - which we call misallocation - in an experimental setting and shed light on the questions how people can be influenced to reduce misallocation. Furthermore, we generalize our findings on borrowing and investment situations.

Chapter I presents the paper "What could possibly go wrong? Triggering misallocation", a conjoint work with Florian Gärtner, and Christina E. Bannier. As a first step to find reasons and predict patterns for misallocation we establish situations in which misallocation theoretically should occur stronger. Following Amar et al. (2011), we design an experiment in which participants have to repay two credit cards with a given amount of money on a checking account. The experiment consists of several independent situations - called "scenarios" - which only vary in the values of the credit card and checking account balances, as well as in the credit card interest rates. We set these values in the scenarios specifically to steer misallocation by triggering heuristics from previous literature as well as novel heuristics that lead to at least partly repayments on the credit card with the lower interest rate. For four of seven heuristics we find the predicted patterns. This implies that misallocation cannot be reduced to mere random noise, but strongly depends on the credit situations consumers are situated in.

In the study "What could possibly go wrong? Nudging and the Cuckoo Fallacy" in chapter II - which is a conjoint work with Florian Gärtner and Christina E. Bannier - we switch perspectives from different situations to framing effects (Tversky and Kahneman, 1981) which we use to increase or decrease the usage of one particular fallacy. We focus specifically on a novel fallacy - we call it the "Cuckoo Fallacy" - that leads people to allocate money to the credit that produces more new debts in the next interest phase. In case of strongly uneven credit balances, more precisely if the low interest credit card has a sufficiently high balance, this heuristic leads to misallocation. In three treatments the participants deal with two credits with a fixed income. The treatments only differ in the presentation of balances and interest rates to either favor misallocation utilizing the Cuckoo Fallacy or nudge people to optimality. While we could not increase misallocation, our framing was effective in decreasing misallocation when we show people how much money they can save with their repayment. This shows that framing can help to resolve this financial puzzle.

After having established situations that favor misallocation and that misallocation can be steered, in the paper "Addressing consumer misunderstanding in credit card debt repayment: Policy suggestions beyond the CARD Act" of chapter III, which is written together with Florian Gärtner and Yannik Bofinger, we now focus on ways to resolve sub-optimal repayment behavior. We run another experiment in the same style, but design several financial interventions to educate the subjects before and during the experiment rounds. All interventions reduce misallocation to a notable amount. Adapted interventions that are situation-specific reduce misallocation more than general advice. The best results are achieved by an interactive assistant that could be used as inspiration for FinTechs to design a cell phone application for consumers. We also discuss regulatory implications and analyze how our interventions improve the way people handle credits beyond legal interference at the example of the U.S. Credit Card Accountability Responsibility and Disclosure (CARD) Act of 2009.

Even after analyzing and fixing situations leading to misallocation the fundamental cognitive principle why misallocation occurs at all still remains unclear. In the paper "Elemental Financial Decisions" in chapter IV Florian Gärtner and I provide one explanation following the idea of "cognitive uncertainty" by Enke and Graeber (2021a). According to this principle even in situations with an objectively optimal solution, people encounter uncertainty whether their solution is correct or not. They anticipate potentially being wrong by leaning to a mediocre choice, for example in our case an even split between credit cards. We run two experiments where people have to spend money in two interest bearing accounts that differ in whether people may or may not split their money between the accounts. We manipulate cognitive uncertainty with three independent variables. The participants encounter all combinations of situations in which they deal with credits or assets, negative or positive interest rates and percentages or absolute values. We find that cognitive uncertainty indeed increases misallocation under divisible money. In borrowing situations misallocation as well as uncertainty strongly increases. The same holds true in case of negative interest rates, but the effect is stronger when combined with borrowing situations. Percentages in comparison to absolute values do not show robust effects. Overall, cognitive uncertainty predicts non-optimal behavior partly, but cannot explain all the differences in our treatments. Therefore, the interplay of different mechanisms leading to non-optimal financial allocation might be more complex and cognitive uncertainty is only one piece of the puzzle.

The last paper "Ethnic Diversity and the Glass Cliff - An examination of French CAC40 Boards" in chapter V, coauthored by Benjamin Fiorelli and Jan Niklas Reinschmidt, is only remotely related to behavioral finance, but supplements the other papers with an approach based on field data. We focus on the development of ethnic diversity in boardrooms of French firms in the CAC40 index. As the composition of a board is the result of an attitude to the capabilities of its members, it is also related to behavioral decision making, especially since we investigate behavior induced by stereotypes. More specifically we investigate the "glass cliff" (Ryan and Haslam, 2005), a phenomenon according to which the probability of female appointments to a firm's board increases when stock prices decrease, and which for example has been found for Fortune 500 companies (Cook and Glass, 2014). Transferred to ethnic minorities which we define by appearance on one hand and by nationality on the other hand, we cannot find any effects of minority appointments on firm performance, neither before nor after an appointment. This study can be seen as indication that a glass cliff for ethnic minorities might not exist in France.

Chapter I

What could possibly go wrong?

Triggering misallocation

Coauthors:

Florian Gärtner

Christina E. Bannier

Relative share:

45%

This chapter has been published in a conjoint version together with chapter II:

Gärtner, Florian, Darwin Semmler, Christina E. Bannier (2023), "What could possibly go wrong? Predictable misallocation in simple debt repayment experiments", *Journal of Economic Behavior & Organization*, 205, 28-43, ISSN 0167-2681, https://doi.org/10.1016/j.jebo.2022.10.032.

A previous version of this chapter has 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 so 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.11). 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;

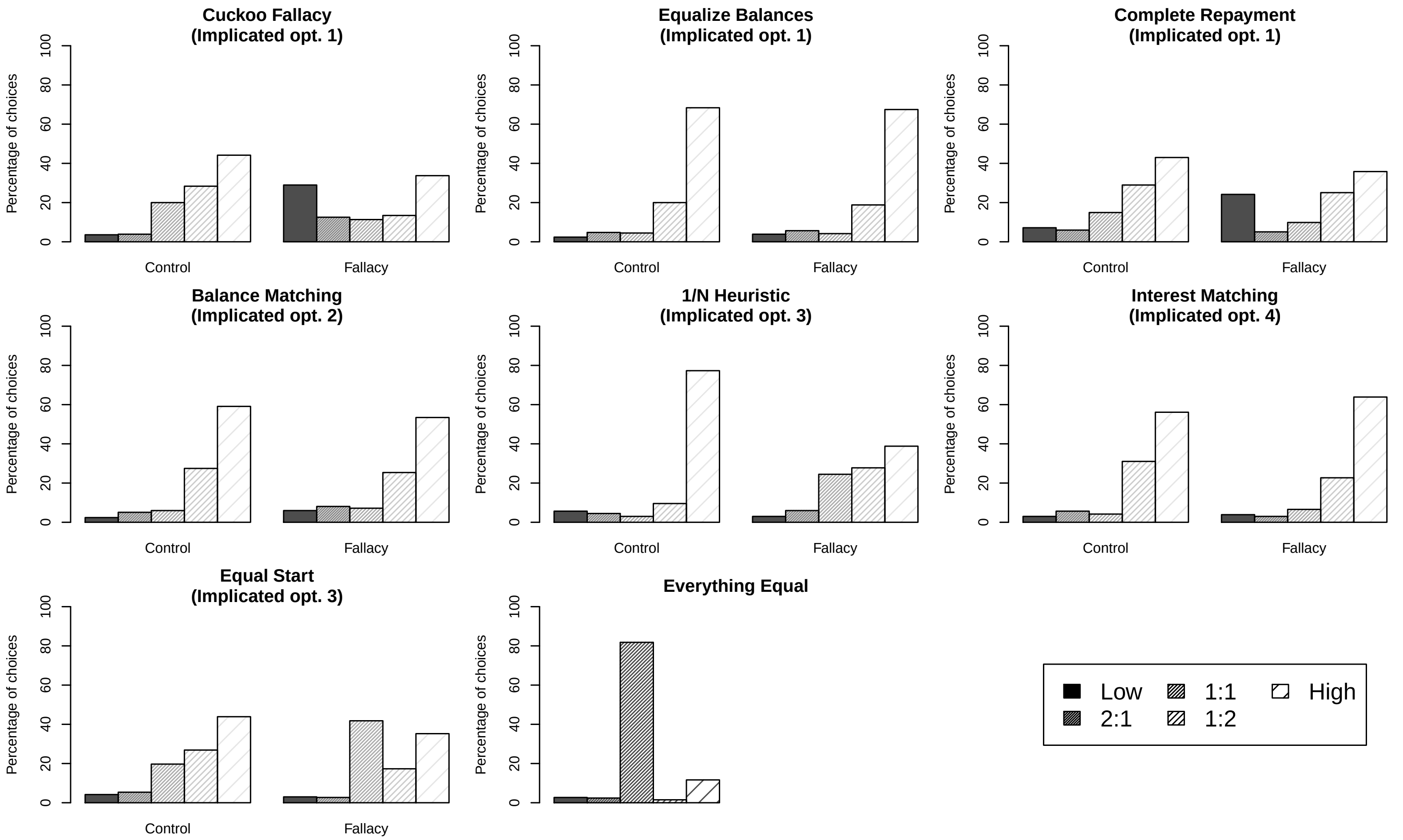


Figure I.1: Comparison of choices between control and fallacy scenarios. The columns show the percentage of choices for every repayment option. They are sorted by option no. 1 to 5. The leftmost column represents the number of participants choosing to repay all the money to the low interest credit card 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.9 and Appendix I.10). 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.9 and Appendix I.10).

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.12 and Appendix I.13). 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.14). 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.15). 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.16 and Appendix I.17. 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.

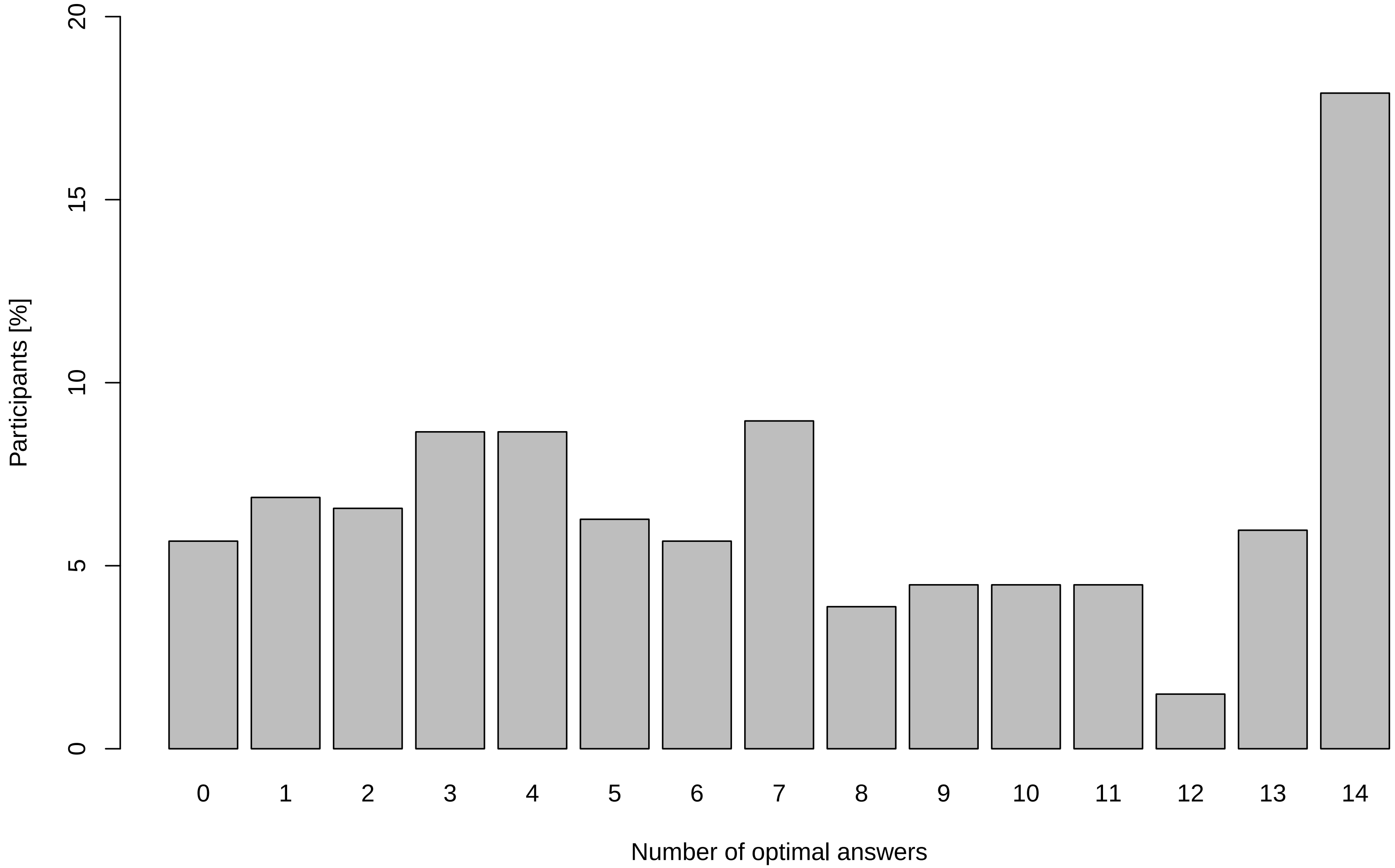


Figure I.2: The bars show which proportion of participants gave a certain amount of optimal answers (option 5) in the 14 scenarios (excluding the Everything Equal scenario).

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

Fallacy	**Cuckoo Fallacy**				
Control	**1**	**2**	**3**	**4**	**5**
1	1.19%	0.00%	0.60%	0.00%	1.79%
2	0.90%	0.30%	1.19%	0.30%	1.19%
3	5.07%	3.28%	4.48%	4.18%	2.99%
4	8.36%	6.87%	3.88%	6.57%	2.69%
5	13.43%	2.09%	1.19%	2.39%	25.07%

Fallacy	**Equalize Balances**				
Control	**1**	**2**	**3**	**4**	**5**
1	1.19%	0.30%	0.00%	0.60%	0.30%
2	0.90%	1.19%	0.90%	0.30%	1.49%
3	0.00%	0.90%	1.19%	1.49%	0.90%
4	0.90%	1.49%	1.49%	8.06%	56.42%
5	0.90%	1.49%	1.49%	8.06%	56.42%

Fallacy	**Complete Repayment**				
Control	**1**	**2**	**3**	**4**	**5**
1	5.07%	0.60%	0.60%	0.30%	0.60%
2	2.09%	1.49%	0.30%	1.49%	0.60%
3	3.28%	1.19%	3.88%	5.37%	1.19%
4	4.18%	0.90%	4.48%	14.33%	5.07%
5	9.55%	0.90%	0.60%	3.58%	28.36%

Fallacy	**Balance Matching**				
Control	**1**	**2**	**3**	**4**	**5**
1	1.19%	0.30%	0.30%	0.30%	0.30%
2	0.30%	2.39%	0.30%	1.19%	0.90%
3	0.90%	0.60%	1.79%	2.09%	0.60%
4	1.79%	2.99%	2.99%	11.94%	7.76%
5	1.79%	1.79%	1.79%	9.85%	43.88%

Fallacy	**1/N Heuristic**				
Control	**1**	**2**	**3**	**4**	**5**
1	1.19%	1.49%	1.79%	0.60%	0.60%
2	0.00%	1.19%	1.19%	1.49%	0.60%
3	0.00%	0.90%	1.19%	0.60%	0.30%
4	0.30%	0.60%	3.28%	4.18%	1.19%
5	1.49%	1.79%	17.01%	20.90%	36.12%

Fallacy	**Interest Matching**				
Control	**1**	**2**	**3**	**4**	**5**
1	1.19%	0.60%	0.30%	0.90%	0.00%
2	0.60%	0.60%	1.79%	1.19%	1.49%
3	0.60%	0.60%	1.49%	0.30%	1.19%
4	0.30%	0.90%	1.49%	14.93%	13.43%
5	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 $< 2.2 \cdot 10^{-16}$								

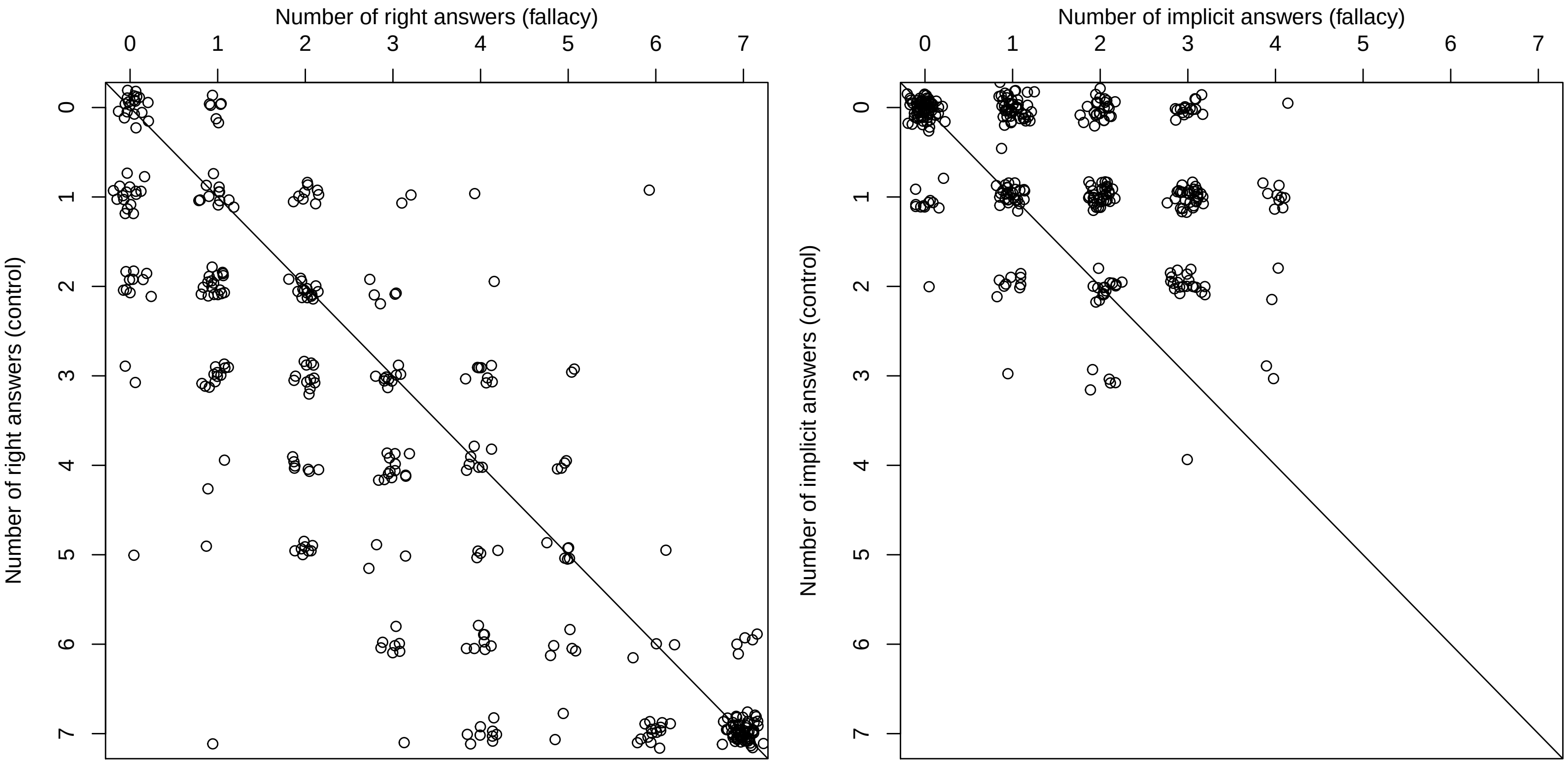


Figure I.3: This graph visualizes how many participants give certain answers in control (y-axis) vs fallacy scenarios (x-axis). The left graphic shows the number of optimal answers (option 5), the right graphic shows the number of fallacy-implicated answers.

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.

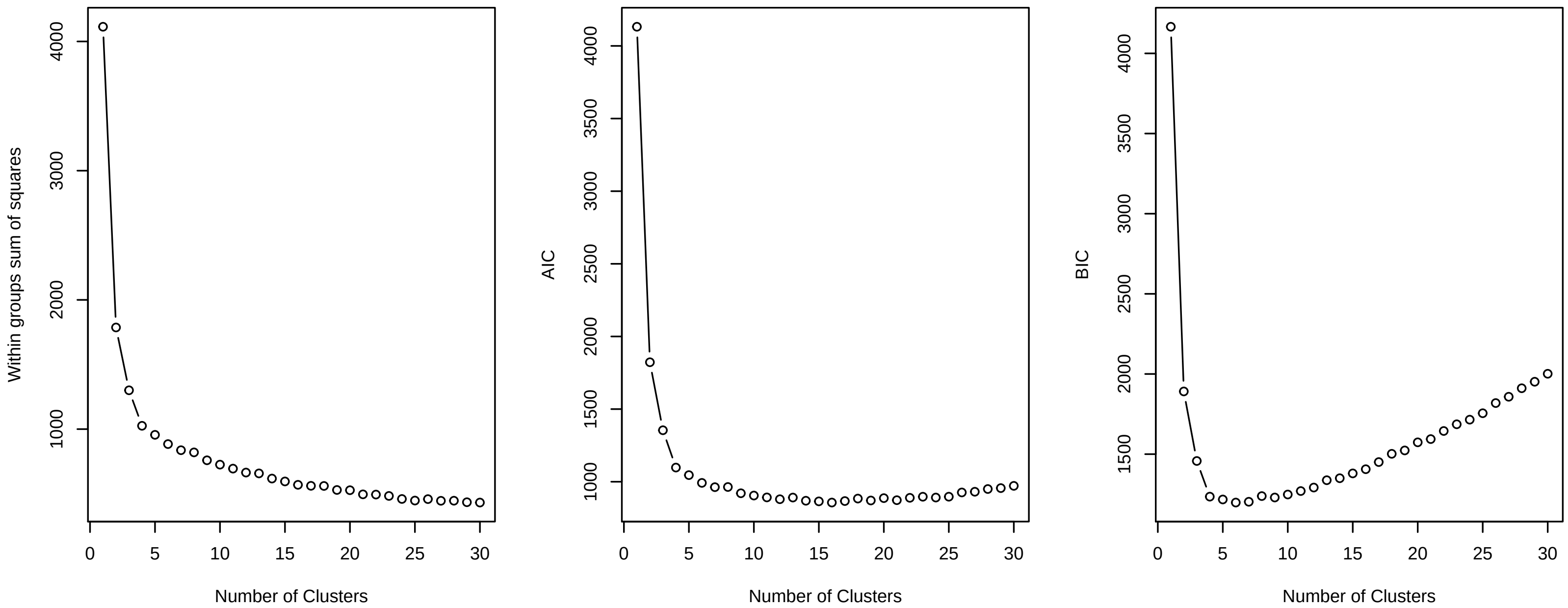


Figure I.4: This Figure shows three criteria to determine the number of clusters for a k-means clustering. The x-axes show the number of clusters between 0 and 30. The y-axes show the value of within groups sum of squares (elbow criterion, left figure), the AIC (Akaike's information criterion, middle 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 (2021b). The authors use a distinct model of a mental gap, which they call "cognitive uncertainty" (see also Enke and Graeber (2021a)). 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:

Florian Gärtner

Christina E. Bannier

Relative share:

45%

This chapter has been published in a conjoint version together with chapter I:

Gärtner, Florian, Darwin Semmler, Christina E. Bannier (2023), "What could possibly go wrong? Predictable misallocation in simple debt repayment experiments", *Journal of Economic Behavior & Organization*, 205, 28-43, ISSN 0167-2681, https://doi.org/10.1016/j.jebo.2022.10.032.

A previous version of this chapter has 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 Risiko-management 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% (ShowNew-Debts 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-

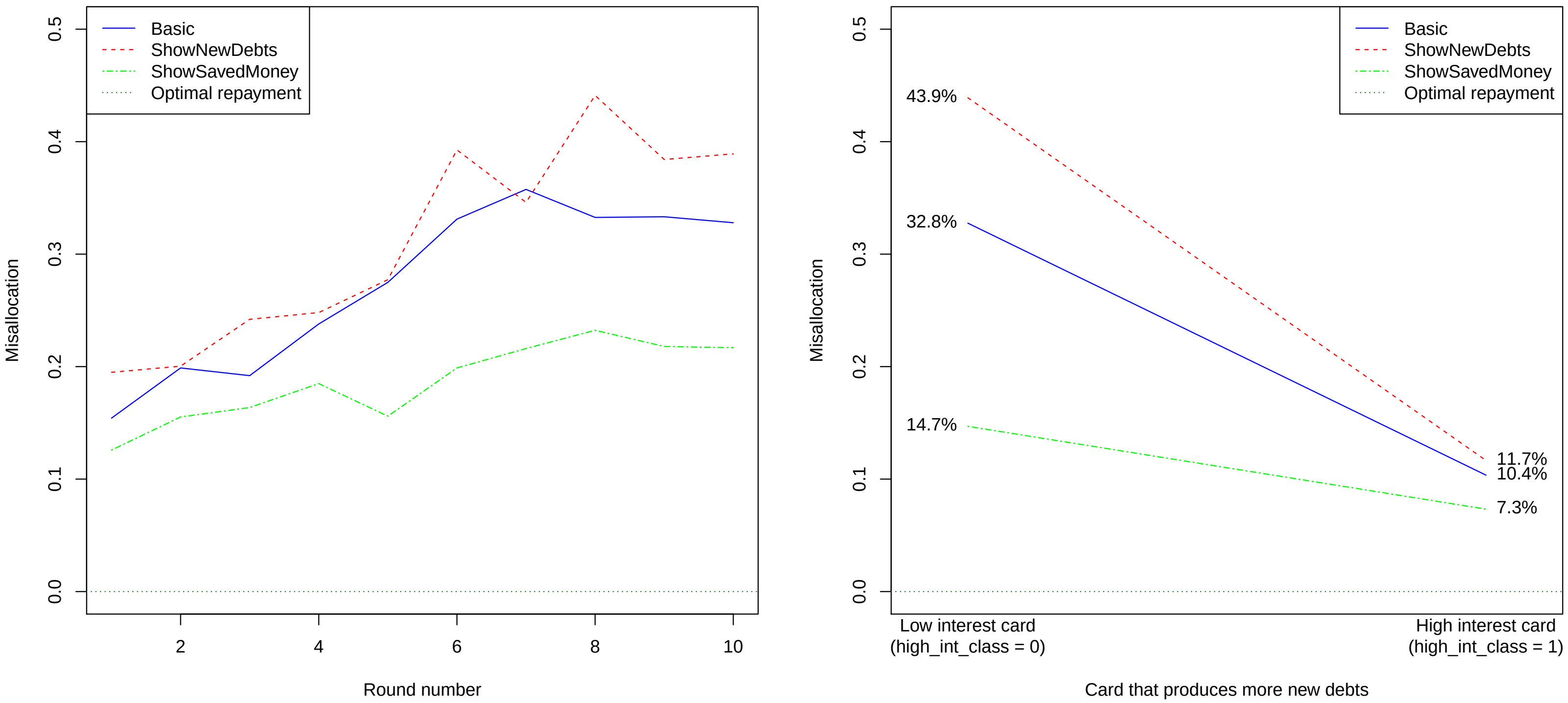


Figure II.1: The left Figure shows the development of the misallocation per treatment in the experiment rounds. The right Figure shows the interaction plot between treatment and interest class (i.e., which credit card accumulates more interest). ShowNewDebts (in red) is the sludge and ShowSavedMoney (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.19 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.19) 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.20). 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.21). 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.22. 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

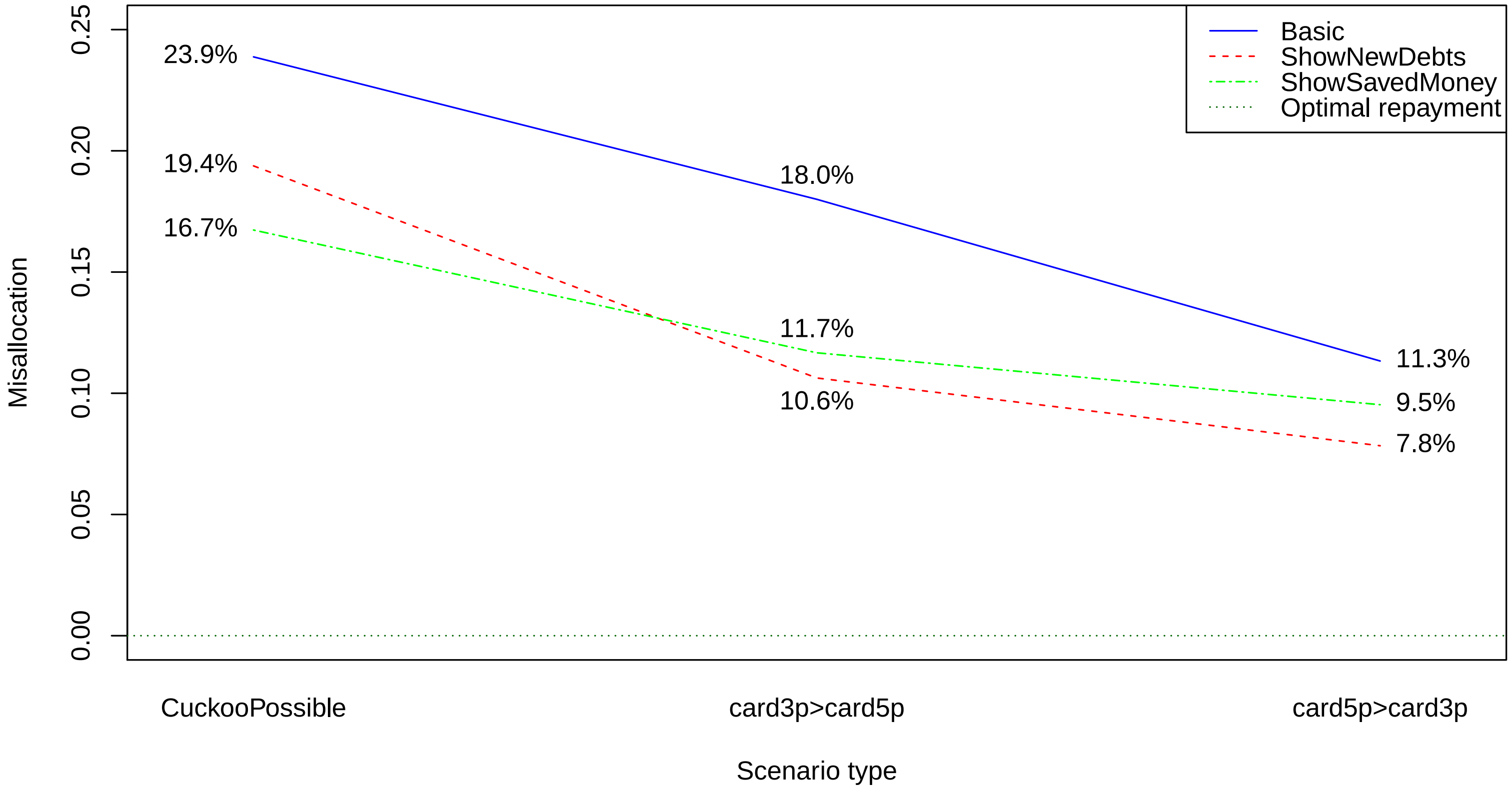


Figure II.2: This Figure shows the interaction plot between treatment and scenario type. That is we differ between scenarios whether the Cuckoo Fallacy was possible (CuckooPossible) or whether the 3% credit card produced more new debts than the 5% card, but not that much that the Cuckoo Fallacy was possible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

Florian Gärtner

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 non-linear 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.24 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.25.

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.7 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.26.

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, the interventions

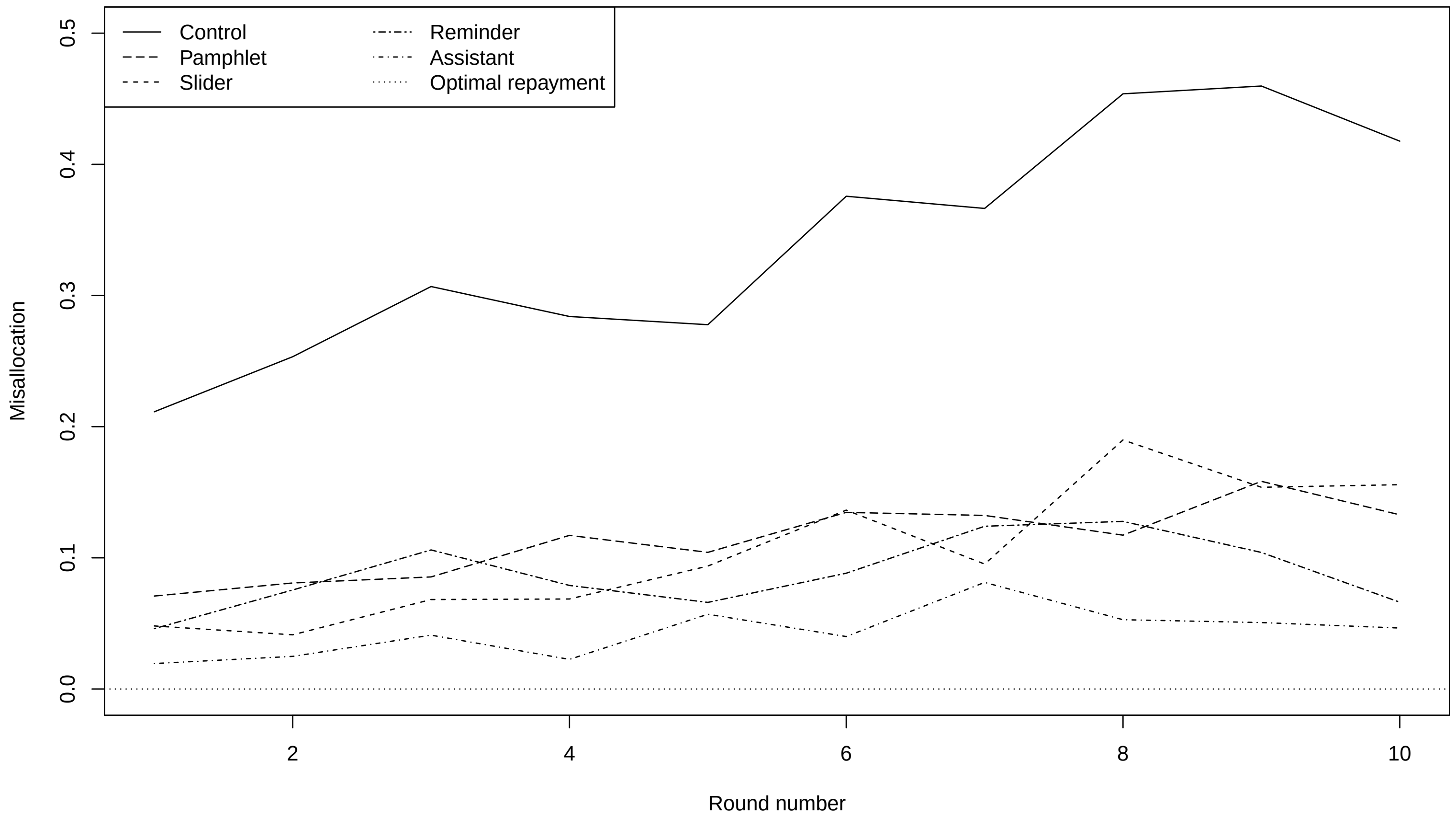


Figure III.1: This figure shows the development of the average misallocation over ten experiment rounds split by treatment. The round numbers are on the x-axis, the mean misallocation on the y-axis. The different values for each treatment are shown as different line types. The classification of the line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

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.27 illustrates the results.

