Are Natural Resource Management Programs Beneficial? Evidence from the POSAF-II case in Nicaragua

De los Santos M., Luis A.[1], Bravo-Ureta, Boris E.[2], Cramon-Taubadel, Stephan[3]

[1] Department of Agricultural Economics and Rural Development, Georg-August-Universität, Germany, Department of Agricultural Economics, University of Talca, Chile, and Dominican Institute of Agricultural and Forestry Research (IDIAF), Dominican Republic. E-mail: ldelossantos@gmail.com

[2] Department of Agricultural and Resource Economics, University of Connecticut, Storrs, USA Department of Agricultural Economics, University of Talca, Talca, Chile.

[3] Department of Agricultural Economics and Rural Development, Georg-August-Universität, Platz der Göttinger Sieben 5, D-37073 Germany.

Contribution presented at the XV EAAE Congress, "Towards Sustainable Agri-food Systems: Balancing Between Markets and Society"

August 29th – September 1st, 2017

Parma, Italy

Are Natural Resource Management Programs Beneficial?
Evidence from the *POSAF-II* case in Nicaragua

Abstract: Understanding the impact of programs designed to improve the management of natural resources in agricultural households is a key task to ensure the adoption of sustainable and profitable practices. In this paper, we analyze the economic impact of natural resource technologies delivered during the implementation of *POSAF-II* in Nicaragua. Results obtained from propensity score matching (PSM), ordinary least squares (OLS), weighted least squares regression (WLS) based on PSM, and instrumental variables (IV) regression indicate that *POSAF-II* has had a positive impact on the total value of agricultural production of beneficiary farmers.

Keywords: natural resource management, impact evaluation, intention-to-treat, spillover, Nicaragua

1. Introduction

Agricultural production worldwide has managed to effectively meet global demand for food and fiber (World Bank, 2008). However, the ongoing rise in food demand stemming from population and income growth along with the uncertainty from climate change is expected to increase pressure on the agricultural system around the globe (Gornall et al., 2010; World Bank, 2016). These challenges pose a significant threat to about 1.2 billion people worldwide that are living below the poverty line, and 70% of this population lives in rural areas. A significant number of these people earn their income directly from agricultural activities or have some reliance on farming for their livelihoods (Cleaver, 2012).

The challenges facing agriculture make it imperative to harmonize the need to promote the sustainable use of natural resources with the choice of policies that can be effective in reducing poverty. In this context, generating compelling evidence regarding the effects of changes in agricultural practices on farmers' incomes has become an important issue for policy makers and donors (Khandker, Koolwal, & Samad, 2010; Kelley, Ryan, & Gregersen, 2008). One of the reasons for assessing this impact is to build accountability in public administration and to guide policy decisions. Along with these reasons, determining what works and why impacts are reached or not reached are additional justifications for producing the "proof" that validates public actions (Gertler, Martinez, Premand, Rawlings, & Vermeersch, 2011).

A number of natural resource management (NRM) programs, designed to simultaneously reduce poverty, increase productivity and protect natural resources, have been implemented in Latin America and elsewhere (Cocchi & Bravo-Ureta, 2007; Dalton, Lilja, Johnson, & Howeler, 2005; District & Kingdom, 2011; Dutilly-diane, Sadoulet, & de Janvry, 2003; Tsiboe, Dixon, Nalley, Popp, & Luckstead, 2016). Moreover, in many cases, NRM technologies have been evaluated in controlled experimental or on-farm technology trials, but these types of data are not useful to evaluate actual farming conditions where many variables are beyond the control of the producer (Del Carpio & Maredia, 2011; Kelley et al., 2008; Pal, 2011; Renkow & Byerlee, 2010). As a result, productivity gains measured under controlled conditions are likely to overestimate the real impact of NRM technologies. Therefore, the performance of specific technologies under real farming conditions needs to be better understood so that robust interventions can be formulated and implemented (Renkow & Byerlee, 2010; Harwood, Kassam, Gregersen, & Fereres, 2005).

Although there is a sizable amount of research illustrating the impact of agricultural growth on poverty reduction, the literature on the impact of different interventions, including technology transfer, is rather limited. Recently, more impact evaluations have been completed, although relatively few have been considered rigorous. A meta-analysis by Del Carpio and Maredia (2011) examined 286 evaluation projects. Among these, only 86 projects met the requirements to be considered in their analysis, referring particularly to the use of a counterfactual group to measure the effects of the interventions. A clear shortcoming of the subgroup of 86 studies is that only three of them examined the spillover effects of the program. Moreover, of the 86 just 12 focused on NRM and a handful conducted a benefit-cost analysis

Latin American countries, like those in other parts of the developing world, have implemented NRM programs delivering sustainable technologies with the intention of reducing rural poverty while improving productivity and protecting natural resources. However, the literature contains limited reliable evidence concerning the effects of these interventions on farmers' incomes. The general objective of this study is to contribute to the literature on impact evaluation of natural resource management programs and the link between these programs and farmer well-being through the evaluation of the Socio-environmental and Forestry Development Program II (*POSAF-II)*. Our study sheds light on the effect of actions that can address the "triangle of poverty", which ties low farm productivity to increased poverty, forcing farmers to place pressure on the environment leading to increasing degradation and in turn to even lower productivity and more poverty. A key contribution of this study is to provide a detailed impact evaluation using different methodologies to reduce biases that stem from both observable and unobservable variables while also accounting for spillover effects. Therefore, this study adds to the limited literature on impact evaluation of natural resource management programs and the link to farmers' well-being in Central America.

