International Development Dissertation Cover Sheet 2020/21

Course Code: DV410 **Candidate Number:**

1	8	1	9	5

Title: Heterogeneous Impacts of Cash Transfers: Applying Machine Learning Methods to Evidence from Uganda

Due Date: 12 pm (Midday) UK time, 23 August 2021

No. of Words: 9995

(including footnotes and/or endnotes; not including compulsory headers, table of contents, title, compulsory abstract, appendices, list of abbreviations or bibliography)

MSc Programme: MSc in Development Studies

Dissertations MUST be submitted with a header in the following format:

DV410 **Page x of y** **Candidate Number**

Exceeding the word limit will AUTOMATICALLY be penalised at the rate of 1 mark per 100 words or part thereof. Word limit is 10,000 with 1% over allowed (i.e 10,100 absolute maximum). Submitting work late without a previously arranged extension will AUTOMATICALLY be penalised at the rate of 5 marks per day or part thereof.

Plagiarism (unacknowledged borrowing, quotation) including self-plagiarism is an examination offence and carries heavy penalties. In submitting his/her essay the student acknowledges it as his/her own work which contains no plagiarism AND that it has not been submitted previously for any assessed unit on the course or any other degree course at the LSE or elsewhere.

Please be sure:
You have uploaded Word or .pdf files as required on the Moodle DV410 page by the deadline of 12 pm (mid-day) on 23 August 2021, or secured an extension.

You have read and complied with all submission requirements for this course as detailed on the Moodle DV410 page.

Please mark 'X' in the box on the right to confirm that you have read, understood and adhered to the conditions stated above. X

Dissertation Release Consent Form

Candidate. No. __18195________________

IMPORTANT:

We would like to make your dissertation available to future students and as an example submission.

International Development undertakes that, if copies are made available, they will be in PDF format only in order to protect your work.

International Development undertakes that dissertations will only ever bear the candidate number. Student names will NEVER be released without specific permission of the writer.

I, candidate number ___18195______________ hereby give permission for my dissertation to be made available to other students/members of LSE as reference material.

Should you not wish for your work to be utilised in this way, you are under no obligation to do so and you will not be affected in any way by making that choice.

If you wish to <u>opt out</u> of giving permission, please mark 'X' in the box ☐

Candidate Number: 18195

MSc in Development Studies 2021

Dissertation submitted in partial fulfilment of the requirements of the degree

Heterogeneous Impacts of Cash Transfers: Applying Machine Learning Methods to Evidence from Uganda

Word Count: 9995

Heterogeneous Impacts of Cash Transfers: Applying Machine Learning Methods to Evidence from Uganda

Abstract

Evidence suggests that average productivity impacts of cash transfers are large. Yet, it is possible that such effects mask heterogeneity. Unfortunately, research on this is thin and its findings ambiguous. Our study adds to this growing literature by postprocessing data from a randomized evaluation of a government program in northern Uganda. Using newer machine learning methods, we show that the program generated outsized effects on investment for wealthier and higher ability types but only in the near term, while heterogeneous impacts on earnings are modest to negligible. We also detect new and unexpected sources of heterogeneity that are fertile for future research.

I. Introduction

Cash transfers have risen in popularity as a poverty alleviation tool in the past decade. Economic theory suggests that because of capital constraints, the poor are limited to less efficient and diminishing returns technology in their microenterprises, creating a vicious poverty cycle (Banerjee et al., 2019). Cash transfers alleviate this by reducing the minimum wealth level at which microentrepreneurs can switch into more efficient technology, allowing them to escape the poverty trap and sparking a virtuous growth cycle (Gertler et al., 2012; McKenzie and Woodruff, 2006). In the long run, it is hoped that the poor can reinvest gains and bootstrap their way out of poverty. Studies have shown that cash transfers consistently produce positive productivity impacts (Bastagli et al., 2016). However, average treatment effects often mask important heterogeneity. This is important for unconditional cash transfer (UCT) programs since usage of funds is not monitored and donors have limited funds to disburse. Knowing who is most likely to benefit allows for targeted transfers that achieve the best bang for the buck. Crucially, policymakers also have a responsibility to ensure that interventions do no harm.

In theory, cash transfers should benefit those who are the most credit constrained. Assuming that production and technological choices are bounded by non-convexities, the poorest of the poor are unable to afford high-fixed cost and more efficient technology and are unable to scale the wealth ladder. Higher wealth individuals already operating at efficient scale face diminishing returns to investment and should choose to save or consume the grant. However, empirical support for this is both thin and ambiguous.

Our paper adds to a growing body of work that uncovers such effects. We look for evidence of heterogeneity by postprocessing data from Blattman et al (2014). In their study, the authors evaluate a government program in northern Uganda known as the Youth Opportunities program (YOP). The program was designed to help poor, unemployed youths find self-employment as artisans through a large cash grant. Young adults were invited to form groups and submit proposals on how they would utilize the grant for vocational training and start their own enterprises. Successful groups received a one-time unsupervised grant of about $382 per member, or about a year's wages. The intervention took place in 2008 and follow up surveys were carried out at two and four-year endlines. Average treatment effects were reportedly high: treated participants invested more in business assets, earned greater profits, and had more employment hours. However, the authors fail to detect heterogeneity for all outcomes.

For our study, we reprocess the same data using newer machine learning methods following Chernozhukov et al (2017). These methods are robust to overfit and allows us to explore multidimensional heterogeneity while overcoming traditional issues related to multiple hypothesis testing (Weeks and Christiansen, 2020). We also extend the authors' framework to account for the presence of multiple investment thresholds and accommodate richer covariates in our heterogeneity analysis. We find that two years after the intervention, wealthier and higher ability males at baseline invest significantly in business assets. Almost all participants are compelled by the treatment to invest in vocational training, though higher ability types invest considerably more than their peers. The most fervent investors also lived closer to education facilities and were slightly older. However, these impacts were only present in the near term and were not detected at the second endline. We also do not detect any heterogeneity in profits, consumption, and savings at both endlines, though our evidence suggests that heterogeneous earnings could take longer to materialize. Crucially, we do not detect any signs of negative impacts for our outcomes of interest.

This paper builds upon evidence of heterogeneous productivity impacts in cash transfer settings and complements similar research in the microcredit literature. In particular, it disputes evidence that credit relaxation yields outsized productivity benefits for poorer types and curbs growing enthusiasm for pro-poor transfers. Our findings are also consistent with non-convexities and threshold effects that lead to poverty trap dynamics (Banerjee et al., 2019; Banerjee and Duflo, 2011; Ghatak, 2015; Kraay and McKenzie, 2014). We think these results should be of interest to policy makers looking to fine-tune targeting mechanisms and design cost-effective interventions, more so as cash transfers become increasingly popular as a tool to spur productivity. For researchers, our study reinforces the need for more attention to be paid to heterogeneous impacts and adds to a burgeoning field of machine learning applications in the social sciences.

The rest of the paper is structured as follows. Section II reviews the literature on heterogeneous productivity outcomes and outlines our conceptual framework. Section III provides a summary of the data and setting from the original experiment. Section IV looks at our empirical strategy, including an overview of our machine learning methods and how model selection is performed. Section V displays our results and robustness checks. Section VI presents a discussion of said results and their policy implications, and section VII concludes.

II. Literature review and conceptual framework

Literature review

Studies show that on average, cash transfers produce large positive productivity impacts for asset and livestock ownership, crop production, and non-farm businesses (Bastagli et al., 2016). programs in Malawi, Zambia, Lesotho, and Ghana have reported large impacts on agricultural assets and inputs (Covarrubias et al., 2012; Daidone et al., 2014; Handa et al., 2014; Karlan et al., 2014; Seidenfeld and Handa, 2011), while cash grants in Uganda and Zambia were found to significantly benefit business and enterprise for beneficiaries (AIR, 2014; Blattman et al., 2015, 2014). Similar impacts have also been reported in Latin America (Gertler et al., 2012; Todd et al., 2010).

Whether such positive effects mask heterogeneity is of greater interest to us. In principle, cash transfers should produce heterogeneous impacts for the talented but most credit constrained who have little access to capital (Banerjee et al., 2019; Blattman et al., 2014). Indeed, de Mel et al (2008) found that randomised shocks to capital stock in Sri Lankan microenterprises produced significantly higher returns for higher ability types who were the most capital constrained. Todd et al (2010) study the Oportunidades program in Mexico and found that poor rural households invested part of their grant in productive assets, and impacts were greatest for those that did not own those assets at baseline. Mckenzie (2017) also studied a national-level grant in Nigeria to identify high potential but credit constrained enterprises via a business plan competition known as 'YouWiN!' with winners randomly selected from a pool of semi-finalists. After five years the study found that winning the competition led to higher survival rates, profits, sales, and employment. While the program was self-selective by nature, the study did find modest evidence of outsized benefits for firms that reported higher capital constraints at time of application.

Yet another strand of research suggests that benefits accrue to those better endowed. Scholars have alluded to differential productive impacts based on a household's existing asset base, with those having access to land and labor more capable of investing productively (Covarrubias et al., 2012; Davis et al., 2002). Prifti et al (2020) evaluate Lesotho's Child Grant program and report greater economic outcomes in both the crop and livestock sector for those with greater land and labor capacity. They further suggest that impacts were greater for those with levels of per capita consumption expenditure above a certain threshold, disputing the idea that the most destitute benefit from extra liquidity. Bandiera et al (2017) investigate the BRAC asset transfer program in Bangladesh which targeted the poorest women through randomised provision of livestock and training. The authors found that gains in productive assets were highest for those at the highest distribution centile at baseline and these effects were mirrored across measures of consumption and savings, though they were unable to untangle the effects of skills training and increased liquidity.

An implication from this is that hidden complementarities between capital endowments and cash grants may only manifest for those operating above a certain threshold. We find comparable evidence of this in microcredit evaluations. Banerjee et al (2019) conduct a six-year follow up of a microcredit expansion program in India and found that treated households that were running a business prior to credit expansion showed persistent benefits that increased over time. We note that individuals who select into entrepreneurship even when credit constrained are likely different from those who do not, holding wealth constant. However, using constructed estimates of baseline wealth, the authors also found that intermediate to high wealth individuals benefited most from credit expansion due to their proximity to high return-high fixed cost technology. Evidence of heterogeneous effects on prior business ownership have also been reported in Meager (2019) and Chernozhukov et al (2017), while other evaluations in Morocco, India, and Mexico have also found positive effects on business profits concentrated in the upper tail of the distribution (Angelucci et al., 2015; Banerjee et al., 2015; Crépon et al., 2015).

On the whole, evidence of heterogeneous productivity impacts remains scarce. Most evaluations focus on social welfare goals and productive outcomes are often of secondary concern, while many studies that do explore productivity do not follow beneficiaries long enough for productive impacts to establish. (Handa et al., 2018; Prifti et al., 2020). What our review shows is that evidence of heterogeneity is strong, but the direction in which it operates is ambiguous. We view this as a potential gap in the literature.

