Measuring the impact of an ongoing microcredit project
Eriksen, Steffen; Lensink, Bernardus

Published in:
Journal of Development Effectiveness

DOI:
10.1080/19439342.2015.1095782

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Measuring the impact of an ongoing microcredit project: evidence from a study in Ghana

Steffen Eriksen\textsuperscript{a} and Robert Lensink\textsuperscript{a, b, *}

\textsuperscript{a}Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands; \textsuperscript{b}Development Economics Group, Wageningen University, Wageningen, The Netherlands

This article uses a mixed method approach to assess the impact of a microfinance organisation in Ghana. By combining propensity score matching with a double-difference method, the authors determine that microcredit has a positive effect on expenditures but does not positively affect a series of other outcome variables. A list experiment further suggests that microcredit loan proceeds often are not spent productively.

Keywords: microcredit; impact evaluation; summary indexation; list experiment; Ghana

1. Introduction

A recent burst of randomised controlled trials (RCT) has sought to investigate the impacts of microcredit (for example, Angelucci, Karlan, and Zinman 2014; Attanasio et al. 2014; Augsburg et al. 2014; Banerjee et al. 2014; Crépon et al. 2014; Tarozzi, Desai, and Johnson 2014). These studies provide empirical evidence of a small, positive effect of microcredit on people’s lives. Yet much remains to be learned, and the results of these RCTs are tentative, because in practice, the ideal conditions required for RCTs virtually never hold (Deaton 2010), such that outcomes differ both between and within countries.

The scope for RCTs also is limited, because most microcredit programmes cannot support random assignments. Moreover, policymakers often want to evaluate ongoing projects. In these cases, a retrospective evaluation that assesses the programme’s impact after it has been implemented, using observational data, can be done. However, questions persist about the reliability of impact studies using observational data. A recent special issue of the European Journal of Development Research summarises this debate about whether nonexperimental designs can and should be used to answer attribution questions. The central question is determining the extent to which observational studies can adequately address self-selection and programme placement biases that lead to inaccurate impact estimates. Lensink (2014) argues that rigorous, nonexperimental evaluations of ongoing projects are difficult. He calls for selectivity, suggesting that expensive impact analyses of ongoing projects should be conducted only in certain conditions, such as when it is possible to obtain pre-intervention data. White (2014) instead argues for more observational studies, in the belief that nonexperimental studies, if designed appropriately, can effectively control for selection biases.\textsuperscript{1}

We respond to White (2014) by considering the results of a retrospective, nonexperimental impact evaluation of a microcredit programme in Ghana. Specifically, we analyse the impact of the microcredit provided by a leading microfinance institution in Ghana,

\textsuperscript{*}Corresponding author. Email: B.W.Lensink@rug.nl

© 2015 Taylor & Francis
Opportunity International. This institution provides different financial services to men and women, but its main product is productive credit offered in the form of group and individual loans to women. Although RCTs offer the greatest rigour for studying attribution questions, the approach we apply in this evaluation constitutes an appropriate alternative evaluation tool for policymakers, if random assignment is not feasible. Therefore, in addition to providing new results about the impact of microcredit in Ghana, which are relevant in their own right, this study suggests a nonexperimental identification strategy when random assignment is not feasible. An important aspect of the underlying theory of change is that the loans are mainly used as a productive investment and not on consumption items. To test this aspect of the theory of change, we examine loan use with list randomisation, which we compare against loan use derived from direct surveys. In line with Karlan and Zinman (2012), we find that a sizeable portion of the borrowers use their loan for consumption purposes rather than as a productive investment, making the microcredit market appear like a consumer loan market.

In the next section, we describe our approach, which includes the theory of change, an explanation of our methodology and a summary of the data sampling process. Section 3 contains the analysis description and the main results. Finally, we conclude with a discussion in Section 4.

2. Approach
2.1. Theory of change

With this article, we primarily aim to examine whether microcredit ultimately contributes to people’s welfare. An elementary approach to model how microcredit can achieve an increase in people’s lives is through the use of a results chain, which for our study looks like the version in Figure 1.

