Online Appendix

The long term impacts of gra.nts on poverty: 9-year evidence from Uganda's Youth Opportunities Program

Christopher Blattman
Nathan Fiala
Sebastian Martinez

Contents

A Existing evidence on microenterprise assistance

Governments and nonprofits commonly grant cash, livestock, or equipment to poor people who propose to start basic businesses. Broadly speaking, evaluations 1–4 years after these programs show that recipients raised their incomes compared to randomized control groups.

- In post-conflict northern Uganda, a program giving women $150 grants, basic training and follow-up led to large income gains 18 months after the grants (Blattman et al., 2016).

- Other studies show that grants of cash and in-kind capital to less poor, existing entrepreneurs in Sri Lanka, Ghana and India lead to sustained increases in earnings 1–5 years later (de Mel et al., 2012; Fafchamps et al., 2014; Hussam et al., 2017).

- Cash grants to poor farmers in Mali raised farm inputs and incomes after 1 and 2 years (Beaman et al., 2018). The authors see substantial heterogeneity in returns to capital among Malian farmers, and high productivity credit constrained farmers select into loans when they become available. In villages where farmers were eligible for loans, those who did not take out loans did not have high returns to subsequent grants, while farmers in villages ineligible for loans did have high average returns.

- Across seven countries, multifaceted programs that give grants of livestock alongside basic training and temporary income support show sustained increases on the incomes and consumption of the poorest rural households four years after grants (Banerjee et al., 2015b; Bandiera et al., 2013, 2017).
- Most conditional cash transfer programs do not target (or measure) investment an earnings, but there is some evidence from a Mexican national program that cash relieves important financial constraints and leads to higher income after 1–2 years (Gertler et al., 2012; Bianchi and Bobba, 2013).
- Studying a multifaceted “microfranchising” program in Nairobi, Kenya, Brudevold-Newman et al. (2017) see income gains after 9 months but none after 1.5 years. This could imply rapid convergence, though the authors interpret the results as evidence of a savings constraint, implying that the participants did not have high returns to this multifaceted program (including capital).

Of course the effects of capital grants are not universally positive.

- Fiala (2018) fails to find income effects from cash grants to existing businesses in Uganda.
- In one of the multifaceted livestock programs, in Ghana, a grant of goats alone (without other program components) has no effect on incomes after 2 or 3 years (Banerjee et al., 2018).
- A cash grant programs to young men living on the streets of Monrovia and engaged in petty crime also had very short-lived impacts (less than one year) on enterprise and earnings, potentially due to the unusual instability and risk of their existence (Blattman et al., 2017).
- Karlan et al. (2014) find that cash grants to Ghanaian farmers had not effect without insurance, also because of the constraints from imperfect insurance.
- And in Kenya, a multifaceted “microfranchising” program that includes capital finds earnings gain in the first nine months but not after 1.5 years (Brudevold-Newman et al., 2017).

Another strain of anti-poverty programs, not discussed so far, provide unconditional cash transfers (UCTs) to the poor. These are not generally targeted or designed to support microenterprise start-up and raise incomes. Consistent with this, medium term studies of two programs find short-term increases in consumption and assets, as recipients spend the money, but they find no evidence of investment or income gains over 2–4 years.

- Baird et al. (2017, 2011) look at small monthly conditional cash transfers (CCTs) and UCTs to adolescent girls in Malawi, sustained over three years, where the CCT is conditioned on school attendance. At the end of this program of transfers, and 2 years afterwards, they see little effect on incomes. They also see convergence after early gains in teen pregnancy, early marriage, and sexually transmitted infections.

- Haushofer and Shapiro (2016, 2018) evaluate UCTs from GiveDirectly in Kenyan villages after 9 months and 3 years, and find sustained increases in assets between treatment and control villages, but no consumption impacts. They do find consumption differences between treated and "spillover" households within treated villages, but it is not clear if this evidence of an income gain or an adverse spillover.

A final common set of programs are those that give poor people access to cheap finance. In principle micro-loans should solve the credit market imperfections that constrain microentrepreneurs. But lending and banking such small sums of money for so many people can be expensive. Also, it is hard to improve institutions that reduce the loan market's information and collection problems. As a result, interest rates on micro-loans tend to be very high. That and the short time periods for repayment of most micro-loans can make them poor vehicles for business investment. This potentially one reason why some recent micro-lending experiments show little or no impact on the earnings of most loan clients.

- See for example Banerjee (2013); Banerjee et al. (2015a).

- Of course, better targeted, cheaper microfinance with longer repayment periods could stimulate the same kinds of entrepreneurship as grants (see for example Feigenberg et al., 2013).

- Thus we view microfinance as operating under the model and assumptions as these other entrepreneurship programs. The theory and evidence of success from the Thai Million Baht Village Fund program is one example (Kaboski and Townsend, 2011).

B Program and experimental design details

B.1 Additional intervention details

One reason for group applications was administrative convenience—it was easier to screen and disburse to a few hundred groups. Another is that officials hoped groups would be more likely to implement proposals. Finally, groups could take advantage of economies of scale in purchasing tools or trainers. But officials were not aiming to create cooperatives. In general, they expected that recipients would set themselves up as independent businesses, although they might share some tools or collaborate.

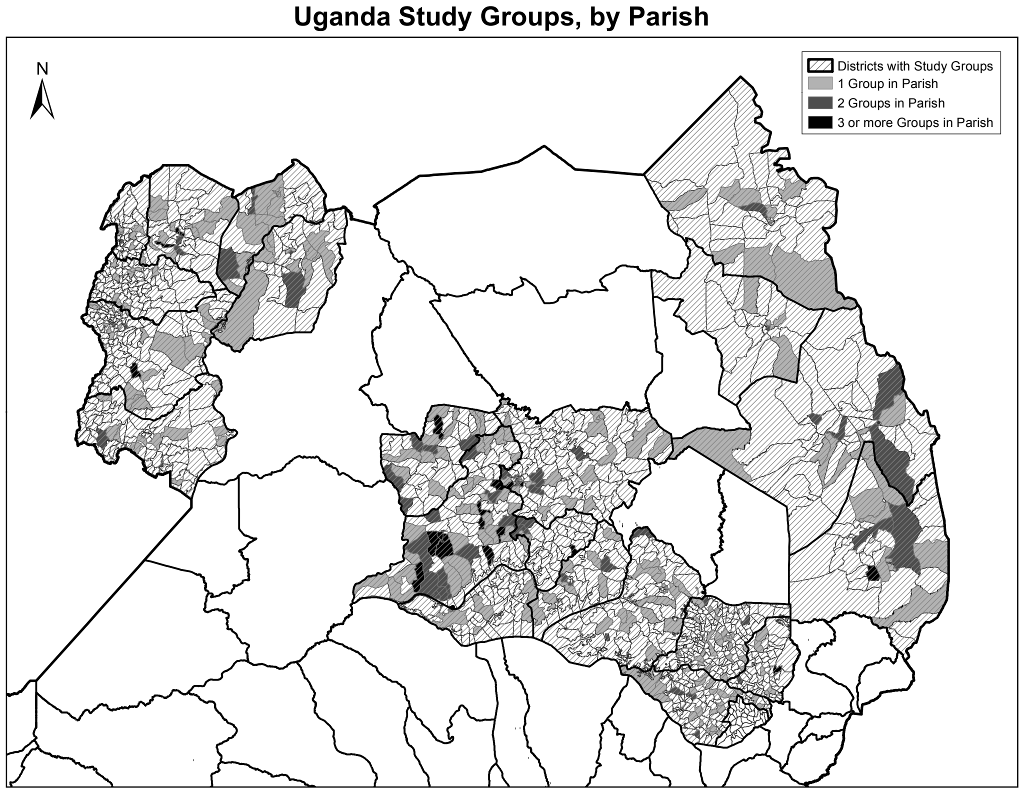
In our sample, 5% of groups are all female and 12% are all male, but most groups are mixed—about one-third female on average. Females and mixed groups often chose trades common to both genders, such as tailoring or hairstyling. Males and a small number of females often chose trades such as carpentry.

The proposal specified members, a management committee of five, proposed trades, and assets to purchase. Decisions were made by member vote, and nearly all members report they had a voice in decisions. In preparing the proposal, groups identified their own trainers, typically a local artisan or small institute. These are commonplace in Uganda (as in much of Africa) and there is a tradition of artisans taking on paying students as apprentices. Most of these artisans and institutes had been in existence more than five years, and most took students previously. In our sample, few were located in the village but the median artisan or institute was within 8 kilometers. Groups would travel to be closer to trainers, or paid transport and upkeep for trainers to come to them.

Proposals had to receive formal advising. Many applicants were functionally illiterate, so YOP also required "facilitators" (usually a local government employee, teacher, or community leader) to meet with the group several times, advise them on program rules, and help prepare the written proposals. Groups chose their own facilitators, and the NUSAF office paid facilitators 2% of funded proposals (up to $200).

Villages or parishes typically submitted one application, and that privilege may have gone to the groups with the most initiative, need, or connections. Village officials passed applications up to districts, which verified the minimum technical criteria (such as group size and a complete proposal) and were supposed to visit projects they planned to fund. Most of the applications were made two years before their actual selection.

Figure B.1: Eligible districts and total number of study communities per parish



B.2 Experimental procedures

Sample selection YOP launched in 2006 and thousands of groups submitted proposals. The government funded hundreds of proposals in 2006-07, prior to our study. By 2008, 14 of the 18 NUSAF-eligible districts had funds remaining for YOP. Figure B.1 maps these study districts along with the number of groups per parish in the experimental sample.[13]

In 2007, the central government asked district governments to nominate 2.5 times the number of groups they could fund. The districts submitted roughly 625 proposals to a central government office. Based on our discussions with district officials, each district reviewed only a small fraction of the thousands of proposals they received. Most said they tried taking a “first come first serve” approach. Others described other ad hoc criteria. Local political considerations might also have come into play, although we heard very few accusations of “political pork” or favoritism.

