
**Revised Impact Evaluation Strategy for the MCA-Nicaragua
Rural Business Development Program**

Michael R. Carter and Patricia Toledo

March 2010



Impact Evaluation Strategy for MCA-Nicaragua Rural Business Program

In 2005, the Nicaraguan government, in cooperation with the Millennium Challenge Corporation (MCC), devised a rural economic growth and poverty reduction program for the high potential Pacific coast departments of León and Chinandega. Three components constituted the program: Construction and improvement of rural trunk and feeder roads; Provision of legally secure land titles to both urban and rural properties; and, Rural business development services.

The BASIS Collaborative Research Support Program was contracted to conduct a rigorous impact evaluation of this third component, the rural business development (RBD) project, which provides technical and other services to rural producers as a way to facilitate their transition to higher-income activities.

This brief document summarizes the impact evaluation strategy that has been developed as well as a description of the sample design, questionnaire design, data editing, and the content of the data file. The impact evaluation 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.

1. Impact Evaluation Methodology

The RBD project is hypothesized to increase incomes and asset values for individuals who benefit from rural business development services and technical and financial assistance activities. 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 the RBP treatment T that is implemented after time period **1**. This impact can be defined using the following difference-in-difference expression:

$$D^T = [(y_2^T - y_1^T) - (y_2^C - y_1^C)]$$

Where

D^T	= the impact of the program
y	= an outcome variable of interest (<i>e.g.</i> , household income, asset value, etc.)
subscripts 1 and 2	= time (1 is the time before the treatment; 2 is the time after the treatment)
superscript C	= values for the counterfactual or control group.
superscript T	= values for beneficiaries of the project or treatment group.



In words, the treatment effect is defined as the difference between the change in y (e.g., income) that an individual experiences following the treatment and the change in y 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. In particular, it was expected that early treatment farmers started their participation at the end of 2007, while late treatment clusters services were not initiated approximately until 18 months later, in early 2009.

Defining treatment and control groups 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. The actual impact evaluation strategy relies on a randomized rollout strategy, that is, early and late status will be randomly determined. Thus, we will be able to use this treatment status as an instrumental variable to estimate program participation in later econometric analysis. The actual econometric methods used to estimate D^T will depend 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.

Regarding the unit of randomization, the RBD project was implemented by “clusters” of producers engaged in the same type of productive activity and in geographic proximity. Based on this, it was logic to randomized “potential clusters” to early and late groups. Therefore, randomization required a previous identification of these potential clusters, which ultimately form the sample. Households belonging to this sample should be invited to participate in the impact study, and completed a baseline survey in late 2007, just as the RBD project was beginning in the early treatment clusters.

2. Sampling Strategy

Ideally, all potential clusters should be identified to select a random sample. However, in practice, that goal is infeasible due to budget and time constraints. The sample strategy was based on the selection of a set of potential clusters in areas where there were high densities of eligible producers. These areas were grouped by type of productive activity. This evaluation team jointly with RBD professionals developed a field work that allowed forming a list of approximately 3000 farmers distributed in 140 potential clusters whose eligibility was



determined by a set of criteria defined by the RBP (see Table 1). Most clusters were formed by 24 eligible farmers.

By considering the evaluation methodology previously described and using data from farmers of Chinandega and Leon, used in a previous study (Carter and Chamorro 2000), we evaluate how different sample sizes could affect the accuracy of the estimated program impact. Based on this analysis, it was desirable to include in the sample at least 120 clusters and, within every cluster, 12 farmers should be randomly selected to be interviewed.

Table 1. Specific criteria used to identify eligibility of a farmer in Agricultural and Livestock groups:

	SESAME	BEAN	VEGETABLES	CASSAVA	LIVESTOCK
1	Farmer had to plant in the past at least:				Farmer owns between 10 and 100 cows in age to produce milk.
	2 blocks with sesame	1 block with bean	any size with vegetables	2 blocks with cassava	
2	Minimum area of the farm with soil suitable for agricultural crops:				Livestock activity is developed in the farm. No minimum farm size.
	10 Blocks	5 blocks	2 blocks	5 blocks	
3	If there is an irrigation system, minimum area of the farm with soil suitable for agricultural crops is 625 square varas				Water access is located inside the farm.
4	Maximum area of the farm with soil suitable for agricultural crops:				There are permanent available roads to access the farm in any season of the year.
	50 blocks	50 blocks	20 blocks	100 blocks	
5	He is in possession or has a title of the farm				
6	Farmer is a least 20 years old				
7	Farm is located out of the "protected natural areas"				

The baseline survey round started as soon as the identification of eligible farmers finished (November 2007). The second survey round took place in early 2009, just before the participation of the late group started. Finally, a third survey round of data will be done during the final year of the program (mid-2011). By that time, households located in control group (later treatment areas) should have been treated. This will open the door to 'continuous treatment' methods in which variation in the extent of treatment (i.e., months with business services) can be used to identify program effects. This method 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 productive assets as well as human capital assets (e.g., children's education).



