Impact Evaluation Strategy for the MCA-Nicaragua Program

Michael R. Carter
University of Wisconsin-Madison

June 2007

Impact Evaluation Strategy for the MCA-Nicaragua Program

This brief document summarizes the impact evaluation strategy that has been developed for the MCA-Nicaragua land titling and rural business services program. This strategy is constantly evolving as the work progresses and hence the information presented here should only be understood as an approximate guide to the current state of the strategy. Up to the minute details may be obtained from program manager Anne Rothbaum (MCC/Washington) or from Michael Carter (University of Wisconsin).

1. Impact Evaluation Methodology

Basic Hypothesis: Both the Property Regularization Project and the Rural Business Development Project are hypothesized to increase incomes and asset values for individuals who benefit from these programs. Letting ***y*** indicate an outcome variable of interest (*e.g.*, family income, land value, etc.), the goal of the evaluation is to estimate the impact of a project treatment ***T*** (land titling only, business services only or both together) that is implemented after time period ***1***. This impact can be defined using the following difference-in-difference expression:

$$\boldsymbol{D^T} = \left[(y_2^T - y_1^T) - (y_2^C - y_1^C) \right]$$

where $\boldsymbol{D^T}$ is the project treatment impact, the subscripts indicate time (1 is the time before the treatment; 2 is the time after the treatment) and the superscript *C* indicates values for the counterfactual or control group. In words, the treatment effect is defined as the change in ***y*** (e.g., income) that an individual experiences following the treatment less the change in less the change in ***y*** that that the same individual (or an adequate control person) would have experienced over the same time period without the treatment.

As this expression makes clear, identification of the treatment effect requires observation of treatment and control groups both before and after the treatment. The actual econometric method that will be utilized to estimate $\boldsymbol{D^T}$ will be determined later, depending on the exact character of the control group and other considerations. Additional information (of the sort found in standard living standards measurement surveys) will be required for estimation.

Treatment Regimes and Detailed Hypotheses

A number of countries, including Nicaragua, have invested in land titling programs with the idea that land titles will promote broadly based growth. The evidence on these programs *in isolation* is mixed. By combining a land titling with a business services program, the Nicaragua program opens the door to understanding the impact of land titling both in isolation and in combination with business services. In

particular, the program will permit observation and evaluation of the following four treatment regimes:

	Without Business Services	*With Business Services*
Without land title	(i) [late/late]	(ii) [late/early]
With land title	(iii) [early/late]	(iv) [early/early]

Ideally, the goal would be to randomly allocate eligible program beneficiaries between these four treatment regimes. Comparison of treatment group (iv) with (control) group (i) using the difference-in-difference estimator sketch above would permit identification of the full impact of the Nicaraguan program. Comparison of groups (iv) and (iii) would permit us to see the additional value added to land titling efforts when they are combined with business services. The other pair-wise comparisons (iv with ii; and, ii with iii) would also provide valuable information on program effectiveness.

Control Group Strategy

The challenge of this and any other impact evaluation is to obtain an adequate control group. Because the benefits of the land titling program will be extended to everyone, a quasi-randomized program implementation strategy should make it possible to obtain adequate controls. Those receiving land titles early in the life of the program will serve as the treated (row 2 in the table above), while those receiving later in the program will serve as controls (row 1). PRODEP has already identified a geographic rollout strategy that will be adequate for this purpose. Unfortunately, this seemingly straightforward approach may have to be modified slightly at the analytical stage as roughly 40% of the households in the late treated areas already hold more or less clear titles. These earlier titles were not randomly distributed in some municipalities of Chinandega and emerged from a demand-driven framework.

Identification of a control group for business services is more challenging as this program is demand driven—that is, services have to be requested and will not be extended to everyone. However, parallel to the land titling component, business services will be rolled out at different times in different (quasi-randomly selected) areas of Leon and Chinandega. Households treated with business services (column 2 in the table above) will thus come from early treated communities, while those from late treated communities will form the controls. Since early and late status will be randomly determined, we will use this treatment status as an instrumental variable to estimate program participation in later econometric analysis. For the final analysis,

we should have the option of using continuous treatment estimation methods as the extent of treatment received by any households will depend on their randomly distributed early or late treatment status.