2. Description of *POSAF-II*

The Nicaraguan Ministry of the Environment and Natural Resources (*MARENA*) implemented *POSAF-II* between 2002 and 2008. The aim of this Program was to improve socio-economic conditions by boosting the productivity of natural resources among small- and medium-sized farmers, primarily through the adoption of soil conservation and water management practices. *POSAF-II* destined a budget of US$20.2 million to promote investments in sustainable agricultural production systems at the farm level based on watershed plans established by Nicaraguan authorities. The central focus was the promotion of a set of alternative production systems designed to increase farm profits and environmental sustainability. Two major production systems were defined: i) Agroforestry (SAGF), including the planting of fruit trees, introduction of soil conservation practices (stone barriers, terraces, live barriers, among others) and silvopastoral sub-systems; and ii) Forestry (SFOR) including forest plantation and regeneration, and management of natural resources. To induce adoption of these systems the program provided technical assistance and materials to participating farmers. The program covered 69,767 hectares and financed a total of 13,477 farmers. To be eligible, farmers had to demonstrate land ownership or any clear documentation of possession of a farm size larger than 1.06 hectares (1.51 Manzanas), located in one of the mentioned river basins, without previous participation in similar programs, and be committed to participate in all *POSAF-II* activities.

3. Analytical Framework and Data

Analytical Framework

Impact evaluations can be conducted through a randomized design where the treated and control groups are assigned before the intervention to ensure that, on average, both groups have the same characteristics in terms of observable and non-observable variables (Angrist and Pischke, 2009; Duflo et al., 2007; Ravallion, 2008). In cases where there is neither an experimental design nor a baseline, as is the case with *POSAF-II*, an alternative methodological approach is to use quasi-experimental methods (Hirano & Imbens, 2001; Khandker et al., 2010). In studies that rely on quasi-experimental methods, careful attention is needed to deal with possible biases stemming from observable and non-observable variables. If one can assume that the source of bias comes only from observable variables, then PSM provides a relatively simple way to mitigate such biases (Bernard and Gabre-Madhin, 2007;

Dehejia and Wahba, 2002). To implement this approach, it is necessary to have a set of covariates associated with program eligibility requirements and other time-invariant variables that are not affected by the intervention. In addition, endline data for a suitable sample of beneficiaries and non-beneficiaries is also required (Khandker et al., 2010). PSM makes it possible to construct statistically a group of non-treated or control units, which is very similar to a group of treated or participating units. This is typically accomplished using a Logit or Probit model to estimate the probability of participating in the program (B = 1) conditioned on a set of observable variables (*X*), and this can be expressed as (Khandker et al., 2010):

$$P(X) = \Pr(B = 1|X) \quad (1)$$

The model makes it possible to calculate propensity scores and then match beneficiaries and control groups based on these scores or probabilities. There is a fairly extensive menu of matching criteria and in this paper we use 1-to-1 nearest neighbor (NN) without replacement. This matching method has a straightforward interpretation and applies the matching based on the common support assumption (Caliendo & Kopeinig, 2008). In addition, it is good practice to apply alternative matching criteria to examine the robustness of the results (Cavatassi et al., 2011; Khandker et al., 2010) and to this end we use the Genetic Matching method.

After matching, the effect of the program is equal to the average difference of the outcome indicator(s) between the beneficiary and the control group. This difference, known as the Average Treatment Effect on the Treated (ATET), can be expressed as:

$$\alpha = E(Y_{Bi} - Y_{Ci} \,|\, \Pr(X), B = 1) \quad (2)$$

where Y_{Bi} and Y_{Ci} represent the value of the pertinent indicator, the total value of agricultural production (TVAP) in this study, for beneficiaries (*B*) and the control group (*C*), respectively.

A second approach to evaluate the impact is a standard OLS regression, where the program's impact on the outcome variable Y_i is determined by the following equation:

$$Y_1 = \alpha_0 + \alpha_1 B_i + \sum \gamma_i X_{ij} + \varepsilon_i \quad (3)$$

where α_1 measures the ATET of the Program, $B_i = 1$ if households participate, 0 otherwise, γ_i are the parameters to be estimated associated with covariates X_{ij}, and ε_i is the typical error term. A problem with this simple approach is that participation (B_i) is likely to be correlated with the error term, which yields biased estimates (Angrist and Pischke, 2009).

