Replication of and Extensions to 'The Miracle of Microfinance? Evidence from a Randomized Evaluation.'

James (Zuo Min) Goh

Overview

Introduction

My project replicates some of the key findings from "The Miracle of Microfinance? Evidence from a Randomized Evaluation." by Banerjee et al. (2015) and performs two extensions using randomization inference and causal random forests.

Microfinance is often hailed as a promising means for the poor to escape the poverty trap, in particular through financing businesses and stimulating entrepreneurship among the so-called unbanked. The 2006 Nobel Prize for Peace was given to Mohammed Yunus and the Grameen Bank for their successful introduction of microfinance into the developing world. Despite the enthusiasm for microcredit, there has been a dearth of evidence, especially rigorous, causal analyses, on the impact of microcredit on the poor. In this context, this study was seminal in that it was the first randomized evaluation of microcredit.

In 2005, 52 of 104 poor neighborhoods in Hyderabad, India, were randomly selected for the opening of a branch by a microcredit institution (MFI) named Spandana, while the remainder were not. After about 1.5 years, a household survey was conducted with an average of 65 households in each neighborhood, for a total of 6,863 households. The survey included outcomes on borrowing, consumption, business income, education, health, and more.

My project will focus only on an intention-to-treat analysis of assets, investments, expenses, revenue, and profits (due to space constraints). The key methods for this project are weighted regressions using clustered standard errors, weighted quantile regressions using cluster-bootstrapped standard errors, matched pairs randomization, randomization inference, and causal forests.

Data

The data is obtained from Banerjee et al. (2015) through the *American Economic Journal*. The unit of observation is the household and the key variables included in the data are as follows. Further explanations of the variables are included with the analyses, where necessary.

- Whether a household was treated (Yes/No)
- Whether a household borrowed from (Yes/No)
  - Spandana
  - Other MFI
  - Any MFI
  - Any other bank
  - Informal lenders like moneylenders, friends/ family, or on credit from sellers
  - Any entity
- Business outcomes (in 2007 Rs)
  - Assets (stock)
  - Investments in last 12 months
  - Expenses (per month)
  - Revenue (per month)

    - Profits (per month)

- Whether a household (Yes/No)
    - Had old businesses
    - Started new businesses

- Area-level baseline control variables
    - Population
    - Total debt (per month, in 2007 Rs)
    - Total businesses
    - Average per capita expenditure (per month, in 2007 Rs)
    - Fraction of household heads who are literate
    - Fraction of all adults who are literate

Replication

Regression Model

To estimate the impact of introducing microfinance, a regression of the following form is conducted

$$y_{i\alpha} = \alpha + \beta \times \text{Treat}_{i\alpha} + X'_{\alpha}\gamma + \epsilon_{i\alpha}$$

where $y_{i\alpha}$ is an outcome for household i in area a, $\text{Treat}_{i\alpha}$ is an indicator for living in a treated area, and β is the intent-to-treat effect. X'_{α} is a vector of control variables, calculated as the aforementioned area-level baseline values.

Treatment was randomized at the neighborhood area level to address spillover and general equilibrium effects. Hence, standard errors for clustering at the area level are used. Briefly, clustered standard errors correct for serial correlation, given observations from the same neighborhood are likely to be similar to one another. Instead of assuming that all observations are randomly sampled, clustered standard errors assume that clusters are randomly sampled, without assuming that the observations within the clusters are randomly sampled. This creates larger but more accurate standard errors.

Spandana borrowers were over-sampled to increase power, as the authors believed that the heterogeneity in treatment effects would have introduced more variance in outcomes among Spandana borrowers than among non-borrowers. Therefore, observations are weighted to correct for this oversampling.

Borrowing from Spandana and other MFIs

(The summary statistics are not shown due to space constraints.)

The treatment effects of the percentage of households who borrowed from Spandana and from any MFI are 12.7 points (Table 1, Column 1) and 8.4 points (Table 1, Column 3) respectively. In other words, households that lived in a treated area were 8.4 percentage points more likely to have borrowed from any MFI.

The difference in treatment effects between Spandana and any MFI implies that some households who borrowed from Spandana in treatment areas would have borrowed from another MFI, in the absence of the intervention. Nonetheless, given only 18.3 percent of the control group borrowed from any MFI, the 8.4 percentage point increase implies a 46 percent increase in take up and is economically significant.