Second, we use fractional regressions as illustrated in Table Appendix III.28 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.28 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.29. 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 = 0.008), which is surprising since the pamphlet intervention is weaker in terms of re-

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

ducing 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 it in the interventions.

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.30 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 implementation restricts and defines caps on various fees or increases in fees and requires

[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.

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 and have additional access to another 0.75 credit cards.

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 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

Elemental Financial Decisions

Coauthor:

Florian Gärtner

Relative share:

50%

A previous version of this chapter has been presented at:

- Society for Experimental Finance Conference 2022

Elemental Financial Decisions

Abstract

We investigate elemental 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 are partly driven by cognitive uncertainty, which is a particular form of confusion.

Keywords: Household finance, investing, experimental finance, elemental 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 stays 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 elemental, and 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 *elemental 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 elemental financial decisions. If people do not consistently behave optimally in the elemental 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 elemental 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, so 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, 2021a). We assume that our participants have monotonous preferences for money, and we use Enke and Graeber (2021a)'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 (2021a) is that this uncertainty leads to a "shrunken action", i.e. an action which 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 elemental 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 elemental 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 strong 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. 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 sub-optimal 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 which 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. We show that one of these supplemental theories can be Enke and Graeber (2021a)'s model of cognitive uncertainty, as this models 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 might not be enough, because cognitive uncertainty is not strong enough to explain all the differences between our treatments.

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 "elemental 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 elemental 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 for 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 elemental 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 elemental financial decision as we understand it exists, if any at all. This is why we use an experimental approach.

We conceptualize elemental financial decisions as situations of cognitive uncertainty (Enke and Graeber, 2021a). Enke and Graeber (2021a) 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 (2021a)'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 (2021a) 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 (2021a) 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 (2021a) 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

(2021a)'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 the 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 the 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 decisions 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 trial to remind the participants to pay attention which variant of the decisions they are dealing with. We exclude 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 Java script, we exclude participants who disabled Java script 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. (Forthcoming) 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 a 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 that 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 (2021b), 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 as 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 which 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 an 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 check and 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 about misallocation in general. In Table IV.2 we report the average misallocation of participants in different treatments. For a graphical representation of misallocation and uncertainty in general as well as in different treatments see Figures Appendix IV.8 and Appendix IV.9. 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

[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	

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.

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 investigate the influence of our treatments on cognitive

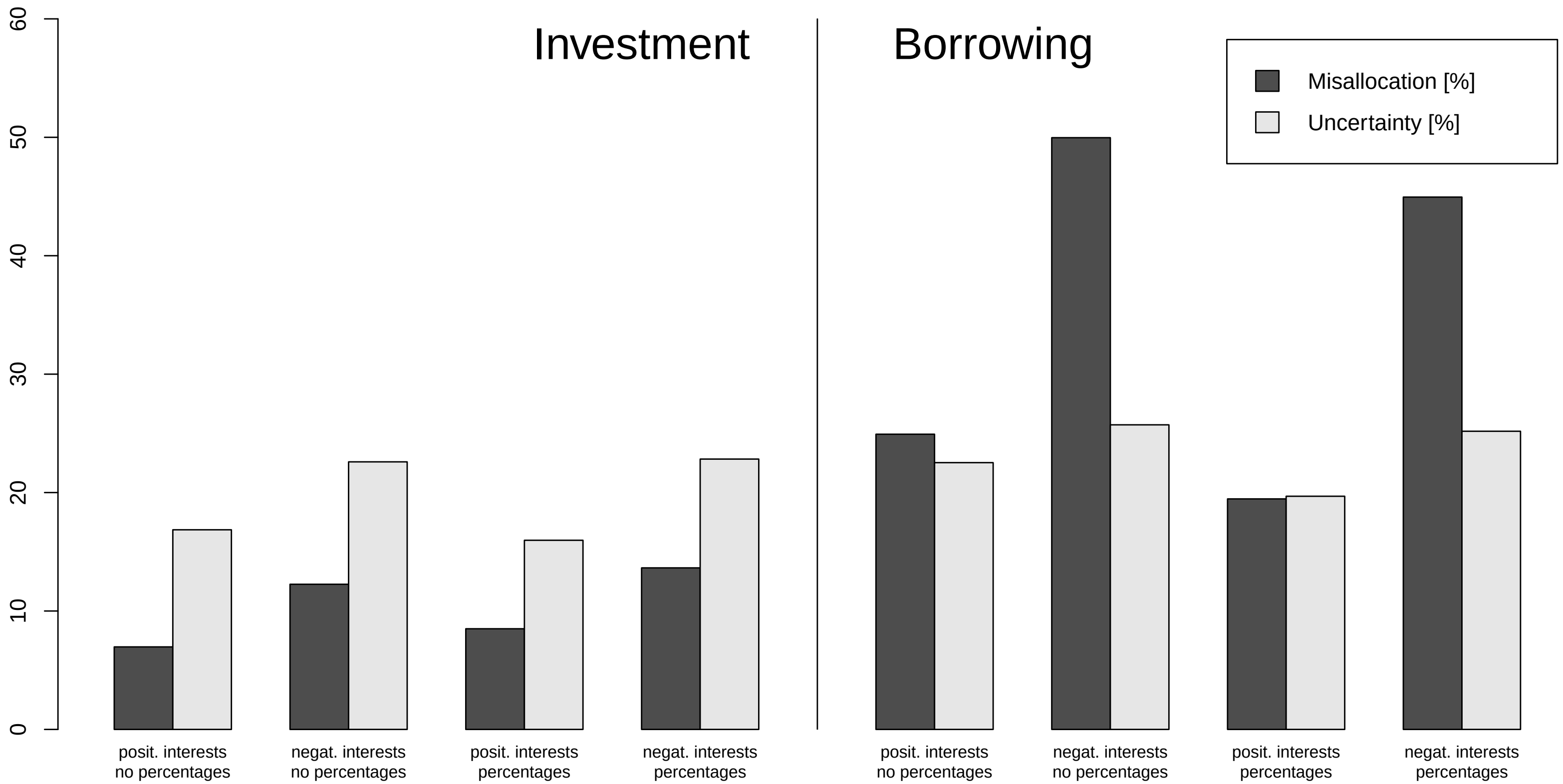


Figure IV.1: The Figure shows barplots of average percentage points in misallocation and uncertainty split by the 8 treatments. The barplots on the left side correspond to investment in assets, the one on the right correspond to borrowing.

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. 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.31.

The columns (1), (3) and (5) describe the models with uncertainty as the dependent variable, measured on a scale between 0 and 100. The 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 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. The 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* 2*nd*, 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.[2] 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

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

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.

money values. The 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 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 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.5, 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.9, 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 avoiding misallocation, which is the opposite of H4b.

We finally turn to the question whether cognitive uncertainty is a mediator on the path to misallocation. We follow the approach of Baron and Kenny (1986) 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 percentage frame. The latter result is unsurprising as we also do not measure any effect of 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 decision, but it is far from explaining non-optimal decisions completely.

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.32 shows the results. We do not display the regressions with uncertainty as 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 works in no 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.33 shows our results. The interpretations for the hypotheses still hold. For concerns regarding participants who took too long or were to quick to complete the experiment, we created a subset of our data which 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.34 remain the same as in the main analysis.

We also apply several hypotheses 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 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 "Elemental 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[3]. In this sense, experiment #1 is not the sim-

[3]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,

plest 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 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

which gives you 10,001 alternatives to choose from.

the amount 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[4]. 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 experiment #1.

Note that we will investigate the 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.

[4]For example, if you think that you should invest 85% of you 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.

IV.4.2 Results

We conducted the experiment in February 2022. We recorded 301 participants starting our study. We lose 28 participants who returned the study, and 2 participants due to time-out. We reject another 2 participants because they failed multiple attention checks. Out of the remaining 269 approved participants, we remove 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.4.

We again start by exploring the average misallocation in different treatments. The results are shown in Table IV.5 and Figure IV.2. We show further depictions of uncertainty and misallocation in general and in different treatments in Figures Appendix IV.10 and Appendix IV.11. 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 whether the results from experiment #1 hold, we replicate the random effects models. Table IV.6 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

Table IV.4: 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	

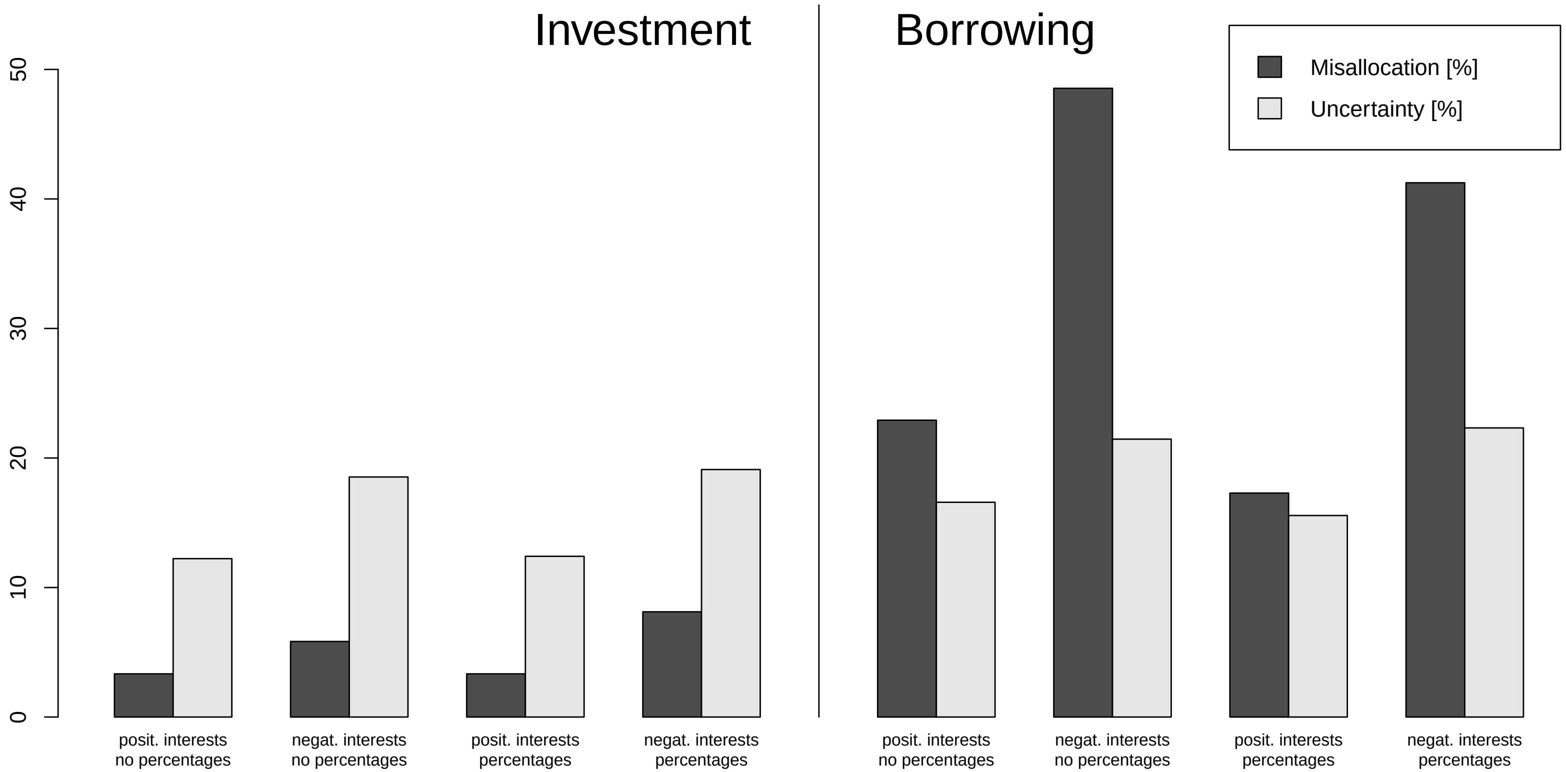


Figure IV.2: The Figure shows barplots of average percentage points in misallocation and uncertainty split by the 8 treatments. The barplots on the left side correspond to investment in assets, the one on the right correspond to borrowing.

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.5: 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.6: 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.

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.36. 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 screen out the 5% of participants with extreme experiment time, that is all participant 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.37. 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.7 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 × Uncertainty is insignificant, so we cannot confirm H5c. However, recall that in Table IV.6 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.7 would be best interpreted as a true null result, this would imply significant effects of uncertainty on misallocation in Table IV.6 - 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.7, 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 × 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.39 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.40, but do not detect considerable deviations from the main results.

In a 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.8. 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.7. 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 hypotheses 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 dependant variable and the duration of each experiment round as 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.7: 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.8: 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 elemental 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 predicable, 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 in actual behavior under divisible money. So in general, our predictions were reasonable.

However, there are some exceptions. While cognitive uncertainty plays a role, it can explain no more than around 12% points of misallocation (comparing the estimates for 0% and 100% uncertainty, respectively), many treatment effects stay significant even after controlling for it, and the mediation analyses also does not show complete mediation except in one case for negative interest rates, so we seem to miss important aspects in our analysis. 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 which 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 which 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 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 which 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. But 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[5]. We believe that this effect dominates the misallocation that stems from the possibility to express 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 to express uncertainty compared to a variable that does not, as long as the smoothing effect is not too strong.

Finally, it is notable that the results often do not hold for the complete model. We are not sure why, but given our within-subjects design, random differences between the treatments cannot explain these vanishing effects. However, we lose roughly a third of our observations, so we think the best interpretation for these insignificant coefficients are the power problems this loss causes.

We finish by drawing some additional conclusions for future research. First, we want to highlight the importance of cognitive uncertainty, or more general, confusion. Since we find that cognitive uncertainty matters for elemental financial decisions with different returns, it is reasonable to assume that other elemental dimensions, such as time or

[5]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 realize, because \$4 is a too small incentive for the seller to produce, while \$5 is too expensive for the buyer.

risk, might be influenced by it as well, and more complex decisions even more. Second, our results suggest that in times of negative interest rates, the average decision quality should decrease. And third, there might be other relatively elemental decisions, such as selling assets, where we might find misallocation, and it should also be fruitful to investigate who exactly misallocates, and why. Research which 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

Ethnic Diversity and the Glass Cliff - An examination of French CAC40 Boards

Coauthors:

Benjamin Fiorelli

Jan Niklas Reinschmidt

Relative share:

25%

Ethnic Diversity and the Glass Cliff - An examination of French CAC40 Boards

Abstract

We investigate a potential glass cliff for ethnic minorities in boards of French CAC40 companies. Using a complete data set of appointments between 2002 and 2018, we identify ethnic minorities via appearance on one hand and citizenship on the other hand. For both definitions of ethnic minority affiliation, we do not find a connection with firm performance, neither before, nor after an appointment. These null results are robust for different accounting based measures of firm performance, constraints in the data set and time period, an alternative matching algorithm to balance the data set, and further minor changes of the model. This can be seen as indicator that a glass cliff for minorities in French firms might not exist.

Keywords: Board structure, Diversity, Minorities, Glass cliff

JEL-Codes: G34 - J15 - J21 - J71 - M14 - M51

V.1 Introduction

Firm boards unite various capabilities and perspectives of their individual members. A high diversity in boards broadens the cultural and mental scope of corporate management and helps firms to deal with the fast-changing requirements of a growing market. In recent times, the attention towards the topic of diversity in boardrooms of large companies is growing. The focus in research so far largely aimed on gender diversity. Other types of diversity, such as cognitive or demographic diversity hitherto only got minor attention (Khatib et al., Forthcoming). Already in the early 2000s Higgs et al. (2003) recommended to find more directors of diverse backgrounds. Considering the special case of ethnic minorities, this topic increasingly gains importance, as ethnic diverse boards fulfill the new challenges arising in times of growing globalization. Singh (2007) finds that boards with higher ethnic diversity have higher market capitalization, a larger workforce, more independent and also more gender-diverse boards. Minority appointees exhibit higher levels of human capital and advanced education. There is a rise of their representation in boards alongside women over the last decade. In Fortune 500 companies there were 856 female board members in 2010, which is equivalent to 15.7% and 700 board members of ethnic minority (12.8%). In contrast to that, in 2020 the seats of women rose to 1,559 (26.5%) and the seats of minorities to 1,027 (17.5%) (Deloitte, 2021). While the increase of minorities is not as strong as the increase of women, there is a clear tendency towards more diversity in boardrooms that goes beyond equality of gender.

In this study we focus on the factors surrounding the appointment of minority board members and whether they face similar obstacles as women. The inequalities between gender and minorities can be linked together, as minorities also have a lower probability to get promoted to managerial positions and earn less wage (McGuire and Reskin, 1993; Maume, 1999, 2012). Regardless of their performance, decision makers view women and minorities as less capable of leading compared to white men (Rosette et al., 2008; Carton and Rosette, 2011) and both women and minorities face lower odds to achieve higher positions at work (Elliott and Smith, 2004). Therefore, prejudices might not only reduce the probability of appointment for women, but also for ethnic minorities.

We focus specifically on a phenomenon previously only attributed to female board members, namely the "glass cliff" and investigate its applicability on ethnic minorities, where research is sparse. The "glass cliff" emerged from gender research over the last two decades and is underpinned with data from the US or UK (e.g. Adams et al. (2009); Brady et al. (2011); Cook and Glass (2014); Ryan and Haslam (2005); Mulcahy and Linehan (2013)). The term of the glass cliff was initially defined by Ryan and Haslam (2005) who found that British firms in the FTSE 100 were more likely to appoint women as board members when stock prices abate. The chances for women to get promoted to a leading position in a firm increase in times of a crisis, when the position holds a higher risk to fail (Hill, 2016). One reason for this to happen might be that firms stick with the conservative notion of "white male leader", but are in need to try different approaches in times of poor performance to signal change (for example to investors) (Kulich et al., 2015).

While the glass cliff originally focuses on gender diversity, we investigate in this study whether we can transfer the glass cliff to ethnic minorities, another underrepresented group in corporate boards. We consider this question as particularly interesting as the argument of the signaling effect not only applies to women, but also to general minorities in firm boards. We restrict our research on a minority glass cliff to boardrooms in France, a subgroup that is particularly suitable for such an investigation, because the French population exhibits a higher mix of different ethnicities than other countries (Dignan, 1981) and an immigration history existing for several generations (Algan et al., 2010), whilst France is the third largest economic power in Europe (IMF, 2021). Furthermore, despite as in the US, firms in France can have an unitary or alternatively a dual board system, where there may be differences in the effects of the board composition; e.g. Jungmann (2006) reported differences in one-tier and two-tier board systems in the effectiveness of corporate governance. Thus, an investigation of a potential glass cliff over different board systems might reveal additional information about determinants of this phenomenon.

Our research goal is to find patterns in the firm performance explaining fluctuations in the proportion of minorities in boards. In particular, we search for evidence for the glass cliff hypothesis concerning ethnic minorities in the CAC40 companies. We tackle this question with a complete data set of appointments for the long time span of 17 years in French boardrooms. We summarize our goal in the following research question:

RQ: Does the phenomenon of the glass cliff also apply to ethnic minorities?

We investigate the presence of the glass cliff by tracking the development of firm performance 24 months before an appointment (and subsequently 24 months after an appointment) and light up differences between appointments of ethnic minorities compared to the other appointments. We measure firm performance by four different measures - cumulative raw return per month, cumulative market-adjusted return per month and cumulative risk-adjusted return per month (for these three see Adams et al. (2009)) and additionally on the systematic risk beta. In robustness checks, we also apply Tobin's Q as well as the accounting based measures return on assets and return on equity to ensure that our results are not driven by the choice of firm performance measures. We identify the affiliation to a minority by two different measures: the appearance as a subjective measure representing the point of view of decision makers, and the nationality as objective measure.

Following this approach, we use propensity score matching to generate comparable sets of ethnic minority and non-minority appointments and compare both groups with panel regressions on appointment level. We cannot establish a negative connection between the prior firm performance and the type of appointment (minority vs. non-minority) and therefore have no indication for the existence of a glass cliff for ethnic minorities in French CAC40 companies. Concurrently we apply the same models to firm performances after an appointment and thus rule out that there are actual objective differences in quality of the appointees. Using alternative firm performance measures, definition of variables, changes in control variables and also changing the matching algorithm cannot alter these null-results.

We contribute to the literature by taking a logical next step in the long strand of glass cliff research. We extend the mechanics of a signaling effect from a mere gender perspective to ethnic minorities on the example of a country with a particularly interesting immigration history. Our findings show that the underlying mechanics do not necessarily apply to minorities in boards beyond the gender-topic.

The remainder of the study is structured as follows. In section V.2 we present the literal background of this study and formulate hypotheses from our research question, section V.3 explains our data set and sets forth the methodology. We show our results in section V.4 and test their robustness in section V.5. Section V.6 concludes.

V.2 Background and hypotheses development

Ethnic minorities are an underrepresented group in firm boards. There are 17.5% non-white people in board seats of the Fortune 500 companies in 2020 (Deloitte, 2021). In the UK, the Parker Report of 2017 revealed that 8% of board members in FTSE 100 companies were non-white, while they represent 14% of the UK population (Parker, 2017). In our own data set, we detect 6.9% non-white and non-Hispanic board members and 31.6% board members without French citizenship at the end of 2018 in French CAC40 boards.

Our study aims to investigate the presence of ethnic minorities in boards by extending research that has been established to explain missing gender diversity in firm boards. The first idea, the so-called "glass ceiling" originated in the middle 1980s and describes the invisible barriers women cannot surpass in the career ladder to upper management or leadership positions (Boyd, 2008). However, Judge (2003) argued with regard to FTSE 100 companies in 2003 that a large number of women in high firm positions lead to worse firm performance and lower share prices. As the finding is mere correlational, this sparked the discussions about the reasons for the decline in performance and led to the phenomenon of the "glass cliff". The glass cliff first was defined by Ryan and Haslam (2005) for FTSE100 companies in the UK and refers to women having a higher chance of being promoted to high corporate positions when a firm encounters a crisis and therefore encountering a higher risk of failure.

Kulich et al. (2015) named the signaling effect as a reason for the glass cliff. Poor performing firms are in need to make structural changes to signal that they deal with the problem and take action to put the firm performance on course again. A female appointment signalizes a movement to workplace equality, the consideration of demographic sub-groups in service and product development as well as progressiveness in the sense of a future-oriented restructuring of the corporate management (Krawiec and Broome, 2009).

Indeed Kaplan and Minton (1994) show that outsider's chances to rise to leader positions increase when stock performances are poor. The first empirical evidence for the glass cliff for women comes from Ryan and Haslam (2005) with data from the UK. They examined the performance of FTSE 100 companies before and after an appointment and find that in times of overall stock-market decline companies that appointed women were more likely to perform consistently bad five month prior to an appointment. Brady et al. (2011) study board compositions between 2001 and 2005 on Fortune 500 companies in the US. Using a large-sample analysis, they find that firms with previous scandals were more likely to appoint female CEOs. Haslam and Ryan (2008) and Bruckmüller and Branscombe (2010) both show the existence of a glass cliff for women in an experimental setting. Cook and Glass (2014) confirms and extends these findings to Black people, also for Fortune 500 companies. While they do not find significant differences in tenures, they establish the "savior effect", where minority CEOs are replaced by white men when company performance abates. Additionally in a direct comparison of all 52 female CEOs in Fortune 500 companies until 2014 to their male colleagues, Glass and Cook (2016) find a higher probability for women to reach the CEO position in times of higher risk. Elsaid and Ursel (2018) find for North-American firms between 1992 and 2014 that female CEOs are 40% less likely to face turnover after their appointment than their male colleagues.

However, the literature on the glass cliff is divided. Cook and Glass (2013) cannot identify a connection between female CEO appointments and firm performance measures in Fortune 500 companies between 1990 and 2011. Bechtoldt et al. (2019) do not find evidence for the glass cliff for British and German companies, Haslam et al. (2009) only find a glass cliff for market-based measures, not for accounting-based measures of firm performance. Brinkhuis and Scholtens (2018) do not find differences in market reaction of female CEO and CFO appointments compared to male appointments when investigating stock returns of global firms between 2004 and 2014.

This ambiguity in the findings, which is mainly noticeable when observing different countries, suggests that the glass cliff is not an universally valid phenomenon, but rather depends on how and when attributes of appointees add value to a firm, which might differ for different countries as well as for appointees of different cultural background (Adams, 2017). Thus, we consider it interesting to analyze an underrepresented group bringing various cultural foci to the firm board.

However, there seem to be differences in the career paths of white woman and minority men and women (Powell and Butterfield, 2002; Bell and Nkomo, 2001), and different minority groups might encounter different dimensions of disadvantages (Chung and Lankau, 2005). Thus, we cannot directly carry over results from gender equality research to minorities. However, besides different reasons, both women and minorities encounter higher obstacles to reach positions with a higher level of power (Elliott and Smith, 2004). McGuire and Reskin (1993) find with survey data from 1980 that women and Black people earned less wage than white men. Maume (1999) adds to these results by finding that women and Black people have decreased chances and longer waiting times to rise to management positions. Rosette et al. (2008) find in experimental settings that white people are seen as more effective and capable leaders. Cook and Glass (2014) find that women as well as Black people are less likely to rise to a CEO position in the Fortune 500 companies between 1996 and 2010 when accounting performance had declined before. So the mechanism hindering women from reaching higher firm positions might also apply to ethnic minorities. Furthermore, the line of reasoning surrounding the signaling effect carries forward to any underrepresented subgroup. Thus, we assume it to be likely to also find a glass cliff for ethnic minorities.

We base our research on French data, because France has a particular high mix of diverse ethnicities compared to other countries (Dignan, 1981). Moreover, France introduced a 40% women's quota in 2011 (Assemblée nationale, 2011). Consequently, female board members are no longer an exception, which annihilates a potential signaling effect for women. This could put ethnic minorities in the spotlight as the new group to signal change. While the attention on gender diversity in boards in France increased, research on the glass cliff in France before the introduction of a 40% women's quota in 2011 is sparse (e.g. Maclean and Harvey (2008); Moulin and Point (2012); Nekhili and Gatfaoui (2012)). Novel research find that gender diversity has a positive effect on firm performance in the CAC40 firms (Ahmadi and Bouri, 2017; Ahmadi et al., 2018) and also on CEO compensation (Benkraiem et al., 2017). Evans (2014) predicts an increase in the importance of women in the future. While they reject the glass cliff hypothesis for SBF120 companies between 2000 and 2009, Dang et al. (2014) find an increase in proportion of women, which is also confirmed by Singh et al. (2015) using demographic board data. There was further research on the impact of female presence in SBF120 companies on firm level (Dang and Nguyen, 2018), as well as investigating profiles of female board members in France (Singh, 2015).

Considering the research on a glass cliff especially for ethnic minorities, Aelenei et al. (2020) show for political left-wing parties that minority candidates have higher chances for hard-to-win seats. While the mechanics of this effect in a political context is similar to the glass cliff, the research gap on corporate level has yet to be filled. To the best of our knowledge, we are the first to study the glass cliff specifically for ethnic minorities in France.

To set up our hypotheses it is indispensable that we formulate a definition of when an appointee in a board is considered as affiliated to an ethnic minority. This largely depends on the perception of decision makers in firms and there could be multiple channels to transport the impression of a minority appointee. Therefore, we employ two different definitions: First decision makers for the board structure might be influenced by subjective impressions like appearance. With reference to Santee et al. (2199), a picture can be used to distinguish among those ethnic groups. Therefore, we classify board members as "white" when their appearance is Caucasian or Hispanic, or "non-white" otherwise. We also apply an alternative definition with Hispanics as "non-white" in the robustness checks. As the classification via a picture is highly subjective, we use the citizenship as second and objective measure: We define a board member affiliated to an ethnic minority when they do not possess the French citizenship. We also control for several personal features of like age, gender or different education since minority appointees might have different career paths than their non-minority peers.

To recap our reasoning so far, we argue that ethnic minorities encounter similar obstacles on the path to corporate leadership positions than women. Since the positive effect of signaling a change in corporate structure also applies to non-whites or foreign citizens as underrepresented groups in French boardrooms, this strongly hints that they also might get promoted to higher corporate positions with a higher probability in times of poor firm performances, and therefore face a glass cliff. This leads us to the following hypotheses:

H1: Non-white board members are more likely to be appointed when firm performance declines prior to an appointment.

H2: Non-French board members are more likely to be appointed when firm performance declines prior to an appointment.

In a next step, we check the actual firm performance in a two-year period after a minority appointment in order to investigate if concerns of a performance difference in minority or non-minority appointees are justified or not. Gender-related papers for the glass cliff already established that women do not perform worse compared to men (e.g. (Haslam et al., 2009)). But also ethnic minority appointees show no difference in firm performance, as Carter et al. (2010) find for the US based on return on assets and Tobin's Q. Therefore, we do not have any expectations for significant differences after appointments of minorities and refrain from formulating a hypothesis, but nevertheless think this is an interesting question to investigate exploratory.

Since we capture a large time span of 17 years, another interesting analysis is the development of ethnic minorities over the years in our data set. We are interested in whether we find a statistically significant change in the appointments of foreigners, such as for example Deloitte (2021) finds an increase in Fortune 500 companies. An effect might be observable on a yearly base or also as side effect of board restructurings after the introduction of the women's quota in 2011. We consider this topic as an interesting background analysis to better understand the situation in the boardrooms we observe. But since this question is independent from the existence of a glass cliff and an exhaustive investigation is well beyond our scope, we refrain from developing a hypothesis.

V.3 Data and method

V.3.1 Data on appointment level

We cover all firms that were listed in the French CAC40 index between 2002 and 2018 and create a sample of board appointments during this period using annual firm reports. For each firm we only cover the period of time between its first date of entry in the CAC40 since 2002 and its last date of exit up to 2018. This leads us to a sample of 58 firms, of which 24 firms were continuously in the CAC40 between 2002 and 2018. The firms Air France and Hermes technically were listed in the CAC40 between 2002 and 2018, but Air France only briefly and Hermes at the very end of our observed time span, such that we record no appointments. Thus, these firms do not appear in our data set. For our main analysis, we ignore breaks from the CAC40 listing in our observed time span to avoid gaps, but remove these firms in a robustness check.
Our final sample consists of 1,183 appointments. 325 (27.5%) board members were woman and 383 (32.4%) had a non-French citizenship. At the date of the appointments, the appointees were on average 54.2 years old. You can view the summary statistics of the appointments in Table V.1.

We employ two measures of minority membership: a dummy variable for the ethnic background rated by images in the annual reports (non-white, equals 0 if the appearance is Caucasian or Hispanic, 1 otherwise; subjectively judged by a rater based on photographs), and a dummy indicating whether the nationality is French (non-French, equals 0 if the nationality is French, 1 otherwise). With this definition we count 77 appointees who were non-white and 383 appointees without French citizenship. For simplicity we summarize both variables under the term "foreign" and always mean analyses in both variables when we use this term. We show a complete list of firms with number

Table V.1: Summary statistics of appointments

Appearance		Nationality		CEOs		Age	
Caucas.	1,006	FR	771	CEO appointments	27	N	1,073
Hispanic	59	US	63	Later CEOs	28	Mean	54.234
Asian	31	DE	50			Std. Dev.	8.239
Near-East	31	UK	41	Boardtype		Min.	24.418
African	3	BE	40	Unitary board	889	Median	54.764
Mixed	12	ES	25	Supervisory board	200	Max.	79.385
		Other	164	Management board	81		
NA	41	NA	29	NA	13	NA	110

Binary variables

Extern		Female		Academic		Grande Ecole	
N	1,183	N	1,183	N	1,183	N	1,183
Yes	831	Yes	325	Yes	175	Yes	603
No	345	No	853	No	975	No	542
NA	7	NA	5	NA	33	NA	38

of appointments in different categories in Table Appendix V.43.

Figure V.1 shows the development of appointments and shares of these variables as well as the development of female appointments from 2002 to 2018. We see that the appointment of women drastically increased since 2010 caused by the introduction of a women's quota.[1] There is no such clear breaking point in the appointments of non-white or non-French people, but the share of both variables appear to increase over time. While the share of non-French board members is in the range between 22% and 36%, the share of non-whites is clearly lower between 2% and 7%. The share of women starts with values around 7% and rises to slightly above 40% since July 2018, just reaching the defined goal of the women's quota.

[1] There might be anticipation effects that caused female appointments to rise slightly earlier than 2011.

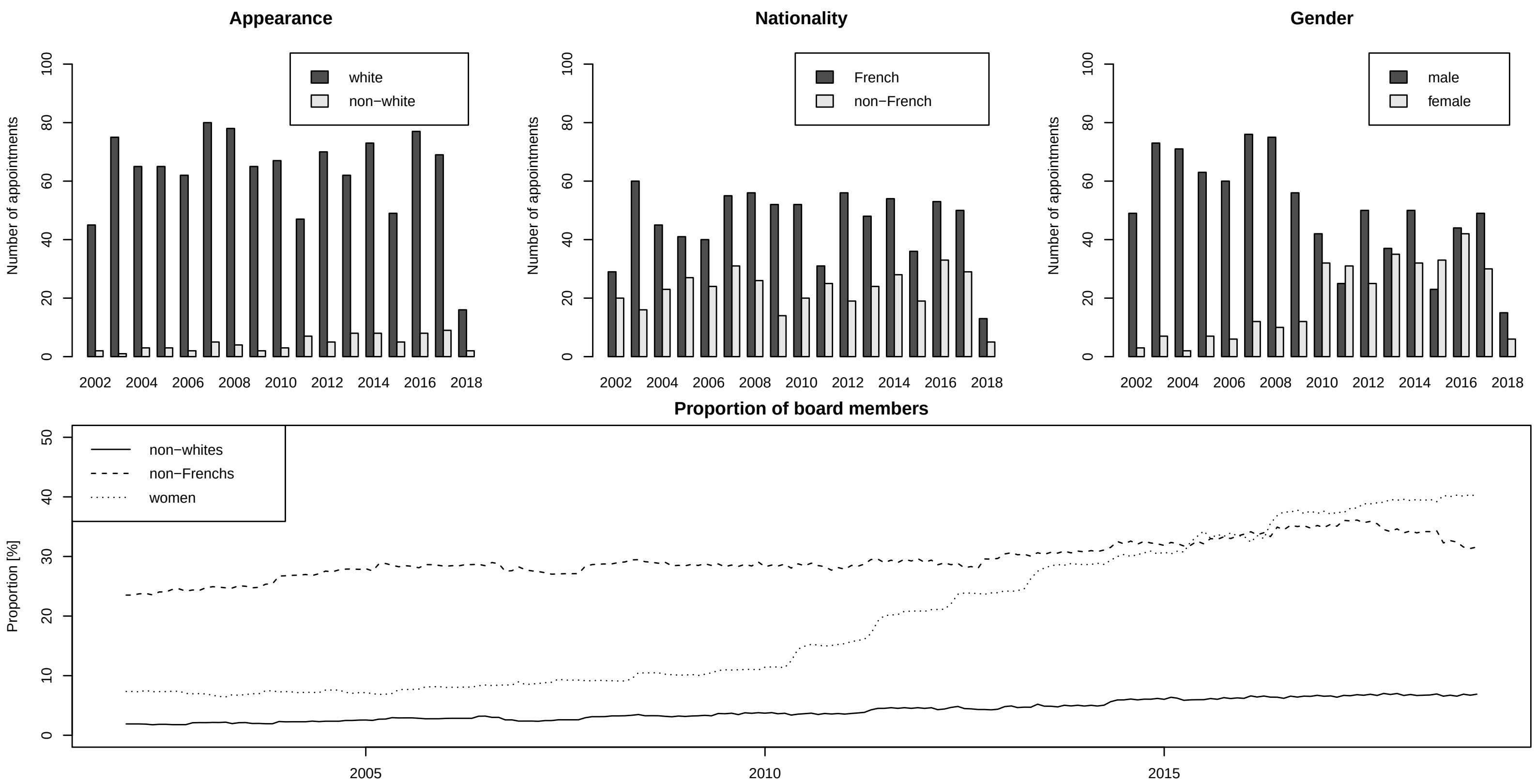


Figure V.1: The figures above show the number of non-white and non-French appointments as well as the number of appointed women between 2002 and 2018. The graphic below shows the development of the share of these three variables in the board. Per construction the numbers of appointments do not consider board members leaving the board, but the share does.