To minimize chances of misuse and corruption, the central government reviewed proposals for completeness and validity. They also sent out audit teams to visit and verify each group.

[13]Gaps in administrative data mean that 20 villages are linked to a district but not a parish. Of the 26 parishes with three or more groups per parish, just six parishes have 4 or more groups.

The government disqualified about 70 applications, mainly for incomplete information or ineligibility (e.g. many group members over 35 years, or a group size more than 40). The government also asked that 22 groups of underserved people (Muslims and orphans) be funded automatically.

Randomization procedures and balance Randomization took place without baseline data because of the central bank's need to create hundreds of bank accounts for the groups in time for program launch. Procurement processes then delayed the baseline survey by 1–2 months.

Table B.1 displays summary statistics and tests of balance for 38 baseline covariates. There is balance across a wide range of measures, but a handful show imbalance. They suggest higher levels of initial wealth among the treatment group. [14]

While this imbalance may be chance, it could also be due to 13 missing control groups. At baseline enumerators could not locate 13 groups (3% of the sample). Unusually, after the survey it was discovered that all 13 had been assigned to the control group. We investigated the matter and found no motive for or evidence of foul play, and no signs of communication between the central bank and local branches of government. District officials, enumerators, and the groups themselves did not know the treatment status of the groups they were mobilizing. We were only able to find one of the 13 at endline. We estimate that if the missing control groups had baseline values just 0.05 standard deviations above the control mean, we would fail to find statistically significant imbalance. If 0.1 standard deviations above the control mean, the mean differences between the treatment and control groups would be close to zero.[15] If so, this would imply that the observed control group may be poorer than the treatment group, and will lead us to overstate true program impacts. We model alternative attrition scenarios below.

B.3 Survey attrition

The national statistics agency conducted the baseline survey, and Innovations for Poverty Action (IPA) and private survey firms conducted all endline surveys.

Table B.2 reports survey response rates and sample size at each round. At baseline, enumerators and local officials mobilized group members to complete a survey of demographic data on all members as well as group characteristics. Virtually all members were mobilized.

[14] The treatment group report 2 percentage points more vocational training, 0.07 standard deviations (SD) greater wealth, 56% greater savings (though only in the linear, not in log form), and 5 percentage points more access to small loans.

[15] Not shown here. These results were originally reported in Blattman et al. (2014).

YOP applicants were young and mobile. 30–40% of respondents had moved or were away temporarily at each endline survey. To minimize attrition, we used a two-phase tracking approach (Thomas et al., 2001). Table B.2 summarizes the phases. In Phase 1, we attempted to interview all 2,675 people in their last known location. We did not find 37% after 2 years, 39% after 4 years, and 29% after 9 years. Most had migrated away from their original home. In Phase 2, we selected a random sample of the unfound—53% after 2 years, 38.5% after 4 years, and 36% after 9 years.[16] We made at least three attempts to find this subset in their new locations.

In total, we found 75% of the selected sample in after 2 years, 70% after 4 years, and 74% after 9 years. Those found in Phase 1 receive unit weight, those selected in Phase 2 are weighted by the inverse of their selection probability, and those not selected in Phase 2 are dropped. Fewer than 2.5% of people refused to answer. Our response rate was 97% at baseline, and effective response rates at endline (where individuals found in phase 2 tracking were given higher weights) were 86% after 2 years, 82% after 4 years, and 87% after 9 years. There is some correlation between attrition and treatment, statistically significant in 4-year endline but not in the 2 or 9-year endline.

There is also some correlation between attrition and potential indicators of economic success. If so, and uncorrected, selective attrition could bias our treatment effects upwards, as estimated in rows 9 and 10 of D.10. Table B.3 reports baseline correlates of attrition (excluding the 3% not found at baseline). Arguably, people with slightly more entrepreneurial potential are slightly less likely to be found: town dwellers, non-farmers, and those with higher literacy, initial employment, earnings and loan access, and members of the group management committee. To attempt to correct for any possible bias, we also weight individuals by the inverse of their predicting probability of attrition, calculated using a Leave-One-Out logistic regression. The two weights, sampling and selective attrition, are multiplied together such that found members of the sample who look more like the attritors will get slightly more weight in the estimated treatment effect.

B.4 Receipt of other programs

We would expect treatment-control convergence if control group members were more likely to receive another government or charitable program in the 9 years after YOP. Broadly speaking, we do not find any evidence that the control group received a greater number of transfers since baseline. In the 4-year survey, we asked about non-YOP programs received in the 12 months before the survey. In the 9-year survey we ask respondents to report all

[16]The proportions mainly varied according to our available resources. The Phase 2 randomization stratified by district and by the proportion unfound in the group.

"major programs" they received from the government or an NGO since baseline.[17] Table B.4 reports results.

Our estimates vary somewhat depending on what types of programs we include or exclude, and so we report several alternative measures. Our preferred measure omits very minor programs (such as mosquito nets, vaccinations, packets of seeds, and unspecified agricultural extension) as well as any programs reported by the treatment group that sound like YOP (such as tailoring training, repair training, and business grants whose source was unspecified). In the decade before the survey, about 17% of the control group reports at least one non-minor program. Some of these are substantial (such as a goat or a training course) but most appear to be modest in nature. In the 4-year survey, for example, we calculated that the average program received by the control group was worth less than $20, which is less than 5% of the value of the YOP grant. After 9 years, the treatment group is 2.9 percentage points less likely to report such assistance (not statistically significant).

If we change our preferred measure to include unspecified agricultural extension as "non-minor", include any YOP-like programs reported by the treatment group, or include all minor programs, the program effect shrinks or reverses. As a result, we do not regard differential treatment with other programs to be a major driver of convergence over the 9 years.

B.5 Measurement

Our primary outcome, income, is notoriously difficult to measure in situations where households have multiple streams of irregular, informal employment. We collected data on three proxies for income and combined them, with equal weights, into an additive index that is normalized to have mean zero and standard deviation one separate within each survey round. The three indices are:

1. *Monthly net earnings (i.e. business profits plus wages).* To measure earnings (and employment levels), survey enumerators went through a list of roughly three dozen common occupations, from farm wage work to skilled trades, asking if the person had worked in that occupation in the past year. Respondents usually answered yes to two to five occupations. For each of these affirmatives, the electronic survey guided the enumerator and respondent through a series of questions about each occupation, including days worked in the past month, average hours per day, and gross and net earnings in the previous month. We used these data to develop estimates of earnings

[17] Specifically we asked respondents to recall all major government or NGO programs they had received since the first multiparty elections in 2006. We used this election as a starting point because it is one of the most significant and memorable events that decade, and we believed it was the best way to have a consistent recall period across respondents.

and hours worked in the previous month. We also report a measure of monthly *gross* earnings (business revenues plus wages).

2. *Nondurable consumption.* The survey included an abbreviated consumption module, measuring the approximate value of food consumed in the past week, and less frequent expenditures (such as clothing or entertainment) in the past month. Thus the sum of these consumed items is not a complete measure of consumption. Nonetheless, the items should account for the majority of non-durable consumption and (so long as treatment does not have a major effect on the composition of consumption) the treatment effect on the abbreviated measure should approximate the treatment effect on full consumption.

3. *Durable assets.* An index of durable assets has been shown to be a reliable proxy of consumption and poverty (Filmer and Pritchett, 2001). We use this as an alternate measure to the abbreviated consumption module, which did not include consumption of durable assets. We do not have imputed rental or resale values of these items, as we deemed these data noisy and unreliably reported. Instead, the survey asked people several questions about their housing quality (such as roof material or dwelling size) as well as the quantity of a number of assets, ranging from various livestock, fruit-producing trees, household goods and furnishings, electronics, and tools and equipment. We have indicators for housing quality and counts of these durable assets. We combine these into a single index using the first principal component of all these asset variables. Our principle component analysis includes observations from baseline, 2, 4, and 9-year surveys. After predicting "durable assets" using the weights form the PCA, we normalize the variable such that the our z-score variable has a mean zero and standard deviation one across our combined baseline, 2, 4, and 9-year surveys. Because we use the same weights across all survey rounds (an object adds the same to one's durable assets score in baseline as it does in the 9-year endline), we can track increases in assets across years, as shown in 1, Panel C and the change in control means across columns 1, 2, and 3 in Table 1 . We consider it a proxy for poverty and permanent income.

We deflate all UGX-denominated measures by a national measure of inflation. Note that Ugandan inflation rose dramatically between the 4- and 9-year surveys. For this reason real consumption levels look fairly flat over this period. These national inflation rates may exaggerate inflation in rural areas, however, though without official data it is difficult to know.

Employment was measured in the same module as monthly earnings, as described above.