Table 1: Credit

	Spandana	Other MFI	Any MFI	Other Bank	Informal	Any Entity
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.127^{***}	-0.012	0.083^{***}	0.003	-0.052^{**}	-0.022
	(0.020)	(0.024)	(0.027)	(0.012)	(0.021)	(0.014)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,811	6,657	6,811	6,811	6,811	6,862

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Cluster-robust standard errors in parentheses.
Results are weighted to account for oversampling of Spandana borrowers.
Columns 1-6: probability of having at least one loan from source listed.

Business Outcomes

Given the substantial increase in the take up of loans from any MFI in the treatment group, one would expect an improvement in the business outcomes for these households. However, there were no statistically significant treatment effects in most business outcomes, namely assets, investments, expenses, revenue, and profits, both among households with old businesses (Table 2), and households with new businesses (Table 3). (Analyses for all households show similar results but are not shown due to space constraints.) Households with old businesses refer to households who already operated businesses before the microcredit intervention, while those with new businesses refer to households who opened new businesses after the microcredit intervention. It is, therefore, broadly untrue that microcredit helps to expand business opportunities among the poor.

Table 2: Business Outcomes (Households with old businesses)

	Assets	Investments	Expenses	Revenue	Profits
	(1)	(2)	(3)	(4)	(5)
Treatment	898	1,119	5,266	1,640	2,105*
	(1,063)	(698)	(3,721)	(3,257)	(1,100)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	2,083	2,083	1,955	2,020	1,624

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Cluster-robust standard errors in parentheses.
Results are weighted to account for oversampling of Spandana borrowers.
Columns 1-5: amounts in 2007 Rs.

Table 3: Business Outcomes (Households with new businesses)

	Assets	Investments	Expenses	Revenue	Profits
	(1)	(2)	(3)	(4)	(5)
Treatment	−873	−706	−8,167	−5,013	−3,548
	(2,201)	(1,324)	(7,314)	(4,049)	(3,813)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	356	356	332	339	270

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Cluster-robust standard errors in parentheses.
Results are weighted to account for oversampling of Spandana borrowers.
Columns 1-5: amounts in 2007 Rs.

It is, however, still worth noting that households with old businesses achieved a marginally significant increase in profits of Rs. 2,105 per month (Table 2, Column 5) and those with new businesses achieved an insignificant decrease in profits of Rs. 3,548 per month (Table 3, Column 5). These estimates are economically significant, given the average consumption of households in the treatment group is only Rs. 1,490 per month. Performing a quantile regression on profits, the increase in profits was concentrated in quantiles 95 and above for households with old businesses, while the decrease in profits was concentrated in quantiles 90 and above for households with new businesses (Figure 1). Notably, the decrease in profits for households with new businesses is statistically significant between quantiles 35 and 65.

Collectively, these findings suggest that microcredit does not help the vast majority of businesses, except perhaps the most-profitable businesses that had already existed. On hindsight, this is unsurprising. Less-profitable businesses and marginal businesses that get created due to microcredit may not have the ability or opportunity to channel credit into productive assets that generate high returns. Moreover, microcredit lowers the profitability threshold to start a new business, such that the marginal businesses created are inherently less profitable and may in fact encourage less savvy households to participate in businesses.

Overall, this study, along with many subsequent randomized evaluations, robustly refutes the exaggerated claim that microcredit supports poor households by broadening their business opportunities.