Conceptual framework

Our conceptual framework builds upon existing work by Blattman et al (2014). In the absence of cheap credit, an unexpected windfall should benefit high-ability but capital-starved types more compared to the low-ability and less credit-constrained. High ability types will invest in enterprise until efficient scale is reached, while low ability types may not invest[1] even under perfect credit markets and will likely consume or save their windfall. Wealthier, high ability individuals will already be investing at efficient scale and will likewise save or consume their windfall. As outlined above this is not always the case. Regardless, the expectation is that such heterogeneous impacts only surface if the windfall relaxes a capital constraint for a subset of the population or if – absent functioning insurance markets – enterprise is not inherently riskier than traditional labor. If enterprise is unduly risky then missing insurance markets may discourage risk-averse individuals from investing (Hennessy, 1998; Mendola, 2007). Whether this is the case depends largely on context. Evidence from a cash grant experiment to farmers in Ghana found that insurance market imperfections were great enough to dull returns to capital (Karlan et al., 2014). On the flipside, de Mel et al (2008) found that treatment effects from their experiment in Sri Lanka did not vary significantly with measures of risk-aversion or uncertainty despite imperfect insurance markets. As discussed in section III, our setting in Uganda had neither formal finance nor functioning insurance markets at time of intervention. However, we discount the role of missing insurance at the outset for two reasons. First, our sample is underemployed and non-enterprise

[1] In the absence of employment, low ability types may invest reluctantly but will produce below efficient scale. If returns are very low they may divest and either consume or save. Whether and when this happens will also depend on any flypaper effect that makes investment 'sticky.' However, in all scenarios we expect high ability types to outperform low ability counterparts.

hours are very low to begin with. Second, our estimates of post treatment earnings are less variable than at baseline. Under these circumstances microenterprise is unlikely to be riskier than traditional labor and may even be less so. Nevertheless, we do not dismiss this wholesale and include a measure of risk-aversion in our analysis.

Our framework is consistent with technological non-convexities and poverty traps and builds upon previous theoretical and empirical scholarship (Banerjee and Duflo, 2011; Banerjee and Newman, 1993; Carter and Barrett, 2006; Gertler et al., 2012; Ghatak, 2015; Kraay and McKenzie, 2014; McKenzie and Woodruff, 2006; Phimister, 1995). We also borrow ideas of heterogeneous technological adoption from Banerjee et al (2019). Because credit markets are imperfect, talented but low-wealth individuals are limited to less efficient and diminishing returns technology in their microenterprises, creating a vicious poverty cycle. Seed adoption is one such example: in Uganda, home-saved seeds account for 85-90% of all crop production. Switching to high-yield varieties is likely to generate greater productive outcomes, but farmers face liquidity constraints and relatively high cost of agricultural inputs (Jelliffe et al., 2018). Cash transfers and credit access alleviate this by reducing the minimum wealth level at which farmers can switch into more efficient technology (Phimister, 1995). In this case there exists an investment threshold beyond which the growth cycle is virtuous. Figure 1 shows a simple illustration of a poverty trap with a singular threshold.

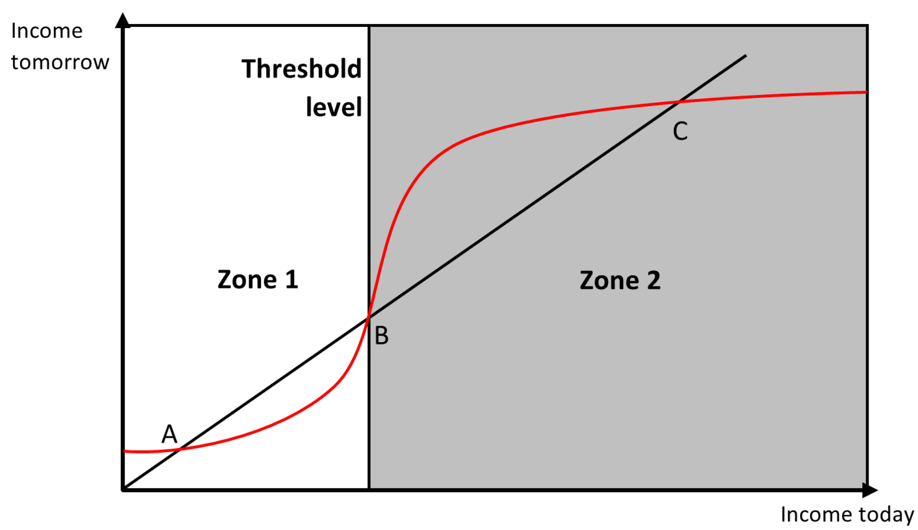

Figure 1: Single threshold

Under a singular threshold we predict the following:

1. An unexpected windfall should induce poorer, high ability types in zone 1 to invest and profit more, assuming the windfall sufficiently unlocks a capital constraint by taking them past point B.

2. Wealthier, high ability types are likely to already be operating at efficient scale in zone 2, in which case any unexpected windfall should be saved or consumed.

We extend this framework to account for multiple thresholds. Multiple evaluations of productive cash grants find that investment levels plateau despite high positive initial treatment effects. Research on Zambia's Child Grant program found that the intervention led to significant increase in livestock assets and ownership of non-farm enterprise 24 months in, though impacts were unchanged following an additional year of exposure (AIR, 2014). De Mel et al (2008) conduct an experiment in Sri Lanka where microentrepreneurs were randomly given 10,000 or 20,000 rupees (about \$100 or \$200) in cash or kind. Despite high returns, individuals who received the larger grant did not invest significantly greater than those who received less, instead electing to consume the surplus. A plausible explanation is that the presence of a subsequent investment threshold discourages microentrepreneurs from scaling their enterprise despite high and positive initial treatment effects.

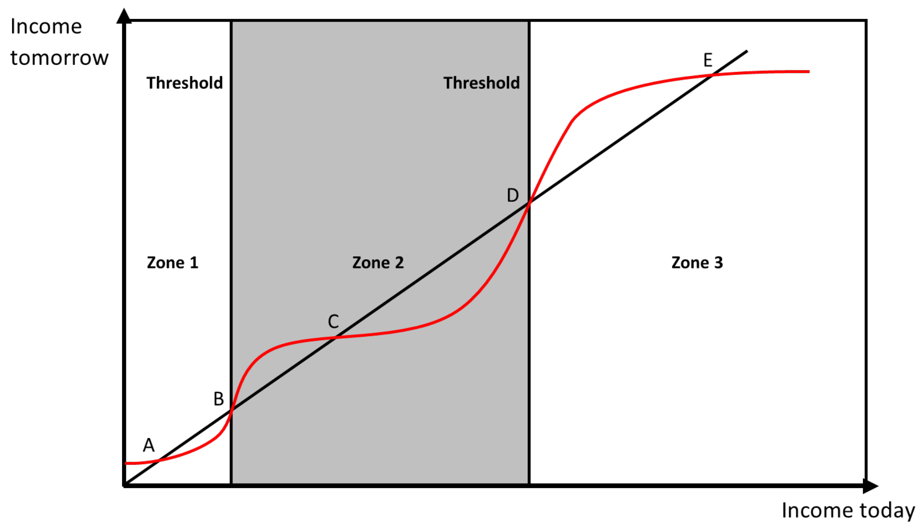

Figure 2: Multiple thresholds

Figure 2 shows a model with multiple thresholds. Under these circumstances it is not immediately obvious the direction and magnitude in which heterogeneous impacts manifest. In the presence of technological non-convexities, wealthier types can benefit more from a cash grant if it takes them past a subsequent investment threshold, and if poorer types are so capital-starved that they are unable to substantially scale their enterprise even with the windfall. We build several scenarios to illustrate this point.

1. If the subsequent threshold (point D) is prohibitively high, a windfall may be insufficient to take individuals at zone 2 to the next equilibrium. However, it may be possible to take poorer types in zone 1 into zone 2. In this case poorer types should invest in enterprise while richer types in zone 2 save or consume their windfall.
2. If a windfall is sufficient to push wealthier types beyond zone 2 and poorer types beyond zone 1, both poor and rich types are likely to invest. Which heterogeneity dominates depends on the size of investment required to cross each threshold. Assuming monotonically increasing thresholds, an unexpected windfall should induce wealthier types at zone 2 to invest more compared to poorer types. Surplus capital for those in zone 1 is either saved or consumed.

3. A third scenario is one in which the initial investment threshold far exceeds subsequent ones. Analogous to this would be industries with prohibitively large upfront costs, but less costly thresholds subsequently. We do not consider this possibility in our setting. The YOP encourages participants to invest in vocational training with the aim of seeking skilled vocational work. As shown in section VI, there is little evidence to suggest that vocational training is prohibitively costly; indeed, the treatment induces almost all participants to invest in training regardless of initial wealth.

Long run heterogeneous effects

We also extend our framework to account for long run heterogeneity. It is not always obvious that heterogeneous profits should surface by endline, even if it is expected to exist. It is possible that adopting new technology may incur adjustment costs that lead to sub-optimal production in the short run even for high ability types. Firm level studies show that such costs can be substantial and delay technological adoption by workers or management (Atkin et al., 2017; Pavlova, 2001). If a cash transfer is restrictive, recipients will be unable to delay technological adoption and are likely to produce inefficiently in the short term. Alternatively, if high-ability types have extremely low levels of initial wealth then a windfall may be insufficient to overcome fixed cost of enterprise in the short run. In this case patient types will remain in traditional labor and accumulate savings until they reach the required threshold. In both cases the 'motivated poor' should reach optimal production only in the medium to long run.

Our ideas of thresholds and poverty traps are not new and build upon existing work. Our main contribution is the introduction of multiple thresholds and long run heterogeneity to our cash transfer framework. To our knowledge this is not widely studied in the literature.

From our conceptual framework we form the following hypotheses:

(1) Under imperfect credit markets, prior economic and human capital endowment are determinants of cash transfer outcomes, though the direction of heterogeneity is ambiguous.
(2) Heterogeneous profits may only manifest in the medium to long term due to high adjustment or startup costs.

III. Empirical data and setting

To test our hypotheses, we postprocess experimental data from Blattman et al (2014), hereafter referred to as BEA. The authors study a government cash transfer in northern Uganda that took place in 2008 and followed this up with endline surveys in 2010 and 2012. For reasons of brevity this section only outlines details of prime relevance from BEA.

Setting in Northern Uganda

Uganda is a poor but growing country located in East Africa. In 2007, the country had gross domestic product (GDP) per capita of roughly $330 with annual GDP growth at about 6.5% a year from 1990 to 2007 (Government of Uganda, 2007). However much of this growth was concentrated in the south-central regions. The economy in the north – being distant from trade routes – was still broadly concentrated around agriculture and livestock rearing. As a bed of opposition support it also received less public investment from the 1980s. From the late 1980s to mid 2000s the region was also plagued by insurgency in the form of rebel militias and insecurity in neighbouring Sudan and Democratic Republic of Congo.

By 2006, the military pushed rebels out of the country and peace came to Uganda's neighbours. Years of destabilisation meant that nearly two-thirds of northern Ugandans were unable to meet basic needs, just over half were literate, and most were underemployed in subsistence agriculture (ibid.). Access to formal finance was almost non-present. Village savings and loan groups were common but carried extortionate interest rates at 100-200% with loan duration seldom exceeding three months. The setting thus closely resembles the credit-constrained environment ubiquitous to many poor communities.

From 2003 to 2010, the central Government headed the Northern Uganda Social Action Fund (NUSAF) as part of its northern development strategy. It allowed poor communities to apply for cash grants for either community infrastructure construction or livestock. In 2006, the government announced a third component in a bid to boost non-agricultural employment: the YOP.

