Replication of The Miracle of Microfinance? Evidence from a Randomized Evaluation by Banerjee et al. (2025)

Satyajith Bavisetty

May 4, 2025

Abstract

Banerjee et al. (2015) evaluated the impact of microfinance on household and business outcomes using a randomized controlled trial conducted in Hyderabad, India. The study examined effects of microcredit on consumption, business activity, income, and social indicators such as health and education. The key-independent variable was a binary treatment assignment at the neighborhood level, and follow-up data was collected at two endlines. The authors found limited effects of microfinance on most outcomes, with small increases in durable goods spending and business investment but no significant improvements in consumption or welfare indicators. In our replication, we utizile the original datasets and methods in python to visualize the quantile treatment effects on informal borrowing and business profits. Similarly the original dataset was used to visualize the regression analysis of the treatment variable at endpoints 1 and 2. The data replicated had different values due to discrepancies in data for some of the variables but produced similar results to the original paper.

1 Introduction

In recent decades, microfinance has emerged as a widely promoted development tool, designed to extend small-scale credit to the world's poor, particularly to women who are often excluded from formal banking systems. The concept has garnered widespread enthusiasm, culminating in the awarding of the 2006 Nobel Peace Prize to Muhammad Yunus and the Grameen Bank. Advocates of microfinance suggest that access to small loans can stimulate entrepreneurship, increase household consumption, and improve social outcomes. However, critics have questioned both the effectiveness and ethics of microfinance, pointing to issues such as over-indebtedness, minimal business impact, and potential exploitation.

Against this backdrop, the 2015 paper "The Miracle of Microfinance? Evidence from a Randomized Evaluation" by Banerjee, Duflo, Glennerster, and Kinnan offers one of the first large-scale randomized evaluations of a group-based microcredit program. Conducted in Hyderabad, India, the study assesses the effects of introducing microfinance in 52 randomly selected neighborhoods, compared to 52 control neighborhoods. The paper evaluates a range of outcomes, including borrowing behavior, business activity, consumption patterns, and women's empowerment, with follow-up surveys conducted approximately 15–18 months and then three years after the intervention.

This reproducibility report aims to assess the transparency and robustness of the original study by attempting to replicate its main findings using the publicly available dataset provided by the authors. In doing so, the report not only evaluates the replicability of the key statistical results, but also reflects on the broader implications of reproducibility in development economics, particularly when policy recommendations rest on experimental evidence.

2 Reproducibility

There were no errors found in the original STATA code or datasets for the reproducibility files from the authors. However, original methods for reproducing code, filtering data, and running regression were not entirely disclosed for all figures. As such, values differed for some regression tables, but results derived from values remained the same. Additionally, there was no difference in the replicated graphs from the paper.

3 Replication

3.1 Figure 1: Summary Table

Did not replicate as the graph showcased the basic timeline of the experimental design and was not a part of the results section.

3.2 Figure 2: Treatment Effect on Informal Borrowing

This graph shows the treatment effect of microfinance on informal borrowing of money (Endline 1). The x-axis showcases the percentiles of the distribution of informal borrowing amounts, such as from friends, family, and moneylenders. The y-axis shows the quantile treatment effect of access to microfinance on informal borrowing at each percentile. The graph showcases the OLS Estimate (dashed line), QTE curve (solid line), and 90 percent confidence interval (dot-dashed line). The OLS equation used for this graph is:

$$InformalBorrowing_{ih} = \alpha + \beta \cdot \text{Treatment}_{ih} + \epsilon_{ih} \tag{1}$$

3.2.1 Challenges We were able to reproduce this graph identically with no challenges.

3.2.2 Results From the graph, it can be seen that QTE is negative across the middle section of the distribution (25th - 80th percentile). From this, we can in-

fer that households that borrowed informally borrowed less once microfinance was introduced. We can infer that microfinance substituted for informal borrowing. However, this effect was not seen at the lowest (5th - 20th percentile) and highest percentiles (90th - 95th percentile). This suggests that microfinance did not affect borrowing informally for minimal amounts or high-borrowing households.

