Journal of Resources Development and Management www.iiste.org
ISSN 2422-8397 An International Peer-reviewed Journal
Vol.27, 2016

The Impact of Farmer Field School Training on Net Crop Income: Evidence from Smallholder Maize Farmers in Oromia, Ethiopia

Admassu Tesso Huluka (PhD)
Wollega University, Department of Economics, Nekemte, Ethiopia
P.O. Box. 395 Nekemte

Workneh Negatu (PhD)
Addis Ababa University, College of Development Studies, Addis Ababa, Ethiopia

Abstract
This study examines the impact of Farmer Field School (FFS) training program on the net crop income of the smallholder farmers. The FFS program was sponsored by the Ethiopian government and launched in 2010. The study aims to compare the impact of the training on net crop income of those FFS graduate and non-FFS graduate maize farmers in Oromia, Ethiopia. For this, panel data were collected in two rounds from 446 randomly selected households of three districts consisting of 218 FFS graduate farmers and 228 non-FFS graduate farmers. The analytical procedure has involved three stages: in the first stage, descriptive analyses were used to detect existence of difference in the outcome indicators between the two farmer groups. In the second stage, we have applied a semi-parametric impact evaluation method of propensity score matching with several matching algorithms. In the third stage, we have used Difference-in-Difference as robustness check in detecting causality between program intervention and the change in outcome indicators. The result of both PSM and DID estimates shows that net crop income of the FFS graduate farmers was not statistically different from those of non FFS graduates. Accordingly, a number of policy recommendations were also suggested.
Keywords: impact evaluation, accounting income, economic income, propensity score matching, difference in difference

1. Introduction

The agricultural sector has always been an important component of the Ethiopian economy. Recognizing the socio-economic roles of the sector, the Ethiopian government has also issued agriculture sector policy and investment Frame work (PIF) in 2010. The policy provides a clear statement of the goal and development objectives of the country spanning the over ten years of 2010 to 2020 with the aims to sustainably increase rural incomes and national food security through increased crop production (FDRE, 2010). In this regard, however, the policy makers seem to consider Farmer Field School (FFS) training program as panacea for increasing crop income with little understanding of the existing situations of diverse groups of smallholder farmers. In effect, the FFS training is merely considered as the best strategy to scale up the 'best practices used by the model farmers whose productivity has been more than two times higher than the average' (FDRE, 2010).

The FFS aims to give special training to some purposively selected 'model farmers,' who, in turn, are supposed to transfer the knowledge to others through their farmers' networks that are administratively organized rather than using the existing social relationship. Therefore, the selection of the 'model farmers' into the training program was made by the district level government officials in collaboration with the Kebele[i] level development agents. Although there was no as such transparent criterion guiding the selections of the model farmers, the past performance of the farmers with the adoption of technological packages, increased agricultural outputs, accessibility of the farmers in terms of geographical location and educational level were mainly considered as selection principle. Ultimately, those who were administratively sampled have attended the training sessions lasting for 15 days. There was a minimum of eight hours of training per day thereby making the total of 120 hours of training. After the completion of the model farmers' training, there were again series of meetings held with all farmers within each Kebele with the aim of briefing the essences of the training and how to organize all farmers into 1 to 5 networks called "sub-development team". The aim of the farmers' networks was to facilitate the diffusion of knowledge and the best practices from the FFS participant farmers (from now onwards, referred to as "FFS graduates") to non FFS participants. The desired outcome of FFS was to improve the crop income of the farmers. This was expected to occur through a number of means including increased awareness among the smallholder farmers, change in perceptions towards new technologies, empowerment of smallholder farmers, greater input use, adoption of new technologies, improved natural resource management, increased productivity, increased efficiency and increased access to lucrative output markets for higher output prices. In effect, policymakers have assumed as if increased crop income is necessarily a linear function of increased knowledge, increased farm technology adoption, increased efficiency, increased productivity (Admassu et al, 2015).