The YOP

The YOP invited young adults aged 16-35 to apply for cash grants to start a skilled trade such as metalworking or carpentry, among others. Applicants had to form groups and submit a written proposal to receive the cash grant. In this sense the cash grant was restrictive, though usage of funds was not monitored. Proposals had to outline how participants planned to use the grant for skills training and business start-up costs. Most participants selected their own trainers in their proposal. These included local artisans or vocational institutes which are commonplace in the country, most of which had been in existence for at least five years at time of the intervention. For administrative purposes, the cash grant was disbursed at the group level and split among members upon approval. Successful groups received one-time, unsupervised grants worth $7,500 on average—about $382 per group member, roughly their average annual income.

Experimental design

Selection for the YOP was carried out by government officials with funding randomised among screened and eligible groups. In total, 265 of 565 groups were randomised to treatment and 270 to control, stratified by district. Spillovers were deemed unlikely due to the large geographical distance between villages. Of the 270 in control, 13 groups could not be located at baseline, though BEA suspected no foul play.

Five members from each group were surveyed three times over the course of four years from 2008 to 2012. In total 2677 individuals were surveyed over four years – two additional members were inadvertently surveyed in one group. A baseline survey was carried out in 2008 before the intervention while subsequent endlines were carried out two years apart in 2010 and 2012 respectively. Response rates for our outcomes of interest were at about 75% for endline 1 and 70% for endline 2.

Participants

Most of the participants were young, rural, credit constrained, and involved in traditional, low skilled labor. About half reported no employment in the past month and most were underemployed with employment hours averaging just 11 per week. They were also credit constrained: less than 10% reported that they had access to large loans of UGX 1,000,000 (about $580). More tellingly, less than 8% had prior vocational training. This even though YOP applicants were slightly wealthier and more educated compared to the population average, based on the Northern Uganda Survey carried out in 2008 (Blattman et al., 2014). Table 1 shows average baseline characteristics of treatment and control groups and regression differences. Despite

Table 1: Pre-intervention test of balance

	Control		Treatment		Regression difference	
	Mean	Std. dev.	Mean	Std. dev.	Mean	p-value
Grant amount applied for, USD	7527.042	2104.467	7276.023	2024.378	108.393	0.407
Group size	22.431	6.860	21.426	6.802	0.279	0.587
Grant amount per member, USD	357.749	151.082	381.468	170.738	15.506	0.187
Group existed before application	0.449	0.498	0.491	0.500	0.034	0.420
Group age, in years	3.793	1.951	3.828	1.890	-0.024	0.887
Within-group heterogeneity (z-score)	-0.028	0.923	0.027	1.061	-0.026	0.757
Quality of group dynamic (z-score)	-0.017	1.016	0.019	0.980	0.054	0.494
Distance to educational facilities (km)	6.830	6.523	7.269	5.708	0.494	0.337
Unfound at baseline	1.000	0.000	1.000	0.000	0.000	0.310
Age	24.756	5.225	25.132	5.314	0.163	0.558
Female	0.349	0.477	0.323	0.468	-0.022	0.395
Large town/urban area	0.232	0.422	0.200	0.400	-0.016	0.642
Risk aversion index (z-score)	-0.022	1.002	-0.034	1.007	-0.013	0.765
Any leadership position in group	0.279	0.449	0.286	0.452	-0.003	0.843
Group chair or vice-chair	0.104	0.305	0.117	0.321	0.010	0.307
Weekly employment, hours	10.724	15.842	11.377	15.485	0.519	0.518
All nonagricultural work	6.028	12.497	5.741	11.387	-0.460	0.426
Casual labor, low skill	1.033	5.201	1.075	5.000	-0.112	0.629
Petty business, low skill	2.251	6.970	2.416	6.817	0.205	0.522
Skilled trades	1.788	8.431	1.543	7.787	-0.341	0.388
High-skill wage labor	0.045	0.578	0.140	1.028	0.080	0.019
Other nonagricultural work	0.910	4.776	0.567	3.763	-0.292	0.098
All agricultural work	4.696	10.122	5.636	10.520	0.979	0.053
Weekly household chores, hours	9.009	17.622	8.714	16.149	0.284	0.747
Zero employment hours in past month	0.477	0.500	0.419	0.494	-0.037	0.202
Main occupation is nonagricultural	0.258	0.438	0.279	0.449	0.001	0.953

Table 1: *(continued)*

	Control		Treatment		Regression difference	
	Mean	Std. dev.	Mean	Std. dev.	Mean	p-value
Engaged in a skilled trade	0.061	0.239	0.058	0.235	-0.007	0.557
Currently in school	0.045	0.207	0.038	0.191	-0.007	0.441
Highest grade reached at school	7.954	2.923	7.823	3.034	-0.070	0.631
Literate	0.751	0.433	0.715	0.452	-0.029	0.152
Prior vocational training	0.073	0.261	0.085	0.279	0.022	0.053
Digit recall test score	4.165	2.000	4.010	1.961	-0.040	0.619
Index of physical disability	8.684	2.527	8.629	2.217	-0.138	0.308
Durable assets (z-score)	-0.159	0.956	-0.072	1.051	0.073	0.135
Savings in past 6 mos. (000s 2008 UGX)	19.353	98.450	33.052	137.331	10.924	0.016
Monthly gross cash earnings (000s 2008 UGX)	62.467	129.337	67.869	135.374	6.806	0.311
Can obtain 100,000 UGX ($58) loan	0.335	0.472	0.400	0.490	0.052	0.010
Can obtain 1,000,000 UGX ($580) loan	0.099	0.299	0.118	0.323	0.012	0.426
Human capital (z-score)	0.037	0.992	-0.021	1.036	-0.024	0.628
Working capital (z-score)	-0.062	0.867	0.083	1.136	0.115	0.005
Patience index (z-score)	-0.021	1.040	-0.012	0.980	0.041	0.340
Aggression index (z-score)	-0.010	1.012	0.010	0.988	-0.013	0.782

Measures of human capital, working capital, patience, risk-aversion, and aggression were aggregated from questionnaire responses, weighted, then standardized. We detail this in Appendix A1.

randomization, treated individuals began wealthier (higher working capital) than control at baseline. BEA attribute this imbalance to chance or missing control groups – they estimate that missing controls could account for imbalance if they were 0.1-0.2 standard deviations above the control mean. We make no such assumptions in this study. Working capital is a main variable of interest and using imbalanced data could seriously undermine our findings. To remedy this, we perform propensity score matching[2] and discard 48 observations – just under 2% of our sample – from our treatment group. We rerun checks for balance and find no differences between treatment and control groups across our main variables of interest. Details of the procedure and rebalanced covariates can be found in Appendix A2.

IV. Empirical strategy

In their heterogeneity analysis, BEA pooled samples from both endlines and found that initial wealth endowment only affected treatment outcomes for females, while human capital endowment produced no heterogeneous treatment effect. Overall, the authors do not detect heterogeneity though the signs point to heterogeneous impacts for poorer and higher ability types at baseline.

Using our extended framework, this study will attempt to detect heterogeneity by postprocessing data from BEA with newer statistical methods and richer covariates. The reasons for postprocessing this dataset are twofold. First, conventional methods used in the paper for detecting heterogeneity like subgroup analysis and interaction terms are sensitive to sample size and model selection. To mitigate this the authors pooled both endlines which raised statistical power, but their model specification precluded them from testing for heterogeneity at each individual endline[3]. Using newer methods, this study seeks to isolate heterogeneous effects at each endline and in turn test hypothesis (2). Second, the use of pre-analysis plans – designed to prevent 'P-hacking' – limits researchers to using only prespecified covariates, wasting rich covariates in the process. Such restrictions can make it difficult to discover strong but unexpected treatment effect heterogeneity (Wager and Athey, 2018). While the search for heterogeneity should be driven by theory, an ex-ante approach to covariate selection risks omitting important sources of heterogeneity, which, given the cost and difficulty in implementing randomised experiments, constitutes a substantial waste.

This study will employ a generic machine learning (GML) method by Chernozhukov et al (2017) to detect treatment effect heterogeneity. In a standard OLS model, introducing all available covariates and their various permutations for interaction terms will likely lead to overfitting. Limitations in statistical power limited BEA's heterogeneity analysis to five key variables: human capital, working capital, patience, gender, and an indicator for whether a participant was operating in a skilled trade. By contrast the advantage with GML is its ability to handle models with large number of covariates, allowing us to include a rich set of characteristics. In effect it is possible to run all covariates through the model in our search for heterogeneity. This allows us to test a third, auxiliary hypothesis:

(3) Heterogeneity in cash transfer outcomes is not limited to human capital and wealth endowment.

[2]If imbalance is purely due to chance then balancing observable covariates may risk generating imbalance in unobservable variables. As we outline in our empirical strategy, covariates are used to perform prediction and not causal inference. Potential imbalance in unobservables should not jeopardize our findings. In the worst case, we merely generate noisy predictions that penalize our estimates.

[3]We unpool the estimates and attempt to replicate our results below using interaction terms but do not detect strong heterogeneity (Appendix A3).

Generic Machine Learning

Since ML is largely focused on prediction, a fundamental issue is that it does not produce consistent estimates and valid confidence intervals, thereby making causal inference difficult. GML sidesteps this problem by using ML solely as a proxy for predicting *S(Z)*, the Conditional Average Treatment Effect (CATE). Inference is made not on CATE but on the features of CATE itself. Multiple sample splits are utilized to mitigate estimation uncertainty conditional on the data split, while median values are taken to account for splitting uncertainty[4]. Alternative ML methods for heterogeneity like Meta-learners (Künzel et al., 2019) and Causal Forests (Wager and Athey, 2018) were considered, but the choice of GML is due to its greater ease of interpretation and agnosticism to ML algorithm used.

The GML method is split into three components: Best Linear Predictor (BLP), Sorted Group Average Treatment Effect (GATES), and Classification Analysis (CLAN).

BLP: To find the BLP, the data is split equally and randomly into a main and auxiliary sample n times. A ML model is fitted onto the auxiliary sample and used to predict treatment and baseline outcomes for the main sample. We then estimate the following regression:

$$Y = \beta_0 + \beta_1(D - p(Z)) + \beta_2(D - p(Z))(S - E(S)) + \epsilon,$$

where Y refers to the outcome of interest, D is the treatment dummy, $p(Z)$ is the propensity score, S is the CATE and $E(S)$ is the expectation of S. Coefficients and p-values are obtained from the median of all splits. Rejecting the hypothesis $\beta 2 = 0$ means that there is both heterogeneity and $S(Z)$ is its relevant predictor, in which case we move on to GATES.

GATES: We split the main sample into K quantiles based on $S(Z)$ and impose the following monotonicity restriction:

$$E[S(Z)|G_1] < ... < E[S(Z)|G_K]$$

To recover the GATES parameters, we assign a dummy to Gk and run the regression:

$$Y = \beta_0 + \beta_1 B(Z) + \sum_{k=1}^{K} \gamma_k (D - p(Z))(G_k) + \epsilon$$

for each k, where $B(Z)$ is the predicted baseline outcome. If GATES is significant and there is heterogeneity, we should see a large, significant coefficient for γK and a small insignificant coefficient for $\gamma 1$.

CLAN: Once heterogeneity is established, we report the median of the average characteristics of choice for $G1$ and GK and compare differences between the two. Analysis will be performed on both endlines, which allows us to test if heterogeneity manifests in the short or long term, if at all.

[4] Confidence intervals and p-values are derived from median values and are 'sample split-adjusted' according to $P = 2\underline{Med}(p\theta|Data) \leq \alpha$.

