Essays in Development Economics

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

vorgelegt von

Matthias Stelter

Frühjahrs-/Sommersemester 2021

Abteilungssprecher	Prof. Volker Nocke, Ph.D.
Referent	Prof. Dr. Markus Frölich
Koreferent	Dr. Alexandra Avdeenko

Tag der Verteidigung	29.07.2021

Acknowledgments

I am indebted to my supervisors Markus Frölich and Alexandra Avdeenko, for all their guidance and support. They gave me the opportunity to conduct empirical studies in the context of development economics, and without their help and support, this dissertation would not exist. Despite the troubles involved in such research and several failed projects, I am grateful for each project and their efforts to provide me with opportunities to gain experience in field research. In particular, I am thankful to Alexandra Avdeenko, who gave me direction during a critical time when I was about to give up.

I thank the University of Mannheim's Graduate School of Economic and Social Sciences (GESS) for supporting all of this work.

The work in chapters 1, 2, and 3 is build on the shoulders of many research assistants in Mannheim and Zambia. I thank all of those who were involved over the years and made this project possible. The authors thank the International Initiatives for Impact Evaluation (3ie) and Sonderforschungsbereich (SFB) 884 (funded by the Deutsche Forschungsgemeinschaft (DFG)) for funding this project. Further, chapter 2 benefited from discussions with Clementine Sadania.

For chapter 4, I thank Natascha Haitz and Alexandra Avdeenko for their feedback and discussions that greatly contributed to the chapter. Further, I am particularly grateful for the interviewers in Zambia who implemented the project and whose perspective and feedback improved the study. I thank the Sonderforschungsbereich (SFB) 884 for funding the data collection used in this study, as well as Markus Frölich for his support.

For chapter 5, Alexandra Avdeenko and I thank the Center for Evaluation and Development (C4ED), Mannheim Germany, for financially supporting this research. Further, we thank Daniel Bruns, Paula Navarro, and Katharina Kaeppel for their research assistance on this project, as well as the Center for Development and Evaluation, in particular Sharafat Hussain, for the feedback on the tools, the support with the clearances, and all other necessary data collection preparations. We thank Human Design Studios Pakistan for the development of the video. We thank Markus Frölich for supporting this research and Dean Karlan for providing valuable feedback during the preparation of this study.

Finally, I thank my fellow students and friends. Claudia, your door was always open, and I believe our talks were crucial for my mental state. Your contribution to my dissertation is not reflected in these pages but in the time I saved, not chasing after every stupid idea. Thank you for being an essential part of this journey. Raphael, talking with you made doing research and being a Ph.D. student more interesting. I learned that a critical ingredient for research is the motivation to continue and follow through. Thank you for providing some of it. Fabian, I enjoyed the time we shared an office and I am grateful to you for pointing me towards development economics. Without you, I might not have come to believe that there is science in economics and that it can be found in development economics. Last but not least, I thank all my friends who greatly improved my doctoral student experience and life in Mannheim and Berkeley.

There are many omissions that I did not forget. But the one thing I learned during my Ph.D. is that completeness and perfection are not ideals but illusions not worth pursuing.

Contents

List of Figures

List of Tables

Introduction

My dissertation is about impact evaluation in the sphere of development and behavioral interventions. The first part of my dissertation consists of impact evaluations; the second part is about data collection and research ethics. Rigorous impact evaluations allow policymakers to make better, evidence-based decisions on policies to improve people's lives. These tools are specifically suited in the context of the Sustainable Development Goals, which outline clear and measurable targets, and thus find increasingly wide adoption, particularly in development economics. The randomized-controlled trial (RCT) is often quoted to be the gold standard of empirical scientific inquiry. The Nobel prize committee acknowledged this in 2019 when they granted the Prize in Economic Sciences to Abhijit Banerjee, Esther Duflo, and Michael Kremer "for their experimental approach to alleviating global poverty." In chapters 2 and 3, we conduct impact evaluations based on an RCT in Zambia, which is outlined in chapter 1. In most cases, RCTs require primary data collections. Thus data collections are an integral part of development economics, as it is practiced today, and hence deserves the attention of development economists. Throughout my studies, I was involved in several data collections, during which I was confronted with issues related to research ethics and data quality. Chapters 4 and 5 present experimental studies that aim to address some of these issues and provide an empirical evidence base for further discussions.

The RCT described in chapter 1 consists of two cross-randomized interventions which are only loosely related and therefore evaluated separately in chapters 2 and 3. In 2015, Markus Frölich and Niels Kemper successfully applied to funds by the International Initiative for Impact Evaluation (3ie) to launch the project in cooperation with the Rural Finance Expansion Programme in Zambia (RUFEP). Further financial support was provided by the Deutsche Forschungsgemeinschaft (DFG) through the Sonderforschungsbereich (SFB) 884, in particular for the phone surveys. Since 2016, P. Linh Nguyen managed and supervised the implementation of the RCT and related data collections, supported by a team of research assistants. A pre-analysis plan for this project was prepared by Markus Frölich, Andreas Landmann, and P. Linh Nguyen in 2018 with the support of research assistants. The study is registered on the AEA RCT registry as AEARCTR-0002640 (https://doi.org/10.1257/rct.2640-1.0). Frölich and Nguyen (2020) report the initial findings of the RCT. While I only joined the project after the endline data collection, I conducted three phone surveys for this project in April and November 2020 and Mai 2021. The impact evaluations discussed in chapters 2 and 3 built on the work and support of all those involved in the past. Chapter 2 is joint work with Markus Frölich and P. Linh Nguyen; and chapter 3 is joint work with Markus Frölich, Andreas Landmann, and P. Linh Nguyen.

In chapter 2, we evaluate the impact of an intervention aimed at facilitating the linkage between savings groups and formal financial institutions in rural Zambia. Four different NGOs implemented the intervention with support from RUFEP across seven districts in three regions of Zambia (for details on this background, see chapter 1). The goal of the intervention was to increase the uptake of bank or mobile money accounts by savings groups to provide them with safer storage technology and eventually increase their access to formal financial services such as loans. We find that the intervention increased uptake and usage of bank accounts for two of the four NGOs. However, we do not find effects on outcomes further down the theory of change, i.e., we find no evidence of changes in saving and lending activities nor welfare-related outcomes.

The evaluation of the second intervention is presented in chapter 3. This intervention was aimed at strengthening the savings groups' internal insurance mechanism. The savings groups in our study usually maintain something called a social fund, i.e., a fund to which each member contributes and which eventually is used for purposes that benefit the group beyond saving and lending. Often these funds are used to support group members in need. The social fund intervention aimed to strengthen this support by explaining concepts related to insurance and its benefits to the savings groups. Again, this intervention showed only a few results. We find some changes related to the social fund and risk-coping for one NGO, but the evidence is limited.

Chapter 1 provides a detailed overview of all the data collections conducted for the impact evaluations discussed in chapters 2 and 3. During the endline data collection, respondents were paid compensation. Our interviewers reported that some group members believed that this compensation was unfairly distributed, causing disputes in some savings groups. Group members reportedly thought it was unfair that some members were selected for an interview receiving the compensation, while others were not. I implemented a survey experiment to study fairness perceptions about survey participation compensation and lotteries to investigate this issue. During the phone surveys after the endline, we conducted an experiment to investigate how compensation affects the willingness to participate in future survey waves. For this purpose, we compensated respondents with a lottery. I took this as an opportunity to study perception related to the fairness of (unequally distributed) survey participation compensation. This study is presented in chapter 4. I find that most participants appreciate survey participation compensation, but that all, and not only some participants should benefit. Further, they reported that lotteries are fair, but this opinion is considerably affected by the outcome of the lottery.

Since I conducted and supported several survey data collections over the time of my studies, I became sensitized to the importance of data quality. When working with primary data, one realizes how important this research input is and that there is too little attention on data quality both in an economist's education and many studies. Good quality of survey data relies on, and research ethics dictate, the voluntary cooperation of study participants. While, at least for survey data-based studies without coercion, we might assume implicit consent of study participants, the topic of informed consent is no trivial matter. With the rise of big data and ever more powerful computers and statistical tools, data protection came to the forefront of policymakers' minds around the globe. Today, we understand that explicit and informed consent is essential. But it is not trivial to inform and be informed when there are so many unknowable consequences.

In a field experiment with Alexandra Avdeenko, discussed in chapter 5, we study approaches to improve the informed consent process. This study is registered at the AEA RCT registry under AEARCTR-0006829 (https://doi.org/10.1257/rct.6829-1.0). Alongside a large survey data collection in rural Pakistan, we experimentally test to augment the standard approach to obtaining informed consent. Usually, information about the purpose, use, and storage of survey data and the rights of the participants is presented as a text. This text is read by potential survey participants or read out by the interviewer if the potential participant is illiterate. We augment this process in two ways: The first approach shows the potential study participants a video that illustrates how their data is used if they consent to participate in the interview. In addition to the video, the second approach makes the standard process interactive. Instead of only reading the information, the potential participant is asked questions about the information. Depending on her answers, relevant information might be repeated. We study the effects of these two augmentations on the consent rate, the informedness of the respondents, and measures of data quality. We find that the second approach successfully increased the understanding that participation is voluntary according to our objective measure. However, it decreased overall understanding as subjectively reported across several aspects. Given that the subjectively reported understanding is very high and unrelated to our objective measures, it is unclear whether this is a worsening or an improvement. Further, we find that neither approach affects the consent rate, and we find no effects on item non-response rates, alleviating concerns that such approaches might reduce data quality or the representativeness of the survey.

The remainder of my dissertation is structured as follows. Chapter 1 provides background for the RCT conducted in Zambia and related data collections. It further outlines the main empirical strategy for the analysis of the RCT reported in chapters 2 and 3. The next four chapters describe one of the studies mentioned above in detail. In a final chapter, I briefly reflect on these studies from a wider perspective.

Chapter 1

An RCT in rural Zambia

Background for chapters 2 to 4

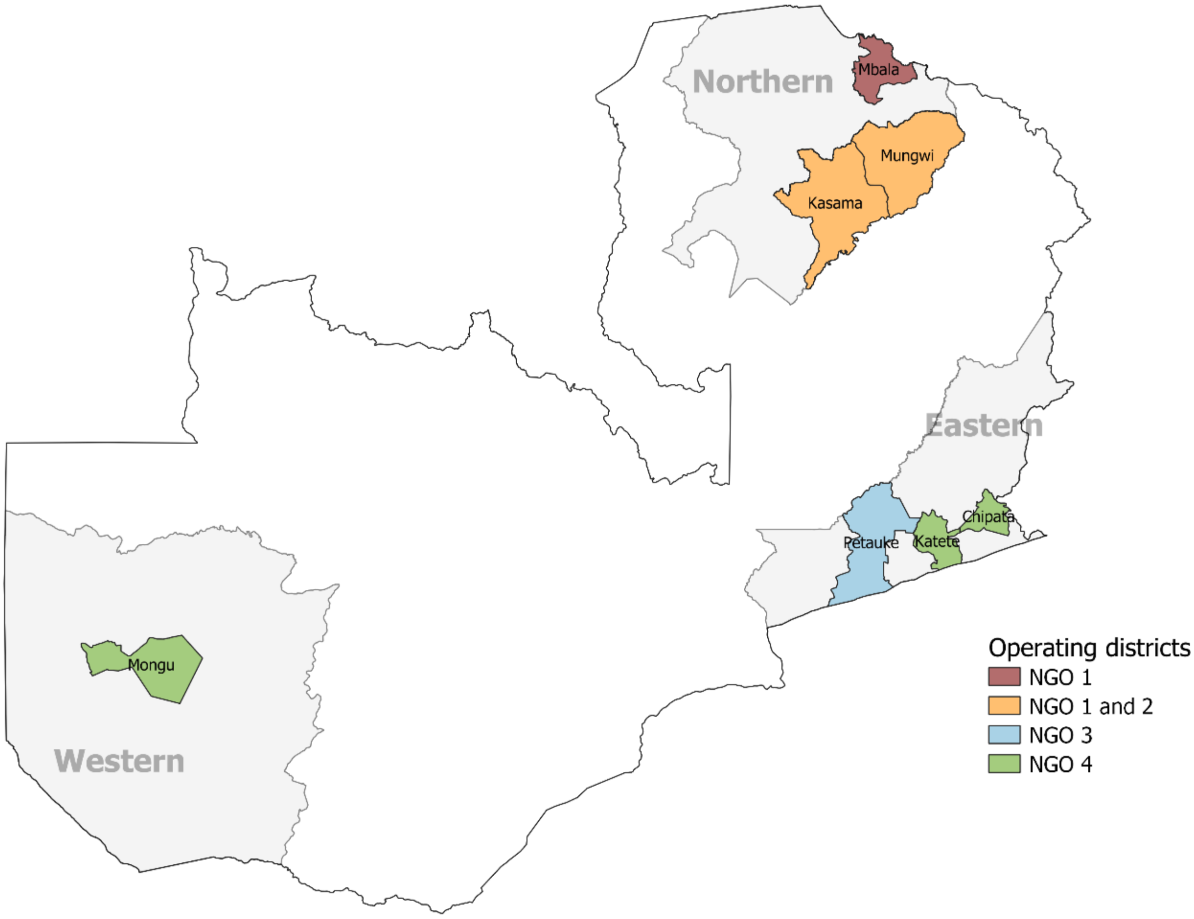
In chapters 2, 3, and 4, we study savings groups affiliated with four different NGOs operating in seven districts in the Northern, Western, and Eastern Province of Zambia (see Figure 1.1). We conducted a baseline survey on 534 savings groups that were active at the time and formed at least one year earlier. After an assessment for eligibility, we excluded 12 savings groups from NGO 3 as they were already benefiting from a different program, leaving a total of 522 savings groups included in the studies.[1]

Savings groups are commonly distinguished by whether they distribute savings on a rotating or accumulating basis. Rotating Saving and Credit Associations (ROSCAs) collect member's regular contributions in a fund and issue the entire fund to a member in each meeting on a rotating basis (Besley et al. (1994)). In contrast, Accumulating Saving and Credit Associations (ASCAs) accumulate member's regular savings and generated interest from loans and distribute the fund to its members after a previously defined period (Bouman (1995)). The savings groups studied in chapters 2, 3, and 4 are predominantly one of three types of accumulating savings groups: Village Saving and Lending Associations (VSLA), Saving and Internal Lending Communities (SILC), and Self-Help Groups. While similar, there are differences regarding meeting modalities, administration and saving, and lending activities.

Table 1.1 gives an overview of some characteristics of the savings groups and their members. We can see that the members are predominantly women with an average age of 46 years. The average household size is 5.7. On average, their monthly savings contribution is 96 ZMW, and the monthly contribution to the social fund is about 5.6 ZMW. There is considerable variation across NGOs. For NGO 1, the averages are 99 and 4.9 ZMW, for NGO 2 91 and 3.1, for NGO 3 84 and 2.7 ZMW, and for NGO 4 101 and 8.8 ZMW for monthly savings and social fund contributions respectively. Most groups meet once a week or once a month, while some meet every other week. The savings are

[1] Note that as only 74 savings groups were affiliated with NGO 2, too few for the planned cross-randomization, these savings groups are excluded from the study presented in chapter 3, leaving 448 savings groups from three NGOs included in that study.

Figure 1.1: Operating districts of implementing partners in Zambia.



primarily either mainly stored in a box or lend out to the members; almost no group in the sample reportedly used a bank account for storage at baseline. Note that there are remarkable differences across NGOs. Savings groups affiliated with NGOs 1 and 4 mostly meet every week, whereas saving groups affiliated with NGOs 2 and 3 almost exclusively meet once per month. Nearly all groups reportedly use a box to store their savings at baseline for NGOs 1 and 2, a majority for NGO 3, and about half of those savings groups affiliated with NGO 4. In contrast, almost all savings groups affiliated with NGOs 3 and 4 loan out their savings to members. A majority of those affiliated with NGO 2 do the same, while only a few savings groups affiliated with NGO 1 do. Also, about half of savings groups affiliated with NGO 4 are active in urban areas, whereas only a few are for NGO 2 and almost none for NGOs 1 and 3.

Table 1.1: Savings group and member characteristics

	All mean	All sd	NGO1 mean	NGO2 mean	NGO3 mean	NGO4 mean
Village level baseline variables						
Number of participating SGs in village	1.5	1.6	1.5	1.5	1.3	1.7
Urban or rural: Urban	.2	.4	.02	.14	0	.53
Urban or rural: Rural > 250	.52	.5	.63	.6	.59	.35
Urban or rural: Rural <250	.28	.45	.35	.26	.41	.12
Mean score of additive food security index	1.2	.81	1.2	.94	1.1	1.5
Mean number of months with food scarcity across HHs	1.5	.82	1.4	1.1	1.2	1.9
Savings group level baseline variables						
Meeting frequency: weekly	.52	.5	.88	.013	0	.71
Meeting frequency: every two weeks	.033	.18	.02	0	0	.074
Meeting frequency: monthly	.45	.5	.1	.99	1	.21
Group uses box to store savings	.74	.44	1	.95	.72	.46
Group loan outs savings as storage	.68	.47	.1	.64	1	.96
Group uses bank account to store savings	.048	.21	.02	.23	.0093	.021
Household level endline variables						
Number of household members	5.7	2.2	5.4	5.3	5.8	5.9
Member level endline variables						
Respondent is female	.8	.4	.73	.73	.77	.89
Age of respondent	46	12	45	51	45	44
Relation to household head: Household head	.43	.5	.43	.54	.41	.41
Respondent is married	.73	.44	.8	.71	.8	.66
Member level baseline variables						
Used mobile money in last 3 months	.32	.47	.24	.45	.37	.31
Trust in financial institutions	.49	.5	.56	.51	.47	.44
Measure of risk aversion	2.6	1.4	2.6	2.6	2.7	2.6
Locus of control (average of z-values)	.0042	.72	.046	-.08	.05	-.02
Monthly savings contribution to group	96	128	99	91	84	101
Value of current savings	702	1098	868	736	638	661
Monthly contribution to SF in ZMW (winsorized)	5.6	7.2	4.9	3.1	2.7	8.8

Notes. The table shows characteristics and outcomes at baseline of savings groups and members overall and by NGO affiliation. Column (1) shows the mean and standard deviation for the whole sample. Columns (2)-(5) display the means for each of the different NGOs. Variables that refer to values are in ZMW and winsorized at the 1% and 99% quantile.

1.1. EXPERIMENTAL DESIGN

The overall intervention is composed of three parts: (a.) general training, (b.) the linkage intervention, and (c.) the social fund intervention. All groups in the study were offered general financial literacy training. The goal of the general training was to enhance the capacity of savings groups through financial education. Despite operating for many years, savings groups face challenges following their principles. Therefore, the implementing partners assessed the needs of the savings groups and trained them accordingly, e.g., in record-keeping, member screening, interest rate calculation, and annual planning.

In addition to this general training, the two other interventions were cross-randomized to be studied. The linkage intervention, discussed in chapter 2, had the goal of facilitating the linkage between groups and financial service providers. The social fund intervention, discussed in chapter 3, aimed at strengthening the groups' informal insurance mechanism, the social fund.

The 522 savings groups were randomized separately according to their NGO affiliation. We defined clusters of savings groups in proximity to avoid spill-over effects, resulting in 351 randomization clusters. Most clusters consist of one savings group, but others comprise up to 14 savings groups. For each NGO, the affiliated clusters were assigned to one of the treatment arms offered by that NGO.[2] To ensure balance across the treatment arms, a re-randomization procedure was conducted: The savings group clusters were randomized repeatedly until a pre-specified degree of balance of 26 variables was reached.

Table 1.2: Design: Cross-randomization within NGOs.

	Control	Linkage	Social fund	Both	Total
NGO 1	35	35	52	30	152
NGO 2	30	44	-	-	74
NGO 3	20	30	30	27	107
NGO 4	55	40	42	52	189
Total	140	149	124	109	522

The table shows the number of savings groups assigned to each of the treatment arms by NGO affiliation. Note that the number of groups differs across treatment arms for each NGO due to different sizes of the randomization clusters.

Table 1.2 shows how many savings groups are affiliated with each NGO and were assigned to each treatment arm. In terms of randomization clusters, the treatment arms per NGO are of equal size, but the number of groups differs across treatment arms due to varying cluster sizes.

The general training was conducted in the second half of 2017, whereas the social fund and linkage interventions were implemented by the different NGOs at different times between the end of 2017 and mid-2018 (see Figure 1.2 for details during which period each NGO reportedly implemented the training). To monitor the treatment implementation, we collected attendance data. The four NGOs were supposed to keep account of attendance at the training sessions. However, some NGOs did not implement attendance lists due to negligence, while others mentioned other reasons for not providing this information (e.g., loss of documents due to moving headquarters). Depending on the type of training, the NGOs provided only up to 82% of lists, but often no lists are available for the interventions. Hence it is difficult to assess fidelity with the randomization protocol.

[2]Due to a small number of groups, NGO 2 did not provide the social fund intervention, such that there are only two treatment arms for this NGO.

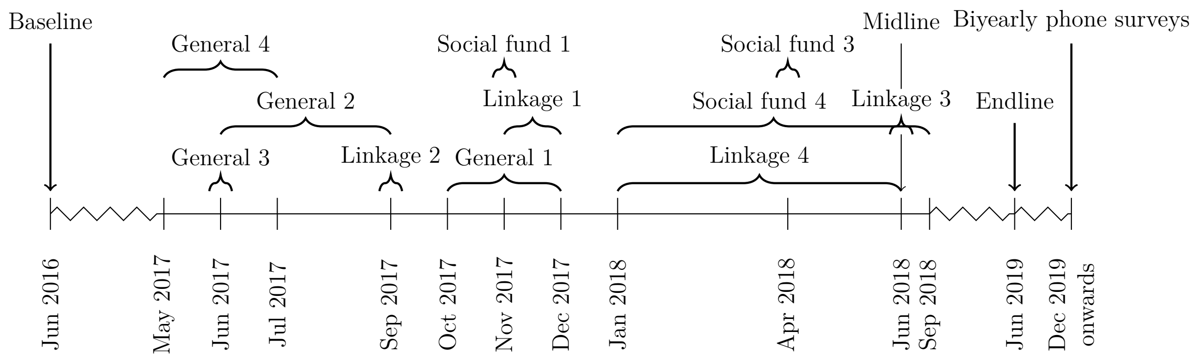


Figure 1.2: Timeline of interventions and major data collections.

1.2. DATA COLLECTION

We take advantage of several data sources for the evaluation. Figure 1.2 displays the timeline for the major data collections in relation to the interventions, for a more detailed illustration refer to Figure A.1 in appendix A. Table 1.3 provides an overview of all data collections, including sample sizes and types of questionnaires. A baseline survey was conducted in June and October 2016, a midline in July 2018, and an endline in July 2019. We collected interviews with randomly selected savings group members and their households during each of these survey waves. Further, we collected interviews with community leaders during the baseline and representatives of the sampled savings groups during the mid- and endline. In addition to the face-to-face interviews, we conducted phone surveys with representatives of the savings groups in each month between April 2018 and April 2019 and in December 2019, April 2020, and November 2020.[3]

[3] While the response rates were high in December 2019 and the subsequent phone survey waves, we reached only about half the groups in each wave of the monthly phone survey.

Table 1.3: Number and type of collected interviews.

	Savings group	Adult	Household	Village
Baseline survey	-	2,085	2,096	316
(June/July 2016)				-
Midline survey	521	1,945	2,070	-
(June/July 2018)		(79.1%)	(83.6%)	-
Endline survey	522	2,604	2,502	-
(June/July 2019)		(76.4%)	(72.2%)	-
Monthly phone surveys				
April 2018	251	-	-	-
May 2018	296	-	-	-
June 2018	274	-	-	-
July 2018	319	-	-	-
August 2018	296	-	-	-
September 2018	308	-	-	-
October 2018	352	-	-	-
November 2018	259	-	-	-
December 2018	259	-	-	-
January 2019	272	-	-	-
February 2019	270	-	-	-
March 2019	294	-	-	-
April 2019	284	-	-	-
May 2019	292	-	-	-
Phone surveys				
December 2019	506	-	-	-
April 2020	515	-	-	-
November 2020	505	-	-	-

Notes. The table shows the sample size of each type of survey questionnaire. Note that during the baseline survey, no savings group were interviewed, but interviews with village leaders were conducted. During the phone survey, we only called representatives of the savings groups to ask about the group's activity. The sample size for the savings groups refers to the number of savings groups reached, regardless of whether they were active, on pause, or dissolved. While we tried to interview the same savings group members and households, this was not always possible; the number in brackets refers to the attrition rate relative to the baseline. Additional households were sampled in both mid- and endline.

1.3. ESTIMATION STRATEGIES.

Our main specification is a regression of the following form:

$$Y_i = \alpha + \beta D_i + \gamma X_i + \delta NGO_i + \epsilon_i$$

where Y_i is the outcome of interest, D_i an indicator for treatment assignment, X_i are covariates which may include outcomes measured at baseline, NGO_i refers to NGO fixed effects which in our context are also stratification dummies. The latter is excluded in the case of estimations for specific NGOs. The standard errors ϵ_i are clustered at the level of the randomization cluster.[4] The reported parameter is the estimate of the average intention-to-treat effect $\hat{\beta}$.

Note that we do not fully saturate the model with respect to the treatment arms. If not mentioned otherwise, in chapter 2, the treatment group refers to the treatment arms which were assigned to the linkage intervention (”Linkage” and ”Both” in Table 1.2), while the control group refers to the remaining treatment arms (”Control” and ”Social fund” in Table 1.2). Whereas in chapter 3, the treatment group refers to the treatment arms which were assigned to the social fund intervention (”Social fund” and ”Both” in Table 1.2), while the control group refers to the remaining treatment arms (”Control” and ”Linkage” in Table 1.2). This is a valid approach as the two interventions are, by construction, independently assigned. Since the two interventions target different aspects and have different goals, we would not expect effects of one intervention on outcomes targeted by the other. Further, we do not expect the interventions to have complementary effects. We nevertheless conducted robustness checks including dummies for all treatment arms. We concluded that our findings are robust with respect to this specification and thus omitted this analysis from the discussion for brevity.

For the main estimation results, we either include no additional covariates or the covariates X_i are selected using a LASSO procedure, taking into account potential unbalance at baseline and selecting predictive covariates to increase power. Note that generally few if any covariates are selected. The reported standard errors are adjusted for this type of model selection (see Chernozhukov et al. (2015)). All variables used in the randomization and savings group characteristics are included in this procedure for all estimations. Household characteristics are included for estimations on the household and group member level, and group member characteristics for estimations on the group member level

[4] Randomization units range in size from 1 to 14, with a majority consisting of a single group. We also conducted the analysis on the level of randomization. For this, we replaced the variables Y_{it} with the randomization unit means.

(for details see Tables A.1, A.2, and A.3 in appendix A.) [5]

Throughout chapters 2 and 3 we make use of indices to analyze the interventions effect on a set of related outcomes. We construct indices for types of outcomes to reduce the number of hypothesis tests and thus avoid spurious findings. Our approach follows Kling et al. (2007) and indices constructed as follows:

$$\mathbb{Y}_i = \frac{1}{n_i} \sum_j \tilde{Y}_{ji} \text{ with } \tilde{Y}_{ji} = \frac{Y_{ji} - \bar{Y}_{j0}}{s_{Y_j 0}}$$

where $\mathbb{Y}_i$ is the index, $\tilde{Y}_{ji}$ the standardized value of related outcome Y_{ji}, $\bar{Y}_{j0}$ and s_{j0} are the control group's mean and standard deviation of Y_{ji} respectively, n_i is the number of non-missing standardized outcomes $\tilde{Y}_{ji}$ for observation i.

Each index is the mean of related outcomes standardized to the control group such that the control group mean and variance for each index are normalized to about 0 and 1, respectively. Thus the coefficient can be interpreted as a change in terms of standard deviations. This approach increases power and is valid if the outcomes the indices are based on are assumed to be affected in the same direction.[6]

[5]This is implemented using the *pdslasso* command in Stata.

[6]One pitfall when using indices can be if there are substitution effects. E.g., assume that the treatment increases income which the household invests in animals and imagine the household would buy a chicken and a pig if they are not treated, with the chicken being considered inferior to the pig, but can afford two pigs (but no chicken) in case they are treated. In this hypothetical scenario there is a treatment effect, but we would no find any effect based on the index. We therefore provide more details on the outcomes the indices are based on in the respective appendices to confirm that such effects are not masqueraded by the presented analysis based on indices.

Chapter 2

Linking savings groups to banks

with Markus Frölich and P. Linh Nguyen

2.1. INTRODUCTION

Saving and investing is commonly believed to be a vehicle that helps households to get out of poverty. At the same time, we observe low rates of financial inclusion among the poor (Demirgüç-Kunt and Klapper (2012)). Evidence, however, suggests that the poor under-save and have a substantial potential demand for financial instruments. This observation might be explained by barriers to efficient adoption and effective usage of financial instruments such as transaction costs, lack of trust and regulatory barriers, information and knowledge gaps, social constraints, and behavioral biases (Karlan et al. (2014)). Therefore, increasing financial inclusion by offering or improving saving technologies is a widely implemented and tested approach for poverty alleviation. One crucial question in this context is the role of formal compared to informal savings technologies.

Our work studies the potential for and of formal financial inclusion in the context of savings groups in rural Zambia. First, we ask whether these savings groups or their members can be formally included by linking them to formal financial institutions. And second, we study the consequences of such a formal inclusion on economic activity and household welfare.

There is extensive literature on the consequences of both formal and informal financial inclusion of the poor in developing countries. While there are many studies investigating barriers to uptake and benefits of formal financial services (see Dupas et al. (2018) for a recent overview), the evidence is generally mixed, and findings seem highly context-dependent. Our first contribution is to provide experimental evidence for a new context regarding both the target population and country. To our knowledge, we are the first to study this type of intervention in Zambia. While Jamison et al. (2014) already studied a similar intervention in the context of groups, we are the first to investigate it in the context of savings groups.

Our context is of particular interest as it highlights the distinction between formal and informal financial inclusion. The target population of our intervention is savings groups which in themselves are informal vehicles of financial inclusion. Thus the second contribution of our study is the focus on potential benefits of formal beyond other types of financial inclusion.

In addition to analyzing intermediate outcomes on savings and lending activity, we further investigate broader welfare outcomes and study effects over a longer time frame than many previous experimental studies.

Further, savings groups in themselves have been subject to both a theoretical (e.g., Besley et al. (1994)) and an emerging empirical literature (e.g., Beaman et al. (2014), Brunie et al. (2014), Ksoll et al. (2016), Karlan et al. (2017)) studying the workings and benefits of such savings groups across the developing world. Given how widespread savings groups are and the evidence that speaks to their benefits, studies started to investigate specific features of these groups, e.g., Burlando and Canidio (2017) who study how group composition matters. Our analysis also contributes to this literature in that it investigates linking savings groups with formal financial service providers.

To answer our research questions, we implemented a randomized control trial. We randomly assigned savings groups to receive what we call a linkage intervention. The particulars of this intervention varied across four implementing partners, but the overall goal was to facilitate linkage between the groups and formal financial service providers. One example would be that a representative of a financial service provider visits the savings groups, informs them about their products, and helps with the registration. The experiment was an encouragement design, i.e., financial services such as bank accounts were not offered directly or for free. Therefore we can only study the linkage intervention directly and formal inclusion only indirectly as the likely mechanism of this intervention on downstream outcomes. Since encouraging the uptake of financial services usually goes hand in hand with increasing financial literacy, we tried to isolate formal inclusion from financial literacy by assigning all the groups, including the control group, to receive general financial literacy training.

In line with previous experimental studies on formal financial inclusion, we find (i) it can be difficult to link savings groups to financial service providers, (ii) the effect of such a financial inclusion is limited in terms of savings and lending activity, and (iii) there are no effects on welfare-related outcomes.

We find that the intervention successfully increased the share of groups that open a bank account by approximately doubling it from 15% and the share of groups that actively use these accounts from 4.6% by 9 percentage points in the mid-term. We further find that the effect on active usage dissipates in the long run. These effects are highly heterogeneous across NGOs, and we only find positive effects for two of the four implementing partners. While these effects might be considered small in absolute terms, i.e., far from linking a majority of savings groups, they are large in relative terms, in line with what we would expect from related studies in many of which accounts are offered for free or are substantially subsidized.

We do not find effects of the intervention on savings or lending activities, which might not come as a surprise, given the limited results on active use of bank accounts, the most plausible channel through which the intervention might affect these activities. While we find negligible point estimates, the variance of all the various measures is too large to conclude that there are no meaningful effects.

Given that we do not find effects on savings or lending activity, we expect to find no effects on welfare-related outcomes. This is indeed the case, and we find that the intervention has no meaningful effects on welfare-related outcomes as measured by our survey instruments.

The remainder of this chapter is structured in three parts. Section 2 provides further information on the background and intervention of the study. In section 3, we present the analysis structured by types of outcomes alongside the theory of change. Finally, section 4 briefly concludes.

2.2. BACKGROUND AND INTERVENTION

While the Zambian financial sector has shown moderate development over the past decade, lack of consumer awareness, low-cost products, and financial literacy keeps financial inclusion at low levels: 48.7% of the rural farming population is not financially included (FSD (2015)). Savings groups could potentially be a vehicle that could help to increase access to formal financial services.

From our baseline survey we get that less than 5% of savings group members have any savings at a formal bank, reflecting low levels of formal financial inclusion in rural Zambia. While including savings group members directly in the formal financial system is challenging, it might be more feasible to include savings groups as a whole. The group mechanism might surmount barriers for individuals to open an account. According to FSD (2015), the most indicated reason for neglecting bank services is insufficient money to justify using a bank. Opening an account as a savings group instead of individual accounts could overcome this threshold.

The linkage intervention had the goal of linking the groups to formal financial service providers by opening and using bank or mobile money accounts to deposit their savings. Storing their savings at formal financial service providers could improve their safety and provide a commitment to save, leading to increased savings or investment activity by the members. In the long term, this could further increase access to credit through formal financial institutions.

While the overall concept was similar, there were substantial differences in implementation between the four NGOs. In addition to informational meetings conducted by all NGOs, NGO 3 provided bank account opening forms and technical support. At the same time, implementers of NGO 1 accompanied groups to registration offices and banks to support the process pro-actively. NGO 4 decided to partner with and focus solely on mobile money providers. But since these providers did not offer group accounts, the intervention focused on individuals - breaking with the intervention's design. As shown in Figure 1.2, the linkage intervention was implemented by the different NGOs at different times between the end of 2017 and mid-2018. Further note that while the implementation differed across NGOs, their operating districts differ as well (see Figure 1.1). In addition, the type of group also varies as the different NGOs established, trained, and supported different concepts of savings groups (compare Table 1.1). So we want to keep in mind that these highly correlated factors might also explain any difference in effect according to NGO affiliation.

Based on monitoring data, we know of at least three groups that attended the linkage-related training but were not assigned to the linkage intervention. Note that this number reflects a lower bound. According to reports by the four NGOs, they implemented the training following the agreed guidelines with a few exceptions, e.g., when savings groups came uninvited to training sessions, and the field officer could not reject them.

2.3. EMPIRICAL ANALYSIS

In this section we discuss the analysis of the linkage intervention. We start with assessing the balance of covariates between the linkage treatment group and the control group followed by the analysis of outcomes ordered according to a theory of change. As first set of outcomes, we discuss whether the intervention achieved its direct goal of creating linkage between savings groups and formal financial institutions. We continue with trust outcomes which can both affect and be affected by such a linkage. This is followed by the analysis of savings and lending activity. Linkage is a purpose not in and of itself, but only matters as far as it affects savings or other behavior. Finally, we investigate the impact on welfare-related outcomes as the ultimate goal of the intervention.

2.3.1. Balance

Table 2.1 shows t-tests between treatment and control group of selected characteristics and outcomes measured at baseline to assess balance.[1] While all differences are small, the treatment savings group are statistically significantly less likely to have used a bank account at baseline (3 percentage points), and respondents in the treatment group are less likely to be female at endline (4 percentage points).

Table 2.1: Savings group and member characteristics

	Control		Treatment		
	mean	sd	mean	sd	β
Village level baseline variables					
Number of participating SGs in village	1.5	1.89	1.5	1.34	-0.018
Urban or rural: Urban	.21	0.41	.19	0.39	-0.014
Urban or rural: Rural > 250	.53	0.50	.51	0.50	-0.028
Urban or rural: Rural <250	.25	0.44	.3	0.46	0.042
Savings group level baseline variables					
Meeting frequency: weekly	.57	0.50	.46	0.50	-0.036
Meeting frequency: every two weeks	.034	0.18	.031	0.17	-0.001
Meeting frequency: monthly	.39	0.49	.51	0.50	0.037
Group uses box to store savings	.73	0.45	.75	0.43	0.026
Group loan outs savings as storage	.66	0.48	.7	0.46	-0.013
Group uses bank account to store savings	.057	0.23	.039	0.19	-0.031*
Household level endline variables					
Number of household members	5.6	2.22	5.7	2.19	0.089
Member level endline variables					
Respondent is female	.82	0.39	.78	0.42	-0.039**
Age of respondent	46	12.52	46	12.29	-0.507
Relation to household head: Household head	.41	0.49	.45	0.50	0.035
Respondent is married	.73	0.44	.74	0.44	0.009
Member level baseline variables					
Used mobile money in last 3 months	.32	0.47	.32	0.47	-0.013
Trust in financial institutions	.49	0.50	.49	0.50	0.004
Monthly savings contribution to group	98	132.00	93	123.12	-3.626
Value of current savings	697	1058.25	708	1136.98	21.109

Notes. The table shows characteristics and outcomes at baseline of savings groups and members by treatment status. Columns (1)-(2) and (3)-(4) show the mean and standard deviation for the control and treatment group. Column (5) informs about balance by displaying the coefficient β from the regression $X_i = \alpha + \beta D_i + \delta NGO_i$, where X_i is the respective variable, D_i treatment assignment, and NGO_i refer to NGO dummies. Significance of a t-test for $\beta = 0$ is referenced by $+: p < 0.10$; $*: p < 0.05$; $**: p < 0.01$. Variables that refer to values are in ZMW and winsorized at the 1% and 99% quantile.

[1] Note that we conducted these t-tests for 19 additional variables for none of which we found statistically significant differences and which are excluded for brevity.

2.3.2. Linkage

The first thing we want to analyze is whether or not the linkage intervention successfully linked the savings groups to formal financial service providers, i.e., whether or not the groups opened bank accounts and used them. It is essential to distinguish in our analysis the active usage from the opening of bank accounts (Karlan et al. (2014)). This analysis is based on reported data, and as such, some issues need to be discussed.

We want to be careful with self-reported data as the group's spokesperson might say they opened a bank account, even if they did not. This can be especially problematic when the intervention encouraged opening a bank account. Hence, to avoid experimenter demand effects, we only ask about bank accounts indirectly.[2] Respondents are asked about all places their groups used as storage for savings in the past, instead of asking about specific locations. But interviewers are supposed to probe for whether there are multiple locations. Only if the respondent mentions a bank account in her answer to this question, further and more detailed questions about the bank account are asked.

