RUHR
ECONOMIC PAPERS

Nathan Fiala
Julian Rose
Filder Aryemo
Jörg Peters

The (Very) Long-Run Impacts of Cash Grants during a Crisis

rwi #961

Imprint

Ruhr Economic Papers

Published by

RWI – Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #961

Responsible Editor: Manuel Frondel

ISSN 1864-4872 (online) – ISBN 978-3-96973-125-3

Ruhr Economic Papers #961

Nathan Fiala, Julian Rose, Filder Aryemo, and Jörg Peters

The (Very) Long-Run Impacts of Cash Grants during a Crisis

Bibliografische Informationen der Deutschen Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at **http://dnb.dnb.de**

RWI is funded by the Federal Government and the federal state of North Rhine-Westphalia.

http://dx.doi.org/10.4419/96973125
ISSN 1864-4872 (online)
ISBN 978-3-96973-125-3

Nathan Fiala, Julian Rose, Filder Aryemo, and Jörg Peters[1]

The (Very) Long-Run Impacts of Cash Grants during a Crisis

Abstract

The economic consequences of COVID-19 lockdowns were significant for poor households in the Global South. In this crisis period, we investigate the very long-run impacts of a randomized cash grant in Uganda on three pre-specified outcomes, including a heterogeneity analysis by gender. In 2008, the program supported young adults through a one-time grant of 380 USD, labelled to invest in vocational training and tools to start a business. The program revealed considerable effects after four years, which vanished after nine years. We now find, 12 years after the intervention, during the COVID-19 pandemic, positive effects on income for the full sample, which are entirely driven by men. Treated men are also significantly more likely to be engaged in an income generating activity, though this does not translate into higher food security. We find no effects for women. Our findings of re-surfacing positive effects are important for the growing literature on long-run impacts of programming as we show that the timing of a follow-up matters. The presence of economic shocks should especially be taken into account when planning long-run follow-ups.

JEL-Codes: C93, J24, O12, H53, I38

Keywords: Cash transfers; long-run impacts; randomized controlled trials

August 2022

1 Nathan Fiala, University of Connecticut, USA, and RWI; Julian Rose; RWI and University of Passau; Filder Aryemo, Gaplink Uganda; Jörg Peters, RWI and University of Passau. *- Data collection was funded by the IZA Institute of Labor Economics (Grant Agreement GA-5-698), the German Federal Ministry of Education and Research (BMBF, funding code 011A1807A DECADE), and RWI – Leibniz Institute for Economic Research. We are grateful for valuable comments and suggestions by Christopher Blattman and participants at the G²LM|LIC COVID-19 Research Meeting, the 2022 German Development Economics Conference, 5th IZA/FCDO G²LM-LIC ONLINE Research Conference as well as seminars at IFPRI, UC Irvine, and UConn. The analysis underlying this paper is preregistered at: www.socialscienceregistry.org/trials/6158.* *– All correspondence to: Nathan Fiala, RWI, e-mail: nathan.fiala@uconn.edu*

1. Introduction

Cash transfers have become a widely used tool to combat poverty, for example shaped as one-time entrepreneurial grants such as in-kind equipment or training subsidies. The underlying theory is that the poor are trapped in poverty because, in spite of high returns to capital, they are subject to market imperfections preventing them from profitable investments. Existing short-term evaluations of entrepreneurial grants document mixed results (Banerjee et al. 2015; Blattman et al. 2014; Brudevold-Newman et al. 2018; De Mel et al. 2008; De Mel et al. 2012b; Fafchamps et al. 2014; Fiala 2018; Hussam et al. 2022), recent long-term evaluations (ten and eleven years) show that the poor can sustainably escape poverty (Banerjee et al. 2021; Balboni et al. 2022). In contrast, Blattman et al. (2020) find that substantial effects after four years observed in Blattman et al. (2014) do not seem to sustain nine years after the intervention.

The present paper examines the same intervention as in Blattman, Fiala, and Martinez (2014; 2020, henceforth BFM), the Youth Opportunities Program (YOP), a one-time cash grant in Uganda, now 12 years later and during the early months of the COVID-19 crisis. The economic consequences of the pandemic and measures to contain the spread of the virus have been devastating in the Global South (Egger et al. 2021). The economic situation in Uganda was dreadful too: according to our data over 50 percent of our study population had to reduce food portions or skipped a meal at least once in the past seven days.

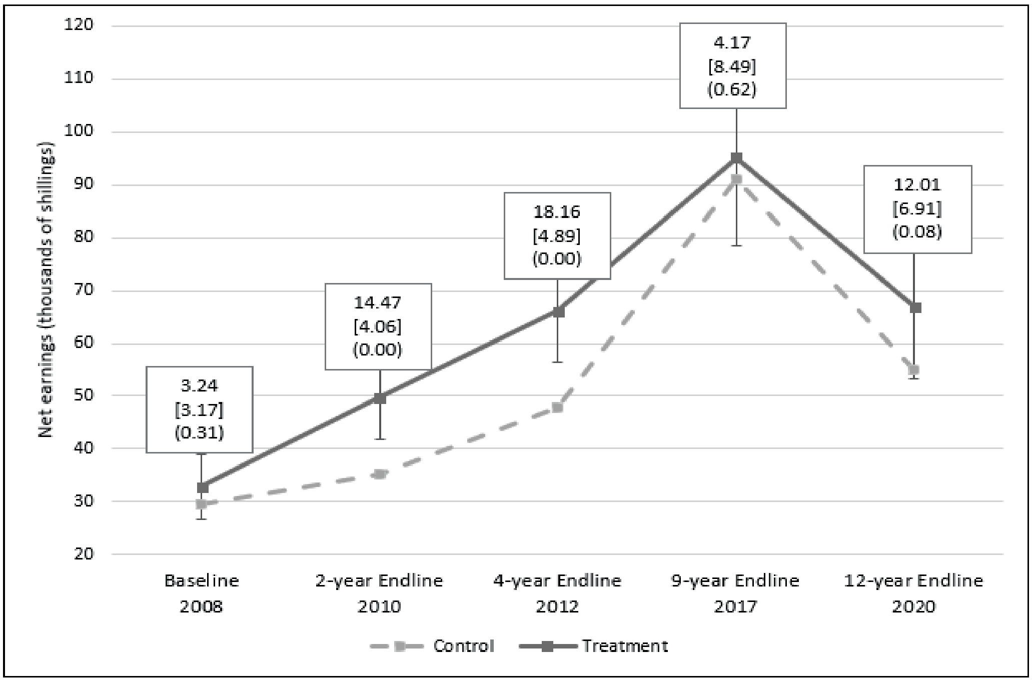
In 2007/2008, the Ugandan government implemented YOP to help poor and unemployed adults finance vocational training and equipment to start a small business.[1] YOP invited young adults to apply for cash grants averaging $380 (in 2008 USD)[2] per person. While the cash grant was labelled as a business creation subsidy, there was no enforcement after the disbursement. Individuals applied in groups with a proposal and grants were paid to the groups. Of 535 eligible groups, 265 were randomly assigned to treatment and 270 to control.

[1] The World Bank provided funding for the program, but the Ugandan government was responsible for implementing the program.

[2] The market exchange rate in 2008 was 1,720 shillings to $1, and the purchasing power parity exchange rate was 862 shillings to $1. The vast majority of recipients, 80 percent, received between 200 and 600 USD.

We pre-specified three primary outcomes before data collection in July 2020 (see Table A5 for detailed description of all outcomes)[3]: *employment*, defined as whether the responded worked for remuneration in the past seven days[4]; total *income* in the last four weeks; and *food security*, defined as whether the household reduced number of meals or portion size in the past seven days. We find that the treatment group reports a 20 percent higher income. Figure 1 illustrates the income development across time for the treatment and control group. It underlines that the COVID-19 led to an income decrease for both groups, yet the treatment group experienced a less serve income shock. Both groups are equally likely to be employed. Moreover, a pre-specified heterogeneity analysis by gender reveals that the income effect is entirely driven by men who are also significantly more likely to be employed and report a 24 percent higher income than men in the control group. The income effects, though, do not translate into higher food security.

Figure 1: Progression of earnings across time



Notes: We plot the mean value of reported monthly income for the control group and the sum of the control mean and the ITT estimate of YOPs impact. Income is in thousands of 2008 Ugandan shillings using the 2008 exchange rate of 1,720 shillings to $1. Results from 2008, 2010, and 2012 are published in BFM (2014); 2017 in BFM (2020); and 2020 in this paper.

[3] Additionally, Table A9 provides the exact questions and coding of the outcomes. The pre-analysis can be accessed here: https://www.socialscienceregistry.org/trials/6158

[4] For employment and income, we consider daily labor, working for wages or in-kind, and self-employment including agricultural businesses. Individuals that have produced crops or animal products for sale in the past four weeks are considered as being employed. See Table A9 or the pre-analysis plan for more details.

Our paper contributes to the understanding of (semi or unconditional) cash transfer effects in the long run. According to Bouguen et al. (2019), the current prior in the small literature is that these interventions "initially help the poor to accumulate assets [and] evidence from the limited number of studies at hand is broadly consistent and indicates that these assets are generally gradually run down over time, generating little permanent impact on poverty". Persistent long-term effects only occur, as it is summarized in Bouguen (2019), if transfers are tied to very intensive support programs (as opposed to the relatively light and untied vocational training in YOP), like the multifaceted assistance in Bandiera et al. (2017) and Banerjee et al. (2021) where people received weekly consumption support and visits by trainers. Yet, our paper showcases that the *timing* of a long-term study matters. Furthermore, Bouguen et al. (2019) emphasize that the effect heterogeneity along gender lines is large in the few existing studies. We add another important observation by the considerable positive effects on men and the absence thereof among women.

We explore (i.e., not pre-specified) potential mechanisms leading to the positive effects on income and find evidence that both the quality and the quantity drive these results: like in the 4- and 9-year evaluations, men in the treatment group are substantially more engaged in skilled trades such as carpentry, tailoring, or metal fabrication. At the same time, our explorative analysis also shows that the treatment group, in particular men, report more working hours in the past four weeks. Hence, the higher income for men can stem from better jobs, more working hours, or a combination of the two.

Unlike the 4- and 9-years evaluations in BFM (2014, 2020), where a broad range of outcomes were elicited, this study focuses on outcomes that are arguably immediately affected by the crisis. Table A7 in the appendix offers a comprehensive overview of all outcomes in the 4-, 9-, and 12-years study including whether the outcome was pre-specified. BFM (2014) find substantial positive impacts on capital stock, income, consumption, and engagement in skilled trades after four years. After nine years, BFM (2020) observe that the control group has converged to the treatment group's income and consumption levels over time, but also that treatment group members have higher assets and are still systematically more engaged in skilled trades and work more – hence, the deeper structural changes we deem to be responsible for the income and employment effects during the COVID-19 crisis.

In response to the spread of the COVID-19 virus, the Government of Uganda gradually imposed a strict lockdown with an overnight curfew from the end of March until the end of May. The measures also included a ban on public transport and restrictions on private movement, closing of international borders, and a restriction on non-food activities. While the lockdown led to a standstill of economic activities in the non-food sectors, farming and food vendors were less affected (Mahmud and Riley 2021; Hartwig and Lakemann 2020). For Kampala, Hartwig and Lakemann (2020) document that 81 percent of businesses in their sample were closed during the lockdown. This is also reflected in a substantial drop in profits and income. Similarly, Mahmud and Riley (2021) report a 60 percent drop in non-farm income in rural areas of western Uganda. Their findings suggest that households shifted labor supply to agriculture during this period, though some were still able to earn incomes. Our qualitative interviews with YOP participants and community leaders offer similar evidence. Moreover, the responses underline that the lockdown was strictly enforced in our communities.[5]

From July to September 2020, we conducted phone- and in-person interviews. The data collection happened in two phases shortly after the government eased the strict measures in May 2020 (Figure 2 outlines all major events). Because the pandemic situation did not allow in-person interviews, we used phone surveys in the first phase, trying to contact all 2,598 YOP participants originally surveyed for BFM (2014) and BFM (2020) and successfully retrieved 1,242 of them. In the second phase, we implemented an in-person tracking of a random sample of those not reached via phone and additionally reached 414 YOP participants.[6] As the second round is a representative draw of those not reached in the first round, we have an *effective response* rate of 83.2 percent. Overall, our effective response rate is relatively high and performs well compared to many other long-term studies reviewed in Bouguen et al. (2019).[7] We nevertheless conduct thorough robustness checks which, even under conservative assumptions, confirm a non-zero effect.

[5] Transcript of qualitative interviews can be obtained from authors upon request.

[6] For the in-person interviews we adhered to all hygiene regulations in Uganda (e.g., enumerators wore face masks and used hand sanitizer).

[7] Orgill-Meyer et al. (2019) conduct a long-term follow-up from a clustered randomized controlled trial in India and replace attrited households with neighboring households. Given that the intervention took place at individual level instead of households, we could not implement this approach.

Our study also speaks to the literature on economic development during and after large shocks. De Mel et al. (2012a) randomly granted microenterprises cash and in-kind transfers after the 2003 Tsunami in Sri Lanka. They find that the additional capital increases the speed of recovery substantially. Bandiera et al. (2019) study the impact of a randomized women empowerment program implemented shortly before the Ebola outbreak in Sierra Leone and observe substantial positive effects during the epidemic. Similarly, Christensen et al. (2021) evaluate the impact of a randomized community monitoring program for health clinics implemented two years before the Ebola outbreak in Sierra Leone and show that the intervention successfully improved health clinics' performance before and during the epidemic. Casey et al. (2021) conduct an 11-years follow-up of a community driven development program in Sierra Leone finding suggestive evidence that the program improved communities' response to the Ebola epidemic in 2014.

Like many other studies conducted during the COVID-19 crisis we use phone calls to conduct the interviews. Concerns are widespread regarding data quality in such surveys, which, the argument goes, facilitate misreporting and errors. Moreover, the respondents' attention span is probably more limited and, hence, phone interviews must be shorter, which in turn might have implications for data accuracy. To address these concerns, we implemented a survey experiment to compare phone and in-person data collection. We randomly selected YOP participants for different questionnaires and varied whether the interview took place in-person or via phone. We thereby also contribute to a growing literature on phone surveys specifically (Arthi et al. 2018; Garlick et al. 2020; Heath et al. 2020) and how data is collected more generally (Kilic and Sohnesen 2019; Gaddis et al. 2019; Di Maio and Fiala 2020; Fiala and Masselus 2022).[8] The findings suggest that our main outcome variables are consistent across survey modes and, hence, we have no indication for bias in the phone survey responses. This is reassuring given that a larger part of our effective response rate sample size is coming from the phone survey.

[8] De Weerdt et al. (2020) provide a comprehensive overview on the role of survey methods.

2. Young Opportunities Program (YOP): Implementation, sampling, and randomization

The Ugandan government implemented YOP in Northern Uganda in 2007/2008 as part of the Northern Uganda Social Action Fund (NSUAF), a strategy to develop and stabilize Northern Uganda after an insurgency that afflicted the region in the early 2000s. According to estimates, two-thirds of the population in Northern Uganda could not meet basic needs in 2006 (Government of Uganda 2007). In response, YOP targeted poor, rural, and unemployed young adults to facilitate self-employment as craftsmen and -women. The government designed YOP as a community-driven development approach where groups could apply for cash grants to start a skilled trade.