However, studies reveal that although knowledge is important as predisposition in adopting farm technologies, there are other conditioning factors which influence the timing and amount of technology adoption and hence increased crop income (Feder, Just and Zilberman 1985; Rola et al, 2002; Feder et al., 2004; Duflo, et al, 2006; Todo and Takahashi, 2011; Admassu et al, 2015). These authors suggest that lack of knowledge is just one of the factors hindering technology adoption but not necessarily the only factor. Admassu and Workneh (2016) have empirically examined the impact of FFS on the farmers' knowledge and farm technology adoption and found no evidence of linear relationship between increased knowledge and increased technology adoption. Study by Admassu et al., (2015) also revealed that there is little evidence supporting the impact of FFS on farmers' technical efficiency. It could be argued that FFS program would not have as such significant positive impact on the farmers' crop income in the absence of the desired impacts on technology adoption and poor effect on technical efficiency. However, as the FFS training program may increase crop income through increasing the bargaining power of the farmers and hence earn relatively higher price for their limited outputs, it is reasonable to assess the program impact on farmers' crop income. Accordingly, this paper aims to empirically examine the impact of FFS on the crop income of the two farmer groups: FFS graduates vs. non FFS graduates. To this end, we have employed two estimation methods: Propensity Score Matching (PSM) and Difference-in-Difference (DID). The former method helps to match program participating farmers and non-participating farmers based on their baseline similarities and clear out those factors to single out only program impacts. The latter approach (DID) help to difference out unobservable factors from the impact analysis process. The result reveals that the crop income of the two farmer groups is not statistically different.

2. Materials and Methods

Study area and sampling: - this study was conducted in three purposively selected major maize producer districts in the Oromia region, East Wollega zone: Guto Gida, Gida Ayana and Boneya Boshe districts. These three districts were purposively selected from the zone on the basis of their land under maize production and the role that maize crop plays in their socio-economic developments. In essence, maize crop is purposively selected because of the fact that it is Ethiopian's largest cereal commodity in terms of total production, productivity, and the number of its smallholder coverage (IFPRI, 2010).

Sample size: Following the procedures employed by IDB (2010) and World Bank (2007), we have employed power analysis for sample size determination and selected equal number of 246 smallholder farmers both from FFS graduates and non FFS graduates thereby making total sample size of 492.

Sampling strategy: First, we have selected three districts with good maize growing records. Second, from each district, we have purposively selected one kebele, from which households were randomly selected. Following the FFS program design, we have stratified our households from each Kebele into two excludable groups as: (i) FFS graduate farmers who were selected for the FFS training program, and; (ii), non-FFS graduate farmers who were exposed to the FFS training via the FFS graduates and hence supposed to follow their best practices. Finally, we made six sampling frame for the three kebeles as we have two strata in each kebele. Stratified probability-proportional-to-size sampling offers the possibility of greater accuracy by ensuring that the groups that are created by a stratifying criterion are represented in the same proportions as in the population (Bryman, 1988). Accordingly, we have divided the total samples of 492 across the Kebeles as well as between the FFS graduates and non-FFS graduates following probability-proportional-to-size sampling technique. However, although 492 questionnaires were distributed to the sampled households, we have collected 446 properly filled questionnaires with distribution across the selected study districts as 142, 160 and 144 from Guto Gida, Gida Ayana and Boneya Boshe districts respectively.

Data sources and collection techniques: Data collection was classified into two stages. In the first stage, qualitative data were collected using key informant interviews and focus group discussions. In the second stage, detailed quantitative data were collected using structured questionnaires prepared with full understanding of the nature of the program. The questionnaires were pre-tested and ensured that all included items were relevant and the questionnaire contained the correct format for the data collection. The survey was conducted in two rounds using the same questionnaire format, the same enumerators and during the same season of June to July in 2012 and 2013.

Analytical Approach: the main challenge of this study, as it is the case for other impact evaluation studies, is to decide on the correct counterfactual: *what would have happened to the crop income of those farmers who participated in the training program if the program had not existed?* Given the non-random selection of farmers for the program participation, estimating the outcome variables by using the OLS would yield biased and inconsistent estimate of the program impact due to some confounding factors: purposive program placement, self-selection into the program, and diffusion of knowledge among the program participant and non-participant farmers. Thus, our impact evaluation design should enable us to control for such possible biases. For this, we have employed two impact assessment methods: Propensity Score Matching (PSM) and Difference-in-Difference (DID). The former method helps to match program participating farmers and non-participating farmers based on their baseline similarities and clear out those factors to single out only program impacts while the latter approach (DID) helps to difference out unobservable factors from the impact analysis process. The combined use of these alternative estimation techniques is expected to lead to consistent results.