3.3 Figure 3: Treatment Effect on Business Profits

This graph shows the treatment effect of microfinance on business profits (Endline 1) for households that had old businesses before the treatment. The x-axis showcases the percentiles of the distribution of business profit amounts. The y-axis shows the quantile treatment effect of access to microfinance on business profits at each percentile. The graph showcases the OLS estimate (dashed line), QTE curve (solid line), and 90 percent confidence interval (dot-dashed line). The OLS equation used for this graph is:

$$BusinessProfits_{ih} = \alpha + \beta \cdot \text{Treatment}_{ih} + \epsilon_{ih} \tag{2}$$

3.3.1 Challenges We were able to reproduce this graph identically with no challenges.

3.3.2 Results From the graph, it can be observed that the QTE curve is around zero for the majority of the percentiles. Additionally, the 90 percent CI around the QTE includes zero across the distribution. This means that there is no significant treatment effect across the distribution. The OLS line also lies around zero, meaning the OLS estimate is small and statistically insignificant. As such, we cannot determine a positive or negative effect of microfinance on HHs' old businesses at endline 1.

3.4 Figure 4: Treatment Effect on Business Profits

This graph shows the treatment effect of microfinance on business profits (Endline 1) for households that created new businesses after receiving treatment. The x-axis presents the percentiles of the distribution of business profit amounts, while the y-axis shows the quantile treatment effect of access to microfinance at each percentile. The graph displays the OLS estimate (dashed line), the quantile treatment effect (QTE) curve (solid gray line), and the 90 percent confidence interval (dot-dashed lines). The OLS equation used for this graph is:

$$BusinessProfits_{ih}^{new} = \alpha + \beta \cdot \text{Treatment}ih + \epsilon ih \tag{3}$$

3.4.1 Challenges We were able to replicate this figure without challenges, using the provided quantile regression and plotting functions.

3.4.2 Results The graph shows that for households that started new businesses, the quantile treatment effect (QTE) is consistently negative across most percentiles, especially between the 40th and 90th percentiles, where the QTE curve dips even further below zero. Importantly, the 90 percent confidence intervals for these mid-to-upper percentiles do not include zero, indicating a statistically significant negative treatment effect on profits for many new businesses. This suggests that, conditional on having started a business, the marginal new businesses created due to access to microfinance tend to earn lower profits compared to control households.

3.5 Figure 5: Treatment Effect on Business Profits

Figure 5 reports the quantile treatment effects (QTE) of microfinance access on business profits for any business at Endline 2. By this point, both treatment and control households had similar access to microfinance loans – though borrowers in treatment areas generally held larger loans and had been borrowing longer. The regression estimates the treatment effects across the distribution of business profits,

using quantile regression to capture heterogeneous impacts that might be masked by OLS. The blue line represents the OLS estimate, the gray line shows the QTE across percentiles, and the dashed blue lines mark the 90% confidence intervals. The OLS equation used for this graph is:

$$BusinessProfits_ih = \alpha + \beta \cdot \text{Treatment}_ih + \epsilon_ih \tag{4}$$

3.5.1 Challenges We were able to replicate this figure without challenges, using the provided quantile regression and plotting functions.

3.5.2 Results Across most percentiles (up to 85th), the quantile treatment effect is near zero, and the 90% confidence intervals comfortably include zero – this indicates no statistically significant impact on profits for the majority of businesses. Only in the upper tail, starting around the 85th percentile and rising sharply beyond the 90th percentile, does the QTE become positive and statistically different from zero. This suggests that after several years of borrowing, the most profitable businesses in treatment areas gained disproportionately, while average and lower-tier businesses saw no meaningful profit growth compared to controls. The OLS point estimate is positive (roughly Rs. 6,000, about 60% of the sample mean), but with a t-statistic around 1.5, it lacks statistical significance.

These results reinforce the paper's conclusion that microfinance has selective benefits. While it allows the most successful existing businesses to expand, it does not raise profits or employment across the average or marginal firm.

3.6 Table 1A:

Table 1A presents baseline summary statistics comparing households in the control group to those assigned to receive microfinance (Spandana) in the study by Banerjee et al. (2015). It assesses whether the two groups were balanced prior to the intervention by showing means, standard deviations, and treatment–control differences for key variables. The high p-values across most characteristics suggest

successful randomization, with no significant pre-existing differences. At baseline, informal loans were the dominant source of credit, and no households had loans from Spandana, confirming it had not yet entered. Households showed modest levels of business activity and consumption, with wide variation in outcomes. This table is crucial for validating that any post-treatment effects observed can be credibly attributed to the microfinance intervention rather than underlying differences between groups.

