Development in the Midst of Drought: Evaluating an Agricultural Extension and Credit Program in Nicaragua.

Conner Mullally
Inter-American Development Bank
mullally@primal.ucdavis.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011.

Abstract:

This paper measures the impact of year one of the Millennium Challenge Corporation’s Rural Business Development program for small rice farming households on the Pacific Coast of Nicaragua. The program was rolled out in the 2009-2010 agricultural year, which was the driest year on record in the region, likely due to an El Niño event. Estimated impacts show that the program at best had no effect, and at worst led to a 10 percent reduction in yields. These impacts are estimated using an econometric model which uses selection on observables as its identifying assumption, and robustness checks suggest that this is a reasonable approach in this case. Inference accounts for spatial correlation across households of the unobserved determinants of agricultural outcomes. The program appears to have been almost exclusively focused on increasing yields through better and greater application of chemical fertilizers, and minimization of losses in the post-harvest stages of production. If the pessimistic estimates of program effects are true, then the program could have been improved by incorporating risk management strategies into extension advice. On the other hand, farmers may be well insured against climatic risk, in which case they may have selected into the program knowing that they would be trading greater risk for higher expected returns. Survey data offers some evidence that the latter is indeed the case.

1 Introduction

When thinking of interventions designed to combat rural poverty, agricultural extension and credit appear to be natural complements. By delivering knowledge to farmers about productivity enhancing techniques and the proper use of inputs, extension can increase returns to capital invested in agricultural activities or diminish risks associated with agriculture. At the same time, including credit as a component of an agricultural extension program can give farmers the resources necessary to fully exploit the knowledge gained via extension services, and bring households into the market for extension services that otherwise could not afford to participate.

This essay evaluates year one of the two-year Rural Business Development (RBD) program for rice farmers in León and Chinandega, located on the Pacific Coast of Nicaragua. The program combines credit in the form of agricultural inputs with agricultural extension services tailored to individual farms. The RBD program is funded jointly by the U.S. and Nicaraguan Governments, and is administered by the local office of the Millennium Challenge Corporation (MCC), a development agency of the U.S. Government.

Estimated impacts indicate that participants in the program…on average. XXX There are reasons to believe that these results are due to the combination of the nature of the benefits offered by the RBD program and the severe drought that occurred in the 2009-2010 agricultural year in the study area due to an El Niño event. The timing of rice planting decisions in León and Chinandega are such that the magnitude of the 2009-2010 El Niño event was not known until quite late in the growing season; the vast majority of farmers plant in July, which is when the presence of an El Niño event in 2009 was first confirmed, but its magnitude was not known until much later in the growing season (IRI 2009). xxx more on nature of extension advice and impact on yields they start planting in june, done in august. harvest November and december

Failing to detect a positive impact in a drought year does not necessarily mean that farmers will not benefit from the program. Deciding whether or not to join an extension program that also offers credit may require weighing a tradeoff between higher expected returns and greater risk. Farmers might elect to participate in the RBD program because of gains from participation that occur over time in years characterized by favorable production conditions, while output in years with poor conditions for rice production could be unaffected or even decrease due to enrolling in the program.

In the case of the RBD program, the skills learned via extension agents could be applied in future years in which conditions are more suitable for rice. Thus the complete stream of benefits due to the program cannot be captured in a static framework. However, the negative aspects of poor outcomes among participants also have dynamic implications. If farmers cannot meet their debt obligations out of income or by cutting consumption, and cooperatives enforce debt repayment, then a long-term deepening of poverty may occur if households sell off assets to meet debt obligations (Carter and Barrett 2006).

Whether this outcome could obtain depends on how well insured households are against shocks; participants did XXX did not sell off assets to a significantly higher degree than non-participants, and there was little change in asset holdings in general, suggesting that in this particular case households were sufficiently diversified. Alternatively, if the leaders of the various farmers' cooperatives do not strictly enforce repayment, the result may be a decrease in the services that each cooperative can offer its members. This would also hurt farmers over time, but the impact would be spread out over the membership of each cooperative. XXX In either case, the implication is that a program combining extension services with credit ought to take the risks faced by households and the degree to which they are insured into account in program design.

This paper adds to the literature on agricultural extension and credit interventions in developing country agriculture. Much has been written about agricultural extension in developing countries, and earlier work in this area in the context of developing economies is surveyed by Anderson and Feder (2004). When econometric methods have been employed, much of this literature reports high returns to investments in extension services, e.g., Bindlish and Evenson (1997). But as noted by Anderson and Feder, data quality and issues of econometric methodology give reason to doubt some of these results. As shown by Gautam and Anderson (1999), small changes to model specifications can drastically reduce high estimated returns to extension investments.

Later studies have made improvements to econometric methodology, and several of these are summarized in Cerdán-Infantes, Maffioli, and Ubfal (2008). Studies such as those by Praneetvatakul and Waibel (2007) and Godtland et al. (2004) tend to find that extension services have had success with respect to knowledge transfer but mixed effects on productivity and income. Overall, the evidence for the benefits of extension to agriculture in developing countries is mixed, and this conclusion extends to the various modalities by which extension services can be delivered (Anderson and Feder 2007).