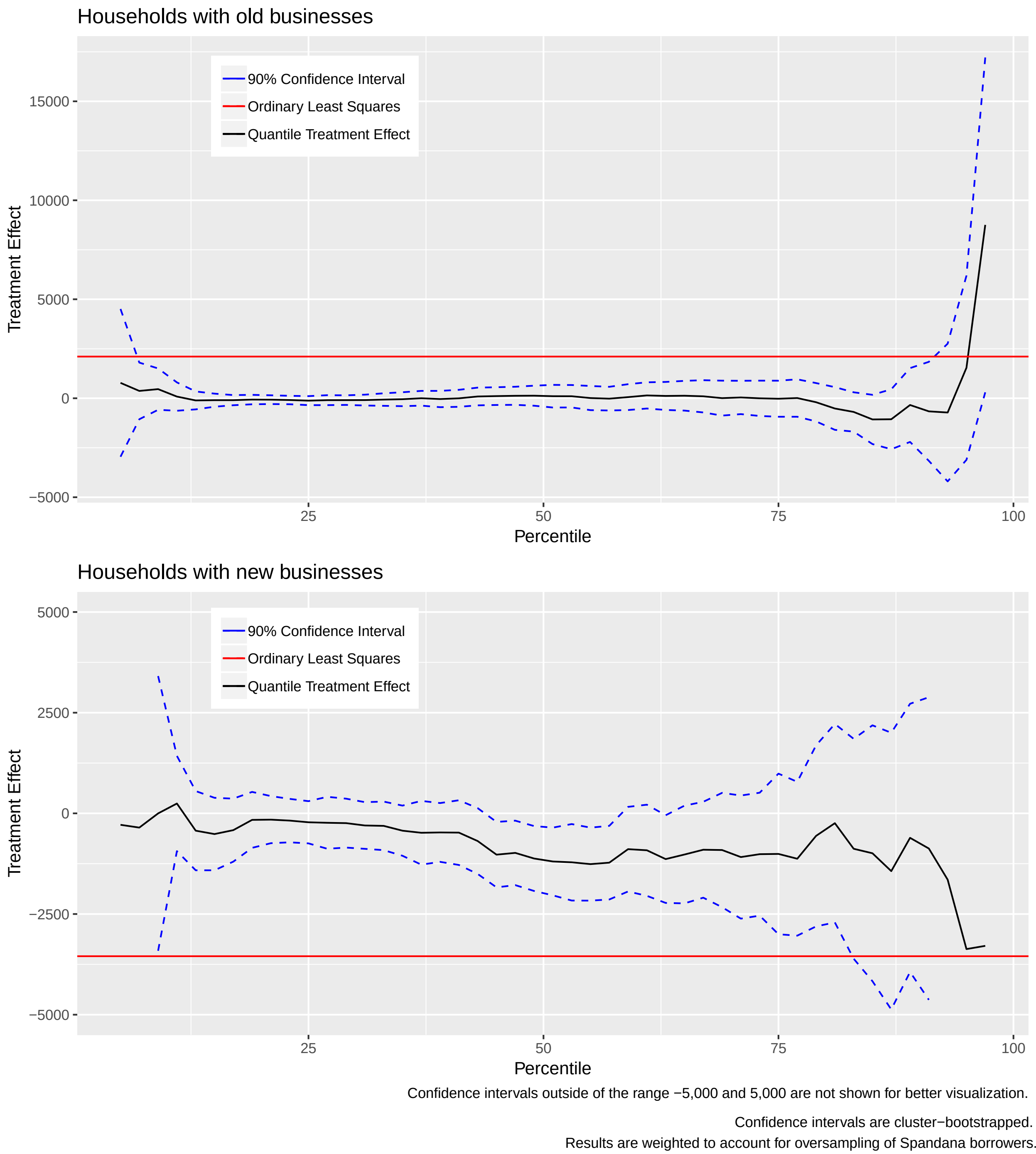

Figure 1: Quantile Treatment Effect on Profits

Extensions

Randomization Inference

The analyses performed thus far uses regression analyses to estimate treatment effects, which relies on asymptotic properties. An alternative method is randomization inference, which makes no such assumptions. While the sample size in this study is relatively large, randomization inference may be insightful given the treatment effects found have been very skewed. This is illustrated by Young (2018) who re-analyzed 2003 regressions in 53 experimental papers drawn from the *American Economic Association* using randomization inference. His tests reduced the number of regression specifications with statistically significant treatment effects by 30 to 40 percent. Given the lack of robustness of regression analyses and the fact that Young did not extend his analysis to this study, I am interested in performing randomization inference for this project.

Briefly, randomization inference tests the sharp null hypothesis $H_0 : Y_i(0) = Y_i(1)$ for all i. Using the test statistic of the absolute difference in means between treatment and control, under the null hypothesis, this test statistic should remain the same even if the treatment and control assignments are re-assigned. By calculating the test statistic under every possible treatment assignment scenario, a p-value can be estimated by calculating the proportion of hypothetical test statistics that are larger than the observed test statistic.

The authors randomized through pair-matching and formed pairs that minimized the sum across pairs A, B (area A average loan balance $-$ area B average loan balance)^2 + (area A per capita consumption $-$ area B per capita consumption)^2. Within each pair, one neighborhood was randomly allocated to treatment. Unfortunately, the authors did not publish the assignment of pairs to the observations. Using the optimal full matching algorithm, which employs an efficient search method for the global minimum, from an R package developed by Hansen and Klopfer (2006), I infer the pair assignments from the existing treatment and control assignments and re-randomize them 1,000 times.

The results obtained are qualitatively identical (Table 4, Figure 2). However, what is extremely surprising is the bi-modal distribution of treatment effects on expenses, revenue, and profits among households with new businesses (Figure 3). This is an unfortunate realization of the hypothetical, extreme scenario Deaton and Cartwright (2018) raise. Depending on whether the outliers are assigned to the treatment or control group during the re-randomization process, the average test statistic estimated for that group will be particularly extreme. As a result, large outliers can produce bi-modal distributions, which in turn leads to over-rejections or spurious significance.