During the high-frequency survey (before the endline survey), this question was often reportedly understood to mean where the group stored their money between the current and last meeting. Thus this question might be more indicative of active bank account use compared to having opened and ever used a bank account. If we look at the share of active bank accounts of all bank accounts reported within a given survey wave, we can see that more than 40% are reported to be in active use.[3] At the same time, less than 25% of all groups that reportedly used a bank account in any of the survey waves use them actively in a given survey.[4] This suggests that bank accounts are more likely to be reported when they are actively used, which explains why we need to consider the information from several survey waves to construct the outcome of ever using a bank account, as probing techniques were not sufficient to collect this information.[5]

[2]During the monthly phone surveys, enumerators reported that respondents asked whether they should use a bank account after being asked about it. This might indicate that directly asking about bank accounts might already be a form of treatment, which would be undesirable from the researcher's perspective.

[3]Active use is defined as reportedly using the account monthly or more frequently.

[4]Note that this holds for midline, endline, and the phone survey waves in December 2019 and April 2020. In November 2020, we changed the survey instrument to account for this issue, and 33% of groups to ever report a bank account reported active usage in this wave.

[5]This also has implications for empirical researchers in general, who rely on reported data subject to recollection bias. High-frequency surveys inquiring about key variables might be an important tool to alleviate issues of recollection bias.

Opening of bank accounts. To account for the issues mentioned above when analyzing the opening of bank accounts, we combine several data sources: The mid- and endline surveys, the monthly phone survey in between, and the phone surveys after the endline. If a respondent mentions a bank account as described above and further can give an approximate date of when the account was opened, the group is considered to have opened a bank account. In most cases, the bank account is mentioned several times, but it is mentioned only once in a significant share. This might not be surprising, as there was a year between the mid- and endline, during which the groups could have opened an account, the respondent might have joined the group, or forgot these details. Further, the high-frequency phone survey only had a low response rate in each given month (see Table 1.3), such that requirements such as several mentions might be too strict.[6] From the information about opening dates, we construct a dummy for whether the group reported using a bank account that was opened during or before a given month for each month since the baseline survey.[7] This means if the group reported using a bank account which was opened in November 2017, the indicator is 1 for November 2017 and all months that follow and 0 for all months before. We then ran the following regression model to get treatment effects in each month:

$$Y_{iT} = \sum_{T}(\alpha_T + \beta_T D_i \cdot T) + \sum_{i} \delta_i + \epsilon_{iT}$$

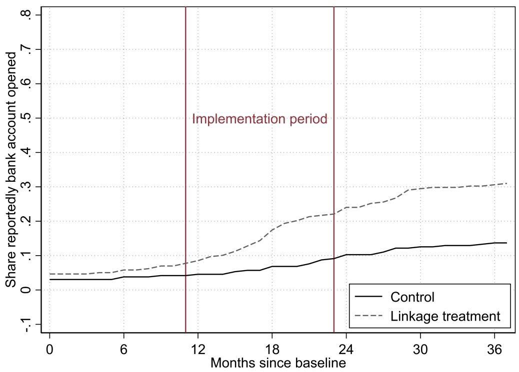
where T refers to the month since the baseline in July 2016, D_i is a treatment indicator, δ_i are savings group fixed effects. The index i refers to savings groups, and the standard errors are clustered at the level of the randomization unit. Months T include each month up to 3 years after the baseline, after which all dates are accumulated. $\hat{\beta_T}$ can be interpreted as an estimate of the average ITT effect for a given month T and is the reported estimate in the discussion that follows.

Figure 2.1 displays the raw information for treatment and control group separately (Subfigure 2.1a) as well as the estimated $\hat{\beta_T}$ (Subfigure 2.1b). The X-axis displays months since the baseline survey. On the Y-axis, we have the share of groups that had opened an account that was reportedly used and the treatment effect, respectively. The gray shaded area around the point-estimates on the right respond to 95% confidence intervals. We can see that the treatment was effective in the sense that groups opened bank accounts after the treatment. We can see significant estimates from around November

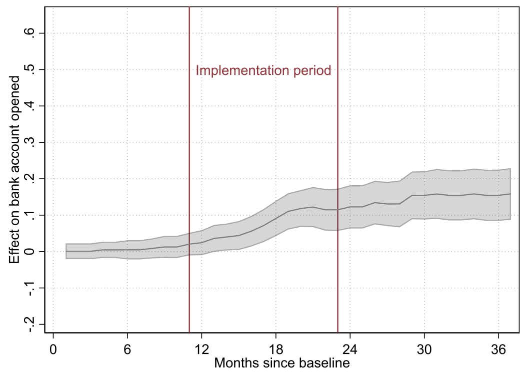
[6]We conducted robustness checks considering only accounts that where mentioned at least twice with similar findings.

[7]Note that the reported dates are sometimes imprecise, i.e. only the year is given, in which case July is assumed. Further, sometimes the dates given across survey waves are contradictory. In these cases, as a default, the oldest date was used if there were no good reasons against it.

Figure 2.1: Opening of bank accounts



(a) Raw shares by treatment



(b) Effect on likelihood

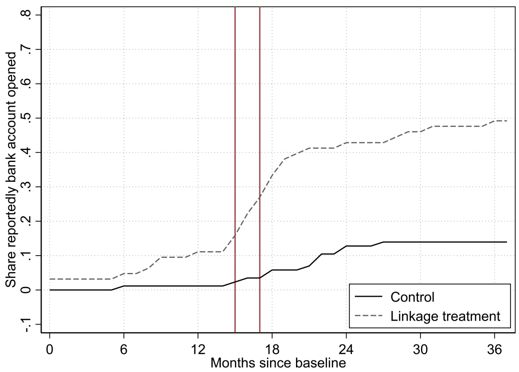
and December 2017 onwards, when NGO 1 implemented the linkage intervention. After this, the treatment effect slowly but steadily rises from around 5 to 15 percentage points. This implies that even up to 2 years after the intervention, the control group did not catch up in opening bank accounts. The share of groups that opened an account in the control group rises from less than 5% to about 15% over the three years, and thus the intervention doubled the share of groups that opened a bank account. Note, however, that this graph does not represent the use of bank accounts but only whether they opened a bank account they ever reportedly used (remember that $Y_{it} \geq Y_{is} \; \forall \; t \geq s$).

As already mentioned, the effect took off after NGO 1 started implementing the intervention, so next, we will look at the effect over time for each NGO separately.

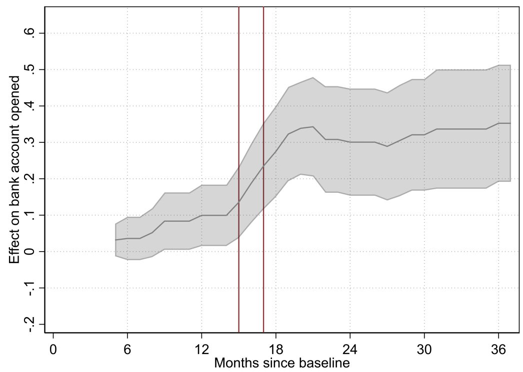
If we take a look at each of the NGOs (see Figures 2.2 and 2.3), we can see that the effect of the intervention was mainly driven by NGO 1 (Subfigure 2.2b) and that there was no effect for NGO 4 (Subfigure 2.3d).[8] We can further see an effect for NGO 2 (Subfigure 2.2d) which is only realized slowly compared to the exemplary development for NGO 1. The effect of the intervention by NGO 1 is estimated at around 30-35 percentage points and was fully realized around the time of their intervention (from a control group mean of around 10%). For NGO 2, the effect looks to be slowly increasing when they started the linkage intervention. However, it took one and a half years to reach an estimated effect of around 35-40 percentage points (compared to around 35% in the control group). For NGO 3, we cannot detect a statistically significant effect, and the point estimates are only at about 5-10 percentage points during the time of the intervention, after which they taper off (Subfigure 2.3b). For NGO 4, we find no effect at all. The lack of an effect for NGO 4 is expected since their implementation of the intervention focused solely

[8]While all NGOs drive the overall effect except NGO 4, NGO 1 and 2 have the largest effects, but twice as many savings groups are affiliated with NGO 1 compared to NGO 2.

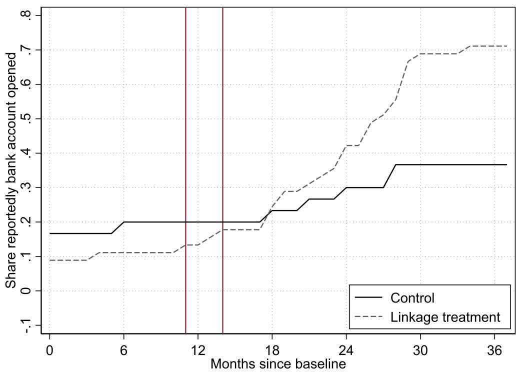
Figure 2.2: Opening of bank accounts: NGO 1 and 2



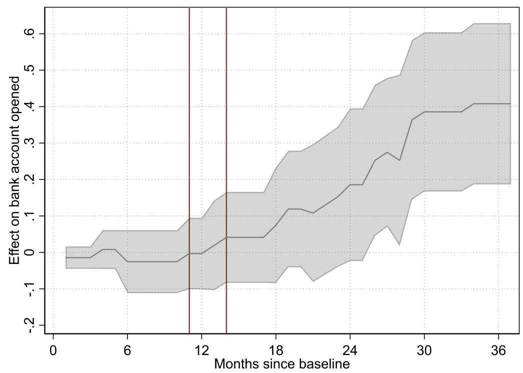
(a) Raw shares by treatment: NGO 1



(b) Effect on likelihood: NGO 1



(c) Raw shares by treatment: NGO 2



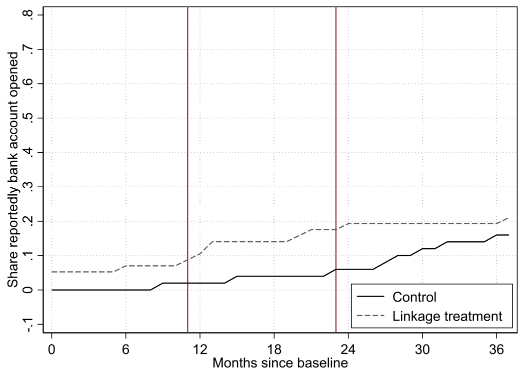
(d) Effect on likelihood: NGO 2

on mobile money rather than bank accounts.[9] Note that while these differences across NGOs might be attributed to the differences in implementation, we have to keep in mind that some group characteristics are highly correlated with NGO affiliation. E.g., both successful NGOs operate only in the Northern province, whereas the other NGOs operate in different provinces.

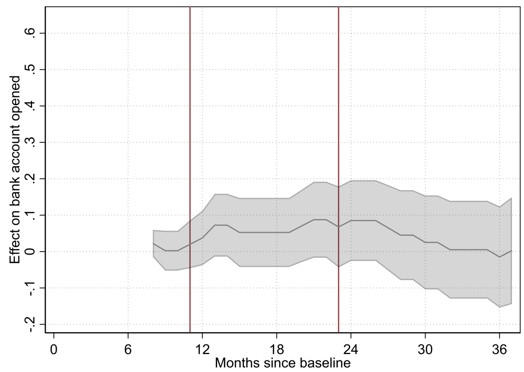
To put the effect sizes in perspective, we first look at studies that offered free bank accounts. Take-up for free bank accounts was 69%, 54%, and 17% for unbanked individuals in Malawi, Uganda, and Chile respectively (Dupas et al. (2018)), 87% for entrepreneurs in Kenya (Dupas and Robinson (2013a)), 69% for households in Kenya (Dupas et al. (2017)), 23% for unbanked individuals in the Philippines (Karlan and Zinman (2018)), 53% for MFI members in Chile (Kast and Pomeranz (2014)) and 85% for female household heads in Nepal (Prina (2015)). In a study in which crop proceeds were deposited directly into the bank account, Brune et al. (2016) find a take-up of 20% for farmers in

[9]We discuss the use of mobile money later, but the intervention was unsuccessful for mobile money uptake as there were no mobile money accounts available for groups.

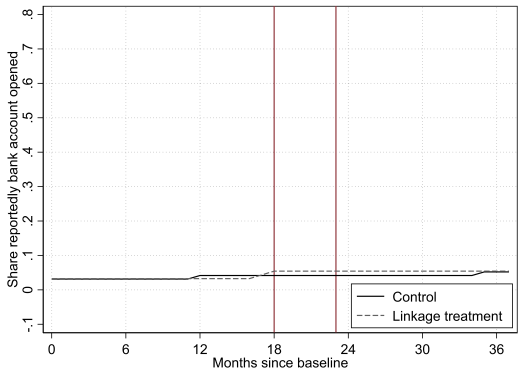
Figure 2.3: Opening of bank accounts: NGO 3 and 4



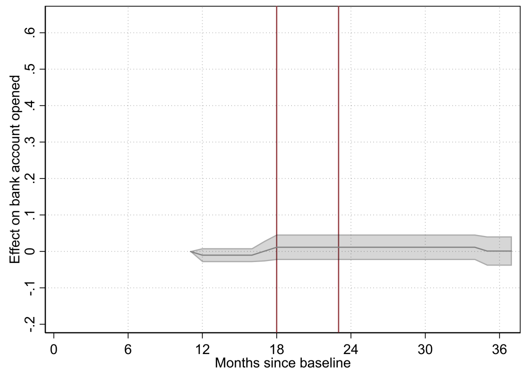
(a) Raw shares by treatment: NGO 3



(b) Effect on likelihood: NGO 3



(c) Raw shares by treatment: NGO 4



(d) Effect on likelihood: NGO 4

Malawi. While these studies all refer to bank accounts for individuals, Jamison et al. (2014) find a take-up rate of 66% for group accounts offered to youth clubs in Uganda.

Considering studies that did not offer free accounts directly, bank account opening rates are naturally lower. For accounts that are only subsidized, Cole et al. (2011) find a take-up rate of 9% for unbanked individuals in Indonesia. Flory (2016) finds that sending bank account promoters to villages increases account opening rates by 2 percentage points among farmers in Malawi. Lee et al. (2017) find an increase in bank account opening of 10-20 percentage points due to promotion to high school youth in Ghana.

In light of the literature, the effect sizes of 30-40 percentage points of the successful implementation partners (NGOs 1 and 2) are comparably large. They fall short of take-up rates of free accounts but are considerably higher than other studies that only promote bank accounts. However, most of these reference studies are based on administrative data, whereas our results are based on reported data with the risk of misreporting. On the other hand, accounts are likely only reported if they were ever used, whereas the above numbers refer to account openings, regardless of whether any deposit was made.

In the study by Dupas and Robinson (2013a) for example, only 60% of entrepreneurs in Kenya made any deposit in their account.

Active use of bank accounts. Next, we have a look at the active use of bank accounts. This analysis relies on the mid- and endline surveys and the phone surveys after the endline. The monthly phone survey reached too few groups in each wave to allow for a proper analysis. We define active use as reportedly using the account at least monthly for deposits, withdrawals, or transfers.[10] The analysis of active bank account use is summarized in Table 2.2.

We can see from Table 2.2 that the intervention was not only successful in increasing the share of groups to open a bank account for storage but also in increasing the share of groups that actively use these accounts. Naturally, the effect on active use is smaller at 9 percentage points at midline, which then drops to about 5 percentage points in April 2020, after which we do no longer find a treatment effect. While these effect sizes seem small in absolute terms, they are significant in terms of percentage of the control group mean, which is about 5% for all survey waves until April 2020, i.e., active usage more than doubles. This effect size seems to be stable over two years, suggesting sustained long-term effects of the intervention for active bank account use. Note that while the treatment effect vanished in November 2020, this is accompanied by an increase in the control group mean, thus suggesting that the control group caught up rather than the treatment group falling back to initial levels.[11] However, upon closer inspection, the effect seems to be less stable and to dissipate earlier. Only about one-third of the groups that reported actively using a bank account in any of the five survey waves also reported actively using it in a specific wave.[12] Across the five waves reported in Table 2.2, 125 groups reported using the bank account actively. Still, only 28 groups did so in all of the waves after they first reported actively using an account.[13] If we take a look at each NGO separately, we find that the effect is mainly driven by NGO 1, the only NGO for which we find significant effects in three of the four survey waves for which we find an effect overall. The point estimates of the effect size for groups affiliated with NGO 1 is

[10] Irregular usage, despite potentially being more frequent than monthly usage, is coded as inactive usage. The findings are robust with respect to this coding. Further note that while active use can mean deposits, withdrawals, or transfers, it usually corresponds to deposits.

[11] Given initiatives such as the Rural Financial Expansion Programme in Zambia and potential spillover effects across groups, it might not be surprising that more and more savings groups in the control group open and use bank accounts.

[12] Again, this differs for November 2020 with a share of 45%.

[13] These numbers only refer to active groups (gave an interview) at the time of a specific survey wave. E.g., if a group reports actively using the bank account in mid-, endline, the December 2019, and April 2020 phone survey but was inactive in the November 2020 wave, it is considered using it actively in all survey waves.

declining over time, from 31 percentage points in the midline to 16 percentage points in the phone survey in December 2019, becoming insignificant and small after that. On the other hand, as with the opening of bank accounts, the estimated effect for NGO 2 affiliated groups is varying over time. Though not being significant, it is large in April 2020. These opposing trends between NGO 1 and 2 can explain why the overall effect seems to be stable even in April 2020. For both NGOs 3 and 4, we don't find any effects, and the point estimates are close to zero, consistent with what we saw for the opening of bank accounts.

Again we can compare this effect size to other studies, which find effect sizes on active use of 10, 17, and 3 percentage points for unbanked individuals in Malawi, Uganda, and Chile respectively (Dupas et al. (2018)), 17 percentage points for entrepreneurs in Kenya (Dupas and Robinson (2013a)), 15 percentage points for households in Kenya (Dupas et al. (2017)), 8 percentage points for MFI members in Chile (Kast and Pomeranz (2014)), and 80 percentage points for female household heads in Nepal (Prina (2015)). These studies are based on administrative data and define active usage by the number of deposits made (4 or more for most of these studies), so it is difficult to compare these numbers directly. However, we can see that while bank account take-up rates are high, usage rates are comparably low in these studies. In our study, the discrepancy between opening and usage is smaller both in absolute and relative terms, as one would expect considering that the groups were not offered free accounts. Nevertheless, we also find that the groups stop using the accounts actively over time. When we asked the groups, which reported using an account in the past, why they stopped using it, they mention reasons such as high fees (48%), high minimum balance on the account (38%), distance to the bank (11%), or that the group would not get a loan (8%). Brune et al. (2016) also report distance and Dupas et al. (2018) high fees as reported reasons for low uptake and usage of bank accounts.

Table 2.2: Active bank account use

	Midline	Endline	Phone surveys		
	(1)	(2)	(3)	(4)	(5)
	July 2018	July 2019	December 2019	April 2020	November 2020
	b/se	b/se	b/se	b/se	b/se
All NGOs					
Linkage treatment	0.091**	0.077**	0.046+	0.047*	0.0053
	(0.029)	(0.027)	(0.028)	(0.023)	(0.032)
Control group mean	0.046	0.042	0.059	0.043	0.12
Observations	518	519	463	459	455
NGO 1					
Linkage treatment	0.31**	0.21**	0.16*	0.038	0.033
	(0.076)	(0.065)	(0.071)	(0.039)	(0.059)
Control group mean	0.058	0	0.026	0.038	0.066
Observations	149	150	128	139	128
NGO 2					
Linkage treatment	-0.15	0.24+	-0.0072	0.21	-0.091
	(0.11)	(0.13)	(0.15)	(0.14)	(0.13)
Control group mean	0.20	0.23	0.32	0.18	0.33
Observations	75	75	65	70	68
NGO 3					
Linkage treatment	-0.0025	-0.067+	-0.043	-0.023	-0.020
	(0.026)	(0.034)	(0.030)	(0.023)	(0.090)
Control group mean	0.020	0.060	0.043	0.023	0.26
Observations	107	107	97	93	98
NGO 4					
Linkage treatment	0.011	0.00057	-0.0098	-0.013	-0.024
	(0.010)	(0.014)	(0.019)	(0.013)	(0.023)
Control group mean	0	0.010	0.022	0.013	0.024
Observations	187	187	173	157	161

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Outcome in all estimations is active use of bank account. Active use is defined as reportedly using the bank account monthly of more frequently for deposits, withdrawals, or transfers. Different columns refer to different survey waves. For the phone surveys, only groups that were active at the time of the survey, or in the months just before, are included. Upper panel includes one estimation for the whole sample, lower panel includes four estimations, one for each NGO. Estimations are done using pdslasso command in Stata. The included covariates can be found in Tables A.1, A.2, and A.3 in appendix A. For estimations on the whole sample, NGO fixed effects are always included, but excluded for NGO specific estimations.

Opening and use of mobile money accounts. Popular in many SSA countries, mobile money could be another way to link savings groups with formal financial services. One of the NGOs even focused its efforts on mobile money providers. However, these providers did not offer mobile money accounts suitable for groups, such that individual accounts would have to be used. In total, across all survey waves, only 16 groups reported having used mobile money to store the group's savings, and only two were categorized as actively using the account. We, therefore, omit any analysis and conclude that the intervention was not effective in promoting the use of mobile money by savings groups. One might think that the approach by NGO 4 might have convinced some of the group members to used mobile money. But we do not find positive effects on the usage of or knowledge about mobile money of group members for the whole sample or NGO 4 (see Table B.2 in appendix B). If anything, the treatment might have reduced the likelihood that a group member reportedly used mobile money previous to the survey. We can further see that mobile money was relatively widespread, with about 50% of respondents having used it before and about 30% having used it in the last three months before the survey in the control group. We also see an increase of about 5% between the mid- and endline of both measures. This is consistent with findings that mobile money is mainly used for remittances and as an insurance mechanism rather than for saving (see, e.g., Jack and Suri (2014), Batista and Vicente (2018), Alinaghi (2019), or Wieser et al. (2019)).

2.3.3. Trust in safety of savings and financial institutions

One reason to open a bank account could be that it is perceived to be a safer way to store the group's savings. One obstacle to open a bank account could be a lack of trust in financial institutions (Karlan et al. (2014), Dupas et al. (2018)). Table B.3 in appendix B reports the treatment effect on these intangibles based on reports from the mid- and endline surveys.

There is no change in how the members reportedly trust that their savings are safe due to the intervention. Even for members who are part of groups affiliated with NGOs 1 and 2 that were successful in terms of increasing bank account usage (see Table B.3 in appendix B). This suggests that using bank accounts as storage did not increase the safety of funds as perceived by the group members. Note, however, that the trust in the safety of savings is already relatively high in the control group, such that there was little room for improvement. Further note that trust declined between mid- and endline on average. A plausible explanation for this pattern would be that the largest perceived risk to the savings is defaulting members. For most savings groups in our sample, the largest share of savings is given out as loans to group members. When every time a savings group member defaults, the trust in the safety of the savings decreases, we would expect to see such a decline in trust over time. We would, however, not expect that using a bank account as storage of savings would mitigate this decline in trust.

We further see that the intervention did not increase the trust in financial institutions of savings group members. Note that we only collected information about trust for financial institutions the respondent has experience with. We find that reported trust in financial institutions generally tends to be lower than in the safety of group savings. Since trust can both increase the willingness to open an account but simultaneously might only be built after using an account, this outcome is hard to interpret. However, the fact that trust in financial institutions seems comparatively low and that it was not measurably affected by the intervention might explain the limited effect of the intervention in terms of account openings and usage. Further building trust with the savings groups does not seem to be the primary channel through which the intervention successfully increased bank account uptake.

2.3.4. Savings and lending activity

While the intervention was to some extent successful in linking the groups to formal banks, financial linkage in and of itself was not the goal of this intervention. Financial inclusion only matters if it results in increased economic activity or improved financial conditions and ultimately in higher welfare. In the following, we thus analyze whether the intervention translated into more savings or lending activity of the group members.

Savings activity. We first want to answer the question of whether or not the savings group members saved more and whether the groups were able to accumulate more savings. We construct measures of the monthly value of savings within the group and elsewhere from the adult questionnaires.[14] And from the savings group questionnaire, we construct measures of total savings within the group. We can see from Table 2.3 that for both questionnaires and both survey waves, reported savings have a large variance, even conditional on predictive covariates. While including covariates helps to reduce the standard errors significantly (compare columns (1) with (2) and (3) with (4) of Table 2.3), it is not enough to find significant effects or reject economically meaningful effects.

The first outcome in Table 2.3 is the reported monthly savings contribution as reported by the group members. The treatment effect estimate is statistically insignificant and small in size in both mid- and endline. As an alternative measure of savings, we asked the members about the value they received at the last share-out.[15] Since the share-out refers to their or the group's accumulated savings plus interest, it can be affected by both

[14]In case we see increased savings within the groups, we might be worried about the crowd-out effects of other savings activities.

[15]Note that for some groups, this can refer to the same share-out for both survey waves.

savings and lending activity. Again the point estimates are statistically insignificant and small in size, measured both in midline and endline. Measuring total savings and the value of the last share-out at the group level (Table 2.3 lower panel), we also do not find significant effects. Since we do not find effects on within-group savings activity, it comes as no surprise that we do not find any crowd-out effects on the members' savings in other groups or outside any groups (third and fourth outcome in the upper panel of Table 2.3).

The lack of findings on savings activity might be expected, given the limited results on active use of bank accounts, the most plausible channel through which the intervention might affect the savings activity. As we have seen in section 2.3.2, there are highly heterogeneous effects on bank account use with respect to NGO affiliation, such that we also analyze the savings activity separately by NGO. We find a sizable negative effect for NGO 2 for the monthly savings contributions at midline and a sizable increase for the value of the last share-out at endline (compare Table B.4 in appendix B). For NGO 1, we find a negative effect on contributions to other savings groups. While the point estimates on the overall sample are relatively small, the variance of the various measures is too large to conclude that there are no meaningful effects on the savings activity of the group members. Particularly the findings for specific NGOs suggest that the intervention might have affected the savings activity even though we detect no clear and consistent patterns.

Table 2.3: Savings activity

	Midline July 2018		Endline July 2019	
	(1)	(2)	(3)	(4)
Monthly savings contribution to group				
Linkage treatment	3.63	5.96	-5.65	-6.10
	(6.38)	(13.2)	(6.11)	(11.7)
Control group mean	105.6	105.6	124.3	124.3
Value received at last share-out				
Linkage treatment	25.0	28.7	43.7	56.0
	(70.4)	(185.3)	(117.1)	(186.2)
Control group mean	1414.9	1414.9	1713.3	1713.3
Monthly savings contributions to other groups				
Linkage treatment	2.20	2.41	-0.73	-0.21
	(2.59)	(2.90)	(2.58)	(2.52)
Control group mean	7.59	7.59	10.5	10.5
Value of savings outside of savings groups				
Linkage treatment	-48.8	-28.6	-1.67	23.0
	(40.5)	(47.8)	(34.1)	(45.2)
Control group mean	244.9	244.9	217.8	217.8
Observations	1945	1945	2604	2604
Total savings in group				
Linkage treatment	2589.6^{+}	2346.6	623.3	82.3
	(1516.5)	(2214.6)	(2172.0)	(2836.2)
Control group mean	14095.9	14095.9	15762.1	15762.1
Value of last share-out				
Linkage treatment	1424.6	1296.0	1228.7	2642.2
	(2909.3)	(4243.7)	(2098.3)	(3635.1)
Control group mean	24576.9	24576.9	23524.2	23524.2
Observations	519	519	520	520
PDS LASSO	Yes	No	Yes	No

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. The upper panel is based on savings group member's responses about their savings behavior in and outside the savings group. The lower panel is based on responses from a savings group representative about the group as a whole. All outcomes are in ZMW and winsorized at 1% and 99%. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A)) or without any covariates. NGO fixed effects always included. Observations refer to total number of observations available not accounting for missing values, thus actual number of observations used varies with outcome. Missing values: Total savings is missing for 69 and 89 and the value of last share-out for 40 and 3 savings groups at mid- and endline. Monthly savings contribution to group is missing for 81 and 30 members at mid- and endline. Monthly contributions to other groups is missing for 30 members at endline.

Lending activity. Even if the savings contributions are not higher, it could be that the bank account serves as an additional commitment device, such that the savings are spent less on temptation and more on investment goods (Dupas and Robinson (2013b)). We thus want to analyze whether we can find a difference in the lending or investment activity of the savings groups or their members' households. The results of lending and investment activities related to the savings groups are shown in Tables 2.4 and 2.5.

From the first two outcomes in Table 2.4, we can see that the members of treatment groups do not take out more loans, neither in number nor value, compared to the control group in both the mid- and endline (Table 2.4 columns (1) and (3)). This is also true if we only take a look at NGOs 1 and 2, which were successful in linking some of the savings groups with the formal financial sector (Table 2.4 columns (2) and (4)). We can infer that there was no economically meaningful change in the number of loans taken. On average, the members take out less than one loan at midline and about one loan at endline. In case of an effect on the loans taken from the group, one might expect displacement effect, for which we included outcome relating to the number of loans from other sources in Table 2.4. Further, one might expect that if the money is less accessible to the group when stored in a bank account, members take out more loans from other savings groups or other sources for which we find no evidence.

Table 2.4: Loan activity

	Midline July 2018		Endline July 2019	
	(1) All NGOs b/se	(2) NGOs 1 and 2 b/se	(3) All NGOs b/se	(4) NGOs 1 and 2 b/se
Number of loans from savings group				
Linkage treatment	0.015	0.026	-0.0090	0.0025
	(0.040)	(0.067)	(0.061)	(0.097)
Control group mean	0.64	0.69	0.95	1.03
Value of loans from savings group				
Linkage treatment	-36.4	-31.8	-59.1	-26.1
	(35.1)	(51.5)	(57.1)	(54.7)
Control group mean	379.4	385.7	561.4	528.5
Number of loans from all savings groups				
Linkage treatment	0.025	0.013	-0.00080	-0.0038
	(0.042)	(0.070)	(0.061)	(0.093)
Control group mean	0.68	0.74	0.97	1.05
Value of loans from all savings groups				
Linkage treatment	-11.6	-78.4	-38.9	-25.1
	(37.5)	(65.6)	(60.5)	(57.7)
Control group mean	413.0	422.9	585.2	543.8
Number of loans taken				
Linkage treatment	0.031	-0.010	-0.016	-0.019
	(0.039)	(0.069)	(0.063)	(0.093)
Control group mean	0.74	0.79	1.03	1.07
Value of loans taken				
Linkage treatment	10.3	-81.6	49.1	-37.7
	(44.3)	(67.3)	(79.5)	(60.0)
Control group mean	450.6	455.3	640.9	572.3
Observations	1945	839	2604	1120

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. Values in ZMW and winzorized at 1% and 99%. All outcomes refer to the last 12 months before survey. Missing values: 22 missing for value of loans at endline and otherwise 11 or less across all NGOs.

Savings group as source of finance of household expenditures. But how do the savings group members' households finance their expenditures? In addition to asking about the value of their (agricultural) investments and medical, educational, and funeral expenses, we inquired about the different sources of financing these expenditures. In Table 2.5 we analyze whether the treatment affected to what extend the households finance these expenditures through their savings group activity. We find no difference in whether they finance through savings or loans from the savings group and not in the amount they finance with either savings or loans from the savings group. This is consistent with the observation that we find no change in neither savings nor lending activity related to the savings group or overall.

Table 2.5: Source of household financing

	Midline July 2018		Endline July 2019	
	(1) All NGOs b/se	(2) NGOs 1 and 2 b/se	(3) All NGOs b/se	(4) NGOs 1 and 2 b/se
Any financing from group savings				
Linkage treatment	-0.014	0.0041	0.0039	0.010
	(0.019)	(0.031)	(0.014)	(0.018)
Control group mean	0.43	0.40	0.19	0.19
Value of financing from group savings				
Linkage treatment	58.4	21.0	1.50	2.35
	(37.5)	(56.6)	(1.10)	(2.57)
Control group mean	132.4	179.1	4.00	5.99
Any financing from group loans				
Linkage treatment	0.011	0.025	0.011	0.025
	(0.017)	(0.027)	(0.015)	(0.021)
Control group mean	0.36	0.34	0.22	0.20
Value of financing from group loans				
Linkage treatment	-4.34	3.31	16.5	2.52
	(13.7)	(22.2)	(14.4)	(17.8)
Control group mean	73.3	67.5	64.1	66.9
Observations	2076	895	2505	1052

$+: p < 0.10$; $*: p < 0.05$; $**: p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Values in ZMW and winzorized within type of expenses at 1% and 99%. All outcomes refer to last 12 months before survey. Outcomes are based on the information given by respondents for the source and value of financing for (agricultural) investments, education, funerals, or health spending.

These findings are in line with what other studies find and do not find. A lack of findings is not the same as finding no effect, and one needs to keep in mind that savings are hard to measure and that savings measures tend to have a large variation, making it difficult to detect changes. While savings groups, in general, are found to increase savings and sometimes lending activity (see Beaman et al. (2014), Ksoll et al. (2016), and Karlan et al. (2017)), there is little support that opening formal savings accounts increases savings or lending. Mobile money accounts for example are mainly used for remittances and as insurance mechanism rather than for saving (see e.g., Jack and Suri (2014), Batista and Vicente (2018), Alinaghi (2019), or Wieser et al. (2019)). For bank accounts, the evidence is mixed: some studies find increased savings, sometimes with a caveat, and many do not find an effect of bank accounts on overall savings. In a small study in Kenya, Dupas and Robinson (2013a) find increased savings only for the subgroup of female market vendors. Brune et al. (2016) find increased savings of farmers in Malawi at planting season when offered a bank account in which harvest proceeds were deposited. However, rarely are any deposits made after the initial deposit. Thus it is more plausible that the bank account served as a commitment device rather than a savings instrument (similarly to Dupas and Robinson (2013b)). De Mel et al. (2018) and Buehren (2011) find increases in formal savings but not overall savings. In a group setting, Dupas et al. (2017), who offered accounts to couples, find that both men and women increase savings when given an account in their own name but do not increase savings if only their spouse is given an account. For group accounts for youth clubs, Jamison et al. (2014) do not find effects on savings from group accounts alone but only coupled with financial education and only for savings in those offered accounts but not on savings overall. Dupas et al. (2018) do not find effects of offering bank accounts to the unbanked on savings in Uganda, Malawi, and Chile.

2.3.5. Welfare outcomes

As discussed in the previous section, we find no evidence of increased savings or lending activity caused by the intervention. The same picture arises for welfare-related outcomes, such as investment, income, and consumption, which is expected given the lack of findings on savings and lending activities, the most plausible channels through which the intervention might affect these outcomes. Table 2.6 gives an overview of treatment effect estimates on a range of investment and welfare-related indices, based on animals purchased and owned, other agricultural inputs and assets, household expenditures and measures of food security (for more details on these indices see Tables B.5, B.6, B.7, B.8 and B.9, B.10, and B.11 respectively in appendix B). Note that we look at indices rather than each specific item to account for multiple hypothesis testing. E.g., for animals owned at endline there is a small negative but statistically significant estimate for the number of poultry owned, but this is one of thirteen categories we included and thus such findings are expected due to chance. One downside is that we neglect potential substitution effects, e.g., moving from one type of asset to another.

As can be seen from Table 2.6 we find no effect of the treatment on any of the welfare-related indices. All of the estimates are small in size. If we take 0.1 standard deviations of the control group as a size threshold for meaningful effects, we can conclude that the treatment did not affect the group members' household welfare in a meaningful way.[16] This is, of course, in line with the previous analysis in section 2.3.4 and we would not expect downstream outcomes to be affected given that we do not find effects on intermediate outcomes in line with related studies (e.g., Dupas et al. (2018)).

[16]Except for the animals purchased index at endline for NGOs 1 and 2, for which the 95% confidence interval covers 0.1 standard deviations.

Table 2.6: Various indices related to household welfare

	Midline July 2018		Endline July 2019	
	(1)	(2)	(3)	(4)
	All NGOs	NGOs 1 and 2	All NGOs	NGOs 1 and 2
	b/se	b/se	b/se	b/se
Index HH animals purchased				
Linkage treatment	-0.034^{+}	-0.043	0.0097	0.065^{+}
	(0.018)	(0.032)	(0.021)	(0.038)
Index HH animals owned				
Linkage treatment	-0.013	-0.031	-0.0017	-0.0059
	(0.020)	(0.022)	(0.020)	(0.011)
Index HH agricultural inputs				
Linkage treatment	-0.020	-0.041	-0.016	-0.050
	(0.029)	(0.050)	(0.028)	(0.043)
Index HH agricultural assets				
Linkage treatment	-0.0060	0.016	-0.017	-0.0028
	(0.015)	(0.016)	(0.018)	(0.016)
Index HH general assets				
Linkage treatment	0.0070	0.017	0.0075	0.014
	(0.016)	(0.028)	(0.013)	(0.019)
Index HH expenditures				
Linkage treatment	0.022	0.039	0.048	0.034
	(0.029)	(0.050)	(0.036)	(0.046)
Index HH food security				
Linkage treatment	0.033	0.027	-0.016	-0.014
	(0.021)	(0.030)	(0.015)	(0.023)
Number of food insecure months				
Linkage treatment	-0.042	-0.079	0.049	0.036
	(0.062)	(0.099)	(0.055)	(0.077)
Control group mean	0.91	0.90	0.88	0.79
Observations	2076	895	2505	1052

$+: p < 0.10$; $*: p < 0.05$; $**: p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. Missing values: 22 and 44 missing for the food security index at mid- and endline. Indices are means of variables related to the index standardized to control group values, thus the control group mean is by construction 0 and omitted. For details on the components for each index refer to Tables B.5, B.6, B.7, B.8, B.9, B.10, B.11 in appendix B.

2.4. CONCLUSION

Overall this chapter provides evidence for whether or not formal financial inclusion is beneficial in the context of savings groups in rural Zambia. As previous studies show, benefits from formal financial inclusion are highly context-specific, and we provide one additional data point. In Kenya, Dupas and Robinson (2013a) find increased savings for female market vendors. In Uganda, Buehren (2011), for MFI clients, and Jamison et al. (2014), for youth club members that also receive financial education training, find increases in savings in formal accounts but not in savings overall. While it seems that financial inclusion can help some people to save more, most studies either do not look for or do not find effects on welfare outcomes (De Mel et al. (2018), Dupas et al. (2017)). Dupas et al. (2018) neither find effects on savings nor welfare outcomes in Uganda, Malawi, and Chile. The cases in which financial inclusion seems to be beneficial in terms of household welfare seem to work through remittances using mobile money in the context of work migration (Jack and Suri (2014), Batista and Vicente (2018)). Our findings are in line with the existing evidence but provide evidence for a new context. Like previous studies, we find (i) it can be challenging to link savings groups to financial service providers, (ii) the effect of such a financial inclusion is limited in terms of savings and lending activity, and (iii) there are no effects on welfare-related outcomes.

