Food Policy 44 (2014) 237–247

Contents lists available at ScienceDirect

Food Policy

journal homepage: www.elsevier.com/locate/foodpol

The impact of high value markets on smallholder productivity in the Ecuadorean Sierra: A Stochastic Production Frontier approach correcting for selectivity bias ☆,☆☆

Mario González-Flores [a,*], Boris E. Bravo-Ureta [b,c,d], Daniel Solís [e,f], Paul Winters [a]

[a] College of Arts & Sciences, Department of Economics, American University, USA
[b] Department of Agricultural and Resource Economics, University of Connecticut, USA
[c] Department of Economics, University of Connecticut, USA
[d] Department of Agricultural Economics, University of Talca, Chile
[e] Division of Agricultural Sciences, CAFS, Florida A&M University, USA
[f] Division of Marine Affairs and Policy, RSMAS, University of Miami, USA

ARTICLE INFO

Keywords:
Stochastic frontiers
Technical efficiency
Sample selection
Impact evaluation
Propensity score matching
Ecuador

ABSTRACT

This paper uses data from small-scale potato farmers in Ecuador to examine the impact of the program *Plataformas de Concertación* on productivity growth. Using propensity score matching combined with a Stochastic Production Frontier model that corrects for sample selection bias, we disaggregate the yield growth attributable to the program into technological change (TC) and technical efficiency (TE). While the results do not exhibit a clear indication of selection bias, the analysis does show that on average beneficiaries exhibit higher yields than control farmers given the same input levels, but lower TE with respect to their own frontiers. These results suggest that while the program raised the technology gap in favor of beneficiaries, it had a negative effect on TE in the short run. The latter finding is consistent with the notion that beneficiaries enjoyed a significant change in production techniques, but it is very likely that they were still in the "learning by doing" stages at the time the data was collected. In fact, the results suggest a fast recovery in TE levels on the part of beneficiaries as time with project increased.

Introduction

Agricultural projects often seek to improve productivity with the expectation that such improvements would lead to higher income and welfare among beneficiaries. Examples of interventions include the introduction of new seed varieties, the adoption of new farming techniques (such as integrated pest management or IPM), linking farmers to markets, better accounting practices, provision of extension services, farmer field schools (FFS), or a combination of various actions. The effective use of a newly adopted technology requires investing the time and effort to become acquainted with the new practices before the full benefits of adoption can be felt by the farmer. This may entail a process of trial and error during several agricultural cycles. While adopting new techniques or inputs can potentially lead to increases in production at the end of an agricultural cycle, this does not necessarily mean that the new procedures are being implemented in an efficient manner. This is particularly true for smallholders, who are typically characterized by having lower levels of education, living in isolated rural areas, and having limited access and exposure to information and markets. Thus, much of the innovative content and techniques can be quite foreign and may entail a challenging process for this type of farmers. Therefore, when evaluating the impact of an agricultural project, it is important to differentiate between indicators of technological change (TC) versus managerial performance (or technical efficiency, TE).

The economic impact evaluation literature has been growing in recent years, and this growth has been mainly focused on the social sectors where the indicators of impact tend to be more easily identifiable (Winters et al., 2011). Rigorous impact evaluations of agricultural projects have been relatively scarce and the evidence on the effectiveness of such projects in developing countries is mostly inconclusive (IDB, 2010; Del Carpio and Maredia, 2009). The relative scarcity of formal evaluations of agricultural projects is likely

☆ This paper was part of M. González-Flores' Ph.D. dissertation in the Department of Economics at American University, USA.

☆☆ Data collection for this study was funded by the FAO-Netherlands Partnership Program and the FAO Norway Partnership Program, and was made possible through the support of FAO-ESA in Rome, farmers and leaders of CONPAPA in Ecuador, the International Potato Center and its Papa Andina Partnership Program, INIAP-Ecuador, the Swiss Agency for Development and Cooperation and its FORTIPAPA Project, CESA-Ecuador, Fundación M.A.R.CO. and IEDECA.

* Corresponding author. Address: College of Arts & Sciences, Department of Economics, American University, 4400 Massachusetts Avenue, N.W. Washington, DC 20016-8029, USA.

E-mail address: mgvoz@aol.com (M. González-Flores).

0306-9192/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.foodpol.2013.09.014

due to several reasons. First, agricultural projects are generally designed to increase output and therefore impact evaluations focus on production-based indicators, typically associated with TC. However, collecting this type of data can be challenging, beginning with the definition of the sample unit, since production is often linked to multiple plots but the decision-making process takes place at the household level. The challenge is greater when attempting to evaluate the impact of a project on different types of households, such as smallholders and large holders, who often have very distinct production systems (Winters et al., 2010).

Second, in analyzing agricultural production, the relationship between inputs and outputs or profitability is often examined through gross margins or total value product functions. Yet, presumably, agricultural projects have an impact not just on inputs and outputs, but also on how these inputs are used and combined. Whether these inputs are being used in an efficient manner to obtain the maximum possible levels of output needs to be considered in an evaluation (Winters et al., 2010). Yet, this is rarely done since, as noted, most project evaluations focus on TC indicators. While such focus allows the researcher to identify impact on different components of production, it does not provide any information on whether farmers made the right use of the available inputs and technology at their disposal, i.e., managerial performance is ignored.

Given these difficulties, combining Stochastic Production Frontier Analysis (SPFA) with impact evaluation methodologies provides a useful avenue for measuring the productivity impact of agricultural projects. SPFA is a widely used econometric technique that estimates the 'best practice' relationship between inputs and output of the farm households in the sample. In addition, SPFA can help identify the levels of efficiency (or inefficiency). Therefore, this approach makes it possible to quantify the potential to increase agricultural output without the need for additional inputs or new technology (Coelli et al., 2005).

Papa Andina, the focus of this paper, is a partnership that worked to address rural poverty in the Andean highlands by fostering innovation and market development for potatoes. The approach recognizes that while agricultural research is a main driver of TC and agricultural development in addressing rural poverty, this research needs to be linked to practical improvements in value chains that are important to smallholders (Horton et al., 2011). A key program within *Papa Andina* is the *Plataformas de Concertación*, hereafter *Plataformas*. This program offered a space for public and private sector partnerships where diverse actors—including farmers, potato processors, supermarkets, national research institutes, universities and non-governmental organizations—could work together to innovate and link small-scale potato producers to commercial interests. *Plataformas* offered a mechanism not just to support agricultural research in the field, through new varieties and different mechanisms to enhance production and marketing, but also served as an experiment in institutional innovation. The question is whether this approach is an effective mechanism to increase farmer production and efficiency and this constitutes the overall goal of this paper.

In implementing impact evaluations of development projects, several researchers have promoted the use of randomized experiments (Duflo et al., 2008). However, it is often the case that experimental designs are costly and difficult to implement; thus, one needs to rely on non-experimental methods (Barrett and Carter, 2010). One common non-experimental approach to assessing project impact is propensity score matching (PSM), which alleviates biases stemming from observable variables (World Bank, 2006). However, in projects where beneficiaries self-select, unobservable variables (e.g., managerial ability) can also be a source of bias. If panel data are available, fixed effects estimators along with PSM can be used to deal with the problem, provided that the unobservables are time invariant (Angrist and Pischke, 2009). Thus, the generation of a counterfactual along with the mitigation of biases from observables and unobservables can be addressed in non-experimental designs as long as one has data on treatment and control groups at both the baseline and the endline. Recent applications of this methodology to agricultural projects include the work of Bravo-Ureta et al. (2011) in Honduras and Cerdán-Infantes et al. (2008) in Argentina.