YOP invited groups of young adults aged between 16 and 35 to apply for cash grants of up to $10,000 in total, labelled for vocational training and tools. The groups, consisting of 22 members on average, submitted a written proposal specifying how they plan to divide and spend the grant for vocational training, tools, and enterprise start-up costs. Although the group submitted the proposal together, members usually applied to set up an independent business. Typically, groups applied for one trade and selected their trainers, e.g., local artisans or small institutes. In most cases, group members came from the same village, and one application per village was submitted. Half of the groups existed already before the intervention as farm cooperatives or other clubs. Most groups are mixed, with only 5 percent of groups being all-female and 12 percent all-male. The main reason for using group applications was to ease the administrative processes and disbursements to a few hundred groups instead of thousands of individuals.

The proposal passed several governmental screening levels, from the village level up to the national NSUAF office, prioritized based on submission date and completeness. The successful groups then received the cash grant as a lump sum in a bank account. After that, no monitoring or enforcement mechanisms were implemented. On average, groups received $7,497 ($382 per individual). Due to group size and requested amounts, the individual amount varied, with 80 percent receiving between $200 and $600.[9] We include covariates capturing the total grant amount applied for by group and grant size per group member to account for grant size heterogeneity.

For the randomization, the program was oversubscribed such that 535 eligible groups were nominated in 14 districts in Northern Uganda. Then, 265 groups were randomly assigned to treatment and 270 to control, stratified by districts.[10] Shortly after the randomization, BFM conducted a baseline survey sampling five members per group resulting in a total sample of 2,677 (see Figure 2 for a detailed timeline of all events).[11] BFM (2014) provide a comprehensive discussion of baseline balance between treatment and control group (see Table 2 in BFM 2014). They show that the randomization led to an overall balanced treatment and control group; the differences that are statistically significant are small in size.[12]. At baseline (2008), YOP group members were on average 25 years old, mainly lived in rural areas, and can be characterized as poor, credit constrained, and underemployed.

BFM (2014) and BFM (2020) evaluate the effects of YOP after 4- and 9-years.[13] Table A6 summarizes the main findings of the previous evaluations and Table A7 offers an overview of all analyzed outcomes in the three studies including an indication on whether the outcome was pre-specified. BFM (2014) provide evidence for YOP's successful implementation and document substantial positive impacts after four years. The treatment group is 53 percentage points more likely to have enrolled in vocational training and received 340 more hours of vocational training. Moreover, the treatment group has a higher capital stock, more durable assets and income, is more engaged in skilled work, and consumes more.

Yet, after nine years, BFM (2020) document only minor sustained effects. The control group managed to find other kinds of work with similar levels of productivity and earnings. The control group's income and consumption converge towards the treatment group over time. The authors

---

[9] See Figure 1 in BFM (2014) which shows the distribution of group size and the average grant size per person.

[10] For more details on the experimental design, see BFM (2014).

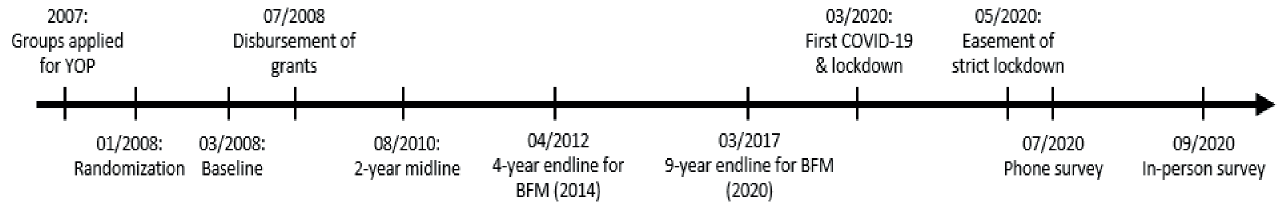
[11] In contrast to the sample used in BFM (2014, 2020), we dropped 79 individuals that were never reached in any data collection (not even at baseline) and our sample therefore consists of 2,598 individuals.

[12] The treatment reports in 2008 $3 more savings and 0.07 standard deviations greater initial wealth. We control for baseline wealth and savings in our analysis. Also, BFM (2014) provide a comprehensive discussion of the baseline balance and conduct sensitivity analyses addressing imbalance concerns.

[13] In Appendix A2 we provide a comprehensive summary of both papers.

still find positive impacts on durable assets, and the treatment group is systematically more engaged in skilled trades. The treatment group reported spending twice as much time in a skilled trade and was twice as likely to be working at least thirty hours per week in a skilled trade.

Figure 2: Timeline of YOP, surveys, and COVID-19 measures



Note: The distance between events does not represent the time that has passed.

3. Data

For our study, we re-visited the sample used in BFM (2014, 2020) and conducted another round of interviews by phone in July 2020 and in-person in September 2020. BFM conducted a baseline in 2008 and follow-ups in 2010, 2012, and 2017 as shown in Figure 2. Given the COVID-19 restrictions at the beginning of the data collection in Uganda, we implemented a hybrid data collection comprising phone surveys for the entire sample in the first phase and in-person tracking for a random subsample in the second phase.

The first phase of interviews took place from mid-July to mid-August 2020. We sought to contact 2,598 YOP participants. Table 1 summarizes the results of the tracking. In the first phase, we were able to successfully contact 47.8 percent of the sample (column 2) via phone. For the second in-person tracking phase, we randomly selected 44.5 percent of those not found during the first phase (column 3).[14] The randomization was stratified by treatment status and district, ensuring that the selected subsample represents the entire sample of those not found in the first phase. The second phase took place in September 2020, and we successfully reached 68.5 percent of those selected for the second phase (column 4). Given the random sampling for phase two, we give higher weight in the analysis to those reached in the second phase using inverse sampling weights and disregard those not selected for intensive tracking. Therefore, we end up with an effective

[14] The decision to track 44.5 percent intensively was based on budget constraints.

response rate of 83.2 percent (column 6). In the analysis, we account for the tracking strategy by weighting those found in phase 1 with unit weight, and those selected in phase 2 are weighted by the inverse of their selection probability. To remedy against a potential bias of attrition, we additionally weight individuals by the inverse of their predicted probability of attrition.[15] We conduct and discuss additional robustness tests for attrition bias in Appendix A3.

Selective attrition is a natural concern for a 12-years follow-up. Columns 7 and 8 of Table 1 show that attrition is slightly higher among controls, yet the weighted difference is not statistically significant (column 10). Additionally, Table A1 in the appendix investigates correlates of attrition for baseline characteristics. Several variables are significant, but the coefficients are very small and are not very strong predictors for attrition (columns 5 and 6). The strongest covariate is whether a YOP participant lives in an urban area at baseline, and it explains only 13 percent of attrition propensity. In columns (1) to (4) of Table A1, the same estimations are performed for the 4- and 9-years evaluations, respectively. The pattern looks quite similar, except that in the 12-years endline, the districts seem to be a strong predictor of attrition. To account for that finding, we use district fixed effects throughout the analysis. Moreover, Table A2 examines whether the 9-years outcomes predict attrition in 2020. The coefficients are very small and only a few are significant.

Table 1: Survey Response Rate

	Selection and tracking by survey phase					Effective response rate				
Survey	Total sought	Found phase 1 (%)	Selected phase 2 (%)	Found phase 2 (%)	Final # of obs.	All (%)	Control (%)	Treated (%)	Weighted difference* (%)	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
12-year endline	2,598	47.8	44.5	68.5	1,656	83.2	82.0	84.3	0.3	0.895

Notes: Column (1) includes only observations that were found at baseline. Columns (6)–(9) report the effective response rates overall, by treatment status, and the treatment-control difference (calculated via regression, controlling for baseline district). Columns (6)–(9) are weighted by the inverse probability of selection in phase 2 of the survey and are referred to as ''effective'' response rates. Unfound respondents randomly dropped in phase 2 receive zero weight. Column (10) reports p-values on the difference term, using robust standard errors clustered at the group level. * The weighted difference between attrited individuals from the control and treatment group is calculated using weights.

[15] The sampling and attrition weights are multiplied together such that retrieved members of the sample who are more similar to the attritors are weighted slightly more in the estimations. As robustness check, we run all regressions without these weights and find no major differences (see Table A11 and Table A13 in the appendix).

4. Results

4.1. Estimation Strategy

We are primarily interested in the effect of YOP on employment, income, and food security. We estimate a simple intent-to-treat effect (ITT) of the program impacts on outcome Y via the weighted least squares regression:

$$Y_{ij} = \beta_{ITT} T_{ij} + \delta X_i + \alpha_d + \varepsilon_{ij} \quad (1)$$

Where Y_{ij} denotes the outcome for individual i in group j. T_{ij} is a dummy variable equal to 1 if the individual was part of the treatment group; X_i is the set of baseline covariates; α_d are district fixed effects and ε_{ij} is an individual error term clustered by group.[16] In addition to baseline controls, we include controls for the timing of the survey and survey mode. We weight observations by their inverse probability of selection into endline tracking to correct for attrition. We also pre-specified treatment effects by gender as heterogeneity analysis.

Table A5 displays the pre-specified primary and secondary outcomes and the way we calculated each outcome. Additionally, Table A9 in the appendix includes the exact question in the survey. The primary outcome *Income* is measured in UGX and is the sum of all reported income in the past four weeks. Since income is a noisy measure and exhibits some outliers in the upper tail, partly caused by obvious enumerator errors, we top-code income at the 99th percentile as pre-specified in our PAP, affecting 15 observations in total. Since top-coding and at which level top-coding takes place influences the impacts, we present the results for different top-coding scenarios in Table A14.[17]

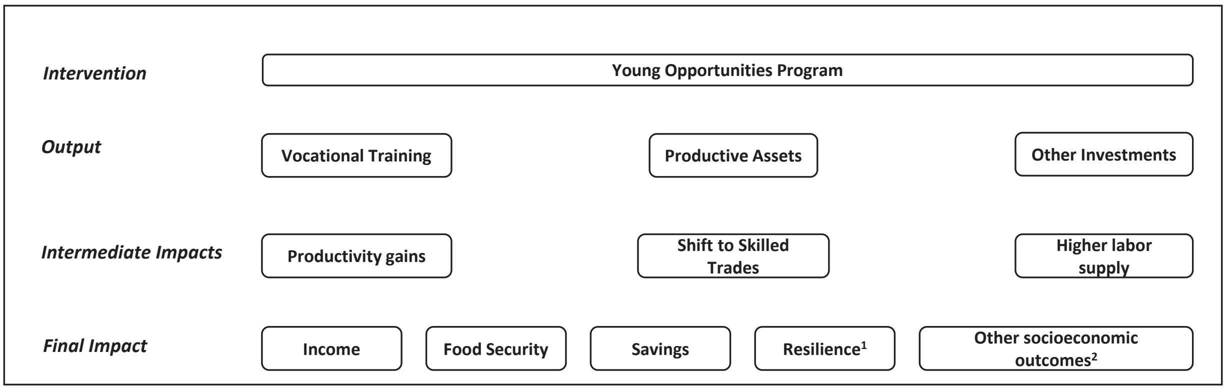
Figure 3 illustrates a simplified result chain for YOP. As documented in BFM (2014), YOP spurred investments into vocational training and productive assets, as it was intended by the program.

[16] We present results without baseline controls in the appendix. Coefficients are very similar, yet the significance level varies slightly given the higher precision with baseline controls.

[17] Top-coding can take place either before aggregating the various income sources to total income or after aggregating, which we did not pre-specify. In Table 2 we use income with top-coding after aggregating. In Table A14, we present the results for top-coding before aggregating and without any top-coding. Overall, the results are very similar across all scenarios. In one specification (no top-coding, all baseline controls) the results turn insignificant. Additionally, in Figure A1 we present the distribution of income for different top-coding scenarios in the treatment and control group.

These investments then translated into a shift to skilled trades and accompanying productivity gains for the treatment group. After four years, YOP impacted several socioeconomic outcomes while only a handful sustained to the 9-years evaluation. For the 12-years evaluation, we hypothesized in our pre-analysis plan that the COVID-19 crisis affected the labor markets (reflected in primary outcome *employment*) and subsequently downstream impacts such as income (primary outcome) and food security (primary outcome). Yet, we hypothesized that occupations are differently affected, and skilled trades are more resilient to the shock. Lastly, we hypothesized that the sustained 9-years impacts, namely assets and savings, led to a higher resilience to shocks in the treatment group, which is reflected in our secondary outcomes (subjective resilience, business resilience, farming resilience, safety net, savings, remittances).

Figure 3: Pathways from intervention to economic impacts



Notes: *We pre-specified several secondary outcomes capturing resilience: subjective resilience, business resilience, farming resilience, safety net, savings, remittances (see Table A9 for an overview). # BFM (2014 & 2020) document several additional socioeconomic outcomes, see Table A7 for a full list.

4.2. Primary Outcomes

Table 2 shows ITT effects for the primary outcomes.[18] The lower part of the table presents the only heterogeneous analyses as pre-specified in the PAP. The presented interaction terms can be directly interpreted as elasticities. We do not find a positive impact on *employed* for the full sample (column 1). The mean of 0.67 in the control group suggests that most of the sample pursued at least one income-generating activity in the past seven days, indicating that economic activities were possible. For reported income, we find a positive and statistically significant effect for the

[18] Table A10 presents the ITT effects without baseline controls.

treatment group. The effect is also economically meaningful, with individuals in the treatment group reporting on average a 20 percent higher income than those in the control group (column 2).

Figure 1 illustrates the development of monthly earnings over the entire 12-year period. While there is a clear increase in monthly earnings until the 9-years follow-up, the consequences of the pandemic are visible with a drop in income for control and treatment group. Yet, the decline is less strong for the treatment group resulting in a statistical different income after 12-years. Finally, we do not find a significant difference between the treatment and control groups for food security. The control mean of 4.6 indicates that food security is at risk since households skipped meals or reduced portions over four times in the past week. Overall, the findings suggest that economic activity had already resumed prior to our survey after the lockdown had ended, and the treatment group seems to be slightly more successful in recovering income.

The pre-specified heterogeneity analysis in the lower panel of Table 2 reveals relevant gender differences. We find that men in the treatment group are significantly more likely to be pursuing an income-generating activity and report a significantly higher income (columns 1 and 2, lower panel). The income effect is even more pronounced than for the full sample, with the income for men in the treatment group being 24 percent higher. Again, we do not find any impacts on food security, and for women, we do not see any significant difference between the treatment and control group.

