The Impact of Farmer Field Schools on Knowledge and Productivity: A Study of Potato Farmers in the Peruvian Andes

ERIN M. GODTLAND
U.S. General Accounting Office

ELISABETH SADOULET and ALAIN DE JANVRY
University of California at Berkeley

RINKU MURGAI
World Bank

OSCAR ORTIZ
International Potato Center, Consultative Group on International Agricultural Research

I. Introduction

The design of agricultural extension programs in developing countries has been the subject of heated debate. Guided by these debates, extension services have undergone several transformations in the past few decades (Byerlee 1994). The main transformation, until recently, was a shift from the transfer-of-technology approach to the training-and-visit, or T&V, system. Under T&V, the extension system was reoriented from a desk-bound bureaucracy with multiple economic and social objectives to a field-based cadre of agents who focused mainly on technology diffusion (Picciotto and Anderson 1997). T&V extension agents would meet with a small group of contact farmers who were expected to disseminate information to the members of their respective communities and convey farmers' opinions back to the agents, thus creating a feedback mechanism absent in the prior system (Birkhaeuser, Evenson, and

We are especially indebted to extension and research staff at CIP and CARE-Peru, who generously offered their time and expertise on behalf of this project. This research was made possible by financial support from the World Bank Research Committee RPO no. 683–56 and the Development Research Group. We thank Gershon Feder for helpful comments on an earlier version of the article. The views expressed in this article are ours and should not be attributed to the World Bank or any affiliated organization. We take full responsibility for any errors.

 0013-0079/2004/5301-0003$10.00

Feder 1991). For nearly 3 decades, international aid donors such as the World Bank promoted T&V as the most cost-efficient extension system.

T&V did, however, have its critics. With continued budgetary crises of less developed countries, some argued that it was too expensive and impossible to implement over extensive regions. Highly dispersed farmers could never establish frequent contact with extension agents. And their needs varied widely and could not be addressed with a single, inflexible technology package (Picciotto and Anderson 1997; Feder, Willett, and Zijp 2001).[1]

In recent years, a number of development agencies have promoted farmer field schools (FFS) as a potentially more effective approach to extend knowledge to farmers. FFS programs were first introduced in East Asia in the late 1980s as a way of diffusing knowledge-intensive integrated pest management (IPM) practices for rice.[2] Farmer field schools have since been adapted to work with other crops and diseases and have spread rapidly across Asia, Africa, and Latin America (Nelson et al. 2001). The FFS approach represents a paradigm shift in agricultural extension: the training program utilizes participatory methods "to help farmers develop their analytical skills, critical thinking, and creativity, and help them learn to make better decisions" (Kenmore 2002). Extension agents, who are viewed as facilitators rather than instructors, conduct learning activities in the field on relevant agricultural practices. Through interactive learning and field experimentation, FFS programs teach farmers how to experiment and problem-solve independently, with the expectation that they will thus require fewer extension services and will be able to adapt the technologies to their own specific environmental and cultural needs (Vasquez-Caicedo et al. 2000). Participants are encouraged to share their knowledge with other farmers and are sometimes trained to teach the courses themselves, thus reducing the need for external support.

Farmer field schools are costly undertakings, making a careful measurement of their impact important. However, empirical evidence on their effectiveness has been mixed. Results of previous impact evaluations have varied greatly according to the setting, the evaluation methods, and the yardstick used to assess impact. The few studies that examine the impact of FFS on farmers' knowledge generally find that FFS participants tend to have higher knowledge

[1] An abundance of empirical research exists on the effectiveness of T&V. See Birkhaeuser et al. (1991) for a review of studies on the economic impact of these and other agricultural extension programs.

[2] IPM is knowledge-intensive because in order to effectively implement IPM—which employs natural predators to combat pests—farmers must be able to understand the origins, cycles, and natural enemies of pests.

test scores after program participation or relative to a group of nonparticipants.[3] Some studies show that FFS participants use less pesticide and have higher yields compared to nonparticipants, while others find little evidence of impact on these outcomes. At the same time, there appears to be little evidence of diffusion of knowledge from FFS graduates to other farmers.[4]

A major drawback of most previous studies is that they do not properly control for potential differences between FFS participants and farmers in the comparison group, making it difficult to draw definitive conclusions. These differences could arise from the nonrandom placement of the program or from the voluntary nature of participation in FFS. For example, FFS villages might be chosen for their relative advantages in land fertility or climate; or farmers who voluntarily participate in FFS might be more productive, on average, than those who do not participate. Selective placement (through individual choice or purposive targeting) means that data on nonparticipants are not an efficient mode of revealing the likely achievements of participants in the absence of the program. Unless proper account of nonrandom farmer and village selection is taken, comparison of outcomes between FFS participants and nonparticipants is likely to yield biased estimates of program impact.[5]

This article uses data from a survey of potato farmers in Cajamarca, Peru, to examine the impact of a pilot FFS program on farmers' knowledge (as measured by a knowledge test score). Since there was no baseline survey documenting the knowledge of farmers prior to their participation in FFS, we rely on methods based on comparison groups. To deal with selection bias, we use propensity score matching (PSM) methods to build a statistical comparison group of farmers comparable to FFS graduates. This allows us to ensure that bias in the impact estimate due to selection on observables is minimized. Any remaining bias in the matching estimator can thus be attributed to unobserved characteristics. That said, given the low participation rate of farmers in this small pilot program, the sample of nonparticipants is very likely to include people who would participate if the program were more widely available.

By assessing impact immediately after participation in FFS, we may be capturing short-term knowledge acquisition that may or may not last over time. However, by restricting the measure of knowledge to the results of a test score on IPM practices, our study does not do full justice to the stated

[3] See, e.g., Rola et al. (2002) in the Philippines, Van de Fliert et al. (1999) in Indonesia, and Praneetvatakul and Waibel (2002) in Thailand.

[4] For a summary of these studies, see Feder et al. (2004).

[5] The only study that properly controls for selection biases finds no evidence of FFS impact (Feder et al. 2004).

purpose of the FFS program—to promote critical thinking and creativity. According to FFS scientists, critical thinking is most valuable in managing problems with pests and weather shocks, when farmers' knowledge on how to react to such problems is useful. Keeping in mind these limitations, our empirical results indicate that farmers who participated in the program have significantly more knowledge about IPM practices than those in the nonparticipant comparison group.

Moving beyond knowledge to the impact of FFS on production decisions and, ultimately, yields would require observing and comparing yields of FFS participants to those of nonparticipants. This survey was designed precisely to be the baseline for such an analysis and therefore was implemented while FFS was only in the first year of operation. Most of the production decisions had been taken either prior to or during the time when the FFS was in operation. Hence, we cannot expect yields from the first year to reflect the knowledge acquired through FFS. However, to get some sense of the importance of knowledge, we use the cross-sectional variation among the subsample of nonparticipant farmers to correlate knowledge with yields. Conditional on observed characteristics, we find that improved knowledge about IPM practices is positively correlated with productivity in potato production. Combining these results, simulations suggest that FFS has the potential to raise productivity substantially, by about 32% of the average value in a normal year. This evidence is merely suggestive, as it relies on the untestable assumptions that (i) knowledge acquired through FFS does not dissipate over time and (ii) the observed relationship between knowledge and yield can be inferred as causal and is not biased upward. These results will need to be confirmed by a rigorous analysis of impacts on yield when suitable data become available.

The article proceeds as follows. Section II describes the FFS program in Peru and the data set. Section III examines how farmers obtain information on potato cultivation and their knowledge levels. In Section IV, we present the research strategy used to test the impact of FFS on knowledge. Sections V and VI apply this methodology to measure impact on knowledge. Finally, Section VII estimates how knowledge affects productivity levels in potato cultivation, and Section VIII concludes.

II. The Program and Data

As the home country for the headquarters of the International Potato Center (CIP), one of the Consultative Group on International Agricultural Research centers, Peru has long been a focal point for the development and deployment of improved potato varieties and cultivation practices. In 1998, CIP scientists, in collaboration with CARE-Peru, launched a pilot farmer field school program for potato farmers in the department of Cajamarca. This department lies in the northern part of the Peruvian Andes, which are known as the Green Andes. Unlike the dry flatlands of the Altiplano, the Green Andes are characterized by steeply sloped, hilly terrain with relatively higher precipitation levels. The elevation of the survey region ranges from 9,000 to 12,000 feet above sea level. The economy in the survey region is dominated by small farms, with potato farming as the main activity. Potatoes constitute the bulk of households' food consumption and are also their most lucrative market crop.