3.7 Table 3a:

3.7.1 Challenges Trying to replicate the data in Table 3 presents several real challenges. First, the study is based on a randomized control trial involving thousands of households, so without access to the original dataset, it is nearly impossible to recreate the exact numbers. Even if we had the data, many of the variables shown, such as the "Index of dependent variables" or whether someone "has started a business in the last 12 months," are not clearly defined in the table. We would need more detail about how these were constructed, which is likely buried in the paper or its appendix. There is also the issue of how the statistics were generated. For example, the table includes Hochberg-corrected p-values, which adjust for multiple comparisons, but without knowing which outcomes were included in that correction, we cannot replicate it accurately. Finally, some values are marked with significance levels, and unless we know the exact statistical models used, including any adjustments for things like clustering or baseline variables, our replication would be incomplete or incorrect.

3.7.2 Results Looking at the table, the results of the intervention appear mixed. In the first follow-up period shown in Panel A, households in the treated area showed slightly higher business investment and were a bit more likely to have started a business. However, most of the results are not statistically significant once the multiple comparison adjustment is applied. By the time of the second follow-up in Panel B, treated households had significantly more assets than the control group.

But interestingly, they spent less on business investment and had lower expenses. This could mean that although people had more resources, they did not necessarily use them to grow their businesses. It might also suggest that the businesses were not very profitable or that people were being cautious. Overall, while the intervention led to some changes, especially in asset levels, it did not clearly result in stronger or more consistent self-employment outcomes over time.

3.8 Table 3b:

3.8.1 Challenges A significant challenge in replicating Table 3B lies in the likely unavailability of the original microdata, as the table only presents summary statistics, hindering independent verification and reproduction of statistical analyses like the Hochberg correction. Furthermore, the precise definitions of key variables and the methodology for defining the treated and control groups are not detailed within the table itself, necessitating a thorough examination of the report's methodology. Understanding how the original researchers handled missing data and potential outliers, which could impact the reported results, is also crucial. Finally, the construction of the "Index of dependent variables" and the specific statistical tests used to generate the reported p-values are unclear from the table alone, requiring further information from the report.

3.8.2 Results The results presented in Table 3B compare self-employment activities between a treated and a control group at two endline periods. At Endline 1 (Panel A), the treated area showed lower mean assets and revenue but higher mean investment and a slightly elevated mean profit compared to the control group. A negative mean change in employee numbers was also observed in the treated area, along with a positive mean value for the index of dependent variables. The Hochberg-corrected p-value of 0.057 suggests a marginal statistical significance in the overall differences between the groups at this first endline. However, by Endline 2 (Panel B), the treated area exhibited lower means across most indicators, including assets, investment, revenue, expenses, and profit, coupled with a negative

mean change in employees and a negative value for the index of dependent variables. The Hochberg-corrected p-value of >0.999 indicates no statistically significant overall differences between the treated and control groups at the second endline after adjustment for multiple comparisons.

3.9 Table 3c:

3.9.1 Challenges Similar to the challenges encountered with Table 3B, replicating Table 3C presents several methodological obstacles. A primary concern is the probable lack of access to the original microdata, as the table provides only summary statistics (means and standard deviations). This limitation impedes independent verification of the reported values and the reproduction of statistical analyses, including the Hochberg correction. Furthermore, the precise definitions of the variables, such as "Assets (stock)" and "Investment in last 12 months," are not explicitly detailed within the table, necessitating a thorough review of the report's methodological section for clarification. Understanding the specific criteria employed to define the "Treated area" and the selection process for the "Control mean" for this particular subsample ("Households with new businesses; ELI only") is also crucial for accurate replication. The methods utilized by the original researchers to address missing data and potential outliers, which could influence the reported statistics, are not evident from the table itself. Finally, the construction of the "Index of dependent variables" and the specific statistical tests conducted to yield the reported p-value require detailed information from the accompanying text within the report.