We collect further demographic data, namely whether appointees are external (dummy $Extern = 1$), appointees with doctorate degree (dummy $Academic = 1$) and with degree from a Grande Ecole (dummy $Grandeecole = 1$) as French board members have elite education in many cases (Singh, 2015). Furthermore we control whether an appointee was CEO at the date of appointment or at a later date after the appointment ($Later\ CEO$) and which board they were appointed to. In most cases this is the unitary board, but in case of a dual board we differentiate between the supervisory board and the management board. We collect this variable for each appointment to account for potential changes of firms from unitary to dual board or vice versa. Additionally, we interpret appointments that are not more than two days apart as simultaneous and add a dummy variable $Jointappointment$ that takes on the value 1 if an appointment was simultaneous to another, and 0 otherwise.[2] Finally, we also control for gender and age of an appointee at the date of appointment.

Since we have complete information about appointments not only in the observed time span, but also prior 2002, we can determine the board size at each time point. We use the board size prior to an appointment as control variable ($PriorBoardsize$). Additionally we include the number of women ($\#Prior\ Females$) as well as the number of non-whites ($\#Prior\ non-whites$) or not French citizens ($\#Prior\ non-Frenchs$), respectively, before an appointment. For the latter variables it is necessary that we impute NA cases with the modal category (which is a white French man), because otherwise we would not be able to calculate prior board sizes. Since NA values are not very frequent in our data set and we use imputation only to calculate these control variables -

[2]An example: Let us assume the board size of a firm was 10 in 2002. Then there is an appointment on the 1.1.2003 and another one on the 2.1.2003. Both appointments are not more than two days apart. Thus, we consider them as simultaneous and would use 10 as prior board size for both appointments. Furthermore, the variable *Jointappointment* would be set to 1 for both appointments.

not for analyzing the actual data - we are certain that this does not reduce data quality, but rather improves the regression analyses.

V.3.2 Data on firm level

To construct our performance measures, we use daily share prices from Thomson Reuters Refinitiv and calculate cumulative raw return per month (CRR), cumulative market-adjusted return per month (CMAR), and cumulative risk-adjusted return per month (CRAR) following Adams et al. (2009) and Bechtoldt et al. (2019). While the first measure only reflects monthly stock return for each firm without accounting for market effects and firm risk, the other two measures are adjusted. CMAR controls for general development in the industry or the market by taking the difference between the monthly raw return for each firm and the monthly market index. CRAR controls for the fact that firms might react differently to changes in the market due to different risk profiles (captured by the coefficient beta) and is calculated via the capital asset pricing model (CAPM, Sharpe (1964); Lintner (1965)) as the difference between the actual monthly raw return and the expected stock return.

Furthermore we employ the measure of systematic risk, beta, to also capture potential influences on volatility. These four variables are the dependent variables in our analyses. For simplicity, we summarize all these measures under the term "firm performance".[3]

[3] As a mere measure of volatility, beta does not capture firm performance per se. But we line it up with the other firm performance measures nonetheless, because it captures instability that might translate into concerns about future firm performances.

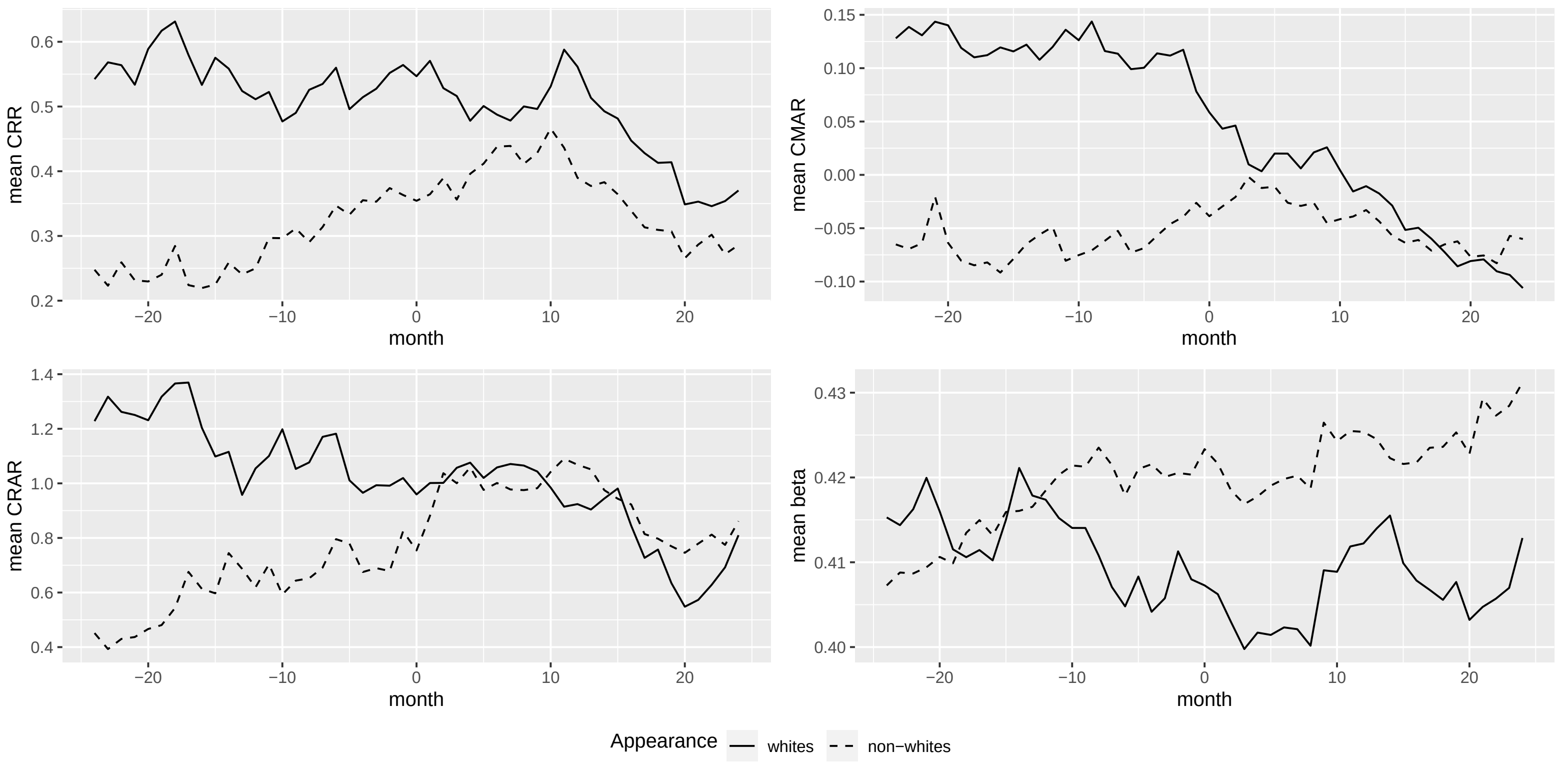


Figure V.2: These figures show the mean of the firm performance measures two years before and after each appointment split by appearance. The graphics are centered at zero, which is the month of appointment. The four figures differ in the measures of firm performance.

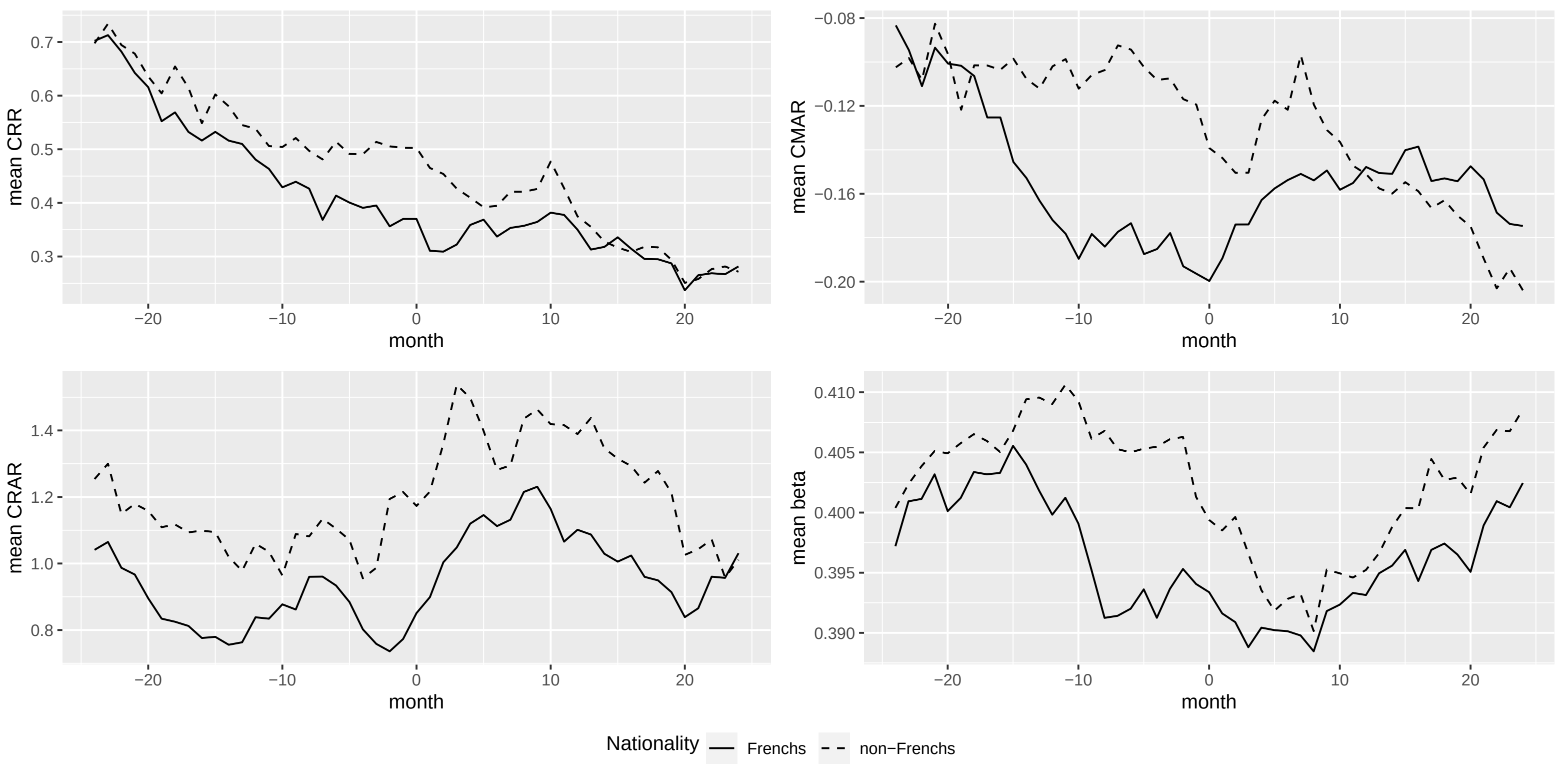


Figure V.3: These figures show the mean of the firm performance measures two years before and after each appointment split by nationality. The graphics are centered at zero, which is the month of appointment. The four figures differ in the measures of firm performance.

Figure V.2 shows the average development of these four firm performance measures relative to appointments for whites as well as for non-whites, Figure V.3 does the same for nationality. Figure V.2 seems to hint that there might be differences in the firm performances between whites and non-whites before an appointment, but we show in the statistical analyses in the results (section V.4) that this is not significantly the case.

We additionally verify our results in the robustness checks for another set of dependent variables. We employ the accounting based measures of firm performance return on assets (ROA) and return on equity (ROE), as well as Tobin's Q (TQ) as it was previously used to investigate the glass cliff (Adams and Ferreira, 2009; Dang and Nguyen, 2018).

As additional control variables we collect the variables industry, number of employees, the natural logarithm of the revenue, EBIT and Debt-equity-ratio (DE) to control for firm-specific influences. Furthermore, we winsorize the variables DE and CRAR to balance out the effects from strong outliers for few firms (Hastings Jr. et al., 1947; Sherman and Tookes, 2022). We show the average firm parameters in the observed timespan in Table V.2. We make the estimated effects in regression models better visible by considering employees in 1,000 [K] and EBIT in billion [B].

Table V.2: Summary statistics of average firm variables from 2002 to 2018

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
# employees [K]	57	98.20	79.73	1.39	45.05	89.95	126.87	411.39
ln(revenue) [B]	57	23.53	0.99	20.92	22.93	23.52	24.14	25.89
EBIT [B]	57	3.00	3.71	0.34	1.01	1.78	3.14	23.63
DE	56	93.36	62.83	9.59	51.09	72.80	125.02	328.15
Beta	56	0.37	0.10	0.15	0.30	0.39	0.43	0.57

V.3.3 Analyses

We start our analyses with an overview of the general development of foreigners in French boardrooms between 2002 and 2018. More precisely we are interested in the probability for a minority appointment based on the year and on firm-specific control variables. We conduct a logistic regression on the appointment-level data and with the two measures for minority affiliation as dependent variable. We vary our models in terms of whether we use the year of appointment ($AppointmentYear$) a dummy variable if the appointment was up until or after 2010 to catch effects of the women's quota ($post2010$) or no time-dependent variable at all. With $AppointmentYear$ we capture a gradient for the temporal development of minorities in CAC40 boards. However, as this procedure requires us to assume a constant slope over 17 observation years, we also use $post2010$ for a simpler before-after-comparison with the introduction of the women's quota as separation threshold.

To analyze our research question, we are interested in the influence of a foreign appointment on the firm performance. Since we compare minority appointments to non-minority appointments, we have to ensure comparability between these two groups. We aim to reduce imbalance in our data set by conducting propensity score matching (Rosenbaum and Rubin, 1985) with regard to all individual and firm-specific control variables as well as the year of appointment, since we only consider appointments to be comparable if they are not too far apart in time. Note that two measures of minority affiliation lead to two different results of the propensity score matching, such that we have to use different partitions of the appointment data set depending on which measure we use.

We run t tests after the matching process to check if the differences in the matched variables vanish, which is sufficiently the case. You can view the results in the Tables V.3 and V.4. Note that - among few others - especially the variables *Extern* and *Grandeecole* still show significant differences in the nationality groups after matching. This reflects a fact which is already to be expected from the very definition of nationality and therefore cannot be resolved with a matching algorithm: It is mainly the group of non-French people who are external and did not visit a French Grande Ecole.

For each of these two partitions we investigate the firm performance - a placeholder for the four independent measures - 24 months before the appointment to estimate connections between minority appointments and firm performance and test our hypotheses. Additionally we observe the firm performance 24 months after the appointment to investigate possible influences of the new board member on the firm performance.

Therefore, we create a panel data set with 49 observations of firm performances per appointment and employ random effects panel regression. We have to use a random effects model, because all individual-specific variables (like *Extern*, *Age*, ...) do not vary on the individual level of the panel regression. A fixed effects model would not be able to estimate effect sizes for these variables as the effect is completely captured in the fixed intercept per individual (compare Wooldridge (2010)).

We are interested whether the firm performance follows a specific pattern before and after an appointment and therefore also account for a possible interaction between the month and the *Foreign* variable - a placeholder for the two variables to assess the affiliation to an ethnic minority. Firms might want to appoint foreign women especially to increase gender and ethnic diversity simultaneously with as few appointments as possible. Thus, we include an interaction term between *Female* and *Foreign* to account for that possibility.

Table V.3: t tests before and after propensity score matching of non-white[a]

	Before Matching				*After Matching*			
	whites (SD)	non-whites (SD)	Difference (SE)	t value	whites (SD)	non-whites (SD)	Difference (SE)	t value
AppointmentYear	2,009.694	2,011.766	2.072	-4.037***	2,012.267	2,012.483	0.217	-0.288
	(4.646)	(4.328)	(0.513)		(4.190)	(4.040)	(0.751)	
# employees	115.479	125.283	9.804	-0.940	115.718	128.533	12.814	-0.705
	(90.041)	(83.919)	(10.430)		(110.199)	(87.657)	(18.179)	
ln(revenue)	23.787	23.689	-0.098	0.834	23.398	23.662	0.265	-1.365
	(1.025)	(0.985)	(0.117)		(1.168)	(0.943)	(0.194)	
EBIT	3.858	2.675	-1.183	3.606***	2.346	2.519	0.173	-0.492
	(4.477)	(2.583)	(0.328)		(1.999)	(1.858)	(0.352)	
DE	109.971	91.407	-18.564	2.459*	79.272	87.212	7.940	-0.856
	(87.930)	(60.112)	(7.549)		(53.975)	(47.359)	(9.270)	
Female	0.269	0.351	0.081	-1.439	0.467	0.367	-0.100	1.107
	(0.444)	(0.480)	(0.056)		(0.503)	(0.486)	(0.090)	
Extern	0.708	0.870	0.162	-3.961***	0.783	0.867	0.083	-1.198
	(0.455)	(0.338)	(0.041)		(0.415)	(0.343)	(0.070)	
Academic	0.150	0.203	0.053	-1.097	0.150	0.167	0.017	-0.248
	(0.357)	(0.405)	(0.048)		(0.360)	(0.376)	(0.067)	
Grandeecole	0.551	0.230	-0.321	6.232***	0.233	0.267	0.033	-0.418
	(0.498)	(0.424)	(0.052)		(0.427)	(0.446)	(0.080)	
Age	54.473	51.794	-2.679	2.407*	52.464	51.357	-1.107	0.692
	(8.138)	(9.119)	(1.113)		(8.915)	(8.612)	(1.600)	
Jointappointment	0.648	0.584	-0.063	1.087	0.550	0.550	0	0
	(0.478)	(0.496)	(0.058)		(0.502)	(0.502)	(0.092)	
Prior Boardsize	15.375	15.117	-0.258	0.535	14.317	14.717	0.400	-0.568
	(4.642)	(4.039)	(0.482)		(3.721)	(3.992)	(0.705)	
# Prior Females	2.547	2.961	0.414	-1.898	2.783	2.983	0.200	-0.563
	(2.135)	(1.824)	(0.218)		(2.084)	(1.799)	(0.355)	
# Prior non-whites	0.533	1.013	0.480	-3.510**	1	1.050	0.050	-0.228
	(0.884)	(1.175)	(0.137)		(1.164)	(1.241)	(0.220)	
# Prior non-Frenchs	4.070	4.325	0.254	-0.630	4.100	3.983	-0.117	0.221
	(3.497)	(3.412)	(0.403)		(2.502)	(3.229)	(0.527)	
Board type								
Unitary board	0.757	0.779	0.023	-0.458	0.750	0.783	0.033	-0.428
	(0.429)	(0.417)	(0.049)		(0.437)	(0.415)	(0.078)	
Supervisory board	0.171	0.182	0.010	-0.228	0.167	0.167	0	0
	(0.377)	(0.388)	(0.046)		(0.376)	(0.376)	(0.069)	
Management board	0.072	0.039	-0.033	1.400	0.083	0.050	-0.033	0.727
	(0.259)	(0.195)	(0.024)		(0.279)	(0.220)	(0.046)	
CEO type								
CEO	0.024	0	-0.024	5.160***	No observations in this category			
	(0.154)	(.)	(0.005)					
Later CEO	0.026	0	-0.026	5.360***	No observations in this category			
	(0.160)	(.)	(0.005)					
No CEO	0.949	1	0.051	-7.539***	All observations in this category			
	(0.219)	(.)	(0.007)					
Industry								
Basic Materials	0.071	0.066	-0.005	0.173	0.017	0.033	0.017	-0.581
	(0.257)	(0.250)	(0.030)		(0.129)	(0.181)	(0.029)	
Consumer Cyclicals	0.256	0.461	0.204	-3.454**	0.467	0.433	-0.033	0.364
	(0.437)	(0.502)	(0.059)		(0.503)	(0.500)	(0.092)	
Consumer Non-Cyclicals	0.098	0.118	0.021	-0.543	0.167	0.150	-0.017	0.248
	(0.297)	(0.325)	(0.038)		(0.376)	(0.360)	(0.067)	
Energy	0.044	0.013	-0.031	2.127*	No observations in this category			
	(0.206)	(0.115)	(0.015)					
Financials	0.144	0.026	-0.118	5.461***	No observations in this category			
	(0.351)	(0.161)	(0.022)					
Healthcare	0.030	0.013	-0.016	1.155	0.033	0.017	-0.017	0.581
	(0.169)	(0.115)	(0.014)		(0.181)	(0.129)	(0.029)	
Industrials	0.125	0.184	0.059	-1.286	0.183	0.233	0.050	-0.670
	(0.331)	(0.390)	(0.046)		(0.390)	(0.427)	(0.075)	
Real Estate	0.036	0.026	-0.010	0.523	0.083	0.033	-0.050	1.165
	(0.188)	(0.161)	(0.019)		(0.279)	(0.181)	(0.043)	
Technology	0.097	0.053	-0.044	1.603	0.033	0.050	0.017	-0.453
	(0.295)	(0.225)	(0.027)		(0.181)	(0.220)	(0.037)	
Utilities	0.100	0.039	-0.060	2.463*	0.017	0.050	0.033	-1.013
	(0.299)	(0.196)	(0.024)		(0.129)	(0.220)	(0.033)	
Observations	1065	77	(NAs: 41)		60	60		

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

[a] This table shows the difference of means for the control variables when comparing the group of whites to the group of non-whites and the results from the corresponding t tests. The first four columns show the results for the complete data set, the latter four columns for the matched data set.

Table V.4: t tests before and after propensity score matching of non-French[a]

	Before Matching				*After Matching*			
	Frenchs (SD)	non-Frenchs (SD)	Difference (SE)	t value	Frenchs (SD)	non-Frenchs (SD)	Difference (SE)	t value
AppointmentYear	2,009.791	2,009.956	0.164	-0.561	2,010.461	2,010.371	-0.089	0.223
	(4.623)	(4.722)	(0.293)		(4.646)	(4.835)	(0.401)	
# employees	122.010	105.044	-16.966	2.970**	112.766	110.657	-2.109	0.272
	(90.612)	(85.674)	(5.712)		(92.107)	(91.348)	(7.752)	
ln(revenue)	23.839	23.688	-0.151	2.241*	23.723	23.680	-0.043	0.456
	(0.980)	(1.084)	(0.068)		(1.121)	(1.107)	(0.094)	
EBIT	3.989	3.517	-0.472	1.657	4.239	3.656	-0.583	1.335
	(4.477)	(4.412)	(0.285)		(5.608)	(4.683)	(0.437)	
DE	117.405	90.982	-26.423	4.851***	100.602	90.608	-9.994	1.515
	(89.110)	(77.681)	(5.447)		(78.618)	(77.465)	(6.596)	
Female	0.253	0.326	0.073	-2.564*	0.314	0.343	0.029	-0.719
	(0.435)	(0.469)	(0.029)		(0.465)	(0.476)	(0.040)	
Extern	0.649	0.851	0.202	-8.044***	0.686	0.846	0.161	-4.567***
	(0.478)	(0.357)	(0.025)		(0.465)	(0.361)	(0.035)	
Academic	0.103	0.251	0.148	-5.920***	0.154	0.250	0.096	-2.859**
	(0.305)	(0.434)	(0.025)		(0.361)	(0.434)	(0.034)	
Grandeecole	0.747	0.072	-0.674	32.556***	0.504	0.064	-0.439	13.176***
	(0.435)	(0.259)	(0.021)		(0.501)	(0.246)	(0.033)	
Age	53.534	55.860	2.326	-4.475***	54.504	56.112	1.608	-2.431*
	(8.294)	(7.774)	(0.520)		(8.123)	(7.513)	(0.661)	
Jointappointment	0.636	0.655	0.020	-0.663	0.654	0.679	0.025	-0.626
	(0.482)	(0.476)	(0.030)		(0.477)	(0.468)	(0.040)	
Prior Boardsize	15.748	14.478	-1.271	4.330***	14.100	13.629	-0.471	1.273
	(4.395)	(4.835)	(0.293)		(4.197)	(4.557)	(0.370)	
# Prior Females	2.703	2.347	-0.356	2.775**	2.704	2.404	-0.300	1.594
	(2.163)	(1.993)	(0.128)		(2.319)	(2.130)	(0.188)	
# Prior non-whites	0.501	0.695	0.194	-3.152**	0.646	0.711	0.064	-0.743
	(0.822)	(1.055)	(0.062)		(0.935)	(1.106)	(0.087)	
# Prior non-Frenchs	3.485	5.217	1.732	-7.438***	3.750	4.454	0.704	-2.570*
	(3.023)	(4.028)	(0.233)		(2.781)	(3.640)	(0.274)	
Board type								
Unitary board	0.764	0.754	-0.010	0.378	0.768	0.754	-0.014	0.396
	(0.425)	(0.431)	(0.027)		(0.423)	(0.432)	(0.036)	
Supervisory board	0.159	0.194	0.035	-1.452	0.171	0.189	0.018	-0.549
	(0.366)	(0.396)	(0.024)		(0.378)	(0.392)	(0.033)	
Management board	0.077	0.052	-0.025	1.669	0.061	0.057	-0.004	0.179
	(0.267)	(0.223)	(0.015)		(0.239)	(0.233)	(0.020)	
CEO type								
CEO	0.023	0.021	-0.002	0.270	0.025	0.025	0	0
	(0.151)	(0.143)	(0.009)		(0.156)	(0.156)	(0.013)	
Later CEO	0.032	0.008	-0.025	3.146**	0.025	0.007	-0.018	1.681
	(0.177)	(0.088)	(0.008)		(0.156)	(0.084)	(0.011)	
No CEO	0.944	0.971	0.027	-2.275*	0.950	0.968	0.018	-1.064
	(0.230)	(0.167)	(0.012)		(0.218)	(0.177)	(0.017)	
Industry								
Basic Materials	0.038	0.135	0.097	-5.038***	0.025	0.075	0.050	-2.728**
	(0.191)	(0.342)	(0.019)		(0.156)	(0.264)	(0.018)	
Consumer Cyclicals	0.285	0.223	-0.062	2.257*	0.193	0.218	0.025	-0.731
	(0.452)	(0.417)	(0.027)		(0.395)	(0.414)	(0.034)	
Consumer Non-Cyclicals	0.100	0.096	-0.004	0.188	0.100	0.104	0.004	-0.140
	(0.300)	(0.296)	(0.019)		(0.301)	(0.305)	(0.026)	
Energy	0.032	0.072	0.039	-2.606**	0.064	0.079	0.014	-0.655
	(0.177)	(0.258)	(0.015)		(0.246)	(0.270)	(0.022)	
Financials	0.151	0.105	-0.047	2.243*	0.132	0.089	-0.043	1.617
	(0.359)	(0.307)	(0.021)		(0.339)	(0.286)	(0.027)	
Healthcare	0.022	0.041	0.020	-1.677	0.039	0.054	0.014	-0.802
	(0.146)	(0.199)	(0.012)		(0.195)	(0.226)	(0.018)	
Industrials	0.126	0.146	0.020	-0.915	0.196	0.179	-0.018	0.541
	(0.332)	(0.354)	(0.022)		(0.398)	(0.384)	(0.033)	
Real Estate	0.028	0.050	0.021	-1.639	0.057	0.054	-0.004	0.184
	(0.166)	(0.217)	(0.013)		(0.233)	(0.226)	(0.019)	
Technology	0.096	0.096	0.0005	-0.025	0.114	0.104	-0.011	0.406
	(0.295)	(0.296)	(0.019)		(0.319)	(0.305)	(0.026)	
Utilities	0.122	0.036	-0.086	5.540***	0.079	0.046	-0.032	1.572
	(0.327)	(0.186)	(0.015)		(0.270)	(0.211)	(0.020)	
Observations	771	383	(NAs: 29)		280	280		

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

[a] This table shows the difference of means for the control variables when comparing the group of appointees with French citizenship to the group of appointees without French citizenship and the results from the corresponding t tests. The first four columns show the results for the complete data set, the latter four columns for the matched data set.

We estimate the model twice, with restriction of the data set for month before the appointment and for months after the appointment. The full model puts together as follows:

$$\begin{aligned}
Firm\ performance_{month} =& Foreign + month + Foreign \times month + Female+\\
& Foreign \times Female + Extern + Academic + Grandeecole+\\
& Age + Jointappointment + Boardtype + CEO+\\
& Prior\ Boardsize + \#\ Prior\ Females + \#\ Prior\ Foreigns+\\
& \#\ employees_{month} + ln(revenue_{month}) + EBIT_{month}+\\
& DE_{month} + industry + u_a + \epsilon_a,
\end{aligned}$$

where u_a stand for the random intercept terms and ϵ_a for the normally distributed error terms for each appointment a.

V.4 Results

We start with analyzing the general development of the probabilities of an ethnic minority appointment with logistic regression models. You can view the results in Table V.5.

The first three columns describe the influence of the firm variables on the probability of a non-white appointment, the latter three for the non-French variable. The influence of the time is captured in the appointment year (2) and (5) or the post2010-dummy (3) and (6). Considering the models (1) - (3), just as Figure V.1 suggests, the appointment year as well as the post2010 show a significant positive influence. Thus, the number of non-white appointments increases over time and there were more such appointments

Table V.5: Firm level logistic regression[a]

Dependent variable	non-white	non-white	non-white	non-French	non-French	non-French
	(1)	(2)	(3)	(4)	(5)	(6)
AppointmentYear		0.007**			0.001	
		(0.002)			(0.003)	
post2010			0.070***			0.018
			(0.020)			(0.028)
# employees	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln(revenue)	-0.009	-0.013	-0.015	0.024	0.024	0.023
	(0.018)	(0.019)	(0.020)	(0.028)	(0.028)	(0.028)
EBIT	0.001	-0.000	-0.001	-0.008	-0.008	-0.008
	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
DE	-0.000	0.000	0.000	-0.001*	-0.001*	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Industry effects						
Consumer Cyclicals	0.052	0.051	0.056	-0.343***	-0.344***	-0.343***
	(0.048)	(0.052)	(0.051)	(0.090)	(0.090)	(0.090)
Consumer Non-Cyclicals	0.019	0.014	0.015	-0.307**	-0.308**	-0.308**
	(0.055)	(0.059)	(0.057)	(0.102)	(0.102)	(0.101)
Energy	0.000	0.000	0.000	-0.089	-0.091	-0.090
	(.)	(.)	(.)	(0.121)	(0.122)	(0.121)
Financials	-0.047	-0.054	-0.052	-0.280**	-0.281**	-0.281**
	(0.049)	(0.053)	(0.052)	(0.104)	(0.104)	(0.104)
Healthcare	-0.032	-0.026	-0.023	-0.138	-0.137	-0.135
	(0.054)	(0.064)	(0.064)	(0.126)	(0.126)	(0.126)
Industrials	0.044	0.032	0.031	-0.248**	-0.250**	-0.250**
	(0.052)	(0.054)	(0.052)	(0.095)	(0.095)	(0.095)
Real Estate	-0.019	-0.028	-0.026	-0.133	-0.136	-0.135
	(0.062)	(0.062)	(0.060)	(0.145)	(0.145)	(0.145)
Technology	-0.031	-0.034	-0.032	-0.294**	-0.295**	-0.294**
	(0.049)	(0.052)	(0.051)	(0.099)	(0.099)	(0.098)
Utilities	-0.034	-0.037	-0.036	-0.465***	-0.466***	-0.466***
	(0.051)	(0.055)	(0.053)	(0.095)	(0.095)	(0.094)
Observations	943	943	943	996	996	996

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

[a] The first three columns (1), (2) and (3) describe the effects of firm specific values on the probability of an appointment of a non-white board member. The columns (4), (5) and (6) describe the influence on the appointment probability of a board member without French citizenship. The reference category for industry effects is Basic Materials. Robust standard errors are reported in parentheses. All reported effects are margins.

after the introduction of the women's quota than before. There is no detectable influence of the number of employees, ln(revenue), earnings before interest and the debt equity ratio. We also detect no influence of a firms industry on the appointment probability of a non-white. The columns (4), (5), and (6) describe the analogous regression models with non-French as dependent variable in the logistic regressions. Interestingly we cannot confirm a significant influence of the time variables. Although they are only controls, it is notable that the debt equity ratio and several industry branches have a significant negative influence on the probability of appointing a non-French board member compared to Basic Materials, which is the reference category. So in Basic Materials there seem to be more non-French board members than in many other industries.

In our main analysis we test the direct influence of foreign appointments on firm performance, 24 months before an appointment to detect a potential glass cliff and 24 months thereafter to check if firm performances differ subsequent to the appointments. For both variables non-white and non-French we use propensity score matched data sets to better compare foreign to non-foreign board members. We account for monthly repeated firm performance observations belonging to the same appointment by using a panel regression. View the results for non-white in Table V.6.

The first four columns show the influence on the firm performance variables before the appointment. The appointment of a board member with non-white appearance does not significantly affect any of the firm performance measures, neither generally, nor in interaction with month. So we are not able to find any traces of the glass cliff. The last four columns show the panel regressions if we observe the 24 months after an appointment. We use such an analysis to check for differences of potentially different management styles of non-white board members. More precisely we investigate whether a potential glass cliff could be justified. Once again, this variable shows no significant effects. Thus, we cannot detect any differences of non-white appointments to white appointments after the date of appointment.

With only 120 appointments left in the matched data set, this subset could appear to be under-powered for our analysis. But note that even in the complete data set we only observe 77 non-white appointments. So weak effects cannot be caused by poor choice of a subset, but because non-white members are weakly represented in CAC40 boardrooms in general.

We repeat this panel regression for the non-French variable and show the results in Table V.7.

Considering the columns before an appointment (1) - (4), it is interesting to note, that we detect a significant negative influence of month in three of the four dependent variables, but no significant influence of the non-French variable. This means that in this matched data set firm performances declines in the months before an appointment in general, but without any influence of the nationality of the appointee. This results hints a tendency for firms to appoint new board members in response to a crisis. However, this insight has to be interpreted cautiously, because we only observe a significance for the subset of our data tailored for the non-French variable. Since there is no correlation between firm performance and nationality, we once again detect no traces of a glass cliff. Observing the development of firm performance after an appointment, we also cannot confirm negative or positive firm performance developments after a non-French appointment. There is an exception in column (6), where we detect a negative slope for CMAR after a non-French appointment. However, the effect is weak and barely offsets the non-significant positive effect of the non-French variable after two years. This is insufficient to conclude any differences between nationalities after an appointment.