The main aim of the FFS program was to introduce IPM techniques to Andean potato farmers. FFS participants were expected to attend 12 training sessions (typically once a week, with each session lasting for 3 hours). As the training strategy was based on the principle of learning by discovery, during these sessions the facilitator would organize various activities and experiments that the farmers could implement themselves. The curriculum was focused on the biology of late blight, the fungus that caused the Irish Potato Famine and continues to take huge tolls on potato production in Peru. Farmers were taught its symptoms, reproductive cycle, contamination source, and the conditions that foster its growth. On the experimental plot (one per FFS community), they identified potato varieties that are resistant to late blight infection. They learned how to prevent and control late blight with the use of improved varieties and fungicides. The program also introduced, in less detail, IPM for the Andean potato weevil and the potato tuber moth.

There was a two-stage selection process that determined which farmers participated in the program. First, CARE selected the villages in which to introduce the FFS program. These villages were chosen from a set of villages where CARE had already been implementing another rural development project named "Andino." This project worked with farmers' groups to improve farm production by providing technical advice and access to credit and by facilitating links to markets. Technical advice in Andino was imparted through conventional transfer-of-technology approaches. The Andino villages (and, consequently, the FFS villages) were not a random sample of villages in the region. Rather, CARE had conducted a diagnostic survey of all communities within the watershed and, based on this survey, classified communities into three types: subsistence, middle income, and high income. The target population for the Andino program was the set of middle-income communities, and, from this target group, 20 villages that were close to their respective district capitals were selected for participation. CARE planned to introduce FFS in all the Andino villages. However, at the time of the survey, field schools were operating only in four of them. Although no explicit rules were applied for the selection of the four villages, extensive interviews with CARE agents and our field observations did not suggest any clear patterns whereby more or less productive Andino villages were targeted for the program. The phasing-in plan for FFS seems to have been driven by practical considerations rather than expected performance; indeed, four more of the Andino villages included in our analysis were covered by the FFS program in the season immediately following the survey.

Within the FFS villages, all farmers were invited to participate in the program. The only requirement imposed on participants was that they had to attend all the training sessions. In reality, although the call for participation was open to all community members, preexisting groups took advantage of their already existing organization and formed an FFS group. As a result, most FFS participants were also participants in other farmer groups such as Andino, and all but nine Andino farmers participated in FFS. However, the participation rate in FFS remained very low during the first year of implementation, with only 45 farmers out of a population of 900 (or 5% of the farmers) participating in the program. Similarly low participation rates of 2.5% are observed for the Andino program in villages where it is offered. This low enrollment rate is largely due to the limited resources of CARE and the fact that these first FFS were planned as pilot projects. Thus, while FFS participants self-selected into the program, there are most likely many similar farmers in the large population that were not enrolled in the program.

The main objective of this article is to analyze the impact of FFS on knowledge by contrasting FFS participants to a matched control group of nonparticipants in either the FFS or the Andino programs. A secondary objective is to analyze the impact of Andino on knowledge by contrasting Andino participants to the same group of nonparticipants. Because we have a large group of nonparticipants, these two measures of impact can be performed. We are, however, limited by the small number of observations in directly testing the difference in knowledge between FFS and Andino. We thus compare and test for a difference in the two impact measures just described. To limit repetition, we report the results on the impact of FFS in detail but report on the impact of Andino only secondarily.

The data for our analysis come from a 1999 survey of potato farmers in 13 communities within the province of San Miguel, located in the department of Cajamarca. Ten of the 13 villages included in the sample are among the CARE Andino villages, including the four villages that were selected as FFS villages at the time of the survey. The sample includes all of the FFS and Andino participants as well as a random sample of nonparticipant farmers from (*a*) the four villages that have FFS programs, (*b*) six villages that have experience with CARE through Andino but do not have farmer field schools,

TABLE 1
SAMPLE OF HOUSEHOLDS

	CARE Villages with FFS Program	CARE Villages with Andino Program	Non-CARE Villages	Total
FFS participants	45	0	0	45
Work with CARE, nonparticipants in FFS	9	62	2	73
Do not work with CARE	39	181	148	368
Total number of households in sample	93	243	150	486
Total number of households in villages	900	2,337	1,278	4,515
Villages	4	6	3	13

and (*c*) three control villages. The control villages were chosen to be similar to the FFS villages in observable characteristics, such as agroclimatic conditions, distance to district capitals, and infrastructure. The distribution of households in the three types of villages is reported in table 1.

The survey was carried out over two household visits. The first visit gathered detailed plot-level data, including the costs and quantities of seed, chemical, and labor inputs for each agricultural activity (from land preparation through harvest) during the year preceding the survey. It also included a knowledge test, which was based on the curriculum of the FFS. The second visit collected information on each household member's education level and marital status, off-farm activities and credit sources, and the household's experience with agricultural and other extension services. The second visit also included a full household consumption recall for the last year and an itemized account of all household and farm assets.

Examination of the potato output-seed ratio (the quantity of seed harvested divided by the quantity of seed planted per hectare) in the sample suggests that the survey was conducted in an average year (see fig. 1).[6] According to potato experts, in Cajamarca, the distribution of output-seed ratios in figure 1 is typical for the region. A ratio below the value of 3 is considered very bad, while the range 4–6 is bad, 7–9 is regular, 10–15 is good, and above 15 is excellent. The average output ratio for the sample was 7.6 with a standard deviation of 4.2. Thirty-eight percent of the plots had productivity levels rated as bad or very bad. While normal, the wide dispersion in the output-seed ratios also illustrates the tremendous variation in productivity levels in the sample villages. This is the variable that we will use to measure the impact of knowledge on productivity.

Part of the variation in productivity arises from production losses due to

[6] Tuber scientists call this measurement the multiplication ratio. It is one of the two most commonly used productivity measures, the other measure being yield estimates based on harvest sampling (Terrazas et al. 1998).

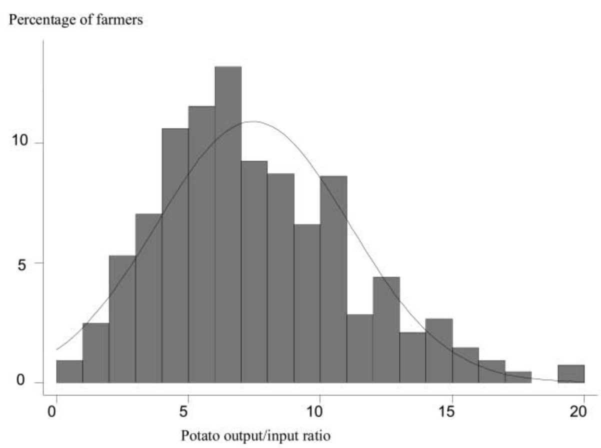

Figure 1. Histogram for potato output/input ratios

late blight, as is evident in table 2, which shows the primary causes of production losses by plot during 1998, as reported by farmers in the sample. It illustrates the need for a curriculum with a heavy emphasis on late blight. Although this was not a wet year, 47% of the potato plots in the sample experienced losses due to late blight; 19% experienced losses from frost. Neither the Andean potato weevil, the potato tuber moth, nor hail was a critical problem in the year the survey was conducted.