Chapter 3

Strengthening social insurance of savings groups

with Markus Frölich, Andreas Landmann, and P. Linh Nguyen

3.1. INTRODUCTION

How households handle the risks they face is a widely investigated topic and gained particular attention in development economics as one potential explanation and driver of poverty and poverty traps. For lack of sufficient insurance, households often engage in risk-management and -coping strategies that might keep them in poverty. Ex-ante households might forsake more profitable income-generating activities, keeping them engaged in low-risk but low-reward activities. Ex-post households might sell productive assets when hit with a shock (see Dercon (2002) for an overview on the topic). Muyanga et al. (2013) provide evidence that idiosyncratic shocks in the form death or illness are important determinants that keep households in poverty and these shocks are among the greatest concerns reported by households across low- and middle income countries (LMICs) (Cohen and Sebstad (2005), De Weerdt and Dercon (2006), Heltberg et al. (2015)). While formal health insurance is associated with a decline in out-of-pocket expenditure (e.g., see Jütting (2004), Dekker and Wilms (2010), Alkenbrack and Lindelow (2015)), fewer sells of assets in case of a shock (e.g., see Scheil-Adlung et al. (2006), Dekker and Wilms (2010), Parmar et al. (2012)), and other benefits (e.g. Strobl (2017) finds a negative association with child labour in favor of schooling), formal insurance coverage remains low in LMICs (Dercon (2002), Cohen and Sebstad (2005), Alkenbrack and Lindelow (2015)). Instead, households often engage in informal insurance schemes. For expenditure risks associated with the funeral of a family member, burial groups are a widespread scheme (Cohen and Sebstad (2005), Dercon et al. (2006)) and networks of friends and relatives play an important role (Fafchamps and Lund (2003), Fafchamps and Gubert (2007)). However, a large literature investigating informal risk-sharing suggests that informal risk-sharing is incomplete and that there is room for improvement (see Townsend (1994) for a seminal paper on the topic).

This chapter investigates if we can strengthen and formalize an informal insurance mechanism built into savings groups in rural Zambia: the social or emergency fund. The function of the fund differs across savings groups, but generally, each member regularly contributes money to the fund, which is then used to help members in need or to buy things for the group. The social fund already functions as an informal insurance mechanism (at least for some groups) but is often not governed by formal rules or depleted and thus unreliable as insurance. We implemented an intervention to improve the social fund as an insurance mechanism to expand the informal insurance available to savings group members. This would allow us to investigate to what extent such an informal insurance expansion affects risk-management and -coping of prevalent risks they face.

The contribution of our study is that we try to build upon a pre-existing informal insurance mechanism and try to improve its functionality. If successful, this would further allow us to investigate the causal effects of an informal insurance mechanism. Hence, our study lies between the literature descriptively exploring what types of informal insurance mechanisms exist and how they work and the literature studying the introduction of (semi-)formal insurance mechanisms or products.

To answer our research questions, we conducted a randomized control trial. Savings groups were randomly assigned to receive what we call a social fund intervention.[1] The particulars of this intervention varied across three implementing partners. Still, they consisted primarily of training on what insurance is and how the social fund can work as insurance against idiosyncratic risks.

We find evidence that the intervention successfully impacted the savings group informal insurance mechanism for one of the implementing partners. The contributions to the social fund increase significantly, and members are more likely to use loans provided by their savings group to finance the funeral expenditures they face. This falls short of introducing or formalizing insurance against such shocks but suggests an improvement in risk-coping strategies. A lack of findings for the savings group members' attitudes towards insurance suggests that more than our intervention is needed to convince members to adopt and formalize insurance through their savings groups. Given the limited effects on the usage of the social fund, it is not surprising that we do not find evidence that members are indeed better at coping with risks as a result of the intervention. We further do not find evidence for impacts on household welfare more generally, which may have resulted as indirect benefits of insurance or improved risk-coping.

The chapter is structured in three main parts. In section 2, we provide background information and describe the intervention. In section 3, we present the analysis for (i) social fund usage, (ii) attitudes and reported impact of shocks, and (iii) welfare-related

[1] Note that this intervention was cross-randomized with another intervention. Further, all the groups, including the control group, received general financial literacy training.

outcomes. Section 4 concludes.

3.2. BACKGROUND AND INTERVENTION

Since many savings group members do not partake in formal insurance schemes, strengthening the savings groups' insurance mechanisms could improve their risk-coping strategies. Initial qualitative and pilot data suggest that grants provided by the social fund were insufficient to cover funeral shocks, i.e., that funerals often impose costs of several hundred ZMW, but savings groups only provide grants of about 50 ZMW.

The social fund intervention had the goal of strengthening the group's informal insurance mechanism by providing information on what an insurable shock is, how insurance works, and how the social fund could provide such insurance. It was recommended that savings groups formalize in their institutions that members regularly have to contribute to the social fund and in return receive support in the form of a cash grant in case they face the funeral of a household member or illness and that the "insurance contributions" are not used for other purposes. Funeral and illness are good examples of insurable shocks and highly relevant in such contexts (Cohen and Sebstad (2005), De Weerdt and Dercon (2006), Heltberg et al. (2015)). While many of the savings groups are located in rural communities and their members are exposed to income risks from droughts or other unfavorable weather conditions, group-based insurance cannot cover shocks that affect all members at the same time. Instead, the focus of insurance should be on idiosyncratic shocks that affect members independently, which is likely for both funerals and illness in many cases. Further, the moral hazard should be considered, and shocks should be observable and verifiable. The group setting plays an important role here, and funerals and severe cases of illness should preclude moral hazard. Finally, the insurance must be reliable. Therefore, the suggestion to formalize it through the group's constitution and discourage using the fund for non-insurance purposes. If in place, such an insurance mechanism could alleviate the financial burden of funerals and illness and reduce the use of (in the long run) more costly risk-coping strategies such as the sale of productive assets or the take-up of loans with high interest.

As shown in Figure 1.2, the social fund intervention was implemented by the different NGOs at different times between the end of 2017 and mid-2018. Further note that while the implementation differed across NGOs, their operating districts differs as well (see Figure 1.1), and the type of group also varies as the different NGOs established, trained, and supported different concepts of savings groups. So we want to keep in mind that these highly correlated factors can also explain any difference in effect according to NGO affiliation.

Based on the monitoring data, we know of at least twelve groups that attended the social fund-related training but were not assigned to the social fund intervention. Note that this number reflects a lower bound. According to reports by the three NGOs, they implemented the training sessions following the agreed guidelines with a few exceptions, e.g., when savings groups came uninvited to training sessions, and the field officer could not reject them.

It should be noted that especially the implementation of NGO 4 is concerning. NGO 4 did not make us aware that they started their implementation such that we were not able to observe any of the training sessions. Further, discussions with representatives of NGO 4 and implementation documentation suggest that the training content was not in adherence with the concept of the intervention.[2] Compared to the other NGOs, there is little correlation between training participation as reported by the NGO and by the survey respondents. For NGOs 1 and 2, respondents in savings groups assigned to the intervention are considerably more likely to report participating in training covering aspects related to the social fund than respondents in savings groups not assigned to the intervention. At the same time, this is not the case for respondents in savings groups affiliated with NGO 4.

3.3. EMPIRICAL ANALYSIS

In this section we discuss the analysis of the social fund intervention. Recall that for this analysis savings groups affiliated with NGO 2 are omitted as they were not included in the randomization of the social fund intervention. First, we assess the balance of covariates between the social fund treatment group and the control group followed by the analysis of outcomes ordered according to a theory of change. We discuss whether the intervention affected any of several aspects related to the usage of the social fund. This continue with the analysis of outcomes related to risk-coping and attitude towards insurance. Finally, we investigate the impact on welfare-related outcomes.

3.3.1. Balance

Table 3.1 shows t-tests between treatment and control group of selected characteristics and outcomes measured at baseline to assess balance.[3] The only statistically significant difference is that slightly more savings groups in the treatment group meet every other week than the control group.

[2]E.g. NGO 4 reports that the savings groups extended the use of the social fund for purposes such as school fees and introduced loans.

[3]Note that we conducted these tests on 26 additional variables for which we found no statistically significant differences.

Table 3.1: Savings group and member characteristics

	Control mean	Control sd	Treatment mean	Treatment sd	β
Village level baseline variables					
Number of participating SGs in village	1.5	1.40	1.6	2.02	0.100
Urban or rural: Urban	.22	0.42	.2	0.40	-0.010
Urban or rural: Rural > 250	.54	0.50	.49	0.50	-0.055
Urban or rural: Rural <250	.24	0.43	.31	0.47	0.065
Mean score of additive food security index	1.4	0.78	1.2	0.81	-0.139
Mean number of months with food scarcity across HHs	1.6	0.84	1.5	0.78	-0.085
Savings group level baseline variables					
Meeting frequency: weekly	.64	0.48	.57	0.50	-0.065
Meeting frequency: every two weeks	.019	0.14	.056	0.23	0.039*
Meeting frequency: monthly	.35	0.48	.38	0.49	0.025
Household level endline variables					
Number of household members	5.8	2.19	5.7	2.24	-0.075
Member level endline variables					
Respondent is female	.81	0.39	.81	0.39	0.005
Age of respondent	45	12.16	45	12.28	0.129
Relation to household head: Household head	.42	0.49	.41	0.49	-0.015
Respondent is married	.74	0.44	.74	0.44	-0.001
Member level baseline variables					
Monthly contribution to SF in ZMW (winsorized)	6	7.41	6.1	7.71	0.292
Measure of risk aversion	2.7	1.39	2.6	1.40	-0.023
Locus of control (average of z-values)	-.0032	0.72	.038	0.74	0.038

Notes. The table shows characteristics and outcomes at baseline of savings groups and members by treatment status and NGO affiliation. Columns (1)-(2) and (3)-(4) show the mean and standard deviation for the control and treatment group. Column (5) informs about balance by displaying the coefficient β from the regression $X_i = \alpha + \beta D_i + \delta NGO_i$, where X_i is the respective variable, D_i treatment assignment, and NGO_i refer to NGO dummies. Significance of a t-test for $\beta = 0$ is referenced by $+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$. Columns (6)-(10) show the means of the variables for each of the different NGOs. Variables that refer to values are in ZMW and winsorized at the 1% and 99% quantile.

3.3.2. Social fund contributions and usage

Contributions. For the social fund to work as proper insurance that significantly reduces the impact of shocks and improves risk-coping, it needs to be large enough, in terms of value, to provide sufficient payouts in case of a shock reliably. Thus the first indicator we investigate is the members' contributions to the social fund. From Table 3.2 we can see that we do not find an effect on social fund contributions overall during mid- and endline. Still, we detect a statistically significant and sizable increase of 2.6 ZMW in the phone survey in December 2019 and an increase of 1.8 ZMW for savings groups affiliated with NGO 1 at endline. These are meaningful effect sizes given the average contributions in the control group of about 8-9 ZMW. Note that reported monthly contributions for NGO 1 and NGO 4 vary a lot across both savings groups and survey waves. Therefore, treatment effect estimates are not precise, and we cannot conclude that there are no meaningful effects. Overall the evidence is consistent with a positive effect of the intervention on contributions to the social fund driven by savings groups affiliated with NGO 1.

Table 3.2: Monthly social fund contributions

	Midline	Endline	Phone surveys		
	(1) July 2018 b/se	(2) July 2019 b/se	(3) December 2019 b/se	(4) April 2020 b/se	(5) November 2020 b/se
All NGOs					
Social fund treatment	1.40	0.93	2.62*	-0.16	-0.14
	(3.78)	(0.81)	(1.27)	(0.92)	(0.97)
Control group mean	12.2	8.75	8.08	9.52	9.45
Observations	424	424	384	379	375
NGO 1					
Social fund treatment	-6.54	1.80*	6.06+	1.12	0.025
	(8.98)	(0.85)	(3.61)	(0.97)	(1.22)
Control group mean	19.6	9.14	9.72	11.0	11.2
Observations	146	146	128	137	128
NGO 3					
Social fund treatment	0.098	0.23	1.05	0.33	0.55
	(1.12)	(0.31)	(0.67)	(0.35)	(0.36)
Control group mean	5.59	3.69	5.16	4.19	4.38
Observations	105	106	94	92	96
NGO 4					
Social fund treatment	3.56	-0.76	-0.49	-1.75	-0.55
	(3.76)	(1.50)	(1.44)	(1.90)	(2.19)
Control group mean	12.7	13.7	12.1	13.7	13.1
Observations	173	172	162	150	151

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$
Notes. Outcome for all regressions is monthly contribution to the social fund in ZMW. This is calculated based on frequency and amount reported in the savings group surveys. Information from the savings group member surveys is used to confirm and clean these reported values. Outcome is winsorized at 1% and 99%. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). For estimations on the whole sample, NGO fixed effects are always included, but excluded for NGO specific estimations.

Usage the social fund. While almost all groups have a social fund, how they use it differs a lot. Generally, savings group members can get grants (which they do not need to pay back) or loans (which they would need to pay back) from the social fund, both in the form of money and in-kind. While in principle, loans need to be paid back and sometimes even carry interest, they are not always fully paid back in practice. Table 3.3 shows results for whether loans or grants from the social fund were given to savings group members in the last twelve months and for the associated value of these pay-outs. We find neither effects on the extensive margin, i.e., whether there were pay-outs, nor the intensive margin, i.e., the value of pay-outs from the social fund. Note that there is a substantial difference in the number and value of pay-outs as measured by the in-person surveys compared to the phone surveys. There are several likely sources of this discrepancy. For the phone survey waves, we asked for the specific months and accumulated across various waves such that it is comparable to the twelve-month reference period used during the in-person interviews. Another explanation for the significant difference in control group means between mid-/endline, and phone surveys could be that the phone surveys covered the time of the COVID-19 pandemic. Further, there is potential selection into the phone survey, as not all groups were reached (in contrast to the in-person-based surveys during which all savings groups were reached), and some savings groups dissolved over the course of the study.

For each social fund pay-out, we also recorded the purpose. Since the intervention tried to strengthen the social fund as an insurance mechanism, we looked specifically at grants given out for insurable shocks in the form of funerals or illness.[4] We do not find effects when restricting the analysis to pay-outs with these purposes (see Table C.2 in appendix C) and also not when looking at each NGO separately (see Tables C.1 and C.3 in appendix C).

Social fund usage to alleviate shocks the households face. Further, we want to measure whether the social fund is used to support the households' financial needs in case of negative events such as funerals, medical expenditures, or business-related shocks on the household side. From Table 3.4 we can see that the social fund is rarely used to cover any shocks. Mainly these shocks are financed by selling agricultural goods or assets. We further do not find that the intervention increased the frequency with which the social fund is used. But at midline households from savings group members affiliated with NGO 1 increasingly used loans to cover shocks after receiving the social fund intervention. Upon closer inspection, this effect is primarily driven by loans provided by the savings group. While the intervention did not intend this, it indicates that some savings groups increased their willingness to financially support their members who faced a shock but not by providing grants. This effect is driven by loans to cover funeral costs (see Table C.5 in appendix C). Looking at the different types of shocks separately, we further find an effect on whether funeral costs were financed with the support of the social fund for NGO 4 at midline (see Table C.4 in appendix C). Further, we need to note that the sample size for any outcomes based on the occurrence of a shock is drastically reduced, as only a share of households faces the specific shocks (less than 30%).

Overall we find limited effects of the intervention primarily driven by NGO 1 that suggests that the intervention could improve the informal insurance mechanism of savings groups, even though not entirely as intended. This is in line with literature that offers a

[4] An insurable shock means that the shock reflects individual rather than common risk and also that there is not too much risk of moral hazard.

lack of financial literacy and trust as an explanation for low take-up of insurance products (e.g., Cole et al. (2013), Cai et al. (2015)) and Dercon et al. (2014), who show that training sessions can impact the uptake of rainfall index insurance and informal insurance among group members. While we find evidence that contributions to the social fund in groups affiliated with NGO 1 increase after the intervention, this does not translate into higher or more frequent grant pay-outs from the social fund. Instead, we find an increase in the role the savings group plays in financing funeral expenditure through loans. Members in treated savings groups affiliated with NGO 1 are more likely to report that they received a loan from the savings group to finance their funeral expenditures. However, we only find evidence for this in the short term after the intervention. This suggests that while the intervention was not able to formalize and introduce the intended insurance mechanism, to some extent, it affected the role the savings group play in risk-coping.

Table 3.3: Various measures for social fund usage

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se	(3) Phone surveys combined b/se
Any pay-outs from social fund			
Social fund treatment	-0.013	-0.045	0.034
	(0.046)	(0.047)	(0.039)
Control group mean	0.47	0.53	0.82
Any grants from social fund			
Social fund treatment	-0.024	-0.027	0.024
	(0.041)	(0.046)	(0.038)
Control group mean	0.33	0.43	0.79
Any loans from social fund			
Social fund treatment	0.016	0.00013	0.027
	(0.037)	(0.035)	(0.040)
Control group mean	0.18	0.14	0.29
Value of pay-outs from social fund			
Social fund treatment	31.3	-2.90	57.5
	(20.9)	(18.5)	(62.7)
Control group mean	94.53	105.75	422.76
Value of grants from social fund			
Social fund treatment	12.6	7.46	9.18
	(11.4)	(12.4)	(29.7)
Control group mean	45.15	59.02	221.37
Value of loans from social fund			
Social fund treatment	16.4	-7.86	49.1
	(15.5)	(12.3)	(47.6)
Control group mean	46.68	39.78	178.96
Observations	446	447	447

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Outcomes are accumulated over all reported pay-outs in a reference period of the last 12 months. For phone surveys this is based on the report from several waves asking about each month separately for the period of July 2019 to June 2020. Value outcomes are in ZMW and winsorized at 1% and 99%. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. For results for each NGO refer to Table C.1 in appendix C. For results related to the reported purpose of funeral or sickness, the target purposes of the intervention, refer to Tables C.2 and C.3 in appendix C for aggregate and NGO specific results respectively.

Table 3.4: Finance sources of various shocks by NGO

		(1) All NGOs b/se	(2) NGO 1 b/se	(3) NGO 3 b/se	(4) NGO 4 b/se
Support from SF for any shocks					
Midline	Social fund treatment	0.024	0.027	0.028	-0.0015
		(0.02)	(0.05)	(0.02)	(0.02)
	Control group mean	0.07	0.13	0.03	0.07
	Observations	1094	304	331	459
Endline	Social fund treatment	-0.0049	-0.022	0.022^{+}	-0.0069
		(0.01)	(0.03)	(0.01)	(0.02)
	Control group mean	0.07	0.12	0.01	0.06
	Observations	1809	532	483	794
Loans to cover any shocks					
Midline	Social fund treatment	0.024	0.12^{**}	-0.013	-0.0095
		(0.02)	(0.04)	(0.05)	(0.04)
	Control group mean	0.18	0.10	0.21	0.22
	Observations	1085	302	329	454
Endline	Social fund treatment	0.017	-0.00098	-0.011	0.011
		(0.02)	(0.03)	(0.03)	(0.03)
	Control group mean	0.18	0.15	0.19	0.19
	Observations	1813	533	484	796
Loans from savings group to cover any shocks					
Midline	Social fund treatment	0.033	0.14^{**}	0.027	-0.011
		(0.02)	(0.05)	(0.04)	(0.04)
	Control group mean	0.07	0.05	0.06	0.09
	Observations	706	214	211	281
Endline	Social fund treatment	0.0071	0.020	-0.016	0.016
		(0.02)	(0.04)	(0.04)	(0.02)
	Control group mean	0.09	0.12	0.11	0.06
	Observations	882	277	229	376
Sold goods to cover any shocks					
Midline	Social fund treatment	0.030	0.00073	-0.0089	0.024
		(0.03)	(0.06)	(0.05)	(0.05)
	Control group mean	0.46	0.45	0.60	0.38
	Observations	1137	328	339	470
Endline	Social fund treatment	-0.011	-0.028	0.016	0.0019
		(0.03)	(0.05)	(0.05)	(0.03)
	Control group mean	0.32	0.40	0.35	0.25
	Observations	1820	543	483	794

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Outcomes refers to reported source used to finance any shock such as funerals, medical expenditures, or business related shocks. Details on the different shocks and the number of times (in most cases this was once) they were reportedly financed with support from the social fund, loans, and selling agricultural goods and assets can be found in Tables C.4, C.5, and C.6 in appendix C respectively. Observations refer to households that reported shocks such as funerals, medical expenditures, or business related shocks. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). For estimations on the whole sample, NGO fixed effects are always included, but excluded for NGO specific estimations.

3.3.3. Attitude and reported impact of shocks

Knowledge and attitude towards insurance. The purpose of the intervention was to strengthen the savings groups' insurance mechanism. Still, social fund usage for the purpose of insurance is hard to measure as households irregularly face shocks. However, the intervention likely needs to change the savings group members' attitude towards insurance to be successful. We asked savings groups members whether they would participate in insurance schemes, their trust towards insurance companies, and we asked questions to assess their knowledge about and attitude towards insurance. Table 3.5 shows that a majority of savings group members are willing to join a hypothetical insurance scheme, but we do not find that the treatment affected this share. We find that about half of the members trust insurance companies but do not find an effect of the treatment on trust. We also do not find effects on knowledge about or attitude towards insurance as measured with our survey instrument. This lack of an effect is in line with the lack of findings that the savings groups use the social fund as insurance. If the intervention did not impact the members' attitudes towards insurance, we might not expect them to introduce insurance to their savings groups.

Reported impact of shocks and locus of control. The primary purpose of insurance is to alleviate the economic impact a shock has on the household. In the previous section, we discussed whether the treatment affected how households react to economic shocks and finance the costs associated with such shocks. Now we turn to the savings group members' subjective perceptions about the severity of the economic impact any funerals or economic shocks had on their household. The control group reports that on average 1-2 shocks in the last twelve months had a profound economic impact[5], suggesting potential benefits from insurance. From Table 3.6 we can see that the number of shocks that reportedly had a (serious) economic impact on the household does not differ between the treatment and control group. Even in the absence of a shock, one might benefit from insurance. Being insured could result in being less worried about potential shock and improve one's feeling of agency. We, therefore, assess both the internal and external locus of control of the savings group members. A high internal locus of control corresponds to feeling agency and control over one's life and circumstances. In contrast, an external locus of control corresponds to feeling as if external forces determine one's circumstances. An internal locus could be both cause and effect of having insurance, but we do not find an effect of the treatment on either locus. If anything, the internal locus of control in the

[5]This is conditional on reporting attending any funeral or any of the following economic shocks: bankruptcy, job loss, loss of harvest, house or equipment damage, loss of livestock, inflation, natural events such as droughts or floods, loan defaults, legal suits, communal and political crisis, theft, divorce, and other shocks.

control group is smaller at midline (significant at the 10% level). This could be explained by the intervention talking about external shocks, bringing uncertainty to participants' attention. While we find limited evidence suggesting improved risk-coping, we do not find evidence that this translates into reduced economic impact of shocks or feelings of more control over life's circumstances.

Table 3.5: Knowledge and attitude outcomes

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Would participate in insurance scheme		
Social fund treatment	0.014	-0.026
	(0.018)	(0.020)
Control group mean	0.85	0.74
Observations	1398	1889
Level of trust in insurance companies		
Social fund treatment	-0.058	0.0097
	(0.083)	(0.059)
Control group mean	2.30	2.24
Observations	748	1450
Trust in insurance companies		
Social fund treatment	0.010	-0.0040
	(0.033)	(0.026)
Control group mean	0.46	0.55
Observations	951	1615
Knowledge about insurance		
Social fund treatment	-0.013	-0.012
	(0.037)	(0.028)
Control group mean	-0.00	-0.00
Observations	1640	2114
Attitude towards insurance		
Social fund treatment	-0.044	0.012
	(0.042)	(0.044)
Control group mean	-0.00	-0.00
Observations	756	545

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. While the hypothetical participation in an insurance scheme and knowledge questions were asked to everyone, the attitude and trust outcomes are asked conditional on the respondent being familiar with the term insurance. Level of trust is on a 4 point scale (1=complete, 4=none at all). Trust refers to the first two levels. Knowledge about and attitude towards insurance are indexes, i.e. the mean of 5 and 8 items each of which is standardized to the control group. For results by NGO refer to Table C.7 in the appendix. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included.

Table 3.6: Reported economic impact of shocks and loci of control

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Average economic impact across shocks		
Social fund treatment	0.012	0.023
	(0.057)	(0.036)
Control group mean	1.85	1.67
Observations	1001	1772
# of shocks with serious economic impact		
Social fund treatment	0.025	-0.019
	(0.082)	(0.075)
Control group mean	1.02	1.72
Observations	1001	1772
# of shocks with economic impact		
Social fund treatment	-0.018	-0.044
	(0.078)	(0.084)
Control group mean	1.65	2.29
Observations	1001	1772
Internal locus of control index		
Social fund treatment	-0.055^{+}	0.031
	(0.029)	(0.030)
Control group mean	0.00	0.00
Observations	1640	2114
External locus of control index		
Social fund treatment	0.000017	0.036
	(0.031)	(0.032)
Control group mean	-0.00	-0.00
Observations	1640	2113

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Shocks comprises both business shocks and funerals. Impact was assessed on a 4 point scale (*1=serious*, *4=none at all*). *Serious economic impact* refers to the first level, *economic impact* refers to the first two levels. Locus of control indexes are the mean of 5 items which are standardized to the control group. For details on the external and internal locus of control indexes refer to Tables C.10 and C.9 in appendix C. For results by NGO and for business shocks and funerals separately refer to Table C.8 in appendix C. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included.

3.3.4. Welfare outcomes

In previous sections, we discussed that we do not find effects of the intervention on the usage of the social fund or the savings group members' attitude towards insurance, the two direct targets of the intervention. We can measure the reaction to shocks only for some shocks and only for households that faced such shocks and do not find evidence of reduced economic impact of shocks. From Table 3.7, we further see no effect on food security. However, savings group members could adapt their behavior even in the absence of a shock, e.g., if they feel better prepared to face shocks. We find evidence that savings group members affiliated with NGO 1 increased their contributions to the social fund and that more of the shocks their households faced were financed through loans given by the savings group. This could potentially lead to a greater feeling of security which could translate into more investment. Table 3.7 presents results of various measures related to household investments and welfare. We do not find effects on assets, investments, or expenditure and not on food security. Hence, while the evidence suggests some limited impact on risk-coping, this does not seem to translate into improved living conditions of the savings group members.

Table 3.7: Various indices related to household welfare

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Index HH animals purchased		
Social fund treatment	0.14	-0.014
	(0.090)	(0.016)
Index HH animals owned		
Social fund treatment	0.017	0.010
	(0.022)	(0.021)
Index HH agricultural inputs		
Social fund treatment	0.018	-0.020
	(0.030)	(0.030)
Index HH agricultural assets		
Social fund treatment	-0.013	0.016
	(0.017)	(0.024)
Index HH general assets		
Social fund treatment	0.026	-0.0078
	(0.016)	(0.014)
Index HH expenditures		
Social fund treatment	0.038	-0.042
	(0.031)	(0.033)
Index HH food security		
Social fund treatment	0.017	0.015
	(0.022)	(0.017)
Number of food insecure months		
Social fund treatment	-0.036	-0.015
	(0.070)	(0.062)
Control group mean	0.91	0.91
Observations	1777	2170

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). Observations reflects maximum number not accounting for missing values. Missing values: 22 and 44 missing for the food security index at mid- and endline. Indices are means of variables related to the index standardized to control group values, thus the control group mean is by construction 0 and omitted. For details on the components for each index refer to Tables C.11, C.12, C.13, C.14, C.15, C.16, C.17 in appendix C. NGO fixed effects are always included.

3.4. CONCLUSION

In this chapter, we study the effects of an intervention that aims to introduce or improve informal insurance mechanisms of savings groups in rural Zambia. The intervention consisted of training sessions explaining the concept of insurance and insurable shocks and giving advice on how the savings groups could implement insurance through their social fund. However, we find no evidence of increased usage of the social fund. While most savings group members reported their willingness to participate in insurance schemes, we do not find that the intervention improved the members' attitudes toward insurance. However, for one of the implementing partners, we find evidence for increased contributions to the social fund and that members are more likely to finance their funeral expenditures with the support of their savings group. While the evidence is limited, strengthened support when faced with shocks could result in improved risk-coping. But we find no evidence suggesting that members of savings groups who received the intervention are less economically impacted by shocks, more food-secure, or increased investments.

Chapter 4

Fairness of lotteries and survey compensation

4.1. INTRODUCTION

With increasing penetration rates of mobile phones in low- and middle-income countries (LMICs), restrictions due to epidemics or conflicts, the need for rapid and high-frequency information, and the interest in hard to reach, highly mobile or displaced populations, phone surveys gain importance and are increasingly conducted in LMICs. One major downside of phone surveys lies in their comparatively low response rates, which might increase cost and invalidate research designs.

Numerous approaches to increase response rates in various survey modes have been tested and there is considerable evidence showing that monetary incentives can increase response rates across modes (for a recent review and meta-analysis see Singer and Ye (2013) and Mercer et al. (2015)). Most of these studies are conducted in high income countries and for mail based surveys, but compensation for phone survey participation in LMICs received more and more attention from researchers in recent years (e.g. Hoogeveen et al. (2014), Leo et al. (2015), Vashistha et al. (2015) and Leo and Morello (2016)). While the primary goal of these approaches is higher response rates, other dimensions of data quality and representativeness are frequently investigated. However, ethical issues are barely touched upon (Singer and Ye (2013)).

With any form of compensation, there is a concern of whether people are coerced into participation (Singer and Couper (2008)). Other ethical concerns arise only with specific compensation schemes. Brown et al. (2006) and Zangeneh et al. (2008) link compensation in the form of lotteries to gambling, elaborating on the potentially exploitative nature of such an approach and how it might undermine informed consent. Further, lotteries create ex-post differential compensation across respondents which raises ethical concerns due to fairness considerations (Brown et al. (2006), Singer and Ye (2013)). This study aims to inform this discussion with empirical evidence and include the perspective of survey participants.

I investigate perceptions about survey participation compensation focusing on aspects of fairness with respondents in a phone survey panel in Zambia. In particular, I study

(i) how respondents feel about compensation that is determined by a lottery and thus leads to ex-post differential compensation across respondents, (ii) how winning or losing a lottery affects these perceptions, as well as (iii) appropriate levels of compensation based on the respondents' perspective.

To my knowledge, this is the first study to investigate survey respondents' fairness perceptions related to survey compensation in the form of lotteries. The only closely related study considers another common practice leading to differential compensation. Singer et al. (1999) study refusal conversion payments, i.e., an offer of extra incentives in case of an initial refusal to convert reluctant respondents and find that a large majority of respondents deem this approach unfair. While there are similarities to lotteries, e.g., both approaches lead to differential compensation, there are meaningful differences as well. In the case of lotteries, the differential compensation is based on luck, whereas with refusal conversion payments, it depends on the respondent's actions, favoring uncooperative behavior. Therefore it is worth investigating how fair or unfair respondents deem lotteries as compensation, in light of Brown et al. (2006) who argue that fair compensation implies equal outcomes.[1] Another differentiating factor from previous studies is that respondents in my study are asked questions about their savings group rather than themselves. This raises additional questions relating to fairness because only the respondent, and not the whole group, is compensated. The information collected in the survey belongs to the group, and in principle, many of the members could provide this information.[2]

In addition to investigating fairness perceptions about lotteries, I try to elicit an appropriate level of compensation from the respondent's point of view. Previous studies have investigated the effect of the level of compensation on response rates and data quality. A meta-analysis conducted by Mercer et al. (2015) finds a non-linear relationship between response rates and the size of the compensation. While these studies are essential for data collectors regarding the cost-effectiveness of different approaches, they neglect the respondent's opinions. Kropf and Blair (2005) investigate what role norms of cooperation play in achieving high response rates and stress their importance. This might be especially relevant for panel survey respondents, such as in my study, who have experience and

[1] By "equal outcomes," they mean not arbitrarily different compensation; arbitrarily being the keyword. E.g., respondents with a higher burden of participation can receive more compensation. Still, the luck involved in a lottery would be arbitrary. Thus, such a procedure results in arbitrarily unequal outcomes. Note that Brown et al. (2006) do acknowledge that sometimes unequal outcomes are necessitated by research design (e.g., in case of experiments), which is also the motivation behind the lottery compensation investigated in this study. In another study, coauthors and I investigate if respondents in a panel survey develop expectations of future compensation after receiving compensation for the first time.

[2] This would likely be a non-issue if the compensation only compensates for the respondent's time and effort, but in this context, at least for some of the respondents or their savings group members, the compensation might well exceed the opportunity cost of participation.

ongoing engagement, potentially leading to trust and feelings of reciprocity (as argued in Callegaro et al. (2014)). Given the concern of financial incentives crowding out internal motivation (see, e.g., Zutlevics (2016)), it is important to get an understanding of what respondents deem to be an appropriate amount of compensation.

Further, while there is a long tradition and numerous studies investigating survey compensation in high-income countries, evidence for LMICs is scarce. The practice of providing survey compensation is not uncommon in LMICs, but there are practical issues specific to phone surveys. In such surveys in LMICs, compensation is often paid in the form of airtime (Hoogeveen et al. (2014), Leo et al. (2015), Vashistha et al. (2015), and Leo and Morello (2016)) which might not always be appropriate or effective depending on the context. I showcase that mobile money can be used to transfer payments which might be a better incentive and more relevant in contexts in which respondents do not necessarily own or rarely use mobile phones.

The basis of my research design is the exogenous variation introduced by a lottery compensation. All respondents are offered a lottery in which they can win 80 Kwacha with a 50% chance. Since the lottery draw is random, I can compare respondents who won the lottery with those who lost to test whether perceptions about fairness and survey participation compensation depend on an ex-ante fair lottery outcome.

To study the difference between ex-ante and ex-post perceptions of fairness, I randomize when the respondents are asked about their perceptions. This cross-randomized experiment allows me to contrast a benchmark, i.e., the respondent being asked before the lottery, with their perceptions after winning or losing the lottery. Thus, in addition to capturing the difference in perceptions between respondents who won or lost, I further can investigate how perceptions are affected by the lottery rather than the outcome of the lottery alone.

To investigate an appropriate level of compensation as perceived by the respondents, I employ a survey experiment in which I randomly vary the amount the respondents are supposed to judge. This way, I avoid priming and framing effects which might occur when asking about different amounts in some order or at the same time. A compensation level might be differently perceived when put in direct comparison to another instead of when it is judged in isolation. The latter is more relevant in this context, as survey compensation is usually set by the data collector before the data collection and not in a negotiation. [3]

I find that most respondents believe survey respondents should be compensated, and

[3]Responsive survey compensation, i.e., compensation offered only after refusal, to some extent resembles a negotiation. However, it is more a series of predetermined one-time offers (mostly limited to only one such offer), and the respondent is not necessarily aware of this.

more than half think either all participants or none should be compensated. Further, in an interview about a savings group and its activities, approximately half of respondents believe that any compensation for the interview should be shared with the group. For the lottery offered to the respondents, I find mixed opinions with respect to its fairness. About 50% perceive the lottery as fair, while 40% perceive it as unfair. These perceptions change when respondents win or lose the lottery. Winning increases the share who perceive it as fair by about 10 percentage points, whereas losing decreases it by about the same amount. For an appropriate amount of compensation, I conclude that about half the respondents deem it fair when there is no compensation. I find considerable non-linearity: a compensation of 40 Kwacha is not perceived differently from no compensation, but a considerable share deems 80 Kwacha as fair.

The chapter is structured in three main parts. In section 2, I provide background information and describe the experimental design in detail. In section 3, I present the analysis for (i) general perceptions, (ii) perceptions related to the implemented lottery, and (iii) an appropriate compensation level. Finally, section 4 summarizes and discusses the findings.

4.2. BACKGROUND AND EXPERIMENTAL DESIGN

This study was conducted alongside a phone survey targeting savings group representatives in November 2020. The survey was one of several data collections for an impact evaluation in rural Zambia.[4] The savings groups were affiliated with one of four NGOs in seven districts across three provinces of Zambia (see Figure 1.1). From June 2016 to November 2020, several survey waves, both in person and on the phone, with members of the savings groups were conducted (see Figure A.1 in appendix A). The phone surveys target one representative of each savings group to answer questions about the group in general. The purpose of the survey is to elicit information about the group. We did not collect information about the respondent apart from the questions I analyze in this study. Further, note that the respondent is not necessarily the same for each group across survey waves. Previous respondents might not be available for an interview or no longer be part of their group. During the monthly phone surveys between April 2018 and April 2019, we usually reached less than three-quarters of the 520 groups, while in the surveys after the endline, we managed to reach a representative for almost all of the groups (for more details, compare Table 1.3). Since we knew that some of the initial 520 savings groups dissolved, we tried to reach only 511 in November 2020.

[4] In the following I refer to *we* when talking about activities related to the impact evaluation, and use *I* for everything exclusively related to this study.

During the endline survey, the respondents were paid compensation for participating in the interviews, but initially, there was no compensation for the phone surveys. In the November 2020 phone survey, we offered potential respondents to take part in a lottery in which they could win 80 Kwacha (or about 3.9 US$ at the time of the study[5]). We want to investigate if such compensation can help to keep response rates high in the phone surveys or whether there are adverse effects, e.g., respondents might form expectations and require compensation for participation in the future, or such compensation might lead to disputes within a savings group eventually leading to its dissolution.[6] In the following, I investigate fairness perceptions related to such compensation using survey experiments.

4.2.1. Experimental design

The main purpose of the experimental design is to investigate if fairness perceptions about survey compensation, especially in the form of a lottery, change depending on whether respondents won or lost the lottery. To evaluate this, the fairness perception questions are asked either before or after the lottery. Thus there are three groups to distinguish, (i) those asked the questions before the lottery, which serves as a benchmark, (ii) those asked after having won, as well as (iii) those asked after having lost the lottery. In addition, I want to investigate the appropriate level of compensation. Since I do not want to further[7] prime the respondents, and since I am interested in potential non-monotonic characterizations[8], I employ a survey experiment asking the respondents to judge one compensation level that was randomly selected out of three levels (*no compensation*, *40 Kwacha*, and *80 Kwacha*), instead of asking about several levels at once or openly asking for an appropriate value. Table 4.1 gives an overview of the design.

Lottery incentive. Before the interview starts, when the interviewers inform the potential respondent about the survey and ask for their consent, the interviewers explain that there is a lottery at the end of the interview. The interviewers were instructed to explain that there is a 50% chance to win 80 Kwacha and that any winnings would be transferred to a mobile money account of the respondent's choice but could not be paid out in cash. Then, immediately after the main questionnaire, the interviewer is informed by the CATI software whether or not the respondent won the lottery and passed this

[5]Three reference points: (i) this is the amount we paid the interviewers per interview, (ii) the average monthly savings contribution to their savings groups is reportedly about 105 to 125 Kwacha, and (iii) the average monthly income in 2015 in rural Zambia is about 1,800 (LCMS (2015))

[6]We had one report from an interviewer that this was the case for at least one group after the survey compensation during the endline was not shared with the group.