In addition to OLS, equation (3) can be estimated using weighted least squares (WLS) where the weights are based on the propensity scores obtained from the PSM. This approach was introduced by Hirano and Imbens (2001) and has been used by Todd et al. (2010) and Cavatassi et al. (2011), among others. This method is implemented as follows: a) Propensity scores (PS) are estimated using a Logit or a Probit model; b) B_i and X_{ij} are weighted by $1/PS(X)$ for beneficiaries and $1/(1-PS(X))$ for controls; and c) Equation (3) is estimated using OLS and the weighted data.

The estimates from equation (3), although superior to those obtained from the conventional OLS model, would be biased if program participation is correlated with unobservables captured in the error term. To address this endogeneity problem and thus ensure that the estimated impact of *POSAF-II* is not biased due to unobservables, an instrumental variable (IV) approach is implemented. This method requires finding an instrument *Z*, which is related to the participation in *POSAF-II* but not correlated with the error term, i.e., $cov(Z, \varepsilon) = 0$. Following Cavatassi et al. (2011) and Khandker et al. (2010), we use 'Intention to Treat' (ITT) as an instrument. This method is appropriate given that the "ITT analysis captures the causal effect of being assigned to the treatment" (Angrist & Pischke, 2015, p.119). ITT relies on the fact that some of those assigned to be treated chose not to receive the treatment. Before using ITT as an instrument, we conduct a Hausman and a Durbin-Wu-Hausman tests for exogenous regressors (Khandker et al., 2010).

The IV approach requires a two stage procedure as follows:

$$Stage\ 1: B_i = \rho Z_i + \sum \phi_j X_{ij} + \nu_i \qquad (4)$$
$$Stage\ 2: Y_i = \lambda_0 + \lambda_1 \hat{B}_i + \sum \delta_j X_{ij} + \mu_i \qquad (5)$$

In the first stage (equation 4), the instrument Z_i is introduced in an equation that explains the participation in *POSAF-II* (B_i). In the second stage, the B_i variable is replaced by the predicted participation in *POSAF-II* ($\hat{B}_i$) obtained in the first stage. This model is then estimated to obtain the measure of impact given by λ_1 in equation (5). All Greek characters (i.e., ρ, ϕ, λ and δ) are the parameters to be estimated.

The data collection procedure to evaluate *POSAF-II* followed two-stages. First PSM was used to match treatment and control communities. Information regarding treated communities was obtained from the monitoring and evaluation system implemented by *POSAF-II*, known as SIMOSE. The list of control communities was based on the National Water Resources Plan for Basins, Sub-basins, and Micro-basins, obtained from the *MARENA*. The matching at the community level was based on agro-ecological characteristics including: altitude (*ALT*); temperature (*TEMP*); precipitation (*PRECI)*; and short-term-drought (*STD*). All variables are defined in Table 1. These variables were selected based on data availability as well as on information obtained from local experts and *POSAF-II* personnel who considered such variables as critical in matching communities consistent with the technologies offered.

The Logit model for the communities can be written as:

$$COMU = f\,(ALT, TEMP, PRECI, STD) + error\ term \qquad (7)$$

where *COMU* is equal to 1 for the *POSAF-II* communities and 0 for the control communities. The results of the Logit model were used to match the communities based on the 1-to-1 nearest-neighbor (NN) criterion. After matching, 618 communities (309 treated and 309 control communities) were selected. From this total of 309 pairs, 106 pairs were randomly chosen using the RAND procedure of SQL. The quality of this final selection was evaluated and deemed appropriate by a local panel of experts.

Panel A in Table 2 shows the pre-treatment variables included in the community level Logit model. The predicted probabilities show that communities located at higher altitudes, with higher temperatures, and lingering short-term-drought periods were less likely to be selected for program implementation. In addition, those communities with higher precipitation were more likely to receive the program. Among those characteristics, the parameters for *ALT* and *PRECI* are statistically significant, which is consistent with the program implementation criteria (MARENA, 2005). Furthermore, at the 1% level of significance, the null hypothesis that all parameters are jointly equal to zero is rejected.

In the second stage of the data collection process, SIMOSE was used to create a list including all beneficiaries of *POSAF-II* and a group of eligible non-beneficiaries located in the 106 treated communities selected in the first stage. Hereafter, the beneficiaries are referred to as BENE, and the non-treated in beneficiary communities as CONI for control in. An additional control group was generated from the 106 matched non-treated communities, hereafter referred to as CONO for control out. Having controls outside the program allows for the examination of spillover effects, i.e. whether untreated farmers located in treated communities received indirect benefits by interacting with neighboring beneficiaries (Angelucci & De Giorgi, 2009). These potential spillover effects can be a significant benefit derived from a project and are thus important to quantify. Moreover, this type of design, as alluded to earlier, makes it possible to define the ITT instrument.

Once the sampling frame was defined, households from each group were randomly selected. Following the procedures in Wassenich (2007), the final sample size for beneficiaries was 257 and 327 for SAGF and SFOR, respectively. A sample size of 641 farmers (318 CONI, and 323 CONO) was defined for the control group. In Section 4 below, we describe the matching undertaken at the farmer level.