	Assets	Investments	Expenses	Revenue	Profits
Old	0.48	0.11	0.23	0.66	0.07
New	0.72	0.63	0.27	0.24	0.44

Table 4: Randomization Inference p-values

The residuals against fitted values plots show an obvious outlier (not shown due to space constraints), namely a household in the control group that reported mean expenses of Rs. 634,500, revenue of Rs. 1,100,000, and profits of Rs. 465,500. In comparison, the mean expenses, revenue, and profits are Rs. 4150, Rs. 5181, and Rs. 968 respectively. As the outlier is in the control group, outcomes are under-dispersed, leading to large, negative, and potentially significant treatment effects. Hearteningly, removing the outlier and re-running the regression analysis for households with new businesses, the results remain qualitatively the same, except that the treatment effect on profits becomes positive, although still statistically insignificant (not shown due to space constraints).

Overall, the results are robust to different statistical methods.

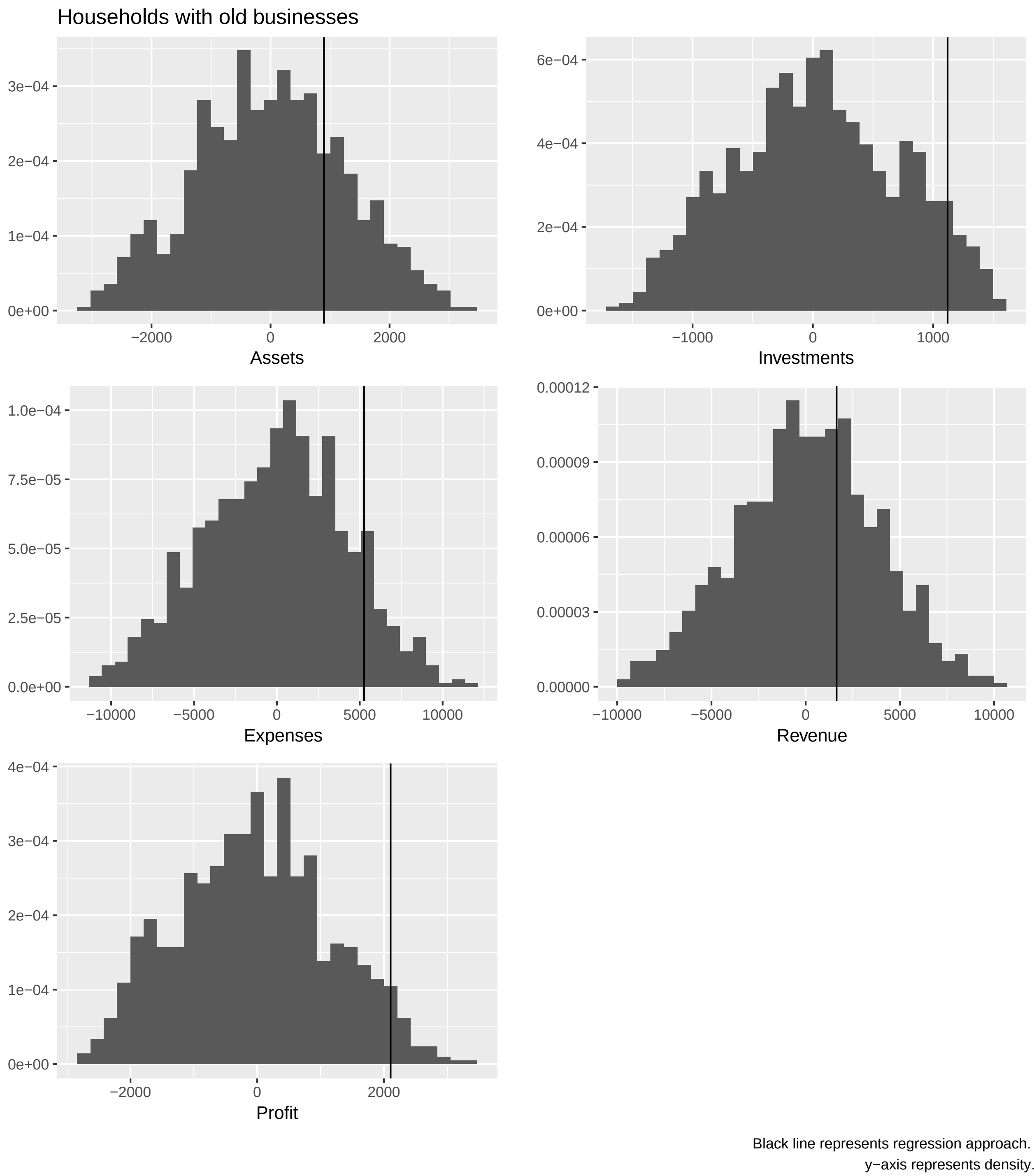

Figure 2: Randomization Inference Treatment Effects (Households with old businesses)