Propensity Score Matching Model: In the absence of random selections, those farmers who participated in the FFS training and those excluded from it may differ not only in their participation status but also in other characteristics that affect both participation and knowledge and their agricultural technology adoption. The Propensity Score Matching (PSM) seeks to find non-participating farmers among farmers not receiving the training that are similar to the participating farmers, but did not participated in the training program. PSM does this by matching participating farmers to non-participated farmers using propensity scores. In other words, this approach tries to replicate the model farmer selection process as long as the selection is based on observable factors (Essama-Nssah, 2006; Ravallion, 2008; World Bank 2010; IDB, 2010). Thus, PSM searches a group of "control" farmers who are statistically "similar" in all observed characteristics to those who participated in the training program.

Under certain assumptions, matching on Propensity Score, $P(X)$, is as good as matching on X. Therefore, rather than attempting to match on all values of the variables, cases can be compared on the basis of propensity scores alone, given that all observable variables which influences program participation and outcome of interest are properly identified and included (for further explanations on PSM, please see, Essama-Nssah, 2006; Heinrich et al., 2010; World Bank, 2010).

PSM constructs a statistical comparison group that is based on a model of the probability of participating in the treatment T conditional on observed characteristics X, or the propensity score is given by:

$$P(x) = pr(T = 1 \mid x) \qquad (1)$$

The propensity score or conditional probability of participation may be calculated by using a probit or a logit model in which the dependent variable is a dummy variable T equal to one if the farmer participated in the FFS training and zero otherwise (Ravallion, 2008; World Bank, 2010; IDB, 2010). Although the results are similar to what would have been obtained by using probit, we have used logit model to estimate participation equation in this study. However, in order to determine if matching is likely to effectively reduce selection bias, it is essential to understand the two underlying assumptions under which the PSM is most likely to work: Conditional Independence Assumption and Common Support Assumption.

Conditional Independence Assumption: states that given a set of observable covariates X which are not affected by the program intervention, potential outcomes are independent of treatment assignment. If Y_1 represents outcomes for participants and Y_0 outcomes for non-participants, conditional independence imply:

$$(Y_1, Y_0) \perp T_i | X_i \qquad (2)$$

This implies that selection is solely based on observable characteristics and that all variables that influence treatment assignment and potential outcomes are simultaneously observed by the researcher. Put in other words, it is to mean that after controlling for X, the participation assignment is "as good as random" and participation in the FFS training program is not affected by the outcomes of interest (Imbens, 2004; Ravallion, 2008; World Bank, 2010; IDB, 2010). This allows the non-participating households to be used to construct a counterfactual for the participating group. This assumption is sometimes called exogeneity or unconfoundedness assumption or ignorable treatment assignment (Imbens, 2004).

Clearly, this is a strong assumption since it implies that uptake of the program is based entirely on observed characteristics, and hence has to be justified by the nature of the program and data quality at hand. Although the nature of the program enabled us to justify that its uptake is based mainly on observable characteristics, we may relax such unconfoundedness assumption since we are interested in the mean impact of the program for the participants only (Imbens, 2004; Essama-Nssah, 2006; Ravallion, 2008; World Bank, 2010).

$$Y_0 \perp T_i | X_i \qquad (3)$$

This equation states that, the outcome in the counterfactual state is independent of participation, given the observable characteristics. Thus, once controlled for the observables, outcomes for the non-participant represent what the participants would have experienced had they not participated in the program.

