

The Impact of Rural Business Services on the Economic Well-Being of Small Farmers in Nicaragua Markus Goldstein

This paper examines the impact of the supply of technical and financial training, supplies and some capital to farmers in Nicaragua as part of the MCC program. Using a randomized roll out design, as well as careful attention to the potential chain of results, the authors are able to trace out the impacts of the program on farming activities, income, investment and household consumption over time.

This evaluation design (and the way it is carefully explained in the paper) have a lot to recommend it as a model for other evaluations (in the MCC and elsewhere):

- It utilizes a *randomized roll out design*, comparing early treatment folks to those who come later. This is a good option when capacity constraints bind (as they often do) and a non-treatment control group presents political and/or ethical problems.
- It compares *treated compliers with soon to be treated compliers*, giving us an estimate of treatment on the treated. The authors managed to pull this off by carefully timing survey rounds with the program identification of beneficiaries. In some instances they are operating with small cells/blocks of observations and this approach is critical to boosting their power.
- The authors capture *program effects over time* – allowing them to trace the evolution of program impacts over time. This is rare – we often assume program effects to be a step wise effect: today you are treated and presto, you have program impacts and these will be the same forever (nb: just because we assume this doesn't mean we believe this – issues of budget, etc tend to force this). In this instance, the authors present a clear theoretical argument as to why some indicators may initially decline and then increase.
- As they line up successive impact variables, the authors *trace out the potential chain of effects* – starting with farming intensity to look at income, investment and, ultimately, household consumption. This gives them another slice at varying effects over time: we often don't know from the start how long a program will take to translate various changes/events into increased food on the table and this approach, combined with the time element above, allows us to potentially see what moves when.
- What's more, the authors also look at *potential displacement/substitution effects* – checking for a decline in crops the program did not focus on and comparing these to those crops which the program did focus on.
- This evaluation does a nice job at looking for *heterogenous program effects* – both across types of individuals (men versus women) and across the distribution of performance.
- This evaluation also presents estimates of the *internal rate of return*, thereby giving us some sense of the return on the money invested in this program – something often lacking in impact evaluations.

Given these careful and creative methods, what can we conclude about the impacts of the project?

- In terms of impacts on production (the first step in the causal chain), the program shows diverse impacts. There seems to be an early boost in planting, use of improved seeds, and prices for bean farmers, with a later drop off. The results for sesame are similar. However for cassava and livestock, there is no appreciable increase in these measures. On one level this is a matter of some concern: if these factors aren't changing, then we would expect little else to move. However, given the diversity of activities/crops in this project, these estimates, more so than others, are plagued by small sample sizes and noise.

- The impacts on income are some of the stronger results in this study. There are positive impacts on income, which seem to flatten out about 15 months after treatment.
- In terms of investment, the authors separate the results into mobile (e.g. tools and equipment) and fixed capital. Mobile capital shows a significant increase when the effects are examined across time. Fixed capital shows no significant increase.
- The program shows no significant impacts on consumption.
- The authors sum these different impacts up, with the conclusion that “the overall large confidence intervals that surround the results indicate that while we cannot reject the hypothesis that the program had no impact, we also cannot reject the hypothesis that it had very large impacts.”
- In terms of heterogeneity, the program appears to have had no significantly different impacts on women. While one result is significant at the 10% level, and the other coefficients are not small, this is likely a power issue as only 10% of the participants were female. However, when the authors look at effects across high and low performers, they find larger impacts for the higher performers, although the confidence intervals on these estimates are large.

So, was the project successful? It depends. In terms of income from targeted activities, there is indeed a significant and respectable rate of return. But the puzzle, as the authors point out, is why this did not translate into improvements in consumption – was it because of increased investment (for which there is a bit of evidence) or because we are lacking data on total income?

In addition, I had a couple of questions:

1. About 20% of business plans were cancelled before the enrollment period. Despite the information in footnote 12, it is still not totally clear to me why these were cancelled.
2. Given the initial plan to overlay the property regularization program that was later cancelled, it would be important to know if participants (differentiated across early and late groups) had expectations regarding improved property rights or not as this may have colored their investment decisions. (One could also think of a more complicated and perhaps less plausible story as well where if/when this cancellation was announced farmers in the early group, who were in fact getting something from “the program” responded differently from farmers in the late group).
3. In terms of benchmarking the poverty levels of the participants, I was wondering if the consumption module used in the evaluation survey was identical to that used in the national survey against which the comparison is made.
4. Would it be possible to use the month of the survey to look at possible expectation/early action effects (the nitty gritty of the timing of the interviews and the recruitment wasn’t totally clear to me).
5. Given potential issues of balancing in the baseline (at least for livestock groups) which could possibly disrupt the assumption of parallel trends, can you redo the analysis without the groups with possible balance issues as a robustness check?