Table B.1: Balance test of all covariates at baseline

Covariate at baseline	Means		Difference	
	Control	Treated	Effect of treatment on covariate	p-value
	(1)	(2)	(3)	(4)
Age at baseline	24.77	25.14	0.111	0.681
Age squared	639.46	660.00	5.334	0.732
Age cubed	17,253.38	18,159.66	185.080	0.802
Male	0.67	0.68	0.004	0.885
Large town / urban area	0.22	0.20	-0.002	0.942
Risk Aversion (z-score)	0.03	-0.03	-0.094	0.050**
Found at baseline	0.94	1.00	0.053	0.000***
Highest grade reached in school	7.90	7.82	-0.091	0.532
Able to read and write minimally	0.76	0.71	-0.045	0.029**
Received prior vocational training	0.07	0.08	0.021	0.066*
Digit recall test score	4.16	4.01	-0.033	0.676
ADL index	8.65	8.63	-0.111	0.397
Durable Assets (z-score)	-0.17	-0.07	0.069	0.146
Savings (000s 2008 UGX)	18.18	32.85	11.838	0.008***
Monthly gross earnings (000s 2008 UGX)	59.59	67.54	9.062	0.161
Could obtain 100,000 UGX (58 USD) loan	0.31	0.40	0.074	0.000***
Could obtain 1,000,000 UGX (580 USD) loan	0.09	0.12	0.019	0.190
Weekly work hours: low skill	0.97	1.07	-0.002	0.991
Weekly work hours: other business	2.11	2.40	0.336	0.279
Weekly work hours: skilled trade	1.68	1.54	-0.260	0.499
Weekly work hours: high skilled trade	0.04	0.14	0.082	0.015**
Weekly work hours: other non-agricultural	0.85	0.56	-0.255	0.132
Weekly work hours: agricultural	4.40	5.63	1.215	0.013**
Weekly hourshold chores, hours	8.46	8.65	0.752	0.382
Zero employment hours in past month	0.45	0.42	-0.017	0.560
Main occupation is non-agricultural	0.24	0.28	0.020	0.357
Engaged in a Skilled Trade	0.07	0.08	0.006	0.613
Crrently in School	0.04	0.04	-0.003	0.734
Grant amount applied for (USD)	7,497.25	7,274.89	144.624	0.285
Group Size	22.52	21.24	0.035	0.946
Grant Amount per Member, USD	363.39	381.71	13.811	0.257
Group existed before application	0.43	0.49	0.053	0.207
Group age, in years	3.76	3.82	-0.005	0.978
Within-group heterogeneity (z-score)	-0.03	0.02	-0.022	0.785
Quality of in-group dynamic (z-score)	-0.01	0.02	0.006	0.939
Management comittee member	0.26	0.29	0.011	0.493
Chairperson or vice-chairperson	0.10	0.12	0.015	0.133
Distance to educational facilities (km)	6.75	7.27	0.638	0.207

Notes: Each row represents a separate regression of the effect of treatment assignement on pre-treatment covariates, including district fixed effects and group-level clustering of standard errors. * implies $p < .1$ ** implies $p < .05$ *** implies $p < .01$

There is balance across a wide range of measures, but a handful show imbalance. They suggest higher levels of initial wealth among the treatment group. While this imbalance may be chance, the missing 13 control groups could also cause the imbalance. We estimate that if the missing control groups had baseline values just 0.05 standard deviations above the control mean, we would fail to find statistically significant imbalance. If 0.1 standard deviations above the control mean, the mean differences between the treatment and control groups would be close to zero (see Blattman et al. 2014). If so, this would imply that the observed control group may be poorer than the treatment group, and will lead us to overstate true program impacts. We model alternative attrition scenarios in the main paper for this reason.

Table B.2: Survey response rates by survey round

Survey	Selection and tacking, by survey phase					Effective response rates				
	Total sought (1)	Found phase 1 (%) (2)	Selected phase 2 (%) (3)	Found phase 2 (%) (4)	Final # of obs. (5)	All (%) (6)	Control (%) (7)	Treated (%) (8)	Difference (%) (9)	p-value (10)
2008 baseline	2,677	97.0			2,598	97.0	94.4	99.8	5.3	0.000
2-year endline	2,677	63.4	53.0	74.7	2,005	85.5	85.6	85.3	-0.8	0.698
4-year endline	2,677	61.0	38.5	58.6	1,868	82.1	79.1	85.5	7.1	0.004
9-year endline	2,677	71.1	36.0	43.2	1,981	87.3	88.5	86.1	-2.8	0.266

Notes: *Effective response rates* refers to the sum of weights of our sample where individuals found in phase 2 are given a weight equal to one over the probability of being selected into phase 2 tracking, described in column 3.

At baseline, enumerators and local officials mobilized group members to complete a survey of demographic data on all members as well as group characteristics. Virtually all members were mobilized. Enumerators could not locate 13 groups (3% of the sample). Unusually, after the survey it was discovered that all 13 had been assigned to the control group. We investigated the matter and found no motive for or evidence of foul play, and no signs of communication between the central bank and local branches of government. District officials, enumerators, and the groups themselves did not know the treatment status of the groups they were mobilizing. We were only able to find one of the 13 at endline.

YOP applicants were young and mobile. 30-40% of respondents had moved or were away temporarily at each endline survey. To minimize attrition, we used a two-phase tracking approach. In Phase 1, we attempted to interview all 2,675 people in their last known location. We did not find 37% in after 2 years, 39% after 4 years, and 29% in after 9 years. Most had migrated away from their original home. In Phase 2, we selected a random sample of the unfound: 53% after 2 years, 38.5% after 4 years, and 36% after 9 years. We made at least three attempts to find this subset in their new locations.

We found 75% of the selected sample after 2 years, 59% after 4 years, and 74% after 9 years. Those found in Phase 1 receive unit weight, those selected in Phase 2 are weighted by the inverse of their selection probability, and those not selected in Phase 2 are dropped. Fewer than 2.5% of people refused to answer. Our response rate was 97% at baseline, and effective response rates at endline (where individuals found in phase 2 tracking were given higher weights) were 90.7% after 2 years, 84% after 4 years, and 87% after 9 years.

Table B.3: Correlates of survey attrition

	Dependent variable: indicator for being unable to find in round					
	2-year endline		4-year endline		9-year endline	
	Coeff	SE	Coeff	SE	Coeff	SE
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned to treatment	0.02	0.02	-0.03	0.02	0.04	0.02
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	-0.00	0.00
Grant Amount per Member, USD	0.00	0.00	-0.00	0.00	-0.00	0.00
Group existed before application	-0.02	0.02	-0.01	0.03	-0.01	0.03
Group age, in years	0.00	0.01	-0.00	0.01	-0.00	0.01
Within-group heterogeneity (z-score)	0.01	0.01	0.03	0.01*	-0.00	0.01
Quality of in-group dynamic (z-score)	0.01	0.01	-0.02	0.02	-0.01	0.01
Distance to educational facilities (km)	0.00	0.00	0.00	0.00	0.00	0.00
Age at baseline	-0.00	0.00*	-0.01	0.00***	-0.01	0.00***
Large town / urban area	0.08	0.03***	0.17	0.04***	0.12	0.03***
Risk Aversion (z-score)	0.04	0.01***	0.02	0.01**	0.04	0.01***
Management comittee member	-0.04	0.02**	-0.03	0.02	-0.01	0.03
Chairperson or vice-chairperson	0.01	0.03	0.02	0.04	-0.05	0.03
Weekly work hours: low skill	0.00	0.00*	0.00	0.00	0.00	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	0.00	0.01	-0.02	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	-0.01	0.00***	-0.00	0.00	0.00	0.00
Weekly hourshold chores, hours	-0.00	0.00	-0.00	0.00	0.00	0.00
Zero employment hours in past month	-0.13	0.03***	-0.02	0.03	0.06	0.03**
Main occupation is non-agricultural	-0.17	0.04***	0.03	0.05	0.10	0.04**
Engaged in a Skilled Trade	-0.06	0.04*	-0.03	0.05	0.02	0.05
Crrently in School	-0.08	0.03**	-0.04	0.05	-0.06	0.05
Highest grade reached in school	-0.00	0.00	0.00	0.00	0.00	0.00
Able to read and write minimally	0.06	0.02***	0.02	0.03	-0.04	0.02
Received prior vocational training	-0.04	0.03	-0.04	0.04	-0.10	0.03***
Digit recall test score	-0.01	0.00**	0.02	0.01***	-0.00	0.01
ADL index	-0.01	0.00***	-0.00	0.00	-0.00	0.00
Durable Assets (z-score)	0.02	0.01	-0.01	0.01	-0.01	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	-0.02	0.02	0.01	0.02	0.03	0.02
Could obtain 1,000,000 UGX (580 USD) loan	-0.01	0.03	0.01	0.04	-0.06	0.03**
Lives in Adjumani	-0.02	0.06	-0.09	0.09	-0.06	0.06
Lives in Apac	0.06	0.05	-0.07	0.08	0.07	0.06
Lives in Arua	0.12	0.06**	-0.02	0.08	0.17	0.07**
Lives in Kaberamaido	-0.03	0.05	0.04	0.10	0.03	0.08
Lives in Kotido	0.24	0.08***	0.10	0.10	0.06	0.08
Lives in Kumi	0.01	0.05	-0.07	0.08	-0.01	0.06
Lives in Lira	0.10	0.06*	-0.09	0.08	0.10	0.08
Lives in Moroto	0.26	0.08***	0.17	0.11	0.06	0.08
Lives in Moyo	0.05	0.06	-0.16	0.08**	0.13	0.09
Lives in Nakapiripirit	0.09	0.06	0.02	0.10	0.11	0.08
Lives in Nebbi	0.02	0.07	-0.10	0.09	-0.00	0.06
Lives in Pallisa	0.01	0.05	-0.17	0.08**	0.04	0.06
Lives in Soroti	-0.06	0.06	-0.05	0.09	0.06	0.07
Mean	0.15		0.18		0.12	
P-value of F-test	0.00		0.00		0.00	
N	2,322.00		2,035.00		2,086.00	
R-squared	0.13		0.16		0.11	

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 B.4: Program impacts on other aid programs received since baseline

Dependent Variable N = 1868 in 4-year N = 1981 in 9-year	Control mean 4-year (1)	 9-year (2)	Treatment effects 4-year (3)	 Difference (4)	 9-year (5)
Treated	0.00	0.00	0.879 [0.020]***	0.002 [0.028]	0.881 [0.019]***
Received a non-YOP transfer in the past 12 months	0.02		0.026 [0.009]***		
Reported a major non-YOP program since 2006		0.17			-0.029 [0.018]
Including all minor programs		0.27			0.001 [0.023]
Inlcuding agricultural extension in non-minor		0.16			0.045 [0.018]**
Including YOP-like programs reported by treatment group		0.17			0.025 [0.019]
Received any NUSAF, YOP, or YOP-like grant since 2006		0.17			0.725 [0.022]***

Notes: Each entry in columns 3 and 5 is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment, district fixed effects, and a vector of baseline covariates. Standard errors are clustered at the group level (of up to 5 people). We report the coefficient on treatment only. All regressions are weighted by inverse probabilities of attrition and selection into the endline tracking sample. Control means in columns 1 and 2 are also calculated using these weights.