A challenge frequently encountered in the field is that analysts and/or policy makers might be interested in measures of impact even when baseline data is not available. In such a situation, which is the case for *Plataformas*, one needs to rely on cross-sectional data along with suitable matching procedures and other econometric techniques, such as instrumental variables, in order to obtain the desired impact measures (Cavatassi et al., 2011a). In this paper, we are particularly interested in separating the effect of TC and TE on farm productivity. To achieve our goals we make use of cross-sectional data collected after the project was underway. A number of steps were taken to ensure that the data on treatment and control groups would have been very similar at the baseline, but it is likely that selection issues still remain. Therefore, we address possible self-selection in a stochastic frontier context using the model recently introduced by Greene (2010) and adapted to the evaluation of development programs by Bravo-Ureta et al. (2012).

In sum, a key feature of this paper is bridging SPFA with impact evaluation methods. Development projects often have a major component intended to improve decision-making and managerial ability along with the transfer of technologies designed to increase output. Thus, for such projects, SPFA methodologies are ideally suited to decompose productivity growth into technological and managerial components; however, these methodologies have hardly been applied for this purpose. A major reason for the absence of such applications is likely to be the challenges posed by selectivity bias, which is a common feature in development projects. In this fashion, this paper adds to a very limited but emerging literature that combines SPFA modeling with impact evaluation techniques.

The remainder of the paper is structured as follows. Section 2 provides an explanation of *Plataformas*, and a description of the data is provided in Section 3. Section 4 presents the analytical framework for analyzing TC and TE and the closely related empirical strategy. Section 5 provides the results and Section 6 concludes.

Plataformas de Concertacion[1]

The *Plataformas* are multi-stakeholder alliances, which bring farmers together with a range of agricultural support service providers, including INIAP (*Instituto Nacional Autónomo de Investigaciones Agropecuarias*), local NGOs, researchers, universities, and local governments. The *Plataformas* pay special attention to expanding the direct participation of low-income farmers in high-value producer chains by providing them with new technologies, by promoting their organizational skills and social capital, and by involving them in a "value chain vision" of production and commercialization that directly links them to the final output markets, thus circumventing intermediaries (Cavatassi et al., 2009). As noted by Devaux et al. (2009, p. 36), "this facilitates knowledge sharing, social learning and capacity building, leading to improvements in small farmer productivity and the quality of potatoes supplied to market." The overall objective of the *Plataformas* is then to "reduce poverty and increase food security, through increasing

[1] More information on the different aspects and activities of *Plataformas* can be found in Cavatassi et al. (2009).

yields and profits of potato smallholders" (Pico, 2006, as cited in Cavatassi et al. (2009), p. 8).

A central component of *Plataformas* was the training provided at FFS, which involved lectures and applied exercises in experimental plots. FFS emphasized improved production technologies and IPM techniques aimed at enhancing the quality and quantity of production. Farmers were taught techniques to manage soil, seeds (renewing and stock selection), insects (Andean weevil and tuber moths), diseases (late blight), and agrochemicals (insecticides, fungicides, pesticides, and fertilizers) efficiently. An important element under the IPM approach was to reconcile improvements in production with the use of these techniques while preserving the environment and protecting human health. The training also exposed farmers to the quality requirements of high-value markets.

Previous evaluations have found positive and significant impacts of the program on technical indicators, such as yields, profits, and gross margins (Cavatassi et al., 2011a,b). However, an issue that has not been evaluated is whether the program had an effect on managerial performance or TE, i.e., are beneficiary farmers producing closer to their best practice production frontier than non-beneficiaries? Given that all participants in the *Plataformas* received the same comprehensive 'package' of 'treatments', one would expect that the program would have a positive effect on farmers' TE. On the other hand, if much of this new information and technologies were too complex to process, comprehend, and master right away, then one would expect that while farmers might attain higher output, TE might remain about the same or might even decrease. The lack of improvement or even a lower TE is plausible if farmers had exposure to the program for only a short period and thus not enough time may have elapsed for them to become proficient with the new methods. Hence, the actual effect of the *Plataformas* on the TE of beneficiaries is an empirical question.

Data

The data used in this paper comes from "*La Nueva Economía Agrícola*" household and community level surveys implemented by the Food and Agriculture Organization (FAO) in collaboration with the International Potato Center (IPC). The *Plataformas* were active in the Ecuadorean provinces of Chimborazo and Tungurahua and surveys were administered in these provinces from June to August of 2007, and they encompass information for the year prior to the survey. Data was collected at the plot, household and community levels.

To ensure a robust evaluation of the impact of the program, the data collection was done in multiple steps and the counterfactual was designed with a great deal of care. First, communities participating in the *Plataformas* program (treated communities) were identified in each of the two provinces and information on these was obtained. Second, using population and agricultural census data, the treated communities and a set of potential control communities that did not participate in the program but who had similar geographic, agro-ecological and socio-demographic characteristics were identified. This provided a list of all possible treatment and control communities to be included in the survey. Third, using PSM, control communities that were most comparable to treated communities from a statistical standpoint were determined. Fourth, the resulting list of potential control communities was discussed with key local organizations that had a central role in *Plataformas* to determine if they were indeed comparable to the treated communities. Some of the key characteristics considered were production and agro-ecological similarities and levels of community and farmer organization. Thus, the selection of treatment and control communities that resulted from the application of PSM was fine-tuned by local agronomists and leaders of organizations with local knowledge. Through this process, the most appropriate control communities were identified. Further, treated communities for which a suitable control community could not be found were excluded from the sample (three in total).

Once the communities for inclusion in the sample were determined, lists of households from treatment and control communities were obtained by *Plataformas* coordinators and community leaders to randomly select those to be included in the final sample. The lists from treated communities included households who participated in the program and those that did not. The final sample includes a total of 35 communities (18 treatment and 17 controls) and contains 1007 households that were randomly selected from control communities and among participants and non-participants in treated communities. Given the focus of this paper, the analysis is restricted to farms that had at least one full potato production cycle (from planting to harvesting) and this final sample consists of 495 households. A test was conducted to check for systematic differences between treated and control farms in their probabilities of having had a complete production cycle and no differences were found.

The data used in this paper makes it possible to create alternative counterfactuals including: (1) non-beneficiaries living in treated communities (non-participants); (2) non-beneficiaries living in non-treated communities (non-eligible); and (3) a combination of 1 and 2. There are two concerns with option 1. First, non-participation is an explicit choice made by those producers and that choice might reflect that they are different from participants, which could lead to self-selection bias, i.e., the estimates may reflect fundamental differences between the two groups rather than the impact of the program. Second, since non-participants live in close proximity to beneficiaries they may obtain indirect benefits from the program (spillover effects). Thus, using non-participants as a control group could be problematic. Nevertheless, this is a potentially useful group because their observable characteristics are likely to be very similar to those of participants since they live in the same communities. Further, previous analysis of these data indicates that there is little evidence of spillover effects (Cavatassi et al., 2009).

The non-eligible farmers can be considered as a 'pure' counterfactual since spillover effects are unlikely. Importantly, given that the program was not offered in these communities, non-eligible farmers are not subject to possible self-selection bias, although they may exhibit program placement bias if the process used to select control communities is imperfect (Baker, 2000). As explained below, the final sample uses option 3 including 340 households who reside in treated communities (171 beneficiaries and 169 non-participants), and 155 in non-treated communities (non-eligible).

The survey has multiple modules containing several variables well suited for the prediction of participation into the program, such as information on land (number of plots, soil quality), socio-demographic variables (household composition and head of household characteristics), welfare and assets, social capital, and community-level variables. Table 1 presents descriptive statistics for the entire sample (pooled) and a test of difference in means between beneficiaries, non-participants and non-eligible as well as all control farmers together (non-participants and non-eligible).[2] Table 1 shows that, as expected, non-participants (column V) have very similar characteristics as beneficiaries (column II)—the two groups exhibit statistically significant differences for only 4 out of 28 variables. The only category in which the two groups differ is

[2] The variables included are theoretically exogenous, i.e., they were not affected by *Plataformas*.

Table 1
Description of the data.