3.9.2 Results Table 3C presents a comparison of self-employment activities between a treated area and a control group, focusing specifically on households with new businesses and meeting the "ELI only" criteria. The treated area exhibits substantially lower means across all reported financial indicators compared to the control mean: assets, investment in the last 12 months, revenue, expenses, and profit are all considerably lower. The treated area also shows a negative mean change in the number of employees, which is statistically significant at some level (indicated

by the asterisk). The mean value for the index of dependent variables is also negative for the treated area. The Hochberg-corrected p-value of 0.280 suggests that, after adjusting for multiple comparisons, the overall differences observed between the treated and control groups across the dependent variables are not statistically significant at conventional levels (e.g., 0.05).

4 Conclusion

This project examined and replicated key parts of Banerjee et al. (2015), a landmark study on the effects of microfinance in India. The original research used a randomized control trial to measure how access to microcredit influenced household business activity, income, and consumption. While the study found some small but statistically significant effects such as increased business investment, asset accumulation, and a higher likelihood of starting a business, the overall impact on profits, consumption, and long-term welfare was limited. These findings suggest that microfinance can expand short-term financial choices but is not a transformational tool for reducing poverty on its own.

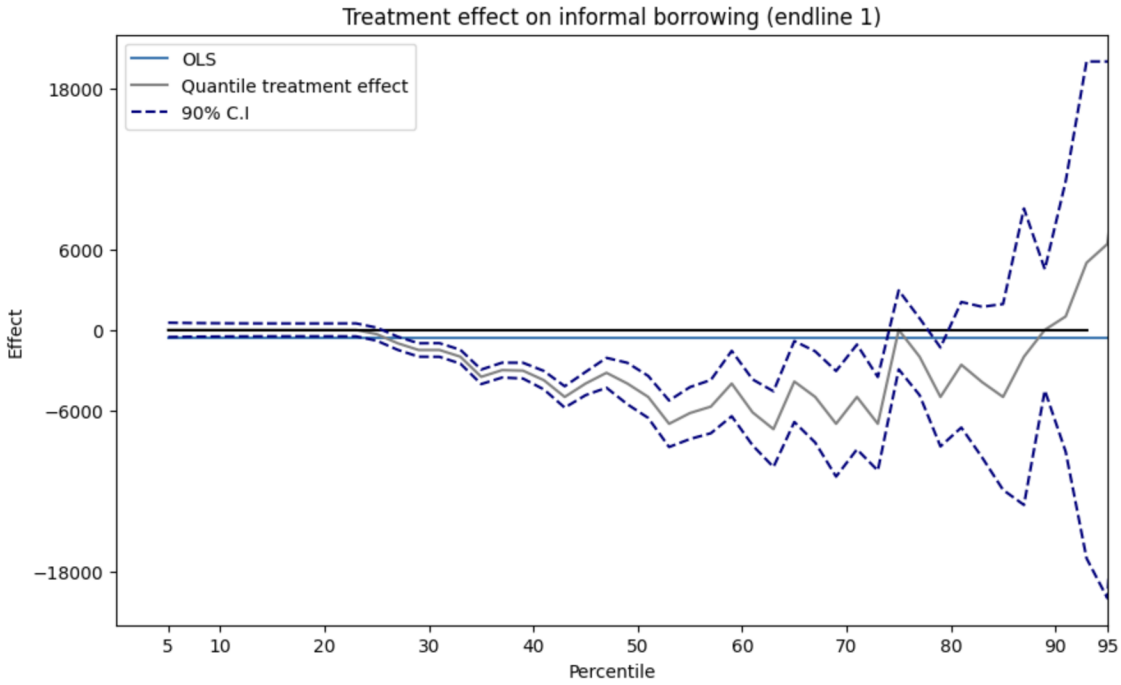
Our replication of the data, particularly in relation to self-employment outcomes, highlighted both the strengths of the original study and the challenges of reproducing complex economic research. Limited data access and unclear variable definitions made it difficult to fully replicate the tables, which emphasized the importance of transparency and open data in research. Overall, our findings support the original conclusion that microfinance may provide some benefits, but its role in improving household livelihoods is modest and should be viewed as one part of a broader development strategy.