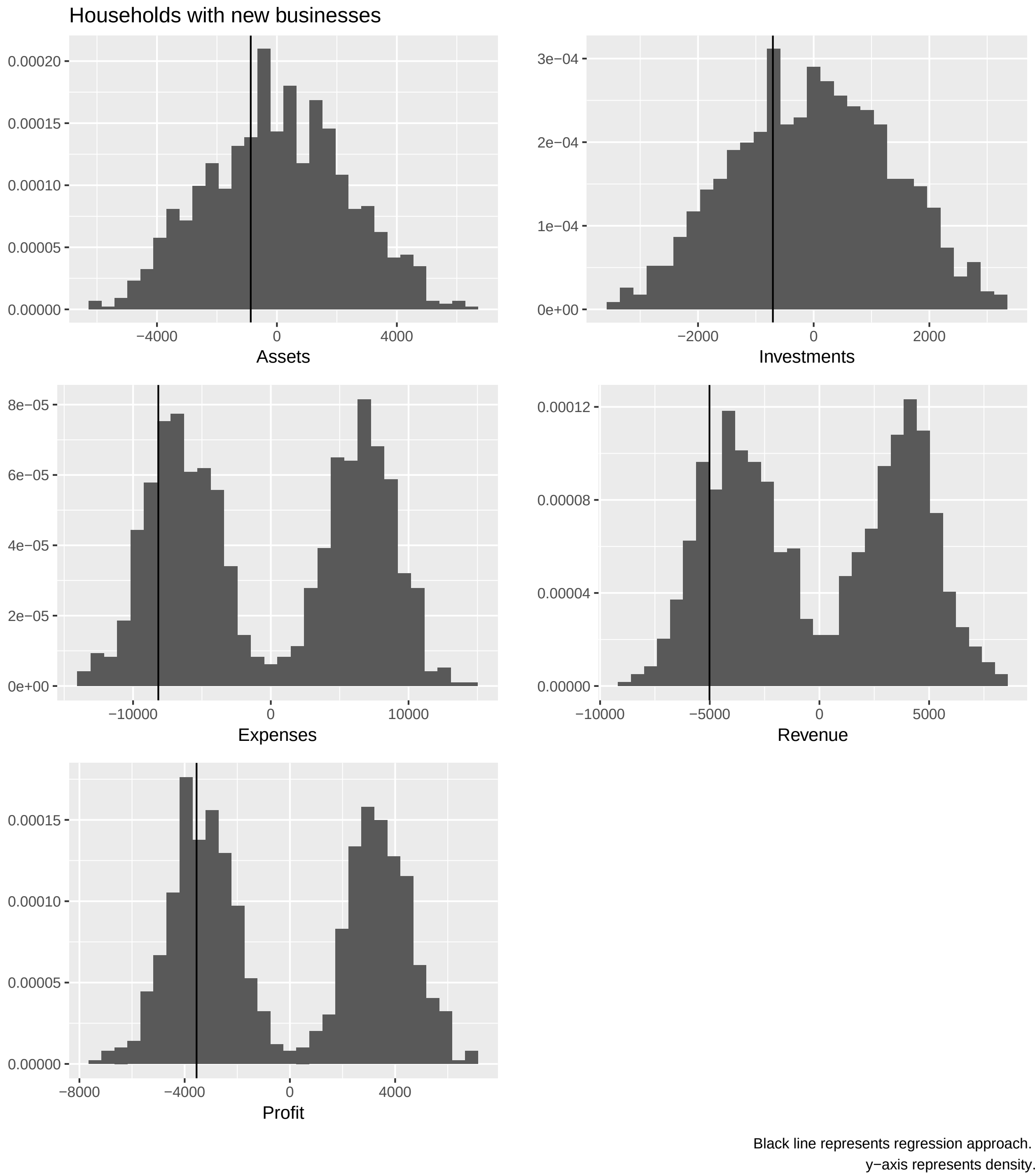

Figure 3: Randomization Inference Treatment Effects (Households with new businesses)

Honest Causal Forests

The results suggest some heterogeneity in treatment effects between households with old businesses and those with new businesses. It is, therefore, possible that neighborhoods with more conducive business environments stand more to gain from microcredit. For example, neighborhoods that have high literacy rates or a large population may be more ripe for business expansion. Such insights will be particularly useful in developing a targeted strategy for introducing microcredit.