Variables (units)	I Pooled	II Benef.	III All controls	Test[a]	IV Non-elig.	Test	V Non-partic.	Test
Altitude (m)	3475	3435	3497	*	3497		3496	
Land owned (ha)	2.77	2.75	2.78		3.33		2.27	
Owned plots (#)	3.04	3.37	2.86	***	3.11		2.63	***
Black soil (%)	81%	77%	83%	**	86%	**	81%	
Flat land (%)	39%	38%	39%		37%		40%	
Irrigated land (%)	60%	57%	62%		63%		62%	
Family size	4.71	4.85	4.64		4.66		4.62	
Max educ. in HH	7.85	8.44	7.53	***	7.86		7.24	***
% of Labor force male	48%	47%	48%		48%		48%	
Dependency share	0.29	0.29	0.29		0.28		0.30	
Credit constrained (1,0)	18%	16%	19%		24%	*	15%	
Average educ. of head	4.91	5.05	4.83		4.59		5.05	
Indigenous head (1,0)	60%	56%	62%		70%	**	56%	
Female head (1,0)	11%	12%	10%		10%		11%	
Single head (1,0)	12%	11%	13%		14%		9%	
Age of head	41.7	42.0	41.5		43.3		39.9	
House (1,0)	87%	83%	89%	*	88%		89%	*
Concrete/brick house (1,0)	87%	82%	90%	**	89%	*	91%	**
Refrigerator (1,0)	16%	12%	17%		18%		17%	
Access to water system (1,0)	94%	93%	95%		97%		93%	
Sewage (1,0)	5%	5%	6%		5%		7%	
Big farm animals (#)	5.67	5.32	5.43		6.35		5.28	
Ag. ass. membership 5 years + (1,0)	8%	6%	9%		9%		8%	
Non-Ag. ass. membership 5 years + (1,0)	66%	63%	68%		70%		66%	
Bus in community (1,0)	46%	43%	47%		45%		50%	
Elementary school (1,0)	87%	88%	87%		85%		89%	
Distance to closest city (km)	29.68	27.36	30.90	**	36.54	***	25.73	
Chimborazo (1,0)	50%	50%	50%		45%		41%	
N	495	171	324		155		169	

[*] $p < 0.10$.
[**] $p < 0.05$.
[***] $p < 0.01$.
[a] Tests are for differences in means with respect to treated farmers.

for welfare variables: on average, non-participants have a higher probability of living in a house, rather than a hut, and have a higher probability of living in a brick house which gives an indication that they may be marginally better off. The comparison between non-eligible (column IV) farmers and beneficiaries indicates that these two groups are also very similar in terms of land-related, social capital, welfare, and community variables, but they do differ slightly in socio-demographic characteristics. As expected, when the two possible control groups are pooled (column III), their differences with beneficiaries fall between what was described for columns IV and V. The two groups differ in 7 out of 28 variables. In sum, the descriptive statistics show that there are relatively minor differences between beneficiary and non-participant and non-eligible farmers. This gives an indication that the rigorous process of identifying control communities prior to data collection was quite successful although not perfect – a fact considered in the empirical strategy discussed below.

The data set also contains information on the value of potato production and input variables, which are used in the SPFA, and are presented in Table 2 on a per hectare basis along with a test of difference in means. Looking at the value of total output per hectare, as expected given the emphasis of the project, the test of means shows that beneficiary farmers have a significantly higher value; and this is true when doing the comparison against non-participants, non-eligible farmers, or the pooled group of controls although the magnitudes vary. Table 2 also shows that beneficiary farmers, on average, spend more per hectare on labor, on seeds and on other inputs, and have a higher likelihood of hiring paid labor.

Analytical framework and empirical approach

Two critical components of productivity growth are TC and TE. The former captures "jumps" in the production function stemming from the application of improved practices that come from research and development efforts, whereas the latter can be interpreted as a relative measure of managerial ability for a given technology. In this paper, these two effects are disentangled making use of the Stochastic Production Frontier (SPF) framework in order to determine if the *Plataformas* had an impact on each component of productivity. Identifying program impact on TC and TE requires ensuring that selection, particularly self-selection, does not bias the estimates. Although the process of creating treatment and control groups, already discussed, attempted to minimize bias, this could still arise from differences in observable and unobservable characteristics of treatment and control households. To address this issue, we follow the framework presented by Bravo-Ureta et al. (2012), where PSM is used to mitigate biases stemming from observable variables when selecting the counterfactual/control groups and then bias from unobservables is addressed using SPF with sample selection.

The SPF model incorporates a composed error structure where a two sided symmetric term captures standard random variability and a one sided component captures inefficiency. In general terms, the model used in this study can be expressed as:

$$\boldsymbol{Y}_{ij} = f(\boldsymbol{X}, \boldsymbol{T_D}) + \boldsymbol{v}_{ij} - \boldsymbol{u}_{ij} \tag{1}$$

where $\boldsymbol{Y}_{ij}$ is the yield (value of output per hectare) of the i^{th} farmer, $\boldsymbol{X}$ is a vector of inputs per hectare, the variable $\boldsymbol{T_D}$ is a farm specific dummy variable that captures the effect of the new technology (i.e., innovations coming from participation in the *Plataformas*), $\boldsymbol{v}_{ij}$ is the two sided error term, and $\boldsymbol{u}_{ij}$ is the one sided error that captures efficiency. The subscript j is equal to B for beneficiaries or C for control farmers. The key technology effect that we are interested in identifying relates to the impact of participating in the *Plataformas* on potato yields. Two null hypotheses (H_0) will be tested:

Table 2
Descriptive statistics for inputs and output used in the SPF models.[d]

Variables	Pooled	Benef.	All cont.	Test[a]	Non-elig.	Test	Non-partic.	Test
Total output ($/ha)	5.72	6.17	5.48	***	5.36	***	5.59	***
Total expenditures on labor ($/ha)[b]	5.81	5.99	5.72	**	5.60	***	5.83	
Total expenditures on seeds ($/ha)	4.22	4.71	3.96	***	3.85	***	4.05	***
Total expenditures on O. inputs ($/ha)[c]	5.11	5.41	4.96	***	4.87	***	5.03	***
Hired labor (1,0)	0.58	0.65	0.53	***	0.61		0.46	***
N	495	171	324		155		169	

* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.
[a] Tests are for differences in means with respect to beneficiary farmers.
[b] Includes Family, Hired, and Minga (community) Labor.
[c] Includes expenditures on insecticides, fungicides (preventative and curative), fertilizer (organic and chemical), tractor, and animal draft.
[d] Values are in natural logarithm.

(1) The parameter of $\boldsymbol{T_D} = 0$; and
(2) Mean TE_B = mean TE_C or equivalently mean $\boldsymbol{u_{iB}}$ = mean $\boldsymbol{u_{iC}}$.

Failure to reject the first null hypothesis indicates that there is no difference in yields, *ceteris paribus*, between control and treatment stemming from TC. Failure to reject the second hypothesis signifies that managerial ability, as measured by mean TE scores, across the two groups is the same. In this context, Fried et al. (2008) argue that although managerial ability is unobservable, it can be inferred from a ranking of TE scores derived from a "best practice" production frontier, which is the concept adopted here.

The empirical strategy to test these hypotheses involves several steps. The first step is to use PSM to construct a counterfactual group of farmers that have time-invariant characteristics similar to farmers that participated in the project (beneficiaries). The PSM procedure uses a Probit model to calculate the predicted probability of treatment based on observable characteristics. These probabilities, or propensity scores, are then used to match similar households in the treatment group with those from the control group. The matching procedure applied here is the 1-to-1 nearest neighbor (NN) with replacement criterion using a caliper width of 0.001, which is more rigorous than the $0.25\sigma_P$ method suggested in the literature (Guo and Fraser, 2010).[3] While other papers rely on the 1-to-1 NN without replacement, arguing that it has the most intuitive interpretation of all alternatives available (e.g., Bravo-Ureta et al., 2011), this paper opts for 1-to-1 NN with replacement since it provides a better quality in the matching and it is more likely to decrease biases, although it adds variance in the estimations (Abadie and Imbens, 2002; Caliendo and Kopeinig, 2008).[4] From this step, we obtain treatment and control groups with a similar range of observable characteristics.

