

## CASE STUDY 2: EXPANDING CREDIT ACCESS

Why Randomize?

---



This case study is based on “Expanding Credit Access: Using Randomized Supply Decisions To Estimate the Impacts,” by Dean Karlan (Yale) and Jonathan Zinman (Dartmouth).

J-PAL thanks the authors for allowing us to use their paper.

## KEY VOCABULARY

---

**Counterfactual:** what would have happened to the participants in a program had they not received the intervention. The counterfactual cannot be observed from the treatment group; can only be inferred from the comparison group.

**Comparison Group:** in an experimental design, a randomly assigned group from the same population that does not receive the intervention that is the subject of evaluation. Participants in the comparison group are used as a standard for comparison against the treated subjects in order to validate the results of the intervention.

**Program Impact:** estimated by measuring the difference in outcomes between comparison and treatment groups. The true impact of the program is the difference in outcomes between the treatment group and its counterfactual.

**Baseline:** data describing the characteristics of participants measured across both treatment and comparison groups prior to implementation of intervention.

**Endline:** data describing the characteristics of participants measured across both treatment and comparison groups after implementation of intervention.

**Selection Bias:** statistical bias between comparison and treatment groups in which individuals in one group are systematically different from those in the other. These can occur when the treatment and comparison groups are chosen in a non-random fashion so that they differ from each other by one or more factors that may affect the outcome of the study.

**Omitted Variable Bias:** statistical bias that occurs when certain variables/characteristics (often unobservable), which affect the measured outcome, are omitted from a regression analysis. Because they are not included as controls in the regression, one incorrectly attributes the measured impact solely to the program.

## INTRODUCTION

---

Access to credit has long been considered an important way to improve the welfare of the poor. Yet the jury is still out on whether expanding access to consumer credit is useful (as opposed to entrepreneurial loans).

In the early 2000s, a South African Bank – which we can call Lender – was operating for twenty years as one of the most profitable micro-lenders in the country. It worked mostly with the “working poor”, and offered small, uncollateralized loans that had a high interest and a fixed repayment schedule. Borrower monthly income was around 2303R (US\$300), and the median loan size around 975R (US\$127).

Lender’s business was transparent and linear. It had no application fees, surcharges or insurance premiums. Most of its borrowers (98 percent) received the standard First Time package: a 4-month maturity at 11.75 percent per month charged on the original balance. (This is equivalent to a 200 percent APR.)

## WHAT HAPPENS IF WE EXPAND CREDIT?

---

What is required in order for us to measure whether these credit packages worked, or whether credit had any impact on household income?

In general, to ask if a program works is to ask if the program achieves its goal of changing certain outcomes for its participants, and ensure that those changes are not caused by some other factors or events happening at the same time. To show that the program causes the observed changes, we need to simultaneously show that if the program had not been implemented, the observed changes would not have occurred (or would have been different). But how do we know what would have happened? If the program happened, it happened. Measuring what would have happened requires entering an imaginary world in which the program was never given to these participants. The outcomes of the same participants in this imaginary world are referred to as the counterfactual. Since we cannot observe the true counterfactual, the best we can do is to estimate it by mimicking it.

The key challenge of program impact evaluation is constructing or mimicking the counterfactual. We typically do this by selecting a group of people that resemble the participants as much as possible but who did not participate in the program. This group is called the comparison group. Because we want to be able to say that it was the program and not some other factor that caused the changes in outcomes, it is important that the only difference between the comparison group and the participants is that the comparison group did not participate in the program. We then estimate “impact” as the difference observed at the end of the program between the outcomes of the comparison group and the outcomes of the program participants.

The impact estimate is only as accurate as the comparison group is successful at mimicking the counterfactual. If the comparison group poorly represents the counterfactual, the impact is (in most circumstances) poorly estimated.

Therefore the method used to select the comparison group is a key decision in the design of any impact evaluation.

That brings us back to our questions: What impact do micro-loans have on the welfare of borrowers?

In this case, the intention of the program is to “improve borrower’s welfare” and household income is the outcome measure. So, when we ask if Lender’s loans “worked”, we are asking if they raised household income. The impact is the difference between the household income of households with loans compared to what those same households’ income would have been had they not taken the loans.

What comparison groups can we use? The following experts illustrate different methods of evaluating impact. (Refer to the table on the last page of the case for a list of different evaluation methods.)

## ESTIMATING THE IMPACT OF ACCESS TO CREDIT

---

### METHOD 1:

*Newspaper article: Huge gains for microloan clients*

Statistics released today reveal that prominent South African bank, Lender, and its new micro-lending program to working class clients in Johannesburg and Cape Town has led to a 78 percent increase in client household incomes. When clients approached Lender, their typical monthly income was 3070R; when households were interviewed 6 to 12 months after taking their loans, their incomes had jumped to 5454R. This provides only further evidence to what has already been recognized – largely via the 2006 Nobel Peace Prize to Muhammad Yunus and his revolutionary Grameen Bank - as one of the powerhouse new methods to fight poverty and raise incomes worldwide.