The main channel through which microcredit can increase incomes or expenditures is through productive investments, because microloans are designed to finance and support businesses. The increase in assets and income that results from expanded businesses can

![Figure 1. Results chain.](image-url)
be used to finance improvements in various welfare indicators, such as housing quality or children’s education. Microcredit also constitutes a promising way to empower women, in that the majority of microfinance services are provided to women. Through their thus supported income-generating activities, women gain the financial means to establish their independence.²

Microcredit also can improve welfare through many other channels, and the impact of microcredit on one outcome indicator might depend on the effects of another indicator. For example, the likelihood that microcredit will improve children’s education is probably higher when women are more empowered. Ideally, we would seek to examine the various possible channels through which microcredit can improve welfare to provide a comprehensive test of an elaborate theory of change. However, our data set cannot support such an analysis, because relevant details are missing. Instead, we focus on the main channel, business investment channel, without addressing the other channels through which microcredit exerts an impact.

A major risk in this setting is that borrowers might not use microcredit productively but instead rely on it for consumption purposes. In this case, expenditures still may increase, through a demand effect, but it is unlikely that assets will increase, so long-term improvements might not emerge from microcredit. In the subsequent analysis, we test this main channel explicitly, according to our theory of change.

2.2. Outcome variables and index construction

In this study we define a set of outcomes that includes expenditures, animals, nonproductive assets, education, housing condition and more (see Table A1 for more). This list may seem long but represents only a small part of all outcomes. Nevertheless, the total number of outcomes reaches 32. Thus, we implement summary index test which pool multiple outcomes into a single test to avoid over-testing. This reduces the total number of outcome indicators to six. To measure the impact of microcredit on education, we assessed two indicators: the education level attained by the oldest son in the household (SEDUC) and the education level attained by the oldest daughter in the household (DEDUC). For income, we used the sum of a list of inflation-adjusted household expenditures (EXPEND). We also constructed an asset index (ASSET), based on 20 household assets; an index measuring the quality of housing (HOUSING), which reflects the roof type, the number of rooms and whether or not the household has toilet facilities and electricity connections; and a female empowerment index (EMPOW), constructed of a list of questions that describe women’s participation in household decision-making processes.

The index construction follows Anderson’s (2008) suggested four-step procedure: (1) code all variables such that a positive change is associated with ‘better’; (2) standardise each outcome by subtracting the mean from the baseline comparison group, then dividing by the standard deviation of the baseline comparison group; (3) assign each outcome variable to the relevant index; and (4) construct the summary index variable by taking a weighted average of each of the outcome variables in the index. The weights are equivalent to the inverse of the covariance matrix of the standardised outcomes for each group. When interpreting the outcome of a summary index test as the one constructed earlier, it is interpreted as summary effect size, as a summary index basically is a weighted mean of several standardised outcomes. Further details about the construction of the outcome indicators are available on request.
2.3. Methodology

To estimate the causal impact of microcredit, we combined propensity score matching (PSM) with a double-difference methodology. Using baseline information about the treated and comparison households, we estimated a model of programme participation, based on observational demographic and location characteristics and baseline information about the main outcome indicators. With propensity scores, we defined a region of common support, and then ignored any households outside this region in the further analyses. Therefore, we ensured that the treated and comparison groups were comparable in their observed characteristics. To determine the impact of microcredit, we next applied a double-difference model to all households in the common space. This procedure ensured that we controlled for selection due to time-invariant observable traits. We considered estimates both with and without covariates. The model with covariates can be specified as follows:

\[ Y = aT + b \times TIME + c \times T \times TIME + d \times Controls + e \]

where \( Y \) is a vector of outcome variables, \( T \) is an indicator variable equal to 1 if a household is in the treatment group and 0 otherwise; \( TIME \) is a dummy equal to 1 for the endline and 0 for the baseline; and \( Controls \) is a vector of the control variables equivalent to those used to estimate the propensity score. The coefficient \( c \) is our parameter of interest; it measures the impact of microcredit. We estimated a balanced panel, so our double-difference specification without controls is similar to a regression model with household- and time-fixed effects. We applied ordinary least squares (OLS) to all models; for the binary outcome variables, we used a linear probability model. We clustered standard errors at the branch level, because households in the same branch are likely correlated.

2.4. Data and sampling issues

Using lists of borrowers, kindly provided to us by Opportunity International, we randomly sampled a group of borrowers that took out their first loan in 2011 but had not yet taken a second loan. The borrowers came from 10 branches of Opportunity International, spread over two regions (Ashanti and Brong-Ahafo). Most of them were from rural areas, but a few represented urban areas as well. The comparison group contained people who lived near the Opportunity International branches but had never borrowed. We obtained a sample of 485 treatment and 511 control entries.