[7]The winning amount of 80 Kwacha might prime the respondents and serve as a reference point.

[8]No compensation might be fine with respondents, whereas a small amount might be considered inappropriate.

Table 4.1: Design of the experiments

	Won	Lost	Questions before	Total
No compensation	40	41	81	162
40 Kwacha	42	42	78	162
80 Kwacha	41	43	85	169
Total	123	126	244	493

The table shows the number of savings groups representatives assigned to each of the randomized treatment arms. All of the following were cross-randomized with equal probability: winning or losing the lottery, questions being asked before or after the lottery, and the compensation level I inquired about. Note that the number of "Questions before" groups, the benchmark group, are roughly twice as large, as there is no distinction between whether the respondents were selected to win or lose the lottery as in the case for those asked the questions after the lottery.

information on.

Since the survey was collected via phone, we faced the challenge of making payments to the respondents. Previous studies (e.g. Hoogeveen et al. (2014), Leo et al. (2015), Vashistha et al. (2015), and Leo and Morello (2016)) paid phone survey respondents in airtime, but in our context, airtime did not seem to be useful to the potential respondents and thus would not be a good incentive. Often respondents do not own the phone they use or mostly use their phones to receive instead of making calls, in which cases airtime would be of little use to the respondent. We, therefore, opted for mobile money instead. The problem with mobile money is that not every potential respondent has access to an account, and there is a withdrawal fee if they want cash.[9] The respondents could therefore also give the account details (number and name) for any account they want their winnings transferred to. In few cases in which even that was not possible, the interviewers sent the money to a mobile money agent from whom the respondents received cash. In the end, every respondent who won the lottery received their winnings.[10]

We randomized the lottery winnings within strata with a probability of 50% before the start of the data collection.[11] The strata are based on the province and NGO affiliation

[9] While in principle, each mobile account can be used for mobile money (and all respondents have access to one; otherwise, they would not be interviewed), in practice, these need to be registered, and the money would first need to be collected from an agent to be of use for most respondents.

[10] At first, a few respondents were reluctant to take the money, but in the end, everyone did. As payment confirmation, I collected payment receipts for every lottery winner from the interviewers, i.e., each transfer they made was documented and included the date, account number, and name of the recipient.

[11] We used Stata's (Version 16.0) random number generator to conduct the randomization and loaded the allocation into the software. 50% is only approximate for strata of an odd size.

of the group as well as the number of previous phone survey waves for which we collected an interview.[12] Each stratum was constructed such that it comprises at least 16 groups.[13]

Position of questions. As I want to investigate the respondents' perceptions about the fairness of the lottery and survey compensation in general, I included related questions at the end of the interview. Since I expected these perceptions to change depending on whether or not the respondent was lucky and won the lottery, these questions were randomly placed before or after the information about the lottery's outcome. The CATI software independently randomized the position of the questions. With a 50% chance, the respondent was asked the question before being informed about whether or not they won and with a 50% chance afterward.[14] Thus leaving us with three distinct experimental groups of interest: those answering the questions after they won the lottery, those responding after they lost, and those answering before being informed. For details on the questionnaire module, refer to section D.A in appendix D.

Compensation level in question. In a survey experiment, the respondents are asked to judge how fair they feel a specific amount of money is as compensation for participating in a survey. This specific amount varied randomly between three levels: 80 Kwacha (the amount they can win in the lottery), 40 Kwacha (the amount that could be paid to every respondent keeping the total amount spent equal), and no compensation (which was the case in all previous phone survey waves in this panel). How the respondents interpret fairness is left up to them.[15] They can say it is fair, unfair, or neither. Further, they are asked to distinguish between it being unfair because the amount is too high or too low. This survey experiment allows me to get at the appropriateness of different levels of survey compensation outside of a direct comparison which might cause framing effects. While this comes at the cost of sample size compared to, for example, asking about all amounts at once, I want to know whether 40 Kwacha is believed to be appropriate as such and not when compared to 80 Kwacha or nothing. Asking about 80 Kwacha might change the answer to the question related to 40 Kwacha, and one might expect that respondents are nudged to give monotonic replies, i.e., if 40 Kwacha is reported to be unfair, then no compensation is also reported to be unfair, whereas in isolation a different answer might be given. Due to the lottery, respondents might already be primed towards 80 Kwacha

[12] The number of survey waves was generally divided into three categories: up to 5, 6-10, and 11-16 successful interviews. We chose to stratify across previous phone survey waves as the response rate is the primary outcome of the investigation for which we implemented the lottery incentive.

[13] In two cases in which strata were too small, two neighboring categories of number of survey waves were combined for these province and NGO combinations.

[14] I used ODK's (Version 1.28) build-in random number generator for the randomization.

[15] E.g., they might take themselves, who so far did not receive any compensation, as a reference, or they could judge fairness with respect to the time it takes to participate in a survey.

(the winning amount). These priming effects might be more pronounced for those asked after the lottery compared to those asked before.

4.2.2. Randomization fidelity and limitations

If everything were implemented according to instructions, respondents would not select into sample based on the treatment. About 97% of the target population is part of the sample, and I detect no differences across treatment arms. However, respondents answering the fairness-related questions after having lost the lottery are 5 percentage points more likely (significant at 10% level) to have had a full interview prior to the experiment. The difference between any and a full interview is that there is only a full interview if the group was active in the months since the last phone survey. I conducted robustness checks including a dummy for full interview. Overall this does not affect the results.

Further, since the position of the fairness-related questions and the compensation level question experiment were independently randomized, I checked whether there is evidence of dependence. I find no such evidence further strengthening the idea that the experiment was implemented as designed and no adjustments are needed. The lottery itself was randomized before the data collection and loaded into the CATI tool, such that the interviewers were not able to affect which respondents win or lose the lottery.[16]

While the lottery treatment was stratified by group characteristics, the interviewer might play an important role in this context. Since the interviewers were not randomized to the groups, there is potentially a correlation between interviewer and treatment assignment. This means that if interviewers affect respondent behavior or if they differently document the respondent's answers this could lead to spurious findings. Including interviewer fixed effects tends to lead to small increases in precision and effect sizes.[17] This suggests that the interviewer indeed affects the respondents' responses and that there is a correlation between the interviewer and treatment assignment.

Since the purpose of this survey was to collect information on the respondent's savings group, I cannot associate any outcomes or effects with respondent characteristics. The only information specifically relating to the respondent are their perceptions about survey compensation studied in the following sections. I plan to remedy this limitation by asking additional questions about the respondent's characteristics and their perceptions unrelated to survey compensation in a future data collection. While I expect this to be fruitful, there is the possibility that not all respondents will be covered because they no

[16] This was also communicated to and by the interviewers that the outcome of the lottery would not change if the interview is repeated and is not affected by the responses provided during the interview.

[17] Note that I present my findings excluding the interviewer fixed effects as the additional precision does not result in more findings and the increases in effect sizes are only small.

longer participate in the survey or are no longer part of the savings group.

4.3. EMPIRICAL ANALYSIS

My main specification is an OLS regression of the following form:

$$Y_i = \alpha + \beta D_{Won} + \gamma D_{Lost} + \sum_j \delta_j + \epsilon_i$$

where Y_i is the outcome of interest, D_{Won} and D_{Lost} are indicators for whether they won or lost the lottery before answering the questions, δ_j refers to strata fixed effects. The standard errors ϵ_i are Eicker-Huber-White standard errors.[18] The reported parameters are the estimates $\hat{\beta}$ and $\hat{\gamma}$ of the average effects.

4.3.1. General perceptions about survey compensation

Table 4.2: General perceptions about survey compensation

	Respondent compensation			Either all or none			Group compensation		
	(1) Level b/se	(2) Agree b/se	(3) Strongly b/se	(4) Level b/se	(5) Agree b/se	(6) Strongly b/se	(7) Level b/se	(8) Agree b/se	(9) Strongly b/se
Lost lottery	0.23*	0.03	0.13*	0.03	0.01	0.01	-0.13	-0.06	-0.08*
	(0.10)	(0.04)	(0.05)	(0.11)	(0.05)	(0.04)	(0.10)	(0.05)	(0.04)
Won lottery	0.09	0.03	0.06	-0.15	-0.05	-0.01	-0.14	-0.09+	-0.03
	(0.10)	(0.04)	(0.05)	(0.12)	(0.05)	(0.04)	(0.12)	(0.05)	(0.04)
F-test: Won=Lost	0.19	0.95	0.25	0.20	0.33	0.58	0.90	0.54	0.20
Adj. R^2	0.08	0.07	0.04	0.06	0.10	0.27	0.34	0.30	0.42
Mean: Questions first	4.02	0.80	0.33	3.66	0.65	0.20	3.37	0.47	0.27
Observations	490	490	490	491	491	491	493	493	493

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Columns (1)-(3) refer to statement "Survey respondents should be compensated for their participation in a survey interview", columns (4)-(6) refer to "Either all or none of the respondents in a survey should be compensated", and columns (7)-(9) refer to "When there is a compensation paid to a respondent for taking part in an interview about her/his savings group, then the compensation should be shared with the group". Respondents were asked to answer on a 5-point scale from *strongly disagree (=1)* to *strongly agree (=5). Level* refers to the numeric value of the scale, *Agree* refers to a dummy for whether the respondent answered *agree* or *strongly agree*, and *Strongly* refers to a dummy for whether the respondent answered *strongly agree*. For more details on the questions and distribution of responses refer to D.A. Number of observations varies due to *don't know* and *refuse to answer*. All estimations presented in this table include strata, but no interviewer fixed effects.

Before going into more detail about the perceptions related to the lottery we implemented, I first want to discuss respondents' perceptions related to survey compensation in general based on three questions relating to (i) should there be compensation, (ii) should it be equal across respondents, and (iii) should it be shared with the savings group.

[18] I used Stata's (Version 16.0) *reg* command with the *robust* specification for standard errors.

The first question asked whether the respondent agrees that survey participants should be compensated. The vast majority agree, and less than 10% disagree. From columns (2) and (3) of Table 4.2 we can see that this belief is strengthened after losing the lottery. When asked after experiencing the loss, 13 percentage points more respondents strongly agree that there should be survey compensation. But I do not find an effect on agreeing, indicating opinions do not change but rather are strengthened.

With the second question, I try to get a better idea about what matters in terms of fairness. The respondents are asked whether they agree that either all or none of the respondents of a survey should be compensated. Agreeing with this statement implies (strong) preferences for equal treatment. Most respondents agree with the statement, but it is a smaller share compared to the question about compensation, and about 20% disagree. From columns (4) to (6) of Table 4.2 we do not find a difference depending on whether they won or lost the lottery before being asked the question. One might expect to see some changes, e.g., the winners having won but knowing that others did not, might prefer this scenario compared to the situation in which no one got compensated. However, this question was not about this specific survey, and all respondents had the chance to win the lottery, so everyone was equally compensated from an ex-ante point of view. Thus, based on stated preferences, equal treatment seems to matter to most of the respondents.

The third question inquires about sharing any compensation with their savings group. Reportedly there were disputes in a few savings groups after respondents did not share the compensation they received for participating in the endline survey with their groups. Opinions are mixed, about 50% of respondents agree, while about 30% disagree and 20% neither agree nor disagree. Columns (7) to (9) of Table 4.2 seem to indicate that being asked this question after the lottery makes respondents tend to disagree with this statement. This might be explained by the lottery being more on top of their head, making the trade-off between self-interest and interest for others more salient to them.

4.3.2. Perceptions about implemented lottery

The respondents were asked how fair they think that some respondents win 80 Kwacha in the lottery while others do not. They were able to answer on a 5-point scale from *Very unfair*(=1) to *Very fair*(=5). When asked before they participate in the lottery, the average value of about 3 corresponds to neither fair nor unfair, but opinions are polarized, with about 50% thinking it is fair and about 45% thinking it is unfair. Columns (1) to (3) of Table 4.3 inform us about how this opinion changes when the respondents are asked after winning or losing in the lottery instead of before. After losing, 11 percentage points fewer respondents think that the lottery is fair, whereas after winning, 12 percentage points more respondents hold this opinion. For the stronger opinion that the lottery

Table 4.3: Perceptions about implemented lottery

	Fairness			Satisfaction		
	(1)	(2)	(3)	(4)	(5)	(6)
	Level	Yes	Very	Level	Yes	Very
	b/se	b/se	b/se	b/se	b/se	b/se
Lost lottery	-0.42**	-0.11*	-0.05+	-0.68**	-0.30**	-0.12**
	(0.13)	(0.05)	(0.03)	(0.11)	(0.05)	(0.04)
Won lottery	0.29*	0.12*	0.07+	0.21*	0.01	0.26**
	(0.13)	(0.05)	(0.04)	(0.08)	(0.03)	(0.05)
F-test: Won=Lost	0.00	0.00	0.00	0.00	0.00	0.00
Adj. R^2	0.31	0.25	0.11	0.30	0.19	0.24
Mean: Questions first	2.92	0.49	0.10	4.04	0.87	0.23
Observations	492	492	492	491	491	491

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Columns (1)-(3) refer to question "How fair do you think it is that some respondents win 80 Kwacha in the lottery while others do not", and columns (4)-(6) refer to "How satisfied are you that you had/have a chance to win 80 Kwacha for your participation in the interview". Respondents were asked to answer on a 5-point scale from *very unfair (=1)* to *very fair (=5)*. *Level* refers to the numeric value of the scale, *Yes* refers to a dummy for whether the respondent answered *fair* or *very fair*, and *Very* refers to a dummy for whether the respondent answered *very fair*. For more details on the questions and distribution of responses refer to D.A. Number of observations varies due to *don't know* and *refuse to answer*.All estimations presented in this table include strata, but no interviewer fixed effects.

is very fair estimated effects are not significant at the 5% level. I find a decrease of 5 percentage points for those losing and an increase of 7 percentage points for those winning. Considerable effects sizes, given the prevalence of 10% among those being asked the questions before the lottery.

One concern is that fairness is a vague concept and can mean different things to different people. To distinguish fairness from their satisfaction about winning, the respondents were asked how satisfied they are about the fact that they have or had the chance to win 80 Kwacha. They gave answers on a 5-point scale from *Very dissatisfied*(=1) to *Very satisfied*(=5), and when asked before the lottery, the average value was 4, which corresponds to being satisfied. A majority of 87% reported being satisfied with being offered the lottery, giving support to the idea that the measured fairness perception is about the unequal treatment rather than the value of potential winnings. From columns (4) to (6) from Table 4.3 we can confirm what we would expect that losing is dissatisfying while winning is satisfying. Losing makes 30% of respondents unsatisfied and 12% not very satisfied. On the other hand, winning makes 26% very satisfied, while it does not make anyone satisfied. This suggests that the minority unsatisfied with the lottery cannot be swayed from their opinion by winning. Whether this is due to fairness considerations, a too-small winning amount or something else is not clear.

4.3.3. Appropriate compensation level

In the survey experiment, the respondents are asked to judge a hypothetical compensation for survey participation, the level of which is randomly varied between no compensation, 40 Kwacha, and 80 Kwacha. The outcome of interest is whether or not the compensation level is deemed fair. Note that those that do not consider the compensation level as fair do not necessarily perceive it unfair; about 10% perceive it as neither fair nor unfair. In the following, I, therefore, look at both the comparison between fair vs. not fair and fair vs. unfair. Since the sample size does not allow me to distinguish between the differences of compensation levels based on lottery outcomes, I consider the following reduced model, which only differentiates between whether the compensation was 40 or 80 Kwacha and whether the question was posed after a win or loss but without interactions.[19]

$$Y_i = \alpha + \alpha_{40} D_{40} + \alpha_{80} D_{80} + \beta D_{Won} + \gamma D_{Lost} + \sum_j \delta_j + \epsilon_i$$

where Y_i is an indicator for whether the compensation is deemed fair, D_{Won} and D_{Lost} are indicators for whether they won or lost the lottery before answering the questions, D_{40} and D_{80} are indicators for whether the question inquired about 40 or 80 Kwacha, δ_j refers to strata or interviewer fixed effects. The standard errors ϵ_i are Eicker-Huber-White standard errors.[20]

When asked about no compensation, 55-65% of respondents report that this is fair. From Table 4.4 we can see that when asked about 40 Kwacha, there is no difference, while when asked about 80 Kwacha instead, 20-30 percentage point more respondents deem this level of compensation to be fair. So while a slight majority believe no compensation to be fair, a significant share only considers it fair when there is a sufficiently high compensation, where sufficiently high means higher than some cut-off between 40 and 80 Kwacha. Further note that respondents who lost the lottery before being asked the question are about 10 percentage points less likely to believe the compensation level is fair, while I detect no effects for those that won.

[19]For the analysis of a model including the interactions, refer to section D.B in appendix D.

[20]I used Stata's (Version 16.0) *reg* command with the *robust* specification for standard errors.

Table 4.4: Reduced specification of survey experiment

	Compensation is fair vs not fair				Compensation is fair vs unfair			
	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se	(7) b/se	(8) b/se
α_{40}	0.01	-0.00	-0.01	-0.01	0.00	-0.00	-0.00	-0.01
	(0.06)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)
α_{80}	0.28**	0.31**	0.30**	0.32**	0.23**	0.26**	0.25**	0.27**
	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)
β	0.00	-0.02	-0.03	-0.03	0.04	0.01	0.00	0.00
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
γ	-0.10+	-0.12*	-0.11*	-0.12*	-0.09+	-0.10*	-0.10*	-0.10*
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)
F-tests:								
$\alpha_{40} = \alpha_{80}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\beta = \gamma$	0.11	0.08	0.17	0.12	0.04	0.03	0.09	0.05
Model description:								
Strata FE	✗	✗	✓	✓	✗	✗	✓	✓
Enumerator FE	✗	✓	✗	✓	✗	✓	✗	✓
Adj. R^2	0.08	0.20	0.18	0.23	0.07	0.28	0.19	0.32
Mean: no compensation	0.56	0.56	0.56	0.56	0.63	0.64	0.63	0.63
Observations	490	490	490	490	448	448	448	448

Notes. Columns (1)-(4) refer to various specifications with an outcome that is 1 if the compensation is deemed fair and 0 otherwise. Columns (5)-(8) refer to various specifications with an outcome that is 1 if the compensation is deemed fair and 0 if it is deemed unfair, omitting respondents that deem it neither fair nor unfair. Specifications differ depending on whether interviewer or strata fixed effects are included. β and γ are estimates for the effect of having won or lost the lottery before answering the questions, α_{40} and α_{80} are estimates for the compensation level being 40 and 80 Kwacha. The omitted categories are *no compensation* and being asked the question before the lottery draw. Rows under *F-tests* display corresponding p-values. A more elaborate model including all interaction terms is presented in Table D.1 in section D.B in appendix D.

4.4. CONCLUSION

To spotlight ethical aspects related to survey participation compensation in increasingly more frequently conducted phone surveys in LMICs, I study respondents' perceptions about such compensation in a phone survey targeting savings group representatives in rural Zambia. The survey under study is part of a panel, and respondents were compensated with a lottery in which they had a 50% chance to win 80 Kwacha for the first time in this panel. About 4 out of 5 respondents agree that survey participants should be compensated, an opinion that is more pronounced in those who have lost the lottery. From a survey experiment, I conclude that about half the respondents deem it fair when there is no compensation. Compensation needs to be at least somewhere between 40 and 80 Kwacha for a considerable share of respondents to change their minds. This is in line with Leo and Morello (2016) who do not find differential effects on response rates across incentive levels in their experiment as the US$ value of their highest incentive level of 1 US$ was well below 40 Kwacha (approx. 1.95 US$ at the time of the study).[21] Note that those who lost the lottery tend to report more often that any of the three compensation levels is unfair, similar to the finding that they are more likely to agree that there should be compensation. Further, more than half of the respondents report a strong preference for fairness and equal outcomes by agreeing that either all or none of the respondents should be compensated. This is similar to Singer et al. (1999) who find that three-quarters of respondents deem refusal conversion payments as unfair. About half of the respondents also agree that if a respondent gives information about their savings group, this respondent is supposed to share any compensation they receive. I find no evidence that these opinions about fairness related to survey participation compensation, in general, are affected depending on whether or not the respondent has won or lost the lottery.

Random allocations, such as the common practice of experimental research, are subject to various criticisms. Many in the research community raise ethical concerns related to the fairness implied by arbitrarily unequal resource allocation (Rayzberg (2019) summarized this discussion recently). More information about how participants evaluate the fairness of such research designs could inform these discussions. For perceptions directly related to this survey's lottery compensation, I do find that fairness perceptions change depending on whether or not the respondent won or lost the lottery. While initially opinions are mixed, with about 50% of respondents agreeing that the lottery is fair, while about 40% disagree, winning makes it about 10 percentage points more likely that a respondent agrees, whereas losing makes it about 10 percentage points less likely. Thus the lottery creates a gap of 20 percentage points in ex-post perceptions. Note that a respondent's fairness perception is difficult to disentangle from their satisfaction with the lottery outcome, and this should be kept in mind as a caveat to my results. This being said, the evidence presented in this study indicates that respondents evaluate the fairness of a lottery differently after its outcome. Extrapolating my findings to a case in which study participants' opinion about the fairness of an experiment changes after the study is implemented might raise concerns about the validity of their initial consent. How should we think about the consent to participate in an experiment when the participant's opinion about the fairness of randomization changes in a predictable manner after the random draw? More research into the persistence of such changes and fairness perceptions in different contexts is needed to evaluate such concerns more thoroughly.

[21] Leo and Morello (2016) implemented the experiments with varying incentives to study response behavior in Tanzania and Ghana in 2015. Of course, it is difficult to compare US$ values across countries and time. Still, the comparison might be justified as the GDP per capita of Zambia in 2020 lies between that of Ghana and Tanzania in 2015.

Chapter 5

How Informed is Consent?

with Alexandra Avdeenko

5.1. INTRODUCTION

With an increase in primary data collections and field research, the development of adequate policies and practices to ensure scientific integrity is critical. Meeting high ethical and legal standards continues to be among field work's core challenges and requirements and is essential to form stronger norms and trust in the results (Asiedu et al. 2021; Gueron 2017).

Formally, research and data protection guidelines command that the purpose of a data collection and information about data processing need to be in clear and simple language, understandable, and easily accessible. Research teams present this information when they first encounter potential survey respondents. To acquire informed consent, the teams usually follow specific procedures. After a short introduction, a standardized text lays out the rights of the respondents and the risks of harm associated with participating in a research study. The interview or intervention only proceeds if consent is given.[1] However, it is not clear whether these procedures sufficiently inform potential participants about their rights and the risks with respect to data protection, especially in largely illiterate populations often studied by development economists.

The topic of informed consent gained little attention in economics, despite its ethical and methodological implications. Carefully explaining (legal) rights to individuals is a core element of numerous empowerment programs in Low- and middle-income countrys (LMICs). Yet less attention has been paid to raising awareness about the value of personal information in this context. Acquiring genuinely informed consent to the collection, use, and disclosure of personal data is essential for at least three reasons: (1) the ethical

[1]For more details on the process, read Glennerster (2017) on"The Practicalities of Running Randomized Evaluations: Partnerships, Measurement, Ethics, and Transparency" which covers the practical ethical issues a researcher conducting randomized evaluations must take into account when designing and carrying out their research. For the far-reaching legal implications of providing consent, please refer to Solove (2012), who also discusses the problems of "the uninformed individual" stemming from common cognitive problems which undermine privacy self-management.

aspects related to different levels of being informed, (2) the data quality, and (3) the external validity of results and policies. First, ethical research requires informed consent of its study participants, not only consent. Data collection, storage, and analysis may expose the poor to a range of new vulnerabilities.[2] Especially when surveys are collected in countries with low literacy rates and low levels of education, practices like "consent forms" might fail to fully and sufficiently inform the respondent.[3] Presented with complex and abstract concepts, terms such as "data protection" and "confidentiality", the often illiterate study population is asked to give consent potentially without fully understanding the implications. Further, the survey may raise wrong expectations of immediate aid.[4] A lack of understanding of the research purpose may effectively coerce vulnerable people in need into participation. Second, methodologically, a better understanding of rights may alter data quality. Yet, once accurately informed, low consent rates could present a real risk to research studies, putting an additional burden on research budgets and timelines. Requiring informed consent may also affect data quality and even conflict with the study design. Finally, the topic is related to a highly relevant aspect of field experiments: External validity. Alternative approaches to inform people about their data protection rights and the purpose of collecting personalized information need to consider potential implications for the representativeness of the research sample.[5] In essence, this study relates to the challenges to find regulations for the usage of private data given the complexity of the topic and behavioral biases involved (Acquisti et al. 2007, 2016; Benndorf and Normann 2018) and to the debate on credibility, replicability, and transparency of economic research which questions the standard economic approaches to conducting empirical research (Asiedu et al. 2021; Christensen and Miguel 2018; Kaplan

[2]Hilbert (2016), for instance, argues that the caveats of the Big Data debate, such as privacy concerns and human resource scarcity, would be aggravated in LMICs by long-standing structural shortages in the areas of infrastructure, economic resources, and institutions.

[3]Even though the respondent is undoubtedly aware that she provides information during an interview, it is not always clear whether she is aware of the consequences of doing so. And even if consent might be assumed to be implicit in a survey setting, this implicit consent applies to the data collection itself, but not necessarily to data storage and analysis.

[4]Instead of considering what happens to their data, respondents often expect that their lives will improve directly or indirectly as an outcome of the interaction with the research teams (Alderman 2013). The large majority of aid programs applies targeting which largely relies on surveys, and not surprisingly, research participants may expect to become eligible for a program.

[5]Duflo et al. (2007) have a whole section on the topic "External Validity and Generalizing Randomized Evaluations" laying out the relevance of the issue. They broadly define "external validity" as "validity whether the impact we measure would carry over to other samples or populations". If, in our setting, the willingness to participate in a study is affected by the consent process, this indicates that requiring and obtaining informed consent might create selection into the sample. If the provision of consent and being informed is related to the characteristics of potential respondents, this could bias statistics derived from the survey for the actual population of interest.

et al. 2020; Ludwig et al. 2019; Olken 2015; Ravallion 2020).[6]

In this study, we ask the following research questions: Are potential respondents more likely to decline an interview if the consent form is presented differently? How informed are respondents, and can amended procedures improve understanding? Does it affect the quality of their responses? The answers are hard to anticipate, and there is little to no empirical evidence to go on. To shed light on this topic, we approach 3,964 potential research participants in rural Sindh, Pakistan, in winter and spring 2021. The illiteracy rates in rural Sindh are high, as are the needs of the people.[7] We randomly varied two alternative approaches of presenting the consent form and assessed their understanding. First, instead of only reading out the consent form, the enumerator presented a short video to the potential respondent, which visualized processes related to confidentiality and data protection and the interview itself. 44% of potential respondents were assigned to this treatment. Additionally, a second approach to increase informedness which combined the video with an interactive scripted process, was assigned to 6% of potential respondents. During this process, the enumerator read the consent form, but between paragraphs asked questions about the information in the paragraphs to check whether the respondent retained the information and to engage her. If the responses indicated misunderstanding, the enumerator repeated the relevant information. This experimental design allows us to study whether alternative approaches change response behavior and the understanding of respondents and whether the two are related.

In the first set of results, we study the level of understanding of the consent form content. Using objective measures, we find that overall, respondents are not sufficiently informed. Less than 20% respondents are well informed, i.e., assess five out of six statements related to informed consent correctly. On average, a respondent assesses only 58% of the questions accurately, only 8 percentage points more compared to random guessing. Study participants have particularly little knowledge related to the voluntary nature of participation in the study, and only about every fifth person understands the purpose of the data collection.

Second, we tested whether the procedures to acquire informed consent should be amended. The goal was to find ways to improve information provision at the beginning of an interview. We find that the combination of watching a video and being presented the consent form in an interactive process improves the informedness of respondents.

[6]The debate has spurred the emergence of alternative approaches, the usage of pre-analysis plans, platforms to predict research results, more piloting, and critical information to be added and discussed in research papers. For example, intending to strengthen the norms, Asiedu et al. (2021) promote the usage of structured ethics appendices in social science papers.

[7]According to the Bureau of Statistics, the literacy ratio in rural Sindh was 44.1% in 2017/8, and Sindh had a (moderate or severe) underweight prevalence of 42 % Link.

Compared to the control group, a treated respondent is about 9 percentage points more likely to assess questions related to the voluntary nature of participation correctly. On the other hand, self-reported understanding decreases with alternative approaches. Given that self-reported understanding is not reflecting the understanding according to our objective measurement (i.e., it could reflect overconfidence in understanding), this detected decrease might be preferred.

Third, we elaborate on potential implications for the survey data quality. While researchers want respondents to be informed about their participation in a study, about their rights, and what will happen with the data, it is well-known that humans - once informed of being studied - might behave differently (the so-called Hawthrown effects). Survey respondents might provide more accurate information or overall more information during the interview, depending on whether the intervention increased or decreased their willingness to provide information. We find that the augmented process of inquiring consent does not affect consent rates, which are overall very high in our sample. Given that the consent rate is close to 100% and not affected by the treatments, we do not investigate any effects on sample composition. Finally, we investigate item non-response rates, i.e., the share of questions for which the respondent did not respond. We find no effects on response behavior on this intensive margin. This provides evidence that concerns related to trade-offs between informing respondents and response behavior are negligible for our process.

Our study contributes to a better understanding of how to improve the adaption of survey research protocols to vulnerable populations, especially in field experiments. The design of our study can be informed by evidence from medical sciences, yet it is unprecedented in the field of development economics.[8] In medicine, several studies look at informed consent. For example, Stanley et al. (1998) compare a routine consent with one that includes a comprehension questionnaire. The evidence indicates that one in four patients have a poor understanding of the risks and complications of the procedures.

[8]Glennerster (2017) discusses the practical aspects of obtaining informed consent, yet this is based on legal and ethical requirements and does not address empirical implications. Alderman (2013) address practical concerns of how to collect consent in difficult situations, yet only by collecting and describing anecdotal evidence. These discussions are critical as they discuss necessary conditions, but it remains unclear whether they are sufficient. In medicine, see on this topic Fitzgerald et al. 2002; Hutton et al. 2008; Miller and Boulton 2007; Stunkel et al. 2010 and a meta-study by Falagas et al. (2009). Overall, many practices in economic research are still behind medical sciences in terms of credibility, replicability, and transparency. See a comparative analysis of the two disciplines by Favereau (2016) and Avdeenko and Frölich (2020). An important distinction to medical science is that economists do often not design the programs they evaluate, i.e., there is a clear difference between the policy content and the research protocols applied, which is demonstrated in "randomizing religion" Bryan et al. (2021). The researchers thoroughly discuss and address the ethical considerations of their research yet are not responsible for the policy that would have been implemented anyway.

Yet, medical science differs from social science in important ways which directly relate to the provision of consent. For medical procedures, a lack of consent might mean that a terminal condition remains untreated. In a situation with no real choice of alternatives, patients might willfully ignore the information provided by the doctor as it might only induce stress and worry without changing their choice. Further, despite being concerned with the same question of "how much information is enough", research in this field is mostly conducted in high-income countries, and studies are often underpowered. Our study also relates to the literature on the importance of privacy and data protection (Acquisti et al. 2016), with its increasing relevance for digital economics (Goldfarb and Tucker 2019). In their overview article, Acquisti et al. (2016) elaborate on the observation that consumers would rarely (if ever) completely be aware of privacy threats and the consequences of sharing and protecting their personal information. Hence, market interactions involving personal data would often occur in the absence of individuals' fully informed consent. Moreover, research on consumer behavior has found that opt-in consent could benefit established firms whom consumers seem to trust more (Campbell et al. 2015). In several laboratory experiments, economists have studied whether data sharing is a rational decision, starting with whether and how much people would value their privacy.[9] For instance, Benndorf and Normann (2018) study the willingness to sell data and find that only a minority of data holders are unwilling to sell their data for commercial purposes. Benndorf et al. (2015) study the voluntariness of private information disclosure and find that voluntary disclosure of private information might result in an unraveling of privacy. Finally, Marreiros et al. (2017) show that study participants would disclose less private information when exposed to information regarding privacy.

This is the first experimental study to investigate the understanding of privacy regulations at a larger scale. We are the first to test whether and how informed consent can be improved with several thousand data holders in a real-world setting. More specifically, we work with a highly vulnerable population in a context of high levels of illiteracy and poverty, describe a lack of an understanding of data regulations, and test modern means of communication to improve informedness. We find that our study participants display a poor understanding of their rights and risks and that this could be improved through paying more attention and time to this topic. We also show that this improvement does not come at a cost for external validity.

If individual consent to research is placed above the greater public good, which could be generated from innovative evidence, access to services may suffer. Again, debates on

[9] See, e.g., Acquisti et al. (2007); Benndorf et al. (2015); Benndorf and Normann (2018); Beresford et al. (2012); Fast and Schnurr (2020); Feri et al. (2016); Marreiros et al. (2017); Schudy and Utikal (2017).

the potential institutionalization of inequalities and evidence from medical research are informative.[10] In Canada, for instance, obtaining written informed consent for participation in a stroke registry led to important selection biases such that registry patients were not representative of the typical stroke patient (Tu et al. 2004). Despite great mobilization efforts to register individuals for a disease register, no general, valid scientific conclusions could be made. In economics, this study relates to the work on selection into surveys, specifically into lab or field experiments, and the credibility of such evidence. Schulz et al. (2019), for instance, find that mentioning financial incentives boosts the participation rate in lab experiments by 50 percent and that more selfish individuals and individuals with higher cognitive reflection scores are more likely to participate in experiments. Our findings further question the motivation for participating in research, which can have clear implications for the external validity of the results generated. Furthermore and more generally, the study is related to research that uses experimental methods to identify biases in survey responses and develop methods to reduce them (Blattman et al. 2016; Karlan and Zinman 2012).

We believe that further investigations on improving the understanding of and decisions about the costs and benefits associated with data sharing in this specific context are far-sighted. The evidence from our work supports the usage of structures ethical consent approaches, which are interactive and carefully adjusted to the population studied. Our results have implications for the potential way ethics and quality of data collections could be improved. First, using additional evidence from a pilot that preceded the study, we show that even well-educated and experienced enumerators are little aware that the respondents are free not to participate in the survey. Thus, the fact that study participants are not aware of their rights starts earlier in the process and could be tackled already by giving enumerators a more in-depth training on data protection, the purpose of use, and processing. Far beyond the research conducted, this could help shape an occupational curriculum that eventually would formalize the work of (often free-lance) enumerators. Second, study participants need to be better informed about the purposes, consequences, and the voluntary nature of their data provision. To explain their voluntary participation, a more interactive and illustrative approach could yield higher levels of understanding and informedness.

The rest of the chapter is organized as follows: In section 5.2, we describe the background and experimental design. Section 5.3 describes our empirical strategy, data, limitations, and findings. Finally, section 5.4 discusses the results further and describes planned or ongoing additions to the study.

[10]See, e.g., Cassell and Young (2002) and Tu et al. (2004).

5.2.1. Background

To design our intervention, we first conducted a review of relevant guidelines and examples. We are interested in current practices of how informed consent is obtained during surveys. We focus primarily on social sciences and, in particular, on the work conducted in LMICs where illiterate and uneducated respondents are often targeted.

Systematically reviewing guidelines and legal codes, we first tried to distill principles of what informed consent should contain in our context. While many guidelines are tailored to medical and experimental research, some deal solely with data protection. Considering the vast amount of different codes across disciplines and countries as well as the great overlap of content, we focused on only three: An international code of conduct in social science research published by United Nations Educational, Scientific and Cultural Organization (UNESCO) (de Guchteneire 2014), the United States (US) code of federal regulations on the protection of human subjects (45 Code of Federal Regulations (CFR) 46), and the European Union (EU) General Data Protection Regulation (GDPR).

Given that our study focuses on survey research, we narrowed the principles of what the respondent should be informed about down to the following: (1) identity and contact information of the research teams; (2) purpose of data collection and research; (3) expected duration of participation, (4) risks, benefits, or consequences of participation; (5) voluntary nature of participation and right to withdraw consent; (6) (limitations of) confidentiality of records. We further divided the latter into (a) the recipients of data (including the risk of transfer into other legal systems), (b) the duration of storage, (c) the right to complain, and (d) the procedures to ensure confidentiality. This narrows in on data protection issues and omits experimental protocols, which we deemed appropriate in our context of data collection.

In addition to these theoretical guidelines, we wanted to conduct a reality check. Unfortunately, we are not aware of any database that consistently collects consent forms. Thus we searched the Datahub for Field Experiments in Economics and Public Policy of Harvard Dataverse for consent forms used by researchers in economics.[11] On September 30th 2019, there was a total of 147 entries. We found 59 entries that provided access to questionnaires, 39 of which included a consent form[12] 34 of which were available in

[11] Harvard Dataverse is an open-source research data repository. The Datahub for Field Experiments in Economics and Public Policy is accessible via the following link: `https://dataverse.harvard.edu/dataverse/DFEEP`. Accessed on November 20th, 2019.

[12] Some of the remainders indicated the existence of such a consent form without giving direct access to it.

English.[13] Of course, this might not be representative of all field research conducted by economists but is a selection of rather popular studies. Further, we collected consent forms from major surveys which are conducted in several countries such as the Demographic and Health Surveys (DHS) or Multiple Indicator Cluster Surveys (MICS). We designed the treatment arms based on this information, starting with the business-as-usual consent form, which represents our reference group.

5.2.2. Pilot of survey instrument

With our survey instrument, we aimed to measure objective and subjective understanding of the study purpose, voluntarism, data confidentiality, and rights w.r.t. data protection.[14] We tested our survey instrument during an enumerator training for a different data collection in the same geographical area. During the training, potential enumerators had to answer questions related to the content of the training.[15] While the survey instrument was the same, the questions referred to a slightly different information text which was used in the survey the enumerators were hired for. Results from the pilot are presented in Table 5.1. For our objective measures of understanding, this might be considered as a benchmark of what is possible. In our context, enumerators are usually more educated and experienced with surveys compared to respondents. Further, they answered these questions as part of a test, so arguably, they were motivated to get it right. For our subjective measures of understanding, on the other hand, the potential enumerators could have been concerned about admitting that they did not understand certain aspects well.

From Table 5.1 we can see that the enumerators had problems with the questions related to voluntarism. This is especially concerning since enumerators need to ensure or facilitate these aspects during a survey. Only 6% of enumerators correctly assessed both that the respondents (i) are free not to participate in the survey and (ii) can decline to answer specific questions (20% and 13% correctly assessed the respective statement). Only every third enumerator correctly understands the purpose of the data collection, despite having participated in extensive training before. Our pilot is the first indicator that even amongst enumerators, who are usually more educated and experienced with surveys than respondents, a general problem with understanding the purpose, rights, and obligations during a data collection exists. When asking about their subjective

[13] Note that there are duplicates, as several entries can relate to the same data collection.