Common Support Assumption: states that for matching to be feasible, there must be individuals in the comparison group with the same value of covariates as the participants of interest. It requires an overlap in the distributions of the covariates between participants and non-participant comparison groups. This assumption is expressed as:

$$0 < \Pr(T = 1|x) < 1 \qquad (4)$$

This equation implies that the probability of receiving FFS training for each value of X lies between 0 and 1. It ensures that persons with the same X values have a positive probability of being both participants and non-participants (Heckman, Ichimura and Todd 1998; Imbens, 2004; Ravallion, 2008). More strongly, it implies the necessity of existence of a non-participant analogue for each participant household and existence of a participant household for each non-participant household. However, since we are interested in estimating the mean effect of the intervention for the participants, as opposed to the mean effect for the entire population, we will use a weaker version of the overlap assumption which is expressed as:

$$P(x) = \Pr(T = 1|x) < 1 \quad \text{.......(5)}$$

This equation implies the possible existence of a non-participant analogue for each participant. It would be impossible to find matches for a fraction of program participants if this condition is not met. Thus, it is recommended to restrict matching and hence the estimation of the program effect on the region of common support. This implies using only non-participants whose propensity scores overlap with those of the participants. In sum, participating farmers will therefore have to be "similar" to non-participating farmers in terms of observed characteristics unaffected by participation; thus, some non-participating farmers may have to be dropped to ensure comparability (Heckman, Ichimura, and Todd, 1998; Ravallion, 2008).

The main purpose of the propensity sore estimation is to balance the observed distributions of covariates across two farmer groups (FFS graduates vs. non-FFS graduates) farmers. Hence, we need to ascertain that (1) there is sufficient common support region (overlapping of the estimated propensity scores) for the two groups of farmers, and; (2) the differences in the covariates in the matched two groups have been eliminated. These two issues are the necessary conditions for the reliability of the subsequent estimate of the program impacts. Although there are many methods of covariate balancing tests, literatures show that the standardized tests of mean differences is the most commonly applied method. Hence, we have employed two methods for this study: standardized tests of mean differences and testing for the joint equality of covariate means between groups using the Hotelling test or *F*-test. The following equation shows the formula used to calculate standardized tests of mean differences (Imbens, 2004).

$$B_{before}(x) = 100.\frac{\overline{X}_T - \overline{X}_C}{\sqrt{\frac{[VT(X) - VC(X)]}{2}}}, B_{after}(x) = 100.\frac{\overline{X}_{TM} - \overline{X}_{CM}}{\sqrt{\frac{[VT(X) - VC(X)]}{2}}} \ldots(6)$$

Where for each covariate, $\overline{X}_T$ and $\overline{X}_C$ are the sample means for the full treatment and comparison groups, $\overline{X}_{TM}$ and $\overline{X}_{CM}$ are the sample means for the matched treatment and comparison groups, and $VT_{(X)}$ and $Vc_{(x)}$ are the corresponding sample variances. Rosenbaum and Rubin (1985) suggest that a standardized mean difference of greater than 20 percent should be considered as "large" and a suggestion that the matching process has failed. In addition to test of covariate balancing, we have also checked that there is sufficient overlap in the estimated propensity scores of the two groups of farmers after matching.

Given that the above specified assumptions holds, and there is a sizable overlap in *P*(*X)* across participants and non-participants, the PSM estimator for the average program effect on the treated (ATT) can be specified as the mean difference in *Y* over the common support, weighting the comparison units by the propensity score distribution of participants (Caliendo and Kopeinig, 2005; World Bank, 2010). A typical cross-section estimator can be specified as follows:

$$ATT_{PSM} = E_{p(x)|T=1,}\left\{E\left[Y_1|\text{T} = 1, \text{p(x)}\right] - E\left[Y_0|\text{T} = 0, \text{p(x)}\right]\right\} \quad \text{.......(7)}$$

This equation shows that, PSM estimator is simply the mean difference in outcomes over the common support, appropriately weighted by the propensity score distribution of participants.

Difference in Difference (DID): this method assumes that program participation is influenced by unobserved household heterogeneity and that such factors are time invariant. Having data collected for both before and after the program on both farmer groups, the unobservable time invariant component can be differenced out by using DID. Accordingly, this section assesses the impact of FFS program on technical efficiency of the farmers using DID.