With the help of M-CRIL, we organised a survey during February–March 2014. Without baseline data about borrowers, we decided to construct pre-intervention data by using recall, which would enable a double-difference methodology. Specifically, during the one-to-one interviews, we asked for information about the participants’ current situation, and we also tried to obtain relevant information referring to 2010, just before they received their first loan.

Prior literature features an ongoing discussion about the usefulness and reliability of recall data. On the one hand, using recall data substantially reduces survey costs, because it is not necessary to interview a panel of households several times, which also should reduce attrition problems. Moreover, it supports the construction of a baseline for ongoing projects. On the other hand, recall is less precise and probably less reliable. As rightly indicated by White (2014) though, it is going too far to argue that recall is by definition...
untrustworthy; all questionnaires use recall, and recall is certainly reliable for variables that are easy to remember, such as births, education enrolments and so forth. Nor is there any a priori reason that recall must lead to a systematic bias between treated and control households. In an interesting article, Nicola and Gine (2014) show, with a sample of small-scale boat owners in India, that the absolute value of recall error increases with the recall period, which implies that errors are not white noise. But their findings also indicate that recall converges to a mean, not to the most recent observations. Thus, recall may be appropriate if researchers are interested in the mean of earnings, but not if they are interested in its volatility. Although it is difficult to say precisely what this recommendation implies for our study, it seems to suggest that our approach may tend to underestimate the impact of microcredit, because the recall values reported by the treatment group may reflect the mean value of the period before the interview, such that they partly include the period after the first loan. However, there are reasons to believe that people who belong to the treatment group and thus self-select into microcredit will recall information related to the outcome indicators better than people from the comparison group. Nicola and Gine (2014) show that it is reasonable to believe that people with a higher monthly income recall earnings better than people with lower monthly income. Furthermore, they find that people tend to overstate their earnings. Their findings are consistent with previous findings by Mullainathan (2002) and Delavande, Giné, and McKenzie (2011), who find that entrepreneurs tend to be optimistic and systematically forecast higher average earnings. Thus, if people in general tend to overstate, but the treatment group recalls better, our impact results may be too positive.

3. Analysis

3.1. Propensity score matching and balancing tests

The analysis starts with the first step of the PSM technique, using baseline data. The aim is to determine a common space and exclude all households outside this common space from subsequent analyses. We calculated propensity scores by using the baseline values of the main outcome variables, as well as several household characteristics and location characteristics. The set of covariates thus was marital status, family type, age, family size (household members), ASSET, HOUSING, EMPOW, EXPEND and location dummies. The 68 observations that appeared outside the common support were dropped from the sample (both baseline and endline). Then, with the sample contained within the common space, we conducted balancing tests of our main variables at the baseline and obtained the results in Table 1.

Although we ignored values outside the common space, 7 out of 16 variables are unbalanced at the baseline. Therefore, differences remain between the control and treatment groups. For this reason we take these unbalanced variables into account as controls in our difference in difference estimation. Fortunately and probably more important, all of the outcome variables, but female empowerment, we consider are balanced at the baseline. Yet, it remains vital to control further for observed and unobserved differences between treatment and comparison groups, to ensure that our impact estimates are unbiased.

3.2. Double-difference estimates

We next conducted double-difference estimates including the controls, such that we corrected for the remaining observed differences between the treatment and control households. The double-difference methodology also ensures that the results are not
confounded by unobserved heterogeneity that does not change over time; our sample is balanced, such that we take the fixed effects at the treatment level into account. The estimation results are shown in Table 2.