Table 2: ITT effects for primary outcomes

	Employed (1)	Income (3)	Food Security (5)
Treatment	0.04	12.01*	-0.02
	(0.17)	(0.08)	(0.95)

District FE	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes
q-value	0.29	0.29	0.46
Control Mean	0.67	59.38	4.62
N	1466	1525	1524
R2	0.174	0.11	0.17
Treatment x Women	-0.05	2.52	-0.02
	(0.31)	(0.73)	(0.97)
Treatment x Men	0.08**	17.15*	-0.02
	(0.02)	(0.08)	(0.95)
Men	-0.08**	17.38*	-0.36
	(0.06)	(0.07)	(0.37)
District FE	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes
q-value Treatment x Women	0.767	0.87	1.00
q-value Treatment x Men	0.095	0.29	1.00
N	1466	1525	1524
R2	0.178	0.12	0.17

Notes: P-values in parentheses. Sampling weights are applied. Standard errors clustered at the group level. In all regressions we control for timing of interview and mode of interview (phone vs. person). To correct for multiple hypothesis testing, we calculate q-values using the Benjamini-Hochberg step-up method. The q-values indicate the smallest false discovery rate at which the null hypothesis of zero effect is rejected. Baseline controls included in columns (1), (3), and (5): Individual characteristics: Age, age squared, age cubed, male (only full sample), urban, risk aversion. Education: Highest grade, literate, vocational training, digit recall test score, ADL Index, distance to educational facilities. Wealth: Wealth Index, savings, monthly income, could borrow $58, could borrow $580. Occupation: Weekly hours in low skill/business/agriculture, in school. Intervention: Grant amount applied for, group size, grant amount per member, group existed before application, group age in years, within-group heterogeneity, group dynamic, group committee member, chair or vice-chair. The q-values are FDR sharpened q-values controlling for multiple hypothesis testing. * implies p < .1, ** implies p < .05, *** implies p < .01.

One potential explanation why the treatment group's income recovers faster, at least for men, might be the type of occupations. Recall from BFM (2020) that the treatment group was substantially more engaged in skilled work, even after nine years. To investigate further whether individuals engaged in skilled work are better off we use the subsample interviewed with the long questionnaire.[19] In contrast to the short questionnaire, we collected detailed information for 35 activities in the long questionnaire instead of summarizing them into five broad categories. In Figure 4 and Figure 5 we compare the share of individuals engaged in skilled trades and the average weekly hours in skilled trades between treatment and control group. Consistent with the previous studies, we define skilled trades as tailoring, weaving, metal fabrication, blacksmith, carpentry, construction work, and running a saloon. Figure 4 shows that both men and women in the treatment group are more engaged in skilled trades compared to the control group. Yet, Figure 4 also underlines a constant decline in engagement in skilled trades since the 4-year follow-

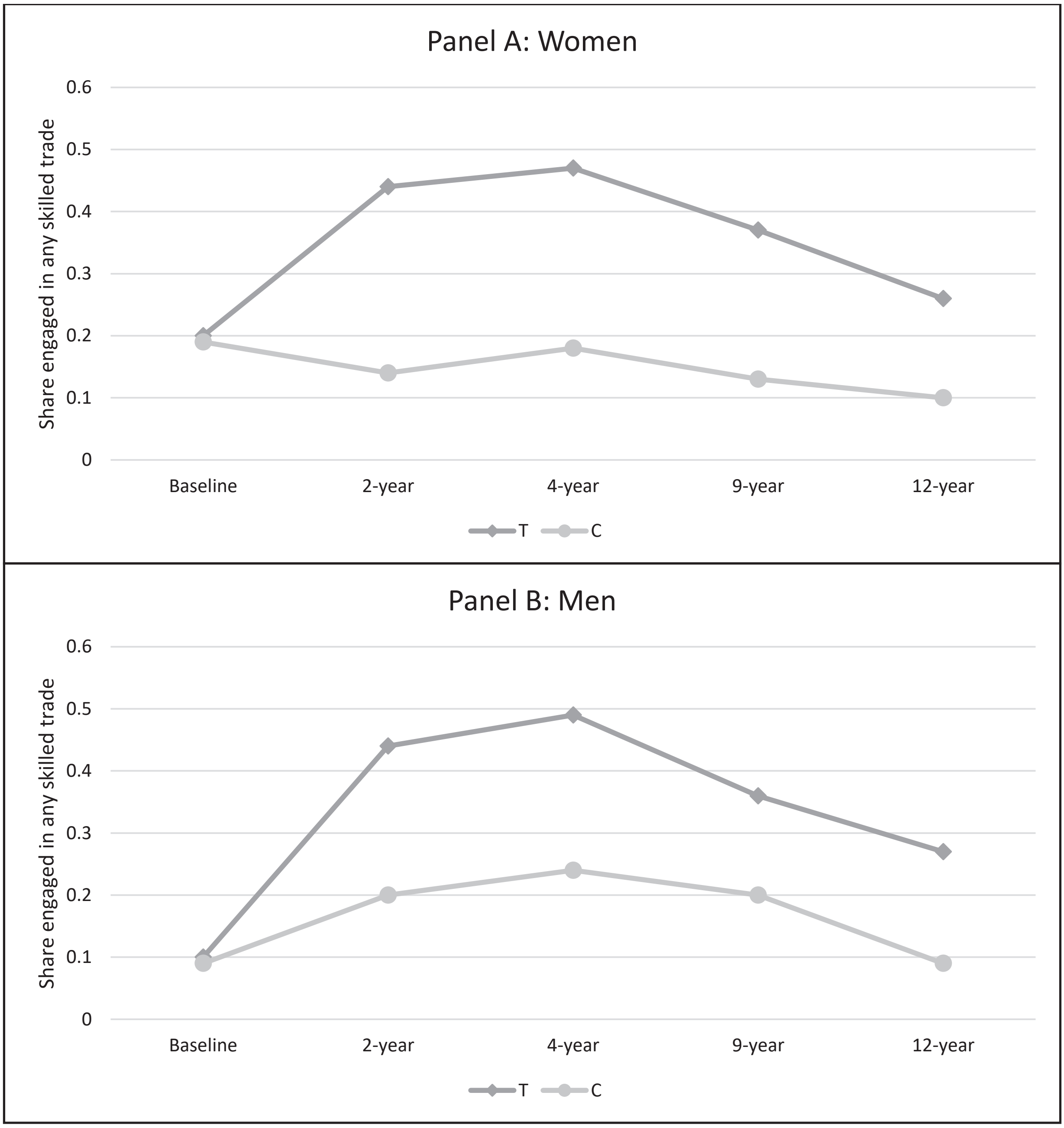
[19] Note that this explorative analysis is not pre-specified.

up. Figure 5 documents that the treatment group is also spending more hours per week on skilled trades. Particularly men in the treatment group report substantially more hours than men in the control group. Figure 5 also demonstrates a strong shift to agriculture, especially women work substantially more hours in agriculture than in the 9-year follow-up. Unfortunately, we have this detailed data only for a small subsample of 194 observations, and it should hence be interpreted with caution. However, statements in the qualitative interviews are in line with these observations.

Another potential mechanism leading to the higher income of treated men is that the treatment group might be working more. We collected data on the number of days and hours worked in the past four weeks. In Table 3, we investigate whether the treatment group works more. We do not find significant differences for working days and hours in the past four weeks between treatment and control group. For one specification we find some evidence that men work more hours. Again, though, the heterogeneity analysis shows that men work more. Our main finding of higher incomes among treated men might, hence, be driven by better jobs or merely by more working hours. Likewise, it is also possible that the quality of work affects the supplied quantity, or vice versa.

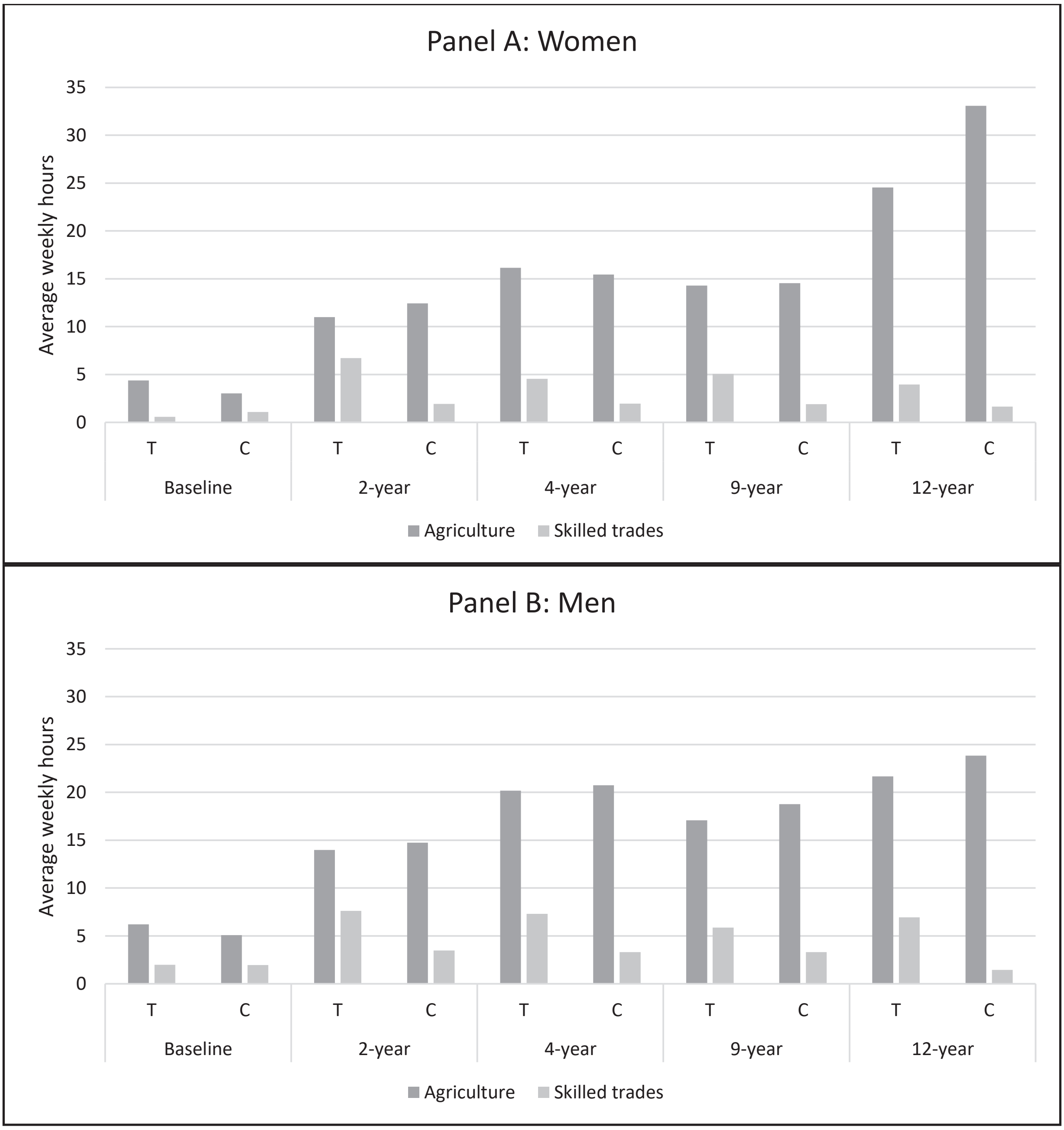
It is also possible that the treatment group received more governmental aid during the crisis. Overall, our data indicates that support from the government and NGOs was very little with only 13 percent reporting receiving any kind of support. We also see no significant differences between treatment and control group. Therefore, we conclude that a difference in governmental support is not driving our results.

Figure 4: Share of respondents engaged in any skilled trade



Notes: The data for baseline, 2-year, and 4-year is from BFM (2014); 9-year from BFM (2020); 12-year is from this data collection. The 12-year sample consists only of 194 individuals from Groups 2 & 3. Skilled trades are tailoring, weaving, metal fabrication, blacksmith, carpentry, construction work, and running a saloon. Sampling weights applied.

Figure 5: Average weekly working hours in agriculture and skilled trades



Note: The data for baseline, 2-year, and 4-year is from BFM (2014); 9-year from BFM (2020); 12-year is from this data collection. The 12-year sample consists only of 194 individuals from Groups 2 & 3. Skilled trades are tailoring, weaving, metal fabrication, blacksmith, carpentry, construction work, and running a saloon. Agriculture includes subsistence and commercial farming. Sampling weights are applied.

Table 3: ITT effects on labor supply

	(1) Total Days	(2) Total Days	(3) Total Hours	(3) Total Hours
Treatment	0.57	0.58	6.25	6.87
	(0.56)	(0.53)	(0.36)	(0.30)
District FE	Yes	Yes	Yes	Yes
Baseline Controls	No	Yes	No	Yes
Control Mean	27.92	27.92	93.29	93.29
N	1525	1525	1525	1525
R2	0.2	0.24	0.09	0.13
Treatment x Women	-1.53	-1.72	-7.72	-6.11
	(0.40)	(0.33)	(0.48)	(0.57)
Treatment x Men	1.71	1.83*	12.92	13.90*
	(0.11)	(0.08)	(0.11)	(0.08)
Men	-1.71	-3.07*	-0.8	-4.16
	(0.30)	(0.09)	(0.94)	(0.70)
District FE	Yes	Yes	Yes	Yes
Baseline Controls	No	Yes	No	Yes
N	1525	1525	1525	1525
R2	0.20	0.24	0.09	0.13

Notes: P-values in parentheses. Sampling weights are applied. Standard errors clustered at the group level. In all regressions we control for timing of interview and mode of interview (phone vs. person). To correct for multiple hypothesis testing, we calculate q-values using the Benjamini-Hochberg step-up method. The q-values indicate the smallest false discovery rate at which the null hypothesis of zero effect is rejected. Baseline controls included in columns (1), (3), and (5): Individual characteristics: Age, age squared, age cubed, male (only full sample), urban, risk aversion. Education: Highest grade, literate, vocational training, digit recall test score, ADL Index, distance to educational facilities. Wealth: Wealth Index, savings, monthly income, could borrow $58, could borrow $580. Occupation: Weekly hours in low skill/business/agriculture, in school. Intervention: Grant amount applied for, group size, grant amount per member, group existed before application, group age in years, within-group heterogeneity, group dynamic, group committee member, chair or vice-chair. The q-values are FDR sharpened q-values controlling for multiple hypothesis testing. * implies p < .1, ** implies p < .05, *** implies p < .01.

4.3. Secondary Outcomes

In addition to the primary outcomes, we defined and pre-registered several secondary outcomes. Table A5 displays the pre-specified secondary outcomes and Table A9 provides the exact questions. Table 4 presents the ITT effects for different measurements for resilience (column 1-3), economic well-being (column 4), safety nets (column 5), savings (column 6), and remittances (columns 7 and 8). The ITT effects in Table 4 suggest minimal effects of YOP on the secondary outcomes. In terms of resilience, we do not find that businesses or farming activities in the treatment and control group are differently affected (columns 2 and 3). Yet, we find that the treatment group is subjectively more resilient in that they are more confident in coming up with UGX 100,000 in seven days (column 1).[20] The lower panel of column 1 suggests that this effect is

[20] UGX 100,000 correspond to $28 which is around 1/20th of the GNI per capita.

driven entirely by men. In line with this finding, effects in column 6 suggest that the treatment group accumulated significantly more savings than the control, and again in particular, men are driving this effect. Having a high amount of savings can lead to more certainty in coming up with UGX 100,000.