Column 4 refers to the difference between coefficients between endlines. P-values for diferences are calculated using a simple t-test using the standard errors of coefficients.

* implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

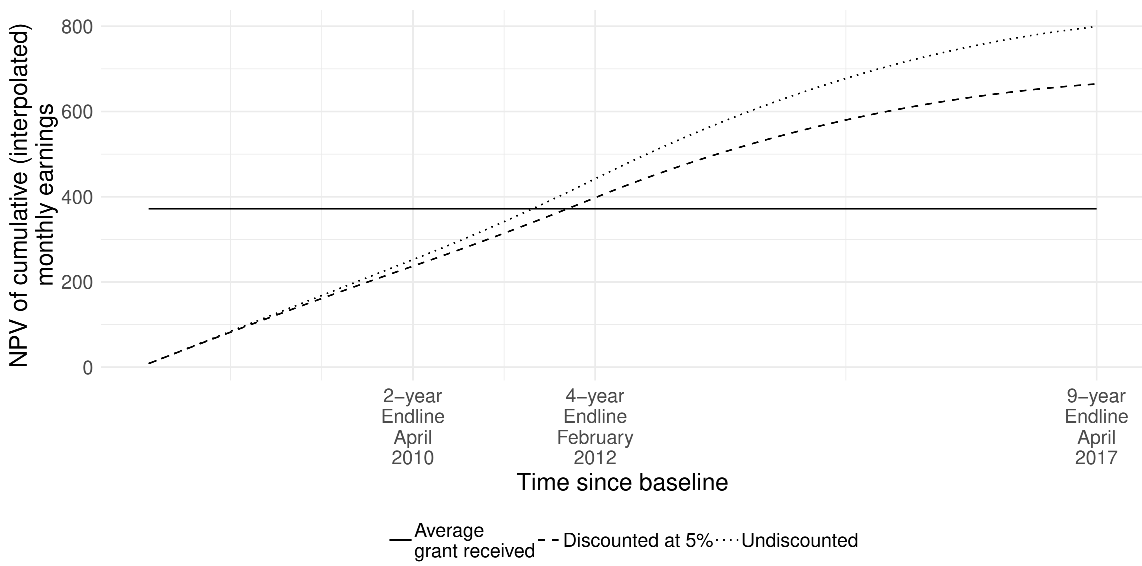
C Additional analysis of economic and human capital impacts

Table C.1: Work opportunities without outside assitance: Progression of control group

Variable	Control group mean (Hours)			
	Baseline	2-year	4-year	9-year
	(1)	(2)	(3)	(4)
Average employment hr/wk	10.7	24.9	32.2	44.7
Agricultural hr/wk	4.4	13.9	18.8	17.3
Non-agricultural hr/wk	5.7	11.0	13.5	27.3
Casual labor, low skill hr/wk	1.0	1.5	2.3	10.9
Petty business, low skill hr/wk	1.3	3.5	3.5	7.6
Skilled Trades hr/wk	1.7	2.9	2.8	2.8
High-skill wage labor hr/wk	0.0	1.2	1.8	2.9
No employment hours in past month	0.5	0.1	0.0	0.0
Main occupation is non-agricultural	0.2	0.2	0.2	0.6
Engaged in any skilled trade	0.1	0.2	0.2	0.2

Notes: The table reports means in the control group, weighted by inverse probabilities of selection into attrition and endline tracking.

Figure C.1: Estimated net present value of intervention using interpolated earnings



Notes: The x-axis represents quantiles of net earnings after 9-years. The left y-axis shows the ITT estimate for that quantile in absolute terms. The right y-axis shows the ITT estimate of each quantile divided by the mean of the control group in that quantile.

We estimate that this temporary earnings gain is roughly $665 or 1.8 times the size of the grant (using 2008 USD and market exchange rates). We base this estimate on a linear interpolation of earnings treatment effects between the four survey rounds, discounted to the baseline year at an annualized interest rate of 5% — the standard social discount rate used in many policy evaluations. (Note: Profits were only measured a the 4-year endline. We model an immediate treatment effect of 18,000 UGX beginning at baseline and spanning 4 years, after which the treatment effect decreases linearly towards our 9-year estimate of 4,800 UGX.)

While imprecise, this estimate implies that the program was cost-effective in the sense of cumulative earnings gains for the directly treated exceeding the average grant size. This should not be taken as a measure of social impact for at least two reasons, however. First, the government has little to no data on implementation costs. Judging by other programs, it would not be unusual for these costs to exceed the value of the grant. Second, there may have been local spillovers, and parishes are too large relative to the grants to measure spillovers with precision. Some of these employment and earnings gains could have come from displacing the work of existing craftspersons. Alternatively, the demand shock could have stimulated local economies. Thus we do not have a full measure of social impacts.

Table C.2: Levels of, changes in, and program impacts on durbale business assets, after 2- and 4-years

	Business assets (000s of 2008 UGX)				
	Mean		Change 2-year to 4-year		
	2-year	4-year	Δ	%Δ	SD
	(1)	(2)	(3)	(4)	(5)
Treatment	725.8	607.8	-135.02	-19%	(83.3)
Control	290.2	392.8	109.9	38%	(53.5)**
ITT, with controls	377	225			
SD	(78.2)***	(62.6)***			
Treatment subgroups (% of total)					
Not funded (11%)	172.9	568.6	375.6	217%	(121.8)
Funded, did not train (22%)	331.4	446.5	91.7	28%	(106.4)
Funded, trained, not practicing in 2012 (29%)	1005.4	301.8	-720.8	-72%	(165.5)
Funded, trained, practicing in 2012 (38%)	1057.0	945.1	-75.4	-7%	(153.4)

Notes: Columns (1) and (2) report treatment and control group means at the 2 and 4-year endline surveys for the full sample. Below this we report the ITT of program assignment and robust standard errors clustered by group (calculated in the same way as the previous table). Column (3) reports the coefficient on a 4-year indicator in a regression of the dependent variable on the indicator and same controls as the ITT. This coefficient represents the change in the dependent variable over time. Columns (4) and (5) report the precentage change (the coefficient relative to the 2-year endline) and robust standard errors, clustered by group. * implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

Table C.3: Program impacts on business outcomes and migration after 9 years

				Treatment effects			
Dependent Variable	Control mean				By gender		
9-year endline	Pooled	Women	Men	Pooled	Women	Men	Difference
N = 1981	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Business outcomes							
Business index (sum)	0.47	0.36	0.52	0.044	0.041	0.045	0.004
				[0.040]	[0.061]	[0.050]	[0.076]
Business maintains formal records	0.21	0.16	0.24	0.030	-0.004	0.047	0.052
				[0.021]	[0.032]	[0.026]*	[0.040]
Business is formally registered	0.09	0.05	0.11	-0.005	0.016	-0.016	-0.032
				[0.013]	[0.021]	[0.017]	[0.027]
Business payes taxes	0.16	0.15	0.16	0.019	0.029	0.014	-0.016
				[0.018]	[0.031]	[0.022]	[0.036]
Has any employees (dummy)	0.57	0.48	0.61	0.032	0.023	0.037	0.015
				[0.023]	[0.042]	[0.028]	[0.050]
Number of employees, all trades	2.17	1.76	2.39	0.320	0.119	0.425	0.306
				[0.165]*	[0.265]	[0.210]**	[0.339]
Number of employees, non-agricultural	0.48	0.25	0.60	0.154	0.033	0.217	0.184
				[0.083]*	[0.086]	[0.112]*	[0.132]
Number of employees, skilled	0.12	0.02	0.17	0.139	0.035	0.193	0.158
				[0.042]***	[0.028]	[0.063]***	[0.069]**
Number of paid employees, all trades	1.14	0.88	1.27	0.257	0.102	0.338	0.237
				[0.136]*	[0.232]	[0.163]**	[0.279]
Number of paid employees, non-agricultural	0.33	0.14	0.43	0.113	0.018	0.163	0.145
				[0.067]*	[0.076]	[0.090]*	[0.113]
Number of paid employees, skilled	0.10	0.00	0.15	0.098	0.034	0.131	0.097
				[0.034]***	[0.025]	[0.051]***	[0.057]*
Number of family employees	0.93	0.76	1.02	0.088	0.119	0.072	-0.047
				[0.096]	[0.159]	[0.112]	[0.182]
Number of non-family employees	1.24	1.00	1.37	0.254	0.047	0.362	0.315
				[0.144]*	[0.239]	[0.177]**	[0.293]
Employee hours in past month, all trades	87.15	58.30	102.00	20.754	7.595	27.614	20.019
				[8.909]**	[12.720]	[11.097]**	[15.930]
Employee hours in past month, non-agricultural	37.88	19.92	47.12	6.305	-3.225	11.273	14.498
				[6.830]	[7.766]	[8.823]	[10.628]
Employee hours in past month, skilled	9.19	1.06	13.37	9.690	2.477	13.450	10.973
				[3.763]**	[2.588]	[5.255]**	[5.322]**
Total payroll in past month (000s of 2008 UGX)	44.67	20.09	57.32	0.846	-4.832	3.806	8.638
				[6.162]	[6.534]	[8.584]	[10.569]
Total payroll in past month, non-agricultural	24.38	5.49	34.11	0.401	-2.272	1.794	4.066
				[5.020]	[3.606]	[7.208]	[7.727]
Total payroll in past month, skilled	6.59	0.13	9.93	3.054	0.672	4.295	3.623
				[2.326]	[1.425]	[3.357]	[3.472]
Has savings account or participates in a savings group	0.69	0.70	0.69	-0.034	-0.060	-0.021	0.038
				[0.024]	[0.038]	[0.028]	[0.045]
Amount of money in savings (000s of 2008 UGX)	125.01	73.67	151.44	-9.167	26.920	-27.979	-54.899
				[12.362]	[14.235]*	[16.577]*	[20.798]***
Log savings	3.29	2.87	3.51	-0.022	0.171	-0.122	-0.293
				[0.106]	[0.167]	[0.123]	[0.192]
Migration outcomes							
Lives in a town or city	0.73	0.71	0.74	0.013	0.022	0.008	-0.014
				[0.021]	[0.034]	[0.025]	[0.040]
Moved from a village to a town/city since Baseline	0.65	0.65	0.66	0.015	0.042	-0.000	-0.042
				[0.017]	[0.028]	[0.020]	[0.033]
Population of town (categorical variable)	3.75	3.71	3.78	-0.006	0.040	-0.029	-0.069
				[0.045]	[0.076]	[0.054]	[0.090]