III. Information Channels and Knowledge Levels

Before evaluating the impact of FFS on farmers' knowledge of IPM practices, it is useful to examine how farmers in San Miguel typically obtain information on potato cultivation. The questionnaire requested farmers to name their primary sources of information on a number of tasks related to potato cultivation. Table 3 summarizes these results. The majority of farmers get information on potato farming from family members. Farmers seek information on new technologies, such as new varieties and pesticides and fungicides, from other neighbors in the community. Given the traditional, rural environment,

TABLE 2
AGRICULTURAL LOSSES CAUSED BY COMMON PESTS AND WEATHER CONDITIONS

	Source of Stress				
Plots	Late Blight	Andean Potato Weevil	Potato Tuber Moth	Frost	Hail
Percentage affected	47.1	5.4	1.0	19.3	1.4
Percentage with 0%–25% loss	31.9	3.8	1.0	7.0	1.4
Percentage with 26%–50% loss	10.8	1.6	0	6.9	0
Percentage with 51%–75% loss	3.0	0	0	2.8	0
Percentage with 76%–100% loss	1.4	0	0	2.6	0

TABLE 3

SOURCES OF INFORMATION ON POTATO CULTIVATION (% OF FARMERS WHO USE THE SOURCE)

	Family Member	Neighbor	Sharecropping Partner	Merchant/ At the Market	CARE	Other NGO	Radio	Own Experience
Agronomic practices:								
Soil preparation	96	3	0	0	0	0	0	2
Planting	96	4	.7	0	0	1	0	1
Fertilizing	92	8	1	0	1	1	0	3
Weeding and uphilling	94	4	5	0	0	0	0	3
Seed selection	88	8	1	.5	1	1	0	0
Technical issues:								
Improved varieties	52	34	1	6	6	1	2	2
Pesticide/fungicide use	73	24	2	3	1	2	1	5
Late blight control	71	23	2	2	2	3	0	q
Andean potato weevil control	28	9	.7	1	1	2	0	5
Potato tuber moth control	15	3	0	2	0	1	0	2

Note. Sources do not sum to 100% since respondents were permitted to list multiple sources.

this makes sense. Using data from several surveys in India, Foster and Rosenzweig (1995) note that information from neighbors on new technologies was as important as information from government extension services. In their study in northern India, Feder and Slade (1986) also note the extensive role of discussions among farmers as a main source of agricultural advice. Ortiz and Valdez (1993) found a similar role for neighbors for information in other Cajamarca communities. Agricultural economists working in developed countries have also noted this phenomenon (Birkhaeuser et al. 1991). For the selection of improved varieties and the control of pests and diseases, which are more technical issues, farmers cite not only family members and neighbors as their primary sources of information but also CARE (either FFS or conventional training) as an important source. Feder and Slade (1986) similarly found that farmers in their sample are more likely to seek information on complex agricultural practices from agricultural extension agents.

How accurate is the knowledge that farmers share with one another? The questionnaire included a test, designed by CIP extension experts, of farmers' knowledge about the control of the three major pests—late blight, the Andean potato weevil, and the potato tuber moth. Farmers were asked how to identify the pest and its cause, how it reproduces, and how to control it. For late blight, farmers were also asked what fungicides are used to control it, how to differentiate categories of pesticides in general and of fungicides in particular, and to name resistant varieties. Finally, farmers were asked how they select pesticides and fungicides, whether they could identify the meaning of different warning labels on the pesticides, and what precautions they take in applying and storing the agrochemicals. The scores for each topic category are presented in table 4. In general, they are very low, with average scores that do not exceed 25% of the total score.

TABLE 4
AGRICULTURAL KNOWLEDGE TEST SCORE COMPARISONS ACROSS GROUPS OF FARMERS

	All Households	FFS Participants	Andino Participants	Nonparticipants	*P*-Value*
Number of observations	486	45	64	329	
Test scores (% of maximum score):					
Knowledge on late blight	24	35	29	24	.06
Knowledge on Andean weevil	10	25	14	9	.02
Knowledge on potato tuber moth	6	15	17	6	.60
Pesticide knowledge	21	29	25	21	.04
Knowledge on resistant varieties	17	49	33	16	.00
Total test score	19	34	26	19	.00

Note. All differences between FFS and nonparticipants and between Andino and nonparticipants are significantly positive at 1%.
* For test of equality FFS = Andino.

This low level of knowledge about important agricultural problems and solutions is what motivates several NGOs to provide agricultural extension services to farmers in Cajamarca and throughout Peru. CARE-Peru works extensively in the Cajamarca region to disseminate information on new technologies through conventional transfer-of-technology agricultural extension programs (Andino) and through experimental extension programs, such as FFS. Table 4 compares test scores of the farmers who participate in the FFS and Andino programs with farmers who do not participate in any program. Farmers who participate in the FFS have significantly higher scores on tests in every area. Farmers who worked with the Andino program also score significantly higher on the tests. Finally, FFS participants have higher scores than Andino program participants overall and in all test scores but one.

IV. Empirical Approach

The purpose of the estimation that follows is to measure the impact of FFS on knowledge levels of those who participated in the program. This is the average treatment effect on the treated (ATE_1), where the treatment is participation in the program. The empirical problem we face is the typical one of filling in missing data on the counterfactual: what would knowledge levels of FFS participants have been if they had not participated in the program? Our challenge is to identify a suitable comparison group of nonparticipants whose outcomes, on average, provide an unbiased estimate of the outcomes that program participants would have had in the absence of the program. Given the nonrandom selection of program villages and farmer self-selection, simple comparisons of knowledge levels between participants and nonparticipants would yield biased estimates of program impact.

Based on program design, there are three potential sources of bias in measuring program impact. First, FFS participants are likely to differ from nonparticipants in the distribution of their observed characteristics, leading to a bias due to "selection on observables." Such a bias is likely to arise because the criteria used for FFS village selection (e.g., distance to the district capital) and participant selection can also be expected to have a direct effect on knowledge levels even in the absence of the program. We control for selection on observables in two ways. First, in the sample design, non-FFS villages were purposively selected to be similar to the FFS villages in terms of observed characteristics such as agroclimatic conditions, prevalence of potato farming, distance to the provincial capital, and so on. Table 5 reports average characteristics of households from FFS and non-FFS villages, including demographic characteristics, assets, whether farmers are credit constrained, and a measure of the severity of the El Niño shock endured the year before the survey (fraction

TABLE 5
COMPARISON OF HOUSEHOLD CHARACTERISTICS IN VILLAGES WITH AND WITHOUT FFS

	FFS Village Farmers (Mean)	Non-FFS Village Farmers (Mean)	*P*-Value*
Number of observations	93	393	
Education of household head (years)	2.4	2.4	.78
Age of household head (years)	46.0	44.8	.62
Number of family members	4.8	5.3	.13
Dependency rate	1.1	1.1	.99
Total land ownership (100 ha)	.12	.11	.66
Value of cattle assets (100 soles)	6.1	5.0	.46
Number of inherited livestock	.11	.44	.02
Value of household assets (100 soles)	1.3	.6	.47
Value of farm assets (100 soles)	.43	.47	.45
Fraction of plot lost from El Niño the previous year	.32	.25	.31
Credit constrained	.31	.25	.42

Note. 1999 exchange rate of 100 soles ≈ US$30.
* For equality of means between villages.

of the plots that were not harvested because of El Niño damage).[7] It shows that the equality in means cannot be rejected for all but one characteristic. Second, as described below, we use both regression and PSM methods to control for differences in observed characteristics between FFS participants and nonparticipants. These approaches provide an unbiased measure of program impact under the assumption of conditional mean independence, whereby preprogram outcomes are independent of participation given the variables used as controls in the regression or for matching. The fact that the FFS were part of a small pilot program makes it more likely that this assumption would be true: the sample of nonparticipants very likely includes farmers who would participate to the program, were it more generally available.[8]

A second source of bias in program impact can arise if there is diffusion

[7] Farmers were categorized as credit constrained if they answered that they did not currently have a loan because they did not have access to, or did not have a guarantee for, loans from both formal banks and NGOs. There were no farmers who are currently receiving loans who responded that they could not obtain more and hence should be categorized as credit constrained.