While traditional regression analyses are often used to identify heterogeneous treatment effects, they present some limitations, including low statistical power, over-fitting, and restrictions to linear relationships. To tackle this, Wager and Athey (2018) introduced honest causal forests, which is an extension to random forests.

Briefly, treatment effects for a particular subgroup are estimated using the difference in means among the treatment and control groups within the terminal node that corresponds to that subgroup. The trees use a unique splitting criterion that balances between finding the largest difference in treatment effects between subgroups and estimating the treatment effects accurately. To prevent over-fitting, honest trees are built, which refer to splitting the training data into two sub-samples, namely the splitting and estimating sub-sample. The splitting sub-sample is first used to estimate the parameters of the forest. The estimating sub-sample is then applied to this causal forest. The results are used to predict new observations. Incredibly, this honest approach results in asymptotically normal distributions of treatment effect estimates, so confidence intervals and variances can be calculated from this approach.

Using the R package developed by Tibshirani et al. (2018) and with guidance from an online tutorial by White (2018), I apply this method and estimate the heterogeneous treatment effects of profits for all households. The covariates used are based on the aforementioned area control variables as well as two indicator variables for households with old businesses and for those with new businesses. Unfortunately, the package is still in beta and the developers have not yet incorporated sample weights, although they have announced their intention to do so. In any case, the sample weights used thus far do not qualitatively change any of the aforementioned analysis and so the exclusion of sample weights here should not impact the findings substantially. The results here can be interpreted as an over-weighting on Spandana borrowers. For reference, the smallest sample weight is 0.51 and the standard deviation of the weights is 0.08.

The results suggest that the top four variables in order of importance are area population, total debt, mean per capita expenditure, and fraction of all adults who are literate (Table 5). Specifically, an increase in the predicted treatment effect of profits is associated with smaller populations and higher literacy rates (Figure 4). The relationships relating to total debt and mean per capita expenditure are less clear and interpretable.

	Variables	Importance
1	Population	0.18
2	Total Debt	0.17
3	Mean Per Capita Expenditure	0.16
4	Fraction of all Adults who are Literate	0.14
5	Total Businesses	0.13
6	Households with Old Businesses	0.12
7	Fraction of all Household Heads who are Literate	0.09
8	Households with New Businesses	0.00
9	Intercept	0.00

Table 5: Causal Forest Variable Importance

The negative relationship with populations exhibits the clearest and biggest relationship and is worth examining in detail. It also runs counter to my initial intuition that large populations offer more business opportunities and hence a greater likelihood of higher profitability. One potential explanation may be that neighborhoods with smaller populations are more credit constrained because there are smaller informal networks of other households to rely on for credit. As such, microcredit interventions help to tackle these issues, capturing the low hanging fruits in these neighborhoods. Although this is entirely speculative, the