Figure A2 depicts the evolution of savings over time for the treatment and control groups. The figure suggests that both groups accumulated substantial savings over time, yet the treatment group's savings decreased slightly from 2017 to 2020, whereas the control group still reports more savings than in 2017. For remittances, we do not detect any meaningful transfers and no difference between the groups. A potential concern for our primary findings is that the treatment group received more aid during the crisis since they might be better connected to governmental programs due to their experience with YOP. Yet, findings on the safety net in column 5 suggest that the treatment group is not more likely to have received any support from the government or any NGO than the control group.

4.4. Addressing challenges of phone surveys

This follow-up study was initiated right after the pandemic started, and, by definition, we had to deal with a dynamic and unclear public health situation. Our response to this was to administer the first wave of interviews via phone surveys in July 2020. Obtaining a meaningful response rate with phone interviews is challenging for several reasons, in particular since the intervention was twelve years ago and the most recent data collection was three years ago at the time of our survey. Phone numbers are often outdated or respondents can be more suspicious on the phone and refuse to participate. In some cases, in spite of multiple attempts, phones were also switched off. To address these challenges, we followed the best practices outlined by J-PAL (see Kopper and Sautmann 2020). Furthermore, we exploited YOP's group structure by asking group members to provide contact details of other group members that are part of our sample. We also compensated each respondent with airtime worth UGX 3,500 (1 USD).

Table 4: ITT Effects for Secondary Outcomes

	(1) Subjective Resilience	(2) Business Resilience	(3) Farming Resilience	(4) Economic Wellbeing	(5) Safety Net	(6) Total savings	(7) Remittances sent	(8) Remittances received
Assigned to treatment	0.180***	0.069	-0.001	0.021	-0.021	132,102*	29,968	-2,379
	(0.01)	(0.35)	(0.75)	(0.51)	(0.49)	(0.04)	(0.49)	(0.86)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
q-value	0.022		1.00	1.00		0.207	1.00	1.00
Control Mean	1.241	0.330	0.335	1.452	0.340	448,185	187,149	98,958
N	1524	483	1524	1524	1466	1525	588	365
R2	0.162	0.228	0.105	0.125	0.108	0.101	0.125	0.257
Treatment x Women	0.057	0.169	-0.004	-0.045	-0.037	23,483	105,531	-4,371
	(0.56)	(0.16)	(0.53)	(0.35)	(0.5)	(0.81)	(0.30)	(0.87)
Treatment x Men	0.247***	-0.006	-0.000	0.057	-0.012	190,944**	-1,295	-1,220
	(0.00)	(0.94)	(0.36)	(0.0.16)	(0.72)	(0.01)	(0.97)	(0.94)
Male	0.149	0.005	-0.040	-0.009	-0.049	104,708	52,550	8,911
	(0.16)	(0.96)	(0.14)	(0.87)	(0.34)	(0.15)	(0.26)	(0.68)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
q-value Treatment x Women	1.00		1.00	1.00		1.00	1.00	1.00
q-value Treatment x Men	0.03		1.00	1.00		0.154	1.00	1.00
N	1524	483	1524	1524	1466	1525	588	365
R2	0.164	0.230	0.105	0.127	0.11	0.102	0.128	0.257

Notes: P-values in parentheses. Sampling weights are applied. Standard errors clustered at the group level. In all regressions we control for timing of interview and mode of interview (phone vs. person). To correct for multiple hypothesis testing, we calculate q-values using the Benjamini-Hochberg step-up method. The q-values indicate the smallest false discovery rate at which the null hypothesis of zero effect is rejected. Baseline controls included in columns (1), (3), and (5): Individual characteristics: Age, age squared, age cubed, male (only full sample), urban, risk aversion. Education: Highest grade, literate, vocational training, digit recall test score, ADL Index, distance to educational facilities. Wealth: Wealth Index, savings,monthly income, could borrow $58, could borrow $580. Occupation: Weekly hours in low skill/business/agriculture, in school. Intervention: Grant amount applied for, group size, grant amount per member, group existed before application, group age in years, within-group heterogeneity, group dynamic, group committee member, chair or vice-chair. The q-values are FDR sharpened q-values controlling for multiple hypothesis testing. * implies p < .1, ** implies p < .05, *** implies p < .01.

In spite of these efforts, our first response rate was at 47.8 percent and we, therefore, conducted a second wave of interviews among those we could not reach via phone, now in-person, once the public health situation allowed. Since integrating phone- and in-person data requires that the survey mode does not systematically affect responses, we implemented a survey experiment in the second phase to compare phone and in-person responses to test for this consistency. Figure 6 depicts how our surveys were sequenced, including the survey experiment. In the first phase, we sought to interview the entire sample via phone with a short questionnaire focusing on employment, income, and coping strategies during the COVID-19 pandemic. Of those we did not reach in the first phase, we randomly selected 603 for the second phase of in-person tracking.[21]

For the experiment, we split the in-person interview sample into two groups and used two different questionnaires: a short questionnaire, identical to the one used in the phone survey phase, and a long questionnaire containing a more detailed income section.[22] The long questionnaire elicits employment and income outcomes for over thirty activities separately. In contrast, the short questionnaire contains only five broad activity categories, requiring respondents to add up numbers for working hours, days, and income. We used the long questionnaire for a randomly selected group (n=100) out of those not reached in the first phase, henceforth group *Long Questionnaire Group*. Another group (henceforth *Short Questionnaire Group*) was also randomly selected for in-person tracking (n=503) from those not reached in the first phase and interviewed with the short questionnaire.

Moreover, we took a random sample of those reached in the first phase and conducted another two interviews with them (henceforth *Phone Group*). First, we re-interviewed *Phone Group* in the first week of September again with the short questionnaire via phone (so, precisely as in the first phase). We need this to correct for potential seasonality effects between the phone and the in-person group. Second, we re-interviewed the same group of people one week later in person with the long questionnaire (*Phone Long Group*).

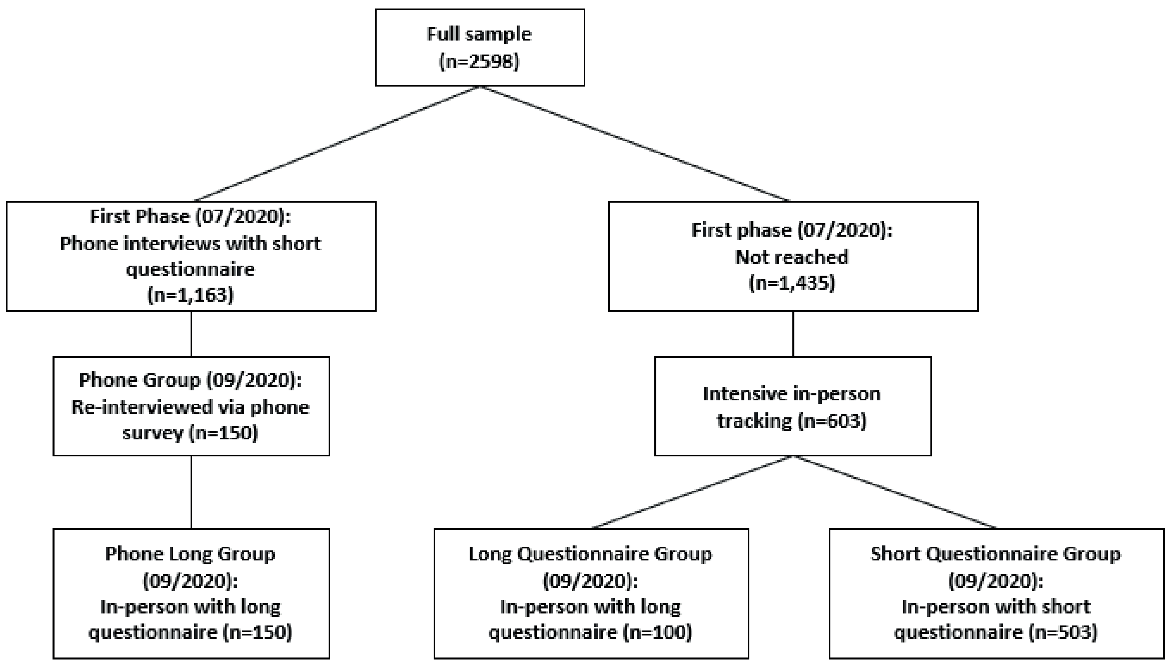
The random allocation of the short and long questionnaires allows comparing the responses between the groups directly. We can, therefore, causally evaluate whether the questionnaire length influences labor market outcomes in our sample. Yet, the comparison between *Phone* vs. *Short Questionnaire*, which shows the effect of phone vs. in-person, rests on the assumption that there is no self-selection into participating in the first phase. Hence, we must assume that respondents reached in the first phase have similar characteristics to those not reached.

[21] Due to budget restrictions, we were not able to conduct intensive in-person tracking for the entire sample.

[22] The income section in the long questionnaire was taken from the 2017 data collection.

Table A8 in the appendix compares baseline characteristics of *Phone* versus *Long Questionnaire* and *Short Questionnaire* and finds some statistically significant differences. It seems that individuals in *Phone* are more likely to consist of men employed in the non-agricultural sector. We include the full set of baseline controls and use district fixed effects in our analysis to control for these differences.

Figure 6: Design of Survey Experiment



4.5. Results of the survey experiment

To test whether the data collection method affects our results, we compare data from phone vs. in-person and short questionnaire vs. long questionnaire. The latter is randomized through the survey experiment, allowing us to test for a systematic bias. We interpret the in-person long questionnaire interviews as the ground-truthed benchmark. Yet, for the comparison between phone and in-person data, we assume that there is no selection bias.

Table 5 presents the difference of in-person interviews with short and long questionnaire, hence, between the groups *Short Questionnaire Group* and *Long Questionnaire Group*. The results suggest that reported income is higher in the short questionnaire, yet the difference is not statistically significant.[23] For food security, we find a substantial and significant difference. This finding is surprising since there was no difference in the questions on food security in the short and long questionnaires. The only difference between the groups is the lengthier income section for *Long Questionnaire Group* leading to a longer overall interview time. In columns 3 and 4, we find that days/hours worked are substantially smaller in the short questionnaire.

Table 5: Survey Experiment: short questionnaire versus long questionnaire

	(1) Income	(2) Food Security	(3) Total Days	(4) Total Hours
Long-Questionnaire	52,180	1.01***	-13.29***	-102.15***
	(0.22)	(0.01)	(0.00)	(0.00)
District FE	Yes	Yes	Yes	Yes
Control Mean	66,784	3.07	40.13	191.78
N	366	366	366	366
R2	0.06	0.28	0.35	0.22

Notes: P-values in parentheses. Standard errors clustered at the group level. Due to a coding error in the long questionnaire, we do not have the employment outcome for this group. Control variable for timing of survey included. * implies p < .1, ** implies p < .05, *** implies p < .01.

Table 6: Survey Experiment: phone versus in-person

	(1) Employed	(2) Income	(3) Food Security	(4) Total Days	(5) Total Hours
In-Person	-0.09	-9,760	0.19	2.20	6.01
	(0.22)	(0.80)	(0.82)	(0.40)	(0.75)
District FE	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Control Mean	0.79	186,437	3.09	25.80	98.55
N	435	435	359	435	435
R2	0.28	0.20	0.37	0.35	0.22

Notes: P-values in parentheses. Standard errors clustered at the group level. Control variable for timing of survey included. Baseline controls included: Individual characteristics: Age, age squared, age cubed, male, urban, risk aversion. Education: Highest grade, literate, vocational training, digit recall test score, ADL Index, distance to educational facilities. Wealth: Wealth Index, savings, monthly income, could borrow $58, could borrow $580. Occupation: Weekly hours in low skill/business/agriculture, in school. Intervention: Grant amount applied for, group size, grant amount per member, group existed before application, group age in years, within-group heterogeneity, group dynamic, group committee member, chair or vice-chair. * implies p < .1, ** implies p < .05, *** implies p < .01.

Table 6 presents the results for comparing phone and in-person outcomes using the groups *Phone Group* and *Short Questionnaire Group*. To reduce concerns of self-selection, we include the full set of baseline controls and district fixed effects. We do not detect any significant differences between

[23] To ensure that our findings for income are not driven by this difference, we estimate the effect excluding all individuals interviewed with the long-questionnaire and find similar effects. This analysis is available upon request.

phone and in-person surveys for our primary outcomes (columns 1 to 3). There are also no differences regarding total days and hours worked. The findings of our survey experiment indicates that rather than the survey mode, the level of detail of the questionnaire plays a role. Overall, the results of the survey experiment increase our confidence in pooling phone- and in-person data.

5. Long-term RCTs: challenges and opportunities

Our paper is part of a nascent but rapidly growing literature that evaluates the long-term effects of randomized interventions. We reconcile our results with some dimensions put forward in Bouguen et al. (2019), a review of long-term evaluations of randomized interventions. Our results emphasize that the timing of a long-run evaluation matters. This is particularly interesting in our case because the positive 4-year YOP effects had largely vanished after nine years. Yet, our findings suggest that the deeper structural effects on the YOP treatment group that were also evidenced after nine years – on assets, occupational choice, and labor supply – apparently led to an increased resilience that only materializes with respect to income and employment in situations of economic decline, namely the COVID-19 crisis.

Bouguen et al. (2019) also raise methodological concerns related to attrition and statistical power, which are potentially aggravated by the long-term character of these studies. While tracking our respondents was challenging during lock-down and via phone, we managed to keep effective attrition rates at about 16 percent, which performs well vis-à-vis many other long-term studies reviewed in Bouguen et al. (2019). Robustness checks suggest that significance levels turn borderline, which is no surprise given that standard errors have increased considerably compared to the 4-year follow-up (although standard errors are lower than for the 9-years follow-up). Appendix A3 offers a comprehensive discussion of the attrition analysis. Yet, also when conservatively accounting for attrition the effects remain, mostly, above zero.

Moreover, Bouguen et al. (2019) voice concerns about a specific type of publication bias in long-term studies: If only interventions with very promising short-term or intermediary evaluations are followed up on in the very long run, this will lead to a misleading picture. This could be aggravated in case statistical power is lower for the long-term follow up than for the short- or mid-term evaluation, leading to an Ioannidis et al. (2017) type power-of-bias problem. Statistical power might decrease for several reasons: attrition is one, but also standard errors for outcome variables could increase, for example because different subgroups of the study population are exposed to a changing environment to varying degrees. The standard error in our population indeed increased between the 4- and 9-years follow-up, but it has decreased again in the 12-years follow-up. In terms of this type of publication bias, we believe our study is important because we decided to return despite a null effect for most indicators after nine years. Our findings thereby also confirm the concern raised by Bouguen et al. (2019) that systematic reviews of (very) long-term evaluations should be careful in assuming that null effects in the short term or after a longer period also imply null effects in the (very) long run.