Descriptive analysis

As aforementioned, the impact of *POSAF-II* is analyzed separately for SAGF and SFOR. Table 1 presents descriptive statistics. The BENE group for each system is compared with the corresponding counterfactual groups to determine whether the means under analysis are the same. On average, the TVAP of BENE is higher for the two systems (SAGF \$1,045, SFOR \$1,041) compared to the respective controls. Another variable to note is *LAND*, which refers to the area used for agricultural production. For the SAGF group, BENE and CONO have similar farm sizes, with 15.8 and 14.9 hectares, respectively. Similarly, average farm size is equal for BENE and CONO in SFOR; hence, these groups are comparable.

The beneficiaries of *POSAF-II* share most of the characteristics of the control farmers. As would be expected, *t*-tests show statistically significant differences among variables affected by the program's implementation, such as TVAP. An exception is *COST*, which does not exhibit any statistical difference between treated and controls for SAGF (BENE \$520, CONI \$825, CONO \$759). In SFOR, BENE is statistically different from the control groups (CONI \$799, and CONO \$580); however, the mean value (\$489) is lower than those in the comparison groups. Even though the program required that beneficiaries worked in the implementation of the various technologies, the cost variable does not display higher means for any of the systems. In addition, *COST*, *EDUC*, and *NET* in the treatment group are not statistically different from CONI and CONO; hence, as already indicated, we have been able to define a suitable counterfactual situation based on observables.

Annual precipitation is between 1,285 to 1,314 millimeters for SAGF, and between 1,281 to 1,304 millimeters for SFOR, and both treatment and control communities received nearly the same amount of rain. The mean values for STD across treatment and control groups are very similar among systems. Again, comparisons based on these variables ensure a reasonable counterfactual.

4. Results and Discussion

Matching beneficiaries with control farmers

As discussed earlier, the first step in defining the samples was to match treatment and control communities. Now we proceed to match farmers for SAGF and SFOR. Two Logit models, one for each system, are estimated to determine the probability of being a *POSAF-II* beneficiary. In each Logit model, the dependent variable is equal to 1 for BENE, and zero for the controls, CONI and CONO.

As depicted in Panel B of Table 2, some of the parameters for the Logit model differ across systems; namely, the parameter for land has a positive and significant effect on the participation in SFOR and a non-significant effect on SAGF. Other covariates include those related to agro-ecological conditions, such as *PRECI*, *ALT*, *STD*, and *TEMP*. Farmers located in areas with higher levels of precipitation are less likely to be SFOR beneficiaries. Farms located at higher elevation are less likely to participate in both SAGF and SFOR. The signs for the respective parameters are as expected given the fact that agricultural activities in these locations are less common due to lack of adequate infrastructure. The percentage of correct predictions for being a beneficiary of *POSAF-II* is 70.4% and 66.7% for farmers in SAGF and SFOR, respectively. Furthermore, chi-squares of 118.8 and 118.9 with 11 degrees of freedom and p-values lower than 0.001 indicate that the parameters in the two models are jointly significantly different from zero. In sum, the statistical results shown in Table 2 suggest that the models are appropriate to explain the participation in the program.

In order to check the common support condition, we provide a graphical balance check with the kernel density estimates of the estimated propensity scores of treatment and control groups for each system (Figure 1). The results show that most of the propensity scores estimated for the BENE and both control groups fall within the common support area for the two systems. In addition, as suggested by Sekhon (2011), we also ran a bootstrapped Kolmogorov-Smirnov test (KS) following Abadie (2002) and the analysis of differences in means shows that matching significantly improved the covariate balance for both SAGF and SFOR implying that beneficiaries and controls are not statistically different. In addition to the nearest neighbor 1-to-1 matching method, we use genetic matching following Diamond and Sekhon (2013) to check the robustness of the matching process. The genetic matching does not improve the covariate balance since the nearest-neighbor 1-to-1 has a smaller KS test statistic with p-values lower than 0.01. However, both matching techniques produce similar p-values for the difference in means. In sum, the t-tests for the two systems show that based on observable characteristics the control groups represent a good counterfactual.

Impact on farmer incomes

As has been indicated above, the economic impact of *POSAF-II* is examined based on four alternative estimation techniques, PSM, OLS, WLS and IV. The indicator of impact is the TVAP for each system. All of the estimated models represent production functions where the dependent variable is expressed in monetary values. To conserve space[1], Table 3 presents the key parameters concerning the estimated impact of *POSAF-II* on the TVAP for SAGF and SFOR. The *F*-statistics in the two regression models are significant at the 1% level; therefore, the joint hypothesis that all coefficients equal zero in each model is rejected.

Our estimates show consistently that *POSAF-II* has a positive and significant effect on the TVAP of beneficiaries relative to controls for SAGF and SFOR based on all four procedures used, i.e., PSM, OLS, WLS and IV. The average increase in TVAP attributable to *POSAF-II* for SAGF farmers is US$330 (PSM), US$343 (OLS), US$695 (WLS) and US$1058 (IV). For SFOR, the average impact of *POSAF-II* on TVAP is US$23, US$604, US$650 and US$913 for PSM, OLS, WLS, and IV models, respectively.