To reflect this basic design, the table above contains a dual early/late designation for each treatment cell. The first indicator refers to the timing of land titling program. The second indicates the timing for treatment with the business services program.

The next section lays out the sampling strategy for filling these different treatment cells and for obtaining reliable estimates of program impacts.

2. Beneficiary Selection and Implementation of Evaluation Methodology

This section will (1) Summarize the impact evaluation strategies about which we agreed in our discussions with MCA Nicaragua; and, (2) Provide a description about the sample frame and the number of clusters/*nuclei* that need to be identified and set aside for late treatment in 2009. In making designing the overall impact evaluation strategy, we have done our very best to impinge as little as possible on program implementation. While we think the proposal outlined below should not cause problems for implementation, we are of course open to further discussion about the specifics of the design.

Logic of the Impact Evaluation Strategy

The impact evaluation strategy relies on a partially randomized rollout strategy—that is, on a comparison of 'early' and 'late' treatment groups. Early treatment communities will begin to participate in ONR programs in 2007. Late treatment communities will begin to participate only after 2009. A random sample of the eligible producers in both groups will be asked to complete a single survey (akin to a standard living standards measurement survey) in 2007. The survey will be designed to take no more than 1-2 hours of the respondents' time. All survey respondents will be revisited by the impact evaluation team in 2009 and again at the end of the MCA project in 2011. No additional interaction with the respondents will be necessary in terms of the impact evaluation study.

Description of the ONR Program

As we understand it, the ONR program will select a group of communities or *nuclei* for each product or *rubro*. Within each community, it will be possible to identify a group of "eligible producers." Eligible producers will be those who produce the relevant product (or who have the capacity to produce the product) and whose resources are neither too small nor too large to make them ineligible for the program. We further understand that ONR will select a single "model farmer" who will be the nuclei for producers within the community. All eligible farmers will be invited to join with the model farmer and

become a participant in the ONR program in their community. We label that fraction of eligible farmers who join the ONR as "participant farmers."

Choosing Clusters for Early and later Roll-out of the ONR Program

In order for the proposed impact evaluation strategy to work, we need to follow the steps listed here:

- ONR will identify by 30 May 2007 a set of *selected communities* where the ONR program will be rolled out in either 2007 or 2009. The exact number of communities needed is specified below.

- The impact evaluation team will randomly divide the set of communities into two groups: the early treatment group (that will receive ONR services in 2007) and a late treatment group that will receive ONR services in 2009. For details, see appendix 2 below.

- ONR will follow the implement its program in accordance with this division. Note that ONR will be free to implement its programs in non-selected communities according to whatever schedule it desires.

- We recognize that in a few cases, ONR may choose NOT to ultimately implement a program in a selected community (for example, the individual who was thought to be a good model farmer, may have moved away and there is no adequate replacement for her or him). It will be good for the impact evaluation analysis if this happens in a *few* communities. However, if it is likely to happen in more than 10% of the selected communities, then additional selected communities will have to be specified by July 2007.

Choosing Households to Survey within Clusters

Selection of households for inclusion in the survey will follow these steps:

1. A visit must be paid to each nucleus. Several local leaders or knowledgeable people (including the nucleus lead farmer if one has been identified) need to be interviewed and asked to identify the full universe of farmers in the nucleus area who meet the eligibility criterion (land size, etc.). Note that this listing does not have to be precise and it is better to include farmers who might be eligible than to exclude them. Also note that being included on the list does not mean that the farmer will or must receive technical assistance. It simply means that the farmer is eligible under program rules to receive assistance if she or he so desires. The study design assumes that only a fraction of the eligible farmers will actually

receive treatment. The resulting listing of eligible farmers (perhaps 20-30 in each nucleus community) will form the sample frame for the survey.

2. Once the sample frame has been formed for each nucleus, a random sample of farmers will be selected from the list. These are the farmers who will be subsequently interviewed for the impact evaluation study (both at baseline in 2007 and in two subsequent years). We will need to discuss in more detail the exact number of farmers, but we assume that approximately 50% of listed farmers will be chosen for interview. These farmers will become our sample farmers.