References

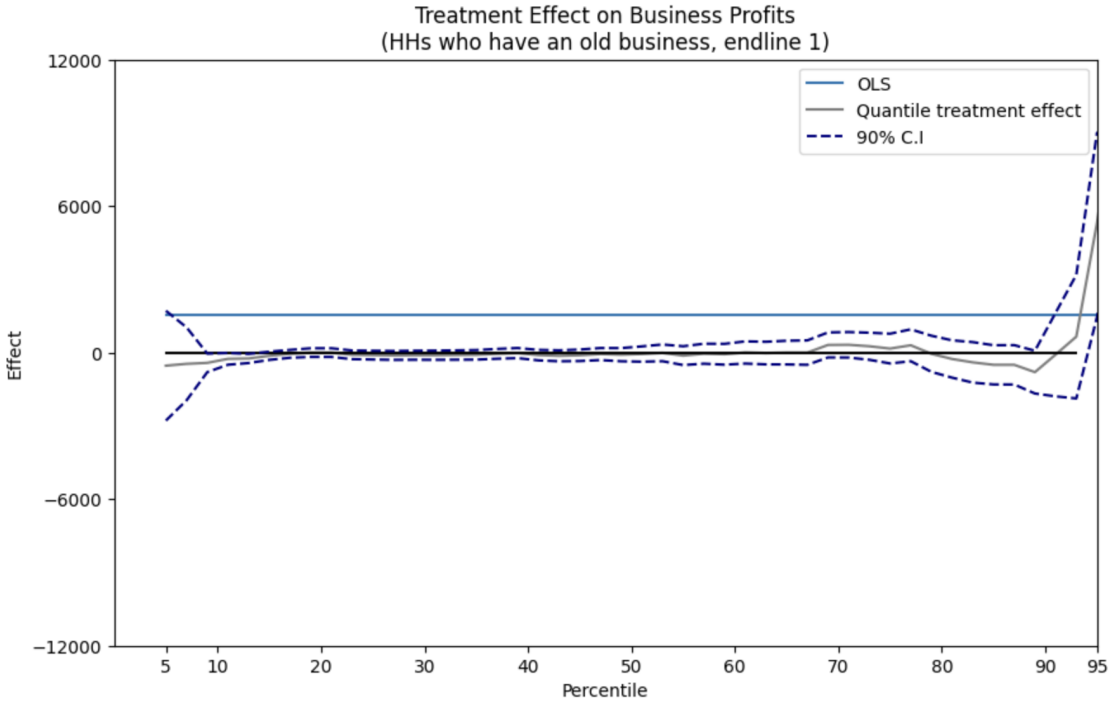
Banerjee, A., Duflo, E., Glennerster, R., Kinnan, C. (2015). The Miracle of Microfinance? Evidence from a Randomized Evaluation. American Economic Journal: Applied Economics, 7(1), 22–53. http://www.jstor.org/stable/43189512

5 Figures

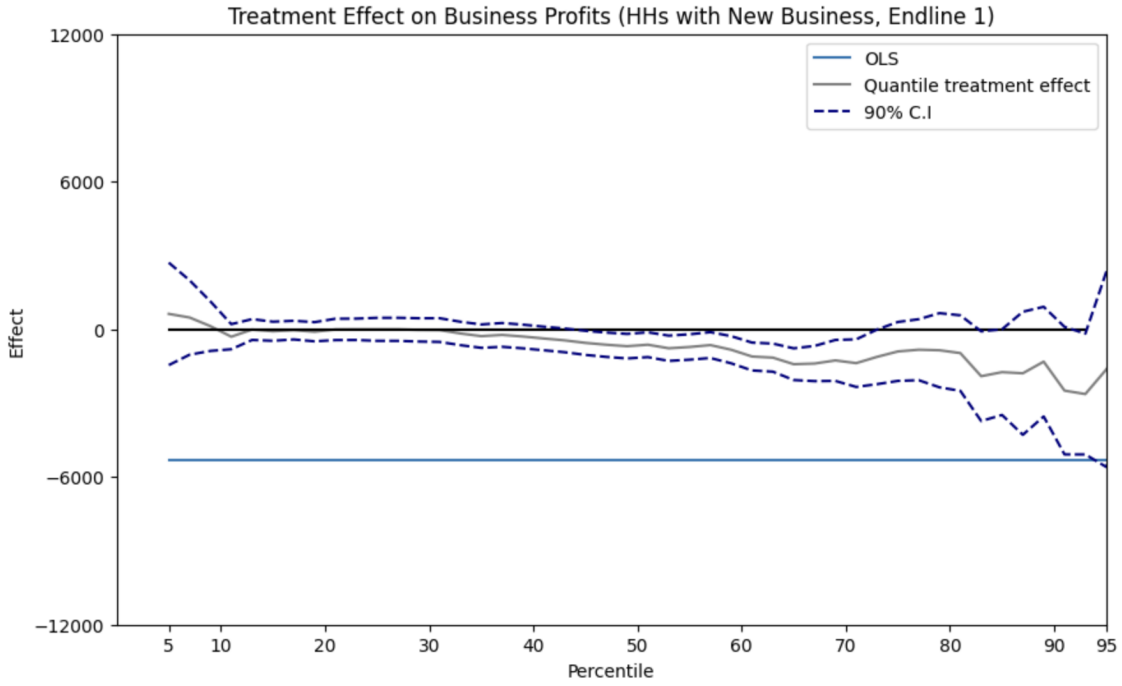
5.1 Figure 2: Treatment Effect on Informal Borrowing (Endline I)