As mentioned previously, to deal with unobservable characteristics between the treated and control groups, the IV approach is used as an alternative estimation method. In order to check for the validity of the ITT as an instrument, we use a weak instrument test (Angrist and Pischke, 2009) and the test results reject the null hypothesis of a weak instrument with an F statistic larger than the rule of thumb of 10, which means that ITT is a valid instrument. Subsequently, we use a modification of the Hausman test (Khandker et al., 2010) in order to check whether participation in *POSAF-II* is exogenous and the result suggests that B_i is indeed exogenous. Hence, there should be no difference between the OLS and IV coefficients, while OLS guarantees a higher efficiency in the estimates (Greene, 2007). Moreover, Wooldridge (2002) argues that a correctly specified WLS leads to more efficient estimates than OLS and this makes the former the more desirable method here. It is worth noting that Cavatassi et al. (2011) also concluded that the WLS approach was the best method in their impact evaluation of the *Plataformas* program in Ecuador.

In the previous estimations, the counterfactual group includes all control farmers, i.e., both CONI and CONO. To examine the possible presence of spillover effects, we now compare the BENE with the CONO groups and the results, presented in Panel-A of Table 4, show that the parameters are positive and statistically significant. For SGAF and SFOR, the impact of *POSAF-II* on TVAP is

[1] Complete estimates of the results are available from the authors upon request.

US$852 and US$938, respectively, and both are higher than the estimates obtained when all controls are used. To further examine the effect of *POSAF-II* on the beneficiary communities, we re-estimate the models contrasting the CONO vs. the CONI groups. Panel-B of Table 4 shows that control farmers from the treatment communities have on average a TVAP of US$425 for SAGF and US$301 for SFOR, which are higher than their counterparts in the non-treated communities. These results suggest that *POSAF-II* had spillover effects on farmers living in proximity to the treated groups.

According to Knowler and Bradshaw (2007), the adoption of NRM technologies in agriculture is correlated with individual motivation, household structure, and agro-ecological characteristics. Among BENE and CONI groups, the latter characteristics are similar; knowledge diffusion is thus likely to occur. Another possible explanation behind these spillover effects is the level of complexity of the technology. Greiner and Gregg (2011) suggest that the adoption of conservation practices is motived by the technological characteristics of the practices. *POSAF-II* delivered some technologies with a relatively low level of complexity and considerable positive effects on production such as fencing, contour plowing, high-quality fruit trees, banana plants with sanitary treatment, and forest trees.

5. Concluding Remarks

In this study, we examine the economic impact of *POSAF-II*, a natural resource management (NRM) program implemented in Nicaragua between 2002 and 2008, on the total value of agricultural production (TVAP). The Program supported small and medium scale producers in improving the use of natural resources, in order to increase productivity and reduce environmental degradation. The econometric analysis relies on methodologies designed to reduce biases that stem from both observable and unobservable variables when only endline data is available. The methodologies implemented include propensity score matching (PSM), ordinary least squares (OLS), weighted least squares (WLS) and instrumental variables (IV). The motivation behind the use of different methods is to evaluate the robustness of the analysis. If different methodologies lead to similar outcomes, then the likelihood that results are reliable is high.

The results for both SAGF and SFOR indicate that *POSAF-II* had a positive and significant impact on the beneficiaries attributable to the program. While the outcomes are consistent among different model specifications, the WLS results are the most robust. These results indicate that the impact of *POSAF-II* on the TVAP of beneficiaries with respect to controls is US$695 for SAGF, US$650 for SFOR. Moreover, the analysis clearly suggests that *POSAF-II* resulted in an overall increase in the total value of agricultural production to beneficiaries and this increase can be attributed to the program.

In addition to the direct impact on the beneficiary groups, *POSAF-II* had positive spillover effects on non-treated farmers living inside treated communities. This impact was estimated by comparing beneficiaries vs. control farmers outside treated communities and control individuals inside treated communities with control individuals living outside treated communities. In both cases, estimates are higher than those obtained when beneficiaries are compared with all control farmers.

A lesson derived from this study is the importance of identifying the most suitable time for carrying out an impact evaluation. The bulk of the data available for *POSAF-II* was collected four years after the program had closed and this is different from the typical case when endline data is collected just before closing the project. The implication of this typical case is that farmers have a very limited time to implement the technologies received and thus benefits can be very limited. In contrast, the four years that had elapsed since completion of *POSAF-II* gave farmers sufficient time to fully adopt the technologies. Ideally, however, one would be able to revisit these farmers again 10 or 15 years after closing to be able to fully gauge the accrued benefits and the long term sustainability of the intervention. In sum, and very importantly, the results for *POSAF-II* suggest that it is possible to have interventions that increase farm income and preserve or enhance environmental conditions.