Rural credit markets are the subject of their own rich literature, but only a small portion of research has been aimed at measuring the effects of credit on agricultural productivity and incomes. Existing studies generally find positive effects of credit receipt and access on agricultural productivity and incomes, but magnitudes vary considerably. Carter (1989) finds weak evidence of a positive relationship between receipt of credit and farm income and productivity in Nicaragua. Feder et al. (1990) and Foltz (2004) find modest effects of relaxing credit constraints on households on output and incomes, the former in the case of rural China and the latter using a sample of Tunisian farms. Guirkinger and Boucher (2008) use a broader definition of credit rationing than that employed by Feder et al. and Foltz, expanding the group of rationed households to include those that exit the credit market due to transaction costs or unwillingness to bear the risk of losing collateral in case of default. They estimate much large impacts of eliminating credit constraints on farmers in rural Peru equal to an increase of 26 percent in the value of output per hectare.

As summarized by Del Carpio and Maredia (2009), there are a relative small number of rigorous impact evaluations of agricultural extension and rural credit market projects in the literature. Their survey of the literature from 2000 to early 2009 identified 20 studies of agricultural extension project and 10 addressing rural credit interventions that satisfied a few basic criteria for categorization as an impact evaluation.[1] When the scope of these studies is limited to evaluations of projects that combine extension services with credit, the number becomes smaller still. One recent example is Ashfar, Giné, and Karlan (2009), who evaluate the impact of DrumNet in Kenya, a program designed to increase participation of horticulturalists in export markets. The authors of that study randomly assign groups of farmers to treatments including extension services, extension with a joint liability loan, and no treatment. They find significant impacts of both versions of the program on production of export crops, formal financial market participation, and significant increases in income among first time growers of export crops.

[1] Basic criteria for inclusion were 1) A focus on agriculture, 2) A defined agricultural intervention, 3) A clearly stated counterfactual (e.g., cannot measure impact simply by using a before and after comparison on a single group).

This essay does not have the benefit of randomized assignment to treatment. Instead, the identification strategy employed is to assume that selection into the program is based on observable characteristics, and program effects are estimated using inverse propensity score weighting combined with linear regression (Wooldridge 2007). The soundness of this assumption is tested to the extent possible using available data, and results suggest that xxx.

The unique features of this paper are the conditions under which the RBD program was rolled out, and the use of spatial methods in conducting statistical inference. By evaluating the RBD program in the context of a severe and unexpected climatic shock, the results of the analysis can serve as a cautionary tale for design of future programs. In conducting inference, standard errors are estimated using the spatial heteroscedasticity and autocorrelation robust covariance (HAC) matrix of Kelejian and Prucha (2007). A review of the literature uncovered no previously published impact evaluations in agricultural development that account for spatial autocorrelation.

In what follows, Section X describes the study area of León and Chinandega and the characteristics of the RBD program. Section X describes the estimation strategy employed. Section X reports estimation results and robustness checks, and Section X concludes.

2 The Rural Business Development program[2]

2.1 The Study Area and the goals and benefits of the RBD program

León and Chinandega are home to around 830,000 persons, 39 percent of which live in rural areas and are involved in agriculture. Nearly all smallholder agriculture is rainfed, with the vast majority of irrigated hectares under the control of large agribusinesses, usually sugarcane or

[2] This section and the one that follows draw from documentation provided by the Nicaragua office of MCC, and are available from the MCC Nicaragua website (http://www.cuentadelmilenio.org.ni) or from the author.

plantain. Along with sesame seeds, maize, and sorghum, rice is one of the primary crops planted by small farms in the region.

There are essentially two levels of benefits to the RBD rice program: cooperative-level benefits and individual-level benefits. Rice farmers participating in the RBD program are all members of cooperatives, and cooperatives with members in the program receive bundles of inputs for rice production sufficient for three manzanas[3] per participating farmer from MCC. These inputs are then leant out to participating members; interest rates on these loans vary across cooperatives, as credit contract details are controlled by cooperatives rather than MCC. While the input packets are meant to spur production in the short term, they are also designed to help each cooperative establish a rotating credit fund that will make liquidity available to farmers at in future years. For each participating cooperative, MCC pays a maximum of 30 percent of the costs associated with the program; the rest is paid for by the cooperative.

At the level of the producer, the RBD program for rice farmers also features benefits in the form of agricultural extension services, focused on tailoring the use of chemical fertilizers to the soil characteristics of each individual farm, more efficient use of agrochemicals meant to control threats to the plant, and on better management of the post-harvest stages of production. The costs of this technical assistance are factored into the portion of the total cost of program participation borne by each participating cooperative.

2.2 Eligibility criteria and participation in the RBD program

For rice farmers, participation in the RBD consists of several stages, the first of which is satisfying eligibility for participation in the program. Eligibility criteria include:

- The producer has planted or currently has at least 2 manzanas of rice.

[3] 1 manzana = 1.72 acres = 0.70 hectares

- Area of farm must be between 2 and 50 manzanas, non-irrigated.
- The main rice parcel must be property of the beneficiary.
- The main rice parcel must be outside environmentally sensitive areas.
- The beneficiary must be at least 20 years of age.

As indicated by the eligibility criteria, the program targeted rice farmers with some degree of experience with the crop, and also focused on small non-irrigated farms. Forcing farmers to own their own land might rule out some of the poorest households in the area, but this restriction makes sense in the context of plot-specific extension services if permanent increases in productivity are to be achieved. As will be discussed in more detail when describing the data set, these criteria were not strictly enforced in the first year, particularly with regards to land tenure status. This evaluation focuses on farmers who did satisfy program participation criteria.