Again note that not all sample farmers will actually receive technical and other services. It is very important that an identical procedure is followed in both early and late treatment communities. That is, in nuclei selected for early treatment, we cannot use only farmers in the ONR program as our sample farmers. Sample farmers in both early and late treatment communities must be randomly drawn from the set of eligible farmers.

Assumptions used in the Statistical Power Calculations

Based on information from the ONR team, we have assumed the following in calculating the number of selected communities that need to be identified in May 2007:

- Averaged Number of eligible producers per-nucleus: 30;
- Average Number of participants per nucleus: 10-15 (or ~33% of eligible producers—this figure is called the "compliance rate");
- Average Number of *early* participants in the late treatment ("control") communities: 0.

Proposed Strategy

The impact evaluation will focus on the 5 rubros listed below. The figure in square brackets after each rubro is the approximate number of nuclei with which the ONR will work over the 5 year life of the MCA.

- Ganaderia [100]
- Ajonjolí [45]
- Frijol [39]
- Yuca [20]
- Hortalizas [6]

Across these 5 rubros, ONR thus proposes to work with a total of 210 communities.

For the impact evaluation study, we proposed that ONR identify roughly 50% (or 104) of these 210 communities for either early or late treatment. As mentioned above, we would randomly allocate these 104 communities into early (26 communities) and late (26 communities) treatment groups. For purposes of the impact evaluation study, these 104 communities need to be divided evenly across the 5 rubros (e.g., 50 ganaderia

communities, 22-23 ajonjoli, etc.). The other 106 communities can be treated according to whatever timetable is convenient for ONR.

Therefore, ONR should roughly identify these 104 communities finding a farmer candidate to be a leader. In practice, and after updating information about the characteristics of farmers, it is possible that some of these communities will never be formed, especially in communities with late treatment. For that reason, ONR has tried to identify more than 104 communities. Up to date, they are determined 142 communities and they are still working on the identification of hortalizas nuclei.

This is the current situation of potential nucleus or community distribution. The lower number of potential nuclei in the row "*with land title*" is due to the low number of municipalities that are receiving early treatment under the land program.

	Without Business Services	*With Business Services*
Without land title	39 [late/late]	39 [late/early]
With land title	30 [early/late]	34 [early/early]

Within each selected community, the impact evaluation team will form a census of eligible producers using criteria provided by ONR for the particular rubro. A random sample of ~12 producers will be drawn from this list. It is these producers who will interviewed in 2007 (and later again in 2009 and 2011) with a living standards survey.

Our analysis of this sample design indicates that it should give us sufficient statistical power to identify program effects that increase household income by at least 14% We feel that this is an adequate 'minimum detectable program effect.' That is, we all hope that the program will have a much greater effect than 14%. If the effect is less than 14%, the survey may not be able to detect it (i.e., statistical results may be insignificant). Greater precision could be had by increasing the number of communities or by increasing the number of eligible producers interviewed in each community. (The latter is less effective if there is a high income correlation within communities.) If the loss rate of nuclei is close to zero, then we could expect to have around 140 formed nuclei. In this case, the effective sample size could be higher and it could possible to identify an increase in household incomes even of 12%.

Details on these statistical power calculations are given in Appendix 1 below.

Contingencies and Concerns

As mentioned above, we have tired hard to design an impact evaluation strategy that can both give reliable estimate of program impact while imposing very few constraints on

program implementation. We can, however, imagine several sorts of difficulties that may make the proposed strategy difficult to implement:

- Program implementation may fall behind schedule and early treatment communities may not begin to benefit from ONR programs until after 2007. If this happens, the impact evaluation team will have the option of delaying the second round surveys. The important thing is that the time lag between early and late treatment be long enough for the effects of the ONR to become visible.

- As mentioned above, the proposed strategy can handle some slippage since ONR has tried to identify additional leader candidates. (e.g., it is not a problem if only 70% of the selected communities ultimately prove to be worthy for an ONR program).

- There are 69 candidates to be a leader. Their nucleus will be formed and treated in 2009 by the ONR. In order to have an approximately even number of early and late communities, the ONR will try to find a new leader if it is necessary. The power of our sample design could be lower if there are too many early treated nuclei than late treated nuclei.