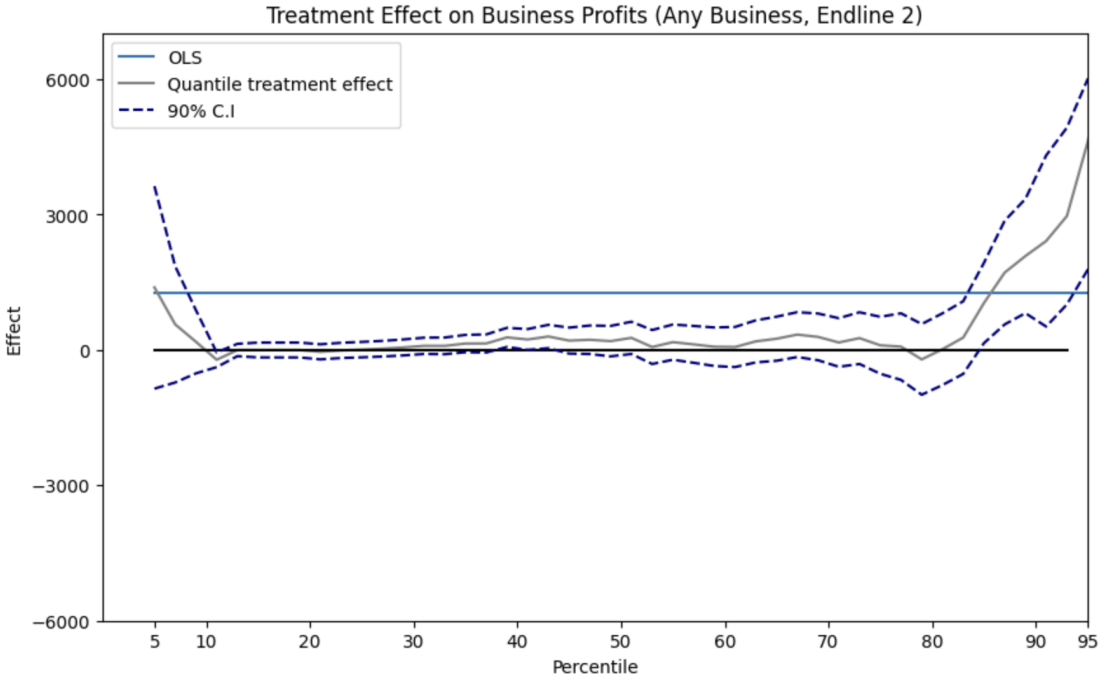
5.2 Figure 3: Treatment Effect on Business Profits (HHs who have an old business, Endline I)