Rice farmers interested in participating in the RBD program submitted requests for assistance to their cooperatives. The cooperatives then organized these requests into a single business plan that was submitted to the MCC office in Nicaragua for approval. The plans business plans themselves are at the cooperative level but are essentially collections of requests made by individual farms to participate in the RBD program. Whether or not an individual farmer participates in the program depends upon the decision made by MCC with regard to the business plan submitted by his or her cooperative.

3 Outcomes and parameters of interest

The goal of this evaluation is to estimate the average impact of the RBD program on participants; that is, the Average Treatment on the Treated (ATT) for a set of outcome variables. Altering the sample to exclude farmers not satisfying program criteria affects the interpretation of the ATT estimate, in that it will capture average effects on participants for the population satisfying program criteria. In addition, estimated impacts will capture effects on farmers who planted rice in 2009, rather than the entire population of farmers who meet program criteria; there are 243 such farmers in the sample of 300.[4]

I focus on three outcomes of interest: rice yields, revenues from growing rice, and net rice revenue; revenues minus costs are referred to as net revenue rather than profit since family labor is priced at the market wage, and it is not clear that this is equal to the opportunity cost of time for all households. Yields are measured in kilograms of output of unprocessed rice (i.e., wet and with the husk still attached) per sown hectare of land. Ignoring program costs not borne by the household, the ideal outcome of interest from the perspective of analyzing impacts of a program on welfare is arguably consumption.

While better measures of welfare exist than the outcomes listed above, there are good reasons for concentrating on agricultural variables. Firstly, the main goal of the program is to address poverty among small rice farmers in León and Chinandega. The program is designed to accomplish this by making information, credit, and high quality inputs available to farmers, thereby removing the constraints keeping them from becoming more commercially successful. If no increase in productivity, revenues, or net revenue were detected among participants, this would not necessarily indicate that households had failed to receive any benefit from the program. But it would at the very least suggest that the program had not worked as intended, and that more effective means of improving household welfare likely exist.

[4] Four farmers that were not members of eligible cooperatives reported being participants in the RBD program. Their names were cross-checked against databases maintained by MCA in Nicaragua, and this could not be verified. These households were dropped from the sample used in the analysis, leaving 243 rice planters. The results reported are robust to their inclusion, however.

Secondly, the survey data contain a measure of rice yields pre and post-RBD. As will be discussed later on in more detail, an implication of the identifying assumptions made in the econometric analysis of the RBD program presented here is that one should not detect any impact of the treatment on outcomes that could not have been affected by participation in the RBD program. For example, suppose we were to estimate the effect of participation in the RBD program on lagged yields. If the model has adequately controlled for differences between treatment and control households, we should detect no significant difference in pre-program yields across these two groups. If we do find a difference, this would strongly suggest the presence of unobservable factors correlated with RBD participation and the outcome of interest being modeled. Detecting significant differences would not require that we abandon all hope with regard to recovering unbiased estimates of program impacts, but it would suggest adjusting the modeling approach to incorporate information on outcomes prior to rollout of the RBD program.

Given that poverty reduction is the goal of the program, one could argue that the emphasis of this evaluation should be on changes in net revenue due to the RBD program, and that inclusion of yields and revenue as outcomes of interest is redundant. These latter two variables are included for two reasons: the availability of data on lagged yields, and because data on yields and revenue are likely to be of substantially higher quality than data on costs. Data collection was done through a single visit to each household shortly after the harvest. Gathering accurate data on input use would likely require multiple visits throughout the duration of the agricultural year. Furthermore, data on costs associated with technical assistance received outside of the RBD program are lacking. xxx I estimate impacts of the RBD program on profits both ignoring outside technical assistance, and valuing it the average price per farm for extension services given to RBD participants as reported in administrative data provided by MCC. The results are not sensitive to inclusion of this component of costs.

Furthermore, while the RBD program paid a maximum of 30 percent of the combined costs of the inputs used in loans and the agricultural extension services delivered to participating farmers, these savings were at the level of the cooperative; there is no information available on the distribution of these discounts across farmers. If we are willing to assume that farmers are able to sell inputs received from the program as they please, then market prices will represent the opportunity cost per unit of each input; given this assumption, using market prices to construct net revenue for the whole sample will not be problematic.

Extension services are more troublesome. Administrative data provided by MCC list an average cost for four months of extension services per farm of $20.00. In calculating profits, this figure was deducted out of revenues for each RBD participant. The bottom line is that there are sources of error in the available data for production costs, and these errors vary by treatment status. As a result, there is good reason to look at several indicators of program impacts on rice production, some of which may be more reliable than others.

4 Identifying assumptions and estimation technique

4.1 Inverse propensity score weighting

The evaluation of programs where participation is not random is complicated by the fact that outcomes of interest may be correlated with household characteristics which are also driving the participation decision. Suppose we would like to estimate the ATT of the RBD program for a given outcome of interest. Merely comparing participants and non-participants will yield a biased estimate of the ATT. For example, if more talented farmers have a higher probability of participating in RBD, and they also have higher crop yields, then a comparison of participant and non-participant households would attribute too great an effect to the RBD program; part of the observed difference in yields ought to be attributed to the difference in farming ability in the two groups.