- In some cases it could be possible that a nucleus can not be formed if there are not a sufficient number of eligible farmers. FIDEG will start an identification of these eligible farmers in order to have a better approximation of the effective number of nuclei or clusters that should be considered for getting an effective sample size. Up to date we could say that if there are **148** potential nuclei/clusters with **12** farmers interviewed per cluster, then <u>the maximum sample size</u> will be **1776** households.

- Alternatively, we could, to some extent, work with a slightly smaller number of communities and increase the number of producers surveyed per-community. However, this will increase impact evaluation costs and we will need to reanalyze the power of our test before we can agree to a smaller number of communities.

3. Disaggregating Program Impacts by Gender, Age and Income

The basic household questionnaire is being designed around the living standards measurement survey instrument currently employed by the National Institute of Statistics (INEC). This survey includes full demographic information on each respondent household, making it possible to econometrically disentangle program impacts by age and gender. Similarly, the survey instrument will also contain a full suite of wealth and income measures so that again program impacts can be calculated separately by economic status.

As detailed above, surveyed households are selected from producer households that meet the eligibility requirements of the ONR program. To the extent that eligible households are preponderantly of one age or gender, the impact evaluation will be less able to precisely disaggregate program impacts. The sampling strategy will however give us a clear picture of the eligibility rules employed by ONR.

4. Timing and Reporting on Implementation Plan

The University of Wisconsin-based impact evaluation team reports constantly to both MCA-Nicaragua and to the MCC project team in Washington. Implementation Plan reports issued to date include:

- *Impact Evaluation Strategy (October 2006)*
- *Two-stage Sample Selection Strategy (February 2007)*
- *Sample Size and Power Calculations (March 2007)*
- *Draft Land Module (April 2007)*
- *Municipality Selection Scheme (April 2007)*
- *Division of Municipalities into Early and Late Treatment Groups (May-June 2007)*

Future reports in the near future will include:

- *Household Selection Procedures (June 2007)*
- *Draft Household Questionnaire (June 2007)*
- *Draft Community Questionnaire (July 2007)*

The next stage of activities will include pre-testing of the questionnaire, final revisions of the questionnaire, enumerator training, survey initiation and survey supervision. Progress reports on these stages will be submitted as the work takes place (roughly September 2007).

5. Data Collection Plan

The first or baseline survey round will take place as soon as possible after the Rural Business Office is able to finalize the geographic program rollout strategy and as soon as data collection and data quality services are contracted. At this writing, it appears that the first survey round will be implemented in September 2007. However, the exact timing of that sample will depend on the timing and speed of the actual program rollout.

The second survey round will take place approximately two years after the baseline (mid-2009). The exact timing will need to be coordinated with the implementation plan of the rural business office. The idea is to have the second round surveys take place before business and titling programs are extended to the ‘late’ areas. Analysis

of the second round data will permit identification of program effects.

Finally, a third survey round of data will be taken during the final year of the program (mid-2011). By this time, households located in control (later treatment areas) should have been treated. This will open the door to 'continuous treatment' methods in which variation in the extent of treatments (e.g., months with title; months with business services) can be used to identify program effect. This method (which requires that the extent of treatment is randomly determined) will permit a more extensive look at the dynamic effects of the Nicaraguan program. This should be especially important in terms of understanding longer term investment effects in both productive assets as well as human capital assets (e.g., children's education).

Appendix 1: Power Calculations

The sample design that we proposed has two levels of randomization. In a first level, we select randomly the clusters or nuclei that will participate in early and late treatment. In a second level, we will select farmers from each of these clusters or nuclei.

As we mentioned before, we will compute the difference-in-difference estimator to evaluate the impact of programs in the farmer income. For the impact of programs be statistically identified, we have to take into account the sample size considering the sample frame. Such a sample frame consists of clusters of farmers in which a proportion of them will participate in the program. Then, the sample design will determine the characteristics of the sample to be interviewed and, consequently, the statistical properties of our estimator in terms of bias of its variance. Note that with a high variance of the difference-in-difference estimator we could conclude wrongly that the program did not have any effect on incomes. In contrast, a lower variance of this estimator could provide a more precise estimation of the program impact.