[8] In the area that we observed, FFS was a small-scale program, with a very low participation rate (5% of the farmers in FFS villages). If it were the case that all farmers that did not participate in the program were genuine nonparticipants in the sense that they would not participate even in a fully developed program, then the average treatment effect of the presence of a farmer field school in a village could be obtained by dividing the average treatment effect on participants by the rate of participation. However, if the very low participation rate in the program were largely due to the fact that the program itself could not expand and hence was not introduced with the same level of information as a full-fledged program, this calculation would lead to a large downward bias of the impact of a fully developed program.

of knowledge in FFS communities. In the presence of diffusion, comparing FFS participants with nonparticipants in the same village is likely to underestimate program impact. Because the program had been in operation for only one year at the time of the survey, the extent of diffusion is likely to have been low. In any event, to avoid all bias from potential diffusion within FFS communities, we exclude nonparticipants in FFS communities from the comparison group. We will return to this choice in the tests of robustness of the results. Hence, the sample that we retain {*P*(FFS) + non-FFS} includes FFS participants (*P*) from the FFS villages and non-FFS villages' farmers (excluding the participants in the Andino program).[9]

A final source of bias is that FFS participants may differ from nonparticipants in the distribution of unobserved characteristics (e.g., in farming ability that affects both the decision to participate in FFS and the desire to seek out new knowledge), resulting in selection on unobservables. In the absence of a suitable instrument for program participation, we are unable to explicitly control for selection on unobservables. However, following Altonji, Elder, and Taber (2002), we use an informal way of assessing the potential bias that could result from unobservables and find that this bias is likely small compared to the estimated impact.

The assumptions underlying the above discussion can be formally expressed as follows.

ASSUMPTION 1. Stable Unit Treatment Value (SUTV) in the retained sample (excluding nonparticipants from the FFS villages): this assumes that the treatment only affects the outcomes of those who participate, that is, there is no diffusion of knowledge from FFS participants (all in FFS villages) to control farmers (all in non-FFS villages).

ASSUMPTION 2. Ignorability of treatment (participation in FFS): conditional on observed village and individual characteristics (x_v, x_i), outcomes (y_0, y_1), and participation w are independent.

This assumption implies the weaker conditional mean independence

$$\mathrm{E}(y_0 \mid x_v, x_i, w) = \mathrm{E}(y_0 \mid x_v, x_i) \ \text{ and } \ \mathrm{E}(y_1 \mid x_v, x_i, w) = \mathrm{E}(y_1 \mid x_v, x_i), \qquad (1)$$

where y_0 and y_1 are the outcomes of interest (farmers' knowledge) without and with participation in the FFS program, w is a binary indicator of participation, and x_v and x_i denote observed village and individual characteristics, respectively.

These two conditions allow us to build a statistical comparison group for

[9] Note that this assumes that there is no diffusion from FFS farmers to farmers in non-FFS villages, which seems reasonable, given the limited time that had elapsed between training and the date of the survey.

FFS participants with similar farmers from the non-FFS villages and to estimate the impact of the FFS program by comparing the observed outcome y_1 of FFS participants with the outcome y_0 of farmers in the comparison group. We use two different estimators.

A. Estimation by Regression

The first method is based on assuming a parametric expression for the conditional mean independence (1):

$$\mathrm{E}(y_0 \mid x) = \alpha_0 + (x - \bar{x})\beta_0 \text{ and } \mathrm{E}(y_1 \mid x) = \alpha_1 + (x - \bar{x})\beta_1,$$

where x is the vector of covariates (x_v, x_i) with average value $\bar{x}$ in the treated population.

This gives the expected knowledge outcome y conditional on a given set of covariates as

$$\mathrm{E}(y \mid x, w) = \mu_0 + \alpha w + x\beta + w(x - \bar{x})\gamma, \tag{2}$$

where $y = (1 - w)y_0 + wy_1$ is the observed outcome (equal to y_1 for participants and y_0 for nonparticipants). Since the regression of y on $x, w, w(x - \bar{x})$ consistently estimates the parameters, we can derive an estimate of the average treatment effect conditional on covariates x,

$$A\hat{T}E^{Reg}(x) = \hat{\alpha} + (x - \bar{x})\hat{\gamma},$$

which can be averaged over any group of observations. In particular, the coefficient α is the average treatment effect on the treated

$$A\hat{T}E_1^{Reg} = \hat{\alpha}.$$

B. Estimation by Matching on Probability Propensity Scores

This method, developed by Rosenbaum and Rubin (1983), is based on modeling the probability of treatment given covariates, called the probability propensity score (PPS):

$$p(x) \equiv \mathrm{P}(w = 1 \mid x).$$

Suppose that two agents from the population have identical PPS. Then under the ignorability condition, the average treatment effect, conditional on the PPS and provided it is not equal to either zero or one, is equal to the expected difference in the observed outcomes between participants and matched nonparticipants:

$$\mathrm{E}[y_1 - y_0 \mid p(x)] = \mathrm{E}[y \mid w = 1, p(x)] - \mathrm{E}[y \mid w = 0, p(x)].$$

Averaging over the distribution of propensity scores in the treated population gives the average treatment effect on the treated:

$$ATE_1^{PSM} = \mathrm{E}\{\mathrm{E}[y \mid w = 1, p(x)] - \mathrm{E}[y \mid w = 0, p(x)] \mid w = 1\}.$$

Implementation of this method relies on having an estimator for the PPS, which we discuss in the next section.

V. Estimation of the Probability Propensity Score

While estimation of the average impact effect is done in the population that excludes the nonparticipants from the FFS villages because of the required SUTV assumption, this need not be the case for the independent estimation of the PPS. In fact, it is within the FFS villages that we have a better identification of the covariates that determine FFS participation, since farmers in these villages were all, to a certain extent, given the opportunity to participate.

Using the subsample {FFS} of farmers living in the FFS villages, we estimate a flexible probit model of participation, where covariates and various functions of these covariates are introduced. The estimated model can be used to predict $\hat{p}(x)$ for the population {P(FFS) + non-FFS} used for the estimation of the average treatment effect. As farmers from the non-FFS villages are not included in the estimation of the propensity score, this constitutes an out-of-sample prediction. Its validity relies on the existence of sufficient overlap of the covariates, and on the assumption that the same participation model would apply in both samples were all villages offered the FFS program. The latter is an assumption of ignorability of the choice of village for participation.

ASSUMPTION 3. Ignorability of the selection of FFS villages for participation choice: conditional on observed village and individual characteristics x_v, x_i, the choice of villages for the placement of an FFS and participation w are independent.

This assumption implies conditional mean independence:

$$\mathrm{P}(w = 1 \mid x_v, x_i, \text{presence of FFS}) = \mathrm{P}(w = 1 \mid x_v, x_i).$$

The results for the probit on FFS participation are reported in table 6. They show the importance of age, the number of family members in a household, and wealth (land and household assets) in influencing FFS participation. Interview with farmers during our fieldwork corroborated the correlation of FFS participation with the availability of labor in the household: many nonparticipants cited the lack of time and availability of labor as their main constraint in participating in the FFS program. In order to improve the prediction of treatment assignment (critical to matching methods), the model is intentionally overparameterized, using many variables and quadratic terms.

TABLE 6
FARMER FIELD SCHOOL PARTICIPATION PROBIT (DEPENDENT VARIABLE: PARTICIPATION [0/1])

	Coefficient	*P*-Value
Education of household head	−.74	.18
Quadratic term for education	.20	.14
Age of household head	−.02	.05
Number of family members	.21	.02
Dependency rate	−.27	.32
Total land ownership (100 ha)	.70	.03
Quadratic term for land ownership	1.05	.41
Value of cattle assets (100 soles)	.01	.76
Number of inherited livestock	.00	1.00
Value of household assets (100 soles)	.22	.00
Quadratic term for household assets	−.01	.00
Value of farm assets (100 soles)	.24	.66
Quadratic term for farm assets	−.23	.21
Fraction of plots lost in El Niño	1.61	.12
Quadratic term for plots lost in El Niño	−2.21	.00
Credit constraint	.16	.69
Constant	.12	.84
Number of observations	93	
Pseudo-R^2	.18	

A similar procedure (results not reported) was applied to participants of the Andino program. The same variables are significant in explaining participation as in the FFS prediction. The only qualitative difference is age, which acts negatively in FFS participation and positively in Andino participation, which is telling of the difference between the two approaches and who might benefit most. Education is insignificant in both cases.

These parameters are used to predict the probability of participating $\hat{p}(x)$, or PPS, for the sample {P(FFS) + non-FFS} that is then used to match FFS participants with observationally similar nonparticipants. Different rules of thumb could be applied to define what constitutes an observationally similar group of nonparticipants. Smith and Todd (2000) demonstrate that program impact estimates calculated using PPS methods are highly sensitive to which method is used, but robustness can be improved by restricting matches only to those participants and nonparticipants who have a common support in the distribution of propensity scores. Therefore, we derive impact estimates by applying the common support condition and further check robustness by using two different methods for selecting matched nonparticipants.