Model selection

To allay concerns of 'P-hacking,' we include all baseline covariates from BEA as well as district-level dummies in our model specifications. As extra precaution we accept only very small p-values as evidence against the null and set $\alpha = 0.01$. Following Chernozhukov et al (2017), confidence intervals are set at $1-2\alpha$ to account for splitting uncertainty. All scores and covariates were obtained at baseline aside from patience which was recorded at endline 1, but which the authors treat as time-invariant. We include this variable nonetheless but treat our results with caution. A list of all covariates can be found in table 1.

We employ GML to predict and test for heterogeneity for the following outcomes: monthly profits, capital stock, consumption (durable and nondurable), training hours, and savings. Of these only monthly profits and capital stock were featured in BEA's heterogeneity analysis. Our choice of outcomes is motivated by our conceptual framework. Under heterogeneous individuals and imperfect credit markets, our framework predicts heterogeneity in profits and investment. High ability and either poorer or wealthier types will invest more in business assets and training and earn greater profits; the rest will either save or consume. We lack a measure for investment in training and include training hours instead as an imperfect proxy. Data for training hours and savings are only available at endline 1, while data on nondurable consumption was only collected at endline 2.

Our ML methods of choice for predicting $S(Z)$ are ridge regression and random forest. Tuning complexities and computational cost severely limited our options and we eventually settled for algorithms with few hyperparameters and ease of tuning. Because each split generates a randomly sampled dataset, hyperparameters should be selected via cross validation at each split for optimal results. Due to computational constraints, hyperparameter selection was performed only once for the entire dataset via 10-fold cross validation. The same hyperparameters were then utilized for all splits. We run 100 splits for all our models with number of groups $K = 5$.

We select our best model by maximizing:

$$\Lambda = |\beta 2|^2 Var(S(Z))$$

which is equivalent to maximizing the correlation between the ML proxy predictor $S(Z)$ and the true score $s0(Z)$.

V. Results

Of the 265 groups assigned to treatment, 236 received the cash grant. Of the untreated, 21 groups had no access to funds due to unsatisfactory proposals, bank complications, or collection delays, plus eight groups reporting missing funds due to either theft or diversion. Because of noncompliance we use intention to treat (ITT) estimates throughout.

Table 2 reports Λ values for our outcomes and ML methods of choice. We find that our random forest models outperform ridge regression in almost all specifications, with the exceptions being durable consumption and savings at endline 1. We choose our BLP based on Λ values and report only coefficients for our method of choice. Table 3 displays our estimates for the coefficients $\beta 1$ and $\beta 2$ in our BLP regression, which correspond to the average treatment effect (ATE) and heterogeneous loading parameter (HET) respectively. In parentheses we report standard errors; in brackets we report adjusted p-values for the hypothesis that the estimate is equal to zero. We find that for all outcomes our ATE coefficients are consistent with findings from BEA. Turning to heterogeneous treatment effects, we are able to obtain very small p-values for the hypothesis that HET is zero for capital stock and training hours at endline 1, suggesting strong presence of heterogeneity. Conversely, our BLP analysis detects no significant heterogeneity for profits, consumption, and savings at all endlines.

Table 2: Comparison of ML Methods

Outcome	Ridge regression (Λ)	Random forest (Λ)
Profits		
E1	9.810	20.642
E2	20.960	128.080
Capital stock		
E1	17326.310	92487.950
E2	2140.553	2584.131
Durable consumption		
E1	0.004	0.002
E2	0.001	0.004
Nondurable consumption		
E1 – Data not available		
E2	0.002	0.003
Savings		
E1	155.674	102.464
E2 – Data not available		
Training hours		
E1	5957.837	8804.406
E2 – Data not available		

We next estimate our GATES parameters. We divide our sample into $K = 5$ groups based on the quintiles of the ML proxy predictor $S(Z)$ and estimate the average effect for each group. We refer to participants in group K as high performers (HPs) and those in group 1 as low performers (LPs). Figure 3 presents the estimated GATES coefficients for capital stock and training hours at endline 1 along with joint confidence bands. Our GATES demonstrate clear monotonicity for outcomes with significant heterogeneity. We further investigate the GATES by comparing the most and least affected groups in table 3.

Overall, our GATES estimates are consistent with our BLP findings: we find extremely small p-values for our most affected groups for capital stock and training hours at endline 1 and the differences are also highly significant. On average HPs invest 872,000 UGX ($506) more in business assets and 482 more training hours than LPs at endline 1. We conclude from this that large average treatment effects of the YOP in terms of investment is driven by heterogeneous individuals but only in the near term. The fact that investment is not sustained may be indicative of a further threshold against which the current transfer is infra-marginal.

Our conceptual framework also predicts that LPs should either save or consume their windfall. Neither our BLP nor GATES display evidence of such, though it should be noted that nondurable consumption data for endline 1 is unavailable. We also do not detect heterogeneity for profits at either endline though we note that our HET estimate at endline 2 is much larger than at endline 1 and its p-value is also greatly reduced from 38.4% to 12.6%. One plausible explanation is that heterogeneity exists though it may require a longer time to manifest. Reassuringly, we also find that the treatment does not produce negative impacts for all our least affected groups. We discuss these findings in greater detail in section VI.

Table 3: BLP and GATES

Outcome	ATE (β_1)	HET (β_2)	G5 (γ_5)	G1 (γ_1)	Difference (γ_5-γ_1)
Profits					
E1	13.142	-0.185	13.870	20.304	15.785
	(5.09)	(0.336)	(11.437)	(11.430)	(13.062)
	[0.009]*	[0.384]	[0.444]	[0.162]	[0.514]
E2	11.679	0.554	25.586	0.386	26.151
	(6.22)	(0.316)	(14.002)	(14.037)	(15.420)
	[0.056]	[0.126]	[0.134]	[1.00]	[0.168]
Capital stock					
E1	408.522	0.803	853.657	16.281	872.024
	(97.466)	(0.263)	(219.350)	(222.249)	(249.556)
	[0.000]**	[0.005]*	[0.000]**	[1.00]	[0.001]*
E2	226.637	0.057	263.046	273.562	280.376
	(85.547)	(0.516)	(191.861)	(192.459)	(200.197)
	[0.007]*	[0.555]	[0.328]	[0.292]	[0.324]
Durable consumption					
E1	0.108	0.255	0.150	-0.021	0.149
	(0.053)	(0.230)	(0.120)	(0.120)	(0.121)
	[0.045]	[0.263]	[0.430]	[1.00]	[0.436]
E2	0.125	0.257	0.170	0.022	0.176
	(0.063)	(0.294)	(0.141)	(0.141)	(0.144)
	[0.045]	[0.342]	[0.460]	[1.00]	[0.480]
Nondurable consumption					
E2	0.122	0.261	0.048	0.212	0.211
	(0.062)	(0.265)	(0.140)	(0.139)	(0.148)
	[0.047]	[0.335]	[0.491]	[0.130]	[0.168]
Savings					
E1	20.343	-0.366	8.099	38.931	7.394
	(13.493)	(0.386)	(30.188)	(30.249)	(32.830)
	[0.127]	[0.360]	[1.00]	[0.384]	[1.00]
Training hours					
E1	348.172	0.882	481.774	196.211	482.570
	(23.779)	(0.229)	(56.831)	(58.551)	(54.881)
	[0.000]**	[0.000]**	[0.000]**	[0.002]*	[0.000]**

**p<.001, *p<.01

Columns (1) and (2) report the ATE and HET estimates respectively. Column (3) reports estimates for our most affected groups Gk based on our GATES regression, column (4) reports estimates for our least affected groups G1, and column (5) reports the difference in means between the most and least affected groups.

All values are obtained from the median of all splits and do not necessarily correspond to the same split.

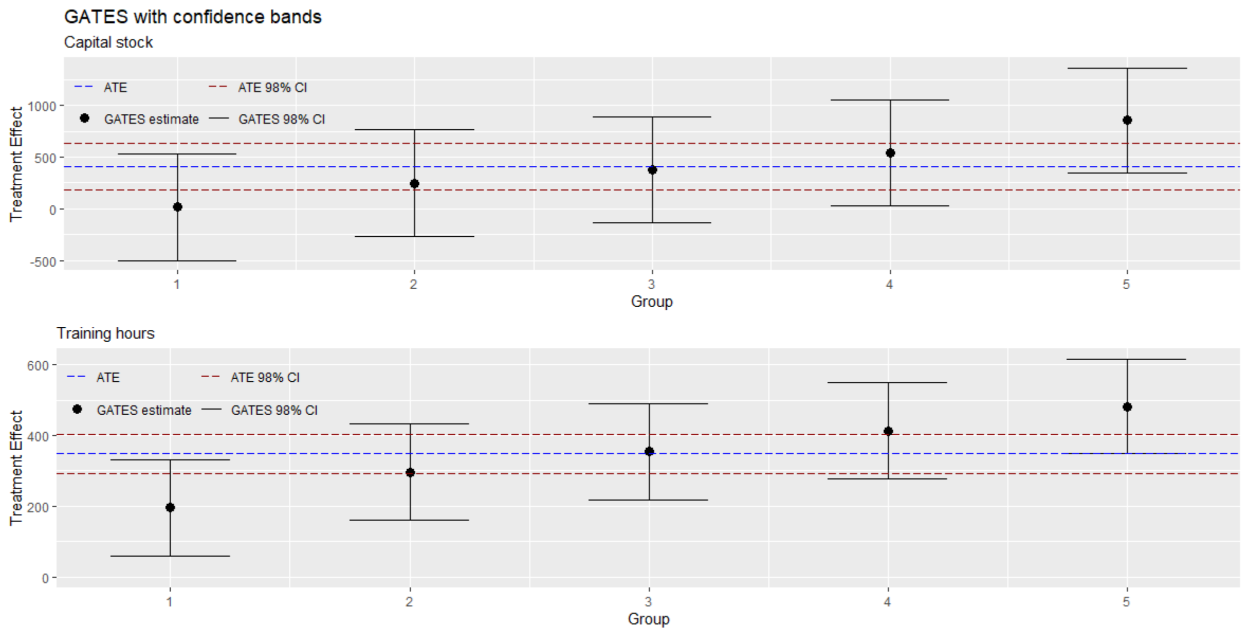

Figure 3: GATES estimates and confidence bands

CLAN:

We proceed by looking at the average characteristics of the most and least affected groups to understand what generates heterogeneity in the treatment effects. Our choice of covariates to report is motivated by our hypotheses in section II and other plausible sources of heterogeneity. We do not report baseline covariates that are already used in index construction. For instance, savings at baseline is included in our working capital index and is thus omitted from our analysis. Table 4 shows our CLAN estimates. Estimates represent sample-split adjusted difference in means for characteristics between our most and least affected groups. Adjusted confidence intervals and p-values are reported in parentheses and brackets respectively.

Our main covariates of interest are prior economic and human capital endowment for which we find strong evidence of heterogeneity. Participants with higher human and working capital at baseline invest more in business assets: for our capital stock outcome HPs score on average 0.52 standard deviations higher on human capital and 0.57 standard deviations higher on working capital than LPs. They also tend to be overwhelmingly male and are slightly older. Participants that report the most training hours score on average 0.89 standard deviations higher than LPs in terms of human capital and live about 2.4 km closer to educational facilities. Interestingly, prior wealth endowment does not determine participants' investment in training though our estimate is positive and the p-value relatively small at 0.068.