[14] For more details about the questions, please refer to appendix E.B.

[15] We included our survey instrument in the test and, after the test, asked the potential enumerators for consent to use their test data in a research study. A total of 115 potential enumerators gave their consent. Whether or not the potential enumerators were hired depended, among other things, on their performance on the test.

Table 5.1: Understanding in enumerator pilot study

	Overall (1)	Rights (2)	Purpose (3)	Voluntary (4)	Confidentiality (5)
Objective understanding					
Share of informed	2%	36% & 72%	29.6%	13% & 20%	99%
Subjective understanding					
Share understanding	73%	87%	94%	85%	93%
Objective and subjective					
High and high	1%	30%	29%	7%	92%
Low and high	72%	59%	65%	78%	1%
High and low	0%	4%	1%	0%	7%
Low and low	27%	7%	5%	15%	0%

Notes. The table displays a summary of responses to our survey instrument of 115 potential enumerators participating in a training for a different survey but in the same region. Columns (2)-(5) each refers to one of four aspects of informed consent we inquired about, and column (1) to a summary measure across aspects. The objective understanding of the aspects *rights with respect to data protection* and *voluntary nature of participation* (columns (2) and (4)) are measured based on two items the shares of which are given in the respective columns. For details on the questions refer to appendix E.B.

understanding, the enumerators reported that they understood the different aspects well. There is little to no relation between how well the enumerators felt they understood the aspects and whether they answered related questions correctly.

5.2.3. Experimental design

In the following, we describe our intervention and treatment arms. The different approaches are integrated into the Computer-Assisted Personal Interviewing (CAPI) survey tool. At the beginning of each interview, the consent form is randomized to either the business-as-usual, Audio-visual Supported Consent Form (ACF), or Audio-visual Supported and Scripted Interactive Consent Form (ASCF) approach described below. Since the study is an independent add-on to a different study, we want to minimize any potential influence on the main survey. Therefore, 50% of potential respondents are assigned to the control group, 44% to the ACF treatment, and only 6% of potential respondents are assigned to the full ASCF treatment.[16] Moreover, we undertook several efforts to contextualize this research design in the local context. The exact phrasing of the various approaches was decided after discussions with local experts and with feedback from several Non-Governmental Organisations (NGOs). All content was translated into the

[16]The survey is collected using SurveyCTO, and the treatment is randomized using the build-in random number generator. One could imagine that an enumerator with a preference over treatments will create new forms to avoid specific treatment arms. However, before the randomization, the enumerator already needs to fill in some information. The effort of creating a new form should keep her from doing so if not necessary. Comparing the actual treatment assignment probabilities to the theoretical ones we only find small differences.

respective local languages.

Control Group: Business-as-usual Consent Form. Currently, the common practice of obtaining informed consent for survey participation is asking the potential respondent to read about one page of information, the so-called consent form. This shall ensure that all aspects of informed consent are addressed. Depending on the survey and who is conducting it, the procedure varies in length. Given that the respondents of surveys in LMICs are often not literate, it is common practice that the interviewers read the consent form to the respondents. Consent is then usually obtained either written, by a signature or a similar practice, or oral. As the survey was collected during the COVID-19 pandemic, consent was obtained through recorded oral statements. We want to test alternative approaches against this benchmark in a real-world survey setting in this research project. The information text presented to the potential respondent and further details for all treatment arms see appendix E.A.

Treatment Group 1: Audio-visual Supported Consent Form (ACF). The first experimental variation is an ACF. In our context, we expect most respondents to be illiterate or to have low reading abilities.[17] Instead of asking the enumerator to read out the consent form, they will first be asked to present a short video to the potential respondent. In addition to reading the text of the consent form, the video tries to visually illustrate processes concerning confidentiality and data protection and the interview itself. This part of the treatment has several potential advantages over the benchmark. The process is normed such that each respondent receives the same information. Further, being visually engaged might increase attention and can make abstract concepts more available to the respondents.

Treatment Group 2: Audio-visual Supported and Scripted Interactive Consent Form (ASCF). The second experimental variation is an ASCF. It combines the ACF with a scripted and interactive reading of the consent form. Even though the respondent is encouraged to ask questions during the benchmark approach, they rarely take advantage of this. Both the benchmark and ACF approach tend to be passive. To actively engage the potential respondent in the process, we include questions about the content of the consent form between paragraphs to check whether the respondent retained the information. If the responses indicate misunderstanding, the relevant information is repeated. We expect this to encourage the respondents to listen actively and keep more information presented in the consent form.

[17] A large-scale data collection was conducted in this area in 2016. Based on evidence from this survey, we expect about 70% of the household heads to be without any formal education.

In the following we describe our estimation strategy, sample and data, limitations and omissions, and present our analysis.

5.3.1. Empirical strategy

Equation 5.1 outlines our main specification capturing the Intention to treat effects (ITTs) of our treatments.

$$Y = \alpha + \beta D_{Video} + \gamma D_{ASCF} + \xi + u \tag{5.1}$$

Y is the outcome variable at the respondent level for all the respondents who provided us with the consent or at the level of potential respondents for consent rates. The ITT for either the ASCF or ACF treatment, i.e. showing at least the video, is captured by the β. γ captures the differential of adding the scripted interactive approach of the ASCF treatment arm. Additionally, ξ corresponds to enumerator fixed effects. The standard errors are Eicker-Huber-White standard errors.

5.3.2. Sample and data

The experiment is implemented alongside an already planned data collection in Sindh, Pakistan. Thus potential effects are established in a real rather than a laboratory setting. To measure the potential respondents' understanding of consent, we developed a survey tool we piloted with enumerators (see section 5.2.2) consisting of a total of ten questions (for more details, refer to appendix E.B). Since the study is an independent add-on to a different study, we want to minimize any potential influence on the main survey. Therefore, we chose to pose all questions only to 12% of respondents, the 6% in the ASCF group and 6% from the control group. The remaining 88%, assigned to either the ACF or control group, are asked only two questions at random (one of each objective and subjective measure of understanding). This implies different sample sizes for different outcomes and comparisons. Table 5.2 gives an overview of the (approximate[18]) sample size for various types of outcomes and comparisons.

Table 5.2 further displays summary statistics of the respondent's age, sex, and relation to the household head across different treatment groups. About 60% of respondents are female, and the average age is 43 years. About 40% of respondents are the household head (almost always male), about 50% are their spouse, and about 10% are in another relationship to the household head. We can see that there are differences in respondent characteristics between the ASCF and control group. Those assigned to the ASCF treatment are about 9% less likely to be female and 10% more likely to be the head of the

[18]Due to the random assignment some questions are asked more frequently than others

Table 5.2: Characteristics of respondents and sample sizes

	Control		Any video		ASCF		Differences	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	mean	sd	mean	sd	mean	sd	V-C	ASCF-C
Female	.6	0.49	.59	0.49	.5	0.50	-0.010	-0.095**
Age	43	11.61	43	11.15	44	11.51	-0.130	0.677
Household head	.39	0.49	.41	0.49	.5	0.50	0.022	0.105**
Spouse	.49	0.50	.47	0.50	.4	0.49	-0.017	-0.082*
Other relation	.12	0.33	.12	0.32	.098	0.30	-0.005	-0.023
Number of observations:								
Meta indicators	2001		1963		215		3964	2216
Objective item level	515		502		215		1017	730
Subjective item level	685		656		215		1341	900
Full module	214		215		215		429	429

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

Notes. The table displays characteristics of study participants as well as approximate sample sizes for various outcomes. Columns (1) and (2) correspond to the control group, columns (3) and (4) to any treatment group, i.e. either the ACF or the ASCF group, and columns (5) and (6) to the ASCF group. Column (7) displays differences between the control and any video group. Column (8) the difference between the control and the ASCF group. Meta indicators refer to outcomes such as consent rates which are mostly observed for everyone. Objective and subjective item level refers to one of the items of our objective and subjective measure of understanding. Each item of the respective category was asked with the same probability such that this gives an approximation for all items. Note that there are only four subjective, compared to six objective items, explaining the discrepancy between the numbers of observations. Full module refers to those who were asked all questions related to our measures of understanding. Note that all in the ASCF, 12% of the control group, and no one in the ACF group was assigned to all questions. For details on the questions, refer to appendix E.B.

household.[19] We do not observe any difference between those assigned to any treatment and those assigned to the control group.

5.3.3. Limitations

The survey was conducted during the COVID-19 pandemic, and thus the data collection was subject to social distancing rules which could have made it inconvenient to properly show the video to respondents. While the various approaches of obtaining consent were trained, and their importance was stressed during the enumerator training, their proper implementation could not be systematically monitored. Further, enumerators are, in principle, able to manipulate the randomization of treatment assignment. For example, if they want to avoid the ASCF treatment, they can check the treatment assignment and delete the corresponding form and start a new one. However, this is cumbersome as all the information about location needs to be re-entered, and it is doubtful whether this would

[19]Note that neither sex nor relationship to the household head are good predictors for any of the outcomes we discuss in the following. Including these as covariates in our specifications does not affect the findings presented in this chapter. Further, we investigated heterogeneous effects with respect to sex and do not find differences.

save the enumerators any time. We do not find any strong indication of manipulation of this process. We observe that only 5.4% instead of 6% of potential respondents were assigned to the ASCF treatment arm. While this difference is statistically significant at the 10% level, it is negligible in absolute terms and small in relative terms.

Finally, since we wanted to measure understanding in a real-world setting at scale, we had to compromise on depth. Most respondents are asked only one question to assess their objective understanding, clearly not enough to give a complete picture. Further, we measure understanding based on the assessment of whether a particular statement is true or not. We found a tendency of respondents to assess a statement as true, regardless of whether it is true or not. This implies that true statements are often by default evaluated to be true, such there is little room for improvement according to our measure. To remedy this limitation, we extended our survey module by including a statement with the opposite truth value compared to the original statement (see E.B). Our extended module will be used in the second half of this survey which takes place in another part of the country and will allow us to investigate this issue further.

5.3.4. Rate of consent

A common concern for researchers is the representativeness of their studies and the number of study participants. Therefore, we test the effect of the treatment approaches on the rate of consent compared to the business-as-usual approach. Are people more or less likely to decline an interview if the consent form is presented differently? Generally, there are two channels to consider. Firstly, the unusual approach might scare off some people or even increase their interest, trust, and willingness to participate. Secondly, the intervention accomplishes its goal to inform the respondents better, and better-informed respondents make different decisions. They might become aware of the consequences and no longer want to provide data because they thought the consequences were less severe - or they might want to provide data in the first place because they thought the consequences were more severe.

The data collection took place over two visits, an initial visit, during which the experiment was implemented, and a second visit, during which a more comprehensive interview took place. The consent is asked in both visits and the respondent in the second visit was not necessarily the same as in the first visit. We collect information on which respondents gave consent in both types of visits and record the number of potential respondents who refused to be interviewed in both types of visits. Table 5.3 presents the results of our analysis. Almost everyone in our study provided consent to be interviewed, and there is no

Table 5.3: Rate of consent

	Consent rates		Response rate
	(1)	(2)	(3)
	1st visit	2nd visit	2nd visit
	b/se	b/se	b/se
At least ACF	0.00046	-0.00067	0.0059
	(0.00046)	(0.0011)	(0.0058)
ASCF	0.000046	0.0016	-0.0033
	(0.00011)	(0.0011)	(0.013)
Model description:			
Adj. R^2	0.01	0.01	0.05
Control group mean	1.00	1.00	0.96
Observations	3964	3626	3963

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. The table displays different rates of consent. Column (1) is the rate of consent asked during the first visit during which the interview took place, everyone gave consent. Column (2) refers to a rate of consent asked during the second visit. Note that this was only properly documented after the data collection already began, such that 334 observations are missing. Finally column (3) refers to the response rate, i.e. an indicator for consent during the second visit conditional on consent during the first.

difference across treatment arms.[20] Thus, a first finding is that the different approaches do not affect the rate of consent in any meaningful way.

Note that the measured consent rate unlikely reflects how many approached people gave their consent to be interviewed. There are two main explanations for this: (i) the timing of measurement and (ii) incomplete documentation.

The timing of measurement is the correct one for this study. We are neither interested in the usual response rate for this study, which also documents unavailable respondents, nor in the overall consent rate. We are only interested in changes due to our intervention and, thus, the consent rate conditional on being part of the experiment. Being part of the experiment means that the formal process of acquiring consent is being conducted; this only happens after an initial buy-in from the potential respondent. Before starting the process, the enumerator already introduced herself and usually informed the potential respondent about the purpose of her visit. The potential respondent agreed (or at least did not effectively object) to start the formal process of acquiring consent. Given this initial buy-in, it is less surprising that almost everyone gave consent. However, it still allows

[20] Only 3 out of the 3,626 potential respondents in the second visit did not provide consent. People who did not provide consent are not directly included in our study, but we document the number of potential respondents without consent by treatment assignment.

us to measure the interventions' effect on the consent rate and potential implications for sample selection.

Incomplete documentation, on the other hand, would be problematic. There was little incentive for the enumerators to document when they did not receive consent. While there was little cost to documenting it, some enumerators might not have submitted forms without consent as they might have feared that this would reflect poorly on them. If the treatment affects the consent rate and forms without consent are not collected, this could be reflected in the shares of the treatment groups and the composition of respondents. We only found negligible differences in treatment group shares and discuss the difference in respondents across treatment arms in 5.3.2.

Given that the two treatment interventions do not affect the rate of consent, we do not need to consider whether the different approaches change our sample composition.[21] Further, since there are no selection effects to consider in the analysis of the remainder of outcomes, this facilitates the interpretation of results.

5.3.5. Understanding of informed consent

Objective understanding. First, we want to assess whether or not survey respondents are informed. We hypothesize that the business-as-usual process of obtaining consent is not sufficiently informative. To assess this hypothesis, we constructed a short questionnaire module consisting of six statements (for details, refer to appendix E.B). Each statement is related to one or more principles of informed consent discussed in section 5.2.1. The respondent is asked to assess whether the statements are *true* or *false*; *don't know* is also offered as an answer option. Table 5.4 presents the results for each statement separately, as well as a summary score and an indicator of being sufficiently informed as measures across statements.

We can see that that on average, the respondents assessed more than half (58%) of the statements correctly, and 18% of respondents assessed almost all statements correctly.[22] About 90% assessed each of the statements related to their rights with respect to data protection and confidentiality correctly, 25% and 33% for the two statements related to voluntarism, and 18% for the statement related to the survey's purpose.[23] We note a pattern here, that statements which are *true* are considerably more likely to be assessed correctly (about 90%) compared to statements which are *false* which less than a third of

[21]Note, however, that we do find differences in sample composition as discussed in section 5.3.2.

[22]Note that the outcomes based on multiple items are only measured for about 12% of the sample, half of which from the control and the other half from the ASCF group.

[23]As a comparison, in our pilot study (see section 5.2.2), similarly almost everyone correctly assessed the statement to confidentiality. But a much smaller share of only 36% and 72% assessed the statements with respect to the data protection rights correctly. Further, about 11% more got the study purpose correct, and about 12% less got the voluntary nature of participation correct.

the control group assesses correctly.

We do not find any effects of showing the video alone for any of the statements. But we find that the ASCF intervention successfully improved the respondents' understanding with respect to voluntarism. The estimated increase is 8.6 percentage points for the statement *I have to participate in the study* and 12 percentage points for *When I give consent, I have to respond to all the questions*. These effect sizes correspond to a significant increase of about one-third relative to the control group means. The tendency to choose *true* as a default, i.e. regardless of whether the statement is *true* or *false*, might explain why we only find significant treatment effects for statements which are *false*.

Table 5.4: Objective measures of understanding

	Overall		Rights		Purpose	Voluntary		Confidentiality
	(1) Score b/se	(2) Informed b/se	(3) Item 1 b/se	(4) Item 2 b/se	(5) b/se	(6) Item 1 b/se	(7) Item 2 b/se	(8) b/se
ASCF	0.0093	0.033	-0.016	-0.022	0.018	0.082*	0.12**	-0.010
	(0.015)	(0.034)	(0.023)	(0.026)	(0.030)	(0.032)	(0.036)	(0.024)
At least ACF			0.0071	-0.015	-0.029	-0.020	-0.00051	0.022
			(0.018)	(0.019)	(0.022)	(0.025)	(0.026)	(0.018)
Model description:								
Adj. R^2	0.55	0.48	0.33	0.32	0.40	0.44	0.44	0.36
Control group mean	0.58	0.18	0.90	0.91	0.18	0.26	0.33	0.87
Observations	429	429	1017	1001	1016	1007	1028	1039

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. The table displays outcomes based on our objective measure of understanding of consent (see appendix E.B). Column (1) is a summary score based on the average number of correctly answered items. Column (2) refers to an indicator for having answered at most 1 of the 6 items incorrectly or at most 2 with *don't know*. These outcomes can only be measured for those receiving the ASCF treatment and part of the control group, the ACF coefficient is thus omitted. Columns (3)-(8) refer to the single items ordered by category and are indicators for a correct response (*don't know* and *refuse to answer* are coded as 0).

Subjective understanding. Next, we analyze our subjective measure of understanding. It would be of concern if respondents feel they have understood, but they do not. Or simply if they do not think they understand what they consented to. Using our subjective measure of understanding, we assess whether the different approaches affect the share of respondents who think they are informed. Again, ex-ante, it is unclear in which direction the effect goes.

From Table 5.5 we can see that about 70% of respondents report that they understood each of the four aspects of consent well, and 57% report that they understood all of the aspects well. Given the low scores on our objective measures, this can be interpreted as overconfidence in their understanding of these aspects. However, since the objective measure is not very comprehensive and based on only one or two items, this subjective measure might capture actual understanding beyond our objective measure in addition to a subjective feeling of understanding.

We further see that the ASCF treatment decreased subjective understanding overall. On average, the ASCF approach decreases the share of well-understood aspects by 6 percentage points (control group mean is 69%). This decrease is spread out across all items, for each of which the reduction is too small to be detected with statistical precision. We cannot disentangle whether this effect is due to the video in general or the ASCF specifically, as we only measure one of the aspects for those respondents who were assigned to the ACF treatment.

Given that the understanding was considerably higher based on the subjective as opposed to the objective measure, which was low for most aspects, this decrease in subjective understanding can be considered an improvement if it leads to an alignment of objective and subjective understanding discussed in the following.

Table 5.5: Subjective measures of understanding

	Overall			Rights	Purpose	Voluntary	Confidentiality
	(1) Share	(2) All	(3) Score	(4)	(5)	(6)	(7)
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
ASCF	-0.060*	-0.088**	-0.14**	-0.0051	-0.037	-0.032	-0.034
	(0.025)	(0.033)	(0.053)	(0.029)	(0.032)	(0.029)	(0.031)
At least ACF				-0.0075	-0.030	-0.0015	0.018
				(0.022)	(0.021)	(0.020)	(0.020)
Model description:							
Adj. R^2	0.63	0.58	0.60	0.45	0.44	0.51	0.50
Control group mean	0.69	0.57	3.67	0.71	0.71	0.70	0.69
Observations	426	426	426	1293	1341	1343	1270

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. The table displays outcomes based on our subjective measure of understanding of consent (see section E.B in appendix E). Column (1) is the share of the four categories (columns (4)-(7)). Column (2) refers to an indicator that all categories are reportedly understood. Column (3) refers to a score based on the average understanding across all categories from *1=not at all* to *5=fully*. Columns (4)-(7) refer to a response of *I undertstood this well* or *I undertstood this fully* for the respective aspect.

Objective vs. Subjective Understanding. Using our objective and subjective measure of understanding, we assess whether the different approaches change the alignment of these measures across respondents. We divide the respondents into four types for each of the four aspects: Respondents who have both a high objective and subjective understanding, respondents who have a low understanding in both, and respondents for which the measures are not aligned. We want to analyze whether these types are differently represented across the different approaches.

From Figure 5.1 we can see a similar picture. For the aspects of *rights with respect to data protection* and *confidentiality*, respondents with both high objective and subjective understanding are the majority, and those with a high objective and low subjective understanding are the second largest group, reflecting the fact that most respondents answered both related items correctly. For the aspects of *purpose of the study* and the *voluntarism*, respondents have a low understanding according to our objective measure and a high understanding according to our subjective measure. This reflects the potential overconfidence pointed out earlier.[24] The second-largest type of respondents is those with both low subjective and objective understanding. The relation between objective and subjective understanding is, however, mainly an artifact of the objective measures. It turns out that the evaluation according to the objective and subjective measures are barely related, i.e., those with high objective understanding are not more likely to have a high subjective understanding and vice versa.

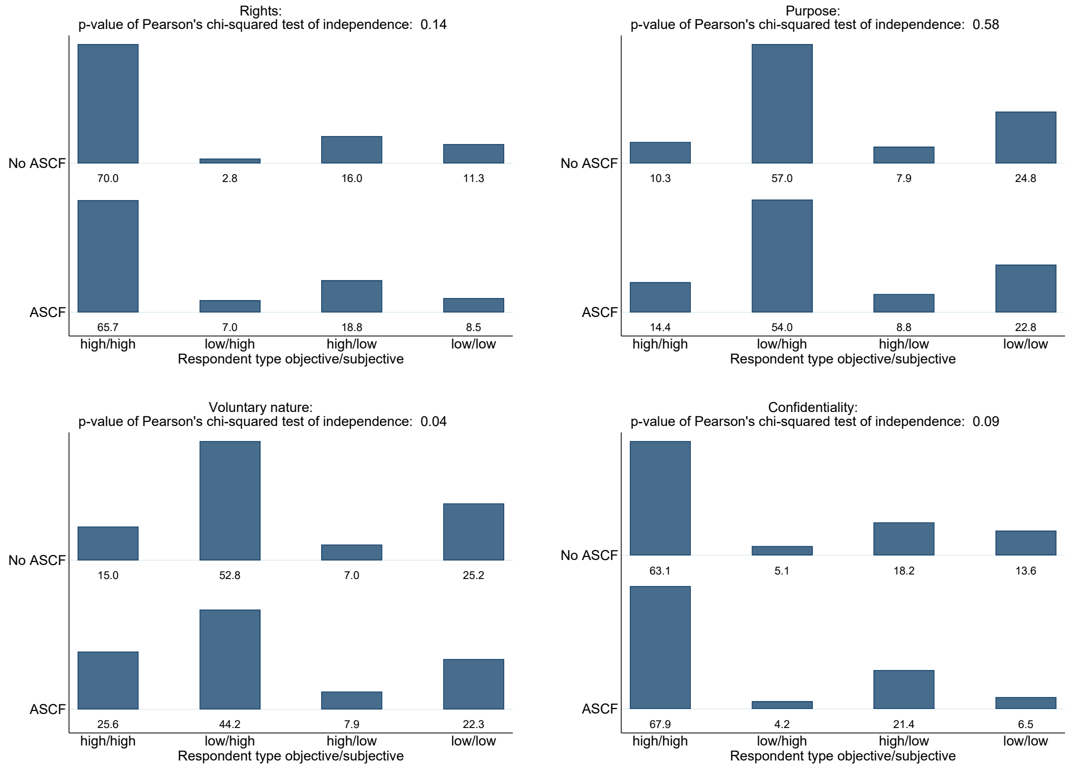
For each aspect, we conduct a Pearson's chi-square test of independence between assignment to ASCF and the distribution across respondent types.[25] For the aspects *rights with respect to data protection*, the *purpose of the study*, and *confidentiality*, we do not reject the null of no difference, but only for the aspect relating to the *voluntarism*.

The change for the *voluntarism* is expected and reflects what we already discussed earlier. The ASCF treatment increased the share of those with a high objective understanding of this aspect. Most of them are those that would otherwise have a low objective but high subjective understanding, thus increasing the alignment between objective and subjective understanding.

[24]Note that we do not find any difference in overconfidence between men and women.

[25]We only analyze this for the ASCF treatment, as for the ACF treatment, only two questions were asked at random, reducing the sample size for outcomes based on three questions to zero and for outcomes based on two questions to less than 5%. Note, however, since we only found effects for the ASCF treatment for the individual measures, we want to look only at this measure in any case.

Figure 5.1: Objective vs. subjective understanding



5.3.6. Item non-response rates

In addition to declining to participate in the interview, respondents can refuse to answer any specific question the interviewer poses. We, therefore, want to assess the effect of the different approaches on the item non-response rate and especially the non-response rate to questions that might be sensitive. Again, the effects could arguably go both ways. The approaches could increase or decrease trust, make respondents aware of the voluntary nature of their participation or how their data is handled.

Since the ASCF treatment increased the share of respondents that are aware that they can refuse to answer specific questions significantly, we might expect more of them to make use of this right. However, we can see from Table 5.6 that this does not seem to be the case.[26] In the first visit, there were hardly any questions any respondent refused to answer; in fact, only 4% of respondents declined to reply to any of the questions. In the second visit, there are two distinctions to make. Respondents who participated in the first visit always answered the household roster questions; however, only some answered the entire questionnaire. For the roster and full interview, the share of those refusing to answer any of the questions was higher at 7% and 11%, respectively. But we find neither

[26]This lack of finding is robust to an alternative definition that includes *don't know* as non-response

Table 5.6: Item non-response rates

	1st visit		2nd visit roster		2nd visit full interview	
	(1)	(2)	(3)	(4)	(5)	(6)
	rate	any	rate	any	rate	any
	b/se	b/se	b/se	b/se	b/se	b/se
At least ACF	-0.066	-0.0034	-0.00077	-0.0027	-0.029	0.00021
	(0.11)	(0.0057)	(0.020)	(0.0080)	(0.040)	(0.018)
ASCF	0.14	0.0071	-0.062^{+}	-0.022	-0.080	-0.017
	(0.23)	(0.012)	(0.034)	(0.015)	(0.072)	(0.039)
Model description:						
Adj. R^2	0.15	0.15	0.06	0.07	0.13	0.12
Control group mean	0.72	0.04	0.15	0.07	0.23	0.11
Observations	3963	3963	3832	3832	1162	1162

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. The table displays results for non-response behavior. Columns (1) and (2) correspond to the first visit, columns (3) and (4) to the roster during the second, and columns (5) and (6) to the full interview during the second visit. Columns (1), (3), and (5) corresponds to the non-response rate in percent among sensitive questions (i.e. 0.1 means the respondent refused to answer 0.1 percent of sensitive questions) and columns (2), (4), and (6) to an indicator of any non-response to a sensitive question. A sensitive question is defined, as per pre-analysis plan, to be any question at least one respondent refused to answer. Note that, to ensure robustness towards outliers, the non-response rates are winsorized to 3 standard deviations from the mean.

that the video alone nor the ASCF treatment affects the share of respondents who refused to answer any question, nor the frequency with which they gave refusals. Note that there is a correlation between our measure of understanding and response behavior. More than twice as many that correctly assessed that they do not have to respond to all questions, refuse to reply to at least one question during the first visit. However, it is not the case that those that incorrectly assessed that they have to reply to all questions do so, about 3% refuse to answer at least one question during the first visit.[27]

5.4. CONCLUSION

The 2019 Nobel Prize acknowledged the work that "laid a solid stepping stone for a new generation of researchers in development economics and other fields" (The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel 2019). These developments and the prize motivate and spur further research in LMICs. Acting like "plumbers", economists continue to take on responsibilities related to social engineering, carefully assisting government to design effective policies to social challenges (Duflo 2017). With

[27]This is expected, as respondents can learn this during the interview when they are asked a question they are reluctant to answer. To refuse to answer, they do not need to be aware of this option at all times actively.

increased attention on the whole discipline, greater attention will be paid to the many ethical aspects of applying experimental methods with vulnerable populations.

Indeed, field experiments improve the credibility of economic research, and we observe an ever-increasing number of experiments in developing countries. These studies are often based on survey data. This article focuses on a specific, little advanced and discussed, ethical aspect: The consent to participate in surveys in LMICs. Research on the ethical aspects of collecting survey data remains scarce and is non-existent in development economics. There is, however, a responsibility to engage with the ethical requirements of survey data collection. We focused here on informational constraints to the consent of potential survey participants. We tested whether survey participants are sufficiently informed about what happens to their data and their rights and found significant gaps in knowledge. Further, we experimentally tested whether an interactive, audio-visual supported approach could improve how well informed they are. In this approach, the content of the consent form was presented in a structured dialog and illustrated with a short video. We showed that augmenting the consent process can improve respondent's understanding without affecting response behavior. The improvements we detect are limited to the aspect of *voluntarism*, and we detect them only for the full ASCF treatment and not the video by itself. Finally, we also investigated implications on response behavior and data quality. We find no evidence of changes in response behavior. On the extensive margin, we find that the consent rate is not affected by either of our approaches. On the intensive margin, we do not find an effect on item non-response rates.

Our study is the first set in the context of survey data collection in LMICs where privacy laws are underdeveloped, and illiteracy further limits informed consent to the collection, use, and disclosure of personal data. Our study informs an emerging debate on the ethical and practical challenges related to conducting field experiments. As Asiedu et al. (2021) argue, an improvement of the norms in the discipline. The authors argue that it would be important for projects to integrate mechanisms to deal with such ethical concerns throughout the project. Our study points to a good time to start this process - the first encounter with the study participants and making sure that they are aware that their participation is voluntary.

Moreover, we hope this study can become a starting point with much potential for future research with high policy relevance. In particular, it is a common belief in development economics that larger field experiments could help to improve external validity or the accuracy with which the estimates of impact from a Randomized Control Trial (RCT) predict the effects of some subsequent policy decision (Duflo et al. 2007; Muralidharan and Niehaus 2017; Peters et al. 2016). Consequently, it is argued that tests of external validity and representativeness of the study sample should be as standard and taken as seriously as tests of internal balance between treatment and control group. Given that development research aims to alleviate poverty and improve people's lives, we need to be careful that this research does not systematically exclude vulnerable parts of the population. Therefore it is crucial to know whether survey respondents differ in an important way from non-respondents. If this difference is partly due to the process through which consent is obtained, this puts the requirement for individual informed consent at odds with the goals of development research. And researchers would need to be aware of this. The high level of initial consent in our context had limited the insights we could gain on this topic, but we believe it is worth further exploration.

Conclusion

Throughout my studies, I was involved in several impact evaluations in various stages. One common theme is problems with implementation. This is not only the reason why some studies did not make it into my dissertation but is also reflected in the studies discussed in chapters 1 to 3.

The linkage intervention discussed in chapter 2 was plagued with difficulties, as the financial institutions were not ready and did not sufficiently design or offer products viable for savings groups. For example, it is impossible to open a group account at any of the mobile money providers. Similarly, for the social fund intervention discussed in chapter 3, reports suggest that the training was improperly implemented, and some trainers were not convinced or understood the concepts they were supposed to convey. To make matters worse, incomplete information on the implementation made it challenging to monitor the extent to which the intervention was implemented, which groups were reached, and the adherence to the experimental protocol. There is considerable room for improvement through sound monitoring systems, and evaluations should be flexible to shift from measuring impact to diagnosing obstacles to implementation. In many projects, there is no need to wait for rigorous evidence of no effect before investigating why a project did not work. This focus on implementation could further help to understand the discrepancy with respect to effects across implementers which we see in chapters 2 and 3 and is commonly observed in many empirical studies.

As it is with research and evaluations, monitoring relies on data, and thus good data quality is crucial. I believe the topic of data quality needs to gain more traction in economics. Particularly development economists increasingly conduct primary data collections, but survey methodology remains essentially the domain of other social sciences. Apart from data quality, processes of data collection can be a topic of investigation. With chapters 4 and 5, I tried to generate empirical evidence related to the ethics of data collection and hope that further research can shine more light on these processes. With econometrics as a sub-discipline, the statistical processing of data is an integral part of economics. There is no reason why it should be different for the gathering of data.

References

Acquisti, A., S. Gritzalis, C. Lambrinoudakis, and S. di Vimercati (2007): *What can behavioral economics teach us about privacy?*, Auerbach Publications.

Acquisti, A., C. Taylor, and L. Wagman (2016): "The economics of privacy," *Journal of economic Literature*, 54, 442–92.

Alderman, Jishnu Rao, V. H. D. (2013): *Conducting Ethical Economic Research Complications from the Field*, Policy Research Working Papers, The World Bank.

Alinaghi, N. (2019): "Mobile money, risk sharing, and transaction costs: a replication study of evidence from Kenya's mobile money revolution," *Journal of Development Effectiveness*, 11, 342–359.

Alkenbrack, S. and M. Lindelow (2015): "The Impact of Community-Based Health Insurance on Utilization and Out-of-Pocket Expenditures in Lao People's Democratic Republic," *Health Economics*, 24, 379–399.

Asiedu, E., D. Karlan, M. Lambon-Quayefio, and C. Udry (2021): "A Call for Structured Ethics Appendices in Social Science Papers," Tech. rep., National Bureau of Economic Research.

Avdeenko, A. and M. Frölich (2020): "Research standards in empirical development economics: What's well begun, is half done," *World Development*, 127, 104786.

Batista, C. and P. C. Vicente (2018): "Is Mobile Money Changing Rural Africa? Evidence from a Field Experiment," Working Paper, NOVAFRICA Working Paper Series.

Beaman, L., D. Karlan, and B. Thuysbaert (2014): "Saving for a (not so) Rainy Day: A Randomized Evaluation of Savings Groups in Mali," Working Paper 20600, National Bureau of Economic Research, series: Working Paper Series.

Benndorf, V., D. Kübler, and H.-T. Normann (2015): "Privacy concerns, voluntary disclosure of information, and unraveling: An experiment," *European Economic Review*, 75, 43–59.

Benndorf, V. and H.-T. Normann (2018): "The willingness to sell personal data," *The Scandinavian Journal of Economics*, 120, 1260–1278.

Beresford, A. R., D. Kübler, and S. Preibusch (2012): "Unwillingness to pay for privacy: A field experiment," *Economics letters*, 117, 25–27.

Besley, T., S. Coate, and G. Loury (1994): "Rotating Savings and Credit Associations, Credit Markets and Efficiency," *The Review of Economic Studies*, 61, 701–719.

Blattman, C., J. Jamison, T. Koroknay-Palicz, K. Rodrigues, and M. Sheridan (2016): "Measuring the measurement error: A method to qualitatively validate survey data," *Journal of Development Economics*, 120, 99–112.

Bouman, F. J. A. (1995): "Rotating and accumulating savings and credit associations: A development perspective," *World Development*, 23, 371–384.

Brown, J. S., T. L. Schonfeld, and B. G. Gordon (2006): ""You May Have Already Won...": An Examination of the Use of Lottery Payments in Research," *IRB: Ethics & Human Research*, 28, 12–16.

Brune, L., X. Giné, J. Goldberg, and D. Yang (2016): "Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi," *Economic Development and Cultural Change*, 64, 187–220.

Brunie, A., L. Fumagalli, T. Martin, S. Field, and D. Rutherford (2014): "Can village savings and loan groups be a potential tool in the malnutrition fight? Mixed method findings from Mozambique," *Children and Youth Services Review*, 47, 113–120.

Bryan, G., J. J. Choi, and D. Karlan (2021): "Randomizing religion: the impact of Protestant evangelism on economic outcomes," *The Quarterly Journal of Economics*, 136, 293–380.

Buehren, N. (2011): "Allocating Cash Savings and the Role of Information: Evidence from a Field Experiment in Uganda," Tech. rep., Working Paper.

Burlando, A. and A. Canidio (2017): "Does group inclusion hurt financial inclusion? Evidence from ultra-poor members of Ugandan savings groups," *Journal of Development Economics*, 128, 24–48.

Cai, H., Y. Chen, H. Fang, and L.-A. Zhou (2015): "The Effect of Microinsurance on Economic Activities: Evidence from a Randomized Field Experiment," *The Review of Economics and Statistics*, 97, 287–300.

Callegaro, M., R. P. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick, and P. J. Lavrakas (2014): *Online Panel Research: A Data Quality Perspective*, John Wiley & Sons.

Campbell, J., A. Goldfarb, and C. Tucker (2015): "Privacy regulation and market structure," *Journal of Economics & Management Strategy*, 24, 47–73.

Cassell, J. and A. Young (2002): "Why we should not seek individual informed consent for participation in health services research," *Journal of Medical Ethics*, 28, 313–317.

Chernozhukov, V., C. Hansen, and M. Spindler (2015): "Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments," *American Economic Review*, 105, 486–90.

Christensen, G. and E. Miguel (2018): "Transparency, reproducibility, and the credibility of economics research," *Journal of Economic Literature*, 56, 920–80.

Cohen, M. and J. Sebstad (2005): "Reducing vulnerability: the demand for microinsurance," *Journal of International Development*, 17, 397–474.

Cole, S., X. Giné, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery (2013): "Barriers to Household Risk Management: Evidence from India," *American Economic Journal: Applied Economics*, 5, 104–135.

Cole, S., T. Sampson, and B. Zia (2011): "Prices or Knowledge? What Drives Demand for Financial Services in Emerging Markets?" *The Journal of Finance*, 66, 1933–1967.

de Guchteneire, P. (2014): "Code of Conduct and Ethical Guidelines," Tech. rep., UNESCO.

De Mel, S., C. McIntosh, K. Sheth, and C. Woodruff (2018): "Can Mobile-Linked Bank Accounts Bolster Savings? Evidence from a Randomized Controlled Trial in Sri Lanka," Working Paper 25354, National Bureau of Economic Research, series: Working Paper Series.

De Weerdt, J. and S. Dercon (2006): "Risk-sharing networks and insurance against illness," *Journal of Development Economics*, 81, 337–356.

Dekker, M. and A. Wilms (2010): "Health Insurance and Other Risk-Coping Strategies in Uganda: The Case of Microcare Insurance Ltd." *World Development*, 38, 369–378.

Demirgüç-Kunt, A. and L. Klapper (2012): "Measuring Financial Inclusion: The Global Findex Database," Working Paper 6025, World Bank Policy Research Working Paper.

Dercon, S. (2002): "Income Risk, Coping Strategies, and Safety Nets," *The World Bank Research Observer*, 17, 141–166.

Dercon, S., J. De Weerdt, T. Bold, and A. Pankhurst (2006): "Group-based funeral insurance in Ethiopia and Tanzania," *World Development*, 34, 685–703.

Dercon, S., R. V. Hill, D. Clarke, I. Outes-Leon, and A. Seyoum Taffesse (2014): "Offering rainfall insurance to informal insurance groups: Evidence from a field experiment in Ethiopia," *Journal of Development Economics*, 106, 132–143.

Duflo, E. (2017): "Richard T. Ely Lecture: The Economist as Plumber," *American Economic Review*, 107, 1–26.

Duflo, E., R. Glennerster, and M. Kremer (2007): "Using randomization in development economics research: A toolkit," *Handbook of development economics*, 4, 3895–3962.