The distributions of propensity scores for FFS participants and nonparticipants are plotted in figure 2. The distribution with the darker bars is the distribution of $\hat{p}(x)$ for participants. For the purpose of matching, observations with very low or very high values of $\hat{p}(x)$ are eliminated, as they may indicate a true value of zero or one. Observations outside the support of the two

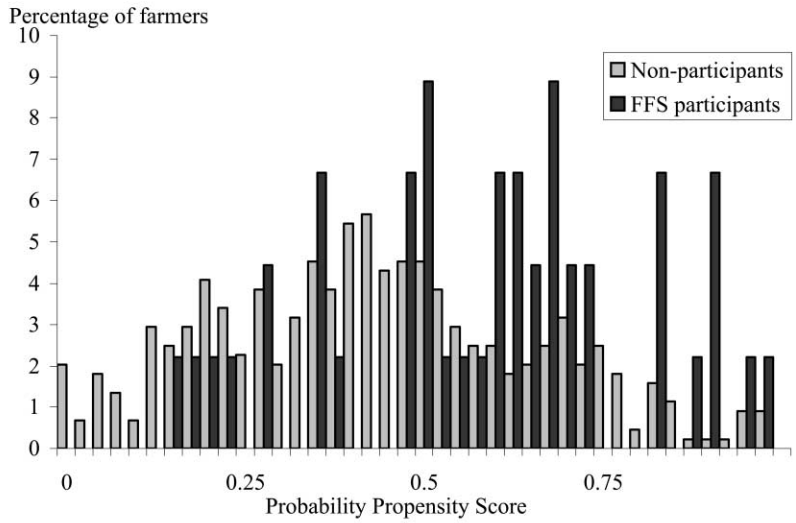

Figure 2. Histogram of probability propensity scores for FFS participants and nonparticipants

distributions of $\hat{p}(x)$ for participants and nonparticipants were also excluded from the analysis. Fifty-one observations among the nonparticipants were dropped in total.

The first method assembles a comparison group by matching each program participant with the five nonparticipants who have the closest $\hat{p}(x)$ (Jalan and Ravallion 2003). The crucial component of this method is to include non-participants with scores that are close to the participants' scores. We restricted matches to those within a 0.01 PPS distance from the FFS participant. After eliminating matches that were not within this range, the mean difference between matches was 0.005, with a maximum of 0.0099.[10] In the second method, the entire sample of nonparticipants (within the common support) is used to construct a weighted match for each participant. We use the non-parametric kernel regression method proposed by Heckman, Ishimura, and Todd (1998) for this construction.[11]

A "balancing test" reveals whether the comparison groups created with these techniques sufficiently resemble the treatment groups by testing whether the means of the observable variables for each group are significantly different (Smith and Todd 2000). For the first method, the balancing test was performed by dividing each comparison and treatment group into two strata, ordered by probability propensity scores. Within each stratum, a t-test of equality of means

[10] One FFS participant did not have a match within this range, and thus the treatment group was reduced to 44 in this method.

[11] We use the Gaussian kernel nonparametric density estimator with a bandwidth of 0.001.

TABLE 7
FFS: BALANCING TEST RESULTS FOR PPS METHODS

	Definition of Control Group		
	Method 1		Method 2
	Stratum 1	Stratum 2	
Education of household head	.42	.77	.46
Age of household head	.55	.15	.10
Number of family members	.58	.47	.23
Dependency rate	.27	.19	.20
Total land ownership	.31	.52	.41
Value of cattle assets	.82	.85	.38
Number of inherited livestock	.85	.63	.99
Value of household assets	.39	.74	.52
Value of farm assets	.30	.48	.31
Fraction of plots lost in El Niño	.76	.73	.31
Credit constrained	.45	.36	.88
Number of observations	22	22	45

Note. In method 1, the control for each participant is the average of the five nonparticipants with closest PPS (within .01 PPS) under common support. In method 2, the control is the kernel-weighted average of all nonparticipant farmers under common support.

in the two samples of participants and nonparticipants was conducted for each variable included in the probit on farmer participation. The results of these tests are reported in table 7. The null was not rejected for any variable. For the second method, we test for the equality of means in the samples of participants and their (weighted) matches. The null was not rejected for all but one variable at the 10% level, which is approximately what could be expected statistically. These results can therefore be taken to indicate no systematic differences between the experiment and comparison groups in their observed characteristics. Balancing tests for the propensity score matching of Andino participants similarly show no systematic differences in observed characteristics with their comparison groups.

VI. Impact of FFS on Knowledge

A. Estimation Based on Regression with Control Variables

As described in equation (2), in order to estimate impact based on the regression method, we regress knowledge test scores on indicators of participation in the FFS and Andino programs and on a set of household and community characteristics. Column 1 in table 8 reports our core specification based on the sample $\{P(\mathrm{FFS}) + \text{non-FFS}\}$ of all FFS participants and all households from non-FFS villages. We employ a flexible specification allowing for a full set of interactions of household and community covariates with both the FFS and

Andino variables.[12] The de-meaned variables used in the interaction terms with FFS and Andino participation are computed around their mean over the FFS participants and Andino participants, respectively. This ensures that the coefficients on indicators for FFS and Andino participation are the average treatment effect on the treated for each of these programs.

Results reported in table 8, column 1, indicate that the estimated average treatment effect on the treated is 14 percentage points for the FFS program and seven percentage points for Andino, and these values are statistically different. The effect on knowledge of traditional extension is thus lower than that of FFS, confirming the observation made on the basis of descriptive statistics in table 4. The impact of FFS on knowledge increases with land ownership, the value of household assets, and the number of family members and decreases with the age of the household head. It is interesting that deriving greater knowledge from participation in FFS is not affected by the level of education of the household head, suggesting that the very few years of formal education (2.4 years on average in the sample) have little bearing on how farmers acquire technical knowledge later. An interesting difference between the impact of the FFS and Andino programs is that in the case of Andino, knowledge is not affected by land ownership and family size and does not increase with the value of household assets. If control over land and household assets proxies for wealth, it suggests that FFS is better taken advantage of by the wealthier, while traditional transfer-of-technology approaches cater to less endowed farmers. The FFS extension method is thus better fit for younger farmers and for farmers with greater endowments.

Since we exclude nonparticipants in FFS villages from the sample, it is possible that part of the estimated FFS impact may be picking up endogenous program placement. For example, if Andino villages where there existed a motivated and knowledgeable group of farmers were selected for FFS programs, comparisons of FFS farmers with those from other villages would pick up this village effect, rather than the impact of the program. To check for this possibility, we reestimate the model with village effects by expanding the sample to include the nonparticipants from FFS villages. Village-fixed effects ensure that measures of program impact are derived essentially by comparing program participants to nonparticipants within villages and have the advantage that the results are not contaminated by village-level fixed unobservables that may

[12] Note that the same specification can be used to estimate the impacts of both FFS and Andino programs since they have the same comparison group (i.e., the farmers who participate neither in FFS nor in Andino).

