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

Candidate number: 18195
Group 16
Word count: 1499

Motivation

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 and credit access alleviate this by reducing the minimum wealth level at which households can switch into more efficient technology, allowing them to escape the poverty trap. 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) programmes 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.

Literature review

Because the poor are often credit-constrained (Berg, 2013), multiple studies have focused on initial wealth and human capital in their search for heterogeneity. In the absence of cheap credit, an unexpected windfall should benefit high-ability but capital-starved types more compared to the wealthier, less credit-constrained. However, empirical literature on heterogeneity in UCTs have turned up conflictual findings. A randomised study of a UCT in northern Uganda (Blattman et al., 2014) found that economic impacts were decreasing in initial wealth and increasing in entrepreneurial ability, though neither of these were statistically significant. Macours et al (2013) study an intervention in Nicaragua and find that earnings from and involvement in non-farm businesses and employment were not affected by initial asset endowment when the programme was switched from universal cash grants to targeted subgroups. Conversely, de Mel et al (2008) found that randomised shocks to capital stock in Sri Lankan microenterprises produced significantly varying returns depending on initial entrepreneurial ability and household wealth. Likewise, a study by Prifti et al (2020) on Lesotho's Child Grant Programme found greater economic outcomes in both the crop and livestock sector for those with greater land and labour capacity. These mixed findings cast doubt on the presence and direction of heterogeneity. Further, preoccupation with economic and human capital endowment may cloud alternative sources of heterogeneity. Related research in microcredit suggests that aside from wealth and ability, ownership of a business prior to credit relaxation is a key predictor of economic success (Banerjee et al., 2019; Meager, 2019). These findings raise important questions: *Does prior economic and human capital endowment determine UCT outcomes for the poor? Otherwise, what other sources of heterogeneity could there be and why so?*

Research design

Data:

To answer these questions, the proposed study will postprocess existing experimental data from Blattman et al (2014). In the experiment the authors conduct a randomised evaluation of a UCT in Northern Uganda known as the Youth Opportunities Programme (YOP). The programme was designed to help poor and unemployed adults become self-employed artisans but carried restrictions: to qualify, applicants had to form groups and submit a business proposal subject to approval, though utilisation of funds was not monitored. The first endline survey was taken in 2010, two years after the start of the experiment, and the second endline was carried out in 2012. Outcomes of interest are individual consumption, earnings, and asset value at both endlines.

An important consideration when selecting the dataset was the study's external validity. In 2006 it was estimated 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 (Government of Uganda, 2007). At the time of the intervention in 2008, beneficiaries of the YOP had recently emerged from regional conflicts and had close to no access to formal finance and savings instruments. The setting thus closely resembles the credit-constrained environment ubiquitous to many poor communities and bolsters generalisability.

While the authors fail to detect heterogeneity in the original paper, this study will attempt to do so using newer statistical methods and richer covariates. The reasons for postprocessing this dataset are threefold. First, conventional methods used in the paper for detecting heterogeneity like subgroup analysis and interaction terms are sensitive to sample size and model selection. Second, because the intervention is self-selective by construction, being able to detect heterogeneity may allow us to extrapolate even stronger effects to the wider population. Third, the use of pre-analysis plans – designed to prevent 'P-hacking' – limits researchers to using only prespecified covariates, wasting rich covariates in the process. In the original experiment the authors limited themselves to using only 40-50 covariates out of over 300 in the full dataset. While the search for heterogeneity should be driven by theory, this 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.

Empirical strategy:

This study will employ a generic machine learning (hereafter known as GML) method by Chernozhukov et al (2018) to detect treatment effect heterogeneity. The advantage with employing ML is its ability to handle models with large number of covariates. In a standard OLS model, introducing all available covariates and their various permutations for interaction terms will likely lead to overfitting. 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 utilised to mitigate estimation uncertainty conditional on the data split, while median values are taken to account for splitting uncertainty. Alternative ML methods for postprocessing data like Meta-learners (Künzel et al., 2019) and Causal Forests (Wager and Athey, 2018) are available, 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 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 treatment propensity, 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 G_k 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, positively significant coefficient for γ_K and a small insignificant coefficient for γ_1.

CLAN: Once heterogeneity is established, we report the median of the average characteristics of choice for G_1 and G_K and compare differences between the two using a two sample t test.

Analysis will be performed on both endlines, which allows us to test if heterogeneity manifests in the short or long term, if at all. To limit attrition bias, robustness checks will be performed using conservative imputations following Manski (1990) and Lee (2009).

Limitations

The 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 younger individuals are more likely to benefit from UCTs but not by how much. Second, because survey responses are self-reported they may be biased upwards. If high-ability or wealthier individuals are more likely to inflate their responses then heterogeneity may be spurious. Third, at $n < 2000$ for each endline the dataset may still be too small for GML to be effective; datasets comprising hundreds of thousands of observations are not uncommon in ML applications. Fourth, the YOP requires participants to self-select into the programme and so may underrepresent the capital-rich and low-ability types. That is, the important heterogeneity may lie outside the sample. Fifth, the study says little about whether efficiency should take precedence over pro-poor targeting in the equity-efficiency debate; it assumes a priori that efficiency is a desirable policy goal and works with that assumption. 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. This is what the paper hopes to deliver.

References

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, Cambridge, MA.

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

Berg, E., 2013. Are poor people credit-constrained or myopic? Evidence from a South African panel. Journal of Development Economics 101, 195–205.

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.

Chernozhukov, V., Demirer, M., Duflo, E., Fernández-Val, I., 2018. Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India (No. w24678). National Bureau of Economic Research.

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

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.

Macours, K., Premand, P., Vakis, R., 2013. Demand versus returns ? Pro-poor targeting of business grants and vocational skills training (No. 6389). The World Bank.

Manski, C.F., 1990. Nonparametric Bounds on Treatment Effects. The American Economic Review 80, 319–323.

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.

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.

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