Another potential caveat is that while RCTs ensure balance at baseline, treatment and control group might be differently influenced by subsequent policy changes, leading to an invalid counterfactual. For a welfare-to-work program in Canada, Riddell and Riddell (2020) showcase that a policy change introduced after the treatment changed the counterfactual, leading to a biased policy conclusion. We have no indication from our repeated data collections for something similar in our sample. In support of this, we collected data on whether the respondent has received support during the COVID crisis and members of both groups are equally likely having received support.

6. Conclusion

This paper has investigated the very long-term effects of a one-time entrepreneurial cash grant program that helped young adults become self-employed artisans twelve years prior to data collection. We conducted phone and in-person interviews from July to September 2020, hence, shortly after the rigorous lockdowns implemented by the Ugandan government to contain the spread of COVID-19. Previous evaluations found substantial effects after four years (BFM 2014), which vanished after nine years because the control group had caught up (BFM 2020). Our key finding is that the treatment group reports significantly higher incomes and employment again after twelve years. The effect is driven by men; for women we do not find this effect. Food security, though, does not differ, neither for men nor for women. We transparently address challenges that are typical for long-term studies, most notably attrition. While we managed to keep the attrition rate low, our robustness checks turn the results insignificant in very conservative scenarios. Yet, even then the patterns support a non-zero effect. Furthermore, we exploratively investigate mechanisms and find evidence for both quality and quantity of work driving the results. The treatment group is substantially more likely to be in skilled trades than the control group, and treated men also report more working hours.

Our result of a re-surfaced effect after twelve years is important as it suggests that the deeper structural changes induced by the treatment materialize in income and employment again in times of a crisis, by making treatment group members more resilient. It generally also emphasizes the pertinence of the *timing* of a long-term follow-up: future long-term studies might particularly examine how effects develop after economic shocks, also for example induced by natural disasters. Indeed, our papers calls for more research on the long-term effects of RCTs because the heterogeneity of contexts and interventions we are anyway facing is certainly even more pronounced for long-term developments. More observations are needed to conclude which effects sustain – under which circumstances.

References

Arthi V., Beegle K., De Weerdt J., & Palacios-López A. (2018). Not your average job: Measuring farm labor in Tanzania. *Journal of Development Economics. 130*: 160–72.

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I. & Sulaiman, M. (2017). Labor markets and poverty in village economies. *The Quarterly Journal of Economics, 132*(2), 811-870.

Balboni, C., Bandiera, O., Burgess, R., Ghatak, M., & Heil, A. (2022). Why do people stay poor?. *The Quarterly Journal of Economics*, *137*(2), 785-844.

Bandiera, O., Buehren, N., Goldstein, M. P., Rasul, I., & Smurra, A. (2019). *The economic lives of young women in the time of Ebola: Lessons from an empowerment program*. The World Bank.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., ... & Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, *348*(6236), 1260799.

Banerjee, A., Duflo, E., & Sharma, G. (2021). Long-term effects of the targeting the ultra poor program. *American Economic Review: Insights*, *3*(4), 471-86.

Blattman, C., Fiala, N., & Martinez, S. (2014). Generating skilled self-employment in developing countries: Experimental evidence from Uganda. *The Quarterly Journal of Economics*, *129*(2), 697-752.

Blattman, C., Fiala, N., & Martinez, S. (2020). The long-term impacts of grants on poverty: Nine-year evidence from Uganda's Youth Opportunities Program. *American Economic Review: Insights*, *2*(3), 287-304.

Bouguen, A., Huang, Y., Kremer, M., & Miguel, E. (2019). Using randomized controlled trials to estimate long-run impacts in development economics. *Annual Review of Economics*, *11*, 523-561.

Brudevold-Newman, A. P., Honorati, M., Jakiela, P., & Ozier, O. W. (2017). A firm of one's own: Experimental evidence on credit constraints and occupational choice. *World Bank Policy Research Working Paper*, (7977).

Casey, K., Glennerster, R., Miguel, E., & Voors, M. J. (2021). *Long run effects of aid: Forecasts and evidence from Sierra Leone* (No. w29079). National Bureau of Economic Research.

Christensen, D., Dube, O., Haushofer, J., Siddiqi, B., & Voors, M. (2021). Building resilient health systems: Experimental evidence from Sierra Leone and the 2014 Ebola outbreak. *The Quarterly Journal of Economics, 136(2),* 1145-1198.

De Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment. *The Quarterly Journal of Economics*, *123*(4), 1329-1372.

De Mel, S., McKenzie, D., & Woodruff, C. (2012a). Enterprise recovery following natural disasters. *The Economic Journal*, *122*(559), 64-91.

De Mel, S., McKenzie, D., & Woodruff, C. (2012b). One-time transfers of cash or capital have long-lasting effects on microenterprises in Sri Lanka. *Science*, *335*(6071), 962-966.

De Weerdt, J., Gibson, J., & Beegle, K. (2020). What can we learn from experimenting with survey methods?. *Annual Review of Resource Economics*, *12*, 431-447.

Di Maio, M., & Fiala, N. (2020). Be wary of those who ask: A randomized experiment on the size and determinants of the enumerator effect. *The World Bank Economic Review*, *34*(3), 654-669.

Egger, D., Miguel, E., Warren, S. S., Shenoy, A., Collins, E., Karlan, D., ... & Vernot, C. (2021). Falling living standards during the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science Advances*, *7*(6), eabe0997.

Fafchamps, M., McKenzie, D., Quinn, S., & Woodruff, C. (2014). Microenterprise growth and the flypaper effect: Evidence from a randomized experiment in Ghana. *Journal of Development Economics*, *106*, 211-226.

Fiala, N. (2018). Returns to microcredit, cash grants and training for male and female microentrepreneurs in Uganda. *World Development*, *105*, 189-200.

Fiala, N., & Masselus, L. (2022). *Whom to ask? Testing respondent effects in household surveys*. (No. 935). Ruhr Economic Papers.

Gaddis, I., Siwatu, G. O., Palacios-Lopez, A., & Pieters, J. (2019). *Measuring farm labor: Survey experimental evidence from Ghana*. The World Bank.

Garlick, R., Orkin, K., & Quinn, S. (2020). Call me maybe: Experimental evidence on frequency and medium effects in microenterprise surveys. *The World Bank Economic Review*, *34*(2), 418-443.

Government of Uganda (2007). '*National peace, recovery and development plan for Northern Uganda: 2006–2009*'. Government of Uganda, Kampala.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., ... & Tatlow, H. (2021). A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature Human Behaviour*, 1-10.

Hartwig, R., & Lakemann, T. (2020). *When the going gets tough: Effects of the COVID-19 pandemic on informal entrepreneurs in Uganda*. German Institute for Global and Area Studies. Retrieved from: https://www.giga-hamburg.de/de/publikationen/22075331-when-going-gets-tough-effects-covid-19-pandemic-informal-entrepreneurs-uganda/.

Heath, R., Mansuri, G., Rijkers, B., Seitz, W., & Sharma, D. (2020). Measuring employment: Experimental evidence from urban Ghana. *The World Bank Economic Review*, *35*(3), 635-651.

Hussam, R., Rigol, N., & Roth, B. N. (2022). Targeting high ability entrepreneurs using community information: Mechanism design in the field. *American Economic Review, 112*(3), 861-98.

Ioannidis, J., Stanley, T. and Doucouliagos, C. (2017). The power of bias in economics research. *Economic Journal*, *127*(605), F236-F265.

Karlan, D., & Valdivia, M. (2011). Teaching entrepreneurship: Impact of business training on microfinance clients and institutions. *Review of Economics and Statistics*, *93*(2), 510-527.

Kilic, T., & Sohnesen, T. P. (2019). Same question but different answer: Experimental evidence on questionnaire design's impact on poverty measured by proxies. *Review of Income and Wealth*, *65*(1), 144-165.

Kopper, S., & Sautmann, A. (2020). Best practices for conducting phone surveys. *Abdul Latif Jameel Poverty Action Lab (J-PAL)*.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, *76*(3), 1071–1102.

Mahmud, M., & Riley, E. (2021). Household response to an extreme shock: Evidence on the immediate impact of the Covid-19 lockdown on economic outcomes and well-being in rural Uganda. *World Development*, *140*, 105318.

Museveni, Y. (2020). *Address on the Corona virus (COVID 19): Guidelines on avoiding the pandemic.* Retrieved from: https://www.yowerikmuseveni.com/address-corona-virus-covid-19-guidelines-avoiding-pandemic

Orgill-Meyer, J., Pattanayak, S. K., Chindarkar, N., Dickinson, K. L., Panda, U., Rai, S., ... & Jeuland, M. (2019). Long-term impact of a community-led sanitation campaign in India, 2005–2016. *Bulletin of the World Health Organization*, *97*(8), 523.

Riddell, C., & Riddell, W. C. (2020). Interpreting experimental evidence in the presence of postrandomization events: A reassessment of the Self-Sufficiency Project. *Journal of Labor Economics*, *38*(4), 873-914.

Appendix

A1. Covid-19 in Uganda

On 18 March 2020, the Ugandan government started implementing the first measure to combat the pandemic's spread even before the first confirmed case. The first measures included the ban of public gatherings and the closing of all educational facilities (Museveni 2020). In the subsequent weeks, the measures became more rigorous, including a ban of public transport and closing of borders until eventually the Ugandan government imposed a nationwide lockdown with a curfew from 7 pm to 6.30 am on 30 March 2020. Initially, the government planned the lockdown for 31 days, but it was not until late May 2020 before the government eased the measures.

The lockdown from March to May affected income-generating activities differently. Farming and food vendors were less affected by the lockdown; farming was possible throughout the strict lockdown without any restrictions. Non-food businesses, though, were severely affected by the complete shutdown of non-essential activities and weekly markets. Hartwig and Lakemann (2020) report for Kampala that 81 percent of businesses in their sample were closed from late March until the end of May. Our qualitative interviews confirm this finding in our sample. Yet, by the time of our interviews, 80 percent of the businesses were re-opened, and the recovery process had started. The complete shutdown of non-food businesses is also reflected in profits and income. Hartwig and Lakemann (2020) show that profits dropped substantially during the lockdown in Kampala. Findings by Mahmud and Riley (2021) show that the non-farm income of rural households in western Uganda dropped substantially (60 percent) during May 2020. Furthermore, they document a shift of labor supply to agriculture during the period of the strict lockdown. Our qualitative interviews with participants and community leaders are in line with those findings. During the strict lockdown from March to May, people focused on farming activities since other businesses were not allowed to operate. Additionally, the qualitative interviews underline that the lockdown was strictly enforced in the communities (transcripts are available upon request from the authors).

In response to businesses' distress, the government launched several programs to assist: (i) obligatory payments to the National Social Security Fund were put on hold; (ii) deferment of tax payments; (iii) economic stimulus package. However, the qualitative and quantitative evidence in our data suggests that these programs hardly reached our communities. Only 13 percent in our sample stated that they received support from the government or an NGO.

A2. Previous Evaluations: BFM (2014) & BFM (2020)

BFM (2014) provide evidence for the YOP's successful implementation and positive impacts after four years. Of the treated groups, 89 percent received the grant.[24] The program successfully increased investments into human capital (skills training) and capital stocks (tools and materials). Table A6 provides an overview of the 4- and 9-year findings. In the treatment group, 68 percent enrolled in vocational training between 2008 and 2010, compared to 15 percent in the control group. Accordingly, the treatment group received on average 340 more hours of vocational training. Moreover, treatment successfully increased capital stocks. After four years, the treatment group reported 57 percent higher capital stocks. BFM (2014) also document a shift in occupation toward skilled work and increased total labor supply. Table A6 shows that total hours worked a week increased by 17 percent relative to the control group. This increase is mainly in skilled trades leading to greater participation and hours in a skilled trade that is 2.5 times greater than in the control group. Lastly, BFM (2014) report effects on income and consumption. Income in the treatment group increased by 39 percent relative to controls after four years. For durable and nondurable consumption, they find 0.18 standard deviations larger consumption for both over the control group. Overall, YOP successfully set the treatment group on a growth trajectory after four years, translating into higher income and consumption levels and more durable assets.

After nine years, BFM (2020) finds overall only minor sustained effects on income. The authors explain the vanishing of effects by the fact that the control group found other kinds of work with similar levels of productivity and earnings. Over time, the control group's income converges towards the treatment group. After nine years, the impact on income is just 4.6 percent of the control mean and not statistically significant. For non-durable consumption, the effect is only 1.4 percent of the control mean, also statistically insignificant. Meanwhile, the nine-year impact on

24 BFM (2014) note two reasons for groups not receiving the grant: (i) 21 groups could not access the funds due to unsatisfactory proposals, bank complications, or collection delays; (ii) 8 groups reported having never received funds due to theft. Baseline characteristics are generally very similar between receivers and non-receivers.

durable assets is significant, with 0.145 standard deviations greater durable assets in the treatment group. Yet, YOP did have lasting effects on occupational choices. The treatment group spent twice as much time in a skilled trade and was twice as likely to be working at least thirty hours per week in skilled trades.

A3. Attrition Analysis

While an effective response rate of 83.2 percent is high for a data collection in the middle of a pandemic and twelve years after baseline, attrition is still a potential threat to the results. Attrition only poses a problem when it is a common effect of treatment and outcome; hence, attrition is non-random. To probe into this, we first compare attrition rates in the treatment and control groups and find no statistically significant difference (see Table 1). Next, we estimate correlates of attrition using baseline data to examine whether baseline characteristics predict attrition. Table A1 shows the result for the 4-, 9-, and 12-year follow-ups. We find a consistent pattern across the waves, with a few variables turning up significant but very small coefficients. One exception is that individuals living in urban areas are significantly more likely to attrit, yet the effect decreases from the 4- to the 12-year follow-up. The districts at baseline are the strongest predictors of attrition in the 12-year follow-up. In Table A2 we conduct the same analysis using 9-year outcomes instead of baseline characteristics. We do not find any strong predictors of attrition, suggesting that the outcomes in the 9-year follow-up are not correlated to attrition in the 12-year follow-up.

We proceed with conducting two robustness checks. First, to gauge the sensitivity of results, we now manually impute missing outcome values instead of dropping attrited observations (see Karlan and Valdivia 2011). We use different scenarios, all conservative, in which we assume lower values for attrited treated individuals than the mean in the treatment group and higher values for attrited control observations than the mean in the control group. We calculate scenarios for +/- 0.1 standard deviations and +/- 0.25 standard deviations.

Table A3 presents the results for re-estimating the primary outcomes' ITT effect. Column (1) shows the results for +/- 0.1 standard deviations. For *employment,* the effect for men turns insignificant while in this scenario women in the treatment group show even a significant negative effect. The more extreme assumption in column (2) amplifies this trend, also turning the effect for the full sample significantly negative. For *income*, the first scenario turns the income effects for the full sample and men insignificant, yet the effects are still very largely positive. In column (2) the overall income effect turns significantly negative for women. The effect for men is still positive but under this assumption strongly insignificant. Lastly, the effect on food security remains insignificant in column (1) but turns significantly negative in column (2).