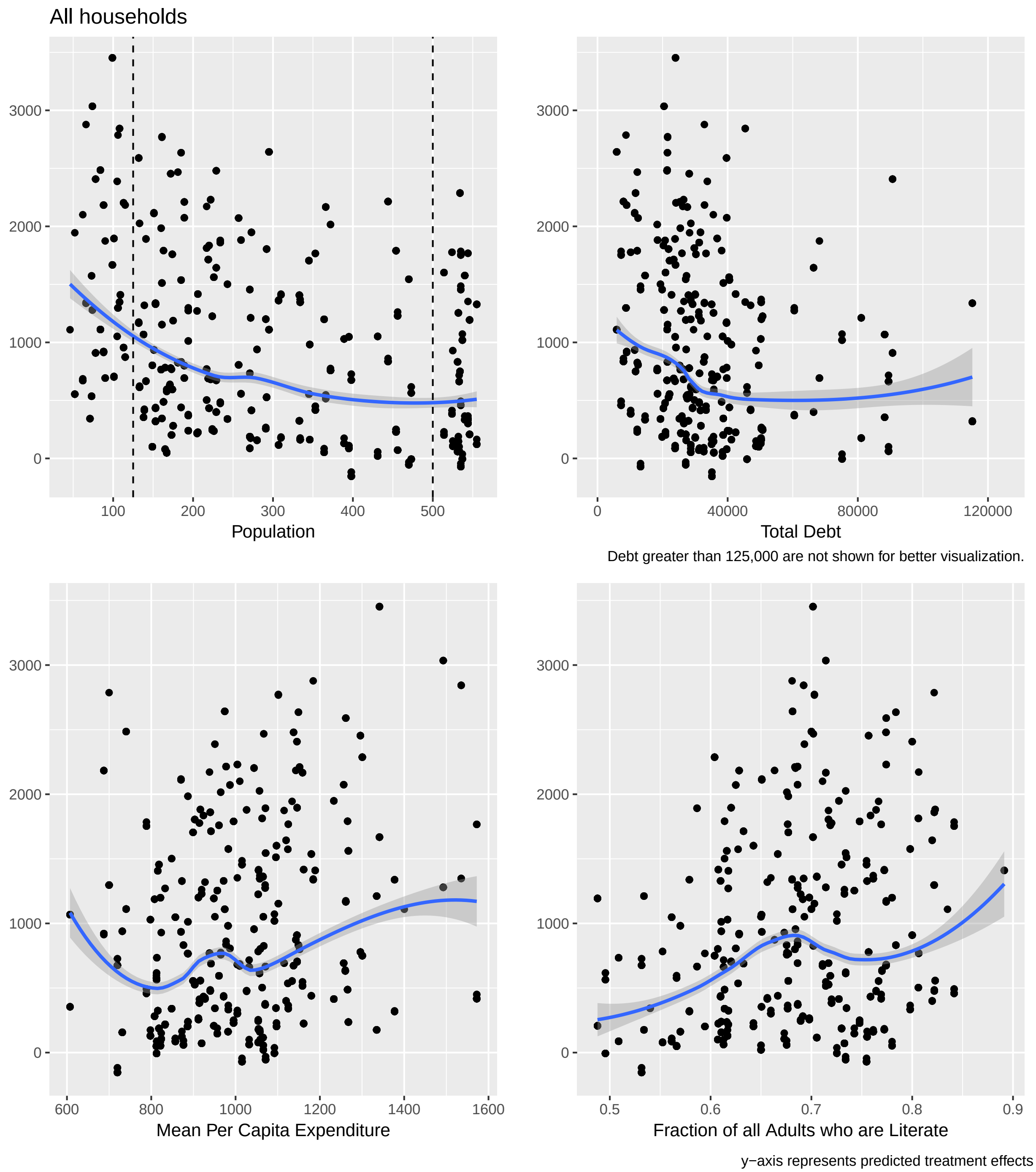

Figure 4: Causal Forests, Relationship between Treatment Effect and Covariates

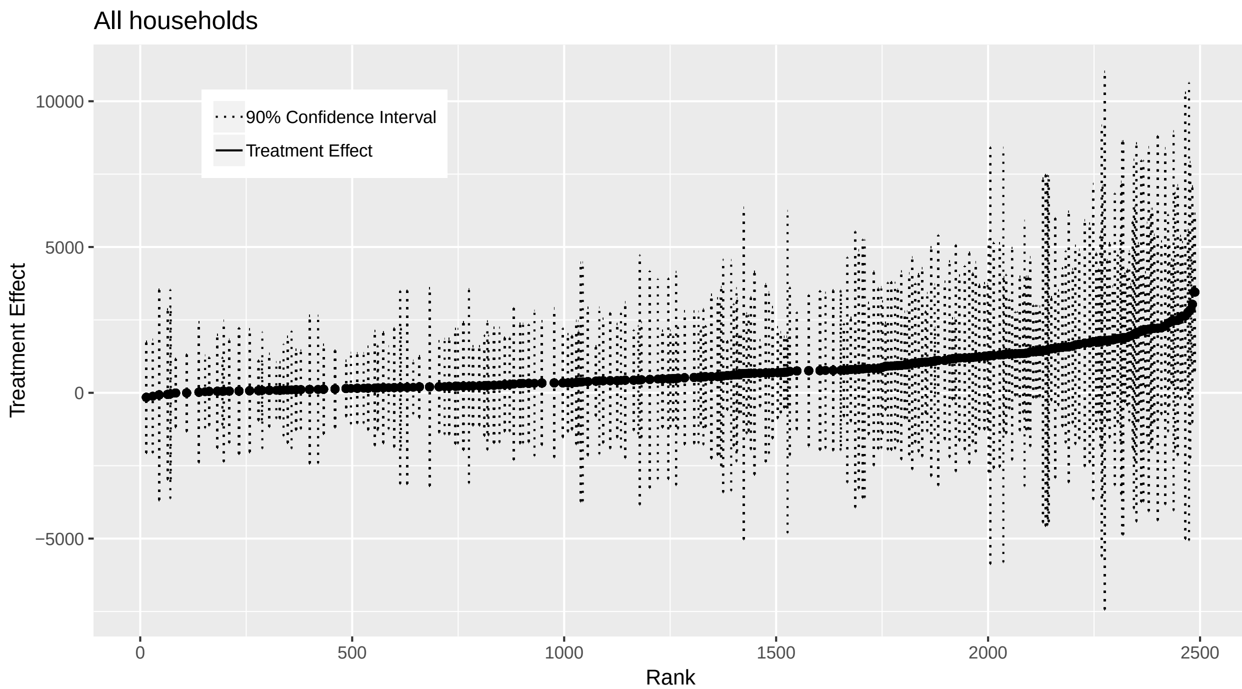

Figure 5: Causal Forests, Heterogeneity of Treatment Effects

results are worth noting. The local regression in Figure 4 suggests that the predicted treatment effect for populations less than 125 is Rs. 1000 or greater, while that for populations greater than 500 is about Rs. 500. Given the substantial difference, neighborhoods with small populations ought to be prioritized.

While these findings are insightful, it is also crucial to put the heterogeneity in context. Arranging the predicted treatment effects by rank order, at first glance, there seems to be large heterogeneity in treatment effects (Figure 5). The largest treatment effect is an increase of about Rs. 3,500 and the smallest treatment effect is a decrease of about Rs. 500. However, attaching confidence intervals to the predictions, it is clear that in reality, the effects are fairly homogeneous in nature, given the existing set of covariates used. The confidence interval for the largest treatment effect overlaps substantially with that for the smallest treatment effect. Consequently, any targeting strategy is only likely to change treatment effects to a small extent relative to the inherent randomness of these effects.

Conclusion

In conclusion, while regressions clearly remain the workhorse of randomized evaluations, as showcased through this project, other statistical methods like randomization inference and causal forests are promising complementary techniques that can offer additional insights. Relatedly, bringing modern statistical techniques to older but rich data sets from randomized evaluations can also yield unexpected and interesting findings.

References

Auguie, Baptiste. 2017. *GridExtra: Miscellaneous Functions for "Grid" Graphics*. https://CRAN.R-project.org/package=gridExtra.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2015. "The miracle of microfinance? Evidence from a randomized evaluation." *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20130533.