**TABLE 1**

	<i>Mean</i>	<i>Standard Error</i>
<i>Baseline monthly income</i>	3070R	250R
<i>Endline monthly income</i>	5454R	1144R
<i>Difference</i>	2384R***	
<i>Difference (percent)</i>	78% increase	

Note: Statistically significant at the 95 percent level. Sample size of 150 households.

**Discussion Topic 1**

*Identifying evaluation*

1. What type of evaluation does this opinion piece imply?
2. What represents the counterfactual?
3. What are the problems with this type of evaluation?

**METHOD 2:**

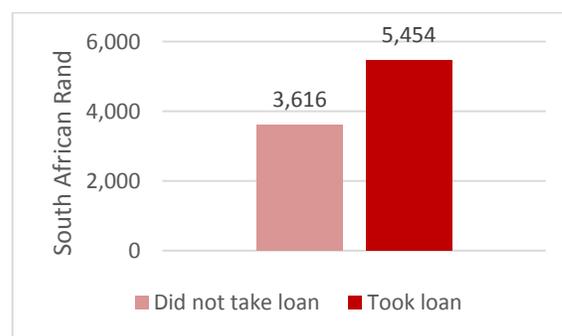
Letter to the Editor: Let’s not jump to conclusions

I have been a follower of the microfinance revolution since its beginnings in the 1970s. And while I’m happy to see South Africa join the rest of the world in exploring this exciting new space, I have some qualms with your newspaper’s article from last week: Huge gains for microfinance clients. I am an economist by training, a dabbler in statistics and an observer of human nature, and I can tell you that – while Lender clients’ gains might certainly be there, they are certainly not “huge”. After all, who knows what happened to those clients between getting the loan and reporting their incomes? South Africa, as you know, has been growing at a healthy 5.4 percent in 2006, with much of the growth occurring in the very “working class” districts Lender targets. Let’s get rid of this confounding temporal aspect, and focus instead

on how households with loans were doing, as compared to households without loans. I’ve gone ahead and done such a calculation – and you’ll see I found a healthy 51 percent impact. Still something to be proud of, but almost 30 percent less than what you found!

**FIGURE 1**

Comparison of Household Income of Households That Took Loans vs. Those That Did Not Take Loans



**Discussion Topic 2**

*Identifying evaluation*

1. What type of evaluation is this opinion piece employing?
2. What represents the counterfactual?
3. What are the problems with this type of evaluation?

**METHOD 3:**

Donor report: Comparing apples to apples

While Lender’s access to credit program has received a good amount of positive press coverage, this report attempts to provide a more accurate measure of its impact. Previous analyses used the wrong metrics to calculate impact – possibly overestimating how much household incomes improved after receiving loans. For example, by comparing those households who took loans with those that did not, prior analyses may be introducing selection bias into the estimate. Those two groups may be very different for many reasons, not just whether they

took loans. Case and point: consider the baseline income levels: non-borrowing households already started out with less – only 2300R – compared to the households that asked for loans, that started with 3069R. This demonstrates that these two groups are different to begin with.

Thus, comparing income *levels* may not be as helpful as comparing *changes in income* of the two groups. That way, we see how *fast* incomes increased for each group. When we repeated the analysis using this more-appropriate outcome measure, *change in income*, non-borrowing households experienced a 57 percent increase in their income, while borrowing households grew by 78 percent. This is a difference of 21 percentage points. And it proves that motivated households, when given access to credit, are able to make their income grow via investments and other activities.

### Discussion Topic 3

#### *Identifying evaluation*

1. What type of evaluation is this letter using?
2. What represents the counterfactual?
3. What are the problems with this type of evaluation?

### METHOD 4:

#### *The numbers don't lie, unless your statisticians are asleep*

Lender celebrates victory; skeptics say it's an exaggeration. A closer look shows that, as usual, the truth is somewhere in between. Lender and its advocates assert that the relaxation of certain lending rules allowed a number of worthy households newfound access to credit, which resulted in substantial gains in household income (up to 78 percent). Detractors instead claim that these gains are inflated, and that gains might be modest and – perhaps more importantly – driven more by the inherent motivation of the households themselves, rather than the credit per se.

Unfortunately, this battle is being waged using instruments of analysis that are seriously flawed. The ultimate victim is the public who is looking for an answer to the question: is Lender helping its intended beneficiaries?

This report uses sophisticated statistical methods to measure the true impact of Lender's loans. We were concerned about other variables confounding previous results. We therefore conducted a survey of these households to collect information on gender, race, marital status, employment and household size. Given that these variables no doubt affect a household's monthly income, controlling for them will allow us to separate out the true effect of the loans.

Looking at Table 1, we find that the results change and our impact estimate drops as we control for additional variables. The results from column (1) suggest that Lender's loans lead to a 1835R (or 67 percent) increase in household monthly income – and this is significant at the 15 percent level. Column (2), which includes controls for a number of other confounding variables, reduces the size of this increase – from 1835R (67 percent) to 1412R (51 percent). Furthermore, the loan is no longer statistically significant – instead, education becomes statistically significant at the 5 percent level.