Notes: Each entry in column 4 is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment, district fixed effects, and a vector of baseline covariates. Standard errors are clustered at the group level (of up to 5 people). We report the coefficient on treatment only. All regressions are weighted by inverse probabilities of attrition and selection into the endline tracking sample. Control means in columns 1, 2, and 3 also use this weight.

Columns 5 through 7 come from a single regression, where we regress treatment on an indicator for treatment, an indicator interacted with gender, and the gender indicator plus the usual controls. We use these coefficients to calculate the treatment effect on each gender separately.

* implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

Table C.4: Program impacts on wages per hour

Dependent Variable N = 1829 in 2-year N = 1808 in 4-year N = 1928 in 9-year	Control mean 2-year (1)	 4-year (2)	 9-year (3)	Treatment effects 2-year (4)	 Difference (5)	 4-year (6)	 Difference (7)	 9-year (8)
Average earnings/hr (000s of 2008 UGX)	0.51	0.39	0.63	0.031 [0.069]	0.041 [0.078]	0.071 [0.037]*	-0.062 [0.072]	0.009 [0.062]
Agricultural earnings/hr	0.16	0.12	0.18	0.035 [0.032]	-0.033 [0.038]	0.002 [0.020]	0.004 [0.051]	0.006 [0.047]
Non-agricultural earnings/hr	1.32	0.94	1.01	-0.361 [0.152]**	0.458 [0.187]***	0.097 [0.109]	-0.054 [0.146]	0.043 [0.098]
Casual labor, low skill earnings/hr	1.36	1.00	0.24	0.559 [0.394]	-0.703 [0.416]*	-0.143 [0.131]	0.148 [0.139]	0.004 [0.048]
Petty business, low skill earnings/hr	3.30	3.21	4.75	1.850 [1.130]	1.180 [2.037]	3.031 [1.695]*	-2.820 [2.101]*	0.211 [1.242]
Skilled Trades earnings/hr	0.64	0.61	0.95	-0.027 [0.103]	0.145 [0.123]	0.117 [0.067]*	0.201 [0.154]*	0.318 [0.139]**
High-skill earnings/hr	0.72	2.05	1.60	-0.220 [0.000]	-0.078 [0.000]***	-0.298 [0.000]	0.371 [0.336]	0.073 [0.336]

Notes: Each entry in columns 4 and 6, and 8 is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment, district fixed effects, and a vector of baseline covariates. Standard errors are clustered at the group level (of up to 5 people). We report the coefficient on treatment only. All regressions are weighted by inverse probabilities of attrition and selection into the endline tracking sample. Control means in columns 1, 2, and 3 are also calculated using these weights. Columns 5 and 7 refer to the difference between coefficients between endlines. P-values for diferences are calculated using a simple t-test using the standard errors of coefficients.

* implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

Table C.5: 9-year program impacts on respondent's own physical and mental health

Dependent Variable N = 2086 9-year endline	Mean Control (1)	 Treated (2)	 ITT (3)
Respondent passed away	0.01	0.00	-0.004 [0.006]
Physical health index (z-score)	-0.03	-0.02	-0.028 [0.047]
Difficulty in hysical functioning at activities of daily life (z-score	-0.01	-0.03	-0.048 [0.048]
Inability perform activities of daily life index (z-score)	0.00	-0.02	-0.077 [0.049]
Tiredness index (z-score)	-0.02	-0.03	-0.003 [0.048]
Illness and injury index (z-score)	-0.01	-0.01	-0.036 [0.049]
Days unable to work in past month	1.29	1.40	0.080 [0.155]
Mental health index (z-score)	0.01	0.00	-0.056 [0.047]
Depression and distress index (z-score)	-0.05	0.04	-0.034 [0.046]
Depression index (z-score)	-0.05	0.04	-0.040 [0.045]
Distress index (z-score)	-0.04	0.03	-0.023 [0.047]
Pro-social behavior index (z-score)	-0.05	0.03	0.022 [0.050]
Hostility index (z-score	0.02	0.00	-0.064 [0.046]

Notes: Column 3 is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment, district fixed effects, and a vector of baseline covariates. Standard errors are clustered at the group level (of up to 5 people). We report the coefficient on treatment only. All regressions are weighted by inverse probabilities of attrition and selection into the endline tracking sample. Means in columns 1 and 2 are setimated using these weights as well.

For each index, missing values were imputed to be the row median of the measures that compose the index, per respondent. Maximum N across all regressions is 2086 because the outcome *Has died* is measured for individuals found, dead or alive, including those who did not complete the full survey. $N = 1981$ for all other regressions.

Physical health is made up of a standardized additive index of (1) a respondent's *difficulty* in performing the Activities of Daily Life (ADL), such as walking, running, going up stairs, or fetching water, as well as how tired they fell from performing such tasks. (2) A additive standardized index of three binary variables for whether a respondent is injured, disabled, or has a chronic illness. And (3) a standardized measure of the days a respondent was unable to work in the past month.

The mental health index is composed of (1) symptoms of depression, for example feeling lonely and not valueing life, (2) symptoms of distress, such as anxiety and discomfort, (4) an (inverse of) pro-social behavior symptoms, such as feeling empathy and helping others, and (5) symptoms of hostility, such as destroying objects and fighting with others.
* implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

Table C.6: Program impacts on compliance, earnings, and employment, gender disaggregated, 4-year

Dependent Variable 4-year endline N = 1868	Control mean Pooled (1)	 Women (2)	 Men (3)	Treatment effects Pooled (4)	By gender Women (5)	 Men (6)	 Difference (7)
Panel A: Compliance & initial investments							
Business assets (000s 2008 UGX)	392.79	153.17	535.40	223.186 [62.679]***	163.410 [54.430]***	255.159 [89.660]***	91.748 [101.706]
Reported a major non-YOP program since 2006							
Panel B: Income							
Standardized Income Index	-0.08	-0.25	0.02	0.224 [0.050]***	0.257 [0.080]***	0.206 [0.061]***	-0.052 [0.097]
Monthly net earnings (000s of 2008 UGX)	47.85	25.46	61.17	18.163 [4.887]***	18.643 [7.207]***	17.907 [6.276]***	-0.736 [9.355]
Nondurable Consumption (000s of 2008 UGX)	202.22	191.84	208.40	21.688 [7.946]***	27.631 [13.215]**	18.509 [9.428]*	-9.122 [15.664]
Durable assets	0.09	-0.01	0.14	0.196 [0.050]***	0.231 [0.085]***	0.177 [0.059]***	-0.053 [0.102]
Panel C: Employment							
Average employment hrs/wk	32.24	23.96	37.19	5.519 [1.286]***	7.157 [2.044]***	4.641 [1.598]***	-2.516 [2.538]
Agricultural hrs/wk	18.77	15.44	20.75	0.422 [0.945]	0.985 [1.486]	0.120 [1.188]	-0.866 [1.872]
Non-agricultural hrs/wk	13.48	8.52	16.43	5.097 [0.999]***	6.172 [1.439]***	4.521 [1.302]***	-1.651 [1.907]
Casual labor, low skill hrs/wk	2.27	1.03	3.00	-0.117 [0.401]	0.335 [0.428]	-0.360 [0.562]	-0.694 [0.692]
Petty business, low skill hrs/wk	3.54	3.18	3.75	0.138 [0.593]	1.449 [0.987]	-0.565 [0.711]	-2.014 [1.185]*
Skilled Trades hrs/wk	2.82	1.97	3.32	3.779 [0.548]***	3.224 [0.695]***	4.077 [0.718]***	0.854 [0.947]
High-skill wage labor hrs/wk	1.84	0.84	2.44	0.898 [0.444]**	1.360 [0.658]**	0.650 [0.557]	-0.709 [0.829]
No employment hours in past month	0.05	0.09	0.02	-0.022 [0.009]***	-0.056 [0.020]***	-0.004 [0.009]	0.052 [0.022]**
Works over 30 hrs/wk in skilled trade	0.03	0.01	0.04	0.037 [0.013]***	0.025 [0.015]*	0.044 [0.017]***	0.019 [0.021]

Notes: Each entry in column 4 is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment, district fixed effects, and a vector of baseline covariates. Standard errors are clustered at the group level (of up to 5 people). We report the coefficient on treatment only. All regressions are weighted by inverse probabilities of attrition and selection into the endline tracking sample. Control means in columns 1, 2, and 3 also use this weight.