Here I will attempt to control for these confounding factors via the Inverse Propensity Score-Weighted Least Squares method (IPS-WLS). The ATT is equal to the average outcome among the subsample of participants when receiving the treatment, minus the average outcome among this same group when the treatment is not available. This first average is observed in the dataset, but the second must be estimated using the subsample of non-participant households. In order to do so, the following assumption is made:

Assumption 1 - Unconfoundedness

Let y_i^1 represent the outcome of interest for household *i* when treatment is received and y_i^0 denote the outcome without participation in the RBD program. Let $d_i = 1$ represent membership in the treatment group and $d_i = 0$ for all non-participant households. Holding observed characteristics constant, the pair $\left[y_i^0, y_i^1 \right]$ is independent of selection into treatment. That is:

$$\left[y_i^1, y_i^0 \right] \perp d_i \mid \mathbf{x_i} = \mathbf{x} \tag{2.1}$$

This is known as the “unconfoundedness” assumption, and it states that potential outcomes are independent of participation in the RBD program conditional on holding $\mathbf{x_i}$ fixed, where $\mathbf{x_i}$ is the vector of observed characteristics (Imbens 2004). Within a group of observationally identical farmers, there are no confounding factors such as higher ability that might affect outcomes of interest while driving some individuals to participate in the RBD and others not to do so. Rather, after controlling for observable characteristics, whether a household is observed to be participating or not participating in the program is random.

As shown by Rosenbaum and Rubin (1983), (2.1) can be restated as follows:

$$\left[y_i^1, y_i^0\right] \perp d_i \mid p(\mathbf{x}) \tag{2.2}$$

where $p(\mathbf{x}) = P(d_i = 1 \mid \mathbf{x_i} = \mathbf{x})$ is the propensity score, or the probability of participating in the RBD program given the observed values of the $\mathbf{x}$ vector. In other words, if unconfoundedness holds, we can recover unbiased estimated of program impacts by conditioning on the scalar propensity score rather than the entire vector of observed characteristics.

In order to condition on the propensity score, an additional assumption must be made:

Assumption 2 - Overlap

$$0 < p(\mathbf{x}) < 1 \text{ for all } \mathbf{x}. \tag{2.3}$$

This is the overlap assumption, and it insures that there are treatment and control households at all values of $\mathbf{x}$ in the support of observable characteristics.

If there are no unobserved factors correlated with both the outcome of interest and selection into the RBD program, then it is only the distribution of observed characteristics along with treatment status that determines the average outcome in any given group. This suggests that we could recover an unbiased estimate of the average outcome without treatment among the group of participating households by applying weights to the subsample of control households. If the weights adjust the distribution of observed characteristics in the control group to reflect that of the treatment group, then the weighted average outcome among control group households would be an unbiased estimate of the average untreated outcome among households participating in the RBD program. This is the intuition behind using weights that are based on the probability of being in the treatment group given observed characteristics, i.e., weights based on the propensity score.

More formally, suppose we construct weights for households that did not participate in the RBD program that are equal to:

$$\frac{p(\mathbf{x})}{1-p(\mathbf{x})} \tag{2.4}$$

We then take the weighted expectation of the outcome *y* among untreated households, multiplied by $(1-d_i)$, holding the $\mathbf{x}$ vector constant:

$$\begin{aligned}
&E\left[\frac{p(\mathbf{x})}{1-p(\mathbf{x})} y_i(1-d_i) \mid \mathbf{x_i}=\mathbf{x}\right]= \\
&\frac{p(\mathbf{x})}{1-p(\mathbf{x})} E\left[y_i(1-d_i) \mid \mathbf{x_i}=\mathbf{x}\right]= \\
&\frac{p(\mathbf{x})}{1-p(\mathbf{x})} E\left[y_i^0(1-d_i) \mid \mathbf{x_i}=\mathbf{x}\right]= \\
&p(\mathbf{x}) E\left[y_i^0 \mid \mathbf{x_i}=\mathbf{x}\right]= \\
&p(\mathbf{x}) E\left[y_i^0 \mid d_i=1, \mathbf{x_i}=\mathbf{x}\right]
\end{aligned} \tag{2.5}$$

The second line is due to holding the **x** vector constant, and the third line comes from the fact that for control households the product of the observed outcome y_i and $(1-d_i)$ is equal to the product of the potential outcome y_i^0 and $(1-d_i)$. The fourth line stems from the fact that the propensity score is equal to the expected value of d_i holding the **x** vector constant. The final term follows from unconfoundedness, i.e., the average untreated outcome conditional on **x** ought to be equal regardless of the decision to select into treatment. By the law of iterated expectations, taking the expected value of this last term over the distribution of **x** yields the average untreated outcome among participating households in the absence of the RBD program, $E\left[y_i^0 \mid d_i = 1\right]$.