Table 1. Summary statistics and balancing tests.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Summary statistics</th>
<th>Balancing test</th>
<th>Treatment–Control (with PSM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Treatment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>N</td>
</tr>
<tr>
<td>Son age</td>
<td>439</td>
<td>5.10</td>
<td>433</td>
</tr>
<tr>
<td>Daughter age</td>
<td>432</td>
<td>5.66</td>
<td>424</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>471</td>
<td>0.27</td>
<td>451</td>
</tr>
<tr>
<td>Married</td>
<td>471</td>
<td>0.61</td>
<td>451</td>
</tr>
<tr>
<td>Separated</td>
<td>471</td>
<td>0.01</td>
<td>451</td>
</tr>
<tr>
<td>Divorced</td>
<td>471</td>
<td>0.06</td>
<td>451</td>
</tr>
<tr>
<td>Widowed</td>
<td>471</td>
<td>0.06</td>
<td>451</td>
</tr>
<tr>
<td>Family type (1 = joint)</td>
<td>472</td>
<td>0.90</td>
<td>452</td>
</tr>
<tr>
<td>Age</td>
<td>472</td>
<td>32.92</td>
<td>452</td>
</tr>
<tr>
<td>Household members</td>
<td>472</td>
<td>3.70</td>
<td>452</td>
</tr>
<tr>
<td>SEDUC</td>
<td>328</td>
<td>0.13</td>
<td>268</td>
</tr>
<tr>
<td>DEDUC</td>
<td>329</td>
<td>0.09</td>
<td>259</td>
</tr>
<tr>
<td>ASSET</td>
<td>472</td>
<td>0.00</td>
<td>452</td>
</tr>
<tr>
<td>HOUSING</td>
<td>472</td>
<td>0.00</td>
<td>452</td>
</tr>
<tr>
<td>EMPOW</td>
<td>472</td>
<td>0.00</td>
<td>452</td>
</tr>
<tr>
<td>EXPEND</td>
<td>472</td>
<td>121.16</td>
<td>452</td>
</tr>
</tbody>
</table>

Notes: The right-most column displays the coefficient from separate ordinary least square regressions of the dependent variable on the treatment dummy. Standard errors are clustered at the branch level. The dependent variables are: son age, age of the oldest son in the household; daughter age, age of the oldest daughter in the household; marital status, group of variables indicating the marital status of the respondent; family type, binary variable equal to 1 if the family is a joint family and 0 if the family is nuclear (living with only spouse and children); age, age of the respondent; household members, number of household members; SEDUC, education level attained by the oldest son in the household; DEDUC, education level attained by the oldest daughter in the household; ASSET, index constructed on 20 household assets; HOUSING, index constructed on house characteristics; EMPOW, index constructed on household decision questions; and EXPEND, total monthly inflation-adjusted household expenses. ***Significant at 1 per cent level. **Significant at 5 per cent level.
The treatment effect appears as the coefficient for the interaction term, year × treatment. The results are mixed: it seems encouraging that microcredit has a positive effect on monthly expenditures, which is our proxy for income. A household belonging to the treatment group experienced a monthly increase in expenditures of 27.105 Cedi (around 6 US dollars), which is a 28% higher increase compared with a household from the comparison group. In light of the relatively short period between the acceptance of the loan and the survey (less than 3 years), it may not come as a surprise that we cannot explain any variation in some of the variables of ultimate interest, such as children’s education level, housing or female empowerment.

More disappointing is that we find a summary effect size decrease on assets of 0.067, thus indicating that households in the treatment group experience a 6.7 per cent point decline in asset holdings compared with a household from the comparison group. Although being a relatively small decrease, it may imply that in the long run, the positive effects of microcredit on expenditures fade away, and the potential positive effects on welfare will not materialise over time.

### 3.3. Explaining the results

How can these results be explained? According to our theory of change, microcredit mainly affects long-run welfare through productive investments. However, if loan proceeds are used for consumption purposes, the immediate positive demand effect will still increase expenditures in the short run, but it cannot improve business growth and thus probably not enhance welfare in the long run. Therefore, it seems relevant to investigate how microfinance clients spend their loan proceeds. We addressed this question in two ways. First, we directly asked each microfinance member to indicate the main expenses financed with the microloan, by providing a list of spending possibilities that included productive investment, education, housing investments and household items (for example food, a television, a radio). If the majority of loans were used to purchase the latter items, microcredit is being spent on consumption, not for business. The results of the direct surveys strongly indicated that the