Columns 5 through 7 come from a single regression, where we regress treatment on an indicator for treatment, an indicator interacted with gender, and the gender indicator plus the usual controls. We use these coefficients to calculate the treatment effect on each gender separately.

* implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

Table C.7: Program impacts on compliance, earnings, and employment, gender disaggregated, 9-year

Dependent Variable 9-year endline N = 1981	Control mean Pooled (1)	 Women (2)	 Men (3)	Treatment effects Pooled (4)	By gender Women (5)	 Men (6)	 Difference (7)
Panel A: Compliance & initial investments							
Business assets (000s 2008 UGX)							
Reported a major non-YOP program since 2006	0.17	0.15	0.17	-0.029 [0.018]	-0.028 [0.032]	-0.029 [0.022]	-0.001 [0.038]
Panel B: Income							
Standardized Income Index	-0.02	-0.17	0.06	0.078 [0.048]	0.096 [0.076]	0.068 [0.059]	-0.027 [0.093]
Monthly net earnings (000s of 2008 UGX)	90.97	57.44	108.23	4.172 [8.491]	7.064 [13.230]	2.664 [10.022]	-4.400 [15.457]
Nondurable Consumption (000s of 2008 UGX)	190.56	182.55	194.68	2.726 [6.298]	10.728 [10.322]	-1.445 [7.528]	-12.172 [12.283]
Durable assets	0.25	0.16	0.30	0.145 [0.047]***	0.107 [0.080]	0.165 [0.058]***	0.058 [0.098]
Panel C: Employment							
Average employment hrs/wk	44.68	46.64	43.67	0.513 [1.593]	1.897 [3.079]	-0.208 [1.705]	-2.105 [3.410]
Agricultural hrs/wk	17.34	14.55	18.77	0.079 [0.856]	1.354 [1.342]	-0.585 [1.095]	-1.940 [1.728]
Non-agricultural hrs/wk	27.35	32.09	24.90	0.434 [1.488]	0.542 [2.756]	0.377 [1.547]	-0.165 [2.949]
Casual labor, low skill hrs/wk	10.93	17.15	7.73	-1.206 [0.990]	-1.015 [1.750]	-1.306 [1.095]	-0.292 [1.952]
Petty business, low skill hrs/wk	7.59	10.03	6.34	-1.589 [1.069]	-3.154 [1.947]	-0.773 [1.077]	2.381 [2.028]
Skilled Trades hrs/wk	2.83	1.90	3.30	2.796 [0.529]***	3.590 [0.886]***	2.383 [0.606]***	-1.207 [1.013]
High-skill wage labor hrs/wk	2.93	2.02	3.40	0.906 [0.582]	0.893 [0.921]	0.913 [0.712]	0.020 [1.125]
No employment hours in past month	0.03	0.03	0.03	-0.004 [0.008]	-0.009 [0.013]	-0.002 [0.010]	0.006 [0.017]
Works over 30 hrs/wk in skilled trade	0.03	0.02	0.04	0.029 [0.011]***	0.042 [0.018]**	0.023 [0.013]*	-0.020 [0.021]

Notes: Each entry in column 4 is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment, district fixed effects, and a vector of baseline covariates. Standard errors are clustered at the group level (of up to 5 people). We report the coefficient on treatment only. All regressions are weighted by inverse probabilities of attrition and selection into the endline tracking sample. Control means in columns 1, 2, and 3 also use this weight.

Columns 5 through 7 come from a single regression, where we regress treatment on an indicator for treatment, an indicator interacted with gender, and the gender indicator plus the usual controls. We use these coefficients to calculate the treatment effect on each gender separately.

* implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

Table C.8: 9-year program impacts on own health, child mortality and fertility, gender disaggregated

Dependent Variable 9-year endline N = 2086	Control mean Pooled (1)	 Women (2)	 Men (3)	Treatment effects Pooled (4)	By gender Women (5)	 Men (6)	 Difference (7)
Panel A: Own health outcomes							
Respondent passed away	0.01	0.00	0.01	-0.004 [0.006]	0.001 [0.007]	-0.007 [0.006]	-0.009 [0.006]
Physical health index (z-score)	-0.03	0.26	-0.17	-0.028 [0.047]	-0.017 [0.092]	-0.034 [0.051]	-0.017 [0.102]
Mental health index (z-score)	0.01	0.33	-0.16	-0.056 [0.047]	-0.122 [0.082]	-0.023 [0.055]	0.099 [0.097]
Panel B: Fertility, household size, and child expenditures							
Number of pregnancies 2007 or later	2.47	2.15	2.64	0.097 [0.101]	-0.031 [0.152]	0.164 [0.121]	0.195 [0.182]
Pct. of births that were live 2007 or later	0.92	0.89	0.94	0.013 [0.010]	0.019 [0.019]	0.011 [0.012]	-0.008 [0.022]
Pct. of pregnancies 2007 or later where child still living	0.89	0.87	0.90	0.009 [0.012]	0.017 [0.020]	0.005 [0.014]	-0.012 [0.025]
Pct. of successful pregnancies 2007 or later where child still living	0.96	0.97	0.96	-0.006 [0.006]	-0.001 [0.010]	-0.008 [0.008]	-0.006 [0.013]
Number of biological children alive born 2007 or later	2.14	1.83	2.30	0.075 [0.083]	-0.006 [0.121]	0.116 [0.100]	0.122 [0.144]
Size of household	5.86	6.15	5.72	-0.127 [0.162]	-0.192 [0.269]	-0.093 [0.183]	0.099 [0.301]
Mean age of children (0-15)	7.50	7.64	7.43	0.014 [0.138]	-0.007 [0.233]	0.025 [0.164]	0.032 [0.277]
Mean age of biological children (0-15)	7.08	7.33	6.95	0.095 [0.147]	-0.084 [0.243]	0.189 [0.173]	0.273 [0.286]
Panel C: Child educational outcomes							
Child age-adjusted educational attainment (6-24)	0.01	0.07	-0.02	-0.012 [0.037]	0.056 [0.061]	-0.047 [0.047]	-0.103 [0.076]
Child age-adjusted educational attainment, (6-24, biological)	0.09	0.19	0.04	-0.046 [0.045]	0.014 [0.067]	-0.079 [0.056]	-0.094 [0.084]
Mean of child enrollment	0.91	0.90	0.91	-0.016 [0.013]	-0.019 [0.021]	-0.015 [0.016]	0.005 [0.026]
Mean of child enrollment, biological	0.91	0.91	0.91	-0.018 [0.013]	-0.028 [0.022]	-0.013 [0.017]	0.016 [0.027]
Current child expenditures (clothes and school)	42.14	50.01	38.09	0.411 [2.784]	-2.241 [5.190]	1.793 [2.974]	4.034 [5.698]
Current child expenditures per child	14.00	16.88	12.48	0.502 [1.071]	-0.587 [2.275]	1.081 [1.139]	1.668 [2.573]
Panel D: Child health outcomes							
Mean health index per child, ages 3-9, family average	-0.03	-0.07	-0.01	0.078 [0.043]*	0.166 [0.076]**	0.036 [0.051]	-0.130 [0.092]
Mean parent-reported health score per child, for ages 3-9, family ave	0.00	-0.06	0.03	0.071 [0.047]	0.144 [0.087]*	0.036 [0.055]	-0.108 [0.101]
Mean malaria cases in past year, ages 3-9, family average	2.96	2.97	2.96	-0.125 [0.087]	-0.163 [0.153]	-0.107 [0.108]	0.055 [0.190]
Mean normalized ADL score per child, ages 3-9, family average	0.01	-0.03	0.03	0.045 [0.041]	0.145 [0.071]**	-0.003 [0.049]	-0.149 [0.085]*

Notes: Each entry in column 4 is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment, district fixed effects, and a vector of baseline covariates. Standard errors are clustered at the group level (of up to 5 people). We report the coefficient on treatment only. All regressions are weighted by inverse probabilities of attrition and selection into the endline tracking sample. Control means in columns 1, 2, and 3 also use this weight.

Columns 5 through 7 come from a single regression, where we regress treatment on an indicator for treatment, an indicator interacted with gender, and the gender indicator plus the usual controls. We use these coefficients to calculate the treatment effect on each gender separately.

* implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

Table C.9: Treatment Heterogeneity by initial working capital, patience, and engagement in a skilled trade

	Standardized Income Index			
Independent Variable	Est. Coeff.	SE	Est. Coeff.	SE
Assigned to treatment	0.031	0.065	0.024	0.065
Female	-0.226	0.080***	-0.100	0.081
Assigned to treatment x Female	0.108	0.102	0.053	0.101
Engaged in a Skilled Trade			0.159	0.183
Assigned to treatment x Engaged in a Skilled Trade			0.035	0.220
Working capital index (z-score)			0.116	0.046**
Assigned to treatment x Working capital index (z-score)			0.020	0.053
Human capital index (z-score)			0.217	0.039***
Assigned to treatment x Human capital index (z-score)			-0.044	0.053
Patience index (z-score)			0.073	0.032**
Assigned to treatment x Patience index (z-score)			-0.011	0.045

This table examines sources of heterogeneity by baseline characteristics, including prior engagement in a skilled trade, initial physical and human capital, and rates of time preference. Each pair of columns reports a separate regression. Estimates include district fixed effects and group level clustering. Observations are weighted by the inverse of the probability of being found in Endline 3 tracking. In principle, returns to grants should be higher among those with lower initial physical capital, no initial business, higher human capital, and the more patient (see Blattman et al. (2014) for these comparative statics). We see no statistically significant evidence of such heterogeneity, perhaps because the sample was selected to be relatively homogenous and the remaining within-sample variation is less relevant. We also examined heterogeneity according to exposure to economic shocks, such as major rainfall shocks, but given the non-farm nature of most employment we generally do not see a strong rainfall-income relationship in this sample (not shown). * implies $P < .1$ ** implies $P < .05$ *** implies $P < .01$

D Robustness and sensitivity analysis

These conclusions are fairly robust to different estimation models and measurement decisions. Appendix Table D.10 reports the original results and 9 tests of robustness for the three measures of income and a measure of employment.[18] We report the 9-year treatment effect as a percentage of the control mean for each estimate.