Equation (2.5) can be estimated using the observed outcomes among the control households, and an estimate of the propensity score. Suppose the population-level model for the decision to enroll in the RBD program follows a logit specification. Then we can write down the propensity score as:

$$p(\mathbf{x}) = \frac{\exp(\pi_0 + \mathbf{x}_\mathbf{i}'\boldsymbol{\pi})}{1 + \exp(\pi_0 + \mathbf{x}_\mathbf{i}'\boldsymbol{\pi})} \tag{2.6}$$

Plugging the logit equation into the equation for the weights given in (2.4) yields:

$$\frac{p(\mathbf{x})}{1 - p(\mathbf{x})} = \exp(\pi_0 + \mathbf{x}_\mathbf{i}'\boldsymbol{\pi}) \tag{2.7}$$

Note that $\mathbf{x}_\mathbf{i}$ could include interactions and higher order terms based on a smaller set of observable characteristics, and as a result the linear term in parentheses in (2.6) need not be restrictive. Once the parameters of (2.6) are estimated, the fitted values $\hat{p}(\mathbf{x})$ are used to construct the weights given in (2.4), and the ATT can be estimated as:

$$\frac{\sum_{i=1}^{N} y_i d_i}{\sum_{i=1}^{N} d_i} - \frac{\sum_{i=1}^{N} y_i (1 - d_i)}{\sum_{i=1}^{N} (1 - d_i)} \frac{\hat{p}(x)}{1 - \hat{p}(x)} \tag{2.8}$$

This is the difference in two sample averages. The first term is the average outcome among the treated households in the sample, and the second term is the sample version of the term in brackets in the first line of (2.5). The difference given in (2.8) will be a consistent estimator of the ATT if the model for the propensity score is correct and a law of large numbers can be applied to the two averages that appear in the formula.

4.2 Linear regression

Inverse propensity score weighting only yields consistent estimates of program impacts if we have the correct model for the propensity score. We may be more confident in our ability to construct a correct regression model for the conditional expectation of a given outcome of interest than in our ability to model the selection process. It turns out that inverse propensity score weighting and regression can be combined in a manner that yields an unbiased and consistent estimate of the ATT, as long as either the model for the propensity score or the regression model of the conditional expectation of the outcome is correct; this is the "double robustness" property of inverse propensity score weighted least squares (IPS-WLS) estimation (Wooldridge 2007).

Consider the following regression model for the conditional expectation of the outcome variable y_i among the group of RBD program participants:

$$\begin{aligned} E\left[y_i^0 \mid d_i = 1, \mathbf{x}\right] &= \alpha_0 + \left(\mathbf{x_i} - \boldsymbol{\mu}\right)' \boldsymbol{\alpha}_2 \\ E\left[y_i^1 \mid d_i = 1, \mathbf{x}\right] &= \alpha_0 + \alpha_1 + \left(\mathbf{x_i} - \boldsymbol{\mu}\right)' \boldsymbol{\alpha}_{2'} \end{aligned} \tag{2.9}$$

The first line of (2.9) specifies the conditional expectation of yields for the group of RBD participants in the absence of the RBD program, and the second line is the conditional expectation of yields for this same group when its members actually participate. Here it is assumed that the $\mathbf{x_i}$ vector that appears in (2.9) is identical to that of (2.6), although there is no need for this to be the case. The vector $\boldsymbol{\mu}$ contains the means of the $\mathbf{x_i}$ variables within the population of participants. The parameter vector $\boldsymbol{\alpha}_2$ is the derivative of the conditional mean of the outcome with respect to the $\mathbf{x}_i$ vector, and it captures how the conditional expectation changes in the absence of treatment as $\mathbf{x}_i$ moves away from its mean. The vector $\boldsymbol{\alpha}_{2'}$ captures this same effect when treatment is received; any difference between $\boldsymbol{\alpha}_{2'}$ and $\boldsymbol{\alpha}_2$ can be attributed to interaction effects between the treatment and observed characteristics.

By the law of iterated expectations, taking the expectation of the first line of (2.9) over the distribution of $\mathbf{x}$ gives the expected value of y_i for the group of participants when not enrolled in the RBD program, while the expected value of the outcome for the group of participants when the treatment is received can be derived similarly using the second line. The difference between these two expectations is the ATT, α_1.

4.3 The double robustness property of inverse propensity score weighted least squares regression

Given the assumption of unconfoundedness, $E\left[y_i^0 \mid d_i = 1, \mathbf{x}\right] = E\left[y_i^0 \mid d_i = 0, \mathbf{x}\right]$, and the first line of (2.9) can be replaced with an equivalent expression that uses the population of non-participant households. This makes it possible to combine the two lines of (2.9) as:

$$E\left[y_i \mid \mathbf{x}\right] = \alpha_0 + d_i\alpha_1 + \mathbf{x}_i'\boldsymbol{\alpha}_2 + d_i\left(\mathbf{x}_i' - \boldsymbol{\mu}\right)'\boldsymbol{\alpha}_3 \qquad (2.10)$$

The ATT is still given by α_1. The vector $\boldsymbol{\alpha}_2$ is interpreted as before, and the sum of $\boldsymbol{\alpha}_2$ and $\boldsymbol{\alpha}_3$ is equal to $\boldsymbol{\alpha}_{2'}$ in (2.9). If the conditional expectation of y_i is indeed equal to (2.10), then the ordinary least squares estimate $\hat{\alpha}_1$ will be consistent for the ATT. Furthermore, we can apply weights to the data and estimate the parameters of (2.10) via weighted least squares. The consistency of $\hat{\alpha}_1$ will be unaffected when the regression model is the correct one for the conditional expectation (Greene 2003, 226).