The second step involves the estimation of SPF models. Conventional SPF models are estimated using the combined sample of beneficiary and control farmers (pooled data) as well as for each group separately. This is done to test whether beneficiary farmers display a different technology than the control group. If no technological difference is found then a single frontier combining all farmers, treated and control, is the more desirable option. At this stage, evaluating alternative functional forms for the SPF is advisable and the Cobb-Douglas (CD) was tested against the Translog (TL), which are the two most commonly used in efficiency studies (Bravo-Ureta et al., 2007). The results of maximum likelihood ratio tests were mixed. The linear parameters for the CD and TL were very similar in magnitude, and the coefficients for the quadratic and interaction terms in the TL where, in most of the cases, not statistically significant. Thus, the Cobb-Douglas (CD) functional form was selected for the analysis.

To control for possible biases based on unobserved characteristics we implement the method recently introduced by Greene (2010). The model assumes that the unobserved characteristics in the selection equation are correlated with the noise in the stochastic frontier model (i.e., the term v_i in Eq. (2)).[5] Greene (2010) frames his model by noting that Heckman's (1979) original sample selection approach was developed for linear models and is not applicable for non-linear cases such as the SPF.[6] Thus, Greene proceeds to develop a selection approach for the SPF, which can be expressed as:

$$\text{Sample selection}: \quad d_i = 1[\alpha' \boldsymbol{z}_i + w_i > 0], w_i \sim N[0,1]$$

$$\begin{aligned} \text{SPF}: \quad & y_i = \boldsymbol{\beta}' \boldsymbol{x}_i + \varepsilon_i, \varepsilon_i \sim N[0, \sigma_\varepsilon^2] \\ & (y_i, \boldsymbol{x}_i) \text{ is observed only when } d_i = 1 \end{aligned}$$

$$\begin{aligned} \text{Error structure}: \quad & \varepsilon_i - v_i - u_i \\ & u_i = |\sigma_u U_i| = \sigma_u U_i, \text{ where } U_i \sim N[0,1] \\ & v_i = \sigma_v V_i, \text{ where } V_i \sim N[0,1] \\ & (w_i, v_i) \sim N_2[(0,1),(1,\rho,\sigma_v,\sigma_V^2)] \end{aligned}$$

where d is a binary variable that takes the value of 1 for beneficiaries and 0 for control farmers, y is output, $\boldsymbol{z}$ is a vector of explanatory variables in the sample selection equation, $\boldsymbol{x}$ is a vector of inputs in the production frontier, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the parameters to be estimated, and the error structure corresponds to that in the

[3] This means using a caliper width for the nearest neighbor that is less or equal to a quarter of one standard deviation ($0.25\sigma_P$) of the estimated propensity scores of the sample.

[4] The alternative for 1-to-1 matching NN without replacement was also used and the results are very similar.

[5] The rationale for the underlying assumption that the bias arises from possible correlation between unobservables in the production frontier with unobservables in the selection equation is that one "…might expect that observations are not selected into the sample based on their being inefficient to begin with" (Greene, 2010, p. 23). As pointed out by one of the reviewers, this assumption is potentially a weakness of the approach but the implications for the empirical results and subsequent analyses are not evident. This is an issue that clearly merits additional methodological and empirical work.

[6] It is worth noting that several papers have applied the Heckman correction method to control for self-selection on unobservables in SPF studies. For instance, Sipiläinen and Oude Lansink (2005) use a distance frontier model to analyze TE for Finnish organic and conventional dairy farms, while Solís et al. (2007) analyze TE for farmers in El Salvador who used different levels of adoption of soil conservation. Wollni and Brümmer (2012) examined productive efficiency for specialty and conventional coffee farmers in Costa Rica, and Rahman et al. (2009) analyzed production efficiency for a sample of rice producers in Thailand using the Greene (2010) method. However, these last two papers do not use matching techniques to control for differences based on observed characteristics. As far as we know, only Bravo-Ureta et al. (2012) use PSM along with Greene's (2010) self-selection model to identify the productivity impact of a development project.

stochastic frontier model. In this model, the parameter ρ captures the presence or absence of selectivity bias.

The log likelihood for the model in (2) is formed by integrating out the unobserved $|U_i|$ and then maximizing with respect to the unknown parameters. Thus,

$$\log L(\boldsymbol{\beta}, \sigma_u, \sigma_v, \boldsymbol{\alpha}, \rho) = \sum_{i=1}^{N} \log \int_{|U_i|} f(y_i|\boldsymbol{x}_i, \boldsymbol{z}_i, d_i, |U_i|)p(|U_i|)d|U_i|. \quad (3)$$

The integral in (3) is not known so it has to be approximated. To simplify the estimation, Greene (2010) uses a two-step approach. The single equation MLE of α in the Probit equation in (2) is consistent but inefficient. The estimation of the parameters of the SPF does not require the reestimation of α so estimates for the latter are taken as given in the simulated log likelihood. Greene (2010) indicates that standard errors are adjusted based on the Murphy and Topel (2002) correction in a manner similar to what is done in Heckman (1979).

Greene (2010) goes onto argue that the non-selected observations (i.e., when $d_i = 0$) do not contribute information about the parameters to the simulated log likelihood and thus the function to be maximized becomes:

$$\begin{aligned}\log L_{S,C}(\boldsymbol{\beta}, \sigma_u, \sigma_v, \rho) = \sum_{d_{i=1}} \log 1/R \sum_{r=1}^{R} [&\exp(-(1/2) \\ &\times ((y_i - \boldsymbol{\beta}'\boldsymbol{x}_i + \sigma_u|U_{ir}|)^2/\sigma_v^2))/\sigma_v\sqrt{2\pi} \\ &\times \boldsymbol{\Phi}(\rho((y_i - \boldsymbol{\beta}'\boldsymbol{x}_i + \sigma_u|U_{ir}|)/\sigma_\varepsilon \\ &+ a_i)/\sqrt{1-\rho^2})],\end{aligned} \quad (4)$$

where $a_i = \hat{\alpha}'z_i$. Model parameters are estimated using the BFGS approach and asymptotic standard errors are obtained using the BHHH estimator. For full details on the model and its estimation see Greene (2010).

The estimation also requires modeling the farmer selection into the project. This selection process can be captured by a criterion function, which is assumed to be associated with exogenous household socio-economic variables, and can be expressed as:

$$B_i = \alpha_0 + \sum_{j=1}^{6} \alpha_j \boldsymbol{Z}_{ji} + w_i \quad (5)$$

where B is a binary variable capturing the i^{th} farmer's participation in the project ($B = 1$ for beneficiary farmers, and 0 for control farmers); $\boldsymbol{Z}$ is a vector of exogenous variables; α are the parameters to be estimated and w is the error term distributed as $N(0, \sigma^2)$.

Within this framework, the predictor of TE can be obtained as the expectation of u_i conditional on the composed error term ε_i following Jondrow et al. (1982). Again, a full description of the estimation of the TE scores can be found in Greene (2010).

Results

Table 3 reports the results of Probit models on participation in the *Plataformas* using data on all controls (I), the non-eligible (II) and non-participants (III). For each model, marginal effects calculated at the sample mean are reported. The models accurately predict 69.9%, 68.7%, and 62.7% of outcomes, respectively. As expected, these results are consistent with those in Table 1. The Probit results are used to calculate propensity scores for the treatment and control groups. Fig. 1 shows the density estimates of the distribution of propensity scores for each group, along with the areas with and without common support. The scores obtained are almost entirely in the area of common support, suggesting that the different controls consistently represent a reasonable counterfactual for the treated group. Indeed, there are very few observations off common support: 7 using all controls; 9 using non-eligibles; and 13 using non-participants.[7]

Table 3
Probit on *Plataformas* participation (marginal effects).