Dupas, P., D. Karlan, J. Robinson, and D. Ubfal (2018): "Banking the Unbanked? Evidence from Three Countries," *American Economic Journal: Applied Economics*, 10, 257–297.

Dupas, P., A. Keats, and J. Robinson (2017): "The Effect of Savings Accounts on Interpersonal Financial Relationships: Evidence from a Field Experiment in Rural Kenya," *The Economic Journal*, n/a.

Dupas, P. and J. Robinson (2013a): "Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya," *American Economic Journal: Applied Economics*, 5, 163–192.

——— (2013b): "Why Don't the Poor Save More? Evidence from Health Savings Experiments," *American Economic Review*, 103, 1138–1171.

Fafchamps, M. and F. Gubert (2007): "The formation of risk sharing networks," *Journal of Development Economics*, 83, 326–350.

Fafchamps, M. and S. Lund (2003): "Risk-sharing networks in rural Philippines," *Journal of Development Economics*, 71, 261–287.

Falagas, M. E., I. P. Korbila, K. P. Giannopoulou, B. K. Kondilis, and G. Peppas (2009): "Informed consent: how much and what do patients understand?" *The American Journal of Surgery*, 198, 420–435.

Fast, V. and D. Schnurr (2020): "The Value of Personal Data: An Experimental Analysis of Data Types and Personal Antecedents," *Working Paper*.

Favereau, J. (2016): "On the analogy between field experiments in economics and clinical trials in medicine," *Journal of Economic Methodology*, 23, 203–222.

Feri, F., C. Giannetti, and N. Jentzsch (2016): "Disclosure of personal information under risk of privacy shocks," *Journal of Economic Behavior & Organization*, 123, 138–148.

Fitzgerald, D. W., C. Marotte, R. I. Verdier, W. D. Johnson, and J. W. Pape (2002): "Comprehension during informed consent in a less-developed country," *The Lancet*, 360, 1301–1302.

Flory, J. (2016): "Banking the Poor: Evidence from a Savings Field Experiment in Malawi," Tech. rep., Working Paper.

Frölich, M. and L. Nguyen (2020): "Impacts of linking savings group to formal financial service providers and strengthening their internal group insurance mechanism in Zambia," Impact Evaluation Report 121, 3ie.

FSD (2015): "FinScope Zambia 2015," Survey report, FSD Zambia and Bank of Zambia.

Glennerster, R. (2017): "The practicalities of running randomized evaluations: partnerships, measurement, ethics, and transparency," in *Handbook of Economic Field Experiments*, Elsevier, vol. 1, 175–243.

Goldfarb, A. and C. Tucker (2019): "Digital economics," *Journal of Economic Literature*, 57, 3–43.

Gueron, J. M. (2017): "The politics and practice of social experiments: Seeds of a revolution," in *Handbook of Economic Field Experiments*, Elsevier, vol. 1, 27–69.

Heltberg, R., A. M. Oviedo, and F. Talukdar (2015): "What do Household Surveys Really Tell Us about Risk, Shocks, and Risk Management in the Developing World?" *The Journal of Development Studies*, 51, 209–225.

Hilbert, M. (2016): "Big data for development: A review of promises and challenges," *Development Policy Review*, 34, 135–174.

Hoogeveen, J., K. Croke, A. Dabalen, G. Demombynes, and M. Giugale (2014): "Collecting high frequency panel data in Africa using mobile phone interviews," *Canadian Journal of Development Studies / Revue canadienne d'études du développement*, 35, 186–207.

Hutton, J. L., M. P. Eccles, and J. M. Grimshaw (2008): "Ethical issues in implementation research: a discussion of the problems in achieving informed consent," *Implementation Science*, 3, 52.

Jack, W. and T. Suri (2014): "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution," *American Economic Review*, 104, 183–223.

Jamison, J. C., D. Karlan, and J. Zinman (2014): "Financial Education and Access to Savings Accounts: Complements or Substitutes? Evidence from Ugandan Youth Clubs," Working Paper 20135, National Bureau of Economic Research.

Jütting, J. P. (2004): "Do Community-based Health Insurance Schemes Improve Poor People's Access to Health Care? Evidence From Rural Senegal," *World Development*, 32, 273–288.

Kaplan, L., J. Kuhnt, and J. I. Steinert (2020): "Do no harm? Field research in the Global South: Ethical challenges faced by research staff," *World Development*, 127, 104810.

Karlan, D., A. L. Ratan, and J. Zinman (2014): "Savings by and for the Poor: A Research Review and Agenda," *Review of Income and Wealth*, 60, 36–78.

Karlan, D., B. Savonitto, B. Thuysbaert, and C. Udry (2017): "Impact of savings groups on the lives of the poor," *Proceedings of the National Academy of Sciences*, 114, 3079–3084.

Karlan, D. and J. Zinman (2018): "Price and control elasticities of demand for savings," *Journal of Development Economics*, 130, 145–159.

Karlan, D. S. and J. Zinman (2012): "List randomization for sensitive behavior: An application for measuring use of loan proceeds," *Journal of Development Economics*, 98, 71–75.

Kast, F. and D. Pomeranz (2014): "Saving More to Borrow Less: Experimental Evidence from Access to Formal Savings Accounts in Chile," Working Paper 20239, National Bureau of Economic Research.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75, 83–119.

Kropf, M. E. and J. Blair (2005): "Eliciting Survey Cooperation: Incentives, Self-Interest, and Norms of Cooperation," *Evaluation Review*, 29, 559–575.

Ksoll, C., H. B. Lilleør, J. H. Lønborg, and O. D. Rasmussen (2016): "Impact of Village Savings and Loan Associations: Evidence from a cluster randomized trial," *Journal of Development Economics*, 120, 70–85.

LCMS (2015): "2015 Living Conditions Monitoring Survey Report," Survey report, Central Statistical Office Zambia.

Lee, Y. S., L. Johnson, D. Ansong, I. Osei-Akoto, R. Masa, G. Chowa, and M. Sherraden (2017): "'Taking the Bank to the Youth': Impacts on Savings from the Ghana YouthSave Experiment," *Journal of International Development*, 29, 936–947.

Leo, B. and R. Morello (2016): "Practical Considerations with Using Mobile Phone Survey Incentives: Experiences in Ghana and Tanzania," SSRN Scholarly Paper ID 2841010.

Leo, B., R. Morello, J. Mellon, T. Peixoto, and S. T. Davenport (2015): "Do Mobile Phone Surveys Work in Poor Countries?" SSRN Scholarly Paper ID 2623097.

Ludwig, J., S. Mullainathan, and J. Spiess (2019): "Augmenting pre-analysis plans with machine learning," in *AEA Papers and Proceedings*, vol. 109, 71–76.

Marreiros, H., M. Tonin, M. Vlassopoulos, and M. Schraefel (2017): ""Now that you mention it": A survey experiment on information, inattention and online privacy," *Journal of Economic Behavior & Organization*, 140, 1–17.

Mercer, A., A. Caporaso, D. Cantor, and R. Townsend (2015): "How Much Gets You How Much? Monetary Incentives and Response Rates in Household Surveys," *Public Opinion Quarterly*, 79, 105–129.

Miller, T. and M. Boulton (2007): "Changing constructions of informed consent: Qualitative research and complex social worlds," *Social Science & Medicine*, 65, 2199–2211.

Muralidharan, K. and P. Niehaus (2017): "Experimentation at scale," *Journal of Economic Perspectives*, 31, 103–24.

Muyanga, M., T. S. Jayne, and W. J. Burke (2013): "Pathways into and out of Poverty: A Study of Rural Household Wealth Dynamics in Kenya," *The Journal of Development Studies*, 49, 1358–1374.

Olken, B. A. (2015): "Promises and perils of pre-analysis plans," *Journal of Economic Perspectives*, 29, 61–80.

Parmar, D., S. Reinhold, A. Souares, G. Savadogo, and R. Sauerborn (2012): "Does Community-Based Health Insurance Protect Household Assets? Evidence from Rural Africa," *Health Services Research*, 47, 819–839, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-6773.2011.01321.x.

Peters, J., J. Langbein, and G. Roberts (2016): "Policy evaluation, randomized controlled trials, and external validity—A systematic review," *Economics Letters*, 147.

Prina, S. (2015): "Banking the poor via savings accounts: Evidence from a field experiment," *Journal of Development Economics*, 115, 16–31.

Ravallion, M. (2020): "Should the randomistas (continue to) rule?" Working paper, National Bureau of Economic Research.

Rayzberg, M. S. (2019): "Fairness in the Field: The Ethics of Resource Allocation in Randomized Controlled Field Experiments," *Science, Technology, & Human Values*, 44, 371–398.

Scheil-Adlung, X., G. Carrin, J. Juetting, and K. Xu (2006): "What is the Impact of Social Health Protection on Access to Health Care, Health Expenditure and Impoverishment? A Comparative Analysis of Three African Countries," SSRN Scholarly Paper ID 916703.

Schudy, S. and V. Utikal (2017): "'You must not know about me'—On the willingness to share personal data," *Journal of Economic Behavior & Organization*, 141, 1–13.

Schulz, J., U. Sunde, P. Thiemann, and C. Thöni (2019): "Selection into experiments: Evidence from a population of students," .

Singer, E. and M. P. Couper (2008): "Do Incentives Exert Undue Influence on Survey Participation? Experimental Evidence," *Journal of Empirical Research on Human Research Ethics*, 3, 49–56.

Singer, E., R. M. Groves, and A. D. Corning (1999): "Differential Incentives: Beliefs About Practices, Perceptions of Equity, and Effects on Survey Participation," *The Public Opinion Quarterly*, 63, 251–260.

Singer, E. and C. Ye (2013): "The Use and Effects of Incentives in Surveys," *The ANNALS of the American Academy of Political and Social Science*, 645, 112–141.

Solove, D. J. (2012): "Introduction: Privacy self-management and the consent dilemma," *Harv. L. Rev.*, 126, 1880.

Stanley, B. M., D. J. Walters, and G. J. Maddern (1998): "Informed Consent: How Much Information Is Enough?" *Australian and New Zealand Journal of Surgery*, 68, 788–791.

Strobl, R. (2017): "Does Health Insurance Reduce Child Labour and Education Gaps? Evidence from Rwanda," *The Journal of Development Studies*, 53, 1376–1395.

Stunkel, L., M. Benson, L. McLellan, N. Sinaii, G. Bedarida, E. Emanuel, and C. Grady (2010): "Comprehension and Informed Consent: Assessing the Effect of a Short Consent Form," *IRB*, 32, 1–9.

The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel (2019): "Understanding Development and Poverty Alleviation," .

Townsend, R. M. (1994): "Risk and Insurance in Village India," *Econometrica*, 62, 539–591.

Tu, J. V., D. J. Willison, F. L. Silver, J. Fang, J. A. Richards, A. Laupacis, and M. K. Kapral (2004): "Impracticability of informed consent in the Registry of the Canadian Stroke Network," *New England Journal of Medicine*, 350, 1414–1421.

Vashistha, A., E. Cutrell, and W. Thies (2015): "Increasing the Reach of Snowball Sampling: The Impact of Fixed versus Lottery Incentives," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Association for Computing Machinery, CSCW '15, 1359–1363.

Wieser, C., M. Bruhn, J. Kinzinger, C. Ruckteschler, and S. Heitmann (2019): *The Impact of Mobile Money on Poor Rural Households: Experimental Evidence from Uganda*, Policy Research Working Papers, The World Bank.

Zangeneh, M., R. Barmaki, H. Gibson-Wood, M.-J. Levitan, R. Romeo, and J. Bottoms (2008): "Research Compensation and Lottery: An Online Empirical Pilot Study," 5.

Zutlevics, T. (2016): "Could providing financial incentives to research participants be ultimately self-defeating?" *Research Ethics*, 12, 137–148.

A. APPENDIX TO CHAPTER 1

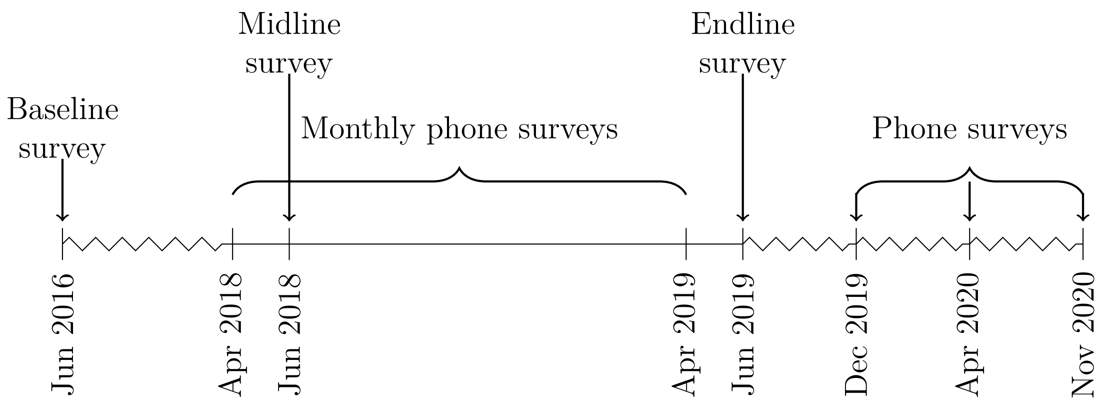


Figure A.1: Timeline of data collections.

Table A.1: Variables used in randomization included in PDS LASSO

	N	Mean	S.D.	Minimum	Maximum
Mean number of children aged 14 and younger per Household	522	2.56	0.79	0	5
Mean number of occupied rooms per Household	522	3.38	0.83	1	8
Mean number of Savings Group members in the Village	522	22.08	6.27	8	47
Mean contribution to Savings Group of members in the Village	522	49.92	60.48	2	600
Mean contribution to social fund	522	2.99	4.85	1	100
Mean Household Size in Village	522	5.90	1.06	3	10
Fraction of interviewees in the village that took a loan from the savings group	522	0.60	0.30	0	1
Total number of savings group members that took a loan from their savings group	522	38.13	46.95	0	198
Mean score of additive food security index	522	1.24	0.73	0	4
Mean number of months with food scarcity across HHs	522	1.47	0.77	0	4
Mean livestock value per household in the village	522	3163.10	4875.66	0	63469
Mean school attendance rate for children born after 1999 and before 2011	522	0.76	0.20	0	1
Mean value of agricultrual output sales per household	522	2871.94	10943.29	0	180443
Additive score of the questions about trust in government and private banks	522	4.28	0.89	2	7
fraction of female headed households	522	0.24	0.20	0	1
# participating SGs in village	522	3.27	4.14	1	19
Mean area of land connected to the household	522	8667.91	35185.54	0	803369
Mean value of loans per person	522	912.69	1787.05	0	17205
Mean value of loans per person for business purposes	522	584.33	1412.82	0	14583
Mean value of loans per person for agricultural purposes	522	32.00	89.78	0	1250
Mean value of loans per person for food purposes	522	40.26	86.31	0	983
Mean value of loans per person for educational purposes	522	100.68	239.87	0	1975
Number of Inhabitants in the Village	522	1002.54	3576.44	0	64683
Electricity in village	522	0.33	0.47	0	1
Urban or rural: Urban	522	0.18	0.38	0	1
Urban or rural: Rural > 250	522	0.56	0.50	0	1
Urban or rural: Rural <250	522	0.26	0.44	0	1

The table displays summary statistics on the savings group level of the variables used in the re-randomization and included in the PDS LASSO procedure. All information was collected during the baseline and includes information from the village questionnaire or from the household and member questionnaires which is collapsed on the randomization unit (usually village) level.

Table A.2: Variables based on baseline survey included in PDS LASSO

	N	Mean	S.D.	Minimum	Maximum
Meeting frequency (mode across surveys): weekly	522	0.52	0.50	0	1
Meeting frequency (mode across surveys): every two weeks	522	0.04	0.19	0	1
Meeting frequency (mode across surveys): monthly	522	0.44	0.50	0	1
Meeting frequency (highest across surveys): weekly	522	0.61	0.49	0	1
Meeting frequency (highest across surveys): every two weeks	522	0.08	0.28	0	1
Meeting frequency (highest across surveys): monthly	522	0.31	0.46	0	1
Meeting frequency (lowest across surveys): weekly	522	0.36	0.48	0	1
Meeting frequency (lowest across surveys): every two weeks	522	0.05	0.22	0	1
Meeting frequency (lowest across surveys): monthly	522	0.59	0.49	0	1
Year of founding: 2012 and earlier	522	0.34	0.47	0	1
Year of founding: 2013	522	0.11	0.31	0	1
Year of founding: 2014	522	0.30	0.46	0	1
Year of founding: 2015	522	0.22	0.41	0	1
Year of founding: 2016	522	0.04	0.20	0	1
Predominantely female group	522	0.33	0.47	0	1
Predominantely female lead group	522	0.29	0.45	0	1
Trained by NGO upon founding	522	0.78	0.42	0	1
Group was trained by NGO upon founding (favor 0)	522	0.57	0.50	0	1
Group uses box to store savings	522	0.74	0.44	0	1
Group loan outs savings as storage	522	0.68	0.47	0	1
Group uses bank account to store savings	522	0.05	0.21	0	1
Mode of storage reported by members (favor loan out): in a box held by one of th	521	0.43	0.50	0	1
Mode of storage reported by members (favor loan out): with one of the members wi	521	0.02	0.15	0	1
Mode of storage reported by members (favor loan out): in a bank	521	0.01	0.11	0	1
Mode of storage reported by members (favor loan out): all savings loaned out	521	0.53	0.50	0	1
Mode of storage reported by members (favor loan out): 8	521	0.01	0.09	0	1
Mode of storage reported by members (favor box): in a box held by one of the gro	521	0.54	0.50	0	1
Mode of storage reported by members (favor box): with one of the members without	521	0.02	0.16	0	1
Mode of storage reported by members (favor box): in a bank	521	0.01	0.09	0	1
Mode of storage reported by members (favor box): all savings loaned out	521	0.42	0.49	0	1
Mode of storage reported by members (favor box): 8	521	0.00	0.04	0	1
Meeting frequency: weekly	522	0.52	0.50	0	1
Meeting frequency: every two weeks	522	0.03	0.18	0	1
Meeting frequency: monthly	522	0.45	0.50	0	1
Average monthly savings contribution in ZMW	521	107.45	196.78	5	3533
Average monthly savings contribution	521	101.04	121.34	8	887
Average value of current savings in ZMW (reported)	521	672.41	1002.08	0	13800
Average value of current savings	521	642.77	767.12	13	4100
Average monthly contribution to SF in ZMW	522	5.92	7.78	0	80
Average monthly contribution to SF in ZMW (winsorized)	522	5.72	6.39	0	35

The table displays summary statistics of variables on the savings group level included in the PDS LASSO procedure. Information was collected during the baseline from the savings group member questionnaire or recall information from the savings group questionnaire in later surveys and is collapsed on the savings group level.

Table A.3: Variables from mid- and endline surveys included in PDS LASSO

	Midline July 2018					Endline July 2019				
	N	Mean	S.D.	Min	Max	N	Mean	S.D.	Min	Max
Savings group variables										
Savings group is dormant	521	0.07	0.26	0	1	522	0.08	0.28	0	1
Number of savings group members	521	19.80	7.20	4	60	522	20.05	7.09	5	68
Social fund accumulates	521	0.12	0.32	0	1	522	0.13	0.34	0	1
Member variables										
Respondent is female	1945	0.80	0.40	0	1	2604	0.80	0.40	0	1
Ethnicity of respondent: Other	1945	0.19	0.39	0	1	2604	0.18	0.39	0	1
Ethnicity of respondent: Bemba	1945	0.31	0.46	0	1	2604	0.31	0.46	0	1
Ethnicity of respondent: Chewa	1945	0.13	0.34	0	1	2604	0.14	0.35	0	1
Ethnicity of respondent: Lozi	1945	0.10	0.30	0	1	2604	0.09	0.29	0	1
Ethnicity of respondent: Mambwe	1945	0.08	0.28	0	1	2604	0.08	0.27	0	1
Ethnicity of respondent: Nsenga	1945	0.18	0.39	0	1	2604	0.19	0.39	0	1
Denomination of respondent: Other	1945	0.35	0.48	0	1	2604	0.54	0.50	0	1
Denomination of respondent: Catholic	1945	0.35	0.48	0	1	2604	0.25	0.44	0	1
Denomination of respondent: UCZ	1945	0.12	0.33	0	1	2604	0.09	0.28	0	1
Denomination of respondent: Pentecostal	1945	0.10	0.30	0	1	2604	0.06	0.25	0	1
Denomination of respondent: New apostolic church	1945	0.08	0.27	0	1	2604	0.06	0.24	0	1
Age of respondent	1936	45.47	12.55	17	85	2578	45.60	12.40	17	91
Relation to household head: Other relation	1945	0.04	0.19	0	1	2604	0.03	0.17	0	1
Relation to household head: Household head	1945	0.42	0.49	0	1	2604	0.43	0.50	0	1
Relation to household head: Spouse	1945	0.54	0.50	0	1	2604	0.54	0.50	0	1
Respondent is married	1945	0.73	0.44	0	1	2604	0.73	0.44	0	1
Household variables										
Number of household members	2070	5.90	2.28	1	15	2501	5.65	2.20	1	15
Number of household members aged 18 or above	2060	3.49	1.60	1	12	2471	3.08	1.43	1	11
Number of household members aged 5 or below	2070	0.53	0.69	0	5	2501	0.58	0.75	0	4
Average age of household members	2063	25.66	9.47	3	85	2487	25.74	10.67	2	90

The table displays summary statistics of variables on the savings group, member, and household level included in the PDS LASSO procedure for respective estimations. Information was collected during the mid- and endline from the respective questionnaires.

B. APPENDIX TO CHAPTER 2

B.A. Characterization of linked groups

To better understand which savings groups are affected by the treatment, we look at how these savings groups differ in terms of characteristics in Table B.1. To compare the characteristics, we assume that the treatment effect is monotonic for each savings group, i.e., $Y_{1i} \geq Y_{0i} \ \forall \ i$ where Y_{1i} refers to the potential outcome (e.g., using a bank account) if group i is assigned to treatment and Y_{0i} refers to the potential outcome if group i is not assigned to treatment. Given random treatment assignment and this assumption, we have:

$$\begin{aligned} & \frac{P(X_i = 1 | Y_{1i} > Y_{0i})}{P(X_i = 1)} \\ = & \frac{P(Y_{1i} > Y_{0i} | X_i = 1)}{P(Y_{1i} > Y_{0i})} \\ = & \frac{E[Y_i | D_i = 1, X_i = 1] - E[Y_i | D_i = 0, X_i = 1]}{E[Y_i | D_i = 1] - E[Y_i | D_i = 0]} \end{aligned}$$

Table B.1 shows the ratio mentioned above of the prevalence of a covariate among those savings groups affected by the treatment compared to the prevalence in the overall sample for NGOs 1 and 2. Affected by the treatment refers to either reportedly using a bank account in any survey wave (Table B.1 columns (2) and (5)) or reportedly actively using a bank account in any of the midline, endline, or phone survey waves after the endline (Table B.1 columns (3) and (6)).[28]

Before discussing Table B.1, please note that while the presented estimates are consistent for their estimand, no measures of uncertainty are given, and we need to be careful with any inference as the estimates are based on rather small sample sizes. For NGO 1 is seems that affected groups tend to be rather urban (less likely to be located in a rural village with less than 250 inhabitants and more likely to be in villages or towns with electricity). In contrast, for NGO 2, the opposite is true: affected groups are more likely located in small villages and less likely in villages with electricity. The affected groups are also more likely to meet monthly (compared to weekly) for NGO 1.[29] There is again

[28]Given that we do not find a treatment effect for NGOs 3 and 4, we omit them from the table. Further, since many covariates are strongly related to NGO affiliation, we want to compare within NGOs instead of pooling them together not to pick up differences between NGOs with larger and lower treatment effects.

[29]Since almost all groups affiliated with NGO 2 meet monthly, there is no difference for this characteristic for NGO 2.

a contrast between NGO 1 and 2 for the year of the founding of the affected groups. For NGO 1, it tends to be older groups, whereas, for NGO 2, affected groups tend to be younger. For both NGOs, the affected groups tend to have a female leader during the years of study. There seems to be no notable difference between reportedly using and actively using an account.

Table B.1: Characteristics of groups affected by the treatment

	NGO1			NGO2		
	(1)	(2)	(3)	(4)	(5)	(6)
	mean covariate	ratio account	ratio active use	mean covariate	ratio account	ratio active use
Urban or rural: Rural >250	0.520	1.02	1.00	.6	0.48	0.58
Urban or rural: Rural <250	0.327	0.69	0.77	.31	1.61	1.81
Electricity in village	0.267	1.50	1.25	.43	0.56	0.41
Meeting frequency: weekly	0.880	0.92	0.93	.013		
Meeting frequency: monthly	0.100	1.56	1.71	.99	0.97	0.94
Year of founding: 2012 and earlier	0.100	1.78	1.56	.83	0.71	0.95
Year of founding: 2013	0.040			.12	1.32	-1.13
Year of founding: 2014	0.533	0.68	0.85	.04		
Year of founding: 2015	0.320	1.18	0.90	0		
Trained by NGO upon founding	0.727	1.09	1.07	.61	1.00	0.87
Predominantely female group	0.220	1.53	1.90	.24	0.83	1.17
Predominantely female lead group	0.173	1.50	1.66	.19	1.06	2.25
Group uses box to store savings	1.000	1.00	1.00	.95	0.83	0.78
Group loan outs savings as storage	0.100	1.25	1.37	.64	0.84	0.91

Notes. Each line refers to a different binary characteristic (covariate) of savings groups at baseline. Predominantly female (lead) group is an indicator based on information of all survey waves about group composition and leadership with respect to sex, a large part of predominantly is exclusively female. Columns (1) to (3) refer to NGO 1 and columns (4) to (6) refer to NGO 2. NGOs 3 and 4 are omitted due to lack of treatment effect. Columns (1) and (4) show the mean of the covariate for the respective NGO. Columns (2), (3), (5), and (6) show estimates for $\frac{E[Y|D=1,X=1]-E[Y|D=0,X=1]}{E[Y|D=1]-E[Y|D=0]}$, where X refers to the covariate, D to the treatment assignment, and Y to ever using a bank account (columns (2) and (5)) or to actively using an account in any of midline, endline, or phone survey waves after the endline (columns (3) and (6)). Under the assumption of monotone treatment effects on Y, this ratio equals the ratio of the prevalence of the covariate among groups that were affected by the treatment and the prevalence of the covariate among all groups. If there are less than 3 groups for a covariate and treatment assignment combination, the respective ratio for that covariate is omitted.

B.B. Supplementary tables

Table B.2: Mobile money use, trust and knowledge

	Midline July 2018		Endline July 2019	
	(1)	(2)	(3)	(4)
	All NGOs	NGO 4	All NGOs	NGO 4
	b/se	b/se	b/se	b/se
Used mobile money before				
Linkage treatment	-0.052*	-0.074+	-0.038+	-0.0016
	(0.024)	(0.040)	(0.023)	(0.037)
Control group mean	0.50	0.50	0.56	0.57
Used mobile money in last 3 months				
Linkage treatment	-0.034	-0.0077	-0.034	-0.0025
	(0.023)	(0.036)	(0.023)	(0.040)
Control group mean	0.29	0.24	0.35	0.35
Uses mobile money at least once a month				
Linkage treatment	-0.010	0.014	-0.018	0.018
	(0.018)	(0.025)	(0.018)	(0.030)
Control group mean	0.14	0.091	0.23	0.21
Level of trust in mobile money agencies				
Linkage treatment	-0.031	-0.018	0.054	0.12
	(0.065)	(0.12)	(0.054)	(0.081)
Control group mean	1.89	2.14	2.03	2.01
Complete trust in mobile money agencies				
Linkage treatment	0.0017	0.025	-0.028	-0.054
	(0.031)	(0.047)	(0.022)	(0.033)
Control group mean	0.43	0.35	0.36	0.44
Level of knowledge of of mobile money				
Linkage treatment	-0.031	-0.026	-0.018	0.041
	(0.073)	(0.11)	(0.065)	(0.087)
Control group mean	2.10	2.06	2.48	2.68
Full knowledge of of mobile money				
Linkage treatment	0.043*	0.042	-0.014	-0.031
	(0.020)	(0.034)	(0.019)	(0.029)
Control group mean	0.20	0.18	0.33	0.35
Observations	1945	711	2604	941

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). For estimations on the whole sample, NGO fixed effects are always included, but excluded for NGO specific estimations. Observations reflects maximum number not accounting for missing values. Missing values: negligible missing values (less than 5 for whole sample) for all outcomes except trust related outcomes. Trust related outcomes have 410 to 560 missing values in the whole sample. Level of trust is on a 4 point scale, level of knowledge refers to 0 to 4 correctly answered questions. Complete trust and full knowledge refer to highest value for respective level. NGO 4 is included separately as their intervention was tailored to mobile money, we similarly find no effects for NGOs 1, 2 and 3.

Table B.3: Trust in safety of savings and financial institutions

	Midline July 2018		Endline July 2019	
	(1) All NGOs b/se	(2) NGOs 1 and 2 b/se	(3) All NGOs b/se	(4) NGOs 1 and 2 b/se
Completely trust that savings are save				
Linkage treatment	-0.00076	-0.0028	-0.017	0.013
	(0.019)	(0.028)	(0.023)	(0.035)
Control group mean	0.82	0.84	0.63	0.53
Level of trust that savings are save				
Linkage treatment	0.018	0.012	0.042	-0.0034
	(0.030)	(0.045)	(0.029)	(0.043)
Control group mean	1.22	1.20	1.42	1.50
Trust in financial institutions (index)				
Linkage treatment	-0.065	-0.039	0.0018	0.0014
	(0.044)	(0.066)	(0.035)	(0.051)
Control group mean	-0.024	-0.33	0.0061	0.10
Observations	1945	839	2604	1120

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. Missing values: 91 and 17 missing for level of trust at mid- and endline. Level of trust is on a 4 point scale, complete trust refers to highest level (=1). The index for trust in financial institutions is the means of variables related to the index standardized to control group values, thus the control group mean is by construction close to 0 for the whole sample.

Table B.4: Savings activity by NGO

			(1) NGO 1 b/se	(2) NGO 2 b/se	(3) NGO 3 b/se	(4) NGO 4 b/se
Monthly savings contribution to group						
	Midline	Linkage treatment	3.70	-23.2*	19.8+	8.62
			(11.39)	(10.69)	(11.62)	(13.09)
		Control group mean	109.8	96.5	59.3	127.1
	Endline	Linkage treatment	-14.0	-7.81	5.53	4.61
			(13.45)	(12.14)	(10.37)	(11.11)
		Control group mean	141.2	99.9	77.5	140.5
Value received at last share-out						
	Midline	Linkage treatment	-43.8	10.3	108.0	-90.2
			(121.13)	(125.74)	(121.70)	(148.49)
		Control group mean	1126.1	1803.4	1145.1	1698.8
	Endline	Linkage treatment	-22.8	245.9*	-171.3	196.8
			(133.15)	(123.05)	(378.78)	(232.21)
		Control group mean	1362.3	1659.3	2221.6	1776.2
Monthly savings contributions to other groups						
	Midline	Linkage treatment	0.13	-0.44	2.05	-0.61
			(2.43)	(4.53)	(5.01)	(4.36)
		Control group mean	5.02	4.50	8.17	10.6
	Endline	Linkage treatment	-8.34*	-1.46	3.52	-0.0080
			(3.80)	(2.78)	(3.38)	(3.55)
		Control group mean	18.2	2.76	5.95	8.34
Value of savings outside of savings groups						
	Midline	Linkage treatment	-35.1	0.72	63.2	-90.3
			(58.81)	(81.14)	(89.09)	(56.90)
		Control group mean	182.0	320.2	190.0	305.5
	Endline	Linkage treatment	32.1	-69.3	52.4	-3.98
			(37.07)	(76.01)	(67.21)	(77.17)
		Control group mean	131.7	348.1	220.7	254.6
Observations Midline			568	271	387	711
Observations Endline			747	373	533	941

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). Observations refer to total number of observations available not accounting for missing values, thus actual number of observations used varies with outcome. All outcomes in ZMW and winsorized at 1% and 99%.

Table B.5: Details on components of the HH animals purchased index

	Midline July 2018		Endline July 2019	
	(1) All NGOs b/se	(2) NGOs 1 and 2 b/se	(3) All NGOs b/se	(4) NGOs 1 and 2 b/se
Index HH animals purchased	-0.034^{+}	-0.043	0.0097	0.065^{+}
	(0.018)	(0.032)	(0.021)	(0.038)
Number of Cow purchased	-0.0065	0.0025	-0.026*	0.0021
	(0.012)	(0.0027)	(0.012)	(0.0021)
Number of Bull purchased	0	0	0.0057	0.011
	(.)	(.)	(0.0051)	(0.011)
Number of Calf purchased	0	0	0	0
	(.)	(.)	(.)	(.)
Number of Horse purchased	0	0	0	0
	(.)	(.)	(.)	(.)
Number of Ox purchased	0.0044	0	-0.00079	0
	(0.0044)	(.)	(0.0036)	(.)
Number of Donkey purchased	0	0	0.0016	0
	(.)	(.)	(0.0015)	(.)
Number of Goat purchased	-0.016	-0.042	0.012	0.069^{+}
	(0.027)	(0.059)	(0.021)	(0.040)
Number of Sheep purchased	0	0	-0.0018	0
	(.)	(.)	(0.0018)	(.)
Number of Pig purchased	-0.0072	-0.032	0.013	0.034
	(0.025)	(0.063)	(0.022)	(0.037)
Number of Chicken purchased	-0.22	-0.17	-0.54	0.11
	(0.24)	(0.36)	(0.86)	(1.48)
Number of Gfowl purchased	-0.046	-0.021	-0.0065	-0.0046
	(0.035)	(0.020)	(0.0057)	(0.0044)
Number of Ofowl purchased	0.0089	0	0	0
	(0.0086)	(.)	(.)	(.)
Number of Poultry purchased	-0.013	-0.021	0.015	0.032
	(0.0086)	(0.017)	(0.021)	(0.043)
Number of Other purchased	0	0	0	0
	(.)	(.)	(.)	(.)
Observations	2076	895	2505	1052

$+: p < 0.10$; $*: p < 0.05$; $**: p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to animals purchased index, the remaining outcomes are the components the index is comprised of, i.e. the number of various purchased animals in the last 12 months.

Table B.6: Details on components of the HH animals owned index

	Midline July 2018		Endline July 2019	
	(1) All NGOs b/se	(2) NGOs 1 and 2 b/se	(3) All NGOs b/se	(4) NGOs 1 and 2 b/se
Index HH animals owned	-0.013	-0.031	-0.0017	-0.0059
	(0.020)	(0.022)	(0.020)	(0.011)
Number of Cow owned	-0.048	0.13^{+}	-0.22^{+}	0.031^{+}
	(0.16)	(0.077)	(0.12)	(0.016)
Number of Bull owned	0.0044	0.0054	0.021	0.043^{+}
	(0.0031)	(0.0054)	(0.032)	(0.023)
Number of Calf owned	0	0	-0.045	0
	(.)	(.)	(0.037)	(.)
Number of Horse owned	0	0	0	0
	(.)	(.)	(.)	(.)
Number of Ox owned	0.0025	0	0.076^{+}	0
	(0.0062)	(.)	(0.041)	(.)
Number of Donkey owned	0.0020	0	0.0070	0
	(0.0020)	(.)	(0.0048)	(.)
Number of Goat owned	-0.083	-0.23	0.0068	-0.026
	(0.16)	(0.17)	(0.11)	(0.14)
Number of Sheep owned	0.013	0	-0.021	0
	(0.013)	(.)	(0.029)	(.)
Number of Pig owned	0.031	-0.14	-0.078	-0.13
	(0.15)	(0.18)	(0.15)	(0.13)
Number of Chicken owned	0.42	0.34	-0.70	-0.15
	(0.52)	(0.89)	(0.65)	(0.99)
Number of Gfowl owned	-0.081^{*}	-0.053	-0.14	-0.019
	(0.039)	(0.044)	(0.093)	(0.014)
Number of Ofowl owned	-0.028	-0.032	0.0088	0
	(0.022)	(0.031)	(0.0081)	(.)
Number of Poultry owned	-0.023	-0.19	-0.32^{*}	-0.27
	(0.15)	(0.13)	(0.14)	(0.20)
Number of Other owned	0	0	0	0
	(.)	(.)	(.)	(.)
Observations	2076	895	2505	1052

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to animals owned index, the remaining outcomes are the components the index is comprised of, i.e. the number of various animals owned.

Table B.7: Details on components of the HH agricultural inputs index

	Midline July 2018		Endline July 2019	
	(1)	(2)	(3)	(4)
	All NGOs	NGOs 1 and 2	All NGOs	NGOs 1 and 2
	b/se	b/se	b/se	b/se
Index HH agricultural inputs				
Linkage treatment	-0.020	-0.041	-0.016	-0.050
	(0.029)	(0.050)	(0.028)	(0.043)
Seeds purchased (ZMW)				
Linkage treatment	-8.83	-29.8	-14.5	-19.5
	(19.8)	(34.6)	(19.8)	(34.7)
Control group mean	186.3	250.6	182.9	230.5
Chemical fertilizer used (ZMW)				
Linkage treatment	-31.9	-112.1	-51.4	-63.0
	(55.7)	(111.5)	(58.7)	(98.4)
Control group mean	560.0	885.0	574.4	777.9
Organic fertilizer used (ZMW)				
Linkage treatment	-1.78^{+}	-1.84	1.29	0.14
	(1.00)	(1.20)	(1.64)	(0.98)
Control group mean	2.69	1.68	4.13	0.63
Insecticides used (ZMW)				
Linkage treatment	4.09	-2.75	0.41	-5.65
	(8.45)	(6.44)	(4.67)	(8.03)
Control group mean	20.6	21.1	24.5	21.8
Use of paid field workers (ZMW)				
Linkage treatment	2.96	-45.1	-10.3	-67.3
	(45.5)	(42.2)	(34.3)	(64.1)
Control group mean	248.3	370.6	253.9	331.0
Observations	2076	895	2505	1052

$+: p < 0.10$; $*: p < 0.05$; $**: p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to agricultural inputs index, the remaining outcomes are the components the index is comprised of, i.e. the value of various agricultural inputs in ZMW.