With a two-period panel data set, impact evaluation using DID method can be estimated just by pooling the two periods' data and use OLS to estimate the performance parameters (Feder, et at., 2004; Lifeng, 2010; World Bank, 2010; Admassu et al., 2015). To specify the equation, assume that a farmer *i* lives in village *j* at a time *t* reporting performance of *y*, while *x* and *z* representing the household and village characteristics that changes over time.

$$\ln Y_{ijt} = \alpha_0 + \alpha D_t + \beta FFS_{ijt} + \mu x_{ijt} \gamma z_{ijt} + \lambda_i + \eta_j + \varepsilon_{ijt} \quad \text{.......(8)}$$

Where, D_t is dummy variable for the second year after the FFS program, FFS showing dummy variable (one if the household is FFS graduate and zero otherwise), λ_i and η_j representing unobserved, time constant factors influencing program participation in household and village respectively while ε_{ijt} showing idiosyncratic error representing the unobservable factors that changes over time. However, given the non random selections of the farmers into the FFS training program, just the naïve estimation of the program impact using OLS may yield biased estimates for the reason that λ_i and η_j may be correlated with some of the explanatory variables thereby violating one of the fundamental assumptions of OLS. Thus, by subtracting the first period observations from the second period observations, equation 8 above can be condensed as:

$$\Delta \ln Y_{ijt} = \alpha + \beta FFS_{ijt} + \mu \Delta x_{ijt} \gamma \Delta z_{ijt} + \varepsilon_{ijt} \text{ ....(9)}$$

The symbol (Δ) in equation 9 above shows the differencing operator between the two periods, while both λ_i and η_j were eliminated by differencing. The dummy variable for the year of observation is also eliminated after differencing. Thus, α measures the before FFS training growth rate in performance for all farmer groups, while β measures the difference in growth rate between the FFS graduates and non FFS graduates after the FFS training program. Note that DID estimator provides unbiased FFS effects under the identifying assumption that change in outcome variable, y, for all groups of farmers would have been the same in the absence of the program although the level of y in any given year may differ (Feder, et al., 2004; World Bank, 2010). Thus, the quality of the DID estimator is that the differencing enabled us to control for the initial conditions that may have a separate influence on the subsequent changes in outcome or assignment to the treatment. As the result, any variations in performance owing to such factors (systemic climate change, price and other policy changes) that affect all farmers are eliminated and hence the individual coefficients in the model actually measure the contributions of each explanatory variable to the growth of the performance indicators.

2.1 Definitions and Measurement of Variables

2.1.1 Variables to estimate the Propensity Score

Participation in the training program (dependent variable) is a dichotomous variable taking the value 1 if the household head has participated and considered as treatment group. The household and takes a value 0 if he or she did not directly participate in the training program but could be exposed to the information conveyed in the training program through interactions with the FFS graduates and hence considered as a control unit. The independent variables include those characteristics that determined project placement in order to replicate the selection process.

2.1.2 Impact Indicator Variables

Crop income: **-** Increasing net crop income of the smallholder farmers is the final intended outcome of the FFS training program. Thus, it is reasonable to evaluate the change in the crop income across the farmers group. Here, we considered only crop income from maize product. We have computed crop income for each household for each year, using information on the price and quantity of the harvested maize as reported by the household. However, we have encountered a problem that prices reported by households vary substantially for the same product within the same village. As the result, we took average village level price for maize during each year and then multiplied the total quantity of the harvested output, including self-consumed quantities of each sampled household and the result gave us estimated gross income from maize production for each year.

Net crop income is estimated as the excess of the estimated crop income over the related costs of production for each year. For most farmers usually consider only their explicit costs in their cost benefit analysis and hence tend to overstate their net crop income, we have classified net crop income as *accounting income* which consider only explicit costs of production and *economic income*, which consider all implicit and explicit costs of production. Implicit costs include opportunity cost of unpaid labour and other materials used for maize farm activities. Such costs include family labour, own oxen labor, estimated value of compost and estimated value of traditional seeds used. Here again, we have considered village level mean opportunity cost of labor used in the maize production measured by average wage rate during each year at each village. Similarly, we have considered opportunity cost of a pair of oxen to be equals to that of opportunity cost of man-day labour at each village for each year. This is because wage rate per man-day is usually considered to be equal to a pair of oxen day labour in the study areas. Opportunity cost of compost prepared and used by each farmer was estimated by taking mean village level estimated prices for each year. The village level mean traditional maize seed price was taken just as the average prices of a kilo of maize during each year. Thus, implicit costs were estimated as the function of the quantity of family labour used for maize production measured as man-days, oxen labour measured as oxen day, and quantity of traditional maize seed used in kg and compost used in quintal multiplied by their respective village level mean prices of each year. Explicit costs include all costs of production which involved out of pocket money payment for specific households during each year. Such costs include costs for chemical fertilizers, hired labour cost, costs of improved seeds, costs of herbicides and pesticides used for maize production, land rental cost, tractor rental costs and shelling cost. As these costs significantly vary across each village, we have taken mean village level prices for each year and multiplied by their respective quantity as reported by the households. Accordingly, the excess of the estimated maize value over the explicit cost was determined as *accounting income* which most farmers consider in judging their agricultural income and the excess of the estimated crop income over the total expense, which is the sum of the implicit and explicit costs for each household for each year estimates *economic income*. Finally, the log form of the *accounting income* and *economic income* was used as dependent variable in estimating program impacts.