Coppock, Alexander. 2018. *Randomizr: Easy-to-Use Tools for Common Forms of Random Assignment and Sampling*. https://CRAN.R-project.org/package=randomizr.

Deaton, Angus, and Nancy Cartwright. 2018. "Understanding and misunderstanding randomized controlled trials." *Social Science and Medicine*. https://doi.org/10.1016/j.socscimed.2017.12.005.

Graham, Nathaniel, Mahmood Arai, and Björn Hagströmer. 2016. *Multiwayvcov: Multi-Way Standard Error Clustering*. https://CRAN.R-project.org/package=multiwayvcov.

Hansen, Ben B., and Stephanie Olsen Klopfer. 2006. "Optimal Full Matching and Related Designs via Network Flows." *Journal of Computational and Graphical Statistics* 15 (3): 609–27.

Hlavac, Marek. 2018. *Stargazer: Well-Formatted Regression and Summary Statistics Tables*. Bratislava, Slovakia: Central European Labour Studies Institute (CELSI). https://CRAN.R-project.org/package=stargazer.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2011. "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference." *Journal of Statistical Software* 42 (8): 1–28. http://www.jstatsoft.org/v42/i08/.

Koenker, Roger. 2018. *Quantreg: Quantile Regression*. https://CRAN.R-project.org/package=quantreg.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Tibshirani, Julie, Susan Athey, Stefan Wager, Rina Friedberg, Luke Miner, and Marvin Wright. 2018. *Grf: Generalized Random Forests (Beta)*. https://CRAN.R-project.org/package=grf.

Wager, Stefan, and Susan Athey. 2018. "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests." *Journal of the American Statistical Association*. https://doi.org/10.1080/01621459.2017.1319839.

White, Mark. 2018. "Explicitly optimizing on causal effects via the causal random forest." https://www.markhw.com/blog/causalforestintro.

Wickham, Hadley, and Evan Miller. 2018. *Haven: Import and Export 'Spss', 'Stata' and 'Sas' Files*. https://CRAN.R-project.org/package=haven.

Young, Alwyn. 2018. "Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *The Quarterly Journal of Economics*. https://doi.org/10.1093/qje/qjy029.

Zeileis, Achim, and Torsten Hothorn. 2002. "Diagnostic Checking in Regression Relationships." *R News* 2 (3): 7–10. https://CRAN.R-project.org/doc/Rnews/.