### Table 2. Regression results.

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1) SEDUC</th>
<th>(2) DEDUC</th>
<th>(3) ASSET</th>
<th>(4) HOUSING</th>
<th>(5) EMPOW</th>
<th>(6) EXPEND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>0.028</td>
<td>-0.072</td>
<td>0.028</td>
<td>0.056</td>
<td>0.386***</td>
<td>-8.134</td>
</tr>
<tr>
<td></td>
<td>(0.564)</td>
<td>(0.316)</td>
<td>(0.729)</td>
<td>(0.351)</td>
<td>(0.000)</td>
<td>(0.269)</td>
</tr>
<tr>
<td>Year</td>
<td>0.005</td>
<td>0.033</td>
<td>0.059*</td>
<td>0.043*</td>
<td>0.092***</td>
<td>63.796***</td>
</tr>
<tr>
<td></td>
<td>(0.808)</td>
<td>(0.334)</td>
<td>(0.085)</td>
<td>(0.077)</td>
<td>(0.008)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Year × treatment</td>
<td>-0.003</td>
<td>0.002</td>
<td>-0.067*</td>
<td>-0.017</td>
<td>-0.017</td>
<td>27.105***</td>
</tr>
<tr>
<td></td>
<td>(0.932)</td>
<td>(0.909)</td>
<td>(0.060)</td>
<td>(0.585)</td>
<td>(0.297)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.373</td>
<td>-0.279*</td>
<td>0.255</td>
<td>0.092</td>
<td>-0.425**</td>
<td>31.645**</td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td>(0.100)</td>
<td>(0.125)</td>
<td>(0.548)</td>
<td>(0.029)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Observations</td>
<td>983</td>
<td>945</td>
<td>1844</td>
<td>1844</td>
<td>1844</td>
<td>1844</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.627</td>
<td>0.317</td>
<td>0.026</td>
<td>0.003</td>
<td>0.080</td>
<td>0.205</td>
</tr>
</tbody>
</table>

Notes: Regressions refer to double-difference regressions. Standard errors are clustered at the branch level. All equations are estimated with the following controls: marital status, family type, age of the respondent and number of household members. For SEDUC (DEDUC), a control variable for the age of the oldest son (daughter) enters the specification. Robust $p$-values are reported in parentheses. ***Significant at 1 per cent level. **Significant at 5 per cent level. *Significant at 10 per cent level.
majority of respondents used the loan productively; only 0.1 per cent of microfinance clients reported that household items were the main expenses financed by their loan.

Second, because of the risk that microfinance borrowers might not be truthful, we used an indirect method as well. Most microfinance lenders, including Opportunity International, require loan proceeds to be spent productively, and it is socially more desirable to claim that all loans are used productively. Therefore, in line with Karlan and Zinman (2012), we used so-called list randomisation to elicit loan spending information indirectly. List randomisation allows the microfinance member to conceal his or her answer, yet the researcher can determine whether loan proceeds have been spent on consumption items. In the list randomisation task, the group of microfinance clients is randomly divided into two groups, A and B. Group A reads a set of statements that do not include any sensitive issues; Group B sees the same set of statements but also reads a statement related to the sensitive issue, such as spending on household items in our case. Both groups must report how many statements, but not which ones, are true. We can estimate the proportion of the sample that engages in the sensitive behaviour according to the difference in the mean number of true statements indicated by the two groups. The list of randomisation item for this study contained the following question:

Please tell me with how many of the following statements you agree. I don’t want to know which ones, just how many.

(1) I used part of my loan to finance productive investment in farming.
(2) I have more than three siblings.
(3) I have a TV in my home.
(4) The government should subsidise food prices.

All respondents saw this question; Group B also read the following sensitive item: ‘The main expense financed by my loan refers to household items, such as food, a TV, a radio, etc.’.

The result of this list experiment revealed a striking contrast with the outcomes from the direct questioning: 41 per cent of the respondents indirectly revealed that they spend the major part of their loan on household items. Thus, direct questioning overestimates productive loan use enormously.6

4. Discussion

Until recently, many studies consistently confirmed the positive impacts of microcredit; recent work instead suggests that these studies were far too positive about microcredit, because they failed to address problems related to self-selection. The impacts of microcredit continue to be widely discussed, and substantial uncertainty persists about its medium- and long-run effects. What seems clear though is the inverse relationship that consistently emerges between the ‘rigour’ of a study and its ‘impact’: the more a study controls for self-selection, the lower the impact of microcredit seems to be. Does this finding imply that only impact studies using experimental approaches such as RCT are valid, whereas those based on observable information are useless? In our view, such a dismissal is too negative and unrealistic; many interventions cannot be randomised simply because they already have taken place. We argue that one should attempt to use nonexperimental identification strategies when an RCT is not feasible. However, measuring the impact of an ongoing project, especially if pre-intervention data are not available, has serious caveats that demand careful consideration.
We sought to measure the impact of an ongoing microcredit project in Ghana, for which no baseline data were available. To address the attribution question as far as possible, we have suggested and applied a method that combines PSM with a double-difference methodology and then constructed pre-intervention data using recall. By calculating propensity scores and constructing a common space, we can identify microfinance clients and comparison households that exhibit substantial differences and thereby exclude them from our analysis. Yet balancing tests showed persistent differences between controls and treatments, even after ignoring households outside the common space. This signalled the need of further controlling for selection biases, and hence the importance of using a double-difference methodology (including control variables) which enables to further control for selection biases due to observables, as well as unobservables that are not changing over time. The crucial importance of baseline data, which is needed to conduct the double-difference method, indicates a clear benefit of recall. Yet, if the treatment group is more precise in recalling than the comparison group, the recall method may bias the results. Additional research is needed to determine whether the potential biases due to recall outweigh the clear benefits.