If the conditional mean is not linear, but we have the correct model for the propensity score, $\hat{\alpha}_1$ will still be a consistent estimate of the ATT if it is estimated via weighted least squares, where the weights for non-participant households are given by (2.4) and the true propensity score is replaced by its estimate. To see why, assume without loss of generality that there is only a single covariate, *x*. The weighted least squares formula for the intercept among treated households is:

$$\hat{\alpha}_0 + \hat{\alpha}_1 = \frac{\sum_{i=1}^{N} y_i d_i}{\sum_{i=1}^{N} d_i} - \hat{\alpha}_2 \frac{\sum_{i=1}^{N} x_i d_i}{\sum_{i=1}^{N} d_i} \tag{2.11}$$

The interaction between d_i and $x_i - \bar{X}$ has dropped out because the latter is evaluated at $x_i = \bar{X}$ when solving for the intercept, where $\bar{X}$ is the average of *x* among RBD participants. The probability limit of the first term of (2.11) is the expected value of the treated outcome among households enrolled in the RBD program. The second term converges in probability to:

$$E\left[x_i d_i\right] = E\left[xE\left[d_i \mid x_i = x\right]\right] = E\left[xp\left(x\right)\right] \tag{2.12}$$

The intercept formula for non-participant households is:

$$\hat{\alpha}_0 = \frac{\sum_{i=1}^{N} y_i \left(1-d_i\right)}{\sum_{i=1}^{N}\left(1-d_i\right)} \frac{\hat{p}\left(x\right)}{1-\hat{p}\left(x\right)} - \hat{\alpha}_2 \frac{\sum_{i=1}^{N} x_i \left(1-d_i\right)}{\sum_{i=1}^{N}\left(1-d_i\right)} \frac{\hat{p}\left(x\right)}{1-\hat{p}\left(x\right)} \tag{2.13}$$

Assuming that $\hat{p}(x) = p(x)$, the probability limit of the first term is the expected value of the untreated outcome among households enrolled in the RBD program. By Slutsky's theorem, the probability limit of the second term is equal to the probability limit of $\hat{\alpha}_2$ multiplied by:

$$E\left[x_i(1-d_i)\frac{\hat{p}(x)}{1-\hat{p}(x)}\right] = E\left[xE\left[(1-d_i)\frac{p(x)}{1-p(x)} \mid x_i = x\right]\right] = E\left[xp(x)\right] \tag{2.14}$$

The second terms on the right hand side of each intercept formula are asymptotically equivalent. Taking the difference between the probability limits of the two intercepts therefore causes the second term to drop out of each, leaving:

$$\widehat{\alpha_1} \xrightarrow{p} = E\left[y_i^1 \mid d_i = 1\right] - E\left[y_i^0 \mid d_i = 1\right] = \text{ATT} \tag{2.15}$$

where y_i^1 and y_i^0 are the potential outcomes with and without treatment, respectively.

4.4 Estimation and inference

Estimating the parameters of the regression model using the IPS-WLS technique is straightforward. First, the logit model is estimated via maximum likelihood, and the fitted values of the propensity score are used to construct the weights for non-participant households. Next, the parameters of the regression model, including the ATT, are estimated by minimizing the weighted sum of squared residuals. Define **w** as the **x** vector augmented to include unity, and **z** as the **x** vector expanded to include unity, the treatment indicator d_i, and the de-meaned covariates used in the regression model. Using this more compact notation, the objective function for the weighted regression can be written as:

$$\sum_{i=1}^{N}\frac{1}{N}\left[d_i + (1-d_i)\exp\left(\mathbf{w}_\mathbf{i}'\hat{\boldsymbol{\pi}}\right)\right]\left[y_i - \mathbf{z}_\mathbf{i}'\hat{\boldsymbol{\alpha}}\right]^2 \tag{2.16}$$

The first term in brackets in (2.16) follows from the fact that the weights for non-participant households simplify to $\exp\left(\mathbf{w}_{\mathbf{i}}'\hat{\boldsymbol{\pi}}\right)$.

How to conduct statistical inference on $\hat{\boldsymbol{\alpha}}$ is less obvious, for two reasons. xxx if first order conditions for alpha depend on pi, must adjustust asump variance. second, nature of outcomes in this study may mean that there is dependence of the residuals across i.

Taking derivatives of and (2.16) with respect to the parameters of each gives the following set of first order conditions:

$$\sum_{i=1}^{N}\frac{\mathbf{z}_{\mathbf{i}}}{N}\left[d_i+\left(1-d_i\right)\exp\left(\mathbf{w}_{\mathbf{i}}'\hat{\boldsymbol{\pi}}\right)\right]\left[y_i-\mathbf{z}_{\mathbf{i}}'\hat{\boldsymbol{\alpha}}\right]=0 \tag{2.17}$$

If there are *m* parameters in the logit model and *l* parameters in the regression, then equations **Error! Reference source not found.** and is and (2.17) are $\left(m\times1\right)$ and $\left(l\times1\right)$ vectors of first order conditions, respectively. Solving these equations for $\hat{\boldsymbol{\pi}}$ and $\hat{\boldsymbol{\alpha}}$ generates the estimates of the model parameters.