In conclusion, in none of the analyses we were able to detect any traces of the glass cliff, nor did we detect differences in firm performances for any foreign board member after an appointment. This is not to say that a glass cliff for minorities does not exist in French CAC40 companies, but considering our long observation time span, our results should be taken as indication that minorities might not face an obstacle like the glass cliff in large French firms.

Table V.6: Panel regression of firm performance 24 months prior and after an appointment (propensity score matched data set of non-white)[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:	Before				After			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-white	-0.127	0.051	-0.313	0.031	0.043	0.033	0.172	0.036
	(0.124)	(0.107)	(0.398)	(0.025)	(0.176)	(0.127)	(0.444)	(0.023)
month	-0.001	-0.001	-0.010	-0.001	-0.008	-0.006**	-0.015	0.000
	(0.006)	(0.003)	(0.013)	(0.001)	(0.004)	(0.002)	(0.014)	(0.000)
non-white × month	0.008	0.002	0.021	0.001	0.002	0.004	0.007	0.000
	(0.007)	(0.004)	(0.016)	(0.001)	(0.006)	(0.003)	(0.018)	(0.001)
Appointment-specific control variables								
Female	-0.032	0.180	-0.648	0.013	0.232	0.080	-0.121	0.021
	(0.165)	(0.101)	(0.467)	(0.029)	(0.207)	(0.153)	(0.512)	(0.026)
non-white × Female	0.017	-0.304	0.400	-0.028	-0.304	-0.110	-0.054	-0.039
	(0.221)	(0.177)	(0.579)	(0.035)	(0.241)	(0.181)	(0.690)	(0.036)
Extern	-0.090	-0.032	0.615	0.002	0.082	0.069	0.445	0.022
	(0.154)	(0.091)	(0.416)	(0.024)	(0.131)	(0.097)	(0.484)	(0.021)
Academic	-0.252*	-0.039	-0.174	0.029	-0.085	-0.093	0.022	0.010
	(0.128)	(0.096)	(0.409)	(0.021)	(0.190)	(0.128)	(0.478)	(0.020)
Grandeecole	0.102	0.040	-0.248	-0.034	0.078	0.120	-0.119	-0.008
	(0.127)	(0.086)	(0.334)	(0.018)	(0.112)	(0.082)	(0.414)	(0.017)
Age	0.011	0.007	-0.027	-0.001	0.016*	0.002	-0.007	-0.002
	(0.007)	(0.004)	(0.019)	(0.001)	(0.007)	(0.004)	(0.025)	(0.001)
Jointappointment	0.011	0.062	-0.138	0.012	-0.105	0.057	-0.225	0.020
	(0.123)	(0.092)	(0.377)	(0.020)	(0.127)	(0.094)	(0.347)	(0.019)
Supervisory board	-0.583**	-0.450**	0.684	0.047*	-0.699***	-0.156	0.003	0.078**
	(0.195)	(0.139)	(0.554)	(0.023)	(0.182)	(0.157)	(0.519)	(0.024)
Management board	-1.257***	-0.783***	0.652	0.085	-0.984***	-0.467*	-0.173	0.122**
	(0.258)	(0.233)	(0.772)	(0.047)	(0.282)	(0.219)	(0.917)	(0.046)
Prior Boardsize	-0.028	-0.018	0.065	-0.008*	0.000	-0.006	0.110	0.001
	(0.023)	(0.019)	(0.061)	(0.004)	(0.021)	(0.013)	(0.073)	(0.003)
# Prior Females	0.059	0.028	-0.267**	0.025***	-0.005	-0.013	-0.143	0.010
	(0.037)	(0.029)	(0.099)	(0.005)	(0.036)	(0.024)	(0.099)	(0.005)
# Prior non-whites	-0.022	-0.021	0.412	-0.016	0.060	0.076	0.101	-0.003
	(0.059)	(0.038)	(0.219)	(0.009)	(0.069)	(0.046)	(0.170)	(0.009)
Firm-specific control variables								
# employees	-0.003*	-0.002**	0.003	-0.000	-0.003**	-0.002**	-0.003	0.000
	(0.001)	(0.001)	(0.002)	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)
ln(revenue)	0.239	-0.038	0.422	0.033*	-0.145	-0.154*	-0.109	-0.011
	(0.136)	(0.133)	(0.408)	(0.016)	(0.089)	(0.061)	(0.355)	(0.010)
EBIT	-0.032	0.024	-0.097	0.007	0.067*	0.082*	-0.273**	0.008
	(0.053)	(0.039)	(0.133)	(0.004)	(0.033)	(0.038)	(0.105)	(0.005)
DE	-0.001	0.001	-0.009*	-0.000	-0.002	0.000	-0.005	-0.000
	(0.002)	(0.001)	(0.004)	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)
Consumer Cyclicals	0.565*	0.139	-0.770	-0.190	0.516	-0.280	-0.054	-0.229
	(0.238)	(0.209)	(1.075)	(0.131)	(0.296)	(0.270)	(0.907)	(0.150)
Consumer Non-Cyclicals	0.106	0.148	-1.007	-0.106	0.413	0.067	0.447	-0.085
	(0.203)	(0.199)	(0.939)	(0.132)	(0.274)	(0.266)	(0.912)	(0.152)
Healthcare	0.316	-0.084	-1.444	-0.229	-0.349	-0.268	0.512	-0.182
	(0.548)	(0.415)	(1.334)	(0.135)	(0.387)	(0.444)	(1.370)	(0.164)
Industrials	0.274	0.224	-1.731	-0.140	0.750*	0.053	-0.083	-0.117
	(0.240)	(0.215)	(0.895)	(0.132)	(0.313)	(0.276)	(0.861)	(0.152)
Real Estate	1.324**	0.634	-1.256	-0.007	0.926**	-0.182	-0.725	-0.160
	(0.500)	(0.370)	(1.487)	(0.137)	(0.342)	(0.325)	(1.204)	(0.154)
Technology	-0.669***	-1.034***	0.578	-0.347*	0.304	-0.500	4.643**	-0.366*
	(0.157)	(0.305)	(1.482)	(0.136)	(0.247)	(0.278)	(1.422)	(0.152)
Utilities	-0.255	-0.755*	-3.377**	-0.268*	0.439	-0.354	-2.872**	-0.319*
	(0.306)	(0.296)	(1.097)	(0.134)	(0.407)	(0.385)	(0.923)	(0.152)
constant	-5.038	0.861	-6.768	-0.111	2.920	3.708**	3.847	0.842**
	(3.036)	(2.924)	(9.020)	(0.369)	(2.002)	(1.292)	(8.203)	(0.256)
Observations	2698	2701	2537	2698	2782	2783	2648	2782

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on the propensity score matched data set for appearance. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board) and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table V.7: Panel regression of firm performance 24 months prior and after an appointment (propensity score matched data set of non-French)[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:		Before				After		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-French	-0.009	0.096	0.171	0.012	0.126	0.087	0.138	0.001
	(0.165)	(0.074)	(0.304)	(0.013)	(0.110)	(0.072)	(0.340)	(0.015)
month	-0.015***	-0.006**	-0.006	-0.001*	-0.004	-0.001	-0.005	0.000
	(0.005)	(0.002)	(0.009)	(0.000)	(0.003)	(0.002)	(0.009)	(0.000)
non-French × month	0.005	0.005	0.004	0.001	-0.007	-0.005*	-0.009	0.000
	(0.007)	(0.003)	(0.012)	(0.000)	(0.004)	(0.002)	(0.012)	(0.000)
Appointment-specific control variables								
Female	-0.236	0.025	-0.362	0.016	-0.077	-0.012	-0.042	0.011
	(0.142)	(0.075)	(0.245)	(0.014)	(0.098)	(0.069)	(0.309)	(0.016)
non-French × Female	0.258	0.077	-0.028	0.021	0.166	0.048	0.130	0.013
	(0.223)	(0.103)	(0.365)	(0.019)	(0.133)	(0.096)	(0.414)	(0.021)
Extern	0.067	-0.046	0.202	-0.008	-0.308**	-0.131*	-0.412	0.010
	(0.149)	(0.069)	(0.254)	(0.013)	(0.095)	(0.061)	(0.291)	(0.013)
Academic	0.047	-0.083	-0.107	-0.001	0.075	-0.069	0.177	-0.007
	(0.159)	(0.068)	(0.227)	(0.011)	(0.087)	(0.063)	(0.242)	(0.011)
Grandeecole	-0.000	0.035	0.254	0.006	0.041	0.031	0.034	0.007
	(0.130)	(0.069)	(0.286)	(0.012)	(0.083)	(0.065)	(0.320)	(0.014)
Age	-0.001	-0.001	0.021	-0.001	0.011*	-0.003	0.037**	-0.002*
	(0.008)	(0.003)	(0.012)	(0.001)	(0.005)	(0.003)	(0.014)	(0.001)
Jointappointment	-0.482**	-0.128*	-0.057	-0.002	-0.070	0.007	0.308	0.008
	(0.175)	(0.055)	(0.209)	(0.010)	(0.080)	(0.055)	(0.215)	(0.010)
Supervisory board	0.167	0.036	-1.189***	0.039**	-0.191	-0.086	-0.297	0.017
	(0.242)	(0.068)	(0.246)	(0.013)	(0.102)	(0.062)	(0.260)	(0.014)
Management board	0.037	-0.186	-1.179***	0.018	-0.264	-0.161	-0.109	0.021
	(0.576)	(0.126)	(0.334)	(0.018)	(0.180)	(0.096)	(0.409)	(0.019)
CEO	0.128	-0.059	0.114	-0.020	-0.154	-0.277*	0.389	-0.016
	(0.268)	(0.151)	(0.543)	(0.026)	(0.172)	(0.128)	(0.522)	(0.021)
Later CEO	-0.554	-0.063	-0.296	-0.027	-0.188	0.091	-0.905	0.029
	(0.326)	(0.146)	(0.805)	(0.025)	(0.195)	(0.190)	(0.819)	(0.032)
Prior Boardsize	-0.015	-0.014	0.032	-0.001	0.004	0.012	-0.018	0.005***
	(0.017)	(0.009)	(0.032)	(0.001)	(0.010)	(0.008)	(0.032)	(0.001)
# Prior Females	-0.086***	0.007	-0.130*	0.011***	-0.010	-0.022	0.041	0.003
	(0.026)	(0.012)	(0.053)	(0.002)	(0.016)	(0.012)	(0.050)	(0.002)
# Prior non-Frenchs	-0.133***	-0.086***	0.145**	-0.009***	0.027	-0.040***	0.191***	-0.010***
	(0.033)	(0.012)	(0.046)	(0.002)	(0.021)	(0.012)	(0.046)	(0.002)
Firm-specific control variables								
# employees	-0.003**	-0.001*	0.001	-0.000	-0.000	-0.001***	0.003	-0.000
	(0.001)	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)
ln(revenue)	-0.022	0.023	0.291	0.031*	-0.373***	0.045	-1.209***	0.035***
	(0.189)	(0.074)	(0.328)	(0.012)	(0.088)	(0.050)	(0.319)	(0.008)
EBIT	0.010	0.012***	-0.074	0.002	0.029***	0.005	0.085	0.003
	(0.008)	(0.004)	(0.046)	(0.002)	(0.006)	(0.004)	(0.052)	(0.002)
DE	-0.002*	-0.001**	-0.005*	-0.000*	-0.001*	-0.002***	-0.004*	-0.000
	(0.001)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)
Consumer Cyclicals	-0.104	-0.557***	2.292***	-0.176***	-0.127	-0.681***	2.274***	-0.213***
	(0.260)	(0.123)	(0.414)	(0.035)	(0.163)	(0.117)	(0.366)	(0.034)
Consumer Non-Cyclicals	0.105	-0.196	0.233	-0.088*	-0.091	-0.331**	1.247**	-0.106**
	(0.274)	(0.127)	(0.487)	(0.037)	(0.154)	(0.118)	(0.454)	(0.036)
Energy	1.028	-0.545***	0.211	-0.169***	-0.119	-0.827***	1.060	-0.175***
	(0.606)	(0.151)	(0.346)	(0.040)	(0.251)	(0.176)	(0.843)	(0.046)
Financials	0.039	-0.627***	0.707	-0.188***	-0.176	-0.668***	1.182*	-0.238***
	(0.258)	(0.128)	(0.499)	(0.037)	(0.166)	(0.121)	(0.500)	(0.036)
Healthcare	-0.199	-0.323**	1.744**	-0.180***	-0.371*	-0.465***	1.542*	-0.198***
	(0.233)	(0.123)	(0.599)	(0.038)	(0.150)	(0.123)	(0.673)	(0.037)
Industrials	0.160	-0.225	0.358	-0.161***	0.456*	-0.248*	1.750***	-0.185***
	(0.266)	(0.129)	(0.395)	(0.037)	(0.180)	(0.113)	(0.376)	(0.035)
Real Estate	0.498	0.290	1.454	-0.010	-0.138	0.009	-0.463	-0.095*
	(0.525)	(0.186)	(0.773)	(0.047)	(0.206)	(0.151)	(0.619)	(0.042)
Technology	-0.727**	-1.579***	1.775***	-0.300***	-0.514***	-1.335***	2.133***	-0.301***
	(0.260)	(0.157)	(0.487)	(0.036)	(0.153)	(0.143)	(0.516)	(0.035)
Utilities	-0.299	-0.868***	0.202	-0.215***	-0.234	-0.840***	1.146	-0.282***
	(0.321)	(0.170)	(0.824)	(0.042)	(0.182)	(0.149)	(0.690)	(0.038)
constant	2.580	0.617	-8.091	-0.119	8.899***	-0.083	25.337***	-0.233
	(4.002)	(1.592)	(7.118)	(0.267)	(1.971)	(1.119)	(7.184)	(0.188)
Observations	12209	12239	10851	12209	12442	12476	11166	12442

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on the propensity score matched data set for nationality. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board), no-CEO-appointment and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

V.5 Robustness checks

We apply several alternative analyses to test the robustness of our results and address the concern that the non-significant results might be only due to a disadvantageous design of the analyses.

First we refer back to the Figures V.2 and V.3, where at least the first Figure seems to hint that the appointment of a non-white might be preceded by a weaker company performance. One could object, that there actually are differences in firm performances which are obscured in our main analyses by the variety of control variables. So we exclude all control variables and even excluded the interaction term between month and Foreign, such that the main effect of the Foreign variable captures the average firm performance in the 24 months before (or after respectively) an appointment. This counters the concern, that the month of appointment could be an unfavorable reference point within the 24 month time span.
In case we detect differences for non-white or non-French board members in firm performances excluding other variables, this would be a sign of multicollinearity in our data, more precisely that there actually is a glass cliff which is not triggered by the affiliation of an ethnic minority itself, but by other variables that go along with it. However, our robustness check (see Table Appendix V.46) shows that this is only case for the cumulative raw return for non-whites. As this result is not reflected in all the other firm performance measures and also not for non-Frenchs, this is too weak to claim a glass cliff. Thus, the differences as shown in the Figures V.2 and V.3 cannot be confirmed statistically.

In the next robustness check we use alternative measures for firm performance. While market-based measures also capture anticipated future development of a firm as perceived by the market (Fama et al., 1969; Fama, 1970), accounting-based measures solely evaluate performance of the past. Thus, we employ return on assets (ROA) besides return on equity (ROE) as accounting-based returns and additionally Tobin's Q (TQ) since these measures are also widely used in prior literature (e.g. Adams and Ferreira (2009); Carter et al. (2010); Dang and Nguyen (2018)). ROA is defined as the EBIT for the fiscal year divided by the book value of total assets at the end of the year. ROE is the EBIT divided by the book value of equity at the end of the year (compare Haslam et al. (2009)). In both cases, we use the EBIT instead of net income for the return numerator since this is not as much influenced by balance sheet optimization as the net income is. In addition, the EBIT values are trailing on a monthly basis[4]. To circumvent the annual measurement of each denominator, total assets and total equity, they are calculated as the sum of the current and following year and divided by 2. This approach is also covered multiple times in financial literature. The trailing values for the denominator were too inconsistent in the data provider such that this method granted more success (Cheema-Fox et al., 2021). Tobin's Q is the ratio of market value and asset replacement costs (Tobin and Brainard, 1977). We show the results for the alternative measures in the Tables Appendix V.47 and Appendix V.48.

The only relevant result in favor of a potential glass cliff is for Tobin's Q before an appointment in the non-French variable and its interaction with month (see column (3) in Table Appendix V.48). So we can state, that Tobin's Q is lower for non-French appointees before an appointment, but since this measure is the only out of in total seven market performance measures to produce a significant result, this is far too weak to claim the existence of a glass cliff.

[4]This means for example, the EBIT of month February starts in the prior year's February such that it always is a twelve-months-period.

We argued earlier, that the introduction of the women's quota in 2011 could spark the employment of ethnic minority board members to signal change and therefore increase a glass cliff for minorities since 2011. However, contrary to that it is also possible that the women's quota caused a rethinking regarding diversity in boardrooms leading to a growth of ethnic minority representation alongside women, but mitigating a potential glass cliff that might be present before. In either case it is interesting to investigate data before and since 2011 separately. We check for a potential glass cliff considering only appointments before 1.1.2011 and present the results in the Tables Appendix V.49 and Appendix V.50. Similarly, we check for a glass cliff using only appointments since 1.1.2011 and present the results in the Tables Appendix V.51 and Appendix V.52. In both cases, considering the effects of the foreign variables and their interaction with months, we detect some significant effects before and after appointments. But again none of these effects are stable throughout multiple performance measures. Moreover, many of the significant effect sizes are positive and therefore even contradict the glass cliff hypothesis. Thus, we cannot verify, that the lack of measuring a glass cliff in the main analysis is in any way caused by the introduction of the women's quota.

For all following robustness checks we change definitions of control variables or exclude firms from our data set. Therefore, we have to repeat the matching algorithm for the modified variables or data sets. That leads to minor changes in the observed data points compared to the main analyses, but increases overall data quality as we find better matching pairs for the robustness check analyses.

For the next robustness check we counter a potential objection to the usage of absolute numbers of women or foreigners in a board to represent prior board composition. We included count variables in the main analyses to be consistent to the board size which is also a count variable, but it is also conceivable that the shares of women or foreigners play a greater role in the appointment decisions of further foreigners. Thus, we replace the counts of women, non-whites and non-French citizens with their shares in the board prior to an appointment and repeat the main analyses. We present the results in Tables Appendix V.53 and Appendix V.54.

Next we account for the concern that there might be differences for firms that were not listed in the CAC40 the entire time span of observation. We exclude these firms and furthermore drop firms that merged with other firms or exhibited other structural changes between 2002 and 2018 as such changes might influence board appointments. We remain with a subset of 725 appointments of 24 firms and repeat the matching algorithm and main analyses. We show the results in Appendix V in the Tables Appendix V.55 to Appendix V.56. Overall the results do not deviate from the observations in the main analyses and confirm the previous findings, except for the fact that we detect a significant positive effect for CRR before an appointment of a non-French board member, and positive effects of CRR and CMAR paired with a negative monthly trend after an appointment of a non-French board member, as Table Appendix V.56 shows. On one hand this means that firms had higher values of CRR as usual directly before the appointment of a board member with a foreign citizenship, but again contrary to our hypothesis and not stable over the other measures. This creates further doubts about the hypothesis that CAC40-firms increase ethnic board diversity in times of crisis. On the other hand, firms appointing non-French board members perform better after an appointment in two measure CRR and CMAR, but the interaction of non-French and month implies that this effect vanishes over the following years (columns (5) and (6) of Table Appendix V.56). We have no hypothesis for such a finding and - although we cannot find this effect in other measures or other analyses - the visibility in at least two performance measures might spark future research in positive short-time effects of appointees from another country.

The next robustness check only refers to the non-white variable. We attempt an alternative definition in which we also rate Hispanic appointees as non-white, such that only people with Caucasian appearance are rated as white. We display the results in Table Appendix V.57. Interestingly we now find a positive influence of non-white before and after an appointment when we take beta as performance measure. Once again this goes against a potential glass cliff.

A further concern might be the matching algorithm to create balanced data sets in the Foreign variables. The data set and therefore our results strongly depend on the matched appointees. In this last robustness check we tackle this concern by employing an alternative algorithm to obtain the matched data sets. We follow Bechtoldt et al. (2019) and use genetic matching (Diamond and Sekhon, 2013), which is a non-parametric algorithm developed by Sekhon and Mebane Jr. (1998) to determine the weight of each covariate. Genetic matching is a generalization of propensity score matching as is searches among a set of distance metrics (also including the propensity score) to find the metric that optimizes the overall balance. The results are presented in the Tables Appendix V.58 and Appendix V.59. The alternative matching algorithm does not find any connections between firm performance and our measures of foreign affiliation before an appointment, and therefore no glass cliff. However, we find significant negative interactions between non-French and month of CRR and CMAR after an appointment of a non-French board member (columns (5) and (6) of Table Appendix V.59). But the effect sizes are economically weak. Even after two years - the end of our observation time after an appointment - the interaction terms cannot even equalize the non-significant positive effects of the non-French variable.

All these robustness checks do not change our overall results. Although we measure some significant effects of the Foreign variable on firm performance, these effects either appear only for a single firm performance variable - that is they are unstable in different measures - or even are positive and therefore hint against the notion that firm performance declines prior to a foreign appointment. Thus, we cannot find any evidence for a glass cliff for minorities in French CAC40 boardrooms.

V.6 Conclusion

In this study, we transfer the well-known phenomenon of the glass cliff (see Ryan and Haslam (2005)), where women are more likely to get appointed to higher positions in times of crisis, to ethnic minorities. As ethnic minorities are also an underrepresented group in corporate boards, they might get promoted to signal change after declining firm performance. We focus on France, a country with an especially high mix of different minority cultures. France introduced a 40% quota for women in 2011 (Assemblée nationale, 2011), which obstructs a potential signaling effect for women (Kulich et al., 2015) and makes ethnic minorities an even more interesting group to bring change to a firm board. Identifying minority affiliation via the subjective measure of appearance or alternatively via the objective measure of citizenship, our analysis could not find a negative correlation between different measures of firm performance before an appointment and the probability of an appointment of an ethnic minority board member compared to a non-minority board member. This result remains stable for several measures of firm performance like monthly raw returns, accounting based measures beta and Tobin's Q. Also various restrictions of the data set and redefining variables did not hint a potential glass cliff.

However, concluding that a class cliff does not exist in French CAC40 firms could be an over-interpretation, not only because null-results always have the problem of an unknown error likelihood, but also because of several limitations in our data set. First of all our data set is rather small after matching, especially for the definition of minority affiliation via appearance, where we only have 120 observations. This shortcoming cannot be resolved by another matching algorithm or another research approach, because our data set already is a complete lineup of appointees in the observed group. Even without matching the pool of non-white appointees is only slightly larger. Future studies may increase the data set by including further French companies with the CAC Next 20, the CAC Mid 60, or the CAC Small. Further approaches may also extent this topic by comparable countries. Ruigrok et al. (2007) come to the conclusion based on Swiss firms that varying national circumstances might lead to the issue, that results from one country could not be comparable to another. Especially for ethnic minorities, we can expect that obstacles and representation in firm boards vary between different countries and cultures. The ideal scenario would include a global comparison although this evokes regulatory complications to align first.

One might also criticize that our data lacks some further informational variables, for example a distinction of shareholders. Additionally, further research could have a closer look on personal features of appointees, such as better overall education.

Another interesting topic for future research could center upon a regulatory implemented quota to enhance representation of minorities. While the women's quota in France forced boards to change drastically, no such law exists for ethnic minorities. As our results show that minorities are distinctly less represented in boards than women, further research could deal with the question whether an extension of quota laws to other minorities could be advantageous for corporate boards.

Despite these shortcomings, overall our findings give interesting insights on ethnic minorities in French boards. On one hand, the representation of ethnic minorities defined by appearance is very small in French firm boards. The lack of a glass cliff could hint a more fundamental problem of a glass ceiling (see Boyd (2008)). A follow-up study could investigate, whether ethnic minorities are represented more widely in lower positions of corporate power and whether they face an invisible wall preventing many of them to reach higher firm positions. On the other hand, even extensive variations of definitions and analyzed data - also including citizenship as definition for ethnic minorities, which is a far more present group in firm boards - could not find evidence for a performance-based increase in minority appointment. This should be taken as hint, that - of all potential hindrances - at least a glass cliff is no relevant obstacle for ethnic minorities aspiring to corporate leadership positions in France.

To summarize some takeaways from our results, the representation of ethnic minorities in French firm boards is still small. Given the positive effects of ethnic minority appointees on workforce and their higher levels of human capital and advanced education according to Singh (2007) as well as the growing importance of ethnic diversity in times of globalization, firms should consider to appoint more board members from other countries. Our results also show that firms do not need to shy away from appointing minority board members, as the firm performance does not suffer from such appointments. However, the strongly different observation sizes in our two definitions of ethnic minorities show that ethnic minority is not a well-defined term. So potential advantages of minorities in boards might - besides the industry of a firm - also depend on the exact definition of ethnic minority. Decision makers in companies trying to implement ethnic diversity will have to carefully evaluate what new beneficial points of views are needed within the management level of their firm and which qualities an outsider should bring with to prepare the firm well for future challenges.

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.8: 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.9: 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.10: 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.11: 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	-

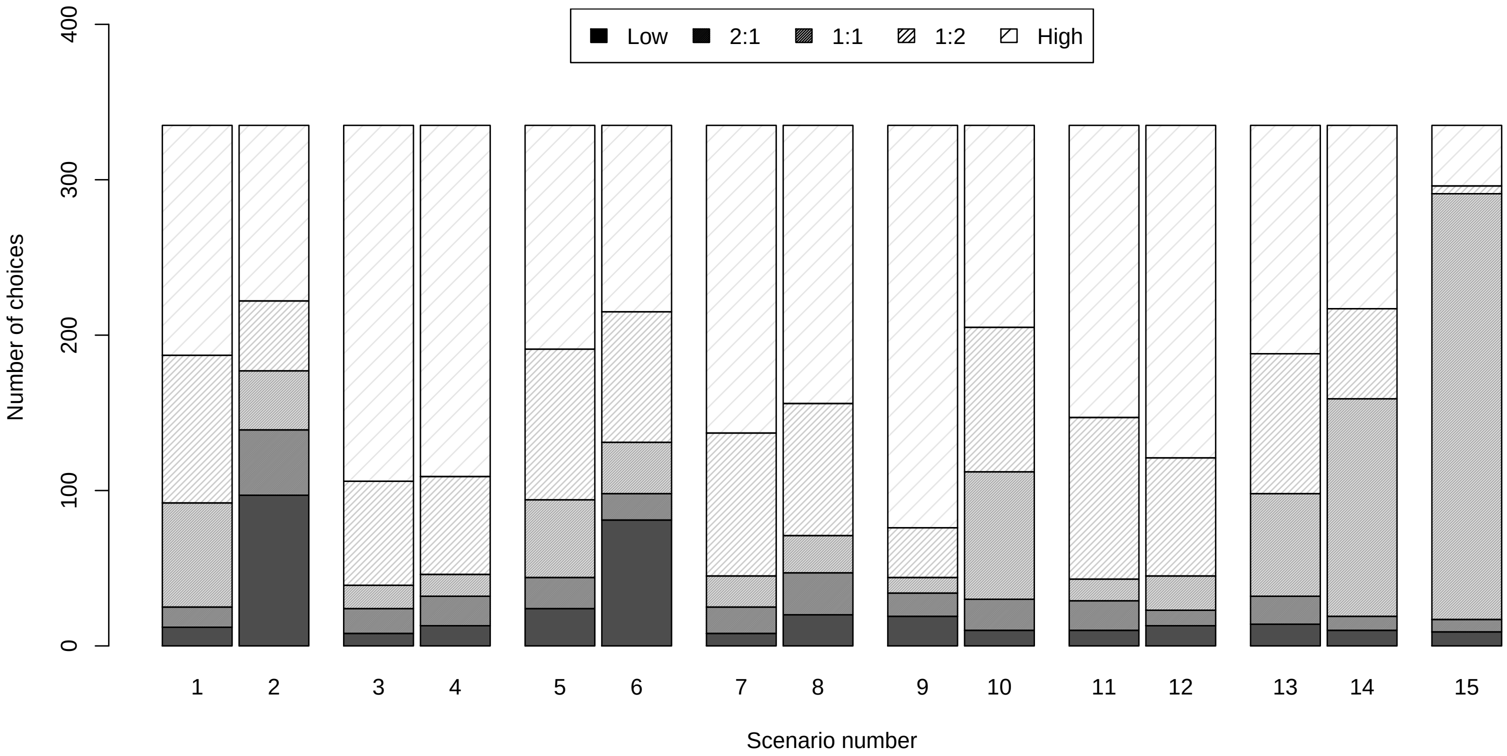


Figure Appendix I.4: Relative proportion of choices in the scenarios

Table Appendix I.12: 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.13: 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.14: 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.15: 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.16: 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.17: 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.[5]

[5]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 Figures and Tables

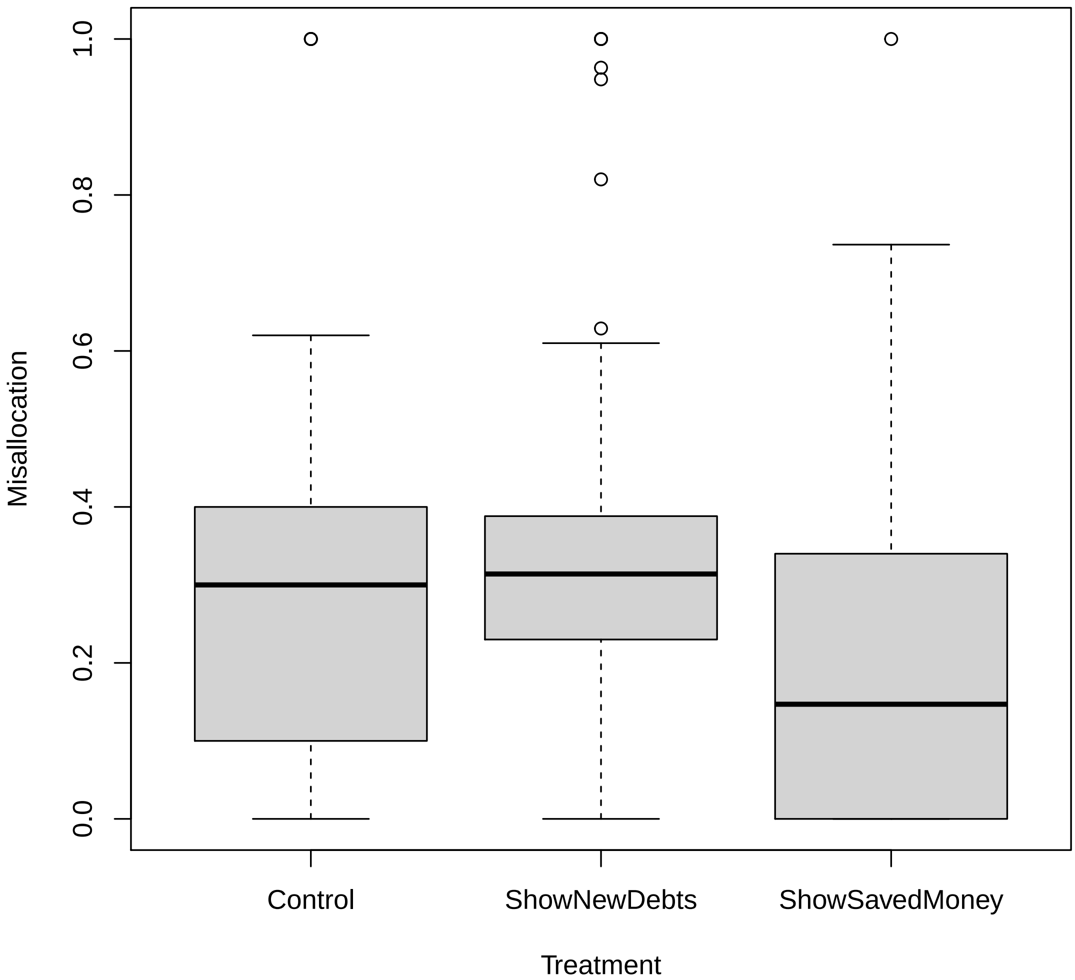


Figure Appendix II.5: Boxplots of misallocation split by treatment. ShowNewDebts is the sludge and ShowSavedMoney is the nudge treatment.

Table Appendix II.18: 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.19: 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.20: 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.21: 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.22: 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.23 and Figure Appendix II.6. 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.

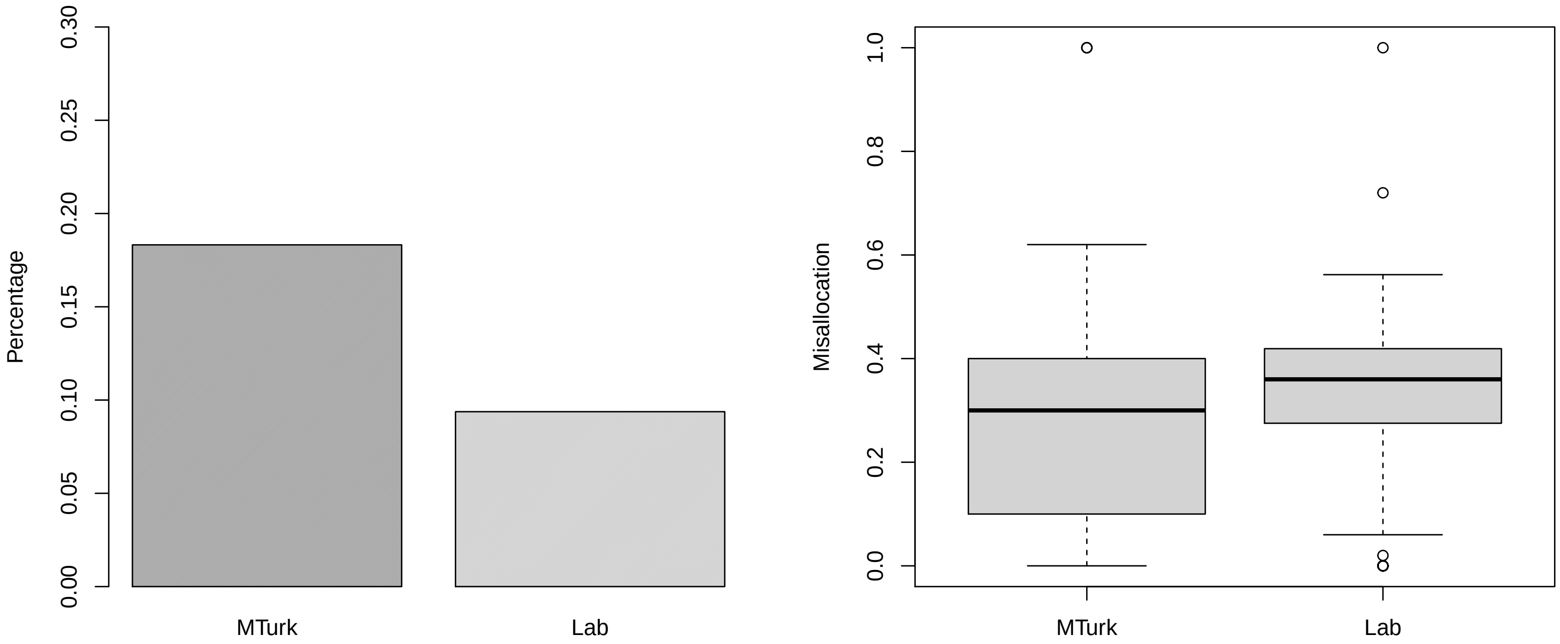


Figure Appendix II.6: The bars in the left Figure show the proportion of subjects without any misallocation in all the experiment rounds. 18.3% of the subjects in MTurk and 9.4% of the subjects in the lab do not exhibit misallocation. The difference is not significant (p=0.0526). The right Figure shows 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.23: 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. $\square$

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 Graphics and Tables

Table Appendix III.24: 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.25: 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.26: 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.27: 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*** (0.019) [0.000] [0.000]	0.242*** (0.018) [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*** (0.019) [0.000] [0.000]	Reference (.) [.] [.]	0.030** (0.011) [0.008] [0.008]
Adapted		-0.242*** (0.018) [0.000] [0.000]	-0.030* (0.011) [0.008] [0.016]	Reference (.) [.] [.]
Financial literacy	-0.055*** (0.014)	-0.055*** (0.014)	-0.028*** (0.007)	-0.003 (0.006)
Interactions between financial literacy and treatment (group)				
Control × FL	Reference (.) [.] [.]	Reference (.) [.] [.]	-0.026 (0.016) [0.099] [0.099]	-0.052** (0.015) [0.001] [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.28: 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.29: 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.30: 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.