Table 1: Descriptive statistics for variables included in the analysis by System

		Agroforestry (SAGF)			Forestry (SFOR)		
Variable	Definition and measurement unit	BENE	CONI	CONO	BENE	CONI	CONO
TVAP	Total value ag. Prod. (US$/hectare)	1044.7 [a,b,c]	878.5 [c]	792.2	1040.8 [a,b,]	613.0 [c]	503.4
AGE	Age of the household head (years)	53.5 [a,b,c]	42.3[c]	49.9	54.7 [a,b,]	42.6 [c]	50.4
EDUC	Years schooling of household head	4.2	4.3	4.9	4.77	4.26	5.2
NET	1 if farmer is member of org.	0.2	0.2	0.2	0.2	0.2	0.2
LAND	Area devoted to production (ha)	15.8 [a,c]	6.4 [c]	14.9	24.42 [a,]	6.9 [c]	26.5
DIST	Plot distance to main town (kms)	44.1[a,b,c]	37.5 [c]	27.4	42.8 [b,]	42.3 [c]	24.4
ALT	Meters above sea level (Meter)	492.9 [b,d]	557.0 [c]	752.9	524.7 [b,c]	563.2 [c]	696.2
PAVE	1 if farm located next to paved road	0.3[b,d]	0.3 [c]	0.5	0.3	0.3	0.3
ACCE	1 if the farm is accessible all year	0.6 [b,c]	0.6 [c]	0.7	0.6	0.6	0.6
TEMP	Avg. temperature in the region (Cº)	24.0 [b,c]	23.6 [c]	22.5	23.7 [b,]	23.6 [c]	22.4
PRECI	Annual precipitation (millimeters)	1284.5	1314.1	1286.3	1280.5	1289.6	1303.9
STD	Drought days during a raining season	23.6	20.3	23.4	21.7	18.6	27.7 [c]
COST	Variable prod. costs, excluding labor	520.0 [b]	825.2	753.79	488.61 [a,b,]	799.4	580.29
FLABOR	Total value of family labor	127.9 [a,b]	171. [c]	82.4	82.7 [a,c]	161.9 [c]	84.80
LABOR	Total value of hired labor expense	368.3 [a,b]	688.2 [c]	458.9	664.1 [c]	573.2	877.8 [a,b]

"**a**" diff. between mean of BENE & CONI is statistically signif. at least at 10%. "**b**" diff. between mean of BENE & CONO is statistically signif. at least at 10% level. "**c**" diff. between mean of CONO & CONI is statistically signif. at least at 10% level.

Table 2: Logit model of *POSAF-II* participation used to match communities and farmers

Variables	Panel A.		Panel B.			
	Community		SAGF		SFOR	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
AGE			0.018 [a]	0.004	0.033 [a]	0.004
EDUC			-0.008	0.015	0.039 [b]	0.013
NET			0.198	0.157	0.005	0.136
LAND			0.003	0.002	0.003 [c]	0.001
DIST			0.012 [a]	0.003	0.008 [a]	0.001
PAVE			-0.199	0.159	-0.175	0.126
ACCE			-0.036	0.152	0.255 [c]	0.126
ALT	-0.003 [a]	0.001	-0.012 [a]	0.003	-0.000 [c]	0.000
TEMP	-0.017	0.054	-0.017	0.054	0.083	0.049
PRECI	0.002 [a]	0.000	-0.001	0.000	-0.001 [c]	0.000
STD	-0.008	0.006	-0.022 [a]	0.005	-0.014 [b]	0.004
Constant	0.796	1.438	1.258	1.437	-2.849	1.374
N total		797		680		643
BENE				239		289
CON				441		354
Log likelihood		*-471.7*		-289.3		-397.8
LR chi^2(11)		*125.5* [a]		118.8 [a]		118.9 [a]
Pseudo R^2				0.17		0.13
Correctly classified				70.4%		66.7%

a = significant at the 1%, b = significant at the 5% and c = significant at the 10%

Table 3: Impact of *POSAF-II* on SAGF and SFOR

	Agroforestry System (SAGF)				**Forestry System (SFOR)**			
	PSM[@]	**OLS**	**WLS**	**IV**	**PSM[@]**	**OLS**	**WLS**	**IV**
POSAF-II	330.32 [b]	342.78 [a]	695.03 [a]	1057.96 [a]	23.19 [c]	603.62 [a]	650.36 [a]	912.92 [a]
	(130.1)	(131.4)	(233.8)	(331.5)	(15.3)	(204.7)	(162.4)	(336.9)
N	478	680	680	680	578	643	643	643
F(Chi^2)		2.75 [a]	2.52 [a]	4.02 [a]		3.04 [a]	2.82 [a]	1.76 [a]
R^2		0.08	0.11	0.06		0.05	0.07	0.05

Robust standard errors for OLS and WLS, standard errors for IV in parenthesis. [@]Values in parenthesis. Bootstrap with 1000 replications is used to estimate the standard errors. [c] $p<0.10$; [b] $p<0.05$; [a] $p<0.01$

Table 4: Spillover effect of *POSAF-II* on the two systems

	Panel A. BENE vs. CONO		Panel B. CONI vs. CONO	
	SAGF	SFOR	SAGF	SFOR
TVAP	852.5 [b]	938.4 [a]	425.2 [c]	301.6 [b]
	(368.5)	(282.1)	(285.9)	(133.8)
N	364	354	441	330
R^2	0.09	0.10	0.09	0.14