3. Results and Discussion

This section presents the survey results and discussions by dividing it into sections. In the first section, comparison of some selected household characteristics and maize production parameters for the baseline year is made by farmer groups. Section two presents comparison of major input and output performance indicators between the FFS graduates and non-FFS graduate farmers before the implementation of the program. Section three presents comparisons of performance indicators before and after the program implementation. Section four then presents impacts of FFS employing PSM method, while section five extends the impact assessment using DID method.

3.1 Household and farm characteristics by farmer groups

Table 1 presents the descriptive statistics for both FFS graduates and non-FFS graduate farmers. Almost in all the cases, FFS graduates were identified with the highest scores in terms of educational levels, non-farm income, family sizes, estimates of asset values, total land size as well as farm size covered by maize. Significant differences were also observed in the proportions of household head owning mobile cell phone, radio ownership, participation in farmers' cooperatives, as well as in the number of contacts with the Kebele level development agent. Those FFS graduate farmers had the highest scores than those non-FFS graduate farmers in all cases.

Table 1 . Household and Farm Characteristics During 2010 (By Farmer Groups)

	Mean		t-test	
Variables	FFS graduate	Non-FFS Graduate	t	p>|t|
Household head age	37.651	38.776	-1.220	0.222
Household head sex	0.92661	0.87719	1.750	0.081
Education level of head	3.211	1.3684	6.940	0.000
Household head literate	0.72018	0.36842	7.950	0.000
Farming Experience of head	20.472	21.395	-1.010	0.315
None farm income	1276.6	824.12	1.720	0.087
Family size	5.7569	5.2895	2.180	0.030
Distance from techno center	0.71353	0.76096	-0.720	0.473
Distance from district town	6.8145	7.1766	-0.800	0.422
Have a pair of oxen	0.73394	0.65789	1.750	0.082
Have mobile cell phone	0.33028	0.2193	2.640	0.009
Have a radio (yes=1)	0.46789	0.39035	1.660	0.099
Estimated asset value	18149	13479	2.040	0.042
Household land size (Ha)	2.0753	1.6758	2.710	0.007
Have land use certificate	0.83871	0.78947	1.330	0.183
Head is member of cooperative	0.84862	0.69737	3.860	0.000
Head received FTC training	0.36697	0.30263	1.440	0.151
Number of DA contact/year	9.5826	6.5965	2.470	0.014
Total maize farm (Ha)	1.4463	1.1012	3.620	0.000
Percent of maize land to total	89.600	86.4000	0.398	0.691

Source: Own calculation from survey data.

This significant difference between the farmers groups could be explained by the intended principles of model farmer selection criteria adopted by the government. Although there was no as such transparent criterion guiding the selections of the model farmers, the educational level of the farmers, the past performance of the farmers with adoption of technological packages, agricultural production outputs, accessibility of farmers in terms of geographical location and history of participation in farmers training centers were some of the factors considered in selecting the participant farmers.

3.2 Maize Production parameters by Farmer Groups

Table 2 presents cost and returns of maize production by farmers' groups. Comparison of costs and returns between the two farmers groups shows that FFS graduate farmers were significantly different from their counterpart non-FFS graduate farmers specifically in terms of total maize obtained, technical efficiency, as well as in income from maize production measured both in terms of accounting and economic profits. However, the difference between the two farmer groups diminishes as we compare their productivity in terms of total maize per hectare; income from maize production measured both in terms of accounting and economic profits per hectare.