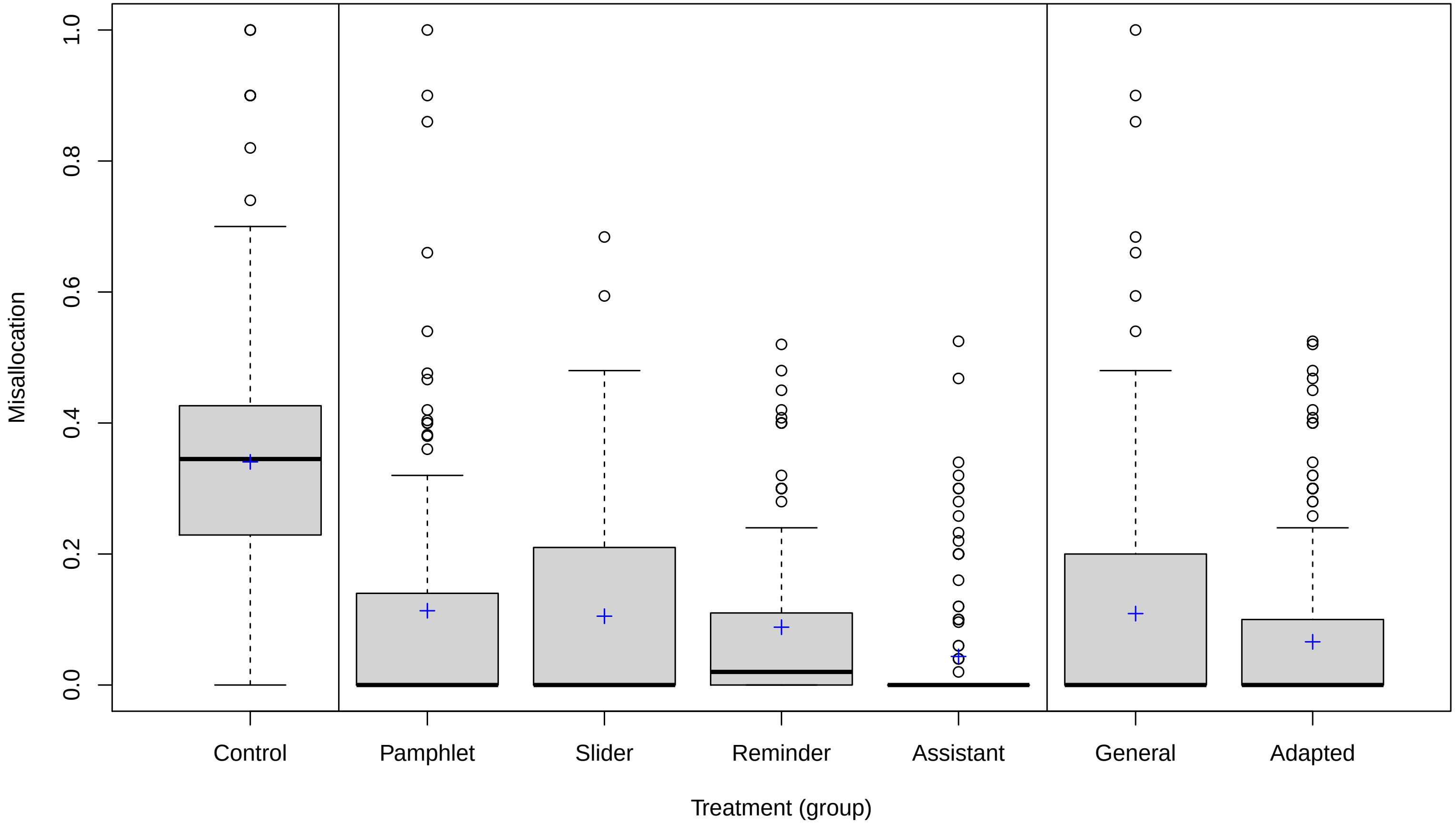


Figure Appendix III.7: This figure shows boxplots of the mean misallocation of all participants split by treatment and intervention group. The thick black line within each boxplot denotes the median, the blue "+"-symbol indicates the mean. The maximum length of the whiskers is 1.5 times the interquartile 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 highest interest rate 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.

Why is that?

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.

An example:

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:

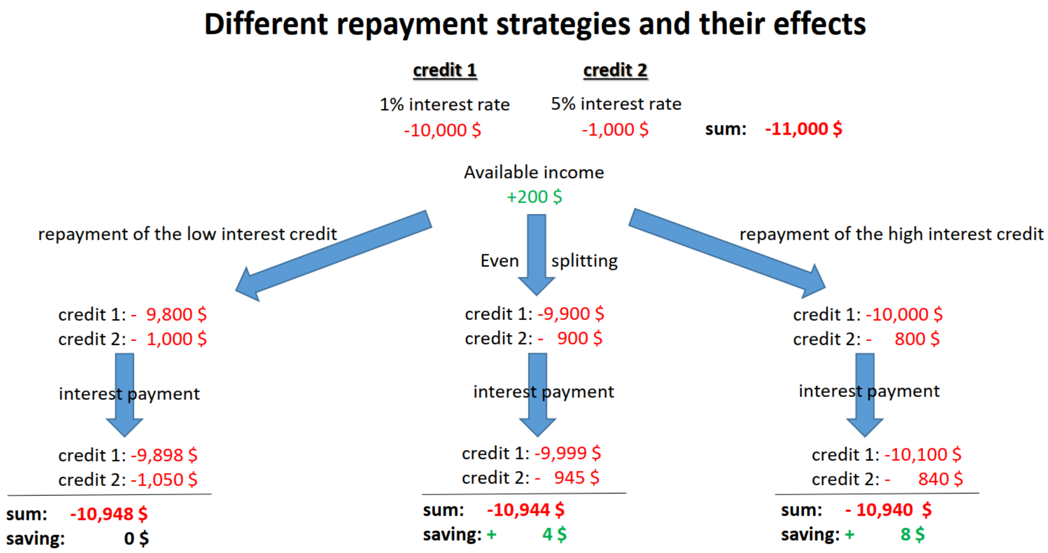
Costly repayment 1: Inclusion of debt levels

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.

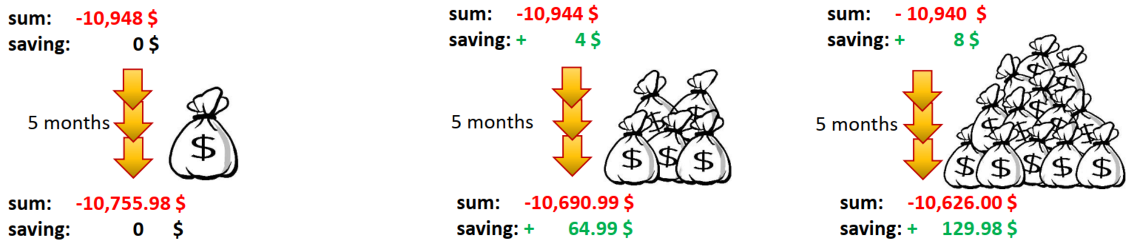


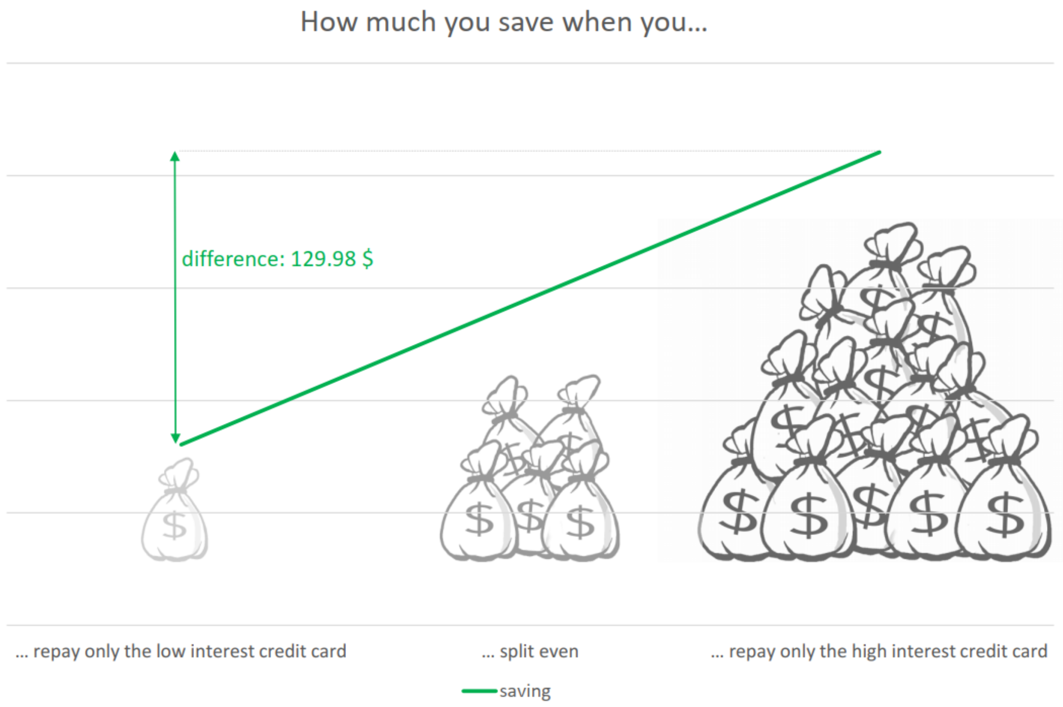
The example on the left shows what happens if we repay the highest debt first: we use the full amount of 200$ to repay part of the debt on the 10,000$-card and we leave the 1,000$-card untouched. If we calculate with the interest, 10,948$ are left after that month.

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

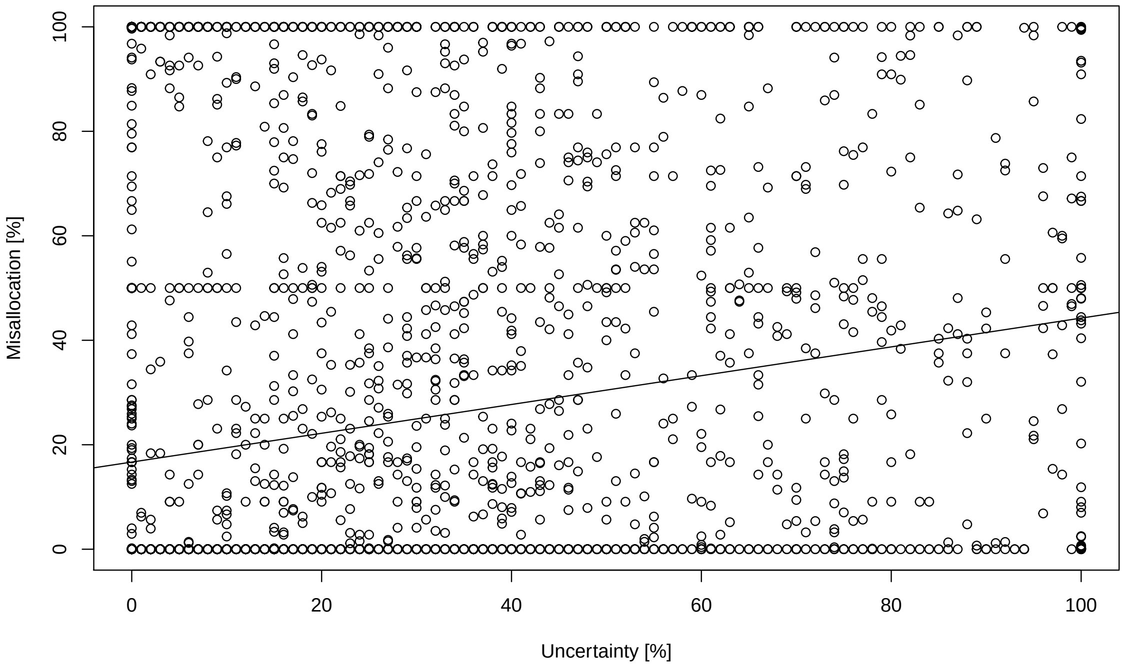


Figure Appendix IV.8: This scatter plot shows the percentage points of uncertainty (x-axis) in relation to misallocation (y-axis). The additional OLS-regression line shows that uncertainty and misallocation correlate positively.

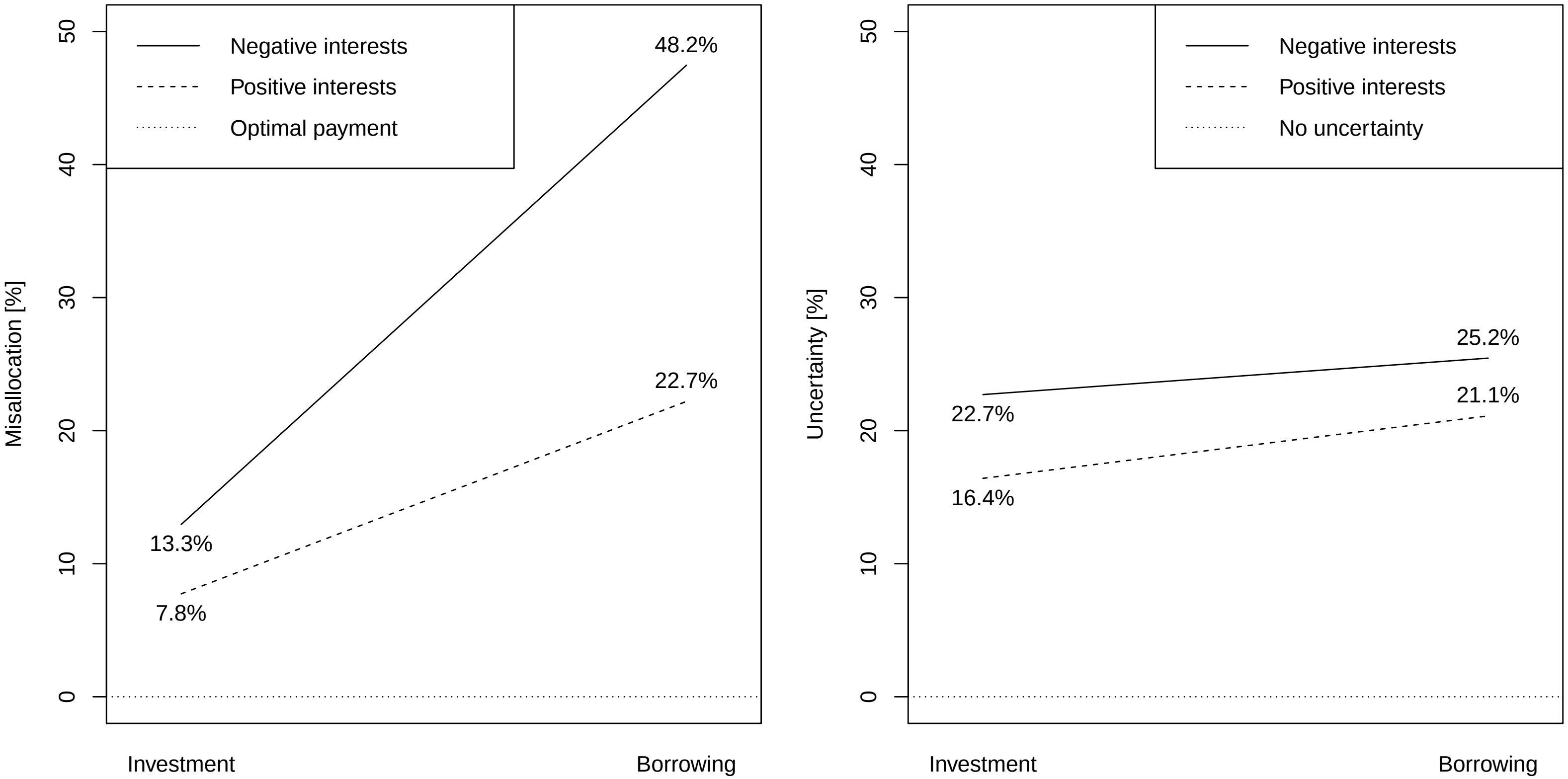


Figure Appendix IV.9: These interaction plots show the average values for misallocation (left graphic) and uncertainty (right graphic) in experiment #1 split by treatment. We differentiate by the borrowing variable (x-axis in both graphics) and the negative int. rates variable (line types in both graphics). 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.31: 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.32: 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.33: 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						
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.34: 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

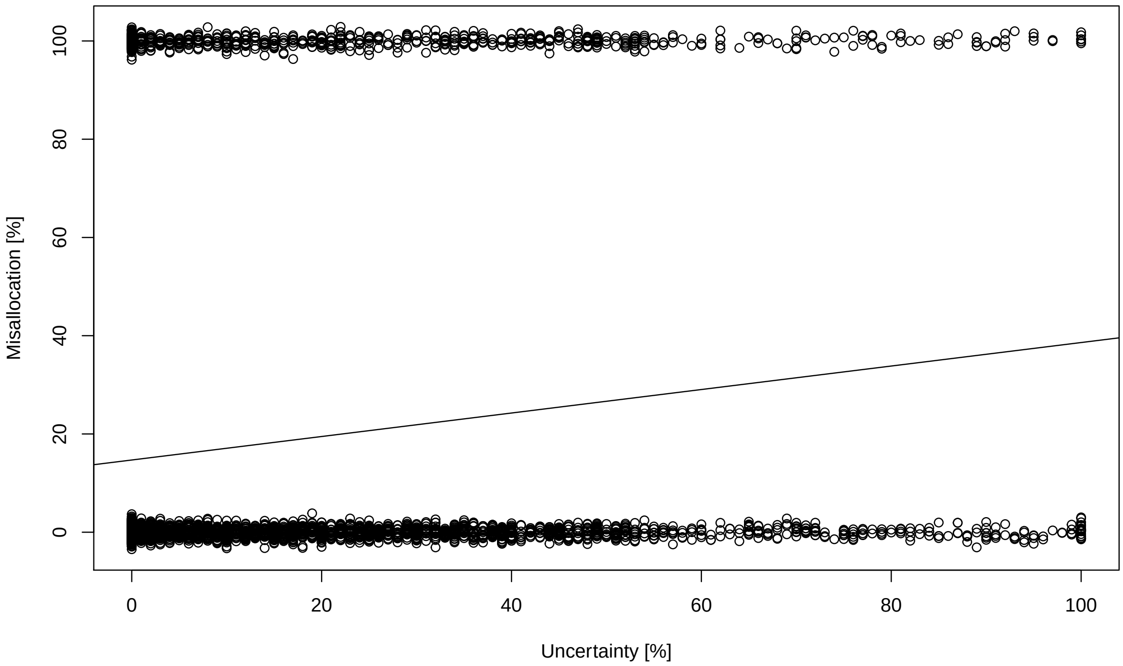


Figure Appendix IV.10: This scatter plot shows the percentage points of uncertainty (x-axis) in relation to misallocation (y-axis). For a better overview the points are jittered randomly in y-direction (although they can only take on the value 0 or 100). The additional OLS-regression line shows that uncertainty and misallocation correlate positively.

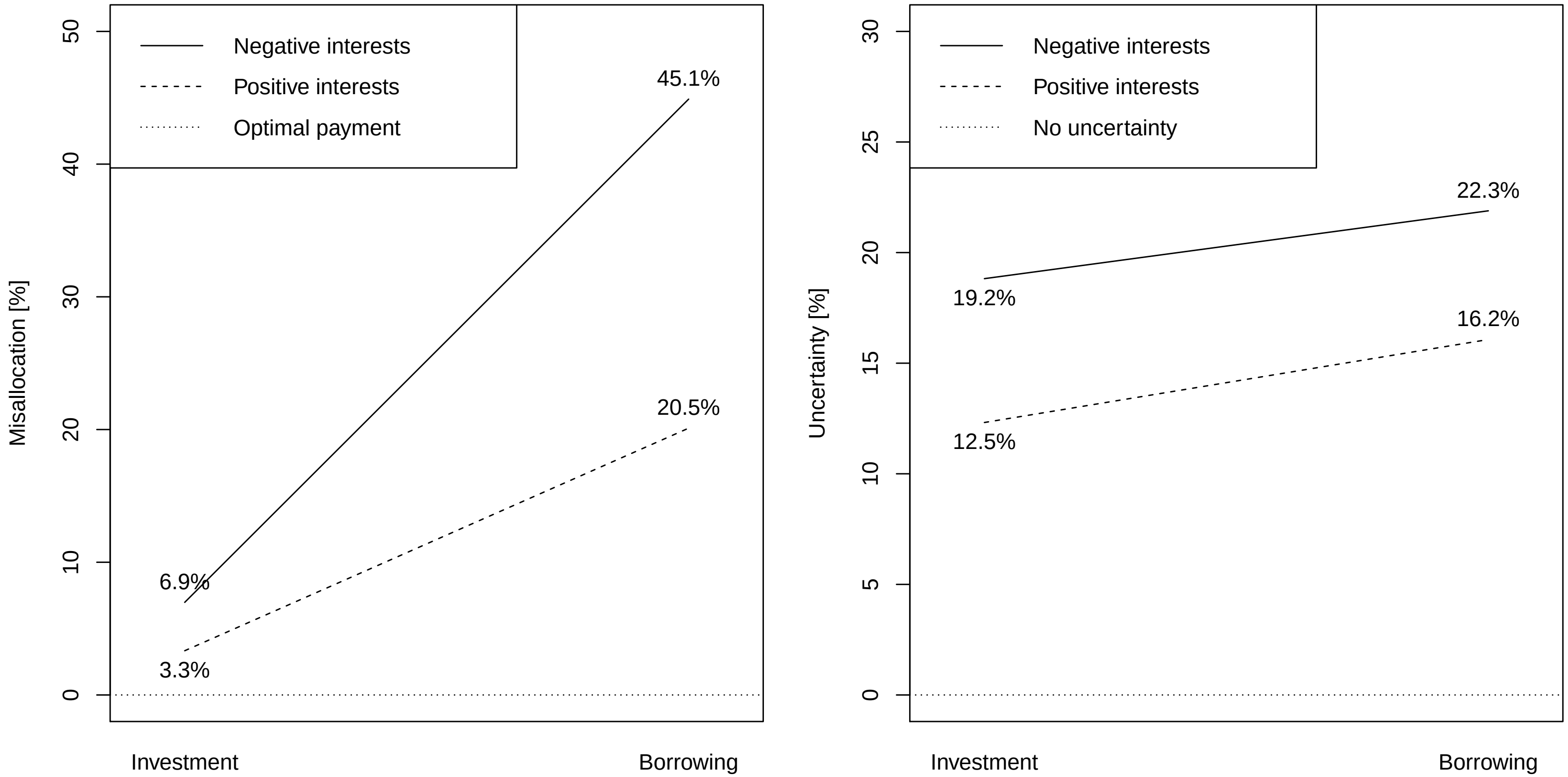


Figure Appendix IV.11: These interaction plots show the average values for misallocation (left graphic) and uncertainty (right graphic) in experiment #2 split by treatment. We differentiate by the borrowing variable (x-axis in both graphics) and the negative int. rates variable (line types in both graphics). 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.35: 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.36: 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						
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.37: 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.38: 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.39: 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.40: 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).

Preregistration: Elemental Financial Decisions

Florian Gärtner* Darwin Semmler†

January 28, 2022

We experimentally investigate elemental investment and borrowing decisions. Examples (in spirit, not in wording or numbers) of such decisions are:

- Example 1: You can invest $250. Do you prefer to invest in a (safe) asset with 5% returns or in a (safe) asset with 10% returns, assuming that all else is equal and that there is no omitted or hidden relevant information?
- Example 2: You need to borrow $500. Do you prefer to borrow for a 15% interest rate or a 17% rate, assuming that all else is equal and that there is no omitted or hidden relevant information?

We design two very similar experiments. In both experiments, each participant has to make 16 decisions. In each decision, they distribute an amount of money over two accounts, similar to the examples above. These decisions differ in the following independent variables:

- Borrowing vs. investing
- Negative vs. positive interest rates
- Framing (either presenting the interest rates, or the already calculated returns/interests)

This design results in 8 combinations; we repeat each combination twice with randomized values for interest rates and the sums of money, which leads to 16 decisions. The order of these 16 decisions is randomized as well. For each decision, we measure the cognitive uncertainty of the participant with the question "On a scale between 0 and 100, how certain are you that your decision maximizes your outcome?" The participants indicate their cognitive uncertainty using a slider; 0 is the maximum uncertainty while 100 is perfect certainty.

*Justus-Liebig-University Giessen, Licher Str. 66, 35394 Giessen, Germany, Phone: +49 641 99 22595, E-mail: Florian.Gärtner@wirtschaft.uni-giessen.de

†Justus-Liebig-University Giessen, Licher Str. 66, 35394 Giessen, Germany, Phone: +49 641 99 22594, E-mail: Darwin.Semmler@wirtschaft.uni-giessen.de

The two experiments themselves differ from each other with respect to the divisibility of money. In experiment #1, the money about which to decide is freely distributable. In experiment #2, the participants have to invest all their money in only one asset or repay only one credit; they cannot split the money in any way. We restrict participants to exactly one of these two experiments.

After the experimental stage, we measure the following variables in a post experimental questionnaire (PEQ):

- Experience with credit cards and investing
- Financial literacy
- Preference for numerical information
- Numeracy
- Consumer confidence
- Risk affinity
- Gender
- Age
- Years of education

In our analysis, we focus on explaining misallocation, i.e. money *not* invested on the high return asset or borrowed from the low interest rate credit. We investigate if misallocation depends on our four treatment variables, and if these effects, assuming they exist, are plausibly mediated by cognitive uncertainty. We also correlate misallocation with the aforementioned variables from the PEQ.

We run these experiments on the internet platform Prolific. We aim for 240 valid participants per experiment (480 in total). We define participants as valid if they pass both of the following data quality measures:

We ask three attention check questions and reject any participant who fails at least two. The screened out participants do not count to the observations.

After the experimental stage, we ask an open question where participants need to briefly describe their strategy for the last decision they made. Two raters independently rate whether these answers are meaningful answers to that question, i.e. if these answer fit the question at all, no matter which strategy is actually described (or if the description matches actual behavior). We screen out participants where both raters agree that the answers are nonsensical. The screened out participants do not count to the observations.

We implement the following analyses:

- Main analysis: Analyses that include all participants who pass the data quality measures
- Robustness check: Analyses that exclude participants who do not fail any of the three attention test questions (point 1)
- Robustness check: Analyses that exclude the outer 2.5% of the response time distribution

We do not rule out additional analyses, but our main analysis and interpretation will focus on the 480 participants that pass the aforementioned data quality measures.

Appendix V (to Chapter V)

Variables

Variable	Description
# employees	Numbers of employees in 1,000 [K].
# Prior Females	Number of women in a board prior to an appointment.
# Prior Foreigns	Placeholder for one of the two dummy variables # Prior non-Frenchs and # Prior non-whites.
# Prior non-Frenchs	Number of non-French board members of a firm prior to an appointment.
# Prior non-whites	Number of non-white board members of a firm prior to an appointment.
Academic	Dummy variable equal to one for an appointment of a member holding a doctorate, and zero otherwise.
Age	The age of an appointee at the date of appointment.
Boardtype	Categorical variable for the relevant board of an appointment. One of either unitary board, supervisory board or management board.
CEO	Categorical variable measuring whether an appointee was a CEO at appointment, a CEO at a later point after the appointment (CEOLate) or no CEO.

`CMAR`	Cumulative market adjusted return per month, difference between CRR and monthly market index.
`CRAR`	Cumulative risk adjusted return per month, difference between CRR and expected stock return via capital asset pricing mode.
`CRR`	Cumulative raw return per month.
`DE`	Debt-to-equity ratio.
`EBIT`	Earnings before interest and taxes in billion [B].
`Extern`	Dummy variable equal to one for an external appointment, and zero otherwise.
`Female`	Dummy variable equal to one if an appointee is female, and zero otherwise.
`Firm performance`	Placeholder for one of the firm performance variables CRR, CRAR, CMAR or beta.
`Foreign`	Placeholder for one of the two dummy variables non-white and non-French.
`Grandeecole`	Dummy variable equal to one for an appointment of a member who visited a Grande Ecole (a French elite school), and zero otherwise.
`Industry`	The industry of a firm in one of the categories Basic Materials, Consumer Cyclicals, Consumer Non-Cyclicals, Energy, Financials, Healthcare, Industrials, Real Estate, Technology or Utilities.

`Jointappointment`	Dummy variable equal to one if there were other appointments in the same firm in a two-day time window of the observed appointment, and zero otherwise.
`month`	Time variable measuring the month relative to an appointment. Month equals zero in the month of appointment and varies between -24 and 24 for two years before and after an appointment.
`non-French`	Dummy variable equal to one if the nationality of an appointee is not French, and zero otherwise.
`non-white`	Dummy variable equal to one if the appearance of an appointee is not Caucasian or Hispanic (subjective measure), and zero otherwise.
`revenue`	Natural logarithm of revenue.
`post2010`	Dummy variable equal to one for an observation since 1.1.2011, and zero for an observation before 1.1.2011 to distinguish effects before and after the introduction of the women's quota.
`Prior Boardsize`	Number of board members of a firm prior to an appointment.
`ROA`	Annual return on Assets.
`ROE`	Annual Return on Equity.
`Share of non-Frenchs`	Share of board members who do not posses French citizenship.

`Share of non-whites`	Share of board members who are not Caucasian or Hispanic in appearance.
`Share of women`	Share of female board members.
`TQ`	Tobin's Q.

Additional tables

Table Appendix V.43: Summary statistics for appointments per firm

	Number of appointments							Summary statistics Age at appointment						
Company	Appointm.	F.Eth	F.Nat	Extern	Fem.	Acad.	Gr.Ecole	Min	%25	Median	%75	Max	Mean	NAs
Accor SA*	34	7	7	29	7	6	25	36.81	45.39	49.71	54.87	66.72	49.73	0
AGF	7	0	3	5	0	1	2	39.84	50.62	55.17	57.89	64.58	53.80	0
Air France KLM	0	No appointments during CAC40 listing												
Air Liquide SA*	17	1	9	15	6	6	6	41.99	54.24	58.71	60.48	64.24	56.80	0
Airbus SE*	27	1	18	23	5	11	10	42.13	50.63	57.10	61.82	70.94	56.13	0
Alcatel	22	1	16	21	7	5	7	45.87	55.93	59.89	61.46	72.12	58.63	0
Alstom	11	2	6	11	6	1	6	45.34	48.75	49	60.45	64.05	53.73	2
Arcelor	9	1	6	2	1	1	1	24.42	42.82	57.39	62.78	69.86	53.50	0
ArcelorMittal SA	9	1	8	7	2	1	3	51.57	54.85	55.46	59.87	61.84	56.68	0
Atos SE	1	0	0	1	1	0	0	54.89	54.89	54.89	54.89	54.89	54.89	0
AXA SA*	29	1	13	23	8	5	13	43.39	50.22	54.55	57.57	61.44	53.84	11
BNP Paribas SA*	26	0	10	16	12	9	12	46.73	51.69	57.81	61.44	69.03	56.93	3
Bouygues SA*	22	0	1	9	7	2	16	30.22	40.96	53.03	59.69	66.21	50.26	0
Capgemini SE*	17	0	4	14	6	4	13	46.87	53.88	58.87	60.27	63.09	56.76	4
Carrefour SA*	40	3	10	27	9	4	19	39.87	48.33	55.45	63.44	79.39	55.55	1
Casino Guichard	3	0	0	3	0	0	3	47.61	50.17	52.73	53.82	54.91	51.75	0
Compagnie de Saint Gobain SA*	26	3	8	19	7	4	14	44.81	50.83	57.68	60.59	68.41	56.03	12
Credit Agricole*	61	1	3	14	10	2	19	45.57	54.42	56.65	60.71	64.05	56.06	40
Danone*	27	5	11	20	10	5	13	38.25	48.91	54.77	60.23	67.52	54.28	0
Dexia	35	0	16	28	5	5	17	43.38	46.88	53.69	58.97	68.77	54.32	16
EDF	35	0	1	25	7	3	29	36.35	50.73	55.55	60.95	68.56	54.52	1
Engie-GDF	34	0	4	26	13	3	23	35.62	45.42	50.06	59.03	66.51	51.60	0
Essilor	25	4	8	15	11	1	12	40.38	44.53	52.61	61.95	67.66	53.12	1
Gemalto	1	1	0	1	1	1	0	51.89	51.89	51.89	51.89	51.89	51.89	0
Hermes	0	No appointments during CAC40 listing												
Kering*	17	1	6	12	9	4	5	33.29	42.90	47.52	52.47	60.73	48.25	0
Klepierre	2	0	0	2	2	0	2	43.80	47.01	50.22	53.42	56.63	50.22	0
L'Oreal*	21	0	6	17	8	1	12	25.97	49.04	53.71	57.60	64.72	52.78	0
Lafargeholcim*	35	2	20	29	6	11	12	41.31	47.50	55.87	60.37	72.90	54.60	1
Lagardère	16	0	3	10	2	2	11	31.65	55.10	59.29	63.20	73.98	58.25	0
Legrand	5	1	3	5	4	0	2	41.90	45.13	47.90	48.01	55.90	47.77	0
LVMH Louis Vuitton*	16	0	4	11	6	2	11	28.44	46.46	55.64	61.61	76.91	53.24	0
Michelin*	15	1	3	13	5	1	6	45.70	50.22	57.42	63.68	68.66	56.85	5
Natixis	2	0	0	1	1	0	2	47.02	48.90	50.79	52.68	54.56	50.79	0
Nokia	6	0	5	6	2	0	1	49.96	58.57	61.92	65.67	68.54	61.12	0
Orange*	49	2	1	32	12	4	33	30.96	46.19	52.18	57.22	75.05	51.83	0
Pernod Ricard SA	18	1	8	13	6	3	8	30.59	43.42	48.24	59.53	70.34	50.12	0
Peugeot	28	2	5	12	6	2	23	38	46.91	52.80	56.82	71.43	52.89	0
Publicis Groupe	12	1	8	7	2	2	2	45.69	50.15	54.20	56.47	66.58	53.71	2
Renault*	44	9	11	33	11	4	23	35.22	47.92	53.77	60.85	67.69	54.02	0
Safran	16	0	0	5	6	2	12	40.96	54.81	57.12	60.77	63.88	56.45	0
Sanofi*	31	1	15	22	8	13	11	34.38	52.81	59.26	64.22	69.85	57.86	0
Schneider Electric*	29	5	13	23	9	2	15	40.84	47.81	54.09	58.11	63.65	53.37	0
Societe Generale*	33	0	9	26	11	6	15	36.96	51.57	55.94	59.23	66.13	54.53	3
Sodexo	3	0	1	3	1	0	1	53.77	56.86	59.96	60.61	61.26	58.33	0
Solvay	8	0	6	8	5	2	1	48.96	57.61	58.52	60.06	62.85	58.11	0
STMicroelectronics	15	0	9	13	2	3	6	37.81	43.81	54.66	59.52	63.28	52.00	0
Suez	6	0	2	5	0	1	1	50.61	54.76	55.99	59.70	61.87	56.62	0
Suez Environnement	2	0	1	1	1	0	0	44.97	46.48	47.98	49.48	50.99	47.98	0
TechnipFMC	16	1	12	14	7	0	3	43.82	50.02	57.35	60.87	72.55	56.67	0
TF1	3	0	1	2	1	0	1	31.07	42.81	54.55	56.54	58.52	48.05	0
Thales	9	0	1	7	0	0	8	51.35	55.39	60.62	63.72	64.96	59.55	0
Thomson	19	1	4	12	2	2	12	38.18	45.42	59.07	61.13	75.90	55.41	7
Total*	23	0	10	15	7	4	12	32.50	52.81	57.21	60.01	66.69	56.18	0
Valeo	6	0	1	5	3	1	4	42.78	49.90	55.85	60.95	64.26	54.88	0
Vallourec	12	0	4	11	4	2	7	41.56	50.80	51.61	56.42	65.69	53.31	0
Veolia Environnement*	30	3	6	22	7	6	19	41.45	49.58	53.31	59.74	66.36	54.03	1
Vinci	25	5	10	17	7	3	13	34.18	51.88	55.96	61.04	65.19	55.27	0
Vivendi*	56	7	15	35	12	9	28	36.27	47.74	53.15	60.70	67.75	53.79	0
WFD Unibail Rodamco NV	37	2	18	28	9	3	18	39.51	47.13	51.98	58.15	73.31	53.92	0

[*] The firm was listed in the CAC40 for the complete time span of observation (01.01.2002 - 31.12.2018).