Second, Table A4 in we provide Lee bounds as suggested by Lee (2009). Lee bounds are derived by trimming the sample such that the share of observation with observed outcome is equal for both groups. It then provides a lower and upper bound for the treatment effect corresponding to the most extreme assumptions about the missing information. Panel A of Table A4 presents the lower and upper bounds for the outcome *employment*. Similar to our ITT estimates, for the full sample and women we do not find any effect. Yet, column (2) supports the finding for men. The upper bound is statistically significant and positive while the lower bound is just slightly negative. Hence, even under the most extreme scenario, the treatment effect becomes virtual zero only. For *income* in Panel B we find that the data is very noisy and extreme assumptions about the missing sample lead to inconsistent results. Overall, the attrition analysis shows that our results are sensitive to attrition, yet only to rather extreme scenarios.

Table A1: Correlates of survey attrition

	4-year endline		9-year endline		12-year endline	
	Coeff	SE	Coeff	SE	Coeff	SE
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned to treatment	-0.03	0.02	0.04	0.02	0.02	0.02
Age at baseline	-0.00	0.04	0.01	0.04	-0.07	0.04
Age squared	-0.00	0.00	-0.00	0.00	0.00	0.00
Age cubed	0.00	0.00	0.00	0.00	-0.00	0.00
Male	0.05	0.02**	-0.01	0.03	-0.04	0.03*
Large town / urban area	0.18	0.04***	0.12	0.03***	0.13	0.03***
Risk Aversion (z-score)	0.02	0.01**	0.04	0.01***	-0.01	0.01
Found at baseline	-0.70	0.06***	0.16	0.06**	0.00	0.00
Highest grade reached in school	-0.00	0.00	0.00	0.00	0.00	0.00
Able to read and write minimally	0.02	0.03	-0.04	0.02	0.02	0.03
Received prior vocational training	-0.04	0.04	-0.10	0.03***	0.00	0.04
Digit recall test score	0.02	0.01***	-0.00	0.01	0.00	0.01
ADL index	-0.00	0.00	-0.00	0.00	-0.01	0.00***
Durable Assets (z-score)	-0.01	0.01	-0.01	0.01	-0.03	0.01**
Savings (000s 2008 UGX)	0.00	0.00***	0.00	0.00	0.00	0.00
Monthly gross earnings (000s 2008 UGX)	-0.00	0.00	0.00	0.00	-0.00	0.00
Could obtain 100,000 UGX (58 USD) loan	0.01	0.02	0.03	0.02	-0.01	0.02
Could obtain 1,000,000 UGX (580 USD) loan	0.01	0.04	-0.06	0.03**	0.00	0.03
Weekly work hours: low skill	0.00	0.00	0.00	0.00	0.01	0.00**
Weekly work hours: other business	-0.00	0.00	-0.00	0.00	-0.00	0.00
Weekly work hours: skilled trade	0.00	0.00	-0.00	0.00	0.00	0.00
Weekly work hours: high skilled trade	-0.01	0.01	-0.00	0.01	0.00	0.01
Weekly work hours: other non-agricultural	-0.00	0.00	-0.00	0.00	0.00	0.00
Weekly work hours: agricultural	-0.00	0.00	0.00	0.00	-0.00	0.00
Weekly household chores, hours	0.00	0.00	0.00	0.00	0.00	0.00
Zero employment hours in past month	-0.02	0.03	0.06	0.03**	0.02	0.03
Main occupation is non-agricultural	0.03	0.05	0.09	0.04**	0.00	0.04
Engaged in a Skilled Trade	-0.03	0.05	0.03	0.05	-0.02	0.05
Currently in School	-0.04	0.05	-0.06	0.05	0.03	0.05
Grant amount applied for (USD)	0.00	0.00	-0.00	0.00	0.00	0.00**
Group Size	-0.00	0.00	-0.00	0.00	-0.01	0.00***
Grant Amount per Member, USD	-0.00	0.00	-0.00	0.00	-0.00	0.00***
Group existed before application	-0.01	0.02	-0.01	0.03	0.05	0.03*
Group age, in years	-0.00	0.01	-0.00	0.01	-0.00	0.01
Within-group heterogeneity (z-score)	0.03	0.01**	-0.00	0.01	0.03	0.01**
Quality of in-group dynamic (z-score)	-0.02	0.02	-0.01	0.01	-0.01	0.01
Management committee member	-0.02	0.02	-0.01	0.03	-0.02	0.03
Chairperson or vice-chairperson	0.02	0.04	-0.05	0.03	-0.03	0.04
Distance to educational facilities (km)	0.00	0.00	0.00	0.00	0.00	0.00
Lives in Adjumani	-0.09	0.09	-0.06	0.06	0.35	0.09***
Lives in Apac	-0.07	0.08	0.07	0.06	0.15	0.05***
Lives in Arua	-0.03	0.08	0.16	0.07**	0.16	0.06***
Lives in Kaberamaido	0.03	0.10	0.03	0.08	0.14	0.07**
Lives in Kotido	0.09	0.10	0.06	0.08	-0.11	0.05**
Lives in Kumi	-0.07	0.08	-0.01	0.06	0.11	0.06*
Lives in Lira	-0.09	0.08	0.10	0.08	0.07	0.06
Lives in Moroto	0.14	0.10	0.06	0.08	-0.02	0.06
Lives in Moyo	-0.16	0.08**	0.12	0.09	0.52	0.08***
Lives in Nakapiripirit	0.01	0.10	0.11	0.08	-0.08	0.05
Lives in Nebbi	-0.10	0.09	-0.00	0.06	0.05	0.05
Lives in Pallisa	-0.16	0.08**	0.04	0.06	0.10	0.05*

Lives in Soroti	-0.05	0.09	0.06	0.07	0.09	0.06
Mean	0.18		0.12		0.17	
P-value of F-test	0.00		0.00		0.00	
N	2,111.00		2,086.00		1,846.00	
R-squared	0.27		0.11		0.17	

Notes: Each pair of columns report the results from a WLS regression of an attrition indicator on baseline covariates and district fixed effects. Standard errors are clustered at the group level. * implies p < .1, ** implies p < .05, *** implies p < .01. Observations are weighted by the probability into selection of endline tracking, and errors are clustered by group.

Table A2: Correlates of survey attrition using 9-year outcomes (full sample)

	Attrited in 2020
Employment Outcomes	
Average employment hrs/wk	-0.000
	(0.96)
Agricultural hrs/wk	-0.000
	(0.96)
Non-agricultural hrs/wk	0.000
	(.)
Casual labor, low skill hrs/wk	0.001
	(0.50)
Petty business, low skill hrs/wk	-0.000
	(0.93)
Skilled Trades hrs/wk	-0.002
	(0.40)
High-skill wage labor hrs/wk	-0.002
	(0.25)
No employment hours in past month	0.000
	(.)
Main occupation is non-agricultural	0.011
	(0.77)
Engaged in any skilled trade	-0.021
	(0.53)
Works over 30 hrs/wk in skilled trade	0.124
	(0.23)
Average hours of chores per week	0.001
	(0.86)
Income and Consumption	
Average earnings/hr (000s of 2008 UGX)	0.033**
	(0.05)
Standardized Income Index	0.000
	(.)
Monthly net earnings (000s of 2008 UGX)	-0.000
	(0.18)
Nondurable Consumption (000s of 2008 UGX)	0.000
	(0.80)
Durable assets	-0.024
	(0.12)
Politics	
Index of political action (z-score)	-0.050
	(0.37)
Attended voter education meeting	0.021
	(0.55)
Discussed Vote	0.047
	(0.22)
Reported campaign malpractice or incident	-0.005
	(0.93)
Voted in presidential election	0.000
	(.)
Attended political rally	-0.004
	(0.93)
Participated in political primary	-0.024
	(0.54)
Worked to get a candidate/party elected	0.062
	(0.47)
Member of a political party	-0.086
	(0.23)
Index of NRM/Presidential support (z-score)	0.093

	(0.18)
Would vote for NRM if election were tomorrow	-0.066
	(0.37)
Like or strongly like NRM	-0.053
	(0.45)
Worked to get the NRM elected	-0.061
	(0.51)
Member of the NRM	0.088
	(0.23)
Voted or supported the president in the last election	0.000
	(.)
Index of opposition support (z-score)	0.008
	(0.90)
Would vote for opposition if election were tomorrow	0.004
	(0.97)
Like or strongly like any opposition party	0.014
	(0.81)
Worked to get the opposition elected	0.000
	(.)
Member of an opposition party	-0.071
	(0.45)
Voted or supported an election party in the past election	-0.006
	(0.93)
Observations	1928

Notes: P-values in parentheses. Results from a WLS regression of an attrition indicator for attrited in 2020 on 9-year outcome and district fixed effects. Includes all observations found in 9-year follow-up. Standard errors are clustered at the group level. * implies p < .1, ** implies p < .05, *** implies p < .01. Observations are weighted by the probability into selection of endline tracking, and errors are clustered by group.

Table A3: Imputing missing dependent variable with mean variations for primary outcomes

	(1) SD +/- 0.1	(2) SD +/- 0.25
	Panel A: Employment	
Treatment	-0.01	-0.06***
	(0.58)	(0.01)
Treatment x Female	-0.08***	-0.14***
	(0.01)	(0.00)
Treatment x Male	0.03	-0.01
	(0.22)	(0.71)
Male	-0.09***	-0.1***
	(0.00)	(0.00)
District FE	Yes	Yes
Baseline Controls	Yes	Yes
Control Mean	0.68	0.70
N	1846	1846
R2	0.10	0.11
	Panel B: Income	
Treatment	5.34	-4.35
	(0.31)	(0.42)
Treatment x Female	-4.76	-15.78***
	(0.39)	(0.01)
Treatment x Male	11.21	2.30
	(0.15)	(0.77)
Male	11.80*	10.50
	(0.09)	(0.14)
District FE	Yes	Yes
Baseline Controls	Yes	Yes
Control Mean	60.67	63.90
N	1846	1846
R2	0.07	0.08
	Panel C: Food Security	
Assigned to treatment	-0.11	-0.45**
	(0.56)	(0.02)
Treatment x Female	-0.11	-0.50
	(0.72)	(0.12)
Treatment x Male	-0.11	-0.43*
	(0.64)	(0.07)
Male	-0.19	-0.23
	(0.51)	(0.43)
District FE	Yes	Yes
Baseline Controls	Yes	Yes
Control Mean	4.61	4.73
N	1846	1846
R2	0.12	0.12

Notes: P-values in parentheses. We re-estimate the ITT making hypothetical assumptions about the missing data. We impute relatively high values for the dependent variables for missing control group individuals, and relatively low values for missing treatment group individuals. Sampling weights are applied. Standard errors are clustered at the group level. Baseline controls included: Individual characteristics: Age, age squared, age cubed, male (only full sample), urban, risk aversion. Education: Highest grade, literate, vocational training, digit recall test score, ADL Index, distance to educational facilities. Wealth: Wealth Index, savings, monthly income, could borrow \$58, could borrow \$580. Occupation: Weekly hours in low skill/business/agriculture, in school. Intervention: Grant amount applied for, group size, grant amount per member, group existed before application, group age in years, within-group heterogeneity, group dynamic, group committee member, chair or vice-chair. * implies p < .1, ** implies p < .05, *** implies p < .01.

Table A4: Lee Bounds for primary outcomes

	(1) Full sample	(2) Treatment x Men	(3) Treatment x Women
Panel A: Employed			
Lower bound	-0.04	-0.002	-0.07
	(0.03)	(0.03)	(0.04)
Upper bound	0.06	0.09**	-0.05
	(0.04)	(0.04)	(0.05)
Panel B: Income			
Lower bound	-22.31***	-7.83	-39.97***
	(7.75)	(8.17)	(13.10)
Upper bound	14.79**	36.57***	-31.48***
	(7.16)	(8.83)	(7.00)
Panel C: Food Security Index			
Lower bound	-0.31	-0.95**	0.86*
	(0.34)	(0.37)	(0.50)
Upper bound	0.91***	0.35	1.11***
	(0.29)	(0.31)	(0.40)

Notes: Lee Bounds are estimated by using the stata command *leebounds*. * implies p < .1, ** implies p < .05, *** implies p < .01.

A4. Tables and Figures.

Table A5 Pre-specified outcomes

Outcome	Indicator	Coding
Primary Outcomes		
Employment	Respondent worked for remuneration last 7 days	Binary with 1 if respondent worked and 0 otherwise
Income	Income of respondent last month	Sum of respondent income in the past month. Coded as zero if respondent did not earn any income in the last month. Coded as missing if one of the subcategories is missing. Top censored at the 99th percentile to contain outliers
Food Security	Number of days with reduced number of meals or reduced portion size (household)	Additive index
Secondary Outcomes		
Subjective Wellbeing	Subjective Economic Status	Index constructed as average of the two ordinal variables
Business Resilience	Change in business operations	Question E 8 will be coded as: 0 business remains open as usual, 1 temporarily closed by government mandate, 2 business temporarily closed, 3 business permanently closed
Farming Resilience	Change in farming practices	Additive standardized index of 6 ordinal variables. All farming variables are coded to missing if off season or if household does not grow crops
Safety Net	Amount of savings	Sum of respondent savings in bank accounts and saving groups. Coded as zero if the respondent does not have any savings
Remittances Received	Amount of remittances received	Total amount of remittances received. Coded as zero if the respondent has not received any remittances
Remittances Sent	Amount of remittances sent	Total amount of remittances sent. Coded as zero if the respondent has not sent any remittances

Notes: We pre-specified the outcomes before the data collection started. The registration and the pre-analysis plan can be accessed here: https://www.socialscienceregistry.org/trials/6158.

Table A6: Summary of 2-, 4-, and 9-years impacts

	(1)	(2)	(3)	(4)	(5)

	ITT Coefficient	Std. Err.	Control Mean	Obs.	Comment
	BFM (2014): 2 and 4-year impacts				
Investments					
Enrolled in vocational training	0.532	0.023***	0.152	1,999	Impact 2010 (2 years)
Hours of vocational training	340.5	22.521***	49.0	1,999	Impact 2010 (2 years)
Business assets (000s 2008 UGX)	225.0	62.601***	392.8	1,868	Impact 2012 (4 years)
Employment					
Avg. Employment hours per week	5.5	1.284***	32.2	1,864	Impact 2012 (4 years)
Engaged in any skilled trade	0.261	0.026***	0.22	1,868	Impact 2012 (4 years)
Works >= 30 hours a week in skilled trade	0.037	0.013***	0.03	1,868	Impact 2012 (4 years)
Income					
Monthly cash earnings (thousands)	18.19	4.898***	47.8	1,868	Impact 2012 (4 years)
Durable assets (z-score)	0.181	0.055***	0.15	1,853	Impact 2012 (4 years)
Nondurable consumption (z-score)	0.18	0.051***	-0.011	1,862	Impact 2012 (4 years)
	BFM (2014): 9-year impacts				
Employment					
Avg. Employment hours per week	0.513	1.593	44.68	1,981	
Engaged in any skilled trade					
Works >= 30 hours a week in skilled trade	0.029	0.011**	0.03	1,981	
Income					
Monthly cash earnings (thousands)	4.172	8.491	90.97	1,981	
Durable assets (z-score)	0.145	0.047***	0.25	1,981	
Nondurable consumption (thousands)	2.726	6.298	190.56	1,981	

Notes: The 2- and 4-year impacts are obtained from Table 3 in BFM (2014). 9-year impacts are obtained from Table 1 in BFM (2020). * implies p < .1, ** implies p < .05, *** implies p < .01.