Table 2 shows the sample of potential clusters by type of productive activity as well as the distribution of clusters in early and late groups. The sample included 140 clusters distributed almost evenly in each treatment group. Even though in most of the clusters the list included 24 or more eligible farmers, in some clusters the list was closely formed by 12 eligible farmers. In those cases, every farmer was selected to be interviewed. This fact jointly with some farmers that rejected to be interviewed, made the number of surveys per cluster slightly lower than 12.

Table 2. Distribution of the Sample

Productive Activity	Selected for Early Treatment	Selected for Later Treatment	Interviewed households per cluster
Vegetables	2	2	12.0
Cassava	7	8	11.5
Sesame	14	15	10.2
Bean	20	19	11.7
Livestock	28	26	11.8
Total	71	69	11.4

3. Data Editing

The *Fundación Internacional para el Desafío Económico Global* (FIDEG) was the firm in charge of the data collection. Data were entered using the Census and Survey Processing System software (CSPRO 4.0). Additionally, the data gathering firm developed procedures to identify mistyped data. This preliminary version was edited by using SPSS statistical analysis software and provided to the evaluation team to do a deeper data checking. To date, the evaluation team has reviewed the data for completeness and internal consistency, and to determine if farmer's household follow-up was correctly done. There were 21 households in which the second interview failed because of rejection (12 households) or because it was not possible to find an adequate household member to answer the questionnaire. Follow-up of the farm -which information is concentrated in Section 6 of the questionnaire- is still under revision. However, preliminary results show that errors, such as interviewers' failure to follow procedures, should not have a significant effect on data quality. Most importantly, land tenure answers could be imprecise because most of the interviewed farmers do not have a formal possession of the farm.

Regarding the variables used to compute the aggregate expenditures, the evaluation team did the following task in the cleaning process:



B A S I S

- 1) Identification of mistyped data by finding extreme values of per capita durable and non durable aggregate expenditures growth.
- 2) Revision of every missing value to verify if it was a mistyped data.
- 3) Consistency between section 3.C, 3.CA and 3.CO to verify if there was information that was not typed.

In most cases, it was identified that the enumerator wrote an incorrect code. However, enumerators were encouraged to write observations if they had some doubt about the farmer's answer. This type of information was key for the cleaning data process.

In other cases, wrong codes of frequency or total value were evident but there was not additional information from the enumerator (e.g., a household consumes 50 pounds of sugar per day). By comparing this information with the other round survey and considering that the size of household had not changed, we concluded that household consumption was the same amount of food but the frequency or the value was not coherent.

Finally, if there was a household with only one missing value in only one round of the survey, we impute a value for this unique missing value. For example, if the missing value was a food value, we take the average of the value of the same food declared by other households living in the same municipality.

4. Questionnaire and Data Files

The survey is composed of seven sections designed to collect information about expenditures and assets of the farmer's household as well as information about his/her farm, agricultural practices, marketing, and prices for their product.

4.1 Aggregate expenditures:

Regarding the modules used to compute aggregate expenditures, the questionnaire has almost the same questions employed by the INIDE living standards measurement survey (LSMS) used to gauge poverty rates in the country as a whole. The next sections follow closely the LSMS in order to estimate comparable aggregate expenditures. Within every section the following expenditures were collected:

- Section 1. Characteristics of the House: housing service expenditures.
- Section 2.2. Educational expenditures per household member.
- Section 3C. Home equipment: estimation of the use of durable goods.
- Section 4. Food and other Non Food Expenditures.

Monthly expenditures per capita can be found in the file “Aggregate Expenditures.sav”. The aggregation of expenditures follows the same steps described in the document “Metodología de Construcción del Agregado de Consumo, de las Líneas de Pobreza y del Agregado de Ingreso en Nicaragua”.¹

Aggregated consumption was not calculated for 42 households because there were more than one missing values for that household or because household could not be interviewed in the second round.

3.2. Data files:

Data files were grouped by section and/or subsections. The last two characters of the name files indicate if data correspond to the baseline (R1) or to the second round survey (R2). Every file contains the household ID named “Formulario”.

Every variable in any file has a brief description. Additionally, labels have been also included in the questionnaire in order to have a better identification of which question represents every variable. Name of the files, as well as the information of every file is described in the next table.