Therefore, controlling for the right variables, we have discovered that Lender has not led to such a large increase in household income, and furthermore, it may not be the most important variable which impacts household income overall.

It looks like Lender is once again a simple lender, rather than a savior of the poor.

### Discussion Topic 4

#### *Identifying evaluation*

1. What type of evaluation is this report utilizing?
2. What represents the counterfactual?
3. What are the problems with this type of evaluation?

**TABLE 1: MONTHLY INCOME**

**Key independent variable:** access to loans is the treatment; the analysis tests the effect of loans on household monthly income.

**Dependent variables:** household monthly income is the main dependent variable, except we test with and without controls in (1) and (2), respectively.

**Control variables:** independent variables other than the loan that may influence a household's monthly income.

**Statistical significance:** the corresponding result is unlikely to have occurred by chance, and thus is statistically significant (credible)

	(1)	(2)
Loan	1837.645* (1234.512)	1411.753 (1261.524)
Female borrower		-2284.92 (1476.018)
Education		690.4036*** (306.133)
Race: African		1070.203 (1277.297)
Race: Coloured		2083.461** (1183.304)
Race: Indian		341.7177 (1070.572)
Marital status		-230.8048 (542.7396)
Household size		438.5256** (263.0638)
Employed		755.9049 (948.3008)
Years employed		-25.96757 (107.5894)
Risk: Green		401.3943 (2814.721)
Risk: Grey		-787.6992 (1703.341)
Risk: Red		-203.2747 (1219.793)
Constant	3616.351*** (463.6122)	-4805.349 (5011.634)

THE DEBATE WERE ARTIFICIALLY PRODUCED FOR THE PURPOSE OF THE CASE.

CASE STUDY 2 • WHY RANDOMIZE • ABDUL LATIF JAMEEL POVERTY ACTION LAB

	Methodology	Description	Who is in the comparison group?	Required Assumptions	Required Data
Quasi-Experimental Methods	Pre-Post	Measure how program participants improved (or changed) over time.	Program participants themselves—before participating in the program.	The program was the only factor influencing any changes in the measured outcome over time.	Before and after data for program participants.
	Simple Difference	Measure difference between program participants and non-participants after the program is completed.	Individuals who didn't participate in the program (for any reason), but for whom data were collected after the program.	Non-participants are identical to participants except for program participation, and were equally likely to enter program before it started.	After data for program participants and non-participants.
	Differences in Differences	Measure improvement (change) over time of program participants <i>relative to</i> the improvement (change) of non-participants.	Individuals who didn't participate in the program (for any reason), but for whom data were collected both before and after the program.	If the program didn't exist, the two groups would have had identical trajectories over this period.	Before and after data for both participants and non-participants.
	Multivariate Regression	Individuals who received treatment are compared with those who did not, and other factors that might explain differences in the outcomes are "controlled" for.	Individuals who didn't participate in the program (for any reason), but for whom data were collected both before and after the program. In this case data is not comprised of just indicators of outcomes, but other "explanatory" variables as well.	The factors that were <i>excluded</i> (because they are unobservable and/or have been not been measured) do not bias results because they are either uncorrelated with the outcome <u>or</u> do not differ between participants and non-participants.	Outcomes as well as "control variables" for both participants and non-participants.
	Statistical Matching	Individuals in control group are compared to similar individuals in experimental group.	<u>Exact matching</u> : For each participant, at least one non-participant who is identical <i>on selected characteristics</i> . <u>Propensity score matching</u> : non-participants who have a mix of characteristics which predict that they would be as likely to participate as participants.	The factors that were <i>excluded</i> (because they are unobservable and/or have been not been measured) do not bias results because they are either uncorrelated with the outcome <u>or</u> do not differ between participants and non-participants.	Outcomes as well as "variables for matching" for both participants and non-participants.
	Regression Discontinuity Design	Individuals are ranked based on specific, measurable criteria. There is some cutoff that determines whether an individual is eligible to participate. Participants are then compared to non-participants and the eligibility criterion is controlled for.	Individuals who are close to the cutoff, but fall on the "wrong" side of that cutoff, and therefore do not get the program.	After controlling for the criteria (and other measures of choice), the remaining differences between individuals directly below and directly above the cut-off score are not statistically significant and will not bias the results. A necessary but sufficient requirement for this to hold is that the cut-off criteria are strictly adhered to.	Outcomes as well as measures on criteria (and any other controls).
	Instrumental Variables	Participation can be predicted by an incidental (almost random) factor, or "instrumental" variable, that is uncorrelated with the outcome, other than the fact that it predicts participation (and participation affects the outcome).	Individuals who, because of this close to random factor, are predicted not to participate and (possibly as a result) did not participate.	If it weren't for the instrumental variable's ability to predict participation, this "instrument" would otherwise have no effect on or be uncorrelated with the outcome.	Outcomes, the "instrument," and other control variables.
Experimental Method	Randomized Evaluation	Experimental method for measuring a causal relationship between two variables.	Participants are randomly assigned to the control groups.	Randomization "worked." That is, the two groups are statistically identical (on observed and unobserved factors).	Outcome data for control and experimental groups. Control variables can help absorb variance and improve "power".