For both outcomes HPs display either similar or significantly higher levels of risk aversion compared to LPs. Participants who invest the most in business assets do not report differences in risk-aversion compared to LPs, while HPs for our training hours outcome score on average 0.53 standard deviations higher in terms of risk-aversion than LPs, and this estimate is highly significant. These results show that absent functioning insurance markets, investment in vocational enterprise – at least for our setting – is not inherently riskier than traditional labor and may in fact be less so. We also find that patience is a non-factor for either outcome though the self-selective nature of the intervention may have filtered out more impatient types.

Table 4: Classification Analysis (CLAN)

	Capital stock (E1)	Training hours (E1)
Human capital	0.519 (0.269, 0.775) [0.000]**	0.885 (0.626, 1.129) [0.000]**
Working capital	0.566 (0.308, 0.819) [0.000]**	0.190 (-0.048, 0.410) [0.068]
Unemployed	-0.071 (-0.186, 0.044) [0.170]	-0.073 (-0.191, 0.045) [0.142]
Risk aversion	-0.119 (-0.358, 0.119) [0.184]	0.528 (0.293, 0.753) [0.000]**
Average distance to education facility	-0.076 (-1.304, 1.253) [0.184]	-2.398 (-3.803, -1.000) [0.000]**
Urban	0.021 (-0.073, 0.119) [0.272]	-0.038 (-0.132, -0.056) [0.402]
Age	1.977 (0.764, 3.124) [0.000]**	-0.551 (-1.762, 0.680) [0.316]
Patience	0.050 (-0.187, 0.288) [0.244]	0.146 (-0.098, 0.386) [0.264]
Aggression	0.232 (-0.005, 0.469) [0.042]	-0.083 (-0.321, 0.154) [0.394]
Female	-0.314 (-0.419, -0.208) [0.000]**	0.002 (-0.110, 0.113) [0.534]
In school	-0.018 (-0.063, 0.028) [0.484]	0.005 (-0.040, 0.051) [0.720]

Robustness checks

Our attrition rate is high and could possibly skew our results. For all outcomes and endlines specified we report response rates of between 70-75%. We assume attrition to be random[5] though we are uncertain the direction in which it operates. Table 5 reports tests of covariate balance for our unfound treated and control groups. We find that differences in observed baseline characteristics between the unfound treated and control are insignificant apart from working capital. On average the unfound treated score about 0.3 standard deviations higher in terms of working capital at baseline and the estimate is highly significant. Results from our main analysis and our conceptual framework suggest that outcomes in this group could be high, in which case our treatment effects could be understated. We also run a regression of the attrition dummy on the treatment dummy and find no difference in the likelihood of attrition between treatment and control groups.

For robustness checks we first perform conservative imputations using Lee bounds (2009). Our choice of Lee bounds was motivated by random assignment and underlying assumptions of monotonicity: that is, assignment to treatment can only affect attrition in one direction. It is plausible that by relaxing capital constraints, our treatment reduces labor search frictions and increases participants' mobility. Indeed, our setting is one in which participants are young, rural, and highly transient: at each endline survey nearly 40% of YOP participants had moved or were away temporarily.

Our concern in using highly conservative imputations is that it may generate spurious results for CLAN given heterogeneous attrition. As reported in table 5, our unfound treated are significantly wealthier at baseline than counterparts in control. In this case imputing high outcomes to control and low outcomes to treatment risks overstating treatment effects for less wealthy participants. We resolve this by generating sharp predictions for missing values using the missForest algorithm (Stekhoven and Bühlmann, 2012). Unlike other imputation methods missForest does not require assumptions to be made about the distribution of the data, only that missingness does not depend on the value of the missing data itself but on observed values. Table 6 shows BLP and GATES estimates after imputation. We find that our results are robust for all model specifications and evidence of heterogeneity is strong.

Of greater interest are our findings from CLAN (table 7). Under missForest we find that our results are consistent for all variables aside from urbanization and the signs are correct. We also run CLAN with Lee bounds and find broadly similar results, though our estimate for working capital is smaller and its p-value slightly insignificant. Estimates for our patience index estimate are also negative and highly significant. We attribute this to heterogeneous attrition and conservative imputations and address this in the following section.

[5]By random we assume that missingness does not depend on the value of missing data itself but on observed values, or MAR assumptions. While we cannot prove this definitively, BEA do not report any survey refusal and no reward was offered for survey completion. Qualitatively, we reason that attrition is more likely due to migration.

Table 5: Test of balance for unfound participants

	Control	Treatment	Difference
Human capital	-0.072 (0.731)	-0.030 (0.890)	0.029 [0.686]
Working capital	-0.173 (0.724)	0.186 (1.34)	0.312 [0.000]**
Unemployed	0.371 (0.483)	0.414 (0.493)	0.031 [0.501]
Risk aversion	0.225 (0.944)	0.039 (0.958)	-0.206 [0.027]
Average distance to education facility	6.650 (6.72)	7.750 (6.61)	0.912 [0.219]
Urban	0.249 (0.433)	0.263 (0.441)	0.022 [0.654]
Age	24.200 (4.55)	24.500 (5.24)	0.044 [0.923]
Patience	0.204 (0.975)	0.269 (0.918)	0.102 [0.249]
Aggression	-0.019 (1.020)	-0.024 (0.962)	-0.036 [0.684]
Female	0.251 (0.435)	0.352 (0.478)	0.078 [0.070]
In school	0.039 (0.194)	0.050 (0.219)	0.011 [0.499]

Column (3) reports the regression difference with district fixed effects and group-clustered standard errors (not shown). P-values are indicated in brackets. We show only covariates included in our heterogeneity analysis.

Table 6: BLP AND GATES with imputations

		ATE (β_1)	HET (β_2)	G1 (γ_1)	G5 (γ_5)	Difference (γ_5 - γ_1)
missForest	Capital stock E1	364.917 (76.579) [0.000]**	0.994 (0.217) [0.000]**	-20.371 (173.917) [1.000]	844.864 (173.341) [0.000]**	867.657 (200.479) [0.000]**
	Training hours E1	329.895 (18.448) [0.000]**	1.079 (0.183) [0.000]**	166.316 (46.188) [0.000]**	492.930 (44.068) [0.000]**	494.055 (44.068) [0.000]**
Lee bounds	Capital stock E1	140.562 (77.995) [0.144]	1.038 (0.246) [0.000]**	-322.526 (177.464) [0.129]	605.530 (175.457) [0.002]*	640.725 (206.839) [0.005]*
	Training hours E1	173.494 (19.617) [0.000]**	1.457 (0.160) [0.000]**	-132.301 (46.265) [0.008]*	393.890 (44.692) [0.000]**	395.058 (43.514) [0.000]**

Table 7: CLAN with imputations

	missForest		Lee bounds	
	Capital stock (E1)	Training hours (E1)	Capital stock (E1)	Training hours (E1)
Human capital	0.552 (0.341, 0.763) [0.000]**	0.795 (0.586, 1.008) [0.000]**	0.368 (0.151, 0.586) [0.000]**	0.700 (0.504, 0.894) [0.000]**
Working capital	0.601 (0.371, 0.844) [0.000]**	0.042 (-0.169, 0.243) [0.212]	0.138 (-0.103, 0.361) [0.066]	-0.035 (-0.227, 0.167) [0.238]
Unemployed	-0.096 (-0.198, 0.005) [0.028]	-0.093 (-0.195, 0.010) [0.018]	0.058 (-0.046, 0.162) [0.034]	0.004 (-0.100, 0.108) [0.366]
Risk aversion	-0.219 (-0.424, -0.013) [0.016]	0.518 (0.313, 0.728) [0.000]**	-0.105 (-0.318, 0.113) [0.152]	0.164 (-0.041, 0.369) [0.092]
Average distance to education facility	-0.490 (-1.673, 0.693) [0.288]	-2.246 (-3.574, -0.879) [0.000]**	-0.057 (-1.2679, 1.177) [0.408]	-1.876 (-3.223, -0.512) [0.002]*
Urban	0.073 (-0.013, 0.157) [0.078]	-0.137 (-0.220, -0.101) [0.000]**	-0.160 (-0.227, -0.048) [0.000]**	-0.216 (-0.306, -0.126) [0.000]**
Age	1.923 (0.810, 3.010) [0.000]**	-0.249 (-1.375, 0.856) [0.402]	1.574 (0.464, 2.707) [0.002]*	1.637 (0.526, 2.753) [0.002]*
Patience	0.087 (-0.125, 0.300) [0.382]	0.063 (-0.148, 0.282) [0.316]	-0.323 (-0.512, -0.126) [0.002]*	-0.577 (-0.765, -0.392) [0.000]**
Aggression	0.294 (0.089, 0.503) [0.002]*	-0.041 (-0.242, 0.160) [0.336]	0.208 (-0.004, 0.420) [0.044]	-0.043 (-0.247, 0.163) [0.592]
Female	-0.334 (-0.426, -0.241) [0.000]**	0.004 (-0.095, 0.103) [0.542]	-0.296 (-0.389, -0.203) [0.000]**	-0.040 (-0.137, 0.058) [0.470]
In school	-0.006 (-0.044, 0.033) [0.724]	-0.023 (-0.066, 0.020) [0.334]	-0.028 (-0.069, 0.011) [0.186]	-0.035 (-0.077, 0.007) [0.098]

Columns report CLAN estimates with 98% confidence intervals in parentheses and p-values in brackets.

VI. Discussion

Main hypotheses

We began our analysis with a conceptual framework of investment thresholds and long run heterogeneity. Our two main hypotheses were 1) prior human and economic capital are determinants of cash transfer outcomes and 2) heterogeneous profits may only arise in the long run due to high upfront or adjustment costs. We find strong evidence for the former. Participants with high human and working capital endowment overwhelmingly invested more in vocational training and business assets. We offer this as suggestive evidence for the presence of multiple investment thresholds. As per our conceptual framework, it is plausible that wealthier types at baseline faced a secondary threshold against which the cash transfer was supra-marginal, whereas poorer types were only able to cross an initial threshold and either saved or consumed surplus windfall. The fact that investments were made early into the intervention but were not sustained in the long run suggests a further threshold that may have discouraged business expansion.

What is puzzling is that we do not detect heterogeneity in savings and consumption. It is worth noting however that our data is both imperfect and incomplete: our measure of savings records only data up to six months before time of survey, and we also lack nondurable consumption data at endline 1. It is possible that the least enthusiastic investors directed surplus capital to nondurable consumption early into the intervention, more so given low or negative real interest rates resulting from high fees that could have discouraged saving. Such scenarios would be consistent with findings on hyperbolic discounting and consumption preferences in poor settings (Banerjee and Mullainathan, 2010; Banerjee and Duflo, 2007; Duflo et al., 2011; Moav and Neeman, 2012). However, we think this is unlikely the case. As mentioned, YOP is self-selective and likely filtered out the highly impatient, while the presence of borrowing constraints should motivate even non-prudent types to engage in precautionary savings over short-term consumption. Indeed, Bazzi et al (2015) study a large-scale UCT in Indonesia and find evidence of how imperfect credit markets encourage precautionary savings despite expectations of future transfers. Given our lack of data much of this remains speculative. We view this as a limitation of our study that can be addressed with more and better data.