- Column 1 reproduces original ITT estimates.
- Column 2 reports a difference-in-difference estimate where baseline data are available.
- Columns 3 and 4 report treatment effects unadjusted by control variables but including district fixed effects. Column 3 is a difference-in-difference estimate and Column 4 is an ITT estimate.
- Column 5 is an ITT estimate controlling for baseline variables but without inverse probability weights for attrition.
- Column 6 changes the unit of clustering to the parish rather than the group.
- Columns 7 to 10 display a bounding exercise. We model treatment effects under extreme assumptions about the unfound. It is common to trim distributions and recalculate treatment effects. But such "Lee bounds" only correct for differential attrition by treatment arm.

  Instead, we calculate bounds in the spirit of Manski bounds: we recalculate the ITT where unfound respondents are imputed with the mean of each variable for found controls plus or minus 0.1 standard deviations (Columns 7 and 8), and 0.25 standard deviations (Columns 9 and 10). Even the +/-0.1 bounds are extreme, as they imply a 0.2 standard deviation gap between the outcomes of unfound treatment and control group members—far more selective attrition than we observed in section B.3 or Appendix Table B.3.

  With these extreme scenarios, we can obtain a statistically significant treatment effect, but these estimates are not economically significant. Take consumption. With a 0.25 standard deviation difference between the treatment and control group attritors, the ITT would only be a 6% increase, and for a 0.25 standard deviation difference the ITT would only be an 11% increase. These attrition scenarios are highly unlikely and we

[18] Appendix Table D.12 reports results without capping variables at the 99th percentile and Appendix Table D.13 shows that results are unaffected by restricting attention to the 1546 individuals found in both survey rounds.

report them primarily for full transparency. Even the upper bound on monthly earnings (Column 8) is UGX 15,084. This is below the average 4-year earnings treatment effect of UGX 18,186.

Appendix Table D.12 shows program impacts on pre-specified economic outcomes without any capping. While our main specification caps earnings, consumption, and assets variables at the 99th percentile, the existence of high earners in either the treatment or control group may still be economically meaningful. While the combined income measure is now significant at the 5% level, this is due to the program impact on durable assets, which is also significant in Table 1. Appendix Figure D.2 shows heterogeneity by individual's position in the distribution of income at endline by estimating quantile treatment effects. We do not see substantive variation in treatment effects according to this metric, though it is plausible that those with higher incomes saw slightly more benefits. It is clear that a large portion of our sample has zero or almost zero income, so ITTs as a percentage of control mean can be quite large, even if the absolute magnitude of the effect is effectively zero. Appendix Table D.13 deals with the issue that due to attrition and non-response, the populations analyzed at the 4-year endline and 9-year endline are not the same.

Table D.10: Sensitivity analysis of 9-year program impacts on economic outcomes to alternate models and missing data scenarios

	Main specification	Difference in Differences	DiD, omit covariates	ITT, omit covariates	Omit selective attrition weights	Cluster by parish	+/- .1 SD		+/- .25 SD	
							control outperforms	treated outperforms	control outperforms	treated outperforms
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Monthly net earnings (000s of 2008 UGX)	4.172	5.357	5.698	7.712	4.764	3.991	-1.451	15.084	-13.853	27.485
	[8.491]	[9.780]	[9.848]	[8.843]	[8.325]	[9.006]	[6.200]	[6.278]**	[6.209]**	[6.401]***
	4.6%	5.9%	6.3%	8.5%	5.2%	4.4%	-1.6%	16.6%	-15.2%	30.2%
Nondurable Consumption (000s of 2008 UGX)	2.726			8.015	4.685	1.592	-2.238	11.209	-12.324	21.295
	[6.298]			[6.645]	[6.465]	[6.185]	[4.722]	[4.755]**	[4.754]***	[4.837]***
	1.4%			4.2%	2.5%	0.8%	-1.2%	5.9%	-6.5%	11.2%
Durable assets	0.145	0.091	0.120	0.192	0.161	0.141	0.086	0.194	0.006	0.275
	[0.047]***	[0.058]	[0.061]*	[0.053]***	[0.047]***	[0.048]***	[0.036]**	[0.036]***	[0.037]	[0.037]***
Average employment hrs/wk	0.513	1.036	-0.433	0.879	0.962	0.492	-0.784	2.237	-3.050	4.503
	[1.593]	[1.656]	[1.875]	[1.645]	[1.526]	[1.533]	[1.192]	[1.193]*	[1.203]**	[1.205]***
	1.1%	2.3%	-1.0%	2.0%	2.2%	1.1%	-1.8%	5.0%	-6.8%	10.1%

Notes: Each entry is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment and district fixed effects. Column 1 reports the estimated ITT program effects from the previous tables. Subsequent columns omit baseline covariates, calculate difference-in-difference ITTs, or omit inverse probability regression weights generated by our LOO logit regressions. We also show clustering by Parish as opposed to group. Finally, the last four columns omit attrition weights and instead impute semi-extreme values of the dependent variable, so that missing treatment and control group members have a 0.2 or .25 standard deviation difference in their outcomes (as we impute a mean + .1 or .25 standard deviation outcome for treatment and mean – .1 or .25 standard deviatioon for controls, or vice versa). * implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

Table D.11: Sensitivity analysis of 9-year program impacts on health outcomes to alternate models and missing data scenarios

	Main specification (1)	ITT, omit covariates (2)	Omit selective attrition weights (3)	Cluster by parish (4)	+/- .1 SD control outperforms (5)	+/- .1 SD treated outperforms (6)	+/- .25 SD control outperforms (7)	+/- .25 SD treated outperforms (8)
Respondent passed away	-0.004	-0.004	-0.003	-0.004	-0.003	-0.003	-0.003	-0.003
	[0.006]	[0.006]	[0.005]	[0.006]	[0.008]	[0.008]	[0.008]	[0.008]
Physical health index (z-score)	-0.028	-0.030	-0.035	-0.031	-0.072	0.032	-0.150	0.110
	[0.047]	[0.050]	[0.048]	[0.045]	[0.035]**	[0.035]	[0.035]***	[0.036]***
Mental health index (z-score)	-0.056	-0.057	-0.048	-0.059	-0.088	0.016	-0.166	0.094
	[0.047]	[0.050]	[0.046]	[0.046]	[0.035]**	[0.035]	[0.035]***	[0.036]***
Child age-adjusted educational attainment (6-15)	-0.017	-0.011	-0.021	-0.016	-0.072	0.050	-0.164	0.142
	[0.043]	[0.047]	[0.043]	[0.043]	[0.027]***	[0.028]*	[0.028]***	[0.028]***
Mean health index per child, ages 3-9, family average	0.078	0.093	0.071	0.079	0.006	0.158	-0.109	0.273
	[0.043]*	[0.045]**	[0.043]	[0.042]*	[0.025]	[0.025]***	[0.025]***	[0.025]***

Notes: Each entry is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment and district fixed effects. Column 1 reports the estimated ITT program effects from the previous tables. Subsequent columns omit baseline covariates or omit inverse probability regression weights generated by our LOO logit regressions. We also show clustering by Parish as opposed to group. Finally, the last four columns omit attrition weights and instead impute semi-extreme values of the dependent variable, so that missing treatment and control group members have a 0.2 or .25 standard deviation difference in their outcomes (as we impute a mean + .1 or .25 standard deviation outcome for treatment and mean – .1 or .25 standard deviatioon for controls, or vice versa). * implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$

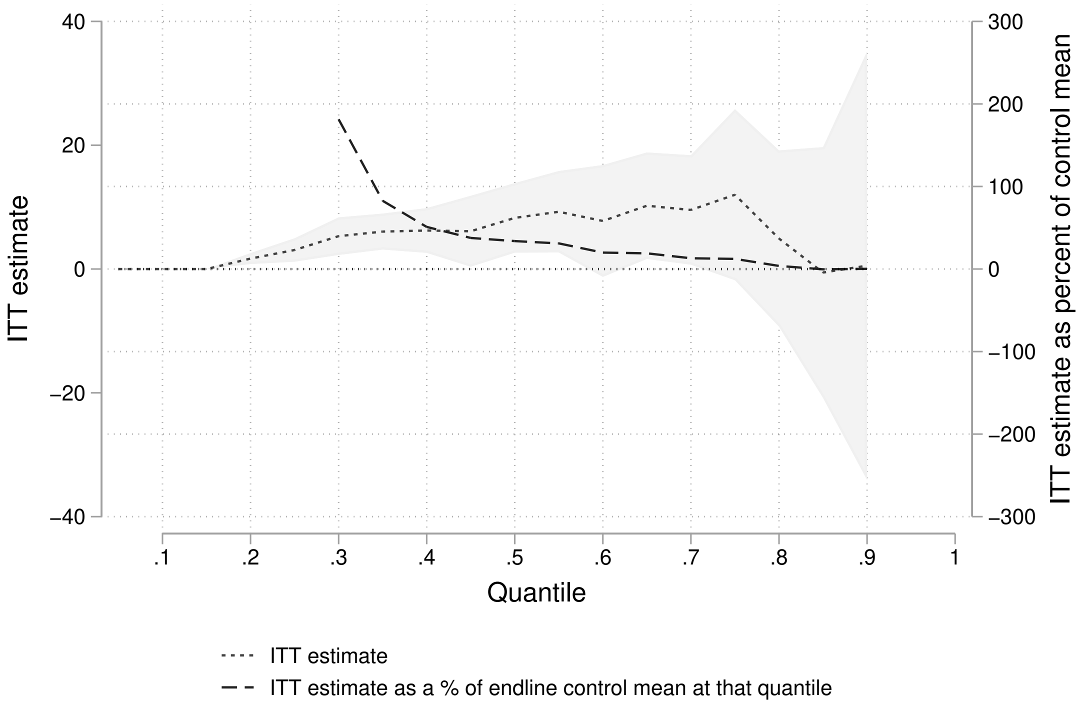
Table D.12: Effect of treatment on income index, without any capping