TABLE 8

IMPACT OF FFS AND ANDINO ON AGRICULTURAL KNOWLEDGE TEST SCORES

	Excluding Non–FFS Participants in FFS Villages		Full Sample	
	Coefficient	*P*-Value	Coefficient	*P*-Value
Participation in FFS	13.8	.00	14.3	.00
Participation in Andino	7.0	.00	8.9	.00
Community characteristics:				
Distance from Cajamarca (km)	−.01	.00		
Dairy delivery station in community (0/1)	1.11	.23		
Household characteristics:				
Education of household head	.33	.65	.79	.28
Age of household head	−.03	.46	.00	.97
Number of family members	.07	.70	−.04	.82
Total land ownership (hectares)	7.00	.18	5.47	.29
Quadratic total land	−7.88	.03	−8.42	.04
Value of cattle assets (100 soles)	−.01	.91	.04	.54
Number of inherited livestock	.10	.66	.10	.67
Value of household assets (100 soles)	.46	.00	.37	.00
Value of farm assets (100 soles)	.42	.67	1.35	.24
Fraction of plots lost in El Niño	1.18	.33	.80	.47
Credit constrained	1.61	.10	.86	.36
Interaction terms: Participation in FFS × de-meaned community characteristics:				
Distance from Cajamarca	.00	.93		
Dairy delivery station in community	−1.62	.71		
De-meaned household characteristics:				
Education of household head	−.66	.75	−1.42	.47
Age of household head	−.30	.11	−.32	.05
Number of family members	1.33	.07	1.41	.04
Total land ownership	84.0	.08	96.7	.02
Quadratic total land	−179	.02	−197	.00
Value of cattle assets	.18	.65	−.03	.94
Number of inherited livestock	−.41	.62	−.09	.92
Value of household assets	3.33	.05	3.81	.02
Value of farm assets	−6.00	.23	−7.06	.13
Fraction of plots lost in El Niño	−10.27	.01	−9.76	.01
Credit constrained	2.90	.48	4.31	.26
Interaction terms: Participation in Andino × de-meaned community characteristics:				
Distance from Cajamarca	.02	.36		
Dairy delivery station in community	−.28	.92		
De-meaned household characteristics:				
Education of household head	1.00	.68	−.09	.98
Age of household head	.04	.67	.12	.43
Number of family members	.56	.47	−2.05	.09
Total land ownership	23.7	.56	16.1	.75
Quadratic total land	−112	.15	−11	.90
Value of cattle assets	.19	.52	−.37	.50
Number of inherited livestock	.31	.67	−.44	.66
Value of household assets	−1.09	.00	−.61	.26
Value of farm assets	5.65	.38	−11.19	.21
Fraction of plots lost in El Niño	−.08	.98	−1.95	.65
Credit constrained	.21	.95	−4.53	.30
Constant	19.7	.00	13.9	.00

TABLE 8 (*Continued*)

	Excluding Non–FFS Participants in FFS Villages		Full Sample	
	Coefficient	*P*-Value	Coefficient	*P*-Value
Number of observations	438		486	
R^2	.17		.23	
Community fixed effects	No		$F(12, 485) = 2.07$	.02
F-test for interaction terms with FFS	$F(13, 437) = 4.98$	.00	$F(11, 485) = 4.46$	.00
F-test for interaction terms with Andino	$F(13, 437) = 6.81$	.00	$F(11, 485) = 3.13$	.00
Test of $\text{ATE}_1(\text{FFS}) = \text{ATE}_1(\text{Andino})$	$F(1, 437) = 8.27$	.00	$F(1, 485) = 4.37$	.04

be upwardly biasing impact estimates. It does suffer from the disadvantage that diffusion from program participants to nonparticipants will downwardly bias estimates of impact, although as mentioned earlier, given the short time elapsed since the start of the program, this is unlikely to be important. Results of the village-fixed effects specification, reported in table 8, column 3, are remarkably similar to our core estimates: the average treatment effect on the treated is 14 and 9 percentage points for the FFS and Andino programs, respectively. Additionally, other parameter estimates do not differ greatly from the first specification, confirming our ex ante expectation that fixed village unobservables do not explain our estimates of program impact.

The validity of this simple regression method is based on the assumption that there is no selection bias, due to unobservables influencing both the choice of participation in FFS and the outcome. This means that, even though participation in the program is endogenous, conditional on observables, it is not correlated with the error term in the regression. While we have argued that this is a reasonable assumption for a pilot program such as FFS, we also use an informal calculation proposed by Altonji et al. (2002) to evaluate the potential bias that would be implied by selection on the unobservables. The idea is the following. Consider a simplified model without interaction terms,

$$y = \mu_0 + \alpha w + x\beta + \varepsilon, \tag{3}$$

where y is the knowledge score, w is an indicator of FFS participation, the parameter α is the effect of FFS on knowledge (the average treatment effect rather than the average treatment effect on the treated in the full model used above), $x\beta$ captures the role of other observed factors that influence knowledge, and ε combines all unobservables. Under certain conditions, it is possible to show that selection on unobservables is comparable in magnitude to the selection on observables in terms of its influence on the outcome y, in the sense that the normalized difference between the average values of observables and of unobservables in the two groups are the same:[13]

$$\frac{\mathrm{E}(x\beta \mid w = 1) - \mathrm{E}(x\beta \mid w = 0)}{\mathrm{var}(x\beta)} = \frac{\mathrm{E}(\varepsilon \mid w = 1) - \mathrm{E}(\varepsilon \mid w = 0)}{\mathrm{var}(\varepsilon)}. \tag{4}$$

Under these conditions, by estimating equation (3) on the sample of FFS participants and nonparticipants from non-FFS villages, we can calculate how the index of observables in the knowledge equation varies with FFS participation and then ask how large the normalized shift due to unobservables would have to be in order to explain away the entire FFS program effect. Applying this method, we find that the bias due to unobservables on the parameter α would be 2.7 points out of the average 7.6 points for the estimated average treatment effect. This is likely to be an upper bound on the bias, since the condition in equation (4) pessimistically assumes that the selected covariates in the impact regression are a random sample of the full set of covariates. In any event, the bias calculation suggests that selection due to unobservables is unlikely to wipe out the measured level of impact of the FFS program on knowledge.

One approach to correct for selection on unobservables would be to estimate probit models that explain which farmers are selected for participation in FFS and Andino, and then use the Heckman approach to correct for selection bias. This model is only weakly identified in our case, as there are no evident instrumental variables that would explain farmer participation but would have no direct effect on performance. The probit selection correction is identified by relying on a distributional assumption of joint normality of the error terms in the selection and knowledge score equations (Heckman and Robb 1985).

As an additional check for selection on unobservables, we estimate this model separately for FFS and Andino participation.[14] In both cases, we cannot reject the hypothesis that the error terms of the participation and the test score equations are not correlated, further suggesting that selection on unobservables may not be a serious problem. The estimated treatment effects are estimated to be 16.8 (SE 4.8) percentage points for FFS and 9.3 (SE 3.0) percentage points for Andino.

[13] The conditions for equality of selection on observables and unobservables are that the included regressors should be a random subset of all factors that determine the outcome, and none of the factors dominate the distribution of program participation or the outcome.

[14] For the FFS (Andino) treatment effects model, we use the sample of FFS (Andino) participants and nonparticipants in either program. We use the household characteristics retained in the regression model of table 8 as regressors in the selection equation and include village-fixed effects as well in the test score equation.

TABLE 9

FFS: TESTING KNOWLEDGE DIFFERENTIALS USING PPS MATCHING METHODS

	Test Scores (Percentage of Maximum Score)		Difference	Test of ATE_1 = 0 (*P*-Value)
	FFS Farmers	Control Group	= ATE_1	
Method 1:				
Number of observations	44			
Knowledge on late blight	35.1	25.2	9.9	.00
Knowledge on Andean potato weevil	25.3	8.5	16.8	.00
Knowledge on potato tuber moth	14.9	4.1	10.9	.00
Pesticide knowledge	29.1	20.8	8.3	.00
Knowledge on resistant varieties	49.4	16.0	33.5	.00
Total test score	34.0	18.7	15.3	.00
Method 2:				
Number of observations	45			
Knowledge on late blight	35.2	25.8	9.4	.01
Knowledge on Andean potato weevil	24.8	11.3	13.5	.00
Knowledge on potato tuber moth	14.6	6.8	7.8	.03
Pesticide knowledge	28.9	21.3	7.6	.00
Knowledge on resistant varieties	48.9	15.6	33.3	.00
Total test score	33.8	19.9	13.9	.00

Note. Method 1 shows control farmers with the five closest PPS (within .01 PPS) under common support, and method 2 shows the kernel-weighted average of all control farmers under common support.

B. Estimation Based on PPS Matching Methods

Table 9 reports estimates of FFS program impact based on the propensity score matching methods. The average difference in test scores between participants and their matches provides an estimate of the average treatment effect on the treated ($\hat{ATE}_1^{\mathrm{PSM}}$). For both methods, there is a significant difference between the two groups' scores in every category of knowledge. The scores are more than twice as high among FFS participants for knowledge of resistant varieties and knowledge of the Andean potato weevil and the potato tuber moth, and these values are very similar across methods. Gain and knowledge attributable to the FFS are greatest for the more technical issues such as identification of resistant varieties and knowledge of the most important pests (late blight and the Andean potato weevil).