For our second hypothesis, we fail to detect evidence of heterogeneous profits at both endlines. However, cash earnings can understate total earnings since it does not capture nonmarket household production like agriculture. If HPs report greater agricultural production than LPs then real profits will be understated. We perform a quick calculation and find that on average, our participants report substantial increases in agricultural labor hours four years after the intervention, but this is not induced by treatment. Differences in agricultural hours between HPs and LPs at endline 2 are also insignificant.

Despite this we remain optimistic as to the presence of heterogeneity for two reasons. First, because we are able to detect heterogeneity for investment outcomes, it makes sense that the most prolific investors should also earn higher profits in the long run. The exception is if vocational skills were not rewarded by the market though this is unlikely given that average treatment effects were high. Second, we find that our HET estimate increases and its p-value reduces substantially with time, suggesting that heterogeneity may surface at subsequent endlines[6]. It is difficult to say why heterogeneity should take this long to manifest. Adjustment costs could be one reason – trades can take a while to master and apprenticeships are often low-paying. In Uganda, apprenticeships can last from under a year to over 3 years (Livingstone and Kemigisha, 1995); two

[6]BEA conduct a follow up study of the YOP 9 years after the intervention and find that earnings for control groups converge with treatment (Blattman et al., 2020). We cannot say for certain if this invalidates our hypothesis given the lack of data. Further research on this is likely to yield conclusive answers.

years[7] after the intervention more than 40% of participants were still enrolled in vocational training. In addition, the transition to self-employment is often accompanied by a loss in wages that cannot be offset in the short run. Hardy et al (2019) study a government sponsored apprenticeship program in Ghana and find that while the program shifted youth out of wage work and into self-employment, loss of wage income was not offset by increases in self-employment profits 4-5 years after the intervention. It is also possible that the intervention crowded out the market and subsequently dulled earnings for the most gung-ho investors. Finally, we rule out unduly high upfront costs as a reason since treatment effects on investment are high and participants do not delay investment, instead preferring to invest shortly after receiving funds.

Thresholds

Our results allude to the presence of multiple investment thresholds. Knowing what these thresholds represent is also important, since a cash transfer cannot spur productivity if the supply of more efficient technology is non-existent at time of intervention. Our obvious candidates are business assets and vocational training. We find that almost all participants invest in training but only wealthier types invest in capital stock. It is plausible that the windfall pushes participants beyond an initial threshold in the form of training costs, but only wealthier types are able to afford the capital stock necessary to cross a secondary investment threshold. It is also possible that by relaxing capital constraints, the cash transfer lowers other barriers like transport costs that make enterprise costly.

Of our sample, less than 8% reported having prior vocational training, but almost all participants were induced by the treatment to seek training. Clearly, vocational training was not an option in the absence of grants. Yet, training institutes were reported to be common and existed for at least five years before the intervention. YOP participants are also slightly wealthier and more educated than their peers, and so it is unlikely that the cost of training should be unduly high. We find this puzzling. One explanation is that participants view vocational training and business investment as partly complementary. Our setting is one in which the economy is largely agrarian and the labor market for skilled employment may be tight – less than 6% of participants engaged in skilled trades at baseline. Self-employment represents a more viable option for those practicing skilled trades, but capital constraints and high start-up costs may discourage investment in enterprise. Learning a trade makes little sense if one cannot make a living off it. Unsurprisingly, we find that those practicing skilled trades at baseline also score about 0.4 standard deviations higher in terms of working capital than their counterparts, and this difference is highly significant. Fafchamps and Woodruff (2017) find suggestive evidence of such complementarities in Ghana. They ran a small business competition in which winners were selected to receive only individual training but not cash and found no significant impact of the training on firm growth.

The question then arises as to why some participants bother at all to invest in training if they cannot afford capital stock. Our ideas of complementarity assumes weak labor demand at baseline for skilled trade practitioners. It is possible that by encouraging investment in vocational enterprise, the intervention itself generates demand for skilled labor and relaxes complementarity. In this case complementarity binds investment in training only before the intervention. This is largely speculative, however, and should be tested in further settings.

[7] We lack data on enrolment at endline 2.

Alternative sources of heterogeneity

An auxiliary aim of this paper was to detect new and unexpected sources of heterogeneity. We find that HPs for our training outcome lived closer to education facilities, which could point to some costs associated with travelling – both financial and intangible – that may inhibit investment in training. For example, Muralidharan and Prakash (2017) study a government rollout of bicycles for school-going girls in India and suggest that a reduction in travelling time and safety costs was responsible for improved educational outcomes. We are uncertain what travel constraints the YOP alleviates, if at all. That participants living closer to educational facilities train more hours suggests the presence of mobility constraints that a cash grant alone is unable to alleviate. At baseline, the median participant in our sample lived 8 km away from the nearest educational facility – a nontrivial distance by foot. Time, safety, and weather are all plausible constraints. At the same time, YOP participants are highly transient and our results may be spurious.

We also find that HPs for capital stock investment were slightly older and overwhelmingly male. This is unsurprising. Age is often correlated with wealth and human capital[8] – both strong predictors of investment – and as outlined in BEA, women may face traditional norms and household pressures to share funds that bind investment outcomes. What is surprising is that this does not bind investment in training. One explanation is that the intervention generates future demand for female labor and spurs investment in women's educational outcomes, mirroring findings by Qian (2008) and Jensen (2012). At the same time, the fact that women continue investing in training despite social (and capital) constraints may be further proof that training costs are not prohibitively high.

Limitations

Our study is not without limitations. First, even if heterogeneity is detected it is not possible to infer the magnitude of the effect due to ML's poor inferential properties. For instance, it is possible to say that wealthier individuals are more likely to benefit from the intervention but not by how much. Relatedly, GML merely extracts the average characteristics of the best and worst performing groups and does not isolate the treatment effect heterogeneity for each characteristic. Since age and wealth are correlated, the treatment may have generated outsized effects for older participants, and wealth is merely a confounder[9]. To be sure, our findings are supported by extant literature and a conceptual framework. From this we reason qualitatively that they are robust to confounding. Nevertheless, it remains to be seen if these results can be replicated using other ML methods.

Second, the YOP was unsupervised but not unconditional; applicants were screened for initiative through a proposal, though usage of funds was not monitored. In this sense it was neither wholly conditional nor unconditional. It is unclear whether these restrictions played an important role in driving investment and may impair generalizability to truly unconditional settings.

Third, our results do not survive conservative imputations though we attribute this to heterogeneous attrition. This was exacerbated by the self-selective nature of YOP which likely limited the variation of our covariates. Following Karlan and Valdivia (2011), we run less conservative imputations and find results broadly consistent with our main analysis (Appendix A4). Having more computational power and tuning at

[8] We run pairwise correlation tests for all variables and find age to be correlated with both (Appendix A5).

[9] It is also possible that outsized impacts on capital stock investment for wealthier types is driven solely by gender differences. Our pairwise correlation tests in Appendix A5 show insignificant correlation between gender and wealth. Nevertheless, we rerun GML on gendered subgroups and again find no evidence of such (Appendix A6).

each split may improve our findings though this remains to be tested. However, we view our ability to detect heterogeneity – despite limited variation – as an important outcome that should not be quickly overlooked.

Finally, computationally expensive methods like GML are meant to complement experimental findings, not replace them. It is dangerous to prescribe policy based on data that is a further level removed from the main findings. That said, what it can do is highlight new or existing sources of heterogeneity that may be ripe for future research, particularly in the context of northern Uganda.

VII. Conclusion

On the whole, our results adds to the growing literature on threshold effects and offer unexpected sources of heterogeneity in cash transfer settings. We find that high initial ability and wealth leads to more investment for YOP participants. We also find evidence of multiple investment thresholds, likely in the form of business assets and vocational training. We fail to detect heterogeneity in profits though we are optimistic that subsequent endlines may be able to do so.

Our findings carry broader significance. Technical and vocational education and training (TVET) is a contentious policy tool in Africa. Critics argue that vocational education is inflexible and suffers from labor market mismatch (Middleton et al., 1993; Tilak, 2002). By contrast our results show that almost all participants invest in training when they are able to and average impacts on earnings are positive. What is more noteworthy is that the intervention itself may have relaxed social constraints that bound women's investment in vocational training. Under the right circumstances then TVET interventions can have sizeable impacts on productivity and gender parity.

Absent better loan terms a cash grant remains the quickest way to relax capital constraints. On the flipside, developing countries are resource constrained and reliance on grants is unsustainable. In answer to this, our study contributes to existing evidence on possible sources of heterogeneity that can improve targeting and efficiency. In particular, our discovery of multiple thresholds and outsized impacts for wealthier types disputes conclusions from BEA and stands in contrast to growing enthusiasm for cash transfers to the poorest (Bandiera et al., 2017; Blattman et al., 2014). Our point is not that cash transfers should not target the poor; we merely suggest that pro-poor targeting may compromise productive impacts of cash grants. Whether this happens will depend on the size of a transfer in relation to threshold effects. As noted by Prifti et al (2020), the implication for the policymaker is to consider complementary interventions for the destitute that raises future productivity.

Finally, lessons from Uganda are relevant to the wider region. Uganda is a fast growing, young country with high rural populations and rural poverty. YOP participants are not only young, but also credit-constrained and recovering from political instability, traits ubiquitous to many poor communities in sub-Saharan Africa. Young people are key agents of change. So it is everywhere, but more so in Africa. How we assist them is important.

Appendix

A1. Construction of summary indices

Index construction was performed by BEA using a range of baseline covariates. Variables are weighted using regression estimates in order of importance in predicting future economic success in the control group, then standardized. Full details can be found in their Online Appendix.

The human capital index comprises measures of (i) educational attainment, (ii) a literacy indicator, (iii) an indicator for prior vocational training, (iv) performance on a digit recall test measuring working memory, and (v) a measure of physical disabilities assembled from responses how easily the respondent can perform a number of activities of daily life.

The working capital index is a weighted average of baseline measures of (i) savings stock, (ii) the stock of loans outstanding, (iii) cash earnings, (iv) perceived access to a 100,000 UGX loan, (v) perceived access to a 1 million UGX loan, and (vi) indices of housing quality and assets (similar to the index of wealth endline measure).

The patience index comprises of 10 self-reported measures of impulsiveness and patience, including self-reported willingness to wait long periods for material goods, to spend money "too quickly," to put off hard or costly tasks, or to resist temptation.

The risk aversion index is a weighted average of baseline measures of 8 self-reported measures of risky behavior, including (i) walking alone at night, (ii) engaging in unprotected sex, (iii) investing in a risky business that could have high profits, and (iv) choosing not to sleep under a mosquito net, among others.

Finally, the aggression index was aggregated from multiple responses from survey questions rooted in psychological survey instruments on U.S. populations and were adapted to the Ugandan context by the authors.

A2. Propensity score matching

We estimate participants' propensity score using a probit model and the list of covariates in table 1. Following Stuart et al (2011), optimal pair matching was performed by minimizing the sum of absolute pairwise distances in the matched sample. We also attempted to match using nearest neighbours but this returned poorer results. Table 8 displays regression differences after optimal pair matching. We find that all covariates are balanced apart from ability to obtain a 100,000 UGX loan. Reassuringly, we find that our covariates of interest - particularly human and working capital at baseline - are highly balanced.