Dependent Variable N = 1868 in 4-year N = 1981 in 9-year	Control mean 4-year (1)	 9-year (2)	Treatment effects 4-year (3)	 Difference (4)	 9-year (5)
Standardized Income Index (no capping)	-0.07	-0.03	0.191 [0.051]***	-0.072 [0.076]	0.119 [0.057]**
Monthly net earnings (000s of 2008 UGX, no capping)	88.01	193.22	36.275 [19.590]*	27.494 [67.355]	63.769 [64.444]
Nondurable Consumption (000s of 2008 UGX, no capping)	329.50	395.41	37.855 [14.803]**	-18.701 [22.951]	19.153 [17.539]
Durable assets (z-score, no capping)	0.09	0.25	0.196 [0.050]***	-0.035 [0.068]	0.161 [0.047]***

Several of our outcomes have a long upper tail, and some of these large values are potentially due to enumeration errors. Extreme values will be highly influential in any treatment effect estimate, and so we top-code all currency-denominated, hours worked, and employee variables at the 99th percentile. This table reports sensitivity to this top-coding procedure.

Each entry in columns 3 and 4 is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment, district fixed effects, and a vector of baseline covariates. Standard errors are clustered at the group level (of up to 5 people). We report the coefficient on treatment only. All regressions are weighted by inverse probabilities of attrition and selection into the endline tracking sample. * implies $p < .1$, ** implies $p < .05$, *** implies $p < .01$.

Figure D.2: Quantile estimates on net earnings.



Notes: We examine heterogeneity by individual's position in the distribution of income at endline by estimating quantile treatment effects. The x-axis represents quantiles of net earnings after 9-years. The left y-axis shows the ITT estimate for that quantile in absolute terms. The right y-axis shows the ITT estimate of each quantile divided by the mean of the control group in that quantile. We do not see substantive variation in treatment effects according to this metric, though it is plausible that those with higher incomes saw slightly more benefits. It is clear that a large portion of our sample has zero or almost zero income, so ITTs as a percentage of control mean can be quite large, even if the absolute magnitude of the effect is effectively zero.

Table D.13: Effect of treatment on pre-specified variables with only respondents present at both 4- and 9-year follow-ups

Dependent variable	Control means				Treatment effects			
N = 1868 in 4-year	4-year		9-year		4-year		9-year	
N = 1981 in 9-year	Full	Subset	Full	Subset	Full	Subset	Full	Subset
N = 1546 in subset	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Monthly net earnings (000s of 2008 UGX)	47.85	42.75	90.97	83.93	18.163	18.534	4.172	1.618
					[4.887]***	[4.845]***	[8.491]	[8.998]
Nondurable Consumption (000s of 2008 UGX)	202.22	199.54	190.56	187.12	21.688	19.673	2.726	6.853
					[7.946]***	[8.815]**	[6.298]	[6.997]
Durable assets	0.09	0.07	0.25	0.23	0.196	0.184	0.145	0.179
					[0.050]***	[0.053]***	[0.047]***	[0.049]***
Average employment hrs/wk	32.24	32.49	44.68	45.41	5.519	5.610	0.513	0.557
					[1.286]***	[1.406]***	[1.593]	[1.675]

Regression estimates include district fixed effects and group level clustering. *Subset* columns 2, 4, 6, and 8 restrict all estimates to individuals present at both 4- and 9-year endlines. All other estmiates in columns 1, 3, 5, and 7 are identical to results in our main specification. * implies $p < .1$ ** implies $p < .05$ *** implies $p < .01$

E Political impacts

In Blattman et al. (2018a), we found that 3 years after YOP, during the 2011 national elections, YOP recipients reported that they were no more likely to engage in non-partisan political actions, but they were significantly more likely to have supported and voted for opposition candidates. We hypothesized that this was partly due to clientelistic election tactics, where villagers are commonly offered small sums of money to vote, especially on behalf of the ruling party (Larreguy et al., 2018). By earning more income, individuals in the treatment group were freed to publicly demonstrate their support for the opposition.

Appendix Table E.14 reports program impacts on political behavior after 4 and 9 years. We see similar but less statistically significant patterns in the 2016 elections, held 8 years after YOP grants were disbursed. We see little treatment-control difference in political actions not specific to a party, and we see no change in self-reported support for the ruling party. But YOP recipients were 0.08 standard deviations more likely to support the opposition party, such as join the party, vote for the opposition candidate, or actively work to get an opposition candidate elected. This family index is significant at the 10 percent level only, and only slightly smaller than the previous effect of 0.12 standard deviations. Nonetheless, some of the component treatment effects are very large: actively working to get an opposition candidate elected rises 23% relative to the control mean, and voting for the opposition rises 14% relative to the control mean.

The fact that political behavior is more persistent than the income effects could mean that separation from clientelistic networks is not what drives the political behavior change. Of course, changes in political behavior could be persistent if party affiliation in young adulthood is habit-forming or otherwise persistent.

Table E.14: 9-year program impacts on political behavior

Dependent Variable N = 1858 in 4-year N = 1981 in 9-year	Control mean 4-year (1)	 9-year (2)	Treatment effects 4-year (3)	 Difference (4)	 9-year (5)
Index of political action (z-score)	-0.02	-0.06	0.013 [0.012]	0.043 [0.051]	0.056 [0.050]
Attended voter education meeting	0.48	0.54	0.026 [0.026]	0.004 [0.035]	0.030 [0.024]
Discussed Vote	0.56	0.69	-0.028 [0.025]	0.025 [0.035]	-0.003 [0.024]
Reported campaign malpractice or incident	0.10	0.08	0.024 [0.017]	-0.037 [0.021]*	-0.013 [0.012]
Voted in presidential election	0.91	0.92	0.001 [0.014]	0.007 [0.019]	0.007 [0.013]
Attended political rally	0.68	0.69	0.002 [0.024]	0.010 [0.035]	0.012 [0.025]
Participated in political primary	0.43	0.47	0.011 [0.025]	0.003 [0.034]	0.014 [0.024]
Worked to get a candidate/party elected	0.34	0.50	0.045 [0.024]*	-0.008 [0.034]	0.037 [0.025]
Member of a political party	0.49	0.56	-0.001 [0.026]	0.046 [0.036]*	0.045 [0.024]*
Index of NRM/Presidential support (z-score)	-0.05	-0.04	-0.041 [0.052]	0.071 [0.074]	0.030 [0.052]
Would vote for NRM if election were tomorrow	0.75	0.77	-0.019 [0.022]	0.009 [0.031]	-0.009 [0.021]
Like or strongly like NRM	0.81	0.81	-0.023 [0.020]	0.037 [0.028]*	0.014 [0.020]
Feels close to the NRM	0.55		0.011 [0.024]		
Worked to get the NRM elected	0.29	0.43	0.014 [0.023]	0.006 [0.034]	0.020 [0.025]
Member of the NRM	0.40	0.51	-0.017 [0.026]	0.054 [0.036]*	0.037 [0.024]
Voted or supported the president in the last election	0.88	0.81	-0.038 [0.018]**	0.033 [0.027]	-0.006 [0.020]
Approve or strongly approve of President	0.85		-0.020 [0.018]		
Index of opposition support (z-score)	-0.00	0.01	0.117 [0.053]**	-0.035 [0.069]	0.082 [0.044]*
Would vote for opposition if election were tomorrow	0.17	0.12	0.010 [0.020]	0.008 [0.025]	0.017 [0.016]
Like or strongly like any opposition party	0.36	0.35	0.030 [0.023]	-0.000 [0.032]	0.029 [0.022]
Feels close to any opposition party	0.10		0.032 [0.016]**		
Worked to get the opposition elected	0.04	0.06	0.031 [0.011]***	-0.017 [0.015]	0.014 [0.010]
Member of an opposition party	0.05	0.53	0.025 [0.013]**	0.014 [0.027]	0.040 [0.024]
Voted or supported an election party in the past election	0.12	0.28	0.038 [0.018]**	-0.045 [0.029]*	-0.006 [0.022]

Notes: Each entry in columns 3 and 5 is estimated from a weighted least squares regression of the dependent variable on an indicator for assignment to treatment, district fixed effects, and a vector of baseline covariates. Standard errors are clustered at the group level (of up to 5 people). We report the coefficient on treatment only. All regressions are weighted by inverse probabilities of attrition and selection into the endline tracking sample. Control means in columns 1 and 2 are also calculated using these weights.
Column 4 refers to the difference between coefficients between endlines. P-values for diferences are calculated using a simple t-test using the standard errors of coefficients.

* implies $p < .1$, ** implies $p < .05$, *** implies $p < .0$