Table B.8: Details on components of the HH agricultural assets index

	Midline July 2018		Endline July 2019	
	(1) All NGOs b/se	(2) NGOs 1 and 2 b/se	(3) All NGOs b/se	(4) NGOs 1 and 2 b/se
Index HH agricultural assets	-0.0060	0.016	-0.017	-0.0028
	(0.015)	(0.016)	(0.018)	(0.016)
Number of trained oxen/cows	-0.027	0.021	-0.12	0.027
	(0.065)	(0.042)	(0.100)	(0.042)
Number of ox-drawn plough	-0.045	-0.00073	-0.015	0.013
	(0.034)	(0.017)	(0.033)	(0.014)
Number of disc plough	0.080*	0.17^{+}	-0.0042	0.018
	(0.036)	(0.087)	(0.015)	(0.031)
Number of harrows	-0.0037	0.0015	-0.013	0.013
	(0.015)	(0.0025)	(0.014)	(0.0087)
Number of cultivators	-0.00087	-0.0013	0.0018	0.0017
	(0.015)	(0.014)	(0.0028)	(0.0019)
Number of rippers	0.0065	0	-0.0086	0
	(0.0074)	(.)	(0.010)	(.)
Number of ridger/weeder	-0.095^{+}	-0.27^{+}	-0.012	0.0032
	(0.055)	(0.14)	(0.021)	(0.0039)
Number of planter	-0.0072	-0.0096	-0.0013	0.0038
	(0.011)	(0.0088)	(0.0037)	(0.0029)
Number of fitarelli	-0.0049	0	0.00040	0.0020
	(0.0057)	(.)	(0.0093)	(0.0037)
Number of hand driven tractor	0.00018	0.0026	0	0
	(0.0016)	(0.0025)	(.)	(.)
Number of scotch carts	-0.0021	-0.0035	-0.0094	0.0019
	(0.015)	(0.0066)	(0.015)	(0.0073)
Number of wheel barrow	-0.0088	0.0096	0.00037	-0.018
	(0.018)	(0.027)	(0.017)	(0.031)
Number of water or treadle pump	-0.0069	0.00028	0.0081	-0.0068
	(0.0058)	(0.0089)	(0.0094)	(0.0063)
Number of other irrigation equipment	-0.0019	0.0020	-0.014	-0.0089^{+}
	(0.0083)	(0.0042)	(0.0089)	(0.0053)
Number of knapsack sprayer	0.027	0.041	0.017	-0.0054
	(0.025)	(0.036)	(0.025)	(0.028)
Number of boom sprayer	0.00070	0.0042	-0.013	-0.022
	(0.012)	(0.024)	(0.0098)	(0.016)
Observations	2076	895	2505	1052

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to agricultural assets index, the remaining outcomes are the components the index is comprised of, i.e. the number of various agricultural assets.

Table B.9: Details on components of the HH general assets index

	Midline July 2018		Endline July 2019	
	(1) All NGOs b/se	(2) NGOs 1 and 2 b/se	(3) All NGOs b/se	(4) NGOs 1 and 2 b/se
Index HH agricultural assets	0.0070	0.017	0.0075	0.014
	(0.016)	(0.028)	(0.013)	(0.019)
Number of trucks or lorries	0.0030$^+$	0.0024	-0.00022	0.0023
	(0.0017)	(0.0024)	(0.0017)	(0.0039)
Number of pick-ups, vans or cars	-0.0069	-0.029$^+$	-0.0041	-0.025
	(0.011)	(0.017)	(0.010)	(0.020)
Number of trailer	0.00085	0	0	0
	(0.0017)	(.)	(.)	(.)
Number of motorcycles	-0.0053	-0.0069	-0.0094	-0.0052
	(0.0095)	(0.012)	(0.0084)	(0.012)
Number of bicycles	0.028	0.084	0.062$^+$	0.084
	(0.036)	(0.059)	(0.035)	(0.057)
Number of boats or canoes	0.0022	0.0095	-0.0025	0.0029
	(0.0082)	(0.0085)	(0.0075)	(0.0057)
Number of fishing nets	0.013	0.040	0.0045	0.049
	(0.023)	(0.026)	(0.029)	(0.041)
Number of cattle dip or crush pen	-0.0020	0.0024	0.0094	0
	(0.0028)	(0.0024)	(0.017)	(.)
Number of hand mills	-0.00047	0.0024	0.012	-0.0083
	(0.018)	(0.0024)	(0.012)	(0.0052)
Number of hammer mills	-0.0051	-0.011	-0.0042	-0.0026
	(0.0040)	(0.0074)	(0.0050)	(0.0065)
Number of rump presses	-0.0021	0	-0.00088	0
	(0.0014)	(.)	(0.00088)	(.)
Number of hand-operated maize sheller	-0.0022	-0.0050	0.0065	0.0031
	(0.0088)	(0.017)	(0.0069)	(0.0027)
Number of motorized maize sheller	-0.0045	-0.011	-0.000023	0
	(0.0052)	(0.0077)	(0.0011)	(.)
Number of improved brazier	-0.017	-0.0021	0.092*	0.041
	(0.040)	(0.060)	(0.046)	(0.095)
Number of solar panel	-0.020	-0.017	0.058	0.074
	(0.054)	(0.10)	(0.045)	(0.070)
Number of generator	0.0059	0.0043	0.0051	0.0031
	(0.0075)	(0.0080)	(0.0061)	(0.0056)
Number of cell phone	0.25	0.70	-0.012	0.029
	(0.25)	(0.77)	(0.046)	(0.070)
Number of radio	0.056$^+$	0.017	0.011	0.032
	(0.032)	(0.049)	(0.028)	(0.050)
Number of TV	-0.011	-0.042	-0.024	-0.024
	(0.027)	(0.041)	(0.022)	(0.032)
Number of car battery	0.015	0.041	0.0089	0.015
	(0.025)	(0.038)	(0.024)	(0.035)
Number of sewing machine	-0.032*	-0.053*	-0.029*	-0.041*
	(0.013)	(0.024)	(0.013)	(0.019)
Number of water tank	0.0031	-0.00077	-0.0018	-0.0046
	(0.0063)	(0.0053)	(0.0032)	(0.0046)
Number of standard well	0.011	0.020	0.0038	-0.0057
	(0.020)	(0.046)	(0.012)	(0.030)
Number of borehole	0.0090	0.023$^+$	-0.0013	0.00070
	(0.0078)	(0.014)	(0.0046)	(0.0086)
Number of gas or electric stove	-0.0040	-0.047$^+$	0.0019	-0.024$^+$
	(0.016)	(0.028)	(0.014)	(0.014)
Number of electric iron	0.0088	0.0086	-0.0029	-0.020
	(0.016)	(0.023)	(0.017)	(0.018)
Number of non-electric iron	0.033	0.061	0.017	0.0042
	(0.028)	(0.041)	(0.027)	(0.044)
Number of lounge suite or sofa	-0.013	0.00015	0.0061	-0.010
	(0.073)	(0.12)	(0.051)	(0.073)
Number of houses	0	0	-0.044	-0.0047
	(.)	(.)	(0.031)	(0.037)
Observations	2076	895	2505	1052

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A).NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to general assets index, the remaining outcomes are the components the index is comprised of, i.e. the number of various general assets.

Table B.10: Details on components of the HH expenditure index

	Midline July 2018		Endline July 2019	
	(1) All NGOs b/se	(2) NGOs 1 and 2 b/se	(3) All NGOs b/se	(4) NGOs 1 and 2 b/se
Index HH expenditures				
Linkage treatment	0.022	0.039	0.048	0.034
	(0.029)	(0.050)	(0.036)	(0.046)
Schooling expenditure				
Linkage treatment	-209.0	-386.5^{+}	137.5	-22.5
	(191.8)	(232.5)	(143.6)	(135.0)
Control group mean	1246.2	1415.7	851.4	927.6
Medical expenditure				
Linkage treatment	28.2	70.1^{+}	24.6^{+}	34.4
	(17.9)	(42.4)	(13.5)	(22.8)
Control group mean	40.2	28.6	45.3	39.4
Funeral expenditure				
Linkage treatment	-21.2	-46.1	29.7	36.3
	(32.7)	(48.3)	(45.7)	(68.1)
Control group mean	164.2	152.3	215.1	181.7
Other expenditure				
Linkage treatment	130.0	-62.1	-150.0	-204.1^{+}
	(101.4)	(139.1)	(137.3)	(119.7)
Control group mean	860.1	982.7	885.0	739
Observations	2076	895	2505	1052

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to the expenditure index, the remaining outcomes are the components the index is comprised of, i.e. the value of various types of expenditures in the last 12 months in ZMW.

Table B.11: Details on components of the HH food security index

	Midline July 2018		Endline July 2019	
	(1)	(2)	(3)	(4)
	All NGOs	NGOs 1 and 2	All NGOs	NGOs 1 and 2
	b/se	b/se	b/se	b/se
Index HH food security index				
Linkage treatment	0.033	0.027	-0.016	-0.014
	(0.021)	(0.030)	(0.015)	(0.023)
Issues covering needs at least monthly				
Linkage treatment	0.00095	-0.0030	-0.032*	-0.017
	(0.018)	(0.021)	(0.015)	(0.016)
Control group mean	0.14	0.095	0.17	0.11
Issues covering needs last 12 months				
Linkage treatment	0.0063	0.016	0.013	0.0076
	(0.025)	(0.033)	(0.021)	(0.028)
Control group mean	1.59	1.60	1.68	1.72
Skipped meals at least monthly				
Linkage treatment	0.0038	-0.0024	-0.019+	-0.013
	(0.013)	(0.015)	(0.011)	(0.012)
Control group mean	0.063	0.039	0.079	0.062
Skipped meals last 12 months				
Linkage treatment	0.032+	0.038	0.016	0.0027
	(0.020)	(0.028)	(0.015)	(0.022)
Control group mean	1.78	1.81	1.85	1.87
Observations	2076	895	2505	1052

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to the food security index, the remaining outcomes are the components the index is comprised of, i.e. the value of various sub-indices related to food security in the last 12 months.

C. APPENDIX TO CHAPTER 3

Table C.1: Various measures for social fund usage separately by NGO

	Midline July 2018			Endline July 2019			Phone surveys combined		
	(1) NGO 1 b/se	(2) NGO 3 b/se	(3) NGO 4 b/se	(4) NGO 1 b/se	(5) NGO 3 b/se	(6) NGO 4 b/se	(7) NGO 1 b/se	(8) NGO 3 b/se	(9) NGO 4 b/se
Any pay-outs from social fund									
Social fund treatment	-0.026	-0.024	0.019	-0.0030	-0.0088	-0.13^{+}	0.088	-0.023	0.041
	(0.075)	(0.10)	(0.073)	(0.089)	(0.10)	(0.073)	(0.062)	(0.070)	(0.052)
Control group mean	0.70	0.34	0.37	0.65	0.50	0.45	0.81	0.82	0.83
Any grants from social fund									
Social fund treatment	0.036	-0.052	-0.028	0.036	-0.0014	-0.097	0.069	0.014	-0.018
	(0.082)	(0.094)	(0.054)	(0.088)	(0.10)	(0.066)	(0.067)	(0.086)	(0.060)
Control group mean	0.55	0.28	0.20	0.57	0.44	0.32	0.79	0.78	0.78
Any loans from social fund									
Social fund treatment	-0.040	0.063	0.063	-0.024	0.010	-0.013	-0.081	0.028	0.11^{+}
	(0.071)	(0.060)	(0.064)	(0.077)	(0.047)	(0.051)	(0.069)	(0.075)	(0.066)
Control group mean	0.27	0.06	0.18	0.18	0.06	0.14	0.26	0.16	0.36
Value of pay-outs from social fund									
Social fund treatment	21.0	47.6	41.6	12.5	41.4	-47.0	181.4^{+}	53.6	-44.5
	(51.6)	(40.1)	(25.3)	(32.8)	(30.8)	(30.5)	(104.7)	(47.5)	(119.8)
Control group mean	190.76	49.92	50.61	117.72	68.40	116.23	405.74	199.36	544.82
Value of grants from social fund									
Social fund treatment	33.3	7.87	-4.19	26.5	41.1	-29.6^{+}	117.6	66.4	-36.4
	(30.6)	(15.9)	(8.91)	(23.6)	(25.8)	(16.0)	(72.3)	(41.1)	(33.9)
Control group mean	83.45	30.72	25.43	71.25	46.40	57.30	315.22	145.86	196.54
Value of loans from social fund									
Social fund treatment	-16.7	37.3	40.5^{+}	-11.1	1.81	-17.3	48.1	7.65	32.0
	(32.2)	(29.2)	(21.9)	(18.9)	(16.8)	(19.1)	(55.3)	(27.7)	(100.2)
Control group mean	98.73	19.20	25.18	42.06	17.84	48.51	60.51	52.00	319.71
Observations	446	446	446	447	447	447	447	447	447

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Outcomes are accumulated over all reported pay-outs in a reference period of the last 12 months. For phone surveys this is based on the report from several waves asking about each month separately for the period of July 2019 to June 2020. Value outcomes are in ZMW and winsorized at 1% and 99%. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). For overall results refer to Table 3.3. For results related to the reported purpose of funeral or sickness, the target purposes of the intervention, refer to Tables C.2 and C.3 in appendix C for aggregate and NGO specific results respectively.

Table C.2: Various measures for social fund usage related to funeral or sickness

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se	(3) Phone surveys combined b/se
Any pay-outs from SF for funeral or sickness			
Social fund treatment	-0.065	-0.024	0.0043
	(0.043)	(0.046)	(0.041)
Control group mean	0.40	0.46	0.80
Any grants from SF for funeral or sickness			
Social fund treatment	-0.036	-0.025	0.025
	(0.041)	(0.045)	(0.038)
Control group mean	0.32	0.41	0.78
Any loans from SF for funeral or sickness			
Social fund treatment	-0.024	0.00099	0.014
	(0.029)	(0.028)	(0.035)
Control group mean	0.11	0.07	0.16
Value of pay-outs from SF for funeral or sickness			
Social fund treatment	4.69	6.00	28.8
	(14.6)	(13.4)	(27.0)
Control group mean	69.94	68.43	226.81
Value of grants from SF for funeral or sickness			
Social fund treatment	8.67	5.38	11.3
	(9.40)	(11.7)	(25.7)
Control group mean	38.91	55.29	190.06
Value of loans from SF for funeral or sickness			
Social fund treatment	-1.85	0.052	9.96
	(10.1)	(4.20)	(9.05)
Control group mean	30.02	11.00	31.00
Observations	446	447	447

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Outcomes are accumulated over reported pay-outs with reported purpose funeral or sickness in a reference period of the last 12 months. For phone surveys this is based on the report from several waves asking about each month separately for the period of July 2019 to June 2020. Value outcomes are in ZMW and winsorized at 1% and 99%. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. For NGO specific results refer to Table C.3. For results related to any purposes, the refer to Tables 3.3 and C.1 in appendix C for aggregate and NGO specific results respectively.

Table C.3: Various measures for social fund usage separately by NGO

	Midline July 2018			Endline July 2019			Phone surveys combined		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	NGO 1	NGO 3	NGO 4	NGO 1	NGO 3	NGO 4	NGO 1	NGO 3	NGO 4
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Any pay-outs from SF for funeral or sickness									
Social fund treatment	-0.029	-0.037	-0.072	0.044	0.011	-0.083	0.041	0.027	-0.0021
	(0.076)	(0.10)	(0.058)	(0.090)	(0.10)	(0.069)	(0.067)	(0.082)	(0.061)
Control group mean	0.69	0.30	0.24	0.57	0.48	0.36	0.81	0.78	0.81
Any grants from SF for funeral or sickness									
Social fund treatment	0.029	-0.032	-0.052	0.012	-0.0014	-0.070	0.060	0.034	-0.012
	(0.084)	(0.093)	(0.053)	(0.087)	(0.10)	(0.065)	(0.067)	(0.085)	(0.061)
Control group mean	0.54	0.26	0.20	0.54	0.44	0.30	0.79	0.76	0.77
Any loans from SF for funeral or sickness									
Social fund treatment	-0.023	-0.0049	0.011	-0.013	0.013	0.00069	-0.083	0.025	0.12*
	(0.071)	(0.037)	(0.030)	(0.069)	(0.040)	(0.036)	(0.063)	(0.055)	(0.055)
Control group mean	0.25	0.04	0.04	0.12	0.04	0.06	0.22	0.08	0.15
Value of pay-outs from SF for funeral or sickness									
Social fund treatment	17.0	16.2	0.12	25.6	47.5^{+}	-29.4^{+}	81.7	57.8	-18.1
	(40.2)	(23.3)	(11.7)	(25.6)	(27.4)	(17.5)	(60.2)	(40.3)	(37.6)
Control group mean	151.46	33.80	31.49	80.59	55.16	67.30	304.85	148.96	206.81
Value of grants from SF for funeral or sickness									
Social fund treatment	29.8	14.4	-6.42	23.0	40.9	-21.7	96.5	58.6	-44.8
	(26.0)	(13.5)	(8.14)	(22.0)	(25.7)	(14.6)	(60.9)	(38.0)	(33.5)
Control group mean	72.51	22.80	23.30	65.37	45.40	53.57	256.40	131.46	175.48
Value of loans from SF for funeral or sickness									
Social fund treatment	-8.19	2.16	6.54	1.56	8.10	-4.40	-1.52	2.04	15.8
	(26.2)	(13.4)	(9.48)	(10.0)	(6.22)	(5.33)	(18.0)	(12.6)	(15.6)
Control group mean	75.75	11.00	8.19	15.22	5.60	11.06	37.35	17.00	27.61
Observations	446	446	446	447	447	447	447	447	447

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Outcomes are accumulated over reported pay-outs with reported purpose funeral or sickness in a reference period of the last 12 months. For phone surveys this is based on the report from several waves asking about each month separately for the period of July 2019 to June 2020. Value outcomes are in ZMW and winsorized at 1% and 99%. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). For overall results refer to Table C.2. For results related to any purposes, the refer to Tables 3.3 and C.1 in appendix C for aggregate and NGO specific results respectively.

Table C.4: SF used to finance various shocks by NGO

	Midline July 2018				Endline July 2019			
	(1) All b/se	(2) NGO 1 b/se	(3) NGO 3 b/se	(4) NGO 4 b/se	(5) All b/se	(6) NGO 1 b/se	(7) NGO 3 b/se	(8) NGO 4 b/se
Support from SF for any shocks								
Social fund treatment	0.024	0.027	0.028	-0.0015	-0.0049	-0.022	0.022+	-0.0069
	(0.018)	(0.050)	(0.021)	(0.024)	(0.011)	(0.027)	(0.013)	(0.017)
Control group mean	0.07	0.13	0.03	0.07	0.07	0.12	0.01	0.06
Observations	1094	304	331	459	1809	532	483	794
# Support from SF for any shocks								
Social fund treatment	0.013	-0.027	0.018	0.012	-0.012	-0.016	0.010	-0.015
	(0.018)	(0.048)	(0.019)	(0.027)	(0.012)	(0.027)	(0.013)	(0.020)
Control group mean	0.06	0.12	0.02	0.05	0.06	0.09	0.01	0.07
Observations	1069	292	327	450	1782	511	479	792
Support from SF for funeral costs								
Social fund treatment	0.039+	0.0060	0.0043	0.077*	-0.00016	0.052	0.020	-0.014
	(0.023)	(0.050)	(0.036)	(0.037)	(0.019)	(0.058)	(0.020)	(0.030)
Control group mean	0.06	0.10	0.03	0.05	0.07	0.12	0.01	0.07
Observations	513	157	146	210	730	223	190	317
# Support from SF for funeral costs								
Social fund treatment	0.043+	0.0060	0.0043	0.086*	-0.00051	0.045	0.031	-0.014
	(0.024)	(0.050)	(0.036)	(0.040)	(0.020)	(0.059)	(0.026)	(0.030)
Control group mean	0.06	0.10	0.03	0.05	0.07	0.13	0.01	0.07
Observations	513	157	146	210	730	223	190	317
Support from SF for medical costs								
Social fund treatment	0.026	0.093	0.074+	-0.047	0.014	0.11	0.068	-0.10+
	(0.036)	(0.097)	(0.044)	(0.049)	(0.042)	(0.089)	(0.060)	(0.060)
Control group mean	0.13	0.26	0.06	0.11	0.13	0.18	0.04	0.13
Observations	316	89	106	121	263	106	62	95
# Support from SF for medical costs								
Social fund treatment	-0.0096	-0.13	0.051+	-0.0069	-0.030	0.12+	-0.040	-0.12+
	(0.036)	(0.100)	(0.029)	(0.045)	(0.033)	(0.070)	(0.038)	(0.062)
Control group mean	0.08	0.22	0.04	0.04	0.06	0.00	0.04	0.14
Observations	291	77	102	112	236	85	58	93
Support from SF for business shocks								
Social fund treatment	-0.0044	-0.022	0.0067	-0.022	-0.0051	-0.040**	0.0038	0.0073
	(0.010)	(0.020)	(0.0065)	(0.015)	(0.0084)	(0.016)	(0.0076)	(0.013)
Control group mean	0.02	0.03	0.00	0.02	0.02	0.04	0.00	0.03
Observations	794	153	276	365	1630	428	455	747
# Support from SF for business shocks								
Social fund treatment	-0.0069	-0.022	0.0067	-0.028	-0.0064	-0.040**	0.0038	0.0046
	(0.011)	(0.020)	(0.0065)	(0.018)	(0.0086)	(0.016)	(0.0076)	(0.013)
Control group mean	0.02	0.03	0.00	0.03	0.03	0.04	0.00	0.03
Observations	794	153	276	365	1630	428	455	747

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Outcomes refers to support from social fund reportedly used to finance any shock such as funerals, medical expenditures, or business related shocks. Observations refer to households that reported shocks such as funerals, medical expenditures, or business related shocks. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). For estimations on the whole sample, NGO fixed effects are always included, but excluded for NGO specific estimations.

Table C.5: Loans used to finance various shocks by NGO

	Midline July 2018				Endline July 2019			
	(1) All b/se	(2) NGO 1 b/se	(3) NGO 3 b/se	(4) NGO 4 b/se	(5) All b/se	(6) NGO 1 b/se	(7) NGO 3 b/se	(8) NGO 4 b/se
Loans for any shocks								
Social fund treatment	0.024	0.12^{**}	-0.013	-0.0095	0.017	-0.00098	-0.011	0.011
	(0.024)	(0.044)	(0.046)	(0.039)	(0.021)	(0.028)	(0.034)	(0.028)
Control group mean	0.18	0.10	0.21	0.22	0.18	0.15	0.19	0.19
Observations	1085	302	329	454	1813	533	484	796
# Loans for any shocks								
Social fund treatment	0.029	0.14^{*}	0.014	-0.031	0.044	0.034	0.020	0.016
	(0.029)	(0.058)	(0.057)	(0.044)	(0.030)	(0.041)	(0.051)	(0.037)
Control group mean	0.20	0.09	0.22	0.24	0.19	0.14	0.21	0.22
Observations	1059	292	324	443	1779	519	475	785
Loans from savings group for any shocks								
Social fund treatment	0.033	0.14^{**}	0.027	-0.011	0.0071	0.020	-0.016	0.016
	(0.022)	(0.050)	(0.037)	(0.037)	(0.017)	(0.038)	(0.038)	(0.022)
Control group mean	0.07	0.05	0.06	0.09	0.09	0.12	0.11	0.06
Observations	706	214	211	281	882	277	229	376
# Loans from savings group for any shocks								
Social fund treatment	0.044^{+}	0.17^{**}	0.023	-0.0083	0.0036	0.020	0.0090	-0.0066
	(0.025)	(0.057)	(0.037)	(0.038)	(0.033)	(0.080)	(0.065)	(0.046)
Control group mean	0.05	0.03	0.04	0.08	0.14	0.18	0.15	0.10
Observations	690	209	206	275	859	266	223	370
Loans for funeral costs								
Social fund treatment	0.011	0.16^{*}	-0.0075	-0.048	0.016	-0.045	0.035	-0.0040
	(0.032)	(0.067)	(0.058)	(0.051)	(0.025)	(0.055)	(0.052)	(0.034)
Control group mean	0.14	0.10	0.18	0.15	0.14	0.18	0.16	0.10
Observations	513	157	146	210	730	223	190	317
# Loans for funeral costs								
Social fund treatment	0.020	0.16^{+}	0.026	-0.048	0.024	-0.048	0.034	-0.0037
	(0.041)	(0.081)	(0.080)	(0.059)	(0.032)	(0.056)	(0.068)	(0.042)
Control group mean	0.16	0.13	0.19	0.17	0.15	0.19	0.19	0.11
Observations	513	157	146	210	730	223	190	317
Loans from savings group for funeral costs								
Social fund treatment	0.030	0.15^{*}	0.025	-0.024	-0.0015	-0.049	0.022	-0.011
	(0.026)	(0.060)	(0.042)	(0.039)	(0.018)	(0.042)	(0.035)	(0.026)
Control group mean	0.06	0.04	0.05	0.08	0.08	0.12	0.06	0.06
Observations	513	157	146	210	730	223	190	317
# Loans from savings group for funeral costs								
Social fund treatment	0.037	0.18^{**}	0.025	-0.024	-0.0060	-0.10	0.034	-0.023
	(0.028)	(0.067)	(0.042)	(0.039)	(0.039)	(0.087)	(0.077)	(0.051)
Control group mean	0.06	0.04	0.05	0.08	0.16	0.25	0.15	0.12
Observations	513	157	146	210	730	223	190	317
Loans for medical costs								
Social fund treatment	0.027	0.072	0.066	-0.028	-0.049	0.0046	-0.15	0.029
	(0.038)	(0.093)	(0.047)	(0.065)	(0.048)	(0.080)	(0.091)	(0.056)
Control group mean	0.12	0.13	0.06	0.16	0.22	0.22	0.35	0.14
Observations	317	87	107	123	270	99	67	104
# Loans for medical costs								
Social fund treatment	0.020	0.049	0.061	-0.031	-0.0067	0.075	-0.16	0.037
	(0.034)	(0.088)	(0.046)	(0.061)	(0.049)	(0.095)	(0.15)	(0.038)
Control group mean	0.06	0.06	0.02	0.09	0.12	0.10	0.28	0.05
Observations	291	77	102	112	236	85	58	93
Loans from savings group for medical costs								
Social fund treatment	0.0022	0.041	0.041	-0.019	-0.0028	0.073	-0.13^{+}	0.068^{+}
	(0.035)	(0.078)	(0.038)	(0.066)	(0.036)	(0.070)	(0.077)	(0.041)
Control group mean	0.09	0.10	0.04	0.10	0.13	0.13	0.24	0.06
Observations	307	82	107	118	259	96	64	99
# Loans from savings group for medical costs								
Social fund treatment	0.012	0.018	0.034	-0.010	0.026	0.061	-0.053	0.061^{*}
	(0.031)	(0.084)	(0.033)	(0.057)	(0.027)	(0.046)	(0.063)	(0.029)
Control group mean	0.04	0.06	0.00	0.05	0.04	0.02	0.12	0.00
Observations	291	77	102	112	236	85	58	93
Loans for business shocks								
Social fund treatment	0.0093	-0.0015	-0.0045	0.0011	0.026	0.029	0.00087	0.025
	(0.024)	(0.051)	(0.029)	(0.040)	(0.022)	(0.025)	(0.033)	(0.023)
Control group mean	0.14	0.07	0.13	0.17	0.12	0.06	0.12	0.15
Observations	794	153	276	365	1630	428	455	747
# Loans for business shocks								
Social fund treatment	0.00060	0.0043	-0.020	-0.0100	0.037	0.037	0.021	0.030
	(0.027)	(0.051)	(0.033)	(0.044)	(0.026)	(0.026)	(0.038)	(0.029)
Control group mean	0.16	0.07	0.16	0.19	0.14	0.06	0.13	0.18
Observations	794	153	276	365	1630	428	455	747

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Outcomes refers to loans reportedly used to finance any shock such as funerals, medical expenditures, or business related shocks. Observations refer to households that reported shocks such as funerals, medical expenditures, or business related shocks. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). For estimations on the whole sample, NGO fixed effects are always included, but excluded for NGO specific estimations.

Table C.6: Sold agriculture assets or goods to finance various shocks by NGO

	Midline July 2018				Endline July 2019			
	(1) All b/se	(2) NGO 1 b/se	(3) NGO 3 b/se	(4) NGO 4 b/se	(5) All b/se	(6) NGO 1 b/se	(7) NGO 3 b/se	(8) NGO 4 b/se
Sold goods for any shocks								
Social fund treatment	0.030	0.00073	-0.0089	0.024	-0.011	-0.028	0.016	0.0019
	(0.028)	(0.058)	(0.054)	(0.046)	(0.026)	(0.046)	(0.049)	(0.034)
Control group mean	0.46	0.45	0.60	0.38	0.32	0.40	0.35	0.25
Observations	1137	328	339	470	1820	543	483	794
# Sold goods for any shocks								
Social fund treatment	0.11	0.070	0.15	0.19	-0.018	-0.12	0.011	0.10
	(0.097)	(0.17)	(0.17)	(0.17)	(0.052)	(0.10)	(0.11)	(0.075)
Control group mean	0.77	0.68	1.11	0.64	0.48	0.61	0.53	0.37
Observations	962	274	277	411	1656	474	429	753
Sold goods for funeral costs								
Social fund treatment	-0.039	-0.029	-0.14^{+}	0.053	0.0091	0.038	-0.036	0.034
	(0.045)	(0.079)	(0.075)	(0.043)	(0.034)	(0.070)	(0.070)	(0.045)
Control group mean	0.36	0.34	0.51	0.29	0.40	0.40	0.49	0.34
Observations	514	157	146	211	730	223	190	317
# Sold goods for funeral costs								
Social fund treatment	-0.079	-0.058	-0.28^{+}	0.11	0.020	0.076	-0.079	0.068
	(0.090)	(0.16)	(0.15)	(0.086)	(0.067)	(0.14)	(0.14)	(0.090)
Control group mean	0.73	0.68	1.03	0.57	0.79	0.80	0.98	0.69
Observations	514	157	146	211	730	223	190	317
Sold goods for medical costs								
Social fund treatment	0.058	0.036	0.044	0.017	-0.016	-0.049	-0.034	-0.11
	(0.043)	(0.073)	(0.065)	(0.070)	(0.047)	(0.074)	(0.082)	(0.078)
Control group mean	0.73	0.76	0.80	0.64	0.63	0.71	0.67	0.51
Observations	466	131	164	171	400	154	112	134
# Sold goods for medical costs								
Social fund treatment	0.37*	0.74*	0.51*	0.22	0.060	-0.36	0.28	0.27
	(0.18)	(0.31)	(0.25)	(0.35)	(0.14)	(0.26)	(0.25)	(0.22)
Control group mean	1.01	1.00	1.22	0.82	0.55	0.88	0.48	0.30
Observations	291	77	102	112	236	85	58	93
Sold goods for business shocks								
Social fund treatment	0.00023	0.011	-0.0037	-0.0098	0.012	-0.0014	0.025	0.015
	(0.032)	(0.052)	(0.055)	(0.054)	(0.015)	(0.033)	(0.026)	(0.023)
Control group mean	0.21	0.09	0.28	0.21	0.10	0.13	0.09	0.10
Observations	794	153	276	365	1630	428	455	747
# Sold goods for business shocks								
Social fund treatment	0.0068	0.013	0.0065	-0.012	0.013	-0.0041	0.027	0.022
	(0.043)	(0.052)	(0.076)	(0.073)	(0.020)	(0.050)	(0.033)	(0.031)
Control group mean	0.26	0.09	0.34	0.27	0.12	0.16	0.10	0.12
Observations	794	153	276	365	1630	428	455	747

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Outcomes refers to selling agricultural goods and assets reportedly used to finance any shock such as funerals, medical expenditures, or business related shocks. Observations refer to households that reported shocks such as funerals, medical expenditures, or business related shocks. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A).

Table C.7: Knowledge and attitude outcomes by NGO

	Midline July 2018			Endline July 2019		
	(1) NGO 1 b/se	(2) NGO 3 b/se	(3) NGO 4 b/se	(4) NGO 1 b/se	(5) NGO 3 b/se	(6) NGO 4 b/se
Would participate in insurance scheme						
Social fund treatment	-0.0018	-0.041	0.048^{+}	-0.019	0.020	-0.047^{+}
	(0.028)	(0.040)	(0.028)	(0.039)	(0.041)	(0.028)
Control group mean	0.91	0.81	0.83	0.71	0.66	0.80
Observations	462	315	621	588	470	831
Level of trust in insurance companies						
Social fund treatment	-0.011	-0.093	-0.091	-0.12	-0.073	0.11
	(0.14)	(0.15)	(0.12)	(0.10)	(0.12)	(0.089)
Control group mean	1.92	2.27	2.51	2.41	2.27	2.12
Observations	223	160	365	384	379	687
Trust in insurance companies						
Social fund treatment	-0.081	0.054	0.047	0.088^{+}	0.026	-0.060
	(0.060)	(0.077)	(0.051)	(0.046)	(0.049)	(0.040)
Control group mean	0.60	0.44	0.39	0.48	0.51	0.62
Observations	295	213	443	473	426	716
Knowledge about insurance						
Social fund treatment	0.015	-0.068	-0.014	0.0099	-0.028	0.023
	(0.053)	(0.049)	(0.064)	(0.056)	(0.041)	(0.045)
Control group mean	0.08	-0.14	0.02	-0.03	-0.05	0.05
Observations	555	384	701	696	512	906
Attitude towards insurance						
Social fund treatment	-0.037	-0.16^{*}	-0.027	0.095	-0.12	0.0040
	(0.074)	(0.075)	(0.066)	(0.079)	(0.087)	(0.049)
Control group mean	0.11	0.05	-0.08	0.01	0.05	-0.03
Observations	245	138	373	126	122	297

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. While the hypothetical participation in an insurance scheme and knowledge questions were asked to everyone, the attitude and trust outcomes are asked conditional on the respondent being familiar with the term insurance. Level of trust is on a 4 point scale (1=complete, 4=none at all). Trust refers to the first two levels. Knowledge about and attitude towards insurance are indexes, i.e. the mean of 5 and 8 items each of which is standardized to the control group. For overall results refer to Table 3.5. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A).

Table C.8: Reported economic impact of shocks and loci of control by NGO

	Midline July 2018			Endline July 2019		
	(1) NGO 1 b/se	(2) NGO 3 b/se	(3) NGO 4 b/se	(4) NGO 1 b/se	(5) NGO 3 b/se	(6) NGO 4 b/se
Average economic impact across shocks						
Social fund treatment	0.045	-0.037	-0.032	0.018	0.024	0.029
	(0.13)	(0.11)	(0.074)	(0.082)	(0.073)	(0.048)
Control group mean	1.86	1.83	1.87	1.92	1.57	1.57
Observations	261	315	425	509	477	786
# of shocks with serious economic impact						
Social fund treatment	0.018	0.074	0.053	-0.046	-0.045	-0.010
	(0.13)	(0.15)	(0.099)	(0.099)	(0.17)	(0.11)
Control group mean	0.83	1.13	1.05	0.85	2.34	1.86
Observations	261	315	425	509	477	786
# of shocks with economic impact						
Social fund treatment	0.010	-0.11	0.041	-0.16	-0.027	0.037
	(0.13)	(0.17)	(0.11)	(0.14)	(0.17)	(0.13)
Control group mean	1.17	1.92	1.73	1.69	2.87	2.31
Observations	261	315	425	509	477	786
Average economic impact across business shocks						
Social fund treatment	0.23	-0.11	-0.016	-0.029	-0.016	0.026
	(0.16)	(0.089)	(0.080)	(0.083)	(0.078)	(0.050)
Control group mean	1.60	1.71	1.67	1.91	1.58	1.55
Observations	153	276	365	427	452	746
# of business shocks with serious economic impact						
Social fund treatment	-0.15	0.082	0.025	-0.070	0.085	0.0074
	(0.17)	(0.13)	(0.096)	(0.079)	(0.14)	(0.086)
Control group mean	1.01	1.06	1.00	0.77	2.04	1.57
Observations	153	276	365	427	452	746
# of business shocks with economic impact						
Social fund treatment	-0.23	-0.12	0.023	-0.11	0.16	0.041
	(0.15)	(0.15)	(0.11)	(0.12)	(0.15)	(0.11)
Control group mean	1.36	1.72	1.60	1.55	2.47	1.97
Observations	153	276	365	427	452	746
Average economic impact across funerals						
Social fund treatment	-0.21	0.033	-0.032	0.11	0.17	0.10
	(0.20)	(0.17)	(0.16)	(0.16)	(0.13)	(0.091)
Control group mean	2.20	2.09	2.34	1.92	1.58	1.53
Observations	156	149	215	216	190	316
# of funerals with serious economic impact						
Social fund treatment	0.13	0.069	-0.012	0.015	-0.070	-0.077
	(0.10)	(0.11)	(0.071)	(0.10)	(0.11)	(0.082)
Control group mean	0.47	0.42	0.37	0.47	0.90	0.95
Observations	156	149	215	216	190	316
# of funerals with economic impact						
Social fund treatment	0.15	-0.042	-0.023	-0.15	-0.15	-0.078
	(0.11)	(0.11)	(0.088)	(0.10)	(0.12)	(0.083)
Control group mean	0.72	0.88	0.66	0.89	1.16	1.13
Observations	156	149	215	216	190	316
Internal locus of control index						
Social fund treatment	-0.024	-0.11^{*}	-0.040	0.0056	0.19^{**}	-0.026
	(0.056)	(0.051)	(0.045)	(0.058)	(0.060)	(0.045)
Control group mean	0.20	-0.02	-0.13	0.35	-0.11	-0.19
Observations	555	384	701	696	512	906
External locus of control index						
Social fund treatment	-0.0053	0.058	-0.013	-0.016	0.13^{+}	0.052
	(0.041)	(0.076)	(0.052)	(0.055)	(0.070)	(0.046)
Control group mean	-0.19	-0.08	0.18	0.07	-0.21	0.05
Observations	555	384	701	695	512	906

$+ : p < 0.10;\ * : p < 0.05;\ ** : p < 0.01$

Notes. Shocks comprises both business shocks and funerals. Impact was assessed on a 4 point scale (*1=serious*, *4=none at all*). *Serious economic impact* refers to the first level, *economic impact* refers to the first two levels. Locus of control indexes are the mean of 5 items which are standardized to the control group. For details on the external and internal locus of control indexes refer to Tables C.10 and C.9 in appendix C. For overall results refer to Table 3.6. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A).

Table C.9: Details on components of the internal locus of control index

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Internal locus of control index		
Social fund treatment	-0.055^{+}	0.031
	(0.029)	(0.030)
My life is determined by my own actions.		
Social fund treatment	0.022	0.046^{+}
	(0.044)	(0.028)
Control group mean	1.70	1.62
When I get what I want, it is usually because I worked hard for it.		
Social fund treatment	-0.10^{**}	0.035
	(0.038)	(0.024)
Control group mean	1.54	1.56
I am usually able to protect my personal interests.		
Social fund treatment	-0.041	0.020
	(0.045)	(0.032)
Control group mean	1.76	1.74
I can mostly determine what will happen in my life.		
Social fund treatment	-0.063	0.020
	(0.059)	(0.052)
Control group mean	2.65	2.18
When I make plans. I am almost certain to make them work.		
Social fund treatment	-0.064	-0.0034
	(0.042)	(0.032)
Control group mean	1.77	1.76
Observations	1674	2231

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. "Internal locus of control index" refers to an index based on the 5 items presented in the table. Each of these items is on a 5-point scale from *1=strongly agree* to *5=strongly disagree*. It is constructed by averaging over these items after they were standardized to the control group. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). Observations reflects maximum number not accounting for missing values. NGO fixed effects are always included.