Table A7: Summary of all outcomes in the 4-, 9-, and 12-year follow-up

	4-year		9-year		12-year	
	Effect size	Std. Err.	Effect size	Std. Err.	Effect size	Std. Err.
Business assets	225	62.601***	-	-	-	-
Food Security	-	-	-	-	-0.02[1]	0.25
Employment						
Avg. Employment hours per week	5.5	1.284***	0.513	1.593	-	-
Engaged in any skilled trade	0.26	0.026***	-	-	-	-
Works >= 30 hours a week in skilled trade	0.04	0.013***	0.029[3]	0.011**	-	-
Agricultural hrs/wk	0.4	0.945	0.08[3]	0.856	-	-
Nonagricultural hrs/wk	5.1	0.998***	0.43[3]	1.488	-	-
Skilled Trades only hrs/wk	3.8	0.548***	2.8[3]	0.529***	-	-
No employment hours in past month	-0.02	0.009***	0.03[3]	0.011**	-	-
Casual labor, low skill hrs/wk			-1.21[3]	0.99	-	-
Petty business, low skill hrs/wk			-1.6[3]	1.069	-	-
High-skill wage hrs/week			0.91[3]	0.582	-	-
Employment (pursued income generating activity in past week)	-	-	-	-	0.04[1]	0.03
Labor supply – total days					0.58	0.93
Labor supply – total hours					6.87	6.62
Income						
Standardized Income Index	0.22	0.05***	0.08[1]	0.048	-	-
Monthly cash earnings (thousands)	18.19	4.898***	4.172[1]	8.491	21.22[1]	12.418*
Durable assets (z-score)	0.18	0.055***	0.145[1]	0.047***	-	-
Nondurable consumption (z-score)	0.18	0.051***	2.726[1]	6.298	-	-
Migration and urbanization						
Has changed parish since baseline	-0.07	0.026***	-	-	-	-
Lives in large town or city	0.01	0.019	0.01	0.021	-	
Moved from village to town/city	-	-	0.02	0.017	-	-
Business formality						
Maintains formal records	0.12	0.023***	0.03[2]	0.021	-	-
Enterprise is formally registered	0.06	0.019***	-0.01[2]	0.013	-	-
Pays business taxes	0.09	0.023***	0.02[2]	0.018	-	-
Hired labor						
No. of paid/unpaid laborers in past month	0.64	0.243***	0.32[3]	0.165*	-	-
Nonagricultural activities only	0.21	0.108**	0.15[3]	0.083*	-	-
Skilled Trade only	0.09	0.038**	0.14[3]	0.042***	-	-
Total hours of paid/unpaid laborers in past month	210.6	63.915***	20.75[3]	8.909**	-	-
Nonagricultural activities only	34.3	24.711	6.31[3]	6.83	-	-
Skilled Trade only	7.3	3.895*	9.7[3]	3.763**	-	-
No. of paid laborers in past month	0.26	0.148	0.26[3]	0.136*	-	-
Nonagricultural activities only	0.08	0.05	0.11[3]	0.067*	-	-
Skilled Trade only	0.05	0.026*	0.1[3]	0.034***	-	-
Total pay to others on typical working day	2.28	1.414	-	-	-	-
Nonagricultural activities only	0.82	0.743	-	-	-	-

Skilled Trade only	0.42	0.26	-	-	-	-
Estimated total pay to others in past month	32.3	33.99	0.85[3]	6.162	-	-
Nonagricultural activities only	7.2	15.65	0.4[3]	5.02	-	-
Skilled Trade only	5.5	3.174*	3.05[3]	2.326	-	-
Number of family employees	-	-	0.09[3]	0.096		
Number of non-family employees	-	-	0.25[3]	0.144*		
Savings						
Has savings account/savings group	-	-	-0.03	0.024	-	-
Amount of savings in 000s	-	-	-9.17	12.362	109.9[2]	66.035*
Log savings	-	-	-0.02	0.106	-	-
Social Outcomes						
Kin integration	0.04	0.047	-	-	-	-
Community participation	0	0.05	-	-	-	-
Public goods contribution	0.01	0.049	-	-	-	-
Antisocial behavior	0.013	0.046	-	-	-	-
Protest and attitudes and participation	-0.02	0.043	-	-	-	-
Own health Outcomes						
Respondent passed away	-	-	-0.004	0.006	-	-
Physical health index (z-score)	-	-	-0.03[2]	0.047	-	-
Mental health index (z-score)	-	-	-0.06[2]	0.047	-	-
Fertility, HH size, and child expenditures						
Number of pregnancies 2007 or later	-	-	0.1[2]	0.101	-	-
Percent of births that were live 2007 or later	-	-	0.01[2]	0.01	-	-
Percent of pregnancies 2007 or later where child still living	-	-	0.01[2]	0.012	-	-
Percent of successful pregnancies 2007 or later where child still living	-	-	-0.01[2]	0.006	-	-
Number of biological children alive born 2007 or later	-	-	0.08[2]	0.083	-	-
Size of household	-	-	-0.13	0.162	-	-
Mean age of children (0–15)	-	-	0.01	0.138	-	-
Mean age of biological children (0–15)	-	-	0.1	0.147	-	-
Child educational outcomes						
Child age-adjusted educational attainment (6–24)	-	-	-0.01[2]	0.037	-	-
Child age-adjusted educational attainment (6–24), biological	-	-	-0.05[2]	0.045	-	-
Mean of child enrollment	-	-	-0.02[2]	0.013	-	-
Mean of child enrollment, biological	-	-	-0.02[2]	0.013	-	-
Current child expenditures (clothes and school)	-	-	0.41	2.784	-	-
Current child expenditures per child	-	-	0.5	1.071	-	-
Child health outcomes						
Mean health index per child, ages 3–9, family average	-	-	0.08	0.043*	-	-
Mean parent-reported health score per child, ages 3–9, family average	-	-	0.07	0.047	-	-
Mean malaria cases in past year, ages 3–9, family average	-	-	-0.13	0.087	-	-
Mean normalized ADL score per child, ages 3–9, family average	-	-	0.05	0.041	-	-

Political Behavior						
Index of political action (z-score)	-	-	0.06[2]	0.05	-	-
Attended voter education meeting	-	-	0.03[2]	0.024	-	-
Discussed Vote	-	-	-0.003[2]	0.024	-	-
Reported campaign malpractice or incident	-	-	-0.013[2]	0.012	-	-
Voted in presidential election	-	-	0.01[2]	0.013	-	-
Attended political rally	-	-	0.01[2]	0.025	-	-
Participated in political primary	-	-	0.014[2]	0.024	-	-
Worked to get a candidate/party elected	-	-	0.04[2]	0.025	-	-
Member of a political party	-	-	0.05[2]	0.024*	-	-
Index of NRM/Presidential support (z-score)	-	-	0.03[2]	0.05	-	-
Would vote for NRM if election were tomorrow	-	-	-0.01[2]	0.021	-	-
Like or strongly like NRM	-	-	0.01[2]	0.02	-	-
Worked to get the NRM elected	-	-	0.02[2]	0.03	-	-
Member of the NRM	-	-	0.04[2]	0.024	-	-
Voted or supported the president in the last election	-	-	-0.01[2]	0.02	-	-
Index of opposition support (z-score)	-	-	0.08[2]	0.044*	-	-
Would vote for opposition if election were tomorrow	-	-	0.02[2]	0.016	-	-
Like or strongly like any opposition party	-	-	0.03[2]	0.022	-	-
Worked to get the opposition elected	-	-	0.01[2]	0.01	-	-
Member of an opposition party	-	-	0.04[2]	0.024	-	-
Voted or supported an election party in the past election	-	-	-0.01[2]	0.022	-	-
Resilience						
Subjective Resilience	-	-	-	-	0.18[2]	0.06***
Business Resilience	-	-	-	-	0.07[2]	0.07
Farming Resilience	-	-	-	-	-0.001[2]	0.03
Economic Wellbeing	-	-	-	-	0.02[2]	0.03
Safety Net	-	-	-	-	-0.02[2]	0.03
Remittances						
Remittances sent in 000s	-	-	-	-	29.97[2]	43.531
Remittances received in 000s	-	-	-	-	-2.38[2]	13.641

Note: [1] outcome was pre-specified as primary outcome for respective follow-up. [2] outcome was pre-specified as secondary outcome for respective follow-up. [3] outcome was pre-specified as other outcome for respective follow-up. * implies p < .1, ** implies p < .05, *** implies p < .01.

Table A8: Baseline Balance for survey experiment groups

	Group 1 (=2)		Group 3 and 4			
	(1)	(2)	(3)	(4)	(5)	(6)
	Mean	SD	Mean	SD	Difference	p-value
Age at baseline	26.28	5.92	24.97	5.12	1.32	0.02**
Male	0.74	0.44	0.58	0.49	0.16	0.001***
Urban	0.18	0.38	0.15	0.36	0.02	0.55
Risk aversion (z-score)	-0.03	0.94	-0.01	1.06	-0.02	0.84
Highest grade reached in school	7.86	3.00	7.34	3.00	0.52	0.09*
Able to read	0.74	0.44	0.71	0.46	0.03	0.48
Received prior vocational training	0.07	0.25	0.07	0.26	0.00	0.93
Digit recall test score	4.16	2.00	3.93	1.96	0.22	0.27
ADL index	8.63	2.22	8.80	2.67	-0.17	0.52
Durable assets (z-score)	-0.12	0.92	-0.07	1.01	-0.04	0.65
Savings (000s 2008 UGX)	39.04	157.13	22.26	100.72	16.78	0.16
Monthly gross earnings (000s 2008 UGX)	66.61	117.01	65.25	129.82	1.36	0.92
Could obtain 100,000 UGX (58 USD) loan	0.37	0.49	0.38	0.49	-0.01	0.91
Could obtain 1,000,000 UGX (580 USD) loan	0.11	0.31	0.11	0.32	-0.01	0.87
Weekly work hours: low skill	0.72	4.06	0.86	4.40	-0.14	0.75
Weekly work hours: other business	2.57	7.39	2.41	6.88	0.16	0.82
Weekly work hours: skilled trade	3.16	11.33	1.17	6.66	1.99	0.02**
Weekly work hours: high skilled trade	0.24	1.38	0.05	0.60	0.20	0.03**
Weekly work hours: other non-agricultural	0.65	3.78	0.44	3.49	0.21	0.57
Weekly work hours: agricultural	4.79	10.33	5.54	10.38	-0.75	0.48
Weekly household chores, hours	7.37	14.82	9.48	16.97	-2.11	0.21
Zero employment hours in past month	0.39	0.49	0.45	0.50	-0.06	0.24
Main occupation is non-agricultural	0.37	0.49	0.24	0.43	0.14	0.00***
Engaged in a skilled trade	0.11	0.31	0.06	0.24	0.04	0.10
Currently in school	0.03	0.17	0.04	0.20	-0.01	0.59
Grant amount applied for (USD)	7159.00	2159.08	7525.72	2016.51	-366.73	0.08*
Group size	20.76	6.24	22.33	7.26	-1.58	0.03**
Grant amount per member, USD	372.47	165.52	375.93	172.05	-3.46	0.84
Group existed before application	0.44	0.50	0.43	0.50	0.01	0.83
Group age, in years	3.89	2.03	3.71	1.65	0.17	0.34
Within-group heterogeneity (z-score)	-0.07	1.00	-0.08	0.96	0.01	0.91
Quality of in-group dynamic (z-score)	-0.05	0.99	-0.10	1.08	0.05	0.65
Management committee member	0.33	0.47	0.25	0.43	0.08	0.08*
Chairperson or vice-chairperson	0.16	0.37	0.09	0.29	0.07	0.03**
Distance to educational facilities (km)	5.81	3.42	8.69	8.18	-2.88	0.00***
Lives in Adjumani	0.00	0.00	0.04	0.20	-0.04	0.02**
Lives in Apac	0.18	0.39	0.23	0.42	-0.05	0.25
Lives in Arua	0.08	0.27	0.08	0.27	0.00	0.92

Lives in Kaberamaido	0.00	0.00	0.03	0.18	-0.03	0.04**
Lives in Kotido	0.06	0.24	0.09	0.29	-0.03	0.30
Lives in Kumi	0.05	0.23	0.11	0.31	-0.05	0.07*
Lives in Lira	0.08	0.27	0.17	0.37	-0.09	0.01**
Lives in Moroto	0.03	0.17	0.04	0.19	-0.01	0.69
Lives in Moyo	0.01	0.09	0.02	0.16	-0.02	0.24
Lives in Nakapiripirit	0.05	0.21	0.07	0.25	-0.02	0.42
Lives in Nebbi	0.08	0.27	0.01	0.07	0.07	0.00***
Lives in Pallisa	0.19	0.39	0.04	0.20	0.15	0.00***
Lives in Soroti	0.10	0.30	0.04	0.20	0.06	0.01**

Notes: * implies p < .1, ** implies p < .05, *** implies p < .01.