The impact estimates are robust to the different estimation methods: for the overall score, the two methods give a remarkably similar estimate of 14–15 percentage points of program impact. This is also similar to the 13.8 percentage points estimate obtained with the regression method.

Knowledge scores are similarly significantly higher for Andino program participants relative to their control group, with the two matching methods and for all categories of knowledge. We compare FFS and Andino estimates in table 10. For each of the programs, we bootstrap the kernel-weights matching procedure with 100 repetitions and compute the average and standard

TABLE 10
DIFFERENTIALS IN KNOWLEDGE GAINS: FFS VERSUS ANDINO

	ATE_1 FFS		ATE_1 Andino		Test of Equality of Means
	Average	SD	Average	SD	*P*-Value
Knowledge on late blight	13.0	(3.2)	4.0	(3.0)	.00
Knowledge on Andean potato weevil	16.0	(4.7)	4.0	(3.0)	.00
Knowledge on potato tuber moth	10.9	(3.4)	9.0	(2.5)	.00
Pesticide knowledge	6.6	(2.4)	4.4	(1.4)	.00
Knowledge on resistant varieties	26.2	(4.6)	17.4	(3.2)	.00
Total test score	14.1	(2.7)	7.1	(1.5)	.00

Note. Average and standard deviations are computed from bootstrapping the kernel-weights procedures 100 times.

deviation of the estimated $\hat{ATE}_1^{\mathrm{PSM}}$. We report the *P*-value for the test of equality of means from these two samples of estimates. Results show that $\hat{ATE}_1^{PSM}$ is higher for FFS than for Andino. This is the case particularly for technical issues that matter most for farmers in the region—knowledge of late blight, the Andean potato weevil, pesticides, and resistant varieties. For the total test score, the gain in knowledge due to treatment for FFS participants (14.1, with standard deviation 2.7) is double that for Andino participants (7.1 with standard deviation 1.5). All the means of $\hat{ATE}_1^{\mathrm{PSM}}$ are significantly different between the two programs at the 1% level.

C. Robustness Tests on the Matching Results

Given the fact that matching methods are usually applied when there exists a very large population of nonparticipants for choosing proper matches, we might worry that our application of the approach to a small population might lead to results that are not robust to specific choices of variables or samples. Table 11 reports on a number of variations in the estimation procedure. Columns 2 and 3 report $\hat{ATE}_1^{\mathrm{PSM}}$ estimates obtained by extending the propensity score model to include plot and community characteristics (details on included variables are provided in the notes to the table). The extended models have similar explanatory power, but balancing tests suggest that the quality of matches is sensitive to model specification.[15] In columns 4 and 5, we return to the original probit specification but use alternative samples. In the first case, we select a random subsample (80%) of farmers from the groups of FFS participants and non-FFS participants. In the last column, we include non-

[15] For the specification in col. 2, balancing tests are rejected in 4 of the 32 cases in the five closest matching method and for 3 of the 16 variables in the kernel-weights method. In col. 3, balancing tests are rejected for 9 of 36 cases and 3 of 18 variables in the five closest and kernel-weights matches, respectively.

TABLE 11

ROBUSTNESS OF AVERAGE TREATMENT EFFECT FROM PPS MATCHING METHODS

		Extended Participation Probit		Variation on Sample	
	Base*	With Plot Characteristics†	With Plot and Community Characteristics‡	Random Subsample§	Not Excluding FFS Villages‖
Method 1:					
Knowledge on late blight	9.9	10.4	12.7	15.6	12.0
Knowledge on Andean potato weevil	16.8	16.8	16.6	15.7	16.3
Knowledge on potato tuber moth	10.9	10.7	11.6	8.7	11.3
Pesticide knowledge	8.3	8.1	6.3	7.7	6.1
Knowledge on resistant varieties	33.5	32.9	30.3	36.9	29.1
Total test score	15.3	15.3	14.5	15.9	14.1
Method 2:					
Knowledge on late blight	9.4	8.0	14.6	14.8	8.0
Knowledge on Andean potato weevil	13.5	13.5	17.1	13.3	13.7
Knowledge on potato tuber moth	7.8	4.6	11.3	9.4	4.0
Pesticide knowledge	7.6	5.6	6.4	7.7	5.5
Knowledge on resistant varieties	33.3	27.7	26.2	33.6	27.7
Total test score	13.9	11.5	14.3	15.0	11.4

Note. In method 1, the control for each participant is the average of the five nonparticipants with closest PPS (within .01 PPS) under common support. In method 2, the control is the kernel-weighted average of all nonparticipant farmers under common support.

* Base estimation as reported in table 9.

† Average plot characteristics (irrigated, rocky soil, fallowed, steep slope, area) included in participation probit.

‡ Community characteristics (distance from Camajarca and presence of a dairy station) and plot characteristics included in participation probit

§ Matching procedures performed on random subsamples of 80% of the FFS farmers and 80% of the non-FFS farmers.

‖ Nonparticipants from the FFS villages included in the pool of farmers for potential matching.

FFS participants from FFS villages (they had been excluded because of the possibility of spillover effects) in the pool for potential matches with FFS farmers. As with the previous variations, the quality of matches is indeed sensitive to the choice of sample.[16] However, the estimates of $\hat{ATE}_1^{\mathrm{PSM}}$ are remarkably similar.

In conclusion, all the variations on the matching method and the regression method yield similar results. The FFS program increases the overall knowledge test score of participants by 11–15 percentage points, while the Andino program increases knowledge of its participants by seven to nine percentage points. A few caveats are in order when interpreting these results. First, as noted above, the FFS program was introduced in addition to the regular activities of the Andino program, so that what we call FFS effect is effectively the cumulative effect of Andino activity and FFS-specific training. Second, FFS farmers were tested within the year of their specific FFS training (Andino extension activities are permanent), and hence whatever increase in knowledge that is measured is a short-term effect. Only time will tell if this knowledge lasts.

VII. From Knowledge to Productivity

As the FFS program was only in its first year of operation, we cannot expect yields of FFS participants to yet reflect acquired knowledge from it. This is because the output-input ratio is computed for the plots that were harvested during the year in which the FFS was occurring. Planting, and much of the spraying, was carried out at the very inception of the program or perhaps even before participation started. This precludes the measure of an average treatment effect of FFS on yield based on these observations.

For this reason, we choose to establish the relationship between agricultural knowledge and productivity on the 245 plots of farmers from the non-FFS communities. We regress plot-level productivity on knowledge score, controlling for plot characteristics and household productive assets and correcting for clustering at the household level. We also include village-fixed effects to control for village characteristics that may be correlated with both knowledge and productivity. Results, reported in table 12, show that a 10-percentage-point increase in knowledge score is associated with a 1.8 increase in the output-input ratio. A potential problem with this regression is the endogeneity of knowledge. It is indeed likely that more entrepreneurial farmers are both

[16] Only 1 out of 22 tests of equality of means between the treated and their matches is rejected for the five closest matching method, and no test is rejected for the kernel-weighting method in the table 11, col. 4 specification. The balancing tests reject 6 of 22 cases and 1 variable out of 11 in the two methods in column 5.

TABLE 12
IMPACT OF SCORE ON PRODUCTIVITY IN NON-CARE COMMUNITIES (DEPENDENT VARIABLE: PLOT-LEVEL POTATO OUTPUT/INPUT RATIO)

	Mean	Coefficient	*P*-Value
Knowledge score (0–100)	17.3	.18	.00
Plot characteristics:			
Area of plot (hectare)	.27	−4.30	.01
Steep slope	.07	.83	.44
Irrigated	.32	1.62	.01
Fallowed last season	.62	.12	.86
Household characteristics:			
Education of household head	2.4	−.06	.90
Age of household head	47.5	−.01	.80
Number of workers in family	5.4	.11	.53
Workers per hectare of arable land owned	1.21	.01	.97
Value of farm assets (100 soles)	.57	1.67	.08
Fraction of plots lost in El Niño	.15	−1.39	.14
Credit constrained	.28	−.51	.48
Community fixed effects included		Yes	
Dependent variable: potato output/input ratio		7.7 (SD 4.8)	
Number of plots (150 households)		245	
R^2		.15	

Note. *P*-value computed from standard errors corrected for clustering at the household level.

more knowledgeable and more productive. The OLS estimates would thus give an upward biased estimate of the effect of knowledge on yield.