Since we need the lowest possible variance of the estimator, we have to determine some parameters that define the quality of our sample design in terms of the dimension of this variance. Some of these parameters are semi controllable in practice (e.g. early and late treated groups or sample size), but others are the result of the sample design that we are using[1]. Then, we would like to find a combination of these parameters such that our sample design can identify the real program impact. For example, if we have a sample design that only can identify an increase in income higher than 30%, this would indicate a poor sample design. Why? Because if the real effect of the program on incomes were 20%, the statistical analysis would conclude that there was no effect. We call minimum detectable effect (MDE) to the lowest level of an increase in income that can be statistically detected given a sample design. Thus, a poor sample design will imply a high variance of the estimator such that the MDE could be so much high that we could obtain wrong conclusions about the program success.

The first parameter to consider in the variance of our estimator is the intra class correlation that could exist within every cluster, that is, the income correlation between farmers that belong to a same cluster. In statistical terms, a high intra class correlation will increase the variance of the difference-in-difference estimator with respect to a sample design in which individuals have been selected randomly. This higher variance is known as *design effect*. In an extreme example, imagine that every farmer within a cluster or nucleus is identical to each other. We could interview 10 farmers in this cluster but since they are perfectly equal the information provided by these 10 farmers would be equivalent to the information given by just one of them. Therefore, the higher the intra class correlation in income, the lower the gain of increasing the number of interviewed farmers within a cluster will be. With a high intra class correlation the best option is to

[1] For example, our sample design is based on the existence of clusters of farmers, because this is the way that ONR will implement the program.

increase the number of clusters, because we can learn more about the average effect of the program knowing farmers with different experiences in terms of income[2].

There are other two parameters that could increase the variance: the compliance rate (c-s), and the proportion of the sample that will be early treated (P).

The compliance rate is the difference between (i) the proportion of eligible farmers that belong to the early treated clusters that will participate in the program in 2007 (c), and (ii) the proportion of eligible farmers that belong to the late treated clusters but that are being treated wrongly in 2007 (s). Ideally, a perfect compliance rate (i.e. c=1 and s=0) would have no effect on the variance. However, if c-s is lower than 1, the variance of the difference-in-difference estimator will increase. We expect that around one third of the eligible farmers that belong to the early treated clusters will participate in the program (c= 1/3) and that no eligible farmers that belong to the late treatment be treated in 2007 (s=0). This entails a compliance rate around one third (c-s=1/3).

In terms of the proportion P, the variance of the difference-in-difference estimator is minimized when P=1/2. This means that there is an even number of both early and late clusters. In practice, it could be tempting to treat as many farmers in 2007 as possible, leaving a lower amount of farmers to be treated in 2009 (P>1/2). However, in this context, the variance and, consequently, the MDE will be higher.

Finally, we have to also decide how much *power* we want to set with this sample design. In our context, if the program impact is highly successful, the power can be defined as the probability that any statistical test concludes that the hypothesis of zero or small effect is not true. We would like to set a high power to our tests such that the MDE decreases[3].

In the case of the business program, it is feasible to think that the program will be successful if the average farmers' incomes increase, in our view, around a minimum of 10%. Therefore, a useful sample size has to consider all these issues in order to guarantee a MDE around 10%.

To construct an estimator of the degree of correlation that we could find within clusters, we use data of the department of Leon and Chinandega from Carter and Chamorro (1999) database. To compute the intra class correlation we assumed that clusters were the municipalities of these two departments. The intra class correlation using total income is 0.03 approximately.

The following table shows both the level of MDE (i.e. the change in household income) and the MDE measured as a percentage of average household income. We have estimated the MDE that could be statistically detected if we finally get 104 nuclei with 12 farmers interviewed per cluster (n =1248). In an optimistic scenario with 140 effective nuclei (i.e.

[2] However, it is also true that more observations per cluster could allow estimating the intra class correlation with more precision.

[3] However, too much power would increase the size of the statistical test. That is, if the program does have a negligible effect, we could conclude wrongly that the program had a larger effect on incomes.

with almost zero loss of nuclei), the effective sample size could reach to 1680 households. Note that for this estimation we have used the estimated intra class correlation from the data mentioned above (0.03).