Table Appendix V.44: Summary statistics of appointments after matching for appearance

Appearance		Nationality		CEOs		Age	
Caucas.	56	FR	50	CEO appointments	0	N	120
Hispanic	4	US	10	Later CEOs	0	Mean	51.910
Asian	21	DE	6			Std. Dev.	8.746
Near-East	27	IN	5	Boardtype		Min.	24.418
African	2	JP	4	Unitary board	92	Median	50.518
Mixed	10	QAT	4	Supervisory board	20	Max.	69.573
		Other	41	Management board	8		
NA	0	NA	0	NA	0	NA	0

Binary variables

Extern		Female		Academic		Grande Ecole	
N	120	N	120	N	120	N	120
Yes	99	Yes	50	Yes	19	Yes	30
No	21	No	70	No	101	No	90

Table Appendix V.45: Summary statistics of appointments after matching for citizenship

Appearance		Nationality		CEOs		Age	
Caucas.	465	FR	280	CEO appointments	14	N	560
Hispanic	33	US	50	Later CEOs	9	Mean	55.308
Asian	21	DE	41			Std. Dev.	7.858
Near-East	21	UK	33	Boardtype		Min.	24.418
African	1	ES	21	Unitary board	426	Median	56.367
Mixed	8	BE	19	Supervisory board	101	Max.	79.385
		Other	116	Management board	33		
NA	11	NA	0	NA	0	NA	0

Binary variables

Extern		Female		Academic		Grande Ecole	
N	560	N	560	N	560	N	560
Yes	429	Yes	184	Yes	113	Yes	159
No	131	No	376	No	447	No	401

Robustness check tables

Table Appendix V.46: Panel regression of firm performance without control variables[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:		Before				After		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-white								
non-white	-0.260*	-0.185	-0.490	0.003	-0.103	0.002	0.034	0.017
	(0.119)	(0.103)	(0.364)	(0.022)	(0.116)	(0.089)	(0.344)	(0.023)
month	0.002	-0.000	0.001	0.000	-0.007*	-0.005**	-0.017	0.000
	(0.004)	(0.002)	(0.008)	(0.000)	(0.003)	(0.002)	(0.009)	(0.000)
constant	0.570***	0.116	1.143***	0.414***	0.559***	0.025	1.103***	0.402***
	(0.094)	(0.071)	(0.269)	(0.017)	(0.107)	(0.075)	(0.285)	(0.017)
Observations	2821	2824	2660	2821	2832	2833	2698	2832
non-French								
non-French	0.069	0.046	0.221	0.007	0.056	0.013	0.278	0.003
	(0.128)	(0.066)	(0.202)	(0.011)	(0.069)	(0.055)	(0.194)	(0.011)
month	-0.012***	-0.003*	-0.006	-0.000	-0.007***	-0.002	-0.012	0.001**
	(0.003)	(0.001)	(0.006)	(0.000)	(0.002)	(0.001)	(0.006)	(0.000)
constant	0.346***	-0.185***	0.803***	0.396***	0.410***	-0.131**	1.161***	0.389***
	(0.091)	(0.048)	(0.149)	(0.008)	(0.058)	(0.042)	(0.156)	(0.008)
Observations	12725	12755	11350	12725	13000	13034	11699	13000

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on the propensity score matched data set for appearance or nationality, respectively. Robust standard errors are reported in parentheses.

Table Appendix V.47: Panel regression of firm performance with alternative measures (non-white)[a]

Dependent variable:	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
Modus:		Before			After	
	(1)	(2)	(3)	(4)	(5)	(6)
non-white	-0.018	-0.034	-0.112	-0.007	-0.007	0.103
	(0.009)	(0.021)	(0.426)	(0.008)	(0.022)	(0.709)
month	-0.000	-0.001	0.021	-0.000*	-0.001*	0.027
	(0.000)	(0.000)	(0.018)	(0.000)	(0.000)	(0.022)
non-white × month	0.000	-0.000	-0.015	0.000	0.000	-0.022
	(0.000)	(0.000)	(0.023)	(0.000)	(0.001)	(0.021)
Appointment-specific control variables						
Female	0.004	-0.002	0.441	-0.003	0.005	0.879
	(0.011)	(0.027)	(0.452)	(0.009)	(0.027)	(1.066)
non-white × Female	-0.014	-0.038	0.097	-0.009	-0.053	-0.355
	(0.013)	(0.036)	(0.570)	(0.012)	(0.035)	(1.212)
Extern	-0.006	-0.006	-0.279	-0.005	-0.004	-0.306
	(0.014)	(0.030)	(0.301)	(0.010)	(0.026)	(0.612)
Academic	-0.021*	-0.067**	0.078	-0.001	-0.021	0.457
	(0.010)	(0.024)	(0.314)	(0.008)	(0.023)	(0.560)
Grandeecole	0.020*	0.046*	-0.488	0.017**	0.036*	-0.808
	(0.009)	(0.023)	(0.260)	(0.006)	(0.017)	(0.489)
Age	-0.001	0.001	-0.007	-0.001*	-0.000	-0.034
	(0.000)	(0.001)	(0.016)	(0.000)	(0.001)	(0.025)
Jointappointment	0.003	-0.041	-0.711*	-0.002	-0.027	-0.559
	(0.010)	(0.025)	(0.281)	(0.008)	(0.018)	(0.365)
Supervisory board	-0.009	-0.042	1.058*	-0.007	-0.040	2.280
	(0.010)	(0.025)	(0.530)	(0.011)	(0.024)	(1.420)
Management board	-0.035	-0.096*	2.369**	-0.022	-0.077*	3.192**
	(0.018)	(0.039)	(0.733)	(0.013)	(0.030)	(0.986)
Prior Boardsize	-0.003*	-0.007	0.013	-0.002	-0.008	0.020
	(0.001)	(0.004)	(0.056)	(0.002)	(0.006)	(0.091)
# Prior Females	-0.002	-0.002	-0.189*	-0.002	0.000	-0.237*
	(0.002)	(0.006)	(0.076)	(0.002)	(0.006)	(0.108)
# Prior non-whites	-0.001	-0.018*	-0.071	0.002	-0.015	0.179
	(0.003)	(0.008)	(0.125)	(0.003)	(0.008)	(0.210)
Firm-specific control variables						
# employees	-0.000*	-0.000**	0.005	-0.000*	-0.000***	0.011***
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)	(0.002)
ln(revenue)	-0.009	-0.072**	1.035*	0.013	0.015	1.426*
	(0.007)	(0.025)	(0.501)	(0.007)	(0.017)	(0.615)
EBIT	0.017***	0.058***	-0.457	0.001	0.026	-1.089
	(0.002)	(0.008)	(0.265)	(0.009)	(0.019)	(0.587)
DE	0.000	0.000*	0.004	-0.000*	-0.000	0.007
	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)	(0.004)
Consumer Cyclicals	-0.022	0.085*	0.199	-0.021	0.058	-0.763
	(0.026)	(0.041)	(0.517)	(0.029)	(0.045)	(0.678)
Consumer Non-Cyclicals	-0.030	0.009	-0.884	-0.010	0.021	-2.041**
	(0.027)	(0.035)	(0.460)	(0.028)	(0.048)	(0.678)
Healthcare	-0.009	-0.098	1.414	0.050	-0.039	5.006
	(0.067)	(0.129)	(1.836)	(0.093)	(0.178)	(4.462)
Industrials	-0.016	0.102**	-0.033	0.001	0.081	-0.949
	(0.028)	(0.034)	(0.383)	(0.038)	(0.053)	(0.604)
Real Estate	-0.061*	-0.169**	1.202	-0.019	0.001	0.411
	(0.029)	(0.056)	(0.987)	(0.030)	(0.047)	(0.911)
Technology	-0.052	-0.019	1.064	-0.031	-0.044	0.064
	(0.029)	(0.041)	(0.830)	(0.046)	(0.067)	(1.100)
Utilities	-0.010	0.154**	3.219**	-0.027	0.062	0.572
	(0.027)	(0.051)	(1.212)	(0.037)	(0.068)	(1.914)
constant	0.366*	1.790***	-21.041*	-0.079	-0.058	-28.173*
	(0.148)	(0.543)	(10.707)	(0.137)	(0.342)	(12.702)
Observations	1361	1350	2665	1454	1439	2788

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

[a] The models (1)-(3) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (4)-(6) repeat the panel regressions with data 24 months after an appointment. All models are based on the propensity score matched data set for appearance. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board) and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.48: Panel regression of firm performance with alternative measures (non-French)[a]

Dependent variable:	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
Modus:		Before			After	
	(1)	(2)	(3)	(4)	(5)	(6)
non-French	0.001	0.007	-1.642*	0.002	0.000	-1.168
	(0.008)	(0.017)	(0.645)	(0.006)	(0.015)	(0.807)
month	-0.000	-0.001**	0.084***	-0.001***	-0.001***	-0.031
	(0.000)	(0.000)	(0.025)	(0.000)	(0.000)	(0.031)
non-French × month	0.000	0.000	-0.084**	0.000	-0.000	0.077*
	(0.000)	(0.000)	(0.028)	(0.000)	(0.000)	(0.035)
Appointment-specific control variables						
Female	-0.009	-0.016	1.577*	-0.001	-0.017	1.998*
	(0.007)	(0.018)	(0.771)	(0.005)	(0.012)	(0.905)
non-French × Female	-0.002	-0.028	-0.803	-0.006	-0.011	-0.632
	(0.011)	(0.024)	(0.930)	(0.007)	(0.017)	(1.078)
Extern	0.013	0.028	0.896	0.001	0.013	1.221
	(0.007)	(0.017)	(0.734)	(0.005)	(0.013)	(0.949)
Academic	-0.013	-0.015	0.421	-0.003	0.002	-0.344
	(0.008)	(0.016)	(0.500)	(0.005)	(0.011)	(0.544)
Grandeecole	-0.004	-0.004	-0.203	-0.002	0.004	-0.659
	(0.008)	(0.016)	(0.621)	(0.005)	(0.013)	(0.759)
Age	0.000	0.000	0.016	0.000	0.001	0.009
	(0.000)	(0.001)	(0.024)	(0.000)	(0.001)	(0.028)
Jointappointment	-0.017**	-0.049***	0.530	-0.014**	-0.039***	0.394
	(0.007)	(0.015)	(0.485)	(0.005)	(0.011)	(0.516)
Supervisory board	0.036***	0.018	-1.276*	0.003	-0.018	-1.185
	(0.010)	(0.022)	(0.537)	(0.007)	(0.017)	(0.621)
Management board	0.030	0.012	0.814	0.004	-0.021	0.116
	(0.016)	(0.033)	(0.813)	(0.008)	(0.021)	(0.893)
CEO	-0.021	-0.011	-0.794	-0.012	0.011	-0.288
	(0.019)	(0.045)	(1.077)	(0.009)	(0.035)	(1.310)
Later CEO	-0.013	-0.013	0.696	-0.010	-0.014	1.489
	(0.015)	(0.033)	(0.798)	(0.013)	(0.038)	(1.015)
Prior Boardsize	-0.002	0.001	-0.128	0.000	0.002	-0.126
	(0.001)	(0.002)	(0.084)	(0.001)	(0.001)	(0.099)
# Prior Females	-0.006***	-0.019***	0.716***	-0.006***	-0.014***	0.582***
	(0.001)	(0.003)	(0.122)	(0.001)	(0.002)	(0.131)
# Prior non-Frenchs	-0.005***	-0.019***	0.033	-0.003**	-0.015***	-0.093
	(0.001)	(0.003)	(0.107)	(0.001)	(0.002)	(0.149)
Firm-specific control variables						
# employees	-0.000***	-0.000***	0.017***	-0.000**	-0.000***	0.016***
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)	(0.002)
ln(revenue)	0.017	0.012	-0.577	0.009*	0.012	0.352
	(0.009)	(0.021)	(0.390)	(0.004)	(0.010)	(0.330)
EBIT	0.006***	0.015***	-0.213*	0.004**	0.014***	-0.311**
	(0.001)	(0.002)	(0.087)	(0.001)	(0.002)	(0.117)
DE	-0.000*	0.000	0.007**	-0.000***	-0.000	0.014**
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)	(0.005)
Consumer Cyclicals	-0.087***	-0.199***	0.863	-0.051***	-0.123***	0.419
	(0.018)	(0.039)	(0.608)	(0.012)	(0.024)	(0.756)
Consumer Non-Cyclicals	-0.020	-0.111***	-2.048**	-0.009	-0.073***	-3.294***
	(0.015)	(0.028)	(0.643)	(0.012)	(0.020)	(0.877)
Energy	-0.063***	-0.182***	2.872***	-0.056***	-0.228***	3.887**
	(0.016)	(0.038)	(0.822)	(0.017)	(0.034)	(1.500)
Financials	-0.135***	-0.213***	27.399***	-0.082***	-0.126***	29.409***
	(0.015)	(0.032)	(1.598)	(0.014)	(0.026)	(1.713)
Healthcare	0.011	-0.141***	1.207	-0.002	-0.147***	1.277
	(0.022)	(0.042)	(0.679)	(0.015)	(0.028)	(0.972)
Industrials	-0.061***	-0.102**	0.675	-0.020	-0.011	-0.516
	(0.017)	(0.037)	(0.583)	(0.013)	(0.024)	(0.705)
Real Estate	-0.073**	-0.241***	-0.040	-0.051***	-0.167***	2.350**
	(0.023)	(0.050)	(0.859)	(0.014)	(0.028)	(0.882)
Technology	-0.101***	-0.186***	1.197*	-0.061***	-0.145***	0.641
	(0.020)	(0.043)	(0.528)	(0.016)	(0.029)	(0.607)
Utilities	-0.095***	-0.194***	1.588*	-0.050***	-0.085**	-0.049
	(0.018)	(0.041)	(0.713)	(0.014)	(0.033)	(1.042)
constant	-0.219	0.149	12.264	-0.054	0.075	-8.841
	(0.188)	(0.437)	(8.485)	(0.094)	(0.197)	(7.492)
Observations	5434	5260	12513	5442	5280	12632

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

[a] The models (1)-(3) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (4)-(6) repeat the panel regressions with data 24 months after an appointment. All models are based on the propensity score matched data set for nationality. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board), no-CEO-appointment and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.49: Panel regression of firm performance (non-white), only data before 2011[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:	Before				After			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-white	-0.034	-0.077	1.325	0.026	0.320	0.055	1.854**	0.031
	(0.283)	(0.179)	(0.803)	(0.042)	(0.236)	(0.211)	(0.687)	(0.030)
month	-0.016	-0.009	-0.066*	-0.001	-0.007	0.003	-0.005	0.001
	(0.013)	(0.007)	(0.031)	(0.001)	(0.005)	(0.004)	(0.017)	(0.001)
non-white × month	0.025	0.013	0.070	0.005*	-0.011	-0.010	-0.014	0.000
	(0.020)	(0.012)	(0.043)	(0.002)	(0.010)	(0.008)	(0.034)	(0.002)
Appointment-specific control variables								
Female	0.217	0.102	1.406	0.022	-0.022	0.152	0.941	0.075*
	(0.274)	(0.259)	(1.387)	(0.052)	(0.200)	(0.212)	(1.064)	(0.038)
non-white × Female	1.339***	0.446	-5.114*	-0.049	0.133	0.054	-6.847***	-0.072
	(0.326)	(0.273)	(2.507)	(0.056)	(0.274)	(0.257)	(1.521)	(0.056)
Extern	0.481*	0.025	4.778*	-0.168***	0.439	-0.117	7.247***	-0.066
	(0.242)	(0.300)	(2.076)	(0.048)	(0.294)	(0.190)	(1.033)	(0.041)
Academic	-0.767*	-0.937***	-0.053	-0.022	-0.911**	-1.104***	0.655	-0.069
	(0.342)	(0.168)	(0.968)	(0.061)	(0.353)	(0.219)	(0.819)	(0.044)
Grandeecole	-0.125	-0.015	-0.129	-0.032	0.107	0.251*	1.646**	-0.017
	(0.154)	(0.152)	(0.948)	(0.028)	(0.155)	(0.121)	(0.637)	(0.023)
Age	-0.001	0.013	-0.122	0.005	0.008	0.015	-0.302***	0.005***
	(0.014)	(0.012)	(0.078)	(0.003)	(0.016)	(0.010)	(0.038)	(0.002)
Jointappointment	0.258	0.215	-0.516	0.030	-0.222	0.163	1.174*	0.020
	(0.236)	(0.201)	(0.786)	(0.045)	(0.199)	(0.123)	(0.475)	(0.028)
Supervisory board	-0.921**	-0.986***	-0.779	0.020	-0.721*	-0.261	-5.130***	0.088**
	(0.302)	(0.218)	(1.792)	(0.064)	(0.306)	(0.171)	(1.184)	(0.032)
Management board	-1.304**	-1.503***	2.825*	-0.166	-0.595	-0.865**	2.943**	-0.021
	(0.469)	(0.411)	(1.246)	(0.089)	(0.527)	(0.265)	(1.139)	(0.061)
Prior Boardsize	-0.058	0.049	-0.461***	0.013	0.115*	0.110***	-0.377***	0.012*
	(0.036)	(0.025)	(0.116)	(0.008)	(0.045)	(0.028)	(0.106)	(0.005)
# Prior Females	0.414***	-0.027	1.361***	-0.054***	-0.093	-0.211*	-0.390	0.011
	(0.088)	(0.066)	(0.331)	(0.013)	(0.086)	(0.084)	(0.355)	(0.016)
# Prior non-whites	-0.111	-0.136	0.757*	-0.045	-0.078	-0.094	0.979**	0.013
	(0.147)	(0.113)	(0.382)	(0.024)	(0.130)	(0.097)	(0.317)	(0.025)
Firm-specific control variables								
# employees	-0.005*	-0.003	0.005	-0.000	0.001	-0.001	0.009***	-0.000
	(0.002)	(0.002)	(0.004)	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)
ln(revenue)	0.687**	0.066	-1.064	0.101**	-0.392	-0.057	-4.146***	0.024
	(0.213)	(0.188)	(1.133)	(0.038)	(0.324)	(0.168)	(0.542)	(0.029)
EBIT	-0.161	0.038	0.270	0.006	0.053	0.011	0.111	-0.001
	(0.094)	(0.062)	(0.288)	(0.010)	(0.050)	(0.050)	(0.112)	(0.009)
DE	-0.006	0.001	-0.018**	0.000	-0.005*	-0.000	-0.011*	-0.000
	(0.005)	(0.002)	(0.006)	(0.000)	(0.002)	(0.002)	(0.004)	(0.000)
Consumer Cyclicals	-13.816**	-2.147	-0.302	0.317***	0.481	0.079	-0.451	0.009
	(5.027)	(4.338)	(2.132)	(0.094)	(0.442)	(0.405)	(0.967)	(0.078)
Consumer Non-Cyclicals	-14.027**	-2.521	-0.765	0.238**	0.458	0.202	-2.480**	0.070
	(5.199)	(4.421)	(1.796)	(0.088)	(0.418)	(0.502)	(0.903)	(0.074)
Healthcare	-13.137*	-2.608	0.000	0.000	0.000	0.247	-0.880	0.000
	(5.372)	(4.447)	(.)	(.)	(.)	(0.504)	(1.407)	(.)
Industrials	-15.030**	-2.822	-1.188	0.101	0.121	-0.362	0.000	-0.052
	(5.430)	(4.624)	(2.116)	(0.083)	(0.482)	(0.565)	(.)	(0.070)
Real Estate	-10.490*	-0.549	-4.515	0.474*	3.144**	1.552*	0.632	-0.094
	(4.506)	(4.066)	(3.334)	(0.198)	(1.169)	(0.744)	(2.406)	(0.108)
Technology	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Utilities	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
constant	0.000	0.000	34.697	-2.483**	7.371	-0.754	114.376***	-0.650
	(.)	(.)	(29.032)	(0.953)	(7.922)	(3.677)	(13.710)	(0.702)
Observations	764	764	629	764	776	778	666	776

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on the propensity score matched data set for appearance. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board) and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.50: Panel regression of firm performance (non-French), only data before 2011[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:	Before				After			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-French	0.012	0.138	0.882	-0.004	0.190	0.093	0.697	-0.017
	(0.298)	(0.114)	(0.463)	(0.017)	(0.143)	(0.101)	(0.501)	(0.018)
month	-0.037***	-0.009**	-0.066***	-0.001	-0.015***	-0.000	-0.028*	0.001***
	(0.009)	(0.003)	(0.014)	(0.001)	(0.004)	(0.003)	(0.014)	(0.000)
non-French × month	0.005	0.005	0.040*	0.001	-0.006	-0.007*	-0.020	0.000
	(0.013)	(0.004)	(0.020)	(0.001)	(0.006)	(0.004)	(0.018)	(0.001)
Appointment-specific control variables								
Female	-0.565	-0.044	-0.500	-0.011	-0.144	-0.111	0.327	-0.022
	(0.386)	(0.142)	(0.516)	(0.021)	(0.165)	(0.141)	(0.542)	(0.024)
non-French × Female	0.719	0.010	0.667	-0.004	-0.165	-0.004	-0.300	0.012
	(0.517)	(0.203)	(0.754)	(0.030)	(0.215)	(0.174)	(0.739)	(0.032)
Extern	0.328	-0.003	0.806	-0.020	-0.183	-0.103	-0.104	0.006
	(0.289)	(0.114)	(0.422)	(0.014)	(0.125)	(0.087)	(0.450)	(0.018)
Academic	0.051	-0.129	-0.418	-0.012	0.034	-0.066	0.461	-0.017
	(0.258)	(0.096)	(0.332)	(0.012)	(0.101)	(0.077)	(0.341)	(0.013)
Grandeecole	0.139	0.150	0.646	-0.008	0.134	0.148	0.037	0.010
	(0.277)	(0.117)	(0.458)	(0.015)	(0.101)	(0.095)	(0.481)	(0.018)
Age	-0.014	-0.008	0.008	0.001	0.004	-0.008	0.003	-0.000
	(0.013)	(0.005)	(0.020)	(0.001)	(0.006)	(0.004)	(0.020)	(0.001)
Jointappointment	-0.802**	-0.040	-0.991*	0.004	-0.047	0.148	-0.192	0.015
	(0.284)	(0.084)	(0.407)	(0.016)	(0.124)	(0.088)	(0.336)	(0.016)
Supervisory board	0.707	0.127	-1.185**	0.024	0.114	-0.090	0.013	0.009
	(0.396)	(0.097)	(0.387)	(0.017)	(0.138)	(0.084)	(0.448)	(0.020)
Management board	0.433	-0.183	-0.911	-0.020	0.007	-0.164	0.270	0.019
	(0.759)	(0.163)	(0.537)	(0.021)	(0.198)	(0.113)	(0.565)	(0.025)
CEO	-0.081	-0.134	-0.259	-0.005	-0.060	-0.256*	0.714	-0.004
	(0.275)	(0.179)	(0.492)	(0.022)	(0.121)	(0.111)	(0.468)	(0.022)
Later CEO	-0.428	0.075	-1.066	-0.013	0.201	0.297	0.016	0.019
	(0.382)	(0.198)	(0.873)	(0.025)	(0.234)	(0.199)	(1.125)	(0.035)
Prior Boardsize	-0.046	-0.000	0.026	-0.000	0.046***	0.052***	-0.021	0.005**
	(0.030)	(0.013)	(0.047)	(0.002)	(0.012)	(0.010)	(0.042)	(0.002)
# Prior Females	-0.076	0.029	0.157	0.000	-0.222***	-0.084*	-0.713***	0.008
	(0.093)	(0.040)	(0.170)	(0.006)	(0.043)	(0.037)	(0.182)	(0.007)
# Prior non-Frenchs	-0.260***	-0.153***	-0.182*	-0.005	0.017	-0.073***	0.029	-0.010***
	(0.059)	(0.023)	(0.090)	(0.003)	(0.022)	(0.013)	(0.075)	(0.002)
Firm-specific control variables								
# employees	-0.003*	-0.001	0.001	-0.000	0.002**	-0.001*	0.007***	-0.000*
	(0.001)	(0.001)	(0.003)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
ln(revenue)	0.140	0.103	0.862	0.021	-0.614***	0.074	-2.368***	0.043***
	(0.383)	(0.136)	(0.784)	(0.018)	(0.115)	(0.058)	(0.362)	(0.009)
EBIT	-0.003	0.009	-0.053	0.000	0.038***	-0.001	0.215***	0.002
	(0.014)	(0.005)	(0.068)	(0.002)	(0.010)	(0.006)	(0.062)	(0.003)
DE	-0.003*	-0.002**	-0.011***	-0.000	-0.001	-0.001*	-0.003	-0.000
	(0.002)	(0.001)	(0.003)	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)
Consumer Cyclicals	0.783	-0.500*	2.850**	-0.093	-0.919**	-1.456***	3.559***	-0.167***
	(0.486)	(0.203)	(0.895)	(0.052)	(0.321)	(0.175)	(0.751)	(0.043)
Consumer Non-Cyclicals	1.204	-0.018	-0.297	-0.060	-0.655	-0.851***	1.366	-0.086
	(0.625)	(0.236)	(1.088)	(0.054)	(0.344)	(0.187)	(0.813)	(0.048)
Energy	3.657*	-0.105	-1.382	-0.111	0.007	-0.996***	-2.189*	-0.165**
	(1.546)	(0.333)	(0.815)	(0.067)	(0.475)	(0.242)	(0.926)	(0.058)
Financials	1.233	-0.230	0.543	-0.045	-0.452	-1.198***	2.830**	-0.155**
	(0.661)	(0.226)	(1.248)	(0.055)	(0.336)	(0.199)	(1.030)	(0.048)
Healthcare	0.596	-0.091	-0.134	-0.205***	-0.820*	-0.858***	-0.305	-0.218***
	(0.389)	(0.200)	(0.771)	(0.053)	(0.345)	(0.208)	(0.840)	(0.048)
Industrials	0.646	-0.355	-0.006	-0.158**	-0.169	-1.024***	2.675***	-0.224***
	(0.561)	(0.236)	(0.983)	(0.055)	(0.338)	(0.186)	(0.798)	(0.048)
Real Estate	1.290	0.776*	2.550	-0.012	-0.559	0.057	-2.248*	-0.148**
	(0.985)	(0.356)	(2.270)	(0.067)	(0.437)	(0.214)	(0.985)	(0.052)
Technology	0.173	-1.839***	1.292	-0.266***	-1.220***	-2.371***	1.540*	-0.272***
	(0.592)	(0.239)	(0.890)	(0.053)	(0.326)	(0.200)	(0.774)	(0.046)
Utilities	1.442	0.015	2.951	-0.084	-0.583	-1.267***	4.061**	-0.177***
	(0.773)	(0.253)	(1.927)	(0.058)	(0.343)	(0.208)	(1.287)	(0.048)
constant	-0.723	-1.146	-19.768	-0.012	14.676***	-0.504	54.353***	-0.497*
	(8.186)	(2.966)	(17.872)	(0.402)	(2.642)	(1.314)	(8.471)	(0.212)
Observations	5538	5559	4678	5538	5849	5885	4995	5849

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on the propensity score matched data set for nationality. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board), no-CEO-appointment and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.51: Panel regression of firm performance (non-white), only data since 2011[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:	Before				After			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-white	-0.219	-0.054	-0.590	0.031	-0.238	-0.058	-0.330	0.036
	(0.147)	(0.110)	(0.373)	(0.024)	(0.226)	(0.118)	(0.492)	(0.024)
month	0.007	0.003	0.019*	-0.001	-0.007	-0.011***	-0.013	-0.000
	(0.004)	(0.003)	(0.009)	(0.001)	(0.006)	(0.002)	(0.018)	(0.001)
non-white × month	-0.001	-0.003	-0.000	-0.000	0.005	0.010**	0.008	0.000
	(0.005)	(0.004)	(0.013)	(0.001)	(0.007)	(0.004)	(0.021)	(0.001)
Appointment-specific control variables								
Female	0.102	0.091	-0.061	0.012	-0.092	-0.090	-0.338	0.021
	(0.188)	(0.116)	(0.394)	(0.029)	(0.231)	(0.132)	(0.496)	(0.025)
non-white × Female	-0.123	-0.313	0.262	-0.020	-0.084	-0.074	-0.027	-0.028
	(0.227)	(0.183)	(0.512)	(0.035)	(0.261)	(0.160)	(0.630)	(0.036)
Extern	0.245	0.061	0.701	-0.021	0.133	0.112	0.562	-0.008
	(0.144)	(0.095)	(0.372)	(0.019)	(0.127)	(0.124)	(0.372)	(0.020)
Academic	0.031	0.113	0.307	-0.020	0.072	0.030	0.224	-0.020
	(0.120)	(0.088)	(0.359)	(0.018)	(0.195)	(0.108)	(0.438)	(0.023)
Grandeecole	0.009	0.097	-0.183	-0.005	0.227	0.180	0.564	0.003
	(0.118)	(0.086)	(0.326)	(0.018)	(0.153)	(0.113)	(0.376)	(0.020)
Age	0.003	0.005	-0.030	-0.000	0.013	0.010	-0.016	-0.001
	(0.007)	(0.005)	(0.018)	(0.001)	(0.009)	(0.006)	(0.022)	(0.001)
Jointappointment	0.075	0.032	0.501	-0.021	-0.033	0.017	0.246	-0.011
	(0.136)	(0.090)	(0.285)	(0.020)	(0.142)	(0.098)	(0.353)	(0.019)
Supervisory board	-0.300	-0.216	0.478	0.017	-0.742**	-0.211	-0.133	0.029
	(0.259)	(0.149)	(0.545)	(0.022)	(0.237)	(0.160)	(0.565)	(0.022)
Management board	-0.429	-0.269	1.231	0.076	-0.950*	-0.404	-0.479	0.083
	(0.254)	(0.179)	(0.745)	(0.048)	(0.415)	(0.261)	(0.722)	(0.056)
Prior Boardsize	0.025	-0.013	0.117	-0.012**	0.023	-0.007	0.192*	-0.007
	(0.032)	(0.029)	(0.090)	(0.004)	(0.036)	(0.022)	(0.081)	(0.004)
# Prior Females	0.107*	0.012	0.019	0.026***	-0.109	-0.074	-0.085	0.014
	(0.047)	(0.037)	(0.107)	(0.007)	(0.061)	(0.038)	(0.124)	(0.007)
# Prior non-whites	0.032	0.004	0.429*	-0.032***	0.054	0.037	-0.044	-0.021*
	(0.057)	(0.044)	(0.175)	(0.009)	(0.085)	(0.056)	(0.194)	(0.009)
Firm-specific control variables								
# employees	-0.002**	-0.002***	0.002	0.000	-0.004***	-0.003***	-0.004	0.000*
	(0.001)	(0.000)	(0.002)	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)
ln(revenue)	-0.085	-0.045	0.605	0.018	-0.099	-0.235***	0.530*	-0.017
	(0.138)	(0.169)	(0.498)	(0.018)	(0.071)	(0.054)	(0.247)	(0.010)
EBIT	0.006	0.020	-0.452***	0.011	0.094	0.166***	-0.426*	0.010
	(0.040)	(0.048)	(0.123)	(0.006)	(0.050)	(0.049)	(0.186)	(0.008)
DE	0.002*	0.000	-0.001	-0.000	-0.000	0.001	-0.002	-0.000
	(0.001)	(0.001)	(0.003)	(0.000)	(0.001)	(0.001)	(0.004)	(0.000)
Consumer Cyclicals	0.134	-0.051	-1.342	-0.164	0.639*	0.294	-0.280	-0.191
	(0.228)	(0.241)	(0.892)	(0.126)	(0.306)	(0.171)	(0.807)	(0.144)
Consumer Non-Cyclicals	-0.424*	-0.082	-1.427	-0.039	0.265	0.429**	-0.339	-0.019
	(0.214)	(0.235)	(0.859)	(0.126)	(0.297)	(0.158)	(0.752)	(0.144)
Healthcare	-0.167	-0.260	1.757	-0.210	-0.775	-0.533	3.493	-0.068
	(0.348)	(0.429)	(1.291)	(0.138)	(0.549)	(0.445)	(1.811)	(0.169)
Industrials	0.133	0.069	-1.139	-0.096	0.782*	0.418*	-0.164	-0.070
	(0.261)	(0.247)	(0.764)	(0.127)	(0.309)	(0.169)	(0.732)	(0.146)
Real Estate	-0.153	0.139	-1.488	0.010	0.546	-0.067	-0.860	-0.067
	(0.421)	(0.456)	(1.668)	(0.138)	(0.428)	(0.278)	(1.050)	(0.148)
Technology	-0.637***	-1.165***	1.426	-0.350**	0.156	-0.295*	4.551**	-0.368*
	(0.180)	(0.315)	(1.504)	(0.129)	(0.244)	(0.141)	(1.393)	(0.145)
Utilities	-0.848***	-0.926**	-4.198***	-0.233	0.298	0.050	-3.649***	-0.269
	(0.225)	(0.298)	(0.965)	(0.129)	(0.450)	(0.387)	(0.770)	(0.145)
constant	1.487	1.223	-12.763	0.254	2.248	5.063***	-11.543*	1.024***
	(2.904)	(3.562)	(10.879)	(0.400)	(1.631)	(1.162)	(5.753)	(0.241)
Observations	1934	1937	1908	1934	2006	2005	1982	2006