To find the variance of each element of $\hat{\boldsymbol{\pi}}$ and $\hat{\boldsymbol{\alpha}}$, we take a first order Taylor expansion around the true values of the parameters of each model, multiply by $N^{1/2}$, and rearrange terms to yield:

$$\sqrt{N}\left(\boldsymbol{\pi}-\hat{\boldsymbol{\pi}}\right)=\left[\sum_{i=1}^{N}\frac{1}{N}\mathbf{H}_{\mathbf{i}}\right]^{-1}\sum_{i=1}^{N}\frac{\mathbf{w}_{\mathbf{i}}}{\sqrt{N}}\left[d_i-\frac{\exp\left(\mathbf{w}_{\mathbf{i}}'\pi\right)}{1+\exp\left(\mathbf{w}_{\mathbf{i}}'\pi\right)}\right] \tag{2.18}$$

$$\sqrt{N}\left(\boldsymbol{\alpha}-\hat{\boldsymbol{\alpha}}\right)=\left[\sum_{i=1}^{N}\frac{1}{N}\mathbf{G}_{\mathbf{i}}\right]^{-1}\sum_{i=1}^{N}\omega_i\frac{\mathbf{z}_{\mathbf{i}}}{\sqrt{N}}\left[y_i-\mathbf{z}_{\mathbf{i}}'\boldsymbol{\alpha}\right] \tag{2.19}$$

where ω_i are the weights, and $\mathbf{H_i}$ and $\mathbf{G_i}$ are the Hessian matrices for the logit and weighted least squares objective functions, respectively. Note that both Hessian matrices include derivatives of the first order conditions with respect to the $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$ vectors, i.e., the $\mathbf{H_i}$ and $\mathbf{G_i}$ terms are $(l+m)\times(l+m)$ matrices.

Applying the variance operator to (2.18) and (2.19) gives:

$$\left[\sum_{i=1}^{N}\frac{1}{N}\mathbf{H_i}\right]^{-1} Var\left[\sum_{i=1}^{N}\frac{\mathbf{w_i}}{\sqrt{N}}\left[d_i-\frac{\exp(\mathbf{w_i'}\pi)}{1+\exp(\mathbf{w_i'}\pi)}\right]\right]\left[\sum_{i=1}^{N}\frac{1}{N}\mathbf{H_i}\right]^{-1} \tag{2.20}$$

$$\left[\sum_{i=1}^{N}\frac{1}{N}\mathbf{G_i}\right]^{-1} Var\left[\sum_{i=1}^{N}\omega_i\frac{\mathbf{z_i}}{\sqrt{N}}\left[y_i-\mathbf{z_i'}\boldsymbol{\alpha}\right]\right]\left[\sum_{i=1}^{N}\frac{1}{N}\mathbf{G_i}\right]^{-1} \tag{2.21}$$

Wooldridge: IPW estimator is asymptotically normal, with a variance that is weakly lower when the estimated propensity score is used rather than the true propensity score, implying that we can ignore the fact that the p-score is estimated when conducting inference. The formula given for the variance of the estimator assumes that the observations are independent draws. This could be adopted for the case of dependent observations by taking the sum of the score function within each cluster, and then taking the plugging in these super observations into the variance formula. The sum of the outer product of the scores would then be done over the number of clusters, and then divided by the number of clusters. But with spatial correlation the problem is more complex. If we could divide the sampling area into non-overlapping regions within which unobservables may be correlated, but with zero correlation in the residuals between the regions, then this same methodology would apply. But completely eliminating correlations in unobservables at the borders is obviously a strong and unrealistic assumption.

If observations could be ordered as in a time-series, then robust inference could be conducted with minimal assumptions about the nature of spatial dependence by applying the Newey-West variance-covariance matrix formula to the residuals generated by the weighted least-squares estimation procedure described above. The Newey-West formula assumes that the covariance between any two residuals in a time series is a decreasing function of the number of periods between them, and is assumed to be zero beyond a given number of lags that must be selected by the econometrician. The formula for the Newey-West estimator is:

The variance-covariance matrix estimator consists of two terms: an average of the variances for each of the residuals with themselves, and a weighted average of the covariances between each residual and those falling before it in the time series. The weighting function takes a values of 1 for lags of a single period, and then decreases in steps of 1/(1+p) to zero for lags greater than *p*. is a decreasing in the size of the lag, and is equal to zero for lags greater than *p*. The Newey-West estimator is robust to arbitrary heteroscedasticity at the level of each individual observation, as well as autocorrelation of any structure, assuming that dependence between residuals is truly a decreasing observation of distance and the value *p* chosen by the econometrician is large enough to capture the underlying lag structure.

Spatial dependence is more complex that autocorrelation within a time series. While the distance between observations in a time series can be captured by a number of equidistant steps on a line, spatial data will be located in two or three dimensions, with distances between neighboring observations varying across the sample space. Several studies have adapted the Newey-West procedure to spatial data, taking these unique characteristics into account. Examples include Conley (1999), who replaces the weight in the Newey-West formulation with a two-dimensional equivalent, and Kelejian and Prucha (2007), who use the Euclidean distance between observations as the argument in a Gaussian kernel which serves as the weighting function. Analogous to the choice of *p* in the Newey-West formula, these studies choose a maximum distance outside of which the residuals of any two observations are assumed to be uncorrelated.