Given the cross-sectional nature of the data, it is difficult to find valid instruments, that is, household variables that are correlated with knowledge and do not influence productivity. We chose as instruments the average knowledge score on varieties of farmers in the same age group in the community and its interactions with the arable land owned by the household. These instruments together represent the influence of the common knowledge in the age class, mediated by land wealth of the farmers. Although admittedly somewhat ad hoc, these instruments are statistically valid.[17] The results of the instrumental variable regression (not reported) give a higher but not significantly different coefficient of knowledge on yield of 0.29 (SE 0.18), suggesting that the OLS estimate is unlikely to be upward biased.

Using the coefficients from the regression in table 12 and the score differentials reported in table 9, we simulate the potential impact of FFS participation on productivity. Using the calculated score differential of 14 percentage points from FFS participation, this implies that FFS participation would have resulted in an increase of 2.5 points in the output-input ratio. This represents

[17] The instruments are not significant when added directly in the productivity regression. The first stage regression indicated that the instruments were strong predictors of knowledge scores ($F(2, 227) = 10.9$, $P = .000$). The overidentification test failed to reject the null hypothesis that the instruments are statistically valid ($P = .91$).

a 32% increase over the average observed output-input ratio of 7.9, which corresponds to the value in a normal year. Note that nonseed inputs are not taken into consideration in the productivity measure. Therefore, although higher knowledge scores help increase productivity, we do not know if they result in higher profits.

VIII. Conclusions

The challenge of the FFS approach is whether training results in higher knowledge about complex technical issues such as IPM and whether improved knowledge in turn translates into higher productivity. Using data on a small-scale pilot FFS program targeted to Peruvian potato farmers, this article finds that FFS participation significantly enhances knowledge on pests, fungicides, and resistant varieties—all instrumental in implementing IPM practices. The robustness of the positive results of FFS participation on knowledge is demonstrated by the fact that two separate approaches (and several variations on each of them) used for estimating the effect of FFS yield the same result—a 14-percentage-point increase in knowledge score for FFS participants.

We also find evidence that the FFS approach adds to the traditional transfer-of-technology approach in imparting knowledge of technical issues related to IPM to farmers. Gains in knowledge almost double when participants of the Andino program also participate in FFS. These results will need to be confirmed with larger samples of participants in extension programs.

A caveat of our analysis, however, is that the knowledge test was applied to FFS farmers only shortly after they completed their training. Its results thus reflect short-term knowledge acquisition. Resurveying the participants after time has elapsed would be necessary to confirm the effect of FFS in imparting lasting knowledge related to IPM.

We have no direct observation that would allow us to measure the impact of FFS participation on productivity. We therefore resort to a simulation exercise, based on the analysis of the association between knowledge and productivity, among a sample of farmers that do not participate in FFS. If this association can be interpreted as a causal relationship, and the 14-percentage-point increase in knowledge endures over time, then our results indicate that FFS participation would raise the average potato seed output-input ratio by 2.5, or approximately 32% of the average value in a normal year. Given the timing of the survey, the results that we have obtained are only suggestive. Collecting evidence to compare changes in actual productivity over time between FFS-treated and untreated farmers would be necessary to confirm these results.

References

Altonji, Joseph, Todd Elder, and Christopher Taber. 2002. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." Working paper, Northwestern University, Economics Department.

Birkhaeuser, Dean, Robert Evenson, and Gershon Feder. 1991. "The Economic Impact of Agricultural Extension: A Review." *Economic Development and Cultural Change* 39, no. 3:607–50.

Byerlee, Derek. 1994. "Technology Transfer Systems for Improved Crop Management: Lessons for the Future." In *Agricultural Technology: Policy Issues for the International Community*, ed. J. Anderson. Wallingford, UK: Cab International.

Feder, Gershon, Rinku Murgai, and Jaime Quizon. 2004. "Sending Farmers Back to School: The Impact of Farmer Field Schools in Indonesia." *Review of Agricultural Economics* 26, no. 1:45–62.

Feder, Gershon, and Roger Slade. 1986. "A Comparative Analysis of Some Aspects of the Training and Visit System of Agricultural Extension in India." *Journal of Development Studies* 22, no. 3:407–28.

Feder, Gershon, Anthony Willett, and Willem Zijp. 2001. "Agricultural Extension: Generic Challenges and the Ingredients for Solutions." In *Knowledge Generation and Technical Change: Institutional Innovation in Agriculture*, ed. S. Wolf and D. Zilberman, pp. 313–56. Boston: Kluwer.

Foster, Andrew, and Mark Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103, no. 6:1176–1209.

Heckman, James, Hidehiko Ishimura, and Petra Todd. 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65:261–94.

Heckman, J. J., and R. Robb, Jr. 1985. "Alternative Methods for Evaluating the Impact of Interventions: An Overview." *Journal of Econometrics* 30:239–67.

Jalan, Jyotsna, and Martin Ravallion. 2003. "Estimating the Benefit Incidence of an Anti-Poverty Program by Propensity-Score Matching." *Journal of Business and Economic Statistics* 21, no. 1:19–30.

Kenmore, Peter. 2002. "Integrated Pest Management." *International Journal of Occupational and Environmental Health* 8, no. 3:73–74.

Nelson, R. J., R. Orrego, O. Ortiz, M. Mundt, M. Fredix, and N. V. Vien. 2001. "Working with Resource-Poor Farmers to Manage Plant Diseases." *Plant Disease* 85, no. 7:684–95.

Ortiz, Oscar, and A. Valdez. 1993. "Enfoque de sistemas y metodología participativa para desarollar medios escritos de comunicacíón agrícola: El caso de la papa en la agricultura de substitencia." *Revista Latinoamericana de la Papa* 5–6:103–21.

Picciotto, Robert, and Jock Anderson. 1997. "Reconsidering Agricultural Extension." *World Bank Research Observer* 12, no. 2:249–59.

Praneetvatakul, Suwanna, and Hermann Waibel. 2002. "A Socio-Economic Analysis of Farmer Field Schools (FFS) Implemented by the National Program on Integrated Pest Management of Thailand." Presented at the CYMMIT International Conference on Impacts of Agricultural Research and Development, San Jose, Costa Rica, February 4–7.

Rola, A., S. Jamias, and Jaime Quizon. 2002. "Do Farmer Field School Graduates

Retain and Share What They Learn? An Investigation in Iloilo, Philippines." *Journal of International Agricultural and Extension Education* 9, no. 1:65–76.

Rosenbaum, P. R., and Rubin, D. B. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41–55.

Smith, Jeffrey, and Petra Todd. 2000. "Does Matching Overcome Lalonde's Critique of Non-experimental Estimators?" Working paper, University of Pennsylvania, Economics Department.

Terrazas, F., V. Suarez, G. Gardener, G. Thiele, A. Devaux, and T. Walker. 1998. "Diagnosing Potato Productivity in Farmer's Fields in Bolivia." Social Science Department Working Paper no. 1998-5, International Potato Center, Lima, Peru.

Van de Fliert, E., N. Johnson, R. Asmunati, and Wiyanto. 2001. "Beyond Higher Yields: The Impact of Sweetpotato Integrated Crop Management and Farmer Field Schools in Indonesia." In *Scientist and Farmer: Partners in Research for the 21st Century: Program Report 1999–2000*, pp. 331–42, Lima: International Potato Center.

Vasquez-Caicedo, Gloria, Julio Portocarrero, Oscar Ortiz, and Cristina Fonseca. 2000. "Case Studies on Farmers' Perceptions about Farmer Field School (FFS) Implementation in San Miguel Peru: Contributing to Establish the Baseline for Impact Evaluation of FFS." Report to the Development Economics Research Group from the World Bank (May).