Robust standard errors in parenthesis p<0.10; [b] p<0.05; [a] p<0.01

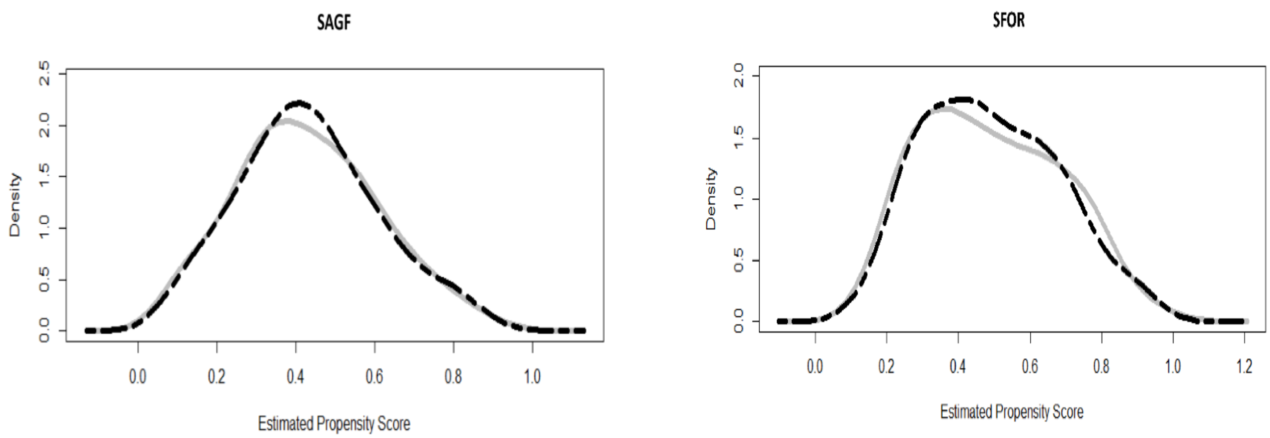

Figure 1: Kernel distribution of propensity scores for BENE (broken black line) and corresponding control groups (continuous gray line).

References

Abadie, A., 2002. Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models. J. Am. Stat. Assoc. 97, 284–292. doi:10.1198/016214502753479419

Angelucci, M., De Giorgi, G., 2009. Indirect effects of an aid program: How Do Cash Transfers Affect Ineligibles' consumption? Am. Econ. Rev. 99, 486–508. doi:10.1257/aer.99.1.486

Angrist, J., Pischke, J.-S., 2015. Mastering 'Metrics: The Path from Cause to Effect. Princeton University Press., New Jersey.

Angrist, J., Pischke, J.-S.P., 2009. Mostly Harmless Econometrics : An Empiricist's Companion. Princeton University Press., New Jersey.

Barrett, C.B., Moser, C.M., Mchugh, O. V, Barison, J., 2004. Better Technology, Better Plots or Better Farmers ? Identifying Changes In Productivity And Risk Among Malagasy Rice Farmers. Am. J. Agric. Econ. 86, 869–88. doi:10.1111/j.0002-9092.2004.00640.x

Bernard, T., Gabre-Madhin, E., 2007. Smallholders' Commercialization through Cooperatives. IFPRI Discussion Paper 00722, Washington, DC.

Bravo-Ureta, B.E., Almeida, A.N., Solís, D., Inestroza, A., 2011. The Economic Impact of Marena's Investments on Sustainable Agricultural Systems in Honduras. J. Agric. Econ. 62, 429–448. doi:10.1111/j.1477-9552.2010.00277.x

Caliendo, M., Kopeinig, S., 2008. Some Practical Guidance for the Implementation of Propensity Score Matching. J. Econ. Surv. 22, 31–72. doi:10.1111/j.1467-6419.2007.00527.x

Cavatassi, R., González-Flores, M., Winters, P., Andrade-Piedra, J., Espinosa, P., Thiele, G., 2011. Linking Smallholders to the New Agricultural Economy: The Case of the Plataformas de Concertación in Ecuador. J. Dev. Stud. 47, 1545–1573. doi:10.1080/00220388.2010.536221

Cleaver, K., 2012. Investing in Agriculture to Reduce Poverty and Hunger, Brief. International Food Policy Research Institute., Washinton, DC.

Cocchi, H., Bravo-Ureta, B.E., 2007. On-Site Costs and Benefits of Soil Conservation Among Hillside Farmers in El Salvador (No. OVE/WP-04/07). Washinton, DC.

Dalton, T.J., Lilja, N.K., Johnson, N., Howeler, R., 2005. Impact of Participatory Natural Resource Management Research in Cassava-Based Cropping Systems in Vietnam and Thailand, in: The Impact of Natural Resource Management Research in the CGIAR, Systemwide Program on Participatory. FAO and Wallingford, UK:CAB International, Italy, pp. 1–27.

Dehejia, R.H., Wahba, S., 2002. Propensity score-matching methods for nonexperimental causal studies. Rev. Econ. Stat. 84, 151–161. doi:10.1162/003465302317331982

Del Carpio, X. V., Maredia, M., 2011. Impact Evaluations in Agriculture: An Assessment of the Evidence. The World Bank, Washinton, DC.