Hypothesis testing using Newey-West style standard errors in a spatial context relies upon the asymptotic normality of model parameters. It is unclear to what extent these asymptotic results will hold when using small samples. Firstly, robust standard errors will have higher sampling variability than conventional standard errors, and in small samples such as the one used in this paper, may lead to over rejection of the null hypothesis of no program effect. Secondly, statistical tests of the null hypothesis of no effect based on asymptotic theory generate significance levels that are biased relative to the true significance level of the test (Cameron, 2005, Chapter 11). For example, in a one-tailed test, we may think that we are rejecting the null at a 5 percent significance level if the calculated t-statistic exceeds the 95^{th} percentile of the standard normal distribution. But the true significance level may be larger. Thus it is no surprise that the performance of robust standard errors is mixed in finite samples. In the context of spatial data, Bester et al. () present simulation evidence with a sample size of over 600 that normal approximations to the true parameter distribution lead to substantial over-rejection of the null hypothesis of no effect. They present an alternative method of variance estimation based upon .

Alternatively, robust standard errors could be used to construct a test statistic which is asymptotically pivotal (i.e., does not rely on unknown parameters) which can then be bootstrapped. The

Conley (1999) constructs an index of "economic distance" for each observation in a data set; the example used in the application presented in the paper is the cost of transporting physical capital between countries. This makes it possible to reduce the dimensionality of spatial data on place observations on a continuous, one-dimensional index. Construction of Newey-West style standard errors is straightforward. Kelejian and Prucha (2007) assume that the influence of the residual of one observation on that of another is a decreasing function of distance, and is equal to zero outside of a certain range. This range is akin to the maximum lag length used in construction of Newey-West standard errors. Kelejian and Prucha then construct a Newey-West style variance-covariance matrix estimator, where the weights in the covariance term are given by the Gaussian kernel with a bandwidth equal to the maximum range within which the residual of one observation may be correlated with that of another.

5 Results

5.1 Data

The sample was drawn from lists of rice producers provided by farmer cooperatives participating in the RBD program. These lists were pooled into a single database of farmers belonging to the 11 cooperatives originally chosen to participate in the RBD program and thought to satisfy the criteria listed in above in section 2.2 for program participation. Of these 11 cooperatives, one was eliminated because it had dropped out of the program partway through the agricultural year, and two others were eliminated because no names of non-participants farmers were made available. The remaining eight cooperatives served as the basis of the sample.

During the process of data collection, a large number of farmers were replaced at the request of MCC due to not satisfying program eligibility criteria; the program was to last for two years, but farmers found to violate program criteria would be disqualified in their first year of participation. MCC wanted to maintain the option of conducting a second round of data collection, and in order to avoid high rates of attrition in the panel or shifting the split between treatment and control observations sharply towards the latter in the second year, it was decided that farmers not meeting program criteria would be dropped from the sample. Nearly 50 percent of the original sample had to be replaced, with the most common cause being failure to satisfy program criteria with respect to land tenure status, followed by households being listed more than once on the roster provided by MCC. To round out the sample, a small number of farmers not belonging to cooperatives but satisfying other program criteria were interviewed; enumerators located a number of such farmers in the field, and a random subsample of this group was chosen to be interviewed.

The data were collected in a single household visit shortly after the post-harvest stage of the agricultural calendar, allowing sufficient time for farmers to have marketed their production of rice. To estimate program impacts, we must adjust for observable differences between treatment and control households. The danger of using data collected after the intervention is that we will hold variables constant that were affected by the treatment and are correlated with outcomes of interest;[5] this would eliminate a portion of the impact from the estimated effect, and potentially introduce other sources of bias (Rosenbaum 1984). Recall questions were asked about purchases and sales of consumer durables, agricultural implements, and land in order to reconstruct the wealth of each household prior to implementation of the RBD program. These are major sources of wealth and it seems reasonable to expect households to remember substantial changes in asset holdings over a one year period.

[5] I will use the term “outcome of interest” to refer to any variable for which we might want to measure impacts of program participation.

These data were used to construct indices of agricultural and non-agricultural wealth via Principal Components Analysis (PCA). The indices explain 26.23 percent and 30.25 percent of variation in agricultural and non-agricultural wealth in the sample, respectively.[6] For data on the agricultural year immediately prior to the RBD program, households were asked about loans taken out for agricultural activities, changes in household membership and demographics, sown area of marketed crops, and rice production. Other potential explanatory variables, such as non-agricultural and unearned income, geographic location, sown rice area suffering unanticipated production shocks, and expectations regarding rice production levels enter into the different models at their reported levels for the 2009-2010 agricultural year.

5.2 Impact on rice yields

5.3 Impact on rice revenue

5.4 Impact on rice net revenue

5.5 Costs

6 Robustness checks

6.1 Indirect test of identifying assumptions

Agricultural extension often has a large public goods aspect to it (Anderson and Feder 2004). If this were true in the context of the RBD program, then this would violate the assumption of no externalities. This assumption can be tested by estimating the impact of non-participation among eligible farmers, i.e., comparing the average outcome of eligible non-participants with the counterfactual outcome among this same group that would have obtained had the RBD program never existed. The latter is estimated using the group of ineligibles.

[6] PCA maps variables into a series of orthogonal components explaining successively smaller shares of the total variation of whatever is being indexed. Härdle (2007) offers a more detailed explanation of PCA with examples of applications.

Estimate two sets of results: One with coop members and coop fixed effects. Another with the whole sample.

7 Seeking an explanation for the results

8 Conclusion