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on the propensity score matched data set for appearance. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board) and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.52: Panel regression of firm performance (non-French), only data since 2011[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:	Before				After			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-French	0.010	0.106	-0.385	0.043*	0.177	0.160*	0.134	0.034
	(0.142)	(0.083)	(0.342)	(0.017)	(0.127)	(0.074)	(0.361)	(0.019)
month	0.004	-0.002	0.037***	-0.001**	0.005	-0.001	0.013	-0.001*
	(0.004)	(0.002)	(0.010)	(0.000)	(0.003)	(0.002)	(0.010)	(0.000)
non-French × month	0.003	0.004	-0.023*	0.000	-0.006	-0.003	0.006	0.000
	(0.005)	(0.003)	(0.012)	(0.000)	(0.005)	(0.003)	(0.015)	(0.000)
Appointment-specific control variables								
Female	-0.035	0.026	0.012	0.041**	-0.123	-0.019	0.039	0.039*
	(0.124)	(0.077)	(0.254)	(0.015)	(0.094)	(0.061)	(0.323)	(0.017)
non-French × Female	0.160	0.022	0.128	-0.024	0.019	-0.106	-0.289	-0.030
	(0.165)	(0.102)	(0.383)	(0.021)	(0.141)	(0.092)	(0.431)	(0.024)
Extern	-0.135	-0.066	-0.168	0.009	-0.210*	-0.102	-0.025	0.004
	(0.101)	(0.069)	(0.241)	(0.014)	(0.095)	(0.055)	(0.303)	(0.017)
Academic	0.104	0.067	0.301	-0.001	0.174	0.050	0.004	-0.006
	(0.096)	(0.054)	(0.267)	(0.015)	(0.095)	(0.052)	(0.257)	(0.015)
Grandeecole	0.083	0.089	0.144	0.006	0.187*	0.088	0.270	-0.006
	(0.108)	(0.067)	(0.289)	(0.014)	(0.084)	(0.054)	(0.326)	(0.016)
Age	0.010	0.006	0.038**	-0.002*	0.015**	0.007*	0.032*	-0.001
	(0.007)	(0.003)	(0.013)	(0.001)	(0.005)	(0.003)	(0.016)	(0.001)
Jointappointment	-0.148	-0.144*	0.251	-0.016	0.013	-0.020	0.299	-0.008
	(0.100)	(0.057)	(0.211)	(0.012)	(0.074)	(0.048)	(0.236)	(0.012)
Supervisory board	-0.284	-0.144	-0.619	0.061**	-0.882***	-0.498***	-0.722	0.028
	(0.148)	(0.079)	(0.336)	(0.021)	(0.109)	(0.077)	(0.402)	(0.021)
Management board	-0.901***	-0.263*	-1.147**	0.145***	-1.191***	-0.569***	-0.990	0.094**
	(0.197)	(0.115)	(0.431)	(0.027)	(0.183)	(0.105)	(0.675)	(0.034)
CEO	0.658	0.369*	0.861	0.034	0.173	0.099	0.688	0.034
	(0.394)	(0.160)	(0.789)	(0.048)	(0.169)	(0.105)	(0.914)	(0.038)
Later CEO	-0.462*	-0.340***	0.152	-0.038	0.025	-0.072	0.045	-0.019
	(0.230)	(0.086)	(0.423)	(0.020)	(0.190)	(0.112)	(0.595)	(0.034)
Prior Boardsize	0.003	-0.020	-0.044	-0.010***	-0.014	-0.005	0.016	-0.006*
	(0.020)	(0.014)	(0.055)	(0.002)	(0.017)	(0.011)	(0.064)	(0.003)
# Prior Females	0.074***	0.011	0.205***	0.012***	-0.045	-0.051***	0.104	0.003
	(0.021)	(0.018)	(0.060)	(0.004)	(0.024)	(0.015)	(0.074)	(0.004)
# Prior non-Frenchs	0.005	-0.042**	0.229***	-0.021***	0.009	-0.029*	0.196***	-0.023***
	(0.024)	(0.014)	(0.052)	(0.003)	(0.020)	(0.013)	(0.059)	(0.003)
Firm-specific control variables								
# employees	-0.001	-0.001*	-0.000	0.000*	-0.004***	-0.002***	-0.005	0.000**
	(0.001)	(0.000)	(0.002)	(0.000)	(0.001)	(0.000)	(0.003)	(0.000)
ln(revenue)	-0.191*	-0.055	0.082	0.040***	-0.011	-0.032	0.462*	0.013
	(0.086)	(0.048)	(0.182)	(0.007)	(0.057)	(0.067)	(0.191)	(0.010)
EBIT	0.011	0.013**	-0.222***	0.006***	0.010	0.021***	-0.147	0.004**
	(0.006)	(0.005)	(0.059)	(0.001)	(0.008)	(0.006)	(0.081)	(0.001)
DE	-0.000	-0.001	0.003	-0.000	-0.001	-0.002***	-0.004*	0.000
	(0.001)	(0.001)	(0.003)	(0.000)	(0.001)	(0.000)	(0.002)	(0.000)
Consumer Cyclicals	0.428*	-0.225	1.955***	-0.307***	1.021***	0.281*	2.313***	-0.328***
	(0.211)	(0.120)	(0.434)	(0.030)	(0.159)	(0.110)	(0.505)	(0.034)
Consumer Non-Cyclicals	0.068	-0.087	1.292**	-0.138***	0.453**	0.245*	1.578***	-0.150***
	(0.160)	(0.103)	(0.406)	(0.029)	(0.151)	(0.110)	(0.477)	(0.034)
Energy	0.002	-0.627***	0.308	-0.255***	-0.247	-0.639**	4.010**	-0.224***
	(0.219)	(0.110)	(0.436)	(0.030)	(0.166)	(0.234)	(1.362)	(0.052)
Financials	-0.488**	-0.953***	1.046	-0.378***	0.222	-0.249*	1.283	-0.352***
	(0.163)	(0.135)	(0.535)	(0.030)	(0.162)	(0.123)	(0.684)	(0.033)
Healthcare	0.165	-0.227*	5.243***	-0.228***	0.054	-0.141	4.240***	-0.221***
	(0.160)	(0.111)	(0.757)	(0.035)	(0.157)	(0.155)	(0.672)	(0.036)
Industrials	0.609**	0.068	0.928**	-0.230***	1.099***	0.485***	1.645***	-0.224***
	(0.206)	(0.125)	(0.333)	(0.030)	(0.185)	(0.133)	(0.473)	(0.033)
Real Estate	0.250	0.114	0.778	-0.012	1.065***	0.405*	1.636*	-0.047
	(0.303)	(0.146)	(0.516)	(0.036)	(0.208)	(0.202)	(0.693)	(0.042)
Technology	-0.156	-0.879***	2.735***	-0.391***	0.450**	-0.144	3.722***	-0.411***
	(0.162)	(0.145)	(0.658)	(0.032)	(0.168)	(0.132)	(0.736)	(0.034)
Utilities	-0.447*	-0.902***	-0.453	-0.372***	0.222	-0.246	-0.870	-0.398***
	(0.210)	(0.149)	(0.553)	(0.037)	(0.212)	(0.180)	(0.502)	(0.038)
constant	4.117*	1.835	-4.990	-0.053	0.241	1.065	-13.702**	0.548*
	(1.811)	(1.026)	(3.919)	(0.150)	(1.204)	(1.429)	(4.279)	(0.217)
Observations	6671	6680	6173	6671	6593	6591	6171	6593

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on the propensity score matched data set for nationality. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board), no-CEO-appointment and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.53: Panel regression of firm performance with shares of board structure variables prior to an appointment (instead of absolute values) for non-white[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:		Before				After		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-white	0.018	0.078	0.284	-0.012	0.138	0.040	0.799	-0.007
	(0.152)	(0.113)	(0.433)	(0.025)	(0.184)	(0.119)	(0.446)	(0.024)
month	0.007	0.002	0.012	-0.000	-0.002	-0.001	-0.011	-0.000
	(0.007)	(0.003)	(0.017)	(0.001)	(0.005)	(0.002)	(0.011)	(0.001)
non-white × month	-0.002	-0.002	-0.000	0.000	-0.003	-0.001	0.004	0.000
	(0.008)	(0.005)	(0.020)	(0.001)	(0.006)	(0.003)	(0.015)	(0.001)
Appointment-specific control variables								
Female	0.522	0.216*	0.013	-0.017	0.334	0.031	0.413	-0.016
	(0.266)	(0.109)	(0.377)	(0.026)	(0.224)	(0.131)	(0.472)	(0.029)
non-white × Female	-0.407	-0.305	-0.262	0.001	-0.476	-0.082	-0.864	-0.003
	(0.309)	(0.165)	(0.535)	(0.033)	(0.273)	(0.176)	(0.577)	(0.037)
Extern	0.052	0.049	0.058	-0.006	0.318*	0.242*	0.576	-0.007
	(0.159)	(0.102)	(0.439)	(0.024)	(0.152)	(0.095)	(0.416)	(0.020)
Academic	-0.117	0.026	0.491	0.021	0.085	0.081	0.677	0.001
	(0.191)	(0.101)	(0.316)	(0.022)	(0.211)	(0.116)	(0.367)	(0.024)
Grandeecole	0.033	-0.128	-0.339	-0.036	0.096	-0.002	0.271	-0.004
	(0.171)	(0.099)	(0.342)	(0.019)	(0.136)	(0.084)	(0.375)	(0.017)
Age	0.010	0.011*	-0.036*	0.001	0.009	0.002	-0.023	0.001
	(0.009)	(0.004)	(0.016)	(0.001)	(0.007)	(0.004)	(0.019)	(0.001)
Jointappointment	-0.229	-0.002	-0.368	0.018	-0.191	0.029	-0.306	0.023
	(0.191)	(0.097)	(0.405)	(0.022)	(0.153)	(0.109)	(0.333)	(0.024)
Supervisory board	-0.557**	-0.405***	0.907	0.055*	-0.705***	-0.174	-0.213	0.091***
	(0.171)	(0.115)	(0.477)	(0.023)	(0.156)	(0.131)	(0.468)	(0.023)
Management board	-0.722*	-0.445*	1.089	0.039	-0.720**	-0.269	0.661	0.075
	(0.336)	(0.218)	(1.041)	(0.064)	(0.224)	(0.187)	(0.805)	(0.058)
Prior Boardsize	0.028	0.007	0.155	-0.004	0.013	0.015	0.209***	-0.002
	(0.026)	(0.015)	(0.085)	(0.004)	(0.026)	(0.014)	(0.061)	(0.003)
Share of women	0.006	0.004	-0.018	0.003***	0.008	0.003	0.004	0.001
	(0.005)	(0.004)	(0.014)	(0.001)	(0.006)	(0.004)	(0.015)	(0.001)
Share of non-whites	-0.019	-0.006	0.049*	-0.002	-0.005	0.004	0.015	-0.001
	(0.015)	(0.006)	(0.024)	(0.001)	(0.015)	(0.008)	(0.024)	(0.002)
Firm-specific control variables								
# employees	-0.001	-0.001***	0.001	0.000	-0.003***	-0.003***	-0.004	0.000
	(0.001)	(0.000)	(0.003)	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)
ln(revenue)	-0.172	-0.153	0.517	0.017	-0.048	-0.164***	0.294	-0.016
	(0.190)	(0.096)	(0.771)	(0.031)	(0.064)	(0.043)	(0.199)	(0.010)
EBIT	0.038	0.035	-0.185	0.013*	0.023	0.044	-0.271*	0.005
	(0.058)	(0.031)	(0.139)	(0.005)	(0.041)	(0.036)	(0.137)	(0.006)
DE	-0.002	-0.002	-0.009*	-0.000	-0.001	0.001	-0.009*	0.000
	(0.001)	(0.001)	(0.004)	(0.000)	(0.001)	(0.001)	(0.004)	(0.000)
Consumer Cyclicals	0.908*	0.078	-0.205	-0.214*	0.329	-0.318	-0.372	-0.234
	(0.440)	(0.196)	(0.410)	(0.095)	(0.236)	(0.260)	(0.609)	(0.121)
Consumer Non-Cyclicals	0.579	0.246	-0.027	-0.144	0.223	0.059	0.232	-0.119
	(0.375)	(0.189)	(0.447)	(0.097)	(0.206)	(0.276)	(0.775)	(0.123)
Healthcare	0.094	-0.290	1.648	-0.273**	-0.270	-0.221	3.446**	-0.113
	(0.562)	(0.258)	(1.577)	(0.104)	(0.358)	(0.458)	(1.193)	(0.148)
Industrials	0.806*	0.278	-0.849	-0.144	0.353	-0.027	-0.396	-0.123
	(0.400)	(0.192)	(0.459)	(0.099)	(0.194)	(0.287)	(0.628)	(0.122)
Real Estate	1.030	0.499	-0.367	-0.085	0.503	-0.431	-1.626	-0.176
	(0.681)	(0.363)	(2.269)	(0.137)	(0.262)	(0.315)	(1.010)	(0.132)
Technology	-0.233	-1.162***	-0.765	-0.341**	-0.149	-0.762*	2.326	-0.354**
	(0.405)	(0.342)	(0.598)	(0.106)	(0.201)	(0.335)	(1.616)	(0.126)
Utilities	0.034	-0.626*	-3.118***	-0.262**	0.039	-0.485	-3.541***	-0.309*
	(0.436)	(0.263)	(0.664)	(0.098)	(0.313)	(0.331)	(0.665)	(0.123)
constant	3.168	3.089	-10.448	0.144	0.888	3.678***	-6.850	0.893***
	(4.122)	(2.165)	(16.557)	(0.654)	(1.415)	(0.930)	(4.722)	(0.233)
Observations	2667	2670	2530	2667	2778	2779	2669	2778

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on a new propensity score matched data set for appearance. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board) and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.54: Panel regression of firm performance with shares of board structure variables prior to an appointment (instead of absolute values) for non-French[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:	Before				After			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-French	-0.010	0.099	0.166	0.014	0.127	0.098	0.131	0.006
	(0.163)	(0.076)	(0.313)	(0.013)	(0.110)	(0.072)	(0.308)	(0.014)
month	-0.018***	-0.005*	-0.007	-0.001*	-0.003	-0.001	0.005	0.000
	(0.005)	(0.002)	(0.009)	(0.000)	(0.003)	(0.002)	(0.008)	(0.000)
non-French × month	0.004	0.004	0.004	0.001	-0.006	-0.005*	-0.009	0.000
	(0.007)	(0.003)	(0.012)	(0.000)	(0.004)	(0.002)	(0.011)	(0.000)
Appointment-specific control variables								
Female	-0.238	0.058	-0.485	0.014	-0.054	0.034	-0.247	0.014
	(0.142)	(0.076)	(0.253)	(0.014)	(0.097)	(0.068)	(0.277)	(0.015)
non-French × Female	0.260	0.027	0.074	0.019	0.143	-0.004	0.277	0.004
	(0.220)	(0.106)	(0.372)	(0.019)	(0.133)	(0.097)	(0.373)	(0.020)
Extern	0.036	-0.029	0.152	-0.010	-0.119	-0.060	0.056	-0.005
	(0.148)	(0.070)	(0.270)	(0.013)	(0.093)	(0.063)	(0.247)	(0.013)
Academic	0.066	-0.087	-0.201	-0.002	0.011	-0.088	-0.049	-0.008
	(0.157)	(0.070)	(0.213)	(0.011)	(0.085)	(0.062)	(0.198)	(0.011)
Grandeecole	0.018	0.031	0.185	0.006	0.104	0.027	0.092	0.005
	(0.138)	(0.072)	(0.302)	(0.012)	(0.085)	(0.065)	(0.285)	(0.013)
Age	-0.001	-0.001	0.014	-0.001	0.004	-0.003	0.014	-0.001
	(0.008)	(0.003)	(0.013)	(0.001)	(0.005)	(0.003)	(0.013)	(0.001)
Jointappointment	-0.465**	-0.101	0.050	0.003	-0.064	0.043	0.373*	0.013
	(0.178)	(0.055)	(0.210)	(0.010)	(0.081)	(0.053)	(0.185)	(0.010)
Supervisory board	0.152	0.058	-1.064***	0.045***	-0.132	-0.096	-0.290	0.020
	(0.235)	(0.071)	(0.260)	(0.014)	(0.104)	(0.061)	(0.227)	(0.014)
Management board	-0.010	-0.198	-1.166**	0.020	-0.208	-0.119	-0.071	0.018
	(0.603)	(0.136)	(0.371)	(0.019)	(0.196)	(0.098)	(0.393)	(0.019)
CEO	0.158	-0.052	0.147	-0.016	-0.193	-0.248*	0.062	-0.000
	(0.287)	(0.159)	(0.555)	(0.028)	(0.156)	(0.123)	(0.584)	(0.019)
Later CEO	-0.775	-0.195	-0.080	-0.011	-0.198	-0.013	-0.382	0.024
	(0.406)	(0.104)	(1.037)	(0.030)	(0.223)	(0.201)	(0.925)	(0.039)
Prior Boardsize	-0.061*	-0.041***	0.022	-0.004*	0.015	0.002	0.069	-0.003
	(0.028)	(0.011)	(0.044)	(0.002)	(0.017)	(0.009)	(0.041)	(0.002)
Share of Women	-0.013**	0.001	-0.012	0.002***	0.001	-0.001	0.018**	0.001*
	(0.004)	(0.002)	(0.008)	(0.000)	(0.002)	(0.002)	(0.007)	(0.000)
Share of non-Frenchs	-0.018***	-0.011***	0.020**	-0.001***	-0.001	-0.007***	0.029***	-0.002***
	(0.005)	(0.002)	(0.007)	(0.000)	(0.003)	(0.002)	(0.007)	(0.000)
Firm-specific control variables								
# employees	-0.002**	-0.001*	0.001	-0.000	-0.001**	-0.001**	-0.001	0.000
	(0.001)	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)
ln(revenue)	-0.092	0.055	0.228	0.036**	-0.022	0.021	0.073	0.025*
	(0.192)	(0.076)	(0.338)	(0.013)	(0.070)	(0.064)	(0.228)	(0.010)
EBIT	0.013	0.011**	-0.061	0.002	0.011*	0.004	0.016	0.003
	(0.008)	(0.004)	(0.046)	(0.002)	(0.005)	(0.004)	(0.052)	(0.002)
DE	-0.002*	-0.001**	-0.004*	-0.000**	-0.001	-0.001***	-0.003	-0.000
	(0.001)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)
Consumer Cyclicals	-0.023	-0.479***	2.128***	-0.189***	-0.329	-0.651***	1.875***	-0.223***
	(0.246)	(0.117)	(0.417)	(0.034)	(0.175)	(0.109)	(0.316)	(0.030)
Consumer Non-Cyclicals	0.153	-0.121	0.094	-0.096**	-0.323	-0.263*	0.636	-0.110***
	(0.269)	(0.126)	(0.526)	(0.036)	(0.168)	(0.120)	(0.461)	(0.032)
Energy	1.129	-0.521***	0.087	-0.186***	-0.505*	-0.778***	2.012*	-0.188***
	(0.588)	(0.149)	(0.358)	(0.039)	(0.249)	(0.191)	(0.799)	(0.045)
Financials	0.068	-0.540***	0.610	-0.195***	-0.527**	-0.628***	0.218	-0.229***
	(0.244)	(0.126)	(0.497)	(0.036)	(0.181)	(0.118)	(0.435)	(0.032)
Healthcare	-0.154	-0.289*	1.748**	-0.193***	-0.549***	-0.402**	1.274*	-0.216***
	(0.230)	(0.123)	(0.609)	(0.038)	(0.157)	(0.126)	(0.629)	(0.035)
Industrials	0.237	-0.165	0.315	-0.170***	0.173	-0.197	1.066**	-0.192***
	(0.254)	(0.124)	(0.416)	(0.036)	(0.178)	(0.114)	(0.356)	(0.032)
Real Estate	0.353	0.381	1.685*	0.008	-0.161	-0.164	0.299	-0.017
	(0.515)	(0.206)	(0.858)	(0.050)	(0.270)	(0.186)	(0.595)	(0.038)
Technology	-0.604*	-1.497***	1.642***	-0.303***	-0.660***	-1.261***	1.801***	-0.310***
	(0.241)	(0.155)	(0.481)	(0.034)	(0.160)	(0.142)	(0.488)	(0.031)
Utilities	-0.240	-0.779***	0.221	-0.213***	-0.682***	-0.803***	0.015	-0.269***
	(0.312)	(0.166)	(0.821)	(0.041)	(0.200)	(0.153)	(0.676)	(0.035)
constant	4.781	0.164	-6.320	-0.189	1.067	0.569	-4.789	0.121
	(4.066)	(1.631)	(7.300)	(0.269)	(1.544)	(1.401)	(5.039)	(0.225)
Observations	11829	11858	10519	11829	12005	12039	10766	12005

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on a new propensity score matched data set for nationality. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board), no-CEO-appointment and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.55: Panel regression of firm performance, only firms in CAC completely between 2002 and 2018 (non-white)[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:	Before				After			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-white	0.232	0.095	0.188	-0.028	-0.023	-0.037	0.021	-0.015
	(0.126)	(0.129)	(0.509)	(0.026)	(0.199)	(0.138)	(0.473)	(0.026)
month	0.008	0.003	0.015	-0.000	0.009	0.003	-0.003	-0.000
	(0.005)	(0.003)	(0.014)	(0.001)	(0.007)	(0.004)	(0.016)	(0.001)
non-white × month	-0.002	-0.001	-0.009	-0.001	-0.014	-0.006	-0.011	0.001
	(0.007)	(0.005)	(0.018)	(0.001)	(0.008)	(0.004)	(0.021)	(0.001)
Appointment-specific control variables								
Female	0.568**	0.272	1.005	0.011	0.009	0.084	-0.133	0.030
	(0.210)	(0.141)	(0.587)	(0.032)	(0.233)	(0.157)	(0.505)	(0.032)
non-white × Female	-0.516*	-0.219	-0.778	-0.003	0.133	0.038	-0.134	-0.030
	(0.237)	(0.186)	(0.678)	(0.036)	(0.285)	(0.195)	(0.574)	(0.038)
Extern	0.177	0.053	0.469	-0.028	0.385*	0.224	0.956*	-0.002
	(0.151)	(0.203)	(0.478)	(0.026)	(0.161)	(0.136)	(0.436)	(0.031)
Academic	-0.003	0.089	0.172	0.003	0.019	-0.056	0.303	0.001
	(0.114)	(0.106)	(0.343)	(0.022)	(0.221)	(0.144)	(0.362)	(0.023)
Grandeecole	-0.219*	-0.182	-0.130	-0.010	0.110	-0.001	0.755	-0.002
	(0.102)	(0.116)	(0.399)	(0.022)	(0.189)	(0.105)	(0.405)	(0.022)
Age	0.001	0.012	-0.018	0.001	0.008	0.005	-0.012	0.001
	(0.006)	(0.007)	(0.023)	(0.001)	(0.012)	(0.007)	(0.021)	(0.001)
Jointappointment	-0.118	0.032	-0.345	0.036	-0.149	-0.039	-0.613	0.035
	(0.134)	(0.104)	(0.359)	(0.021)	(0.173)	(0.109)	(0.371)	(0.019)
Supervisory board	-0.379**	-0.289*	0.052	0.057*	-0.638**	-0.150	-0.466	0.063*
	(0.146)	(0.123)	(0.542)	(0.024)	(0.237)	(0.170)	(0.414)	(0.028)
Management board	-0.546*	-0.635*	0.920	0.053	-0.743*	-0.249	1.365	0.079
	(0.241)	(0.279)	(0.925)	(0.052)	(0.292)	(0.255)	(0.986)	(0.051)
Prior Boardsize	-0.011	-0.002	0.110	-0.001	0.012	0.021	0.116	0.005
	(0.016)	(0.017)	(0.064)	(0.004)	(0.031)	(0.018)	(0.068)	(0.003)
# Prior Females	0.038	0.042	-0.211*	0.017**	-0.027	-0.022	-0.102	0.001
	(0.026)	(0.035)	(0.107)	(0.006)	(0.051)	(0.034)	(0.110)	(0.006)
# Prior non-whites	0.149***	0.060	0.382*	-0.031**	0.100	0.048	-0.025	-0.017
	(0.039)	(0.038)	(0.181)	(0.010)	(0.074)	(0.052)	(0.136)	(0.009)
Firm-specific control variables								
# employees	-0.001	-0.001***	0.003	-0.000	-0.004*	-0.003***	-0.002	-0.000
	(0.001)	(0.000)	(0.002)	(0.000)	(0.001)	(0.001)	(0.004)	(0.000)
ln(revenue)	-0.004	-0.327***	0.388	0.009	-0.008	-0.124*	0.779**	-0.041***
	(0.116)	(0.095)	(0.415)	(0.017)	(0.104)	(0.058)	(0.243)	(0.011)
EBIT	-0.004	0.088**	-0.229*	0.010*	-0.078	-0.060	-0.550***	0.022**
	(0.047)	(0.033)	(0.111)	(0.005)	(0.063)	(0.050)	(0.115)	(0.007)
DE	0.000	0.000	-0.013***	-0.000	-0.002	-0.001	-0.014***	0.000
	(0.002)	(0.001)	(0.004)	(0.000)	(0.001)	(0.001)	(0.004)	(0.000)
Consumer Cyclicals	-0.167	-0.208	0.737	-0.295***	0.314	-0.135	1.458**	-0.352***
	(0.207)	(0.256)	(0.746)	(0.035)	(0.268)	(0.214)	(0.510)	(0.034)
Consumer Non-Cyclicals	-0.210	0.164	0.180	-0.228***	0.497	0.338	1.644*	-0.225***
	(0.203)	(0.249)	(0.677)	(0.036)	(0.258)	(0.194)	(0.652)	(0.041)
Healthcare	-0.509	-0.714*	0.166	-0.353***	0.563	1.053*	5.448***	-0.324***
	(0.423)	(0.331)	(1.147)	(0.055)	(0.582)	(0.432)	(1.010)	(0.070)
Industrials	0.312	0.287	0.394	-0.344***	0.916*	0.390	1.540*	-0.340***
	(0.318)	(0.302)	(0.743)	(0.037)	(0.360)	(0.237)	(0.601)	(0.040)
Technology	-0.688*	-1.034*	0.445	-0.427***	0.394	0.118	2.737**	-0.496***
	(0.292)	(0.439)	(0.788)	(0.056)	(0.359)	(0.271)	(0.935)	(0.053)
Utilities	-0.856***	-0.791*	-2.182**	-0.375***	0.387	-0.030	-1.761**	-0.419***
	(0.252)	(0.336)	(0.772)	(0.040)	(0.406)	(0.306)	(0.573)	(0.043)
constant	0.286	6.932**	-8.323	0.433	0.224	2.809*	-17.106**	1.537***
	(2.593)	(2.150)	(9.051)	(0.377)	(2.241)	(1.292)	(5.639)	(0.253)
Observations	2035	2035	2008	2035	2072	2072	2072	2072

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on a new propensity score matched data set for appearance. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board) and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.56: Panel regression of firm performance, only firms in CAC completely between 2002 and 2018 (non-French)[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:		Before				After		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-French	0.250*	0.153	0.202	-0.002	0.239*	0.162*	-0.029	-0.012
	(0.108)	(0.082)	(0.362)	(0.015)	(0.100)	(0.082)	(0.346)	(0.015)
month	-0.014***	-0.005*	-0.015	-0.001**	0.003	0.002	-0.010	0.000
	(0.003)	(0.002)	(0.010)	(0.000)	(0.003)	(0.002)	(0.010)	(0.000)
non-French × month	0.003	0.005	-0.000	0.001	-0.010**	-0.007**	-0.001	0.000
	(0.006)	(0.003)	(0.014)	(0.000)	(0.004)	(0.003)	(0.012)	(0.001)
Appointment-specific control variables								
Female	-0.063	0.020	0.041	-0.012	0.058	0.085	-0.142	-0.013
	(0.105)	(0.083)	(0.314)	(0.016)	(0.103)	(0.080)	(0.327)	(0.016)
non-French × Female	-0.034	0.089	0.325	0.060**	0.054	-0.002	0.249	0.058**
	(0.168)	(0.112)	(0.474)	(0.023)	(0.149)	(0.117)	(0.434)	(0.022)
Extern	0.082	-0.020	-0.079	-0.011	-0.036	-0.021	-0.178	-0.009
	(0.085)	(0.070)	(0.319)	(0.013)	(0.094)	(0.069)	(0.288)	(0.013)
Academic	-0.096	-0.070	0.074	-0.007	0.027	-0.047	0.312	-0.017
	(0.103)	(0.071)	(0.299)	(0.013)	(0.091)	(0.079)	(0.275)	(0.012)
Grandeecole	0.009	0.022	0.620	0.002	0.109	0.040	0.579	-0.011
	(0.111)	(0.075)	(0.350)	(0.015)	(0.094)	(0.077)	(0.326)	(0.014)
Age	-0.002	0.006	0.018	0.000	0.005	-0.001	0.017	-0.001
	(0.005)	(0.003)	(0.016)	(0.001)	(0.005)	(0.004)	(0.015)	(0.001)
Jointappointment	-0.133	-0.057	-0.380	-0.005	-0.027	0.009	-0.213	0.014
	(0.090)	(0.061)	(0.280)	(0.014)	(0.073)	(0.059)	(0.235)	(0.011)
Supervisory board	-0.294*	-0.081	-1.162***	0.038*	-0.290***	-0.035	-0.626*	0.048**
	(0.136)	(0.069)	(0.306)	(0.016)	(0.085)	(0.066)	(0.249)	(0.017)
Management board	-0.801***	-0.339**	-1.483**	0.022	-0.383**	-0.148	-0.524	0.041*
	(0.148)	(0.126)	(0.455)	(0.023)	(0.133)	(0.102)	(0.444)	(0.020)
CEO	-0.012	0.022	-0.271	0.001	-0.095	-0.066	-0.498	-0.002
	(0.193)	(0.126)	(0.737)	(0.032)	(0.169)	(0.140)	(0.468)	(0.017)
Later CEO	0.140	0.152	-0.218	-0.036	0.149	0.293	-0.172	0.010
	(0.387)	(0.230)	(1.670)	(0.039)	(0.257)	(0.164)	(0.817)	(0.037)
Prior Boardsize	0.019	0.003	0.016	-0.003	-0.008	0.014	0.056	0.002
	(0.014)	(0.008)	(0.043)	(0.002)	(0.011)	(0.010)	(0.039)	(0.002)
# Prior Females	-0.039*	0.004	-0.195**	0.011***	0.035*	-0.007	-0.009	-0.001
	(0.017)	(0.014)	(0.068)	(0.003)	(0.016)	(0.014)	(0.055)	(0.003)
# Prior non-Frenchs	0.065	-0.018	0.050	-0.009*	0.002	-0.044*	0.061	-0.001
	(0.042)	(0.021)	(0.087)	(0.004)	(0.023)	(0.019)	(0.072)	(0.003)
Firm-specific control variables								
# employees	-0.002	-0.001	0.002	-0.000*	-0.001*	-0.001*	0.002	-0.000
	(0.001)	(0.001)	(0.003)	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)
ln(revenue)	-0.408*	-0.167*	0.042	0.050***	0.051	-0.045	-0.254	0.001
	(0.196)	(0.083)	(0.416)	(0.013)	(0.077)	(0.056)	(0.304)	(0.011)
EBIT	0.018**	0.012***	-0.000	0.001	0.009	0.002	0.121*	0.003
	(0.007)	(0.003)	(0.053)	(0.002)	(0.005)	(0.005)	(0.056)	(0.002)
DE	-0.001	-0.001*	-0.009***	-0.000	-0.001	-0.001**	-0.006***	0.000
	(0.001)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Consumer Cyclicals	0.388*	-0.367**	3.079***	-0.260***	-0.002	-0.561***	2.314***	-0.318***
	(0.195)	(0.117)	(0.458)	(0.040)	(0.113)	(0.112)	(0.322)	(0.034)
Consumer Non-Cyclicals	0.626*	0.111	1.079	-0.169***	-0.016	-0.122	1.055*	-0.184***
	(0.272)	(0.150)	(0.668)	(0.041)	(0.132)	(0.142)	(0.513)	(0.037)
Energy	0.579	-0.114		-0.148**	-0.665**	-0.270		-0.092
	(0.439)	(0.202)		(0.053)	(0.209)	(0.178)		(0.053)
Financials	0.289	-0.342**	1.508**	-0.257***	-0.162	-0.447***	-0.052	-0.304***
	(0.213)	(0.120)	(0.549)	(0.041)	(0.112)	(0.112)	(0.486)	(0.035)
Healthcare	0.259	-0.096	1.560**	-0.258***	-0.275*	-0.217	0.540	-0.278***
	(0.193)	(0.108)	(0.582)	(0.042)	(0.114)	(0.113)	(0.665)	(0.038)
Industrials	0.534**	0.132	1.027*	-0.261***	0.435**	-0.038	1.267***	-0.296***
	(0.188)	(0.109)	(0.411)	(0.041)	(0.137)	(0.108)	(0.359)	(0.035)
Technology	0.463	-0.656*	1.313	-0.399***	-0.166	-0.701*	0.328	-0.366***
	(0.430)	(0.258)	(0.839)	(0.046)	(0.236)	(0.287)	(0.619)	(0.039)
Utilities	0.295	-0.387	3.023	-0.325***	0.267	-0.275	1.554	-0.379***
	(0.318)	(0.218)	(1.551)	(0.045)	(0.198)	(0.196)	(1.354)	(0.038)
constant	9.567*	3.892*	-1.807	-0.500	-1.084	1.520	4.241	0.657**
	(4.263)	(1.843)	(9.256)	(0.286)	(1.732)	(1.273)	(6.734)	(0.252)
Observations	7487	7508	6379	7487	7755	7760	6761	7755