Table 3. Data Files

File name	Section name in the questionnaire	Description
Aggregate Expenditures.sav		Monthly per-capita expenditures and number of household members for both survey rounds. Include cluster ID, instrumental variable, and participant’s indicator variable
Cover_R1.sav Cover_R2.sav	Section 0	Codes of departments and municipalities. Result of the survey (completed/uncompleted)
Section I_R1.sav Section I_R1.sav	Section 1	Housing characteristics and services expenditures
Section II_R1.sav Section II_R2.sav	Section 2.1 Section 2.1N (only second round survey)	Demographics characteristics of household members. In the second round of the survey, new members are also included
	Section 2.2 Section 2.2N (only second round survey)	Educational Expenditures per household member. In the second round of the survey, new members are also included
	Sections 2.3 and 2.3N	Distribution of farm animals in the

¹The specific document used by the authors can be downloaded from http://www.inide.gob.ni/pobreza/publicacion/Constr_agring.pdf. INIDE’s SPSS code can be also found in <http://www.inide.gob.ni/pobreza/pubmetodol.htm>



	(only second round survey)	household members. In the first round of the survey, questions of this module were in Section 2.1
Section IIIA_R1.sav Section IIIA_R2.sav	Section 3A	Farm mobile equipment
Section IIIAO_R1.sav Section IIIAO_R2.sav	Section 3AO (only second round survey)	This section collects the farm mobile equipment that was omitted at the time of the first round survey
Section IIIB_R1.sav	Section 3B	Farm animals (total). This module was not included in the second round of the survey
Section IIIC_R1.sav Section IIIC_R2.sav	Section 3C	Home equipment
Section IIICA_R2.sav	Section 3C , “Cuadro Adicional”	This section collects new home equipment or changes in the stock of home equipment in the second interview, respect to the first interview
Section IIICO_R2.sav	Section 3CO	Register the home equipment that was omitted at the time of the first round survey
Section IV1_R1.sav Section IV1_R2.sav	Section 4.1	Food expenditures in the last 15 days before the interview. Include self-consumption, gifts, and donations.
Section IV2A&E_R1.sav Section IV2A&E_R2.sav	Section 4.2A Section 4.2E	No food expenditures in the last week before the interview. Health expenditures.
Section IV2B_R1.sav Section IV2B_R2.sav	Section 4.2B	No food expenditures in the last month before the interview.
Section IV2C_R1.sav Section IV2C_R2.sav	Section 4.2C	No food expenditures in the last 6 months before the interview.
Section IV2D_R1.sav Section IV2D_R2.sav	Section 4.2D	No food expenditures in the last 12 months before the interview.
Section VA_R1.sav Section VA_R2.sav	Section 5A	Work-unrelated incomes
Section VB_R1.sav Section VB_R2.sav	Section 5B	Received remittances
Section VC_R1.sav Section VC_R2.sav	Section 5C	Sent remittances
Section VII1A_R1.sav Section VII1A_R2.sav	Section 7.1A	Agricultural practices, marketing, and prices for bean, sesame and cassava crops.



Section VII1B_R1.sav Section VII1B_R2.sav	Section 7.1B	Agricultural practices, marketing, and prices for vegetable crops.
Section VII1C_R1.sav Section VII1C_R2.sav	Section 7.1C	Field crops problems (bean, sesame, a cassava and vegetables)
Section VII1D_R1.sav Section VII1D_R2.sav	Section 7.1D	Agricultural practices in other annual crops
Section VII1E_R1.sav Section VII1E_R2.sav	Section 7.1E	Agricultural practices, marketing, and prices for permanent or semi-permanent crops
Section VII1F_R1.sav Section VII1F_R2.sav	Section 7.1F	Livestock practices, marketing, and prices
Section VII1G_R1.sav Section VII1G_R2.sav	Section 7.1G	Production practices, marketing, and prices for milk
Section VII1H_R1.sav Section VII1H_R2.sav	Section 7.1H	Production practices, marketing, and prices for cheese, honey, handicraft and brick
Section VII2A_R1.sav Section VII2A_R2.sav	Section 7.2A	Temporary workers hired for the agricultural season. Include unpaid workers
Section VII2B_R1.sav Section VII2B_R2.sav	Section 7.2B	Temporary workers hired for the livestock season. Include unpaid workers
Section VII2C_R1.sav Section VII2C_R2.sav	Section 7.2C	Permanent hired workers
Section VII3_R1.sav Section VII3_R2.sav	Section 7.3	Technical assistance
Section VII4A_R1.sav Section VII4A_R2.sav	Section 7.4A	Characteristics of the farmer's credit
Section VII4BA_R1.sav Section VII4BA_R2.sav	Section 7.4B	Perception of farmer credit constraint
NA	Section 6.	Not available data. Variables under data cleaning