Table 8: Post propensity score matching test of balance

	Control		Treatment		Regression difference	
	Mean	Std. dev.	Mean	Std. dev.	Mean	p-value
Grant amount applied for, USD	7527.042	2104.467	7271.754	2030.932	96.070	0.463
Group size	22.431	6.860	21.501	6.697	0.264	0.605
Grant amount per member, USD	357.749	151.082	372.689	159.427	9.214	0.412
Group existed before application	0.449	0.498	0.485	0.500	0.031	0.474
Group age, in years	3.793	1.951	3.822	1.900	-0.021	0.905
Within-group heterogeneity (z-score)	-0.028	0.923	0.015	1.059	-0.029	0.732
Quality of group dynamic (z-score)	-0.017	1.016	0.024	0.981	0.047	0.556
Distance to educational facilities (km)	6.830	6.523	7.160	5.459	0.412	0.415
Individual unfound at baseline	1.000	0.000	1.000	0.000	0.000	0.309
Age at baseline	24.756	5.225	25.054	5.273	0.101	0.718
Female	0.349	0.477	0.325	0.468	-0.021	0.426
Large town/urban area	0.232	0.422	0.195	0.397	-0.019	0.575
Risk aversion index (z-score)	-0.022	1.002	-0.025	1.000	-0.014	0.747
Any leadership position in group	0.279	0.449	0.278	0.448	-0.010	0.528
Group chair or vice-chair	0.104	0.305	0.113	0.316	0.007	0.510
Weekly employment, hours	10.724	15.842	11.213	15.440	0.396	0.619
All nonagricultural work	6.028	12.497	5.669	11.386	-0.481	0.405
Casual labor, low skill	1.033	5.201	1.070	4.992	-0.100	0.667
Petty business, low skill	2.251	6.970	2.418	6.815	0.216	0.503
Skilled trades	1.788	8.431	1.525	7.734	-0.339	0.388
High-skill wage labor	0.045	0.578	0.071	0.727	0.021	0.442
Other nonagricultural work	0.910	4.776	0.585	3.831	-0.279	0.118
All agricultural work	4.696	10.122	5.545	10.355	0.877	0.081
Weekly household chores, hours	9.009	17.622	8.846	16.301	0.358	0.688
Zero employment hours in past month	0.477	0.500	0.429	0.495	-0.029	0.311

Table 8: *(continued)*

	Control		Treatment		Regression difference	
	Mean	Std. dev.	Mean	Std. dev.	Mean	p-value
Main occupation is nonagricultural	0.258	0.438	0.271	0.445	-0.004	0.866
Engaged in a skilled trade	0.061	0.239	0.058	0.233	-0.007	0.549
Currently in school	0.045	0.207	0.039	0.193	-0.006	0.519
Highest grade reached at school	7.954	2.923	7.797	3.012	-0.110	0.450
Literate	0.751	0.433	0.716	0.451	-0.029	0.153
Received prior vocational training	0.073	0.261	0.081	0.273	0.018	0.108
Digit recall test score	4.165	2.000	4.015	1.963	-0.042	0.610
Index of physical disability	8.684	2.527	8.647	2.254	-0.125	0.359
Durable assets (z-score)	-0.159	0.956	-0.092	1.039	0.054	0.274
Savings in past 6 mos. (000s 2008 UGX)	19.353	98.450	25.651	114.678	4.684	0.268
Monthly gross cash earnings (000s 2008 UGX)	62.467	129.337	67.042	135.593	6.042	0.372
Can obtain 100,000 UGX ($58) loan	0.335	0.472	0.388	0.488	0.043	0.035
Can obtain 1,000,000 UGX ($580) loan	0.099	0.299	0.112	0.315	0.006	0.661
Human capital (z-score)	-0.062	0.867	0.019	0.965	0.060	0.118
Working capital (z-score)	0.037	0.992	-0.029	1.027	-0.038	0.445
Patience index (z-score)	-0.021	1.040	-0.018	0.984	0.032	0.452
Aggression index (z-score)	-0.010	1.012	0.013	0.990	-0.009	0.847

A3. Analysis with interaction terms

As explained in section IV, our choice of ML over conventional methods like interaction terms was motivated by traditional issues of overfit and multiple hypothesis testing associated with interacting covariates. However, for benchmarking purposes, we attempt to replicate our results in section V using interaction terms. We deviate from BEA by unpooling the sample and we only perform regression for outcomes that displayed heterogeneity in our main findings - capital stock and training hours at endline 1. Table 9 displays our findings. For our capital stock and training hours outcomes we detect heterogeneity only for gender and human capital respectively.

Table 9: Interaction terms

	Capital stock E1	Training hours E1
	(1)	(2)
Assigned	418.591	363.135^{***}
	(402.716)	(91.367)
Working capital	$2,414,754.382$	$3,355,430.646$
	$(20,798,085.273)$	$(5,591,096.841)$
Human capital	$-18,733,556.449$	$1,456,808.079$
	$(18,361,118.577)$	$(5,086,413.286)$
Female	-109.908	7.142
	(73.944)	(24.641)
Age	-1.454	-2.478
	(10.264)	(1.879)
Avg distance to education	0.757	3.777
	(5.913)	$(1.520)^{**}$
Assigned:Working capital	-61.575	-4.856
	(115.291)	(17.171)
Assigned:Human capital	24.013	31.079^{**}
	(66.799)	(15.945)
Assigned:Female	-368.684^{**}	-12.418
	(146.266)	(39.895)
Assigned:Age	5.955	0.947
	(14.654)	(3.291)
Assigned:Avg distance to education	-4.558	-4.881
	(12.045)	(3.129)

***p<.001, **p<.05, *p<.10
We regress on all baseline covariates (not shown) with group clustered standard errors.

A4. Less conservative imputations

Following Karlan and Valdivia (2011), we impute bounds in varying orders of magnitude. For the control group we impute the found control mean plus 0.25 and 0.5 standard deviations of the found control distribution and for the treatment group we impute the found treatment mean minus 0.25 and 0.5 standard deviations of the found treatment distribution. Table 10 reports BLP estimates for the average treatment effect (ATE) and heterogeneity loading parameter (HET) with imputations. Table 11 reports CLAN estimates. We find that using less conservative imputations - in particular with 0.25 SD - produces results that are broadly in line with our main findings.

Table 10: BLP AND GATES with imputations

		ATE (β_1)	HET (β_2)	G1 (γ_1)	G5 (γ_5)	Difference (γ_5 - γ_1)
0.25SD	Capital stock E1	182.767	0.904	-204.897	578.414	590.050
		(77.028)	(0.251)	(173.624)	(172.917)	(198.565)
		[0.038]	[0.001]**	[0.457]	[0.002]*	[0.007]*
	Training hours E1	238.994	1.076	64.106	399.663	401.432
		(19.506)	(0.188)	(46.257)	(45.773)	(46.119)
		[0.000]**	[0.000]**	[0.326]	[0.000]**	[0.000]**
0.5SD	Capital stock E1	56.677	1.349	-468.302	597.480	623.920
		(78.167)	(0.287)	(176.174)	(175.749)	(209.167)
		[0.925]	[0.000]**	[0.015]	[0.002]*	[0.006]*
	Training hours E1	213.822	1.059	-14.667	366.921	366.715
		(19.899)	(0.165)	(47.135)	(45.976)	(45.201)
		[0.000]**	[0.000]**	[1.000]	[0.000]**	[0.000]**

Table 11: CLAN with imputations

	0.25SD		0.5SD	
	Capital stock (E1)	Training hours (E1)	Capital stock (E1)	Training hours (E1)
Human capital	0.397 (0.194, 0.609) [0.000]**	0.706 (0.505, 0.901) [0.000]**	0.329 (0.114, 0.542) [0.000]**	0.621 (0.427, 0.814) [0.000]**
Working capital	0.303 (0.065, 0.527) [0.002]*	-0.019 (-0.217, 0.181) [0.212]	0.045 (-0.187, 0.283) [0.064]	-0.007 (-0.192, 0.191) [0.106]
Unemployed	-0.018 (-0.120, 0.086) [0.316]	-0.089 (-0.191, 0.013) [0.018]	0.086 (-0.015, 0.187) [0.024]	-0.086 (-0.189, 0.017) [0.050]
Risk aversion	-0.201 (-0.414, 0.012) [0.048]	0.253 (0.047, 0.459) [0.006]*	-0.160 (-0.369, 0.052) [0.082]	0.161 (-0.043, 0.370) [0.126]
Average distance to education facility	-0.185 (-1.438, 1.044) [0.270]	-1.954 (-3.273, -0.642) [0.002]*	-0.287 (-1.444, 0.892) [0.218]	-1.602 (-2.877, -0.323) [0.008]*
Urban	-0.085 (-0.175, 0.005) [0.038]	-0.170 (-0.257, -0.083) [0.000]**	-0.154 (-0.241, -0.066) [0.000]**	-0.171 (-0.260, -0.083) [0.000]**
Age	1.979 (0.842, 3.078) [0.000]**	1.186 (0.048, 2.307) [0.026]	2.027 (0.888, 3.063) [0.000]**	1.532 (0.343, 2.650) [0.006]*
Patience	-0.282 (-0.488, -0.076) [0.004]*	-0.531 (-0.731, -0.328) [0.000]**	-0.368 (-0.572, -0.162) [0.000]**	-0.678 (-0.864, -0.489) [0.000]**
Aggression	0.264 (0.062, 0.467) [0.004]*	0.002 (-0.202, 0.208) [0.402]	0.240 (0.034, 0.445) [0.014]	0.052 (-0.155, 0.252) [0.534]
Female	-0.279 (-0.373, -0.186) [0.000]**	-0.012 (-0.110, 0.087) [0.620]	-0.285 (-0.377, -0.192) [0.000]**	0.013 (-0.085, 0.110) [0.560]
In school	-0.031 (-0.073, 0.011) [0.164]	-0.020 (-0.062, 0.021) [0.364]	-0.043 (-0.084, -0.003) [0.186]	-0.024 (-0.066, 0.014) [0.026]

Columns report CLAN estimates with 98% confidence intervals in parentheses and p-values in brackets.

A5. Correlation matrix

We perform pairwise correlation tests for all covariates of interest. Table 12 reports correlation scores for all pairwise permutations. We find that as expected, age is correlated with our measures of working and human capital and the p-values are highly significant ($p<0.001$). We also find that our indicator for gender (female) is highly correlated with human capital ($p<0.001$) though there is no significant correlation between gender and working capital ($p=0.116$). We only report correlation scores and do not show p-values here.

Table 12: Pairwise correlation

	Human capital	Working capital	Unemployed	Risk aversion	Avg. distance to education facility	Urban	Age	Patience	Aggression	Female	In school
Human capital	1.000	0.186	0.010	0.042	-0.045	0.141	0.060	0.106	-0.079	-0.211	0.064
Working capital	0.186	1.000	-0.028	0.040	-0.056	0.085	0.125	0.055	-0.037	-0.038	0.030
Unemployed	0.010	-0.028	1.000	0.027	-0.030	0.034	-0.102	0.021	-0.026	0.059	0.047
Risk aversion	0.042	0.040	0.027	1.000	-0.082	-0.087	0.000	-0.032	-0.047	0.046	-0.032
Average distance to education facility	-0.045	-0.056	-0.030	-0.082	1.000	-0.164	0.026	-0.003	-0.028	-0.006	-0.023
Urban	0.141	0.085	0.034	-0.087	-0.164	1.000	-0.044	0.095	0.001	0.050	0.042
Age	0.060	0.125	-0.102	0.000	0.026	-0.044	1.000	-0.010	-0.043	-0.129	-0.122
Patience	0.106	0.055	0.021	-0.032	-0.003	0.010	-0.010	1.000	-0.047	-0.068	0.044
Aggression	-0.079	-0.037	-0.026	-0.047	-0.028	0.001	-0.043	-0.047	1.000	0.053	-0.006
Female	-0.211	-0.038	0.059	0.046	-0.006	0.050	-0.129	-0.068	0.053	1.000	-0.024
In school	0.064	0.030	0.047	-0.032	-0.023	0.042	-0.122	0.044	-0.006	-0.024	1.000

A6. GML with gendered subgroups

Our pairwise correlation tests show no significant correlation between gender and working capital. Nevertheless, it is possible that outsized impacts in investment outcomes are dominated by gender differences. We divide our sample into gendered subgroups and rerun GML. At $n = 659$, our sample size for females is small and the lack of variation prevents us from running GML for our female subgroup. We show only CLAN results for males instead. We find that consistent with our main findings, wealthier types invest in business assets more than their counterparts and this is not driven by gender, though the estimate for human capital is reduced and its p-value slightly insignificant.