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on a new propensity score matched data set for nationality. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board), no-CEO-appointment and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.57: Panel regression of firm performance with Hispanics evaluated as non-white (variable non-white_Cauc)[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:		Before				After		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-white_Cauc	-0.005	0.099	-0.018	0.049**	0.039	0.121	-0.137	0.044*
	(0.141)	(0.103)	(0.417)	(0.018)	(0.130)	(0.095)	(0.329)	(0.018)
month	-0.004	0.002	-0.001	-0.001	0.001	0.001	0.006	0.000
	(0.007)	(0.003)	(0.014)	(0.000)	(0.005)	(0.003)	(0.012)	(0.000)
non-white_Cauc × month	0.004	-0.004	0.013	0.001	-0.007	-0.006	-0.023	-0.000
	(0.008)	(0.004)	(0.016)	(0.001)	(0.006)	(0.004)	(0.016)	(0.001)
Appointment-specific control variables								
Female	0.175	0.322**	-0.097	0.055*	0.242	0.134	-0.274	0.053**
	(0.220)	(0.122)	(0.473)	(0.022)	(0.186)	(0.120)	(0.393)	(0.019)
non-white_Cauc × Female	-0.096	-0.383*	-0.083	-0.059	-0.260	-0.180	0.388	-0.063*
	(0.264)	(0.168)	(0.655)	(0.030)	(0.218)	(0.158)	(0.571)	(0.028)
Extern	-0.062	0.104	-0.206	0.017	-0.132	-0.051	-0.417	0.006
	(0.175)	(0.108)	(0.454)	(0.021)	(0.113)	(0.075)	(0.361)	(0.017)
Academic	-0.065	-0.005	0.088	-0.011	0.074	0.032	0.496	-0.026
	(0.157)	(0.109)	(0.316)	(0.018)	(0.165)	(0.108)	(0.303)	(0.018)
Grandeecole	-0.069	-0.156	-0.276	-0.014	-0.004	-0.072	-0.307	-0.019
	(0.149)	(0.119)	(0.300)	(0.017)	(0.119)	(0.086)	(0.261)	(0.015)
Age	0.015*	-0.002	0.046*	-0.001	-0.001	-0.006	0.015	-0.002*
	(0.007)	(0.005)	(0.020)	(0.001)	(0.005)	(0.004)	(0.017)	(0.001)
Jointappointment	0.062	0.035	-0.007	-0.006	0.054	0.052	0.381	0.000
	(0.116)	(0.076)	(0.306)	(0.014)	(0.100)	(0.077)	(0.273)	(0.013)
Supervisory board	-0.611***	-0.180	-0.637	0.076***	-0.483***	-0.197	-0.360	0.060**
	(0.144)	(0.100)	(0.393)	(0.018)	(0.145)	(0.110)	(0.384)	(0.020)
Management board	-0.613*	-0.122	-0.608	0.074*	-0.690***	-0.270*	-0.714	0.058*
	(0.258)	(0.180)	(0.682)	(0.033)	(0.182)	(0.135)	(0.543)	(0.027)
CEO	0.087	-0.125	-0.687	-0.045	0.050	-0.040	-1.142	-0.013
	(0.526)	(0.476)	(1.108)	(0.041)	(0.333)	(0.310)	(0.674)	(0.028)
Later CEO	0.373	0.031	2.985	-0.014	0.117	0.084	1.769*	-0.052
	(0.463)	(0.178)	(2.025)	(0.045)	(0.145)	(0.206)	(0.746)	(0.037)
Prior Boardsize	0.026	0.010	0.085	-0.005	0.001	0.015	0.077	0.001
	(0.033)	(0.021)	(0.083)	(0.004)	(0.020)	(0.015)	(0.065)	(0.003)
# Prior Females	-0.022	0.030	-0.306**	0.016***	0.012	-0.008	-0.023	0.002
	(0.040)	(0.024)	(0.103)	(0.004)	(0.028)	(0.020)	(0.075)	(0.004)
# Prior non-whites	0.036	-0.037	0.375*	-0.017*	0.106*	0.009	0.028	-0.017**
	(0.055)	(0.033)	(0.149)	(0.007)	(0.051)	(0.027)	(0.138)	(0.006)
Firm-specific control variables								
# employees	-0.001	-0.002**	0.001	-0.000	-0.002*	-0.002*	-0.001	0.000
	(0.001)	(0.001)	(0.003)	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)
ln(revenue)	-0.451	-0.234*	-0.482	0.039**	-0.146*	-0.200***	0.291	-0.011
	(0.247)	(0.118)	(0.553)	(0.015)	(0.069)	(0.044)	(0.220)	(0.010)
EBIT	0.136*	0.108***	0.162	0.007*	0.074*	0.083***	-0.142	0.012**
	(0.058)	(0.026)	(0.126)	(0.003)	(0.036)	(0.021)	(0.112)	(0.004)
DE	0.000	-0.000	-0.003	-0.000	-0.001	-0.000	-0.001	-0.000
	(0.001)	(0.001)	(0.004)	(0.000)	(0.001)	(0.001)	(0.005)	(0.000)
Consumer Cyclicals	0.712**	0.173	0.450	-0.130	-0.065	-0.203	1.261*	-0.150*
	(0.227)	(0.178)	(0.744)	(0.078)	(0.435)	(0.210)	(0.624)	(0.067)
Consumer Non-Cyclicals	0.484*	0.470*	-0.405	-0.023	-0.109	0.084	0.400	-0.021
	(0.219)	(0.187)	(0.720)	(0.079)	(0.400)	(0.228)	(0.662)	(0.070)
Energy	-0.381*	-0.579**	-0.551	-0.169*	0.000	0.000	0.000	0.000
	(0.178)	(0.191)	(0.663)	(0.081)	(.)	(.)	(.)	(.)
Financials	0.293	-0.517*	-0.827	-0.187*	-0.267	-0.510*	0.684	-0.202**
	(0.342)	(0.249)	(0.903)	(0.084)	(0.422)	(0.245)	(0.884)	(0.074)
Healthcare	-1.000*	-0.893**	-2.058	-0.159*	-1.059*	-0.569	0.297	-0.147
	(0.474)	(0.303)	(1.176)	(0.078)	(0.491)	(0.354)	(1.769)	(0.089)
Industrials	0.643*	0.314	-0.707	-0.095	0.225	0.041	0.289	-0.093
	(0.252)	(0.207)	(0.650)	(0.079)	(0.374)	(0.228)	(0.588)	(0.073)
Real Estate	0.018	0.082	-1.174	0.031	-0.017	-0.338	0.672	-0.061
	(0.551)	(0.344)	(1.219)	(0.083)	(0.481)	(0.276)	(0.847)	(0.072)
Technology	0.025	-1.048***	0.978	-0.247**	-0.436	-0.959**	2.887**	-0.276***
	(0.379)	(0.294)	(1.019)	(0.080)	(0.418)	(0.307)	(0.994)	(0.072)
Utilities	-0.032	-0.507	0.123	-0.189*	-0.472	-0.566	-1.720*	-0.205**
	(0.280)	(0.268)	(1.581)	(0.083)	(0.411)	(0.294)	(0.798)	(0.074)
constant	9.368	5.092*	9.447	-0.333	4.092*	4.997***	-7.477	0.828***
	(5.226)	(2.526)	(12.183)	(0.318)	(1.663)	(1.024)	(4.914)	(0.222)
Observations	4410	4419	4143	4410	4554	4563	4338	4554

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on a new propensity score matched data set for appearance measured by the variable non-white_Cauc. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board), no-CEO-appointment and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.58: Panel regression of firm performance (genetic matched data set of non-white)[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:		Before				After		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-white	-0.156	0.036	-0.526	0.007	0.089	-0.008	0.010	0.024
	(0.145)	(0.093)	(0.468)	(0.025)	(0.125)	(0.089)	(0.415)	(0.026)
month	0.003	0.002	0.022	0.001	-0.003	-0.003	-0.002	0.001
	(0.007)	(0.004)	(0.016)	(0.001)	(0.005)	(0.003)	(0.014)	(0.001)
non-white × month	0.003	-0.002	-0.016	-0.000	-0.003	-0.000	-0.008	-0.000
	(0.009)	(0.006)	(0.019)	(0.001)	(0.006)	(0.004)	(0.018)	(0.001)
Appointment-specific control variables								
Female	-0.278	-0.035	-0.530	0.001	-0.048	-0.146	-0.193	-0.002
	(0.157)	(0.112)	(0.473)	(0.025)	(0.146)	(0.116)	(0.402)	(0.026)
non-white × Female	0.257	-0.133	0.145	-0.025	-0.164	0.027	0.091	-0.036
	(0.199)	(0.167)	(0.561)	(0.029)	(0.194)	(0.146)	(0.658)	(0.032)
Extern	-0.503*	-0.169	-0.232	0.010	-0.308*	-0.050	0.045	0.002
	(0.222)	(0.121)	(0.588)	(0.031)	(0.156)	(0.102)	(0.469)	(0.022)
Academic	0.140	0.080	0.583	-0.019	0.100	0.075	0.573	-0.016
	(0.160)	(0.100)	(0.380)	(0.016)	(0.143)	(0.110)	(0.366)	(0.016)
Grandeecole	-0.092	-0.190	0.114	-0.042*	-0.023	0.004	-0.243	-0.002
	(0.132)	(0.101)	(0.384)	(0.021)	(0.114)	(0.074)	(0.372)	(0.020)
Age	0.020*	0.019**	0.003	-0.000	0.005	0.000	-0.010	0.000
	(0.010)	(0.006)	(0.023)	(0.001)	(0.006)	(0.004)	(0.024)	(0.001)
Jointappointment	0.168	0.150	-0.299	0.031	-0.092	0.105	-0.481	0.047*
	(0.121)	(0.099)	(0.376)	(0.022)	(0.150)	(0.101)	(0.352)	(0.021)
Supervisory board	-0.454*	-0.361*	-0.344	0.077***	-0.477*	-0.070	-0.065	0.069**
	(0.184)	(0.150)	(0.538)	(0.022)	(0.201)	(0.144)	(0.489)	(0.023)
Management board	-1.532***	-0.927***	-0.620	0.079	-1.278***	-0.725**	1.183	0.047
	(0.303)	(0.236)	(1.047)	(0.057)	(0.269)	(0.235)	(1.338)	(0.069)
Prior Boardsize	0.045	0.049*	0.214*	-0.012***	0.072*	0.015	0.162	-0.009*
	(0.025)	(0.020)	(0.088)	(0.003)	(0.029)	(0.018)	(0.084)	(0.004)
# Prior Females	0.042	0.057*	-0.209	0.017**	0.022	0.023	0.073	0.008
	(0.031)	(0.026)	(0.128)	(0.005)	(0.033)	(0.024)	(0.128)	(0.006)
# Prior non-whites	-0.013	-0.020	0.104	0.002	0.042	0.053	-0.294*	0.000
	(0.054)	(0.042)	(0.206)	(0.009)	(0.064)	(0.047)	(0.129)	(0.010)
Firm-specific control variables								
# employees	-0.002*	-0.001	0.002	0.000	-0.002	-0.001	0.000	0.000
	(0.001)	(0.001)	(0.002)	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)
ln(revenue)	-0.035	-0.336**	0.546	0.040**	-0.109	-0.122*	0.226	0.008
	(0.114)	(0.111)	(0.492)	(0.014)	(0.082)	(0.049)	(0.221)	(0.014)
EBIT	-0.035	0.075*	-0.225*	0.007	-0.005	0.038	-0.235	0.007
	(0.038)	(0.030)	(0.108)	(0.005)	(0.064)	(0.051)	(0.156)	(0.008)
DE	-0.002	-0.003	-0.018***	-0.000	-0.002	-0.001	-0.010*	-0.000
	(0.003)	(0.002)	(0.005)	(0.000)	(0.001)	(0.001)	(0.004)	(0.000)
Consumer Cyclicals	0.303	-0.223	1.323	-0.381***	-0.731	-0.434**	1.645**	-0.295***
	(0.217)	(0.142)	(0.727)	(0.023)	(0.422)	(0.142)	(0.584)	(0.073)
Consumer Non-Cyclicals	-0.182	-0.156	0.124	-0.283***	-0.992*	-0.208	1.163	-0.170*
	(0.175)	(0.136)	(0.603)	(0.027)	(0.417)	(0.157)	(0.604)	(0.074)
Healthcare	-0.147	-0.871**	0.262	-0.388***	-0.879	0.057	3.124*	-0.160
	(0.375)	(0.320)	(0.930)	(0.044)	(0.695)	(0.454)	(1.262)	(0.099)
Industrials	0.315	-0.007	-0.030	-0.330***	-0.538	-0.196	0.602	-0.221**
	(0.212)	(0.151)	(0.533)	(0.027)	(0.425)	(0.174)	(0.566)	(0.073)
Real Estate	0.427	0.065	0.333	-0.114	-0.290	0.007	-0.406	-0.086
	(0.366)	(0.377)	(1.497)	(0.062)	(0.522)	(0.239)	(1.192)	(0.095)
Technology	-0.500**	-1.082***	-0.146	-0.442***	-1.132*	-0.930***	2.023	-0.352***
	(0.167)	(0.245)	(0.562)	(0.036)	(0.440)	(0.279)	(1.417)	(0.081)
Utilities	-0.491	-0.837***	-1.864**	-0.451***	-1.004*	-0.455*	-1.961**	-0.360***
	(0.282)	(0.218)	(0.619)	(0.035)	(0.470)	(0.213)	(0.620)	(0.078)
constant	0.333	6.663**	-12.392	-0.112	3.241	3.069**	-5.514	0.484
	(2.540)	(2.415)	(10.786)	(0.309)	(1.841)	(1.044)	(5.269)	(0.306)
Observations	2688	2688	2517	2688	2802	2803	2680	2802

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on a genetic matched data set for appearance. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board) and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Table Appendix V.59: Panel regression of firm performance (genetic matched data set of non-French)[a]

Dependent variable:	CRR	CMAR	CRAR	Beta	CRR	CMAR	CRAR	Beta
Modus:		Before				After		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-French	0.059	0.147	-0.037	0.013	0.213	0.144	0.153	0.000
	(0.182)	(0.082)	(0.320)	(0.014)	(0.117)	(0.076)	(0.348)	(0.014)
month	-0.012*	-0.004*	-0.005	-0.001**	-0.003	-0.001	0.002	0.000
	(0.005)	(0.002)	(0.009)	(0.000)	(0.003)	(0.002)	(0.009)	(0.000)
non-French × month	0.001	0.003	0.002	0.001	-0.008*	-0.005*	-0.016	0.000
	(0.007)	(0.003)	(0.012)	(0.000)	(0.004)	(0.002)	(0.012)	(0.000)
Appointment-specific control variables								
Female	-0.003	0.139*	-0.518*	0.030*	0.131	0.082	-0.210	0.014
	(0.144)	(0.070)	(0.238)	(0.015)	(0.109)	(0.069)	(0.279)	(0.014)
non-French × Female	0.047	-0.045	0.130	0.002	-0.042	-0.062	0.293	0.006
	(0.220)	(0.099)	(0.350)	(0.019)	(0.144)	(0.094)	(0.384)	(0.019)
Extern	0.171	-0.066	0.657**	-0.026	-0.287**	-0.162**	0.080	-0.006
	(0.212)	(0.075)	(0.225)	(0.014)	(0.108)	(0.062)	(0.280)	(0.012)
Academic	0.071	-0.064	-0.275	-0.001	0.067	-0.075	-0.027	-0.008
	(0.155)	(0.067)	(0.191)	(0.011)	(0.090)	(0.060)	(0.215)	(0.010)
Grandeecole	0.089	0.099	-0.117	-0.004	0.130	0.072	-0.064	0.001
	(0.131)	(0.066)	(0.255)	(0.012)	(0.081)	(0.058)	(0.276)	(0.011)
Age	0.005	-0.001	0.011	-0.001	0.014**	-0.001	0.027*	-0.002*
	(0.007)	(0.003)	(0.012)	(0.001)	(0.005)	(0.003)	(0.014)	(0.001)
Jointappointment	-0.486**	-0.147**	-0.212	-0.006	-0.090	-0.039	0.159	-0.008
	(0.173)	(0.054)	(0.202)	(0.010)	(0.081)	(0.054)	(0.206)	(0.010)
Supervisory board	0.105	-0.020	-0.755**	0.033*	-0.276**	-0.149*	-0.113	0.023
	(0.232)	(0.073)	(0.255)	(0.013)	(0.103)	(0.066)	(0.270)	(0.013)
Management board	0.299	-0.186	-0.501	-0.017	-0.187	-0.209	0.308	-0.004
	(0.627)	(0.147)	(0.343)	(0.018)	(0.199)	(0.114)	(0.474)	(0.018)
CEO	-0.200	-0.247	-0.091	-0.027	-0.210	-0.455***	0.084	-0.026
	(0.240)	(0.144)	(0.596)	(0.030)	(0.176)	(0.134)	(0.554)	(0.021)
Later CEO	-0.092	-0.150	1.098	-0.061	0.010	-0.393	1.121	-0.069
	(0.408)	(0.182)	(1.485)	(0.032)	(0.306)	(0.264)	(1.854)	(0.062)
Prior Boardsize	-0.012	-0.008	-0.008	0.001	0.018	0.016*	-0.034	0.006***
	(0.018)	(0.009)	(0.032)	(0.001)	(0.012)	(0.008)	(0.034)	(0.001)
# Prior Females	-0.069*	0.010	-0.112*	0.014***	-0.003	-0.013	0.026	0.005*
	(0.028)	(0.013)	(0.053)	(0.002)	(0.018)	(0.013)	(0.048)	(0.002)
# Prior non-Frenchs	-0.131***	-0.077***	0.150**	-0.011***	0.018	-0.034**	0.183***	-0.011***
	(0.033)	(0.012)	(0.047)	(0.002)	(0.022)	(0.012)	(0.046)	(0.002)
Firm-specific control variables								
# employees	-0.002**	-0.001*	0.001	0.000	-0.000	-0.001**	0.003	0.000
	(0.001)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)
ln(revenue)	-0.081	-0.018	0.347	0.030*	-0.407***	0.015	-1.211***	0.019**
	(0.183)	(0.071)	(0.308)	(0.012)	(0.088)	(0.048)	(0.308)	(0.007)
EBIT	0.012	0.014***	-0.067	0.003	0.032***	0.000	0.081	0.008***
	(0.008)	(0.004)	(0.044)	(0.002)	(0.009)	(0.007)	(0.052)	(0.002)
DE	-0.002*	-0.001	-0.004*	-0.000	-0.001	-0.001**	-0.001	-0.000*
	(0.001)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)
Consumer Cyclicals	-0.020	-0.456***	2.145***	-0.198***	-0.095	-0.591***	2.031***	-0.224***
	(0.224)	(0.105)	(0.358)	(0.032)	(0.146)	(0.107)	(0.329)	(0.030)
Consumer Non-Cyclicals	0.130	-0.143	0.485	-0.124***	-0.150	-0.279**	1.131**	-0.108***
	(0.219)	(0.101)	(0.406)	(0.035)	(0.135)	(0.106)	(0.409)	(0.032)
Energy	1.207	-0.416**	0.118	-0.194***	-0.096	-0.666***	0.815	-0.229***
	(0.627)	(0.140)	(0.316)	(0.037)	(0.244)	(0.182)	(0.825)	(0.041)
Financials	0.079	-0.595***	0.652	-0.226***	-0.246	-0.648***	1.188*	-0.273***
	(0.233)	(0.119)	(0.480)	(0.036)	(0.154)	(0.117)	(0.524)	(0.033)
Healthcare	-0.036	-0.191	1.979***	-0.207***	-0.340*	-0.318**	1.603*	-0.235***
	(0.205)	(0.108)	(0.526)	(0.035)	(0.135)	(0.121)	(0.662)	(0.034)
Industrials	0.258	-0.136	0.471	-0.182***	0.544**	-0.163	1.651***	-0.189***
	(0.226)	(0.117)	(0.336)	(0.034)	(0.167)	(0.106)	(0.341)	(0.031)
Real Estate	0.508	0.324	1.231	-0.011	-0.100	0.073	-0.712	-0.135***
	(0.496)	(0.172)	(0.767)	(0.045)	(0.198)	(0.147)	(0.623)	(0.037)
Technology	-0.516*	-1.468***	2.006***	-0.293***	-0.430**	-1.262***	2.018***	-0.291***
	(0.220)	(0.155)	(0.453)	(0.032)	(0.147)	(0.151)	(0.495)	(0.030)
Utilities	-0.210	-0.890***	0.300	-0.263***	-0.312	-0.904***	0.491	-0.325***
	(0.289)	(0.150)	(0.689)	(0.039)	(0.168)	(0.138)	(0.580)	(0.034)
constant	3.193	1.299	-8.650	-0.072	9.209***	0.383	25.812***	0.179
	(3.900)	(1.547)	(6.800)	(0.265)	(1.996)	(1.075)	(6.951)	(0.164)
Observations	12326	12350	10971	12326	12538	12566	11331	12538

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

[a] The models (1)-(4) describe panel regressions with 24 observations in the months up to two years prior to an appointment. The models differ in the measures of firm performance they use as dependent variable. The models (5)-(8) repeat the panel regressions with data 24 months after an appointment. All models are based on a genetic matched data set for nationality. Reference categories are unitary board (compared to supervisory board and management board in case of a dual board), no-CEO-appointment and Basic Materials (for firm industries). Robust standard errors are reported in parentheses.

Bibliography

Adams, Renée B. (2017), "Boards, and the Directors who sit on them", in *The Handbook of the Economics of Corporate Governance*, 1, 291–382: Elsevier.

Adams, Renée B. and Daniel Ferreira (2009), "Women in the boardroom and their impact on governance and performance", *Journal of Financial Economics*, 94 (2), 291–309.

Adams, Susan M., Atul Gupta, and John D. Leeth (2009), "Are Female Executives Overrepresented in Precarious Leadership Positions?", *British Journal of Management*, 20 (1), 1–12.

Aelenei, Christina, Yvette Assilaméhou-Kunz, Vincenzo Iacoviello, and Clara Kulich (2020), "The political glass cliff: When left-wing orientation leads to minority candidate choices for hard-to-win seats", *European Review of Applied Psychology*, 70 (3).

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.

Ahmadi, Ali and Abdelfettah Bouri (2017), "Board of Directors' Composition and Performance in French CAC 40 Listed Firms", *Accounting*, 3 (4), 245–256.

Ahmadi, Ali, Nejia Nakaa, and Abdelfettah Bouri (2018), "Chief Executive Officer attributes, board structures, gender diversity and firm performance among French CAC 40 listed firms", *Research in International Business and Finance*, 44, 218–226.

Akaike, Hirotugu (1974), "A new look at the statistical model identification", *IEEE Transactions on Automatic Control*, 19 (6), 716–723.

Algan, Yann, Christian Dustmann, Albrecht Glitz, and Alan Manning (2010), “The Economic Situation of First and Second-Generation Immigrants in France, Germany and the United Kingdom”, *The Economic Journal*, 120 (542), F4–F30.

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).

Assemblée nationale (2011), “LOI n° 2011-103 du 27 janvier 2011 relative à la représentation équilibrée des femmes et des hommes au sein des conseils d’administration et de surveillance et à l’égalité professionnelle (1)”, `https://www.legifrance.gouv.fr/loda/id/JORFTEXT000023487662`.

Baron, Reuben M. and David A. Kenny (1986), “The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations”, *Journal of Personality and Social Psychology*, 51 (6), 1173–1182.

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.

Bechtoldt, Myriam N., Christina E. Bannier, and Björn Rock (2019), “The glass cliff myth? – Evidence from Germany and the U.K.”, *The Leadership Quarterly*, 30 (3), 273–297.

Bell, Ella and Stella Nkomo (2001), *Our Separate Ways: Black and White Women and the Struggle for Professional Identity*: Harvard Business Review Press, Boston.

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.

Benkraiem, Ramzi, Amal Hamrouni, Faten Lakhal, and Nadia Toumi (2017), “Board independence, gender diversity and CEO compensation”, *Corporate Governance: The International Journal of Business in Society*, 17 (5), 845–860.

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.

Blumenstock, Robin, Michael Callen, and Tarek Ghani (2018), “Why Do Defaults Affect Behavior? Experimental Evidence from Afghanistan”, *American Economic Review*, 108 (10), 2868–2901.

Boyd, Karen S. (2008), “Glass ceiling”, in Schaefer, Richard T ed. *Encyclopedia of race, ethnicity, and society*, 549–552, Thousand Oaks: Sage.

Brady, David, Katelin Isaacs, Martha Reeves, Rebekah Burroway, and Megan Reynolds (2011), “Sector, size, stability, and scandal: Explaining the presence of female executives in Fortune 500 firms”, *Gender in Management: An International Journal*, 26.

Brinkhuis, Eline and Bert Scholtens (2018), “Investor response to appointment of female CEOs and CFOs”, *The Leadership Quarterly*, 29 (3), 423–441.

Bruckmüller, Susanne and Nyla R. Branscombe (2010), “The glass cliff: When and why women are selected as leaders in crisis contexts”, *British Journal of Social Psychology*, 49 (3), 433–451.

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.

Carter, David A., Frank D'Souza, Betty J. Simkins, and W. Gary Simpson (2010), “The Gender and Ethnic Diversity of US Boards and Board Committees and Firm Financial Performance”, *Corporate Governance: An International Review*, 18 (5), 396–414.

Carton, Andrew and Ashleigh Rosette (2011), “Explaining Bias against Black Leaders: Integrating Theory on Information Processing and Goal-Based Stereotyping”, *Academy of Management Journal*, 54, 1141–1158.

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.

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.

Cheema-Fox, Alex, Bridget R. LaPerla, Hui S. Wang, and George Serafeim (2021), “Corporate Resilience and Response to COVID-19”, *Journal of Applied Corporate Finance*, 33 (2), 24–40.

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.

Chung, Beth and Melenie Lankau (2005), "Are We There Yet? An Assessment of Fit Between Stereotypes of Minority Managers and the Successful-Manager Prototype", *Journal of Applied Social Psychology*, 35, 2029 – 2056.

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.

Cook, Alison and Christy Glass (2013), "Women and Top Leadership Positions: Towards an Institutional Analysis", *Gender, Work & Organization*, 21 (1), 91–103.

——— (2014), "Above the glass ceiling: When are women and racial/ethnic minorities promoted to CEO?", *Strategic Management Journal*, 35 (7), 1080–1089.

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).

Dang, Rey and Duc Khuong Nguyen (2018), "Does Board Gender Diversity Make a Difference? New Evidence from Quantile Regression Analysis", *Management international*, 20 (2), 95–106.

Dang, Rey, Duc Khuong Nguyen, and Linh-Chi Vo (2014), "Does The Glass Ceiling Exist? A Longitudinal Study Of Womens Progress On French Corporate Boards", *Journal of Applied Business Research (JABR)*, 30 (3), 909–916.

DellaVigna, Stefano (2009), "Psychology and Economics: Evidence from the Field", *Journal of Economic Literature*, 47 (2), 315–372.

Deloitte (2021), "Missing Pieces Report: The Board Diversity Census of Women and Minorities on Fortune 500 Boards", `https://www2.deloitte.com/content/dam/Deloitte/us/Documents/center-for-board-effectiveness/missing-pieces-fortune-500-board-diversity-study-6th-edition-report.pdf`.

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.

Diamond, Alexis and Jasjeet S. Sekhon (2013), “Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies”, *Review of Economics and Statistics*, 95 (3), 932–945.

Dignan, Don (1981), “Europe’s melting pot: A century of large-scale immigration into France”, *Ethnic and Racial Studies*, 4 (2), 137–152.

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.

Elliott, James R. and Ryan A. Smith (2004), “Race, Gender, and Workplace Power”, *American Sociological Review*, 69 (3), 365–386.

Elsaid, Eahab and Nancy D. Ursel (2018), “Re-examining the glass cliff hypothesis using survival analysis: The case of female CEO tenure”, *British Journal of Management*, 29 (1), 156–170.

Enke, Benjamin and Thomas Graeber (2021a), “Cognitive Uncertainty”, NBER Working Paper.

——— (2021b), “Cognitive Uncertainty in Intertemporal Choice”, Working Paper.

Evans, David P. (2014), “Aspiring to Leadership… A Woman's World?”, *Procedia - Social and Behavioral Sciences*, 148, 543–550.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde (Forthcoming), “The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences”, *Management Science*.

Fama, Eugene F. (1970), “Efficient capital markets: A review of theory and empirical work”, *Journal of Finance*, 25, 383–417.

Fama, Eugene F., Lawrence Fisher, Michael C. Jensen, and Richard Roll (1969), “The adjustment of stock prices to new information”, *International Economic Review*, 10, 1–21.

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.

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.

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.

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.

Glass, Christy and Alison Cook (2016), "Leading at the top: Understanding women's challenges above the glass ceiling", *The Leadership Quarterly*, 27 (1), 51–63.

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.

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.

Haslam, S. Alexander and Michelle K. Ryan (2008), “The Road to the Glass Cliff: Differences in the Perceived Suitability of Men and Women for Leadership Positions in Succeeding and Failing Organizations”, *The Leadership Quarterly*, 19, 530–546.

Haslam, S. Alexander, Michelle K. Ryan, Clara Kulich, Grzegorz Trojanowski, and Cate Atkins (2009), “Investing with Prejudice: the Relationship Between Women's Presence on Company Boards and Objective and Subjective Measures of Company Performance”, *British Journal of Management*, 21 (2), 484–497.

Hastings Jr., Cecil, Frederick Mosteller, John W. Tukey, and Charles P. Winsor (1947), “Low Moments for Small Samples: A Comparative Study of Order Statistics”, *The Annals of Mathematical Statistics*, 18 (3), 413 – 426.

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.

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.

Higgs, Derek et al. (2003), *Review of the role and effectiveness of non-executive directors*: Stationery Office, London.

Hill, A. (2016), ""Glass cliffs" and the female leaders who are set up to fail", *Financial Times*.

Hoffrage, Ulrich, Samuel Lindsey, Ralph Hertwig, and Gerd Gigerenzer (2000), "Communicating Statistical Information", *Science*, 290 (5500), 2261–2262.

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.

IMF (2021), "World Economic Outlook Database", `https://www.imf.org/en/Publications/WEO/weo-database/2021/October/weo-report?c=171,193,122,124,423,935,128,939,172,132,134,174,176,178,136,941,946,137,181,138,142,182,135,936,961,184,144,146,112,&s=NGDPD,&sy=2017&ey=2022&ssm=1&scsm=1&scc=1&ssd=1&ssc=0&sic=0&sort=country&ds=.&br=1`.

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.

Judge, Elizabeth (2003), "Women on board: Help or hindrance", *The Times*, 11 (21), 543–562.

Jungmann, Carsten (2006), "The Effectiveness of Corporate Governance in One-Tier and Two-Tier Board Systems – Evidence from the UK and Germany", 3 (4), 426–474.

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.

Kaplan, Steven N. and Bernadette A. Minton (1994), "Appointments of outsiders to Japanese boards: Determinants and implications for managers", *Journal of Financial Economics*, 36 (2), 225–258.

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.

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.

Khatib, Saleh, Dewi Earnest Abdullah, Ahmed Elamer, Ibrahim Yahaya, and Andrews Owusu (Forthcoming), "Global trends in board diversity research: A bibliometric view", *Meditari Accountancy Research*.

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.

Krawiec, Kimberly D. and Lissa Lamkin Broome (2009), "Signaling Through Board Diversity: Is Anyone Listening?", *University of Cincinnati Law Review*, 77 (1), 431–464.

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.

Kulich, Clara, Fabio Lorenzi-Cioldi, Vincenzo Iacoviello, Klea Faniko, and Michelle K. Ryan (2015), "Signaling change during a crisis: Refining conditions for the glass cliff", *Journal of Experimental Social Psychology*, 61, 96–103.

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.

Lintner, John (1965), "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets", *The Review of Economics and Statistics*, 47 (1), 13–37.

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.

Maclean, Mairi and Charles Harvey (2008), "Women on corporate boards of directors: The French perspective", in Kardes, Frank R., Paul M. Herr, and Norbert Schwarz eds. *Women on Corporate Boards of Directors: International Research and Practice*, Chap. 4: Edward Elgar Publishing, Cheltenham.

Maume, David J. (1999), “Glass Ceilings and Glass Escalators”, *Work and Occupations*, 26 (4), 483–509.

——— (2012), “Minorities in Management: Effects on Income Inequality, Working Conditions, and Subordinate Career Prospects among Men”, *The ANNALS of the American Academy of Political and Social Science*, 639 (1), 198–216.

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.

McGuire, Gail M. and Barbara F. Reskin (1993), “Authority Hierarchies at Work: The Impacts of Race and Sex”, *Gender & Society*, 7 (4), 487–506.

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.

Moulin, Yves and Sébastien Point (2012), “Les femmes dans les conseils d’administration du SBF120: qualités féminines ou affaires de famille?”, *Revue de gestion des ressources humaines*, 1 (83), 31–44.

Mulcahy, Mark and Carol Linehan (2013), “Females and Precarious Board Positions: Further Evidence of the Glass Cliff”, *British Journal of Management*, 25 (3), 425–438.

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.

Nekhili, Mehdi and Hayette Gatfaoui (2012), “Are Demographic Attributes and Firm Characteristics Drivers of Gender Diversity? Investigating Women’s Positions on French Boards of Directors”, *Journal of Business Ethics*, 118 (2), 227–249.

Nielsen, Kirby and John Rehbeck (2022), “When Choices are Mistakes”, *American Economic Review*, 112 (7), 2237–2268.

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.

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.

Parker, John (2017), “A Report into the Ethnic Diversity of UK Boards”, Independent Report.

Paternoster, Ray, Robert Brame, Paul Mazerolle, and Alex Piquero (1998), “Using the Correct Statistical Test for Equality of Regression Coefficients”, *Criminology*, 36, 859 – 866.

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.

Ponce, Alejandro, Enrique Seira, and Guillermo Zamarripa (2017), “Borrowing on the Wrong Credit Card? Evidence from Mexico”, *American Economic Review*, 107 (4), 1335–1361.

Powell, Gary N. and D. Anthony Butterfield (2002), “Exploring the Influence of Decision Makers' Race and Gender on actual Promotions to Top Management”, *Personnel Psychology*, 55 (2), 397–428.

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.

Rosenbaum, Paul R. and Donald B. Rubin (1985), "Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score", *The American Statistician*, 39 (1), 33–38.

Rosette, Ashleigh, Geoffrey Leonardelli, and Katherine Phillips (2008), "The White Standard: Racial Bias in Leader Categorization", *The Journal of applied psychology*, 93 (4), 758–77.

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.

Ruigrok, Winfried, Simon Peck, and Sabina Tacheva (2007), "Nationality and gender diversity on Swiss corporate boards", *Corporate governance: an international review*, 15 (4), 546–557.

Ryan, Michelle K. and S. Alexander Haslam (2005), "The Glass Cliff: Evidence that Women are Over-Represented in Precarious Leadership Positions", *British Journal of Management*, 16 (2), 81–90.

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.

Santee, Jennifer, Kylie Barnes, Nancy Borja-Hart, Cheng An-Lin, Juanita Draime, Akesha Edwards, Nkem Nonyel, and Mark Sawkin, "Correlation Between Pharmacy Students' Implicit Bias Scores, Explicit Bias Scores, and Responses to Clinical Cases.", *American Journal of Pharmaceutical Education*, 86 (1), 57–61.

Schwarz, Gideon (1978), "Estimating the Dimension of a Model", *The Annals of Statistics*, 6 (2), 461–464.

Sekhon, Jasjeet Singh and Walter R. Mebane Jr. (1998), "Genetic Optimization Using Derivatives: Theory and Application to Nonlinear Models", *Political Analysis*, 7, 187–210.

Sharpe, William F. (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk", *The Journal of Finance*, 19 (3), 425–442.

Sherman, Mila Getmansky and Heather E. Tookes (2022), "Female Representation in the Academic Finance Profession", *The Journal of Finance*, 77 (1), 317–365.

Sims, Christopher A. (2003), "Implications of rational inattention", *Journal of Monetary Economics*, 50 (3), 665–690.

Singh, Val (2007), "Ethnic diversity on top corporate boards: a resource dependency perspective", *The International Journal of Human Resource Management*, 18 (12), 2128–2146.

——— (2015), "Legitimacy profiles of women directors on top French company boards", *Journal of Management Development*, 34, 803–820.

Singh, Val, Sébastien Point, and Yves Moulin (2015), "French supervisory board gender composition and quota threat: changes from 2008 to 2010", *Gender in Management: An International Journal*, 30 (7), 551–571.

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.

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.

Tobin, James and William C. Brainard (1977), “Assets Markets and the Cost of Capital, Economic Progress, Private Values and Public Policy”, *Cowles Foundation Paper*, 440.

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.

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.

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.

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.

Signature author   Date

Submitted Papers:

1. Gärtner, Florian, Darwin Semmler, Christina E. Bannier (2023), "What could possibly go wrong? Predictable misallocation in simple debt repayment experiments", *Journal of Economic Behavior & Organization*, 205, 28-43, ISSN 0167-2681, https://doi.org/10.1016/j.jebo.2022.10.032 (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). "Elemental Financial Decisions". Working Paper (Chapter IV)