Table C.10: Details on components of the external locus of control index

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
External locus of control index		
Social fund treatment	0.000017	0.036
	(0.031)	(0.032)
To a great extent my life is controlled by accidental/chance happenings.		
Social fund treatment	0.0044	0.097^{+}
	(0.064)	(0.052)
Control group mean	2.97	2.47
I feel like what happens in my life is determined by others.		
Social fund treatment	0.048	0.045
	(0.059)	(0.056)
Control group mean	3.08	2.95
It is not always wise for me to plan too far ahead ...		
Social fund treatment	-0.057	-0.0068
	(0.061)	(0.056)
Control group mean	2.93	2.57
My life is chiefly controlled by other powerful people.		
Social fund treatment	0.095	0.045
	(0.072)	(0.057)
Control group mean	3.01	2.79
People like myself have little chance of protecting personal interest.		
Social fund treatment	-0.074	0.052
	(0.066)	(0.055)
Control group mean	3.23	2.84
Observations	1674	2230

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. "External locus of control index" refers to an index based on the 5 items presented in the table. Each of these items is on a 5-point scale from *1=strongly agree* to *5=strongly disagree.* It is constructed by averaging over these items after they were standardized to the control group. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). Observations reflects maximum number not accounting for missing values. NGO fixed effects are always included.

Table C.11: Details on components of the HH animals purchased index

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Index HH animals purchased	0.14	-0.014
	(0.090)	(0.016)
Number of Cow purchased	-0.026	-0.0064
	(0.064)	(0.014)
Number of Bull purchased	0	-0.0042
	(.)	(0.0055)
Number of Calf purchased	0	0
	(.)	(.)
Number of Horse purchased	0	0
	(.)	(.)
Number of Ox purchased	0.0049	0.0027
	(0.0049)	(0.0040)
Number of Donkey purchased	0	0.0018
	(.)	(0.0017)
Number of Goat purchased	0.055*	-0.0076
	(0.022)	(0.022)
Number of Sheep purchased	0	-0.0020
	(.)	(0.0020)
Number of Pig purchased	0.016	-0.014
	(0.026)	(0.024)
Number of Chicken purchased	0.090	-0.031
	(0.23)	(0.93)
Number of Gfowl purchased	0.044	0.011+
	(0.035)	(0.0060)
Number of Ofowl purchased	0.010	0
	(0.0097)	(.)
Number of Poultry purchased	0.0056	-0.0045
	(0.0099)	(0.022)
Number of Other purchased	0	0
	(.)	(.)
Observations	1777	2170

$+: p < 0.10$; $*: p < 0.05$; $**: p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to animals purchased index, the remaining outcomes are the components the index is comprised of, i.e. the number of various purchased animals in the last 12 months.

Table C.12: Details on components of the HH animals owned index

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Index HH animals owned	0.017	0.010
	(0.022)	(0.021)
Number of Cow owned	0.061	-0.15
	(0.18)	(0.14)
Number of Bull owned	-0.0053	0.022
	(0.0037)	(0.037)
Number of Calf owned	0	-0.054
	(.)	(0.043)
Number of Horse owned	0	0
	(.)	(.)
Number of Ox owned	-0.0024	0.023
	(0.0071)	(0.048)
Number of Donkey owned	0.0022	0.0061
	(0.0022)	(0.0055)
Number of Goat owned	0.16	-0.0053
	(0.18)	(0.12)
Number of Sheep owned	0.015	-0.033
	(0.014)	(0.033)
Number of Pig owned	-0.053	0.084
	(0.17)	(0.17)
Number of Chicken owned	0.035	-0.83
	(0.55)	(0.68)
Number of Gfowl owned	0.043	0.031
	(0.046)	(0.11)
Number of Ofowl owned	0.025	-0.011
	(0.026)	(0.010)
Number of Poultry owned	0.18	0.045
	(0.17)	(0.13)
Number of Other owned	0	0
	(.)	(.)
Observations	1777	2170

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to animals owned index, the remaining outcomes are the components the index is comprised of, i.e. the number of various animals owned.

Table C.13: Details on components of the HH agricultural inputs index

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Index HH agricultural inputs		
Social fund treatment	0.018	-0.020
	(0.030)	(0.030)
Seeds purchased (ZMW)		
Social fund treatment	20.4	5.75
	(21.3)	(21.9)
Control group mean	171.9	168.9
Chemical fertilizer used (ZMW)		
Social fund treatment	82.1	-28.7
	(60.6)	(62.4)
Control group mean	467.6	526.0
Organic fertilizer used (ZMW)		
Social fund treatment	-0.94	-3.46^{+}
	(1.13)	(1.80)
Control group mean	2.53	7.23
Insecticides used (ZMW)		
Social fund treatment	-4.50	-7.71
	(9.00)	(5.38)
Control group mean	24.8	30.9
Use of paid field workers (ZMW)		
Social fund treatment	16.0	24.9
	(52.7)	(36.0)
Control group mean	224.5	214.1
Observations	1777	2170

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to agricultural inputs index, the remaining outcomes are the components the index is comprised of, i.e. the value of various agricultural inputs in ZMW.

Table C.14: Details on components of the HH agricultural assets index

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Index HH agricultural assets	-0.013	0.016
	(0.017)	(0.024)
Number of trained oxen/cows	0.067	-0.071
	(0.073)	(0.11)
Number of ox-drawn plough	0.034	0.027
	(0.038)	(0.039)
Number of disc plough	-0.082*	0.015
	(0.035)	(0.017)
Number of harrows	0.0026	0.016
	(0.016)	(0.016)
Number of cultivators	0.015	-0.0047
	(0.018)	(0.0031)
Number of rippers	-0.0027	0.0040
	(0.0081)	(0.012)
Number of ridger/weeder	-0.036	-0.0027
	(0.044)	(0.025)
Number of planter	-0.0031	-0.0035
	(0.012)	(0.0044)
Number of fitarelli	-0.0063	0.0082
	(0.0066)	(0.011)
Number of hand driven tractor	-0.0025	0
	(0.0017)	(.)
Number of scotch carts	0.024	0.012
	(0.016)	(0.017)
Number of wheel barrow	0.00064	-0.0096
	(0.019)	(0.016)
Number of water or treadle pump	-0.012*	0.0034
	(0.0058)	(0.010)
Number of other irrigation equipment	-0.0022	-0.010
	(0.0097)	(0.010)
Number of knapsack sprayer	-0.025	-0.011
	(0.026)	(0.027)
Number of boom sprayer	-0.00046	-0.0017
	(0.014)	(0.0098)
Observations	1777	2170

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to agricultural assets index, the remaining outcomes are the components the index is comprised of, i.e. the number of various agricultural assets.

Table C.15: Details on components of the HH general assets index

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Index HH agricultural assets	0.026	-0.0078
	(0.016)	(0.014)
Number of trucks or lorries	0.0023	-0.00014
	(0.0016)	(0.0013)
Number of pick-ups, vans or cars	0.015	-0.0023
	(0.012)	(0.0082)
Number of trailer	0.00090	0
	(0.0020)	(.)
Number of motorcycles	0.0071	-0.0041
	(0.0097)	(0.010)
Number of bicycles	-0.0061	-0.014
	(0.038)	(0.039)
Number of boats or canoes	0.0032	0.0100
	(0.0093)	(0.0086)
Number of fishing nets	0.045^{+}	0.047
	(0.025)	(0.032)
Number of cattle dip or crush pen	-0.00030	0.0086
	(0.0034)	(0.020)
Number of hand mills	-0.0011	-0.0011
	(0.020)	(0.013)
Number of hammer mills	0.0015	0.0069
	(0.0043)	(0.0053)
Number of rump presses	0.000010	0.00089
	(0.0017)	(0.00089)
Number of hand-operated maize sheller	0.0082	-0.0074
	(0.0065)	(0.0073)
Number of motorized maize sheller	-0.011^{+}	-0.000047
	(0.0063)	(0.0014)
Number of improved brazier	-0.082^{*}	-0.054
	(0.041)	(0.046)
Number of solar panel	0.026	-0.054
	(0.058)	(0.050)
Number of generator	0.00090	-0.0033
	(0.0083)	(0.0071)
Number of cell phone	-0.012	-0.0099
	(0.047)	(0.047)
Number of radio	0.013	-0.020
	(0.036)	(0.030)
Number of TV	0.050	-0.046^{+}
	(0.037)	(0.024)
Number of car battery	0.035	0.041
	(0.027)	(0.027)
Number of sewing machine	-0.0098	-0.019
	(0.012)	(0.012)
Number of water tank	-0.00014	0.000054
	(0.0069)	(0.0030)
Number of standard well	-0.012	-0.0076
	(0.022)	(0.012)
Number of borehole	-0.0098	-0.0014
	(0.0074)	(0.0040)
Number of gas or electric stove	-0.019	-0.028
	(0.017)	(0.021)
Number of electric iron	-0.028^{+}	-0.040^{*}
	(0.016)	(0.018)
Number of non-electric iron	0.064^{*}	0.0079
	(0.029)	(0.029)
Number of lounge suite or sofa	-0.057	-0.014
	(0.079)	(0.053)
Number of houses	0	0.015
	(.)	(0.033)
Observations	1777	2170

$+: p < 0.10$; $*: p < 0.05$; $**: p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to general assets index, the remaining outcomes are the components the index is comprised of, i.e. the number of various general assets.

Table C.16: Details on components of the HH expenditure index

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Index HH expenditures		
Social fund treatment	0.038	-0.042
	(0.031)	(0.033)
Schooling expenditure		
Social fund treatment	-191.0	-253.6
	(207.4)	(177.4)
Control group mean	1181.2	1050.6
Medical expenditure		
Social fund treatment	17.6	-5.61
	(12.2)	(15.0)
Control group mean	37.6	62.1
Funeral expenditure		
Social fund treatment	19.2	-39.2
	(30.2)	(50.2)
Control group mean	151.4	259.2
Other expenditure		
Social fund treatment	33.6	-58.2
	(113.1)	(136.8)
Control group mean	942.3	906.8
Observations	1777	2170

$+ : p < 0.10$; $* : p < 0.05$; $** : p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to expenditure index, the remaining outcomes are the components the index is comprised of, i.e. the value of various types of expenditures in the last 12 months in ZMW.

Table C.17: Details on components of the HH food security index

	(1) Midline July 2018 b/se	(2) Endline July 2019 b/se
Index HH food security index		
Social fund treatment	0.017	0.015
	(0.022)	(0.017)
Issues covering needs at least monthly		
Social fund treatment	-0.011	0.0062
	(0.020)	(0.017)
Control group mean	0.16	0.17
Issues covering food needs last 12 months		
Social fund treatment	0.053^{+}	0.026
	(0.028)	(0.022)
Control group mean	1.56	1.66
Skipped meals at least monthly		
Social fund treatment	0.0024	-0.00032
	(0.014)	(0.013)
Control group mean	0.072	0.076
Skipped meals last 12 months		
Social fund treatment	0.0079	0.0060
	(0.023)	(0.018)
Control group mean	1.78	1.84

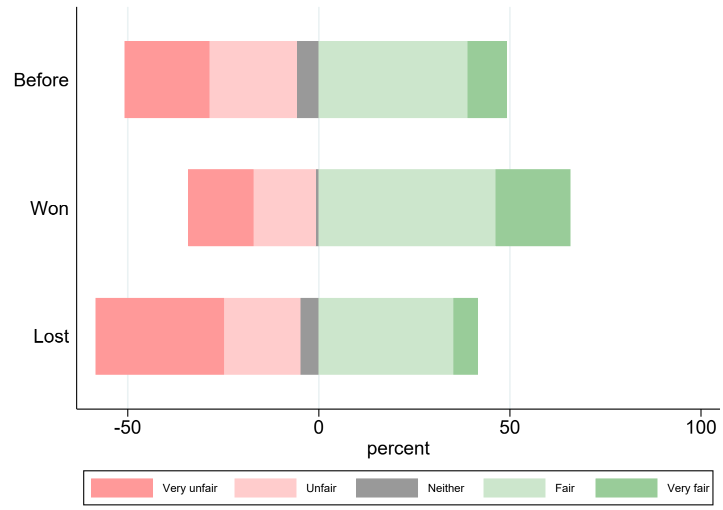
$+: p < 0.10$; $*: p < 0.05$; $**: p < 0.01$

Notes. Estimations are done using pdslasso command in Stata (included covariates can be found in Tables A.1, A.2, and A.3 in appendix A). NGO fixed effects are always included. Observations reflects maximum number not accounting for missing values. First outcome refers to food security index, the remaining outcomes are the components the index is comprised of, i.e. the value of various sub-indexes related to food security in the last 12 months.

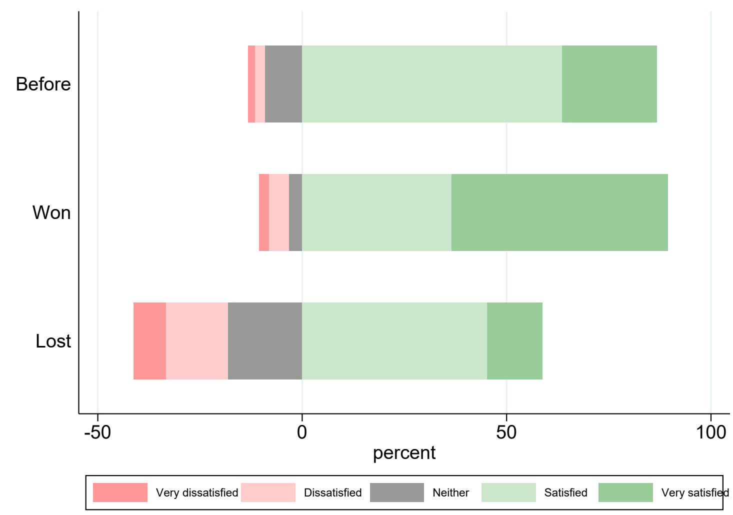
D. APPENDIX TO CHAPTER 4

D.A. Questions and response distribution

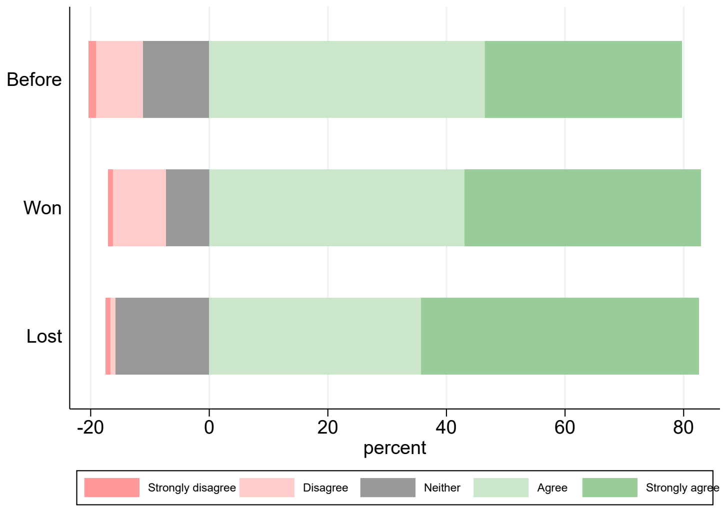
Q1: How fair do you think it is that some respondents win 80 Kwacha in the lottery while others do not?



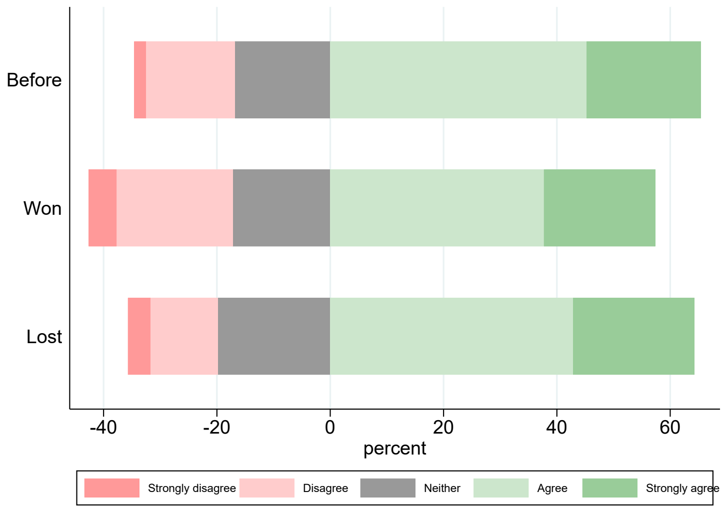
Q2: How satisfied are you that you had/have a chance to win 80 Kwacha for your participation in the interview?



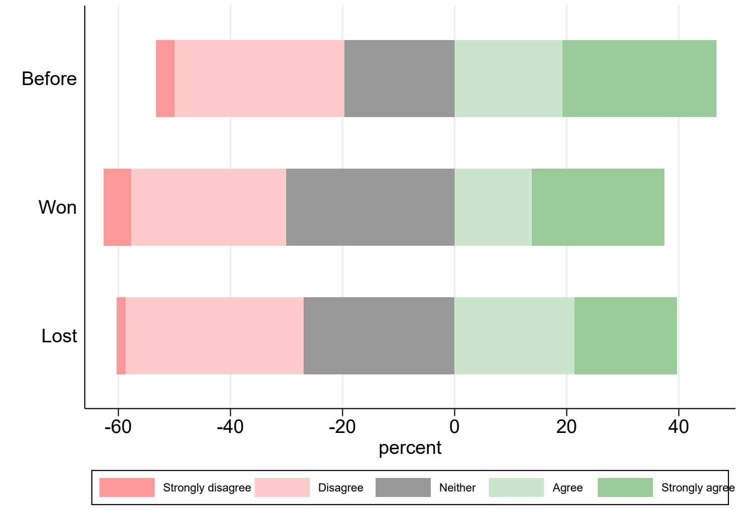
Q3: Do you agree or disagree with the following statement: Survey respondents should be compensated for their participation in a survey interview.



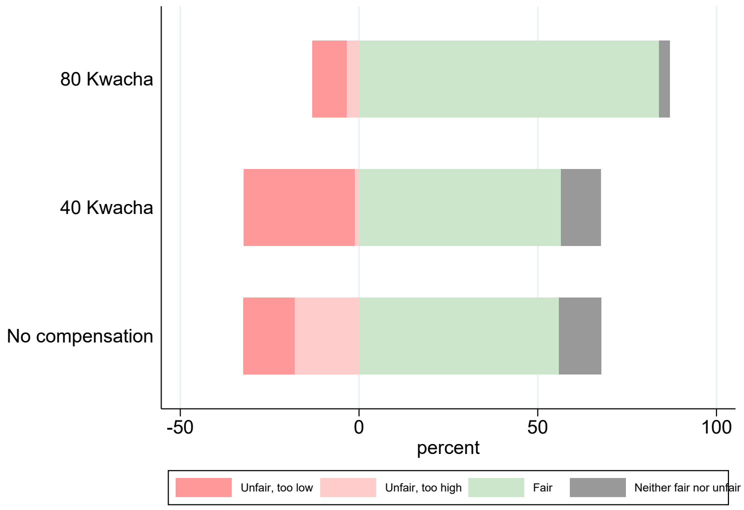
Q4: Do you agree or disagree with the following statement: Either all or none of the respondents in a survey should be compensated.



Q5: Do you agree or disagree with the following statement: When there is a compensation paid to a respondent for taking part in an interview about her/his savings group, then the compensation should be shared with the group.



Q6: Assume a survey respondent is offered *80 Kwacha/40 Kwacha/no compensation* for participating in a survey. Judging the amount offered, do you think this is fair, or unfair because it is too little, or unfair because it is too much?



D.B. Survey experiment full model

In the following I present the analysis of the survey experiment based on the model which is fully saturated with respect to the different randomized groups.

$$\begin{aligned} Y_i = & \alpha_0 D_{Before} \times D_0 + \alpha_{40} D_{Before} \times D_{40} + \alpha_{80} D_{Before} \times D_{80} \\ & + \beta_0 D_{Won} \times D_0 + \beta_{40} D_{Won} \times D_{40} + \beta_{80} D_{Won} \times D_{80} \\ & + \gamma_0 D_{Lost} \times D_0 + \gamma_{40} D_{Lost} \times D_{40} + \gamma_{80} D_{Lost} \times D_{80} + \sum_j \delta_j + \epsilon_i \end{aligned}$$

where Y_i is an indicator for whether the compensation is deemed fair, D_{Won} and D_{Lost} are indicators for whether they won or lost the lottery before answering the questions and D_{Before} an indicator for being asked the question before the lottery, D_0, D_{40} and D_{80} are indicators for whether the question inquired about no compensation, 40 or 80 Kwacha, δ_j refers to strata or interviewer fixed effects. The standard errors ϵ_i are Eicker-Huber-White standard errors.[30]

As shown in Table D.1 I detect little if any differences with respect to whether they lost or won the lottery. This is true for both the compensation levels of no compensation and 40 Kwacha ($\alpha_i = \beta_i = \gamma_i$), as well as the differences between the compensation levels ($\alpha_i - \alpha_j = \beta_i - \beta_j = \gamma_i - \gamma_j$). I only find a statistically significant difference for the compensation level of 80 Kwacha between whether questions are asked before or after the lottery, which is not robust to the definition of Y_i. Note however that standard errors are quite large so this is not a finding of no effect. I further find no difference with respect to whether they are asked about no compensation or 40 Kwacha ($\alpha_0 = \alpha_{40} = \beta_0 = \beta_{40} = \gamma_0 = \gamma_{40}$), however a difference when asked about 80 Kwacha which is considered to be more fair ($\alpha_{80} - \alpha_j = \beta_{80} - \beta_j = \gamma_{80} - \gamma_j = 0$).

[30] I used Stata's (Version 16.0) *reg* command with the *robust* specification for standard errors.

Table D.1: Full specification of survey experiment

	Compensation is fair vs not fair		Compensation is fair vs unfair	
	(1)	(2)	(3)	(4)
	b/se	b/se	b/se	b/se
α_0	0.59	0.33	0.70	0.38
	(0.12)	(0.12)	(0.12)	(0.11)
α_{40}	0.61	0.37	0.73	0.44
	(0.12)	(0.12)	(0.11)	(0.11)
α_{80}	0.95	0.72	1.01	0.71
	(0.10)	(0.11)	(0.10)	(0.10)
γ_0	0.53	0.26	0.64	0.33
	(0.13)	(0.13)	(0.13)	(0.12)
γ_{40}	0.51	0.23	0.62	0.28
	(0.13)	(0.13)	(0.12)	(0.11)
γ_{80}	0.80	0.57	0.89	0.62
	(0.13)	(0.13)	(0.12)	(0.12)
β_0	0.64	0.40	0.79	0.50
	(0.13)	(0.13)	(0.13)	(0.12)
β_{40}	0.58	0.32	0.74	0.42
	(0.13)	(0.14)	(0.13)	(0.12)
β_{80}	0.85	0.60	0.92	0.63
	(0.12)	(0.13)	(0.11)	(0.11)
F-tests:				
$\alpha_0 = \beta_0 = \gamma_0$	0.56	0.37	0.44	0.19
$\alpha_{40} = \beta_{40} = \gamma_{40}$	0.54	0.30	0.42	0.13
$\alpha_{40} - \alpha_0 = \beta_{40} - \beta_0 = \gamma_{40} - \gamma_0 = 0$	0.94	0.83	0.95	0.68
$\alpha_{80} = \beta_{80} = \gamma_{80}$	0.05	0.05	0.11	0.22
$\alpha_{80} - \alpha_0 = \beta_{80} - \beta_0 = \gamma_{80} - \gamma_0$	0.36	0.24	0.32	0.14
$\alpha_{80} - \alpha_{40} = \beta_{80} - \beta_{40} = \gamma_{80} - \gamma_{40}$	0.76	0.82	0.62	0.56
$\alpha_{80} - \alpha_0 = \beta_{80} - \beta_0 = \gamma_{80} - \gamma_0 = 0$	0.00	0.00	0.00	0.00
$\alpha_{80} - \alpha_{40} = \beta_{80} - \beta_{40} = \gamma_{80} - \gamma_{40} = 0$	0.00	0.00	0.00	0.00
Model description:				
Strata FE	✓	✓	✓	✓
Enumerator FE	✗	✓	✗	✓
Adj. R^2	0.72	0.74	0.77	0.81
Observations	490	490	448	448

Notes. Columns (1) and (2) refer to specifications with an outcome that is 1 if the compensation is deemed fair and 0 otherwise. Columns (3) and (4) refer to specifications with an outcome that is 1 if the compensation is deemed fair and 0 if it is deemed unfair, omitting respondents that deem it neither fair nor unfair. Specifications differ depending on whether interviewer fixed effects are included. α_X, β_X and γ_X are estimates for the mean conditional on a compensation level of X Kwacha when asked before the lottery, having lost or won the lottery respectively. Rows under *F-tests* display the corresponding p-values.

E. APPENDIX TO CHAPTER 5

E.A. Interventions in Detail

Consent Form

[*This is an example, wording depends on whether the potential respondent is an adult or not.*]

TO ENUMERATOR: Please let the parent (and child together) read the text on the next screen. If they are not able to read, please read the text to the parent (and child together).

If they have questions, answer them to your best knowledge or direct them to your supervisor.

Hello,
I am *[name]* conducting a survey for *[information of who the principal investigators are]*. We conduct a research study about *[topic of survey or research]*. We are interested in your opinions and general information about you, your family and your household. Your household was randomly selected for an interview. First we would like to ask you about your household and then interview your child about his/her life. The interview with you will take about 40 minutes to complete, the interview with your child will take about 60 minutes to complete.

[*Goal of the study*]

We would be glad if you would support our study with your participation in the interview. We do not expect any negative consequences for you or your family from this study.

[*Data protection*]

The study is for research purposes only. During the interviews personal data about you, your child and your family is collected and stored for several years until the completion of this study. All responses will be treated strictly confidential by the researchers. The data will only be used for this study. For the analysis of the survey all identifying information (such as names and identification numbers) will be replaced with numbers. We keep this information only in case we are interested in following-up with an interview in the future. Any results from this survey will only be reported in aggregate terms and no personal data will be revealed in any of our reports. Third parties and public institutions will not receive access to any personal information. Your name and your family member's names will not be passed on to anyone and will not be made public. All of your data will be deleted upon request.

[*Rights of the respondent)*]

Your and your child's participation in this research study is fully voluntary. If you choose to continue with the interview, you and your child can choose not to respond to any or

all of the questions we ask. You can withdraw your consent for participation in the study at any time, without the need to mention any reasons and without any negative consequences for you or your family. In case you withdraw your consent, all personal data which was collected will be erased. Let me assure you again that all the information provided by you will be kept strictly confidential.
If you want to withdraw your consent, get further information about the survey, or are interested in the results of the study, please contact the person listed on the business card.

Do you have any questions?

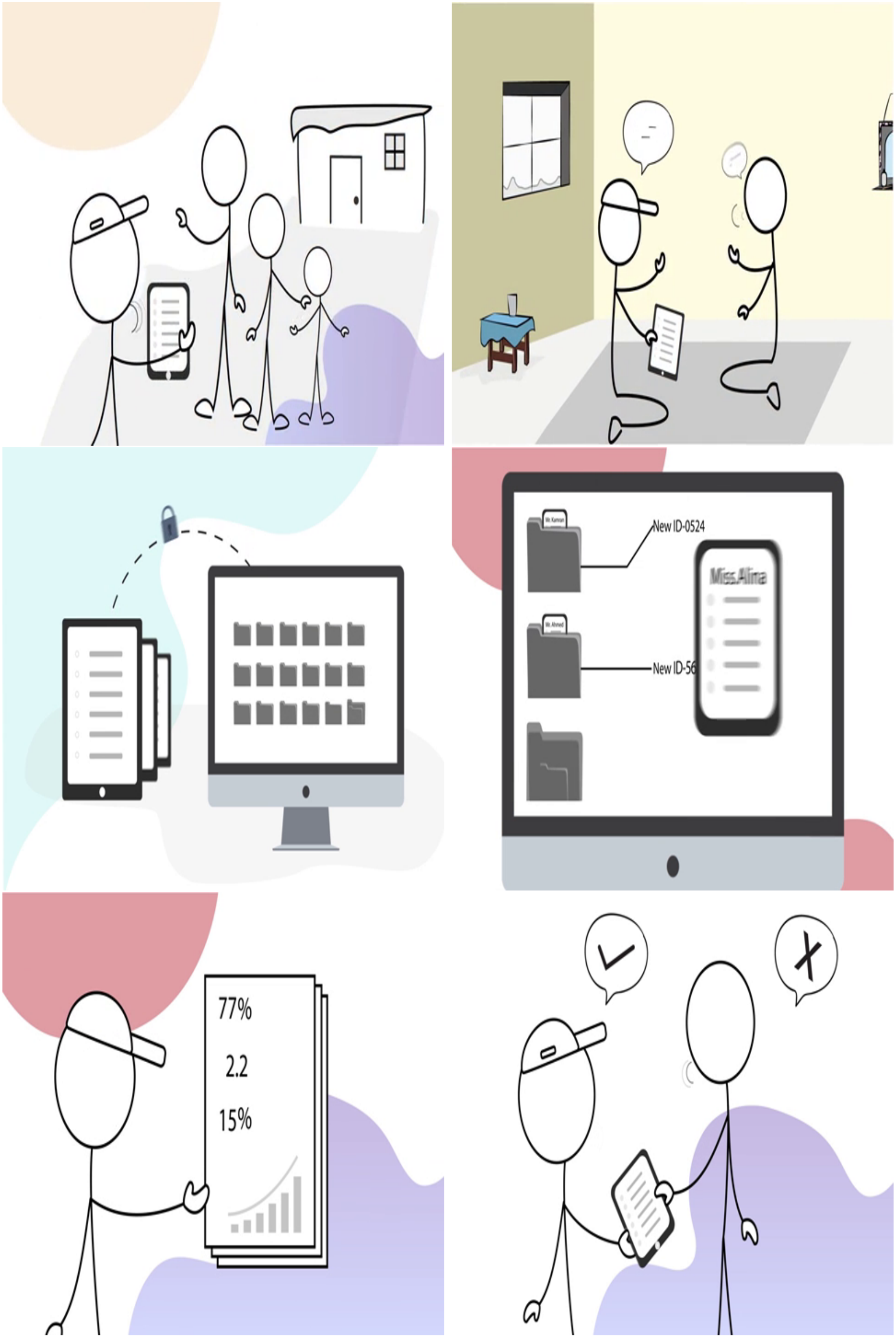
Video for ACF and ASCF

TO ENUMERATOR: Please show the respondent the video on the next screen.
Tell them they can pause or re-watch the video at any time.

Script

Hello! We are conducting a survey for *[information of who the principal investigators are]*. We conduct a research study about *[topic of survey or research]*. We are interested in your opinions and general information about you, your family and your household. We would like to ask you about your household and your life. The interview will take about 60 minutes to complete. We would be glad if you could support our study with your participation in the interview. We do not expect any negative consequences for you or your family from this study. During the interview personal data about you and your family is collected and stored for several years until the completion of the study All responses will be treated strictly confidential by the researchers. This means after the interview is done, the information is send to the data collection company. At the data collection company all information from all the interviews is collected. Then all identifying information (such as names and identification numbers) will be replaced with new numbers. We keep the personal information only in case we are interested in following up with an interview in future. It is stored for several years until the completion of the study. The rest of the information is used for research . The information from all the interviews is then analyzed and reported in aggregated terms, such that no personal data will be revealed in any of the reports. The aggregated information is then shared but third parties and public institutions will not receive access to any personal information. Your name will not be passed onto anyone and will not be made public. Your participation in this research study is fully voluntarily. If you choose to continue with the interview, you can choose not to respond to any or all the questions we ask. You can withdraw your consent for participation at this study anytime, without the need to mention any reasons and without any negative consequences for you and your family. If you want to withdraw your consent, get further information about the survey, or are interested in the result of the studies please tell the enumerator or contact the person listed on the business card. In case you withdraw your consent, all personal data which was collected will be erased.

Examples of screenshots



Scripted Interactive Part of ASCF

[*This is an example, wording depends on whether the potential respondent is an adult or not.*]

TO ENUMERATOR:
On the following screens, there will be either text or questions displayed. If there is text displayed, please read it to the respondent. If a question is displayed, please ask the respondent for an answer. DO NOT READ OUT THE CHOICES, but select all choices which reflect the respondent's answer. There are no right or wrong answers.

Hello,
I am *[name]* conducting a survey for *[information of who the principal investigators are]*. We conduct a research study about *[topic of survey or research]*. We are interested in your opinions and general information about you, your family and your household. Your household was randomly selected for an interview. We would like to ask you about your household and your life. The interview will take about 100 minutes to complete.

Who will be using the information you provide?
[multiple responses possible]

- Government
- NGO
- **Researchers**
- Private company
- Don't know
- Other, please specify

In case of incorrect response[31]: We are conducting the survey for researchers from the University of Mannheim in Germany.

We would be glad if you would support our study with your participation in the interview. We do not expect any negative consequences for you or your family from this study.

The study is for research purposes only. During the interview personal data about you and your family is collected and stored for several years until the completion of this study.

What will the information you provide be used for?
[multiple responses possible]

- Needs assessment (determining the eligibility to a program)
- **Research**
- Marketing
- Criminal prosecution
- Don't know
- Other, please specify

[31] Correct responses are **emphasized** in this illustration

In case of incorrect response: The study is for research purposes only.

All responses will be treated strictly confidential by the researchers. The data will only be used for this study. For the analysis of the survey all identifying information (such as names and identification numbers) will be replaced with numbers. We keep this information only in case we are interested in following-up with an interview in the future. Any results from this survey will only be reported in aggregate terms and no personal data will be revealed in any of our reports. Third parties and public institutions will not receive access to any personal information. Your name and your family member's names will not be passed on to anyone and will not be made public. All of your data will be deleted upon request.

Who will have access to your personal information?
[multiple responses possible]

- **Enumerators**
- Government/Public institutions
- **Researchers**
- NGOs
- **Data collection company**
- Other private companies
- Don't know
- Other, please specify

In case of incorrect response: Only the researchers and the data collection company have access to your personal information. Third parties and public institutions will not receive access to any personal information. Your name and your family member's names will not be passed on to anyone and will not be made public.

Your participation in this research study is fully voluntary. If you choose to continue with the interview, you can choose not to respond to any or all of the questions we ask. You can withdraw your consent for participation in the study at any time, without the need to mention any reasons and without any negative consequences for you or your family. In case you withdraw your consent, all personal data which was collected will be erased. Let me assure you again that all the information provided by you will be kept strictly confidential.

If you want to withdraw your consent, get further information about the survey, or are interested in the results of the study, please contact the person listed on the business card.

What happens if you give consent?
[multiple responses possible]

- **I will be interviewed**
- I will receive money and / or compensation for the interview
- **My (and my child's) responses will be send to the researchers**

- A NGO will help me, my family, or my community
- **My (and my child's) information will be saved for a potential new interview**
- Don't know
- Other, please specify

What happens if you do not give consent?
[multiple responses possible]

- I will be declined services in the future
- I will lose existing benefits
- **The interview stops immediately**
- Someone will punish me
- Don't know
- Other, please specify

In case of incorrect response: Your participation in this research study is fully voluntary. If you choose to continue with the interview, you can choose not to respond to any or all of the questions we ask.

In case you withdraw your consent, all personal data which was collected will be erased.

TO ENUMERATOR: On the next screen, the whole text is displayed in case the respondent wants to read it for themselves. Please ask them if they have any further questions.

E.B. Questionnaire Modules

Objective Measure of Understanding of Consent

In the following you can find the questions used to assess the understanding of the respondent. Correct answers are emphasized in bold font. Questions A1 and A6 relate to the rights of the respondent, question A2 relates to purpose of the study and benefits from participating, questions A3 and A4 relate to the voluntary nature of the respondent, and question A5 relates to confidentiality.

Please indicate whether the statements are true or false.

A1: I can call/tell the data collection team/researchers to have my information deleted.

- **True**
- False
- Don't know

A2: I am interviewed to assess my needs and determine whether I am eligible for a beneficial program.

- True
- **False**
- Don't know

A3: I have to participate in the study

- True
- **False**
- Don't know

A4: When I give consent, I have to respond to all of the questions.

- True
- **False**
- Don't know

A5: Only the researchers and data collection team will know the responses I gave.

- **True**
- False
- Don't know

A6: I can complain about the way the data collection team and researchers handled my data.

- **True**
- False
- Don't know

Alternative version of objective measure

Since we detected a tendency of respondents to assess statements as *true*, regardless of whether they were, we plan to introduce an alternative version of each statement for which the opposite assessment is correct in the second part of the data collection. This allows us to take this default response behavior into account to improve our analysis.

Please indicate whether the statements are true or false.

B1: Once I provided any information, I cannot tell the researchers or data collection team to delete the information.

- True
- **False**
- Don't know

B2: The answers I provide in this interview do NOT affect whether I am eligible for a beneficial program.

- **True**
- False
- Don't know

B3: The participation in the study is fully voluntarily.

- **True**
- False
- Don't know

B4: Even after I gave consent, I can choose not to respond to specific questions.

- **True**
- False
- Don't know

B5: My responses, together with my name and other identifying information, will be shared with third parties.

- True
- **False**
- Don't know

B6: I CANNOT complain about the way the data collection team and researchers handled my data.

- True
- **False**
- Don't know

Subjective Measure of Understanding of Consent

We only focus on three aspects which seem to be the most relevant in this context.

In the following, you will be presented three aspects related to this survey.
Please indicate how well you understood each of these aspects. There are no right or wrong answers.

C1: The purpose of this study.

- I didn't understand this at all
- I didn't understand much of it
- I understood this to some extend
- I understood this well
- I understood this fully

C2: My participation in the interview being fully voluntary.

- I didn't understand this at all
- I didn't understand much of it
- I understood this to some extend
- I understood this well
- I understood this fully

C3: How the confidentiality of my information is ensured.

- I didn't understand this at all
- I didn't understand much of it
- I understood this to some extend
- I understood this well
- I understood this fully

C4: My rights with respect to data protection and storage.

- I didn't understand this at all
- I didn't understand much of it
- I understood this to some extend
- I understood this well
- I understood this fully

Lebenslauf

CURRICULUM VITAE

Ph.D. in Economics	Sep 2015 - Jul 2021
Center for Doctoral Studies in Economics, University of Mannheim	*(expected)*
M.Sc. in Economics	Aug 2018
University of Mannheim	
Visiting student	Aug 2016 - Mai 2017
University of California, Berkeley	
Research Master student in Economics	Sep 2014 - Jul 2015
Tilburg University	
B.Sc. in Economics	Oct 2011 - Jul 2014
University of Bonn	

Mannheim, 12.08.2021 Matthias Stelter