	(I) All controls	(II) Non-eligible	(III) Non-partic.
Land			
Land owned (ha)	−0.00325	−0.00643	0.00769
Owned plots (#)	0.0365**	0.0359**	0.0365**
Black soil (%)	−0.126*	−0.196**	−0.112
Flat land (%)	−0.0127	0.0416	−0.0219
Irrigated land (%)	−0.0728	−0.0936	−0.102
Socio-demographic			
Family size	0.00246	0.00598	−0.00566
Max educ. in HH	0.0186**	0.00722	0.0355***
% of labor force male	−0.0506	−0.0745	−0.0578
Dependency share	0.0336	−0.0220	0.135
Credit constrained (1,0)	−0.0580	−0.155**	0.0244
Indigenous head (1,0)	−0.0571	−0.103	−0.0233
Female head (1,0)	−0.0387	−0.0848	−0.0276
Age of head	−0.00104	−0.00533*	0.00113
Welfare			
House (1,0)	−0.0461	−0.0696	−0.0405
Concrete/brick house (1,0)	−0.158**	−0.142	−0.187**
Refrigerator (1,0)	−0.148**	−0.204**	−0.143
Access to water system (1,0)	−0.0993	−0.313**	0.000203
Sewage (1,0)	−0.0916	−0.0727	−0.136
Big farm animals (#)	−0.00228	−0.00533	0.00223
Social capital			
Ag. ass. membership 5 years + (1,0)	−0.0779	−0.149	−0.0629
Non-Ag. ass. membership 5 years + (1,0)	−0.0800	−0.0490	−0.102
Community variables			
Bus in community (1,0)	−0.160**	−0.177**	−0.147*
Elementary school (1,0)	0.0458	0.158*	−0.0179
Distance to closest city (km)	−0.00537**	−0.0114***	0.000951
Chimborazo (1,0)	0.0780	0.0534	0.153*
N	495	326	340
Correctly classified	69.9%	68.7%	62.7%

* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

Having identified the appropriate groups of control farmers, the next step is to determine if the conventional SPF should be run for the whole sample or if separate frontiers are necessary for beneficiary and control farmers. First, a pooled SPF is estimated that includes a binary variable for participation in *Plataformas*. Next, two separate SPF models, one for beneficiaries and a second one for control farmers, are estimated. The LR tests confirm that treated and control farmers display different technologies; thus, separate SPFs are preferable. These results are corroborated by the pooled models presented in Tables 4 and 5 where the parameters for the variable Beneficiaries (beneficiary farmers) are positive and significant. Then, to correct for the possible bias from unobservables, two separate SPF models are reestimated using Greene's (2010) selection correction framework.

The results of the SPF models are presented in Table 4 for the unmatched samples, and in Table 5 for the matched samples. All models presented in these two tables use the value of potatoes harvested per hectare (US$/ha) as the dependent variable. Each table contains a total of nine models, divided into two groups, one denominated conventional SPF and the second sample selection corrected SPF. As expected, all estimated models present positive

[7] Table A.1 reports the punctual test of means for the pooled sample following Leuven and Sianesi (2003).

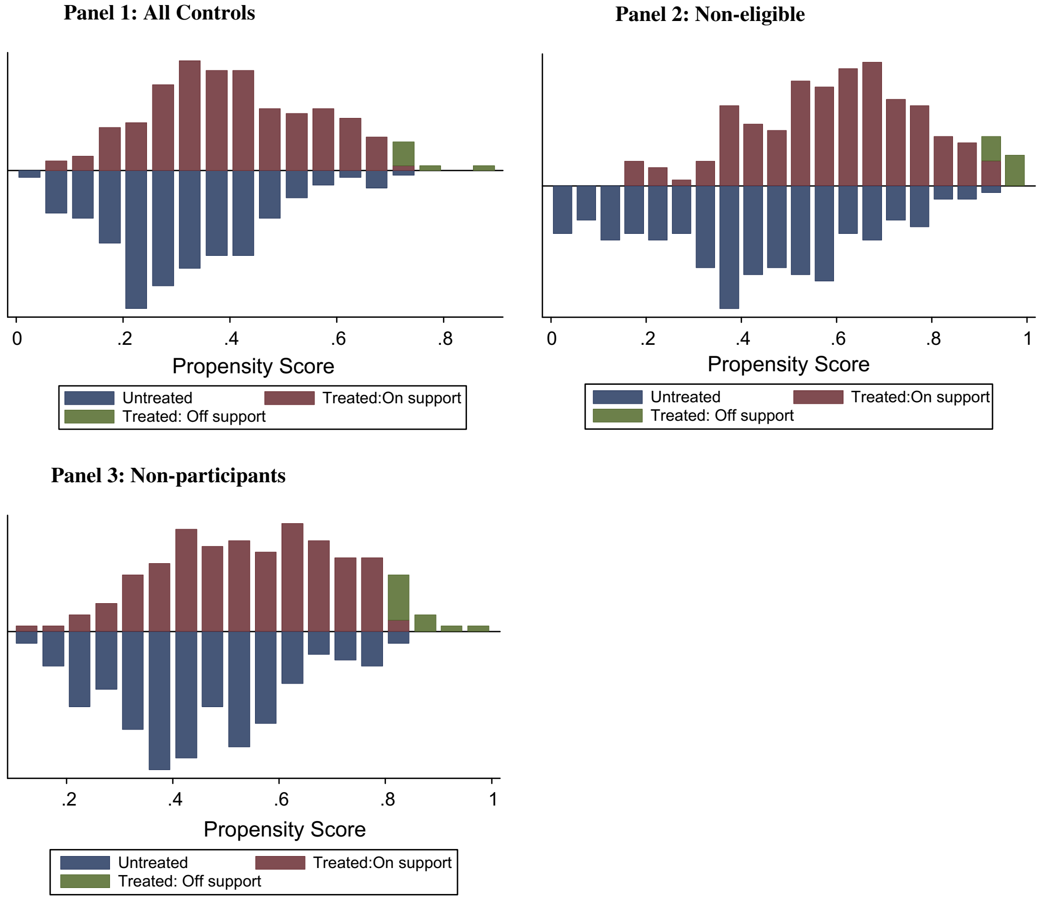

Fig. 1. Common support with different controls.

partial production elasticities; however, their magnitudes and statistical significance differ. In every model, expenditures on seeds make the greatest contribution to yields, followed by total expenditures on other inputs.[8] On the other hand, labor plays a minor role, particularly for beneficiaries and the non-eligible. The results on input expenditures are well in-line with those found by Kalirajan (1991) and Bravo-Ureta et al. (2012). The former argues that the cash required to buy inputs is one of the main production constraints for smallholders in developing countries. The mixed results for labor expenditures, where the coefficient is insignificant for beneficiaries and the non-eligible, are consistent with the Lewis (1954) model. This is the case because the farmers being studied have very limited amounts of land with a relatively abundant work force; thus, the marginal productivity of labor can be zero. In the case of beneficiaries, this might reflect the additional effort, in terms of labor used, put forth to ensure that the new technology is successful. On the other hand, the fact that the coefficient on labor is significant for the non-participants (in treated communities) may reflect their preference to diversify their labor allocation by using it outside potato production.

Given that expenditures on seeds and other inputs make the greatest contribution to farm production, one way to interpret these results is that since many of these farmers spend very little on inputs, the marginal returns to these expenditures are very high. The values for γ, also reported at the end of Tables 4 and 5 are used to test whether inefficiency is statistically significant. As shown in the tables, the null hypothesis that $\gamma = 0$ is rejected, which reveals that technical inefficiency is indeed an important contributor to yield variability (Coelli et al., 2005).

Turning to the results obtained from the matching procedure along with Greene's (2010) sample-selection correction method, we see that these are very similar to those obtained with the conventional SPF method. The results of the selection equation are presented in the Appendix Table A.2, which show that only the parameters for the age variables are significant.