5.3 Figure 4: Treatment Effect on Business Profits (HHs with New Business, Endline I)



5.4 Figure 5: Treatment Effect on Business Profits (Full sample of business owners, endline 2)



6 Tables

Table 1: Table 1A Summary Statistics

Variable	Obs	Control$_{m}ean$	Control$_{s}d$	Difference	p$_{v}al$
hh$_{s}ize$	1220	5.0377	1.6662	0.0951	0.3010
adults	1220	3.4385	1.4656	-0.0115	0.8728
children	1220	1.5992	1.2283	0.1041	0.0952
male$_{h}ead$	1216	0.9071	0.2905	-0.0122	0.3790
head$_{a}ge$	1216	41.1497	10.8394	-0.2435	0.6753
head$_{n}oeduc$	1216	0.3701	0.4830	-0.0077	0.7864
spandana	1213	0.0000	0.0000	0.0066	0.1922
othermfi	1213	0.0107	0.1030	0.0065	0.4508
bank	1213	0.0363	0.1870	0.0015	0.8584
informal	1213	0.6323	0.4824	0.0015	0.9581
anyloan	1213	0.6801	0.4666	0.0021	0.9418
spandana$_{a}mt$	1213	0.0000	0.0000	68.9655	0.1890
othermfi$_{a}mt$	1213	201.1542	2742.3639	170.3565	0.5670
bank$_{a}mt$	1213	7438.1698	173268.3440	-5419.6969	0.2760
informal$_{a}mt$	1213	28460.0165	65312.1557	-570.4598	0.8555
anyloan$_{a}mt$	1213	37892.0033	191291.5692	-5878.5550	0.3407
total$_{b}iz$	1220	0.3197	0.6822	-0.0189	0.5778
female$_{b}iz$	1220	0.1451	0.4003	-0.0066	0.7492
female$_{b}iz_{p}ct$	295	0.4883	0.4824	-0.0056	0.9038
bizrev	295	15990.6441	53488.7483	4501.2292	0.5375
bizexpense	295	3617.4160	26144.0332	640.5524	0.7500
bizinvestment	295	384.7119	3156.8150	14.4783	0.9584
bizemployees	295	0.1695	0.8279	0.2545	0.1448
hours$_{w}eekbiz$	295	76.3153	66.0544	-4.5873	0.4116
total$_{e}xp_{m}o$	1220	4888.4312	4074.3721	269.6827	0.2294
nondurable$_{e}xp_{m}o$	1220	4734.6854	3839.8025	251.8193	0.2318
durables$_{e}xp_{m}o$	1220	153.7457	584.5941	17.8634	0.5298
home$_{d}urable_{i}ndex$	1220	1.9414	0.8291	0.0269	0.6680
bizrev$_{a}llHH$	1220	3866.5902	27146.7868	903.6492	0.6248
bizexpense$_{a}llHH$	1220	874.7031	12932.6650	116.4962	0.8114
bizinvestment$_{a}llHH$	1220	93.0246	1559.0523	-0.0984	0.9988
bizemployees$_{a}llHH$	1220	0.0410	0.4130	0.0575	0.1629
hours$_{w}eekbiz_{a}llHH$	1220	18.4533	46.0537	-1.8011	0.3983

Notes: Significant at the ***[1%] **[5%] *[10%] level.

Table 2: Table 3 Self-employment activities for All households (Panel A, Endline 1)

outcome	coefficient	std$_error$	p$_value$	r$_squared$	control$_mean$	control$_sd$	N
bizassets$_1$	597.510000	383.518000	0.119000	0.002000	2497.549000	10802.223000	6800
bizinvestment$_1$	390.853000	212.695000	0.066000	0.001000	280.069000	4037.973000	6800
bizrev$_1$	926.592000	1182.463000	0.433000	0.008000	4856.380000	33108.469000	6608
bizexpense$_1$	254.664000	1056.330000	0.809000	0.004000	4055.445000	30445.990000	6685
bizprofit$_1$	354.338000	313.531000	0.258000	0.001000	744.898000	10694.572000	6239
any$_biz_1$	0.008000	0.021000	0.699000	0.004000	0.349000	0.477000	6810
total$_biz_1$	0.018000	0.038000	0.636000	0.003000	0.503000	0.854000	6810
newbiz$_1$	0.015000	0.007000	0.043000	0.002000	0.053000	0.250000	6757
female$_biz_new_1$	0.014000	0.005000	0.007000	0.003000	0.026000	0.174000	6762
biz$_index_all_1$	0.036000	0.019000	0.056000	0.004000	0.000000	0.466000	6810

Notes: Significant at the ***[1%] **[5%] *[10%] level.