Table 13: CLAN (Males)

	Capital stock (E1)	Training hours (E1)
Human capital	0.261 (-0.010, 0.534) [0.022]	0.797 (0.516, 1.068) [0.000]**
Working capital	0.789 (0.437, 1.143) [0.000]**	0.389 (0.084, 0.694) [0.006]*
Unemployed	0.082 (-0.062, 0.225) [0.062]	-0.038 (-0.183, 0.105) [0.192]
Risk aversion	-0.202 (-0.507, 0.090) [0.164]	0.450 (0.161, 0.745) [0.000]**
Average distance to education facility	-0.872 (-2.391, 0.633) [0.224]	-3.606 (-5.308, -1.998) [0.000]**
Urban	0.030 (-0.079, 0.134) [0.398]	-0.008 (-0.122, 0.109) [0.694]
Age	1.832 (0.366, 3.284) [0.008]*	0.680 (-0.814, 2.146) [0.316]
Patience	0.081 (-0.198, 0.373) [0.554]	0.257 (-0.037, 0.551) [0.074]
Aggression	0.489 (0.219, 0.769) [0.000]**	-0.106 (-0.397, 0.179) [0.422]
In school	-0.023 (-0.079, 0.032) [0.526]	0.039 (-0.020, 0.100) [0.250]

References

AIR, 2014. Zambia's Child Grant Program: 36-Month Impact Report. American Institutes for Research, Washington, DC.

Angelucci, M., Karlan, D., Zinman, J., 2015. Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco. American Economic Journal: Applied Economics 7, 151–182.

Atkin, D., Chaudhry, A., Chaudry, S., Khandelwal, A.K., Verhoogen, E., 2017. Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan*. The Quarterly Journal of Economics 132, 1101–1164.

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, 811–870.

Banerjee, A., Breza, E., Duflo, E., Kinnan, C., 2019. Can Microfinance Unlock a Poverty Trap for Some Entrepreneurs? (No. w26346). National Bureau of Economic Research.

Banerjee, A., Duflo, E., Glennerster, R., Kinnan, C., 2015. The Miracle of Microfinance? Evidence from a Randomized Evaluation. American Economic Journal: Applied Economics 7, 22–53.

Banerjee, A., Mullainathan, S., 2010. The Shape of Temptation: Implications for the Economic Lives of the Poor (No. 15973). National Bureau of Economic Research, Inc.

Banerjee, A.V., Duflo, E., 2007. The Economic Lives of the Poor. Journal of Economic Perspectives 21, 141–168.

Banerjee, A.V., Duflo, E., 2011. Poor economics: A radical rethinking of the way to fight global poverty, 1st ed. ed. PublicAffairs, New York.

Banerjee, A.V., Newman, A.F., 1993. Occupational Choice and the Process of Development. Journal of Political Economy 101, 274–298.

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T., 2016. Cash transfers: What does the evidence say? Overseas Development Institute.

Bazzi, S., Sumarto, S., Suryahadi, A., 2015. It's all in the timing: Cash transfers and consumption smoothing in a developing country. Journal of Economic Behavior & Organization 119, 267–288.

Blattman, C., Fiala, N., Martinez, S., 2014. Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda*. The Quarterly Journal of Economics 129, 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, 287–304.

Blattman, C., Green, E.P., Jamison, J.C., Lehmann, M.C., Annan, J., 2015. The Returns to Microenterprise Support Among the Ultra-Poor: A Field Experiment in Post-War Uganda (Working {Paper} No. 21310). National Bureau of Economic Research.

Carter, M.R., Barrett, C.B., 2006. The economics of poverty traps and persistent poverty: An asset-based approach. The Journal of Development Studies 42, 178–199.

Chernozhukov, V., Demirer, M., Duflo, E., Fernandez-Val, I., 2017. Generic machine learning inference on heterogenous treatment effects in randomized experiments. The IFS.

Covarrubias, K., Davis, B., Winters, P., 2012. From protection to production: Productive impacts of the Malawi Social Cash Transfer scheme. Journal of Development Effectiveness 4, 50–77.

Crépon, B., Devoto, F., Duflo, E., Parienté, W., 2015. Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Experiment in Morocco. American Economic Journal: Applied Economics 7, 123–150.

Daidone, S., Davis, B., Dewbre, J., Gonzalez-Flores, M., Handa, S., Seidenfeld, D., Tembo, G., 2014. Zambia's Child Grant Programme: 24-month impact report on productive activities and labour allocation. FAO, Rome.

Davis, B., Handa, S., Arranz, M.R., Stampini, M., Sant'Anna, S., Winters, P., 2002. Conditionality and the impact of program design on household welfare: Comparing two diverse cash transfer programs in rural Mexico. FAO, Rome.

Duflo, E., Kremer, M., Robinson, J., 2011. Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. American Economic Review 101, 2350–2390.

Fafchamps, M., Woodruff, C., 2017. Identifying Gazelles: Expert Panels vs. Surveys as a Means to Identify Firms with Rapid Growth Potential. World Bank Economic Review 31, 670–686.

Gertler, P.J., Martinez, S.W., Rubio-Codina, M., 2012. Investing Cash Transfers to Raise Long-Term Living Standards. American Economic Journal: Applied Economics 4, 164–192.

Ghatak, M., 2015. Theories of Poverty Traps and Anti-Poverty Policies. The World Bank Economic Review 29, S77–S105.

Government of Uganda, 2007. National Peace, Recovery and Development Plan for Northern Uganda: 2006–2009. Kampala, Uganda.

Handa, S., Natali, L., Seidenfeld, D., Tembo, G., Davis, B., 2018. Can unconditional cash transfers raise long-term living standards? Evidence from Zambia. Journal of Development Economics 133, 42–65.

Handa, S., Park, M., Darko, R.O., Osei-Akoto, I., Davis, B., Daidone, S., 2014. Livelihood Empowerment Against Poverty Program Impact Evaluation. Carolina Population Center, University of North Carolina, Chapel Hill, NC.

Hardy, M.L., Mbiti, I.M., Mccasland, J.L., Salcher, I., 2019. The Apprenticeship-to-Work Transition: Experimental Evidence from Ghana. The World Bank.

Hennessy, D.A., 1998. The Production Effects of Agricultural Income Support Policies under Uncertainty. American Journal of Agricultural Economics 80, 46–57.

Jelliffe, J.L., Bravo-Ureta, B.E., Deom, C.M., Okello, D.K., 2018. Adoption of High-Yielding Groundnut Varieties: The Sustainability of a Farmer-Led Multiplication-Dissemination Program in Eastern Uganda. Sustainability 10, 1–21.

Jensen, R., 2012. Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India *. The Quarterly Journal of Economics 127, 753–792.

Karlan, D., Osei, R., Osei-Akoto, I., Udry, C., 2014. Agricultural Decisions after Relaxing Credit and Risk Constraints *. The Quarterly Journal of Economics 129, 597–652.

Karlan, D., Valdivia, M., 2011. Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions. The Review of Economics and Statistics 93, 510–527.

Kraay, A., McKenzie, D., 2014. Do Poverty Traps Exist? Assessing the Evidence. Journal of Economic Perspectives 28, 127–148.

Künzel, S.R., Sekhon, J.S., Bickel, P.J., Yu, B., 2019. Metalearners for estimating heterogeneous treatment effects using machine learning. Proceedings of the National Academy of Sciences 116, 4156–4165.

Lee, D.S., 2009. Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. The Review of Economic Studies 76, 1071–1102.

Livingstone, I., Kemigisha, S., 1995. Some Evidence on Informal Sector Apprenticeship in Uganda. The Journal of Modern African Studies 33, 339–342.

McKenzie, D., 2017. Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition. American Economic Review 107, 2278–2307.

McKenzie, D.J., Woodruff, C., 2006. Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises. Economic Development and Cultural Change 55, 3–42.

Meager, R., 2019. Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments. American Economic Journal: Applied Economics 11, 57–91.

Mel, S. de, McKenzie, D., Woodruff, C., 2008. Returns to Capital in Microenterprises: Evidence from a Field Experiment*. The Quarterly Journal of Economics 123, 1329–1372.

Mendola, M., 2007. Agricultural technology adoption and poverty reduction: A propensity-score matching analysis for rural Bangladesh. Food Policy 32, 372–393.

Middleton, J., Ziderman, A., Adams, A.V., 1993. Skills for Productivity: Vocational Education and Training in Developing Countries. Oxford University Press, Washington, D.C.

Moav, O., Neeman, Z., 2012. Savings Rates and Poverty: The Role of Conspicuous Consumption and Human Capital. The Economic Journal 122, 933–956.

Muralidharan, K., Prakash, N., 2017. Cycling to School: Increasing Secondary School Enrollment for Girls in India. American Economic Journal: Applied Economics 9, 321–350.

Pavlova, A., 2001. Adjustment Costs, Learning-by-Doing, and Technology Adoption Under Uncertainty ({SSRN} {Scholarly} {Paper} No. ID 308747). Social Science Research Network, Rochester, NY.

Phimister, E., 1995. The impact of borrowing constraints on farm households: A life-cycle approach *. European Review of Agricultural Economics 22, 61–86.

Prifti, E., Daidone, S., Pace, N., Davis, B., 2020. Heterogeneous impacts of cash transfers on farm profitability. Evidence from a randomised study in Lesotho. European Review of Agricultural Economics 47, 1531–1558.

Qian, N., 2008. Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance*. The Quarterly Journal of Economics 123, 1251–1285.

Seidenfeld, D., Handa, S., 2011. Results of the Three Year Impact Evaluation of Zambia's Cash Transfer Program in Monze District: Final Report. American Institutes for Research, Washington, DC.

Stekhoven, D.J., Bühlmann, P., 2012. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics 28, 112–118.

Stuart, E., King, G., Imai, K., Ho, D., 2011. MatchIT: Nonparametric Preprocessing for Parametric Causal Inference. Journal of Statistical Software 42.

Tilak, J.B.G., 2002. Education and Poverty. Journal of Human Development 3, 191–207.

Todd, J.E., Winters, P.C., Hertz, T., 2010. Conditional Cash Transfers and Agricultural Production: Lessons from the Oportunidades Experience in Mexico. The Journal of Development Studies 46, 39–67.

Wager, S., Athey, S., 2018. Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. Journal of the American Statistical Association 113, 1228–1242.

Weeks, M., Christiansen, T.G., 2020. Understanding the Distributional Aspects of Microcredit Expansions. arXiv:2011.10509 [econ, q-fin].