Moreover, when looking at the results for ρ for the various models, there is no statistical support for selection bias. One implication of these results is that the rigorous process followed to define the counterfactual groups, including the PSM before estimating the SPF models, have eliminated significant biases from unobservables.

The results for average TE are presented in Table 6 separately depending on the control group used along with tests for the difference in means.[9] The average TE for the entire unmatched sample is slightly higher than those found in other similar studies for small-scale farmers in Central America (Bravo-Ureta et al., 2007). More specifically, in all pooled models, the difference in TE between ben-

[8] This variable includes expenditures on insecticides, fungicides (preventative and curative), fertilizer (organic and chemical), tractors, and draft animals.

[9] Average TE is computed for each group of farmers with respect to their respective frontier and these means are then compared across groups. Therefore, these comparisons are not absolute but relative to each group's own frontier. In other words, it is not possible to say which group or which specific farmer exhibits higher or lower productivity for the overall sample. However, it is possible to investigate the effect of selectivity bias on TE within each group and then discuss how close each group, on average, is to its own frontier and thus make relative productivity/efficiency statements.

Table 4
Parameter estimates for the conventional and sample selection SPF models: unmatched sample.

Variables	Pooled	Conventional SPF				Sample selection corrected SPF			
		Benef [a]	All control	Non-eligible	Non-partic.	Benef [a]	All control	Non-eligible	Non-partic.
β1 = Labor ($/ha)	0.160**	0.063	0.180***	0.153	0.170*	0.062	0.181***	0.147	0.188**
β2 = Seeds ($/ha)	0.560***	0.518***	0.599***	0.681***	0.550***	0.515***	0.599***	0.695***	0.532***
β3 = O. Inputs ($/ha)	0.318***	0.436***	0.259***	0.197*	0.341***	0.442***	0.257***	0.193**	0.336***
Hired labor (0,1)	0.081	−0.036	0.121*	0.086	0.131	−0.023	0.121*	0.086	0.122
Chimborazo	−0.416***	−0.386*	−0.460***	−0.404**	−0.604***	−0.386**	−0.479***	−0.426***	−0.584***
Beneficiaries	0.140*	–	–	–	–	–	–	–	–
Controls in PC	−0.001	–	–	–	–	–	–	–	–
Altitude	−0.001**	−0.001	−0.001*	−0.001	−0.001*	−0.001*	−0.001**	−0.001**	−0.001**
Constant	2.671***	3.535***	2.434***	2.740***	2.663***	3.241**	2.443***	2.864***	2.747***
γ	2.192***	2.658***	1.861***	1.667***	2.279***	–	–	–	
L. Likelihood	−441.745	−173.378	−259.873	−130.035	−125.636	−316.064	−316.064	−244.745	−242.025
$\sigma(u)$	–	–	–	–	–	0.967***	0.653***	0.646***	0.721***
$\sigma(v)$	–	–	–	–	–	0.385**	0.388***	0.425***	0.355**
$\rho(w,v)$	–	–	–	–	–	0.446	−0.106	−0.238	−0.592
N	495	171	324	155	169	171	324	155	169

* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.
[a] Estimates from beneficiary sample using all controls for the matching procedure. The other two sets of estimates for beneficiaries show similar estimates and where omitted from this table due to space limitation.

Table 5
Parameter estimates for the conventional and sample selection SPF models: matched sample.

Variables	Pooled	Conventional SPF				Sample selection corrected SPF			
		Beneficiaries[a]	All controls	Non-eligible	Non-partic.	Beneficiaries [a]	All controls	Non-eligible	Non-partic.
β1 = Labor ($/ha)	0.164***	0.074	0.180***	0.153	0.170*	0.067	0.180***	0.147	0.194**
β2 = Seeds ($/ha)	0.550***	0.495***	0.599***	0.681***	0.550***	0.494***	0.609***	0.687***	0.527***
β3 = Oth. Inputs ($/ha)	0.322***	0.447***	0.259***	0.197*	0.341***	0.452***	0.247***	0.196**	0.328***
Hired labor (0,1)	0.08	−0.047	0.121*	0.086	0.131	−0.037	0.126**	0.080	0.128
Chimborazo	−0.4076***	−0.367*	−0.460***	−0.404**	−0.604***	−0.343*	−0.48979***	−0.421***	−0.580***
Beneficiaries	0.153*					–	–	–	–
Controls in PC	−0.00056					–	–	–	–
Altitude	−0.00032***	−0.00045	−0.0002*	−0.0003	−0.0003*	−0.0004	−0.0002***	−0.0003**	−0.0003**
Constant	2.636***	3.494***	2.434***	2.740***	2.663***	3.119**	2.466***	2.804***	2.715***
RTS	1.04	1.03	1.04	1.03	1.06	1.01	1.04	1.03	1.05
γ	2.208***	2.644***	1.861***	1.667***	2.279***	–	–	–	–
L. Likelihood	−434.999	−166.324	−259.872	−130.035	−125.636	−343.465	−396.058	−236.020	−239.961
$\sigma(u)$	–	–	–	–	–	1.0190***	0.621***	0.677***	0.642***
$\sigma(v)$	–	–	–	–	–	0.387*	0.403***	0.405***	0.398***
$\rho(w,v)$	–	–	–	–	–	0.325	−0.139	−0.020	−0.603
N	488	164	324	155	169	164	324	155	169

* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.
[a] Estimates from beneficiary sample using all controls for the matching procedure. The other two sets of estimates for beneficiaries show similar estimates and were omitted from this table due to space limitation.

eficiaries and controls is negligible and non-significant. On the other hand, in all separate models beneficiaries exhibit lower levels of TE than control farmers independent of the control group used for comparison. Another point to highlight is that average TE decreases for beneficiaries when going from the pooled to the separate model, and then to the separate with sample selection models, although the magnitude of the change is small. On the other hand, the opposite is true for most cases for the controls for which generally TE increases when going from the pooled to the separate, and from the separate to the sample correction models. In sum, the main result is that, independent of the group or method used for comparisons, beneficiary farmers exhibit lower levels of average TE. That is, beneficiary farmers tend to use their resources in a less efficient way than the control group with respect to their respective technologies.

To explore the issue of TE further, we focus on the matched sample and estimate TE levels for three different groups of beneficiaries based on the number of years they had participated in *Plataformas* at the time the data was collected: (1) those who had recently joined *Plataformas* (at least one year); (2) those who had been in the program for up to two years; and (3) those that had participated for three or more years.[10] We then conduct a test of means across groups by comparing the TE levels of farmers that had been in the *Plataformas* for up to three years against the other two groups, and against all controls. The results, presented in Table 7, show that beneficiaries who had been in the program for only one year exhibit the lowest levels of average TE (0.49) compared to those who had been in the program for three years or more (0.57), and this difference is statistically significant. Similarly, beneficiaries who had been in the program for up to two years also exhibit lower levels of TE (0.50), although slightly higher than the first group. Finally, when comparing all controls against beneficiary farmers who had been in the program for three years or more, while there are differences between their levels of TE (0.61 versus 0.57), these are not statistically

[10] We thank an anonymous reviewer for suggesting this additional analysis.

Table 6
Technical efficiency levels.

Control group	Sample and method	Pooled	Beneficiaries	All controls	Test of means
All	Unmatched				
	Pooled	0.57	0.55	0.57	
	Separate	–	0.51	0.61	***
	Sample Selection	–	0.51	0.62	***
	Matched				
	Pooled	0.57	0.55	0.57	
	Separate	–	0.51	0.61	***
	Sample Selection	–	0.50	0.61	***
Non-eligible	Unmatched				
	Pooled	0.55	0.54	0.55	
	Separate	–	0.51	0.61	***
	Sample Selection	–	0.52	0.62	***
	Matched				
	Pooled	0.55	0.54	0.55	
	Separate	–	0.52	0.61	***
	Sample Selection	–	0.51	0.61	***
Non-partic.	Unmatched				
	Pooled	0.55	0.54	0.57	
	Separate	–	0.51	0.60	***
	Sample Selection	–	0.50	0.60	***
	Matched				
	Pooled	0.56	0.54	0.57	
	Separate	–	0.52	0.60	***
	Sample Selection	–	0.51	0.57	**

* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

Table 7
Technical efficiency levels by years of participation in *Plataformas*.