Table A9: Variable description for primary and secondary outcomes

Outcome	Indicator	Question	Coding
		Primary Outcomes	
Employment	Respondent worked for remuneration last 7 days	In the past 7 days, have you worked for remuneration for at least one hour? By "work for renumeration" we mean any activities you undertook for remuneration, including daily labor, working for wages or in-kind, or working on your own account or running a business, including an agricultural business.	Binary with 1 if respondent worked and 0 otherwise
Income	Income of respondent last month	Q1: For casual labor/salaried employment, what was your wage/salary in the last 4 weeks? By salary I mean the cash that you earned related to activity. Q2: For commercial farming/self-employed business owner, what was your profit from this farm in the last month? By profits I mean the cash that you earned minus all expenses related to activity.	Sum of respondent income in the past month. Coded as zero if respondent did not earn any income in the last month. Coded as missing if one of the subcategories is missing. Top censored at the 99th percentile to contain outliers
Food Security	Number of days with reduced number of meals or reduced portion size (household)	Q1: In the past 7 days, how many days have you or someone in your household had to… Limit portion size at mealtimes? Q2: In the past 7 days, how many days have you or someone in your household had to… Reduce number of meals eaten in a day?	Additive index
		Secondary Outcomes	
Subjective Wellbeing	Subjective Economic Status	Q1: Compared to last year, would you say the economic situation of your household this year has improved, stayed the same or worsened? Q2: Compared to your neighbors, would you say the economic situation of your household is better than average, about average or worse than average?	Index constructed as average of the two ordinal variables
Business resilience	Change in business operations	What is the current status of your business?	Question E8 will be coded as: 0 business remains open as usual, 1 temporarily closed by government mandate, 2 business temporarily closed, 3 business permanently closed)
Farming resilience	Change in farming practices	For your main crop... Q1: Relative to the same season in the last year, how many days did you and your household members spend on this activity on your farm? Q2: Relative to the same season in the last year, how many days did you hire workers to work on this activity on your farm? Q3: Relative to the same season in the last year, how many seeds and inputs (e.g. fertilizer, chemicals) have you used (do you plan to use) for your farm for this crop? Q4: Relative to the same season in the last year, how much have you harvested (do you expect to harvest) for your farm for this crop? Q5: Relative to the same season in the last year, how are /do you expect prices for this crop? Q6: Are you/do you expect to be able to sell your crop in the locations/markets where you usually sell it?	Additive standardized index of 6 ordinal variables All farming variables are coded to missing if off season or if household does not grow crops
Safety Net	Amount of savings	Q1: How much of your own money do you have saved in this bank account now?	Sum of respondent savings in bank accounts and saving

		Q2: How much of your own money do you have saved with these groups? Q3: How much money do you have saved in other locations (Just to clarify, savings do not have to be deposited in an account or formal institution, and they may or may not gain interest. They can be somewhere at home, hidden in a safe place, or with a friend or family member)?	groups. Coded as zero if the respondent does not have any savings.
Remittances Received	Respondent received remittances	How much (remittances received) in total since the lockdown (March 17th)?	Total amount of remittances received. Coded as zero if the respondent has not received any remittances
Remittances Sent	Respondent sent remittances	How much (remittances sent) in total since the lockdown (March 17th)?	Total amount of remittances sent. Coded as zero if the respondent has not sent any remittances

Table A10: ITT effects for primary outcomes without baseline controls

	Employed (1)	Income (2)	Food Security (3)
Treatment	0.03	12.88*	-0.01
	(0.23)	(0.09)	(0.97)
District FE	Yes	Yes	Yes
Baseline Controls	No	No	No
q-value	0.41	0.41	0.52
Control Mean	0.672	59.38	4.62
N	1466	1525	1524
R2	0.14	0.057	0.14
Treatment x Women	-0.04	2.51	-0.06
	(0.34)	(0.75)	(0.89)
Treatment x Men	0.07**	15.32	0.07
	(0.03)	(0.14)	(0.80)
Men	-0.04	28.24***	-0.68*
	(0.38)	(0.00)	(0.09)
District FE	Yes	Yes	Yes
Baseline Controls	No	No	No
q-value Treatment x Female	0.86	1.00	1.00
q-value Treatment x Male	0.24	0.63	1.00
N	1466	1525	1524
R2	0.15	0.08	0.15

Notes: P-values in parentheses. Sampling weights are applied. Standard errors clustered at the group level. In all regressions we control for timing of interview and mode of interview (phone vs. person). The q-values are FDR sharpened q-values controlling for multiple hypothesis testing. * implies p < .1, ** implies p < .05, *** implies p < .01.

Table A11: ITT effect for primary outcomes without weights

	Employed	Income	Food Security
	(1)	(3)	(5)
Treatment	0.05*	13.54*	-0.15
	(0.09)	(0.06)	(0.53)
District FE	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes
Control Mean	0.67	59.38	4.62
N	1466	1525	1524
R2	0.13	0.11	0.15
Treatment x Women	-0.06	1.91	-0.05
	(0.14)	(0.78)	(0.89)
Treatment x Men	0.09***	18.84*	-0.2
	(0.00)	(0.05)	(0.49)
Men	-0.08**	14.14*	-0.31
	(0.08)	(0.08)	(0.41)
District FE	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes
N	1466	1525	1524
R2	0.14	0.11	0.15

Notes: P-values in parentheses. Standard errors clustered at the group level. In all regressions we control for timing of interview and mode of interview (phone vs. person). To correct for multiple hypothesis testing, we calculate q-values using the Benjamini-Hochberg step-up method. The q-values indicate the smallest false discovery rate at which the null hypothesis of zero effect is rejected. Baseline controls included in columns (1), (3), and (5): Individual characteristics: Age, age squared, age cubed, male (only full sample), urban, risk aversion. Education: Highest grade, literate, vocational training, digit recall test score, ADL Index, distance to educational facilities. Wealth: Wealth Index, savings, monthly income, could borrow $58, could borrow $580. Occupation: Weekly hours in low skill/business/agriculture, in school. Intervention: Grant amount applied for, group size, grant amount per member, group existed before application, group age in years, within-group heterogeneity, group dynamic, group committee member, chair or vice-chair. The q-values are FDR sharpened q-values controlling for multiple hypothesis testing. * implies p < .1, ** implies p < .05, *** implies p < .01.

Table A12: ITT effects for secondary outcomes without baseline controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Subjective Resilience	Business Resilience	Farming Resilience	Economic Wellbeing	Safety Net	Total savings	Remittances sent	Remittances received
Assigned to treatment	0.199***	0.072	0.025	0.032	-0.013	148,331**	11,002	-1,420
	(0.00)	(0.34)	(0.93)	(0.35)	(0.69)	(0.02)	(0.76)	(0.92)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Covariates	No	No	No	No	No	No	No	No
q-value	0.022		1.00	1.00		0.207	1.00	1.00
Control Mean	1.241	0.330	2.541	1.452	0.340	452,003	187,149	98,958
N	1524	483	788	1524	1466	1525	588	365
R2	0.108	0.129	0.015	0.070	0.080	0.039	0.065	0.099
Treatment x Female	0.078	0.160	0.286	-0.037	-0.028	35,119	84,209	3,327
	(0.43)	(0.20)	(0.58)	(0.46)	(0.63)	(0.70)	(0.42)	(0.90)
Treatment x Male	0.239***	-0.027	-0.169	0.063	-0.003	191,570**	-18,588	-5,953
	(0.00)	(0.76)	(0.63)	(0.13)	(0.93)	(0.01)	(0.58)	(0.75)
Male	0.193**	-0.037	0.593	0.023	-0.029	118,109**	52,092	20,437
	(0.05)	(0.65)	(0.13)	(0.60)	(0.48)	(0.04)	(0.17)	(0.35)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Covariates	No	No	No	No	No	No	No	No
q-value Treatment x Female	1.00		1.00	1.00		1.00	1.00	1.00
q-value Treatment x Male	0.03		1.00	1.00		0.154	1.00	1.00
N	1524	483	788	1524	1466	1525	588	365
R2	0.121	0.230	0.157	0.075	0.080	0.047	0.069	0.101

Notes: P-values in parentheses. Sampling weights are applied. Standard errors clustered at the group level. In all regressions we control for timing of interview and mode of interview (phone vs. person). The q-values are FDR sharpened q-values controlling for multiple hypothesis testing. * implies p < .1, ** implies p < .05, *** implies p < .01.

Table A13: ITT effects for secondary outcomes without weights

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Subjective Resilience	Business Resilience	Farming Resilience	Economic Wellbeing	Safety Net	Total savings	Remittances sent	Remittances received
Assigned to treatment	0.16***	0.01	-0.12	0.03	-0.01	136,992**	20,383	-2,384
	(0.01)	(0.90)	(0.66)	(0.20)	(0.83)	(0.03)	(0.54)	(0.86)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	1.24	0.33	2.54	1.45	0.34	452,003	187,149	98,958
N	1524	483	1524	1524	1466	1525	588	365
R2	0.17	0.20	0.14	0.12	0.09	0.1	0.11	0.22
Treatment x Women	0.09	0.10	-0.43	0.01	-0.04	83,284	60,820	-8,595
	(0.37)	(0.44)	(0.36)	(0.83)	(0.43)	(0.34)	(0.42)	(0.72)
Treatment x Men	0.19***	-0.05	-0.35	0.05	0.01	161,490**	5,756	-435
	(0.01)	(0.56)	(0.25)	(0.17)	(0.80)	(0.03)	(0.87)	(0.98)
Male	0.223**	0.01	0.90**	0.01	-0.04	108,848	36,094	17,593
	(0.02)	(0.91)	(0.05)	(0.86)	(0.37)	(0.11)	(0.38)	(0.36)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1524	483	1524	1524	1466	1525	588	365
R2	0.17	0.21	0.14	0.12	0.09	0.1	0.11	0.22

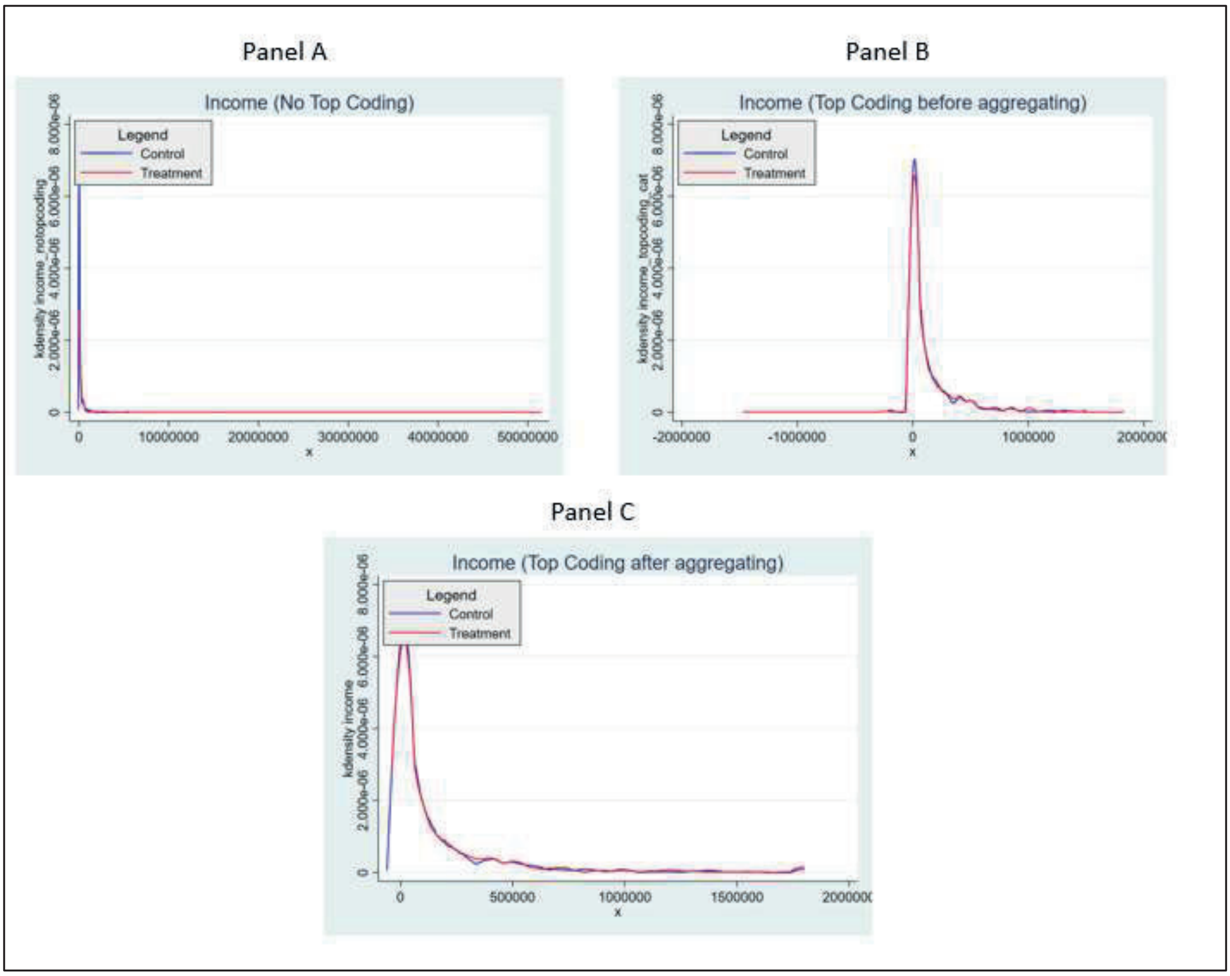
Notes: P-values in parentheses. Standard errors clustered at the group level. In all regressions we control for timing of interview and mode of interview (phone vs. person). To correct for multiple hypothesis testing, we calculate q-values using the Benjamini-Hochberg step-up method. The q-values indicate the smallest false discovery rate at which the null hypothesis of zero effect is rejected. Baseline controls included in columns (1), (3), and (5): Individual characteristics: Age, age squared, age cubed, male (only full sample), urban, risk aversion. Education: Highest grade, literate, vocational training, digit recall test score, ADL Index, distance to educational facilities. Wealth: Wealth Index, savings, monthly income, could borrow $58, could borrow $580. Occupation: Weekly hours in low skill/business/agriculture, in school. Intervention: Grant amount applied for, group size, grant amount per member, group existed before application, group age in years, within-group heterogeneity, group dynamic, group committee member, chair or vice-chair. The q-values are FDR sharpened q-values controlling for multiple hypothesis testing. * implies p < .1, ** implies p < .05, *** implies p < .01.

Table A14: Different top-coding scenarios for income

	(1) Income – After	(2) Income – After	(3) Income - Before	(4) Income – Before	(5) Income – no top-coding	(6) Income – no top-coding
Treatment	12.01*	12.88*	9.60*	10.49*	38.27	42.97
	(0.08)	(0.09)	(0.08)	(0.09)	(0.10)	(0.13)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	No	Yes	No	Yes	No
Control Mean	59.38	59.38	53.58	53.58	63.75	63.75
N	1525	1525	1525	1525	1525	1525
R2	0.11	0.04	0.11	0.05	0.05	0.01
Treatment x Female	2.52	2.51	3.09	3.43	3.94	8.07
	(0.73)	(0.75)	(0.65)	(0.64)	(0.77)	(0.44)
Treatment x Male	17.15*	15.32	13.13*	11.77	56.86	57.30
	(0.08)	(0.14)	(0.08)	(0.16)	(0.10)	(0.15)
Male	17.38*	28.24***	13.72*	23.56***	32.55	24.88**
	(0.07)	(0.00)	(0.09)	(0.00)	(0.12)	(0.04)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	No	Yes	No	Yes	No
N	1525	1525	1525	1525	1525	1525
R2	0.09	0.06	0.1	0.07	0.05	0.01

Notes: P-values in parentheses. In columns (1) and (2) income is top-coded at the category level and the 99th percentile. In columns (3) and (4) the aggregated income is top-coded at the 99th percentile and in columns (5) and (6) income is not top coded. Sampling weights are applied. Standard errors clustered at the group level. In all regressions we control for timing of interview and mode of interview (phone vs. person). Baseline controls included in columns (1), (3), and (5): Individual characteristics: Age, age squared, age cubed, male (only full sample), urban, risk aversion. Education: Highest grade, literate, vocational training, digit recall test score, ADL Index, distance to educational facilities. Wealth: Wealth Index, savings, monthly income, could borrow $58, could borrow $580. Occupation: Weekly hours in low skill/business/agriculture, in school. Intervention: Grant amount applied for, group size, grant amount per member, group existed before application, group age in years, within-group heterogeneity, group dynamic, group committee member, chair or vice-chair. * implies p < .1, ** implies p < .05, *** implies p < .01.

Figure A1: Distribution of income with different top-coding scenarios



Notes: Income in UGX on x-axis.

Figure A2: Total savings of treatment and control group members