Our findings are largely similar to recent experimental research pertaining to microcredit. Various recent studies have also shown that the impacts of microcredit are minor, and often much smaller than expected, at least in the short run. We find a positive impact on expenditures, but we do not uncover any impacts of microcredit on education, housing quality or female empowerment, and even a slightly negative impact on assets. This result at the least provides some confidence in the rigour of our methodology.

We sought to explain our main outcomes by investigating how loan funds are used. The striking result of this assessment was that almost 50 per cent of microfinance members revealed, indirectly, that household items were the main expense they financed with their microcredit loans. This outcome clearly indicates the challenge associated with obtaining information about sensitive question items through direct surveys; when we asked the microfinance members directly about their spending, only 0.1 per cent admitted that they had spent their loans mainly on household items. The extent to which this outcome might also invalidate the answers to other questions, and information obtained from self-reported data obtained by questionnaires in general, is unclear. We posit that the potential bias would be smaller, because most of the other issues addressed are not particularly sensitive. Yet, more research about potential biases related to direct questioning seems important.

Acknowledgements
The authors thank Opportunity International in Ghana for making the field study happen, M-CRIL for conducting the field research and S.R Budjhawan and G. Hieminga, from ING, and I.J. Unger, from AllWays Impact, for initiating the study and their constructive discussions. The authors also thank two anonymous referees for their helpful comments.

Disclosure statement
No potential conflict of interest was reported by the authors.
Notes

1. White’s argument is made more valid when observational studies are undertaken in the context of a causal chain analysis which analyses successive outcomes along that chain.

2. In a recent review article by Vaessen et al. (2014), they conclude that it is very unlikely that microcredit has an overall impact on women empowerment, as they find a lack of evidence for microcredit having an impact on women’s control over household resources. Thus, although microcredit is believed to have a promising effect on women empowerment, impacts should be taken with caution.

3. A potential limitation of the study is that besides from asking the participants to recall their status from 3 years ago right before they received their first loan, we did not anchor the participants to the appropriate time period.

4. It should be noted that this selection criteria may have implications for the impact quantity estimated (especially if spillover effects are present). Optimally, the comparison group should have been selected from areas not linked to microcredit branches.

5. Recall that our balancing tests showed that EMPOW was not balanced at baseline. The double-difference approach partially controls for this potential bias as can be seen by the positive coefficient on the treatment variable, which picks up the difference at baseline (in Table 2).

6. A test for difference in proportions between the two groups showed a significant difference at the 1 per cent level ($p = 0.0011$).

7. Additional analysis was undertaken to ensure that age and marital status did not affect the outcome equations. As long as age is highly correlated with the other factors that could “interact with the productive activities”. Adding age directly controls for these potential biases.

References


Appendix

Table A1. Summary index components.

<table>
<thead>
<tr>
<th>Summary index</th>
<th>Summary index components</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASSET</td>
<td>Bicycle, car, motorcycle, bus, tractor, sewing machine, gas, DVD/CD player, music system, microwave, refrigerator, washing machine, television, water purifier, mobile phone</td>
</tr>
<tr>
<td></td>
<td>Furniture, number of cattle, number of goats, number of pigs, number of poultry</td>
</tr>
<tr>
<td>HOUSING</td>
<td>Electricity connection, toilet facility, number of rooms, roof type</td>
</tr>
<tr>
<td>EMPOW</td>
<td>Five questions on household decisions: day-to-day expenses, business expenses, loan taking, loan spending, children’s education</td>
</tr>
</tbody>
</table>