Length of time in Plataformas 3 or more years	At least 1 yr.	Test[a]	At least 2 years	Test[a]	All controls	Test[a]
0.57	0.49	**	0.50	**	0.61	

** $p < 0.05$.
[a] Tests are for differences in means with respect to treated farmers with 3+ years in the *Plataformas*.

Table 8
Predicted frontier value of output per hectare after sample bias correction at the mean of the data.

Sample	Beneficiaries[b]	All controls	Test[a]	Non-eligible	Test	Non-partic.	Test
Unmatched	1689	940	***	946	***	1185	***
Matched	1738	954	***	859	***	856	***

*** $p < 0.01$.
[a] Tests are for differences in means with respect to treated farmers.
[b] Average frontier output for the 3 matched samples.

significant. One way to interpret these results is that farmers who had applied the new technology package for one year or less experienced a drop in their efficiency, but TE levels started to pick up by the second year, and by the third year they got closer to the level exhibited by the counterfactual group.

The results of Table 6 show that on average beneficiary farmers are less efficient than control farmers, when compared to their own frontiers, and now we are interested in examining which of the two groups (beneficiaries versus controls) has higher value of yields after controlling for biases from observed and unobserved characteristics. These comparisons, shown in Table 8 along with a test of the difference in means, are based on the average predicted frontier output obtained for beneficiaries and controls, using their respective frontiers, after applying both PSM and Greene's (2010) selection correction.[11] These results show that, on average, beneficiary farmers are significantly more productive than control farmers and this finding is consistent for all comparisons made. Thus, an important implication of these results is that participation in *Plataformas* has contributed to a technological advantage for participants *vis a vis* non-participants—that is, the program induced TC. This finding is compatible with previous studies that have looked at the impact of *Plataformas* on farm output using very different conceptual frameworks (Cavatassi et al., 2011a,b).

Overall, the results lead to the rejection of both of the null hypotheses of this study: (1) equality of yields between beneficiary and control farmers; and (2) equality in TE for both groups. However, the rejection of the second hypothesis is opposite to what was anticipated. More precisely, beneficiaries of *Plataformas* have significantly higher yields, but are significantly less efficient. Therefore, beneficiaries operate on a higher production frontier than controls due to the technology transferred by the program; but, on average, they are further away from their frontier compared to controls. However, TE levels rise with the length of participation in the *Plataformas*. An initial lower level of TE for

[11] The results of the matched conventional SPFA model are very similar. We present results using the matched sample with the sample correction since it is more rigorous than the conventional model.

Table A.1
Punctual test of means.

	Mean		% Reduction of bias	P>|t|
	Treated	Control		
Land				
Land owned (ha)	2.75	2.72	9.5	0.944
Owned plots (#)	3.37	3.25	75.5	0.602
Black soil (%)	0.77	0.78	80.2	0.753
Flat land (%)	0.38	0.38	−1.6	0.890
Irrigated land (%)	0.57	0.60	46.9	0.567
Socio-demographic				
Family size	4.85	4.85	97.0	0.977
Max educ. in HH	8.44	8.37	92.5	0.867
% of Labor force male	0.47	0.46	39.6	0.704
Dependency Share	0.29	0.29	−40.0	0.818
Credit constrained (1,0)	0.16	0.19	26.8	0.590
Indigenous head (1,0)	0.56	0.55	76.9	0.793
Female head (1,0)	0.12	0.12	71.4	0.903
Age of head	42.02	42.22	59.5	0.892
Welfare				
House (1,0)	0.83	0.84	75.1	0.720
Concrete/brick house (1,0)	0.82	0.83	92.2	0.885
Refrigerator (1,0)	0.12	0.14	68.4	0.669
Access to water system (1,0)	0.93	0.93	90.3	0.952
Sewage (1,0)	0.05	0.06	−78.7	0.835
Big farm animals (#)	5.43	5.73	17.2	0.614
Social capital				
Ag. ass. membership 5 years + (1,0)	0.06	0.06	86.7	0.912
Non-Ag. ass. membership 5 years + (1,0)	0.63	0.62	96.7	0.972
Community variables				
Bus in community (1,0)	0.43	0.46	42.4	0.677
Elementary school (1,0)	0.88	0.89	−349.5	0.631
Distance to closest city (km)	27.36	27.29	98.2	0.966
Chimborazo (1,0)	0.49	0.50	82.5	0.835
N	164	324		

Table A.2
Selection equation: treated communities only.

	Unmatched		Matched	
	Treated Coeff.	Controls Coeff.	Treated Coeff.	Controls Coeff.
Age of head	.08031**	−.08031**	.0744**	−.07444**
Age of head^2	−.00079**	.00079**	−.00073**	.00073**
Education of head	0.03135	−0.03135	0.03976	−0.03976
Education of head^2	−0.00102	0.00102	−0.0013	0.00133
Family size	0.01074	−0.01074	−0.00225	0.00225
Land owned (ha)	0.01803	−0.01803	0.01228	−0.0123
Constant	−2.03667***	2.03667***	−1.89953**	1.89953**
N	340	340	327	327

* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

beneficiaries is generally consistent with Schultz (1975) who postulates that traditional farmers, when faced with a production shock, such as the introduction of a new technology, could experience an initial drop in productivity.

Conclusions

While there is growing interest in conducting impact evaluations in the agricultural sector, most of the related literature focuses on output indicators, such as increases in yields, total value of product, profits, and gross margins, while attention to managerial performance is rare. Focusing on TE, as an indicator of managerial performance, can be a significant addition to impact evaluation studies of agricultural projects. This paper uses data from small-scale potato farmers in Ecuador to disentangle TE measures from TC indicators by combining impact evaluation tools with the SPF approach. A matched group of beneficiaries and control farmers is generated using PSM techniques to correct for biases based on observed characteristics. In addition, this paper deals with possible self-selection arising from unobserved characteristics using Greene's (2010) selection correction method for stochastic frontiers. The analysis presented in this study, does not exhibit clear indication of selection biases. For instance, the PSM only dropped, on average, 10 out of 495 observations,[12] while Greene's (2010) method does not reveal clear evidence of selection bias from unobservables. These findings suggest that the rigorous process implemented in data collection and in constructing the counterfactual groups was able to mitigate the presence of biases from both observable and unobservable variables.

The analysis presented in this paper suggests that while beneficiary farmers have significantly improved their performance in terms of TC indicators, their levels of TE are lower than those of control farmers. Thus, while *Plataformas* had a positive and significant impact on inducing higher yields, the results suggest that the implementation of the new technology package had a cost in terms of lower managerial performance (TE), at least in the short run.

It is important to put these results in context. Given that *Plataformas* beneficiaries were taught an array of new procedures to implement throughout the production cycle (starting with selecting the right type of seeds, turning the land, applying IPM techniques, etc.), it is very likely that beneficiaries had to spend additional effort to master their newly acquired skills. Moreover, at the time of the interview, 30% of farmers had been with *Plataformas* for only one year, another 30% had been in the program for two years, and the average number of years for the entire sample of beneficiaries was 2.3 years. This means that for about one third of the beneficiaries, their most recent harvest was the first time they had a chance to implement their newly acquired technology. Thus, a significant share of beneficiaries had limited exposure and not enough time to fully incorporate their new know how from learning by doing. Indeed, our results do show losses in TE but only during the first years of participation, while there is evidence that TE starts to pick up the longer farmers are exposed to the new technology package.