In this simplified exercise, we have assumed that a larger sample size does not improve the estimation of the residual variance. However, with a low intra class correlation we should expect that additional observations decrease the residual variance. This means that the MDE could be even lower than the MDE that we are reporting if the sample size increases.

Assumptions used to Estimate the Minimum Detectable Effect (MDE) Reported in Text

- 52 clusters with early treatment and 52 clusters with late treatment
- Intra class correlation = 0.026*
- Average household income = 21,347 córdobas*
- S.D. household income (in córdobas) = 32,211*
- Percentage of the income variance that it is not explained by determinants of income (residual variance) = 50%
- Compliance rate in 2007 (c-s) =1/3
- Level of significance= 5%

	Number of clusters=104 Surveys per cluster=12		**Number of clusters=140 Surveys per cluster=12**	
Power	**MDE**	**% growth of household income**	**MDE**	**% growth of household income**
0.9	2976	14%	2565	12%
0.8	4906	23%	4228	20%
0.7	6297	30%	5427	25%

* Estimators were computed using data of Leon and Chinandega used in Carter and Chamorro (1999). MDE is expressed in córdobas per household.

The following table has as main purpose to evaluate the impact that could have a higher intra class correlation in the MDE using different sample sizes. We show that with a higher intra class correlation (0.1), the MDE could reach up to 18% if the effective number of clusters is 104. In this context, it could be useful to have a sample size with more nuclei.

MDE Calculations using Alternative Assumptions

	Intra class correlation =0.026					
Number of clusters	**104**			**140**		
Surveys per cluster	**MDE**	**% growth of household income**	**Sample size**	**MDE**	**% growth of household income**	**Sample size**
10	3193	15.0%	1040	2752	12.9%	1400
12	**2976**	**13.9%**	**1248**	**2565**	**12.1%**	**1680**
15	2741	12.8%	1560	2363	11.1%	2100

	Intra class correlation =0.1			
Number of clusters	**104**		**140**	
Surveys per cluster	**MDE**	**% growth of household income**	**MDE**	**% growth of household income**
10	3963	18.6%	3416	16.0%
12	3803	17.8%	3278	15.4%
15	3637	17.0%	3134	14.7%

Appendix 2. Municipalities used by ONR in the Identification of Potential Nuclei

Rubro	**Early Titling**			**Late Titling**		
	Eligible Municipalities	*ONR has identified:*	*M&E allocated randomly these potential nuclei*	*Eligible Municipalities*	*ONR has identified:*	*M&E allocated randomly these potential nuclei*
Ajonjolí	La Paz Centr [Chinandega] [León]	→14 nuclei→	8 early ONR 6 late ONR	Larreinaga Somotillo[2008] El Sauce[??] Posoltega[??] El Viejo[?]	→ 16 nuclei→	8 early ONR 8 late ONR
Yuca	[León]	→10 nuclei→	5 early ONR 5 late ONR	Telica Larreynaga	→7 nuclei→	4 early ONR 3 late ONR
Ganadería	Villanueva Nagarote La Paz Centro [León] [Chinandega]	→26 nuclei→	14 early ONR 12 late ONR	Larreinaga [León] El Sauce[??] Achuapa[??] El Viejo[?]	→30 nuclei→	14 early ONR 16 late ONR
Frijol	San Pedro del Norte Santo Tomás del Norte	→14 nuclei→	7 early ONR 7 late ONR	Santa Rosa Cinco Pino[2008] San Frcisco del Norte[2008] Achuapa[??]	→25 nuclei→	13 early ONR 12 late ONR
Hortalizas	[León] [Chinandega]	→3 nuclei→	1 early ONR 2 late ONR	Telica Quezalguaque Chichigalpa[2008] El Sauce[??] El Viejo[?]	→2 nuclei→	1 early ONR 2 late ONR

[ ] Only some communities of these municipalities could be treated in 2007 by the titling program.

[2008] The "conflict solution" stage in the titling program will start in 2008.

[??] This municipality could be treated after 2009 by the titling program

[?] Some communities of these municipalities could not be treated by the titling program.