Diamond, A., Sekhon, J.S., 2013. Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. Rev. Econ. Stat. 95, 932–945. doi:10.1162/REST_a_00318

District, K., Kingdom, U., 2011. The System of Rice Intensification : An Alternate Civil Society Innovation. Kassam, Amir 20, 145–152.

Duflo, E., Glennerster, R., Kremer, M., 2007. Using Randomization in Development Economics Research: A toolkit (No. 6059), Discussion Paper Series, CEPR Discussion Paper Series. London. doi:10.1016/S1043-2760(97)84344-5

Dutilly-diane, C., Sadoulet, E., de Janvry, A., 2003. Household Behaviour under Market Failures : How Natural Resource Management in Agriculture Promotes Livestock Production in the Sahel. J. Afr. Econ. 12, 343–370. doi:10.1093/jae/12.3.343

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B., Vermeersch, C.M., 2011. Impact Evaluation in Practice. The World Bank, Washinton, DC. doi:10.1596/978-0-8213-8541-8

Gornall, J., Betts, R., Burke, E., Clark, R., Camp, J., Willett, K., Wiltshire, A., 2010. Implications of climate change for agricultural productivity in the early twenty-first century. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365, 2973–89. doi:10.1098/rstb.2010.0158
Greene, W.H., 2007. Econometric Analysis, 6th ed. Prentice Hall, New Jersey.
Greiner, R., Gregg, D., 2011. Farmers' intrinsic motivations, barriers to the adoption of conservation practices and effectiveness of policy instruments: Empirical evidence from northern Australia. Land use policy 28, 257–265. doi:10.1016/j.landusepol.2010.06.006
Harwood, R.R., Kassam, A.H., Gregersen, H.M., Fereres, E., 2005. Natural Resources Management Research in the CGIAR : the Role of the Technical Advisory Committee. Exp. Agric. 41, 1–19. doi:10.1017/S0014479704002315
Hirano, K., Imbens, G.W., 2001. Estimation of Causal Effects using Propensity Score Weighting : An Application to Data on Right Heart Catheterization. Health Serv. Outcomes Res. Methodol. 259–278. doi:10.1023/A:1020371312283
Kassie, M., Shiferaw, B., Muricho, G., 2011. Agricultural Technology, Crop Income, and Poverty Alleviation in Uganda. World Dev. 39, 1784–1795. doi:10.1016/j.worlddev.2011.04.023
Kelley, T., Ryan, J., Gregersen, H., 2008. Enhancing ex post impact assessment of agricultural research: the CGIAR experience. Res. Eval. 17, 201–212.
Khandker, S.R., Koolwal, G.B., Samad, H.A., 2010. Handbook on Impact Evaluation. The World Bank, Washinton, DC.
Knowler, D., Bradshaw, B., 2007. Farmers' adoption of conservation agriculture: A review and synthesis of recent research. Food Policy 32, 25–48. doi:10.1016/j.foodpol.2006.01.003
MARENA (Ministerio del Ambiente y los Recursos Naturales), 2005. Establecimiento y manejo de Sistemas. Managua, Nicaragua. doi:10.1017/CBO9781107415324.004
Mendola, M., 2007. Agricultural technology adoption and poverty reduction: A propensity-score matching analysis for rural Bangladesh. Food Policy 32, 372–393. doi:10.1016/j.foodpol.2006.07.003
Pal, S., 2011. Impacts of CGIAR Crop Improvement and Natural Resource Management Research : A Review of Evidence. Agric. Econ. Res. Rev. 24, 185–200.
Ravallion, M., 2008. Evaluating Anti-Poverty Programs, in: Handbook of Development Economics. Elsevier, Amsterdam. doi:10.1016/S1573-4471(07)04059-4
Renkow, M., Byerlee, D., 2010. The impacts of CGIAR research: A review of recent evidence. Food Policy 35, 391–402. doi:10.1016/j.foodpol.2010.04.006
Science-Council, 2006. Natural Resources Management Research Impacts : Evidence from the CGIAR. Rome.
Sekhon, J.S., 2011. Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching package for R. J. Stat. Sofware 42, 1–52. doi:10.18637/jss.v042.i07
Todd, J.E., Winters, P.C., Hertz, T., 2010. Conditional Cash Transfers and Agricultural Production: Lessons from the *Oportunidades* Experience in Mexico. J. Dev. Stud. 46, 39–67. doi:10.1080/00220380903197945
Tsiboe, F., Dixon, B.L., Nalley, L.L., Popp, J.S., Luckstead, J., 2016. Estimating the impact of farmer field schools in sub-Saharan Africa: the case of cocoa. Agric. Econ. 47, 329–339. doi:10.1111/agec.12233
Wassenich, P., 2007. Data for Impact Evaluation (No. 6), Doing impact evaluation series. The World Bank., Washington, DC.
Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data, 1st ed. MIT Press, Massachusetts.
World Bank., 2016. High and Dry: Climate Change, Water, and the Economy. Washington, DC.