These results are in line with economic theory (Schultz, 1975), and are compatible with those reported in related studies that have looked at TE for farmers that have recently entered into new niche markets. For instance, Sipiläinen and Lansink (2005) found that organic dairy farms experienced a decrease in TE when they first converted into organic production. Moreover, these authors estimated that TE started to increase only after 6 years from the switch and concluded that learning by doing as well as experience acquired over time are likely to be important factors in the efficiency of organic production.

Our results do suggest the need to conduct follow up evaluations after a project is completed to address longer term impacts. Such evaluations would make it possible to assess the sustainability of the project and also to examine if the managerial performance of beneficiaries improves as a result of the project compared to that of the control group. If the TE of beneficiaries continues to lag, in relation to the controls, it would be important to understand why that might be the case.

[12] The models drop 7 using all controls; 9 using non-treated communities; and 13 using treated communities only.

Acknowledgements

We would like to thank two anonymous referees, and participants at the *Innovations in Impact Assessment of Agricultural Research: Theory and Practice* Pre-conference Workshop at the 28th Conference of the International Association of Agricultural Economists (IAAE), Foz do Iguaçu, Brazil, for comments and suggestions.

Appendix A

See Tables A.1 and A.2.

References

Abadie, A., Imbens, G., 2002. Simple and bias-corrected matching estimators for average treatment effects. Technical Working Paper 283. National Bureau of Economic Research.

Angrist, J., Pischke, J., 2009. Mostly Harmless Econometrics. Princeton University Press, New Jersey.

Baker, J.L., 2000. Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners. LCSPR/PRMPO, The World Bank, Washington, DC.

Barrett, C., Carter, M., 2010. The power and pitfalls of experiments in development economics: some non-random reflections. Applied Economic Perspectives and Policy 32, 515–548.

Bravo-Ureta, B., Solís, D., Moreira, V., Maripani, J., Thiam, A., Rivas, T., 2007. Technical efficiency in farming: a meta-regression analysis. Journal of Productivity Analysis 27, 57–72.

Bravo-Ureta, B., Almeida, A., Solís, D., Inestroza, A., 2011. The impact of MARENA's investments on sustainable agricultural systems in Honduras. Journal of Agricultural Economics 62, 429–448.

Bravo-Ureta, B., Greene, W., Solís, D., 2012. Technical efficiency analysis correcting for biases from observed and unobserved variables: an application to a natural resource management project. Empirical Economics 43, 55–72.

Caliendo, M., Kopeinig, S., 2008. Some practical guidance for the implementation of propensity score matching. Journal of Economic Surveys 22, 31–72.

Cavatassi, R., González, M., Winters, P., Andrade-Piedra, J., Espinosa, P., Thiele, G., 2009. Linking smallholders to the new agricultural economy: an evaluation of the *Plataformas* program in Ecuador. ESA Working Paper No. 09-06, Food and Agricultural Organization.

Cavatassi, R., González-Flores, M., Winters, P., Andrade-Piedra, J., Espinosa, P., Thiele, G., 2011a. Linking smallholders to the new agricultural economy: the case of the *Plataformas de Concertación* in Ecuador. Journal of Development Studies 47, 1545–1573.

Cavatassi, R., Gonzalez-Flores, M., Salazar, L., Winters, P., 2011b. Do agricultural projects alter crop production technologies? Evidence from Ecuador. Journal of Agricultural Economics 62, 403–428.

Cerdán-Infantes, P., Maffioli, A., Ubfal, D., 2008. The impact of agricultural extension services: The case of grape production in Argentina. OVE Working Papers 2008/6. Inter-American Development Bank, Office of Evaluation and Oversight.

Coelli, T., Rao, D.S.P., O'Donnell, C.J., Battese, G.E., 2005. An Introduction to Efficiency and Productivity Analysis, second ed. Springer, New York.

Del Carpio, X., Maredia, M., 2009. Measuring the Impacts of Agricultural Projects: A Meta-Analysis of the Evidence. Working Paper, Independent Evaluation Group, World Bank. World Bank, Washington, DC.

Devaux, A., Horton, D., Velasco, C., Thiele, G., Lopez, G., Bernet, T., Reinoso, I., Ordinola, M., 2009. Collective action for market chain innovation in the Andes. Food Policy 34, 31–38.

Duflo, E., Glennerster, R., Kremer, M., 2008. Using randomization in development economics research: a toolkit. In: Schultz, T., Strauss, J. (Eds.), Handbook of Development Economics. Elsevier, Amsterdam, pp. 3895–3962.

Fried, O., Lovell, C., Schmidt, S., 2008. The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press, New York.

Greene, W., 2010. A stochastic frontier model with correction for sample selection. Journal of Productivity Analysis 34, 15–24.

Guo, S., Fraser, M., 2010. Propensity score analysis, statistical methods and applications. Advanced Quantitative Techniques in the Social Sciences. SAGE Publications, California, No. 12.

Heckman, J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.

Horton, D., Devaux, A., Ordinola, M., 2011. Highlights of the Papa Andina Experience. In: Devaux, A., Ordinola, M., Horton, D. (Eds.), Innovation for Development: The Papa Andina Experience. International Potato Center, Lima, Peru, pp. 1–36.

IDB (Inter-American Development Bank)., 2010. Development Effectiveness Overview, Special Topic, Assessing the Effectiveness of Agricultural Interventions. Inter-American Development Bank, Washington, DC.

Jondrow, J., Knox, L.C., Materov, I., Schmidt, P., 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. Journal of Econometrics 19, 233–238.

Kalirajan, K., 1991. The importance of efficiency use in the adoption of technology: a micro panel data analysis. Journal of Productivity Analysis 2, 113–126.

Leuven, E., Sianesi, B., 2003. PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. <http://ideas.repec.org/c/boc/bocode/s432001.html>.

Lewis, W.A., 1954. Economic development with unlimited supplies of labour. Manchester School 22, 139–191.

Murphy, K., Topel, R., 2002. Estimation and inference in two stem econometric models. Journal of Business & Economic Statistics 20, 88–97.

Pico, H.A., 2006. La cadena agroalimentaria de la papa a través de la metodología de plataformas de concertación y proyectos compartidos. In: Paper presented at the 1st national Ecuadorian potato congress, INIAP, Ecuador. <http://www.quito.cipotato.org/presentambato/TEMATICAS%20DEL%20CONGRESO/COMERCIALIZACION/HPICO.doc> (accessed 17–19.06.06).

Rahman, S., Wiboonpongse, A., Sriboonchitta, S., Chaovanapoonphol, Y., 2009. Production efficiency of Jasmine rice producers in northern and north-eastern Thailand. Journal of Agricultural Economics 60, 419–435.

Schultz, T.W., 1975. The value of the ability to deal with disequilibria. Journal of Economic Literature 13, 827–846.

Sipiläinen, T., Lansink, A., 2005. Learning in organic farming – an application on Finnish dairy farms. In: Paper Prepared for Presentation at the XIth Congress of the European Association of Agricultural Economists, Copenhagen, Denmark, August 24–27.

Solís, D., Bravo-Ureta, B., Quiroga, E., 2007. Soil conservation and technical efficiency among hillside farmers in Central America: a switching regression model. Australian Journal of Agricultural and Resource Economics 51, 491–510.

Winters, P., Salazar, L., Maffioli, A., 2010. Designing Impact Evaluations for Agricultural Projects, SPD Working Papers 1007. Inter-American Development Bank, Office of Strategic Planning and Development Effectiveness.

Winters, P., Maffioli, A., Salazar, L., 2011. Evaluating the impact of agricultural projects in developing countries: introduction to the special feature. Journal of Agricultural Economics 62, 393–402.

Wollni, M., Brümmer, B., 2012. Productive efficiency of specialty and conventional coffee farmers in Costa Rica: accounting for technological heterogeneity and self-selection. Food Policy 37, 67–76.

World Bank, 2006. Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank. Independent Evaluation Group, World Bank.