Table 3: Table 3 Self-employment activities for All households (Panel B, Endline 2)

outcome	coefficient	std$_error$	p$_value$	r$_squared$	control$_mean$	control$_sd$	N
bizassets$_2$	1260.792000	529.814000	0.017000	0.003000	5002.792000	14422.972000	6142
bizinvestment$_2$	-133.688000	206.581000	0.518000	0.002000	1007.315000	9622.613000	6142
bizrev$_2$	266.109000	526.470000	0.613000	0.001000	5847.052000	16784.047000	6116
bizexpense$_2$	-530.422000	546.714000	0.332000	0.001000	5224.676000	20603.004000	6116
bizprofit$_2$	541.995000	371.709000	0.145000	0.002000	953.132000	11279.838000	6090
any$_biz_2$	0.023000	0.023000	0.332000	0.005000	0.418000	0.493000	6142
total$_biz_2$	0.045000	0.040000	0.261000	0.007000	0.561000	0.787000	6142
newbiz$_2$	0.003000	0.013000	0.834000	0.003000	0.093000	0.328000	6142
female$_biz_new_2$	-0.005000	0.006000	0.450000	0.001000	0.047000	0.232000	6142
biz$_index_all_2$	0.015000	0.019000	0.419000	0.003000	-0.000000	0.507000	6142

Notes: Significant at the ***[1%] **[5%] *[10%] level.

Table 4: Table 3B Self-employment activities for Households with old Businesses (Panel A, Endline 1)

outcome	coefficient	std$_error$	p$_value$	r$_squared$	control$_mean$	control$_sd$	N
bizassets$_1$	361.168000	226.060000	0.110000	0.002000	606.264000	5193.463000	4647
bizinvestment$_1$	65.764000	80.285000	0.413000	0.001000	104.470000	3313.506000	4647
bizrev$_1$	-464.783000	485.862000	0.339000	0.001000	781.842000	23488.674000	4636
bizexpense$_1$	-253.426000	280.102000	0.366000	0.002000	500.331000	13588.108000	4643
bizprofit$_1$	-211.868000	209.162000	0.311000	0.001000	271.558000	9982.472000	4602
bizemployees$_1$	-0.002000	0.004000	0.643000	0.001000	0.007000	0.142000	4651

Notes: Significant at the ***[1%] **[5%] *[10%] level.

Table 5: Table 3B Continued (Panel B, Endline 2)

outcome	coefficient	std$_error$	p$_value$	r$_squared$	control$_mean$	control$_sd$	N
bizassets$_2$	899.008000	458.849000	0.050000	0.003000	2572.523000	9550.605000	4159
bizinvestment$_2$	150.648000	182.542000	0.409000	0.003000	448.589000	4469.494000	4159
bizrev$_2$	128.941000	405.716000	0.751000	0.001000	2789.109000	11725.338000	4152
bizexpense$_2$	171.196000	344.933000	0.620000	0.001000	2043.312000	9975.287000	4152
bizprofit$_2$	397.870000	230.295000	0.084000	0.002000	519.683000	6653.591000	4144
bizemployees$_2$	0.012000	0.033000	0.721000	0.000000	0.116000	1.216000	4159

Notes: Significant at the ***[1%] **[5%] *[10%] level.

Table 6: Table 3C Self-employment activities for Households with new businesses (ELI only)

outcome	coefficient	std$_error$	p$_value$	r$_squared$	control$_mean$	control$_sd$	N
bizassets$_1$	-872.616000	2200.692000	0.692000	0.020000	8410.855000	24129.793000	356
bizinvestment$_1$	-705.549000	1324.094000	0.594000	0.010000	2418.092000	13759.739000	356
bizrev$_1$	-8166.723000	7313.812000	0.265000	0.029000	17423.027000	91781.656000	332
bizexpense$_1$	-5012.906000	4048.930000	0.217000	0.033000	12114.005000	53020.098000	339
bizprofit$_1$	-3547.546000	3813.119000	0.353000	0.026000	6081.093000	43517.223000	270
bizemployees$_1$	-0.195000	0.112000	0.084000	0.037000	0.289000	1.325000	356
biz$_index_new_1$	-0.081000	0.044000	0.068000	0.038000	0.000000	0.571000	356

Notes: Significant at the ***[1%] **[5%] *[10%] level.