Table 2. Costs And Returns of Maize Production Before The FFS Training

	Mean		t-test	
Variable	FFS Graduate	Non FFS Graduate	t	p>|t|
Total maize (kg)	6323.3	4550.7	3.590	0.000
Maize yield (kg/ha)	4048.147	3737.4	1.7977	0.0729
Technical Efficiency (index)	0.6176	0.5676	2.1280	0.0339
Accounting income(Br)	9795.7	6753.4	3.810	0.000
Accounting income/ha	6870.7	6241.5	1.670	0.096
Economic income	7972.3	5262.8	3.600	0.000
Economic income/ha	5422.2	4748.7	1.890	0.060
DAP/ha (kg)	78.893	80.401	-0.450	0.656
UREA/ha (kg)	80.547	80.401	0.040	0.967
Total cost/ha	3807.1	3693.7	0.820	0.412
Total labor/ha	55.794	56.047	-0.110	0.912
Cash cost/ha	2358.7	2200.9	1.360	0.174
Non cash cost/ha	1448.5	1492.9	-0.620	0.537
Family labor/ha	46.635	48.329	-0.680	0.496

Source: Own calculation from survey data.

Given the fact that FFS graduate farmers own larger farm size than those non-FFS graduate farmers, profit margin diminishes as we look at their profit per hectare. There was no as such apparent difference between the two farmer groups in terms of fertilizer use per hectare, total labor application per hectare and total cost per hectare.

3.3 Performance indicators by farmer groups

Table 3 presents comparisons of various input and output performance indicators between the two farmer groups before and after the FFS program intervention.

Table 3. Performance Indicators Before And After FFS By Farmer Groups

Measurement year	2010= y0		2012 = y2		Difference = y2-y0	
Parameters	mean	Std. Err	mean	Std. Err	mean	Std. Err
Maize yield/ha in kg:						
Non FFS Graduates	3737.402	121.88	4042.747	132.91	305.3447	121.86
FFS graduates	4048.147	122.48	4138.464	124.7	90.31728	89.6580
t-test	-1.798^{*}		-0.524^{ns}		1.41^{ns}	
Labor yield(kg/man-day):						
Non FFS Graduates	68.609	2.678	68.507	2.496	-0.103	2.319
FFS graduates	80.050	3.344	82.533	3.696	2.483	2.597
t-test	-2.68^{***}		-3.1698^{***}		-0.744^{ns}	
Technical efficiency:						
Non FFS Graduates	0.57	0.02	0.58	0.02	0.01	.0142
FFS graduates	0.62	0.02	0.61	0.02	-0.01	.0088
t-test	-2.13^{**}		-1.60^{*}		0.7571^{ns}	
Non cash cost/ha:						
Non FFS Graduates	1492.863	51.145	2596.646	98.682	1103.783	71.870
FFS graduates	1448.457	50.288	2373.331	83.831	924.874	48.772
t-test	0.619^{ns}		1.718^{*}		2.042^{**}	
Family labor/ha:						
Non FFS Graduates	48.329	1.778	51.433	1.902	3.104	1.3096
FFS graduates	46.635	1.735	45.964	1.648	-0.670	.901422
t-test	0.681^{ns}		2.165^{**}			2.354^{**}
Act income/ha:						
Non FFS Graduates	6241.53	264.069	11149.0	484.68	4907.50	404.26
FFS graduates	6870.686	268.868	11506.87	441.374	4636.184	315.116
t-test	-1.6693^{*}		-0.544^{ns}			0.526^{ns}
Econ income/ha:						
Non FFS Graduates	4748.664	248.552	8552.382	439.456	3803.718	372.785
FFS graduates	5422.229	255.856	9133.544	410.606	3711.315	303.522
t-test	-1.889^{*}		-0.964^{ns}			0.191^{ns}

Note: *** Significant at 1%, ** significant at 5% and * Significant at 10%, ns non-significant difference.
Source: Own calculation from survey data

A statistical comparison in Table 3 reveals that the increase in productivity achieved by the non-FFS graduate farmers is found to be almost three times the increase in the productivity of FFS graduate farmers between the two time periods. Although the FFS graduates had statistically higher maize productivity before the training year [t=1.798], the difference gradually diminished two years after the training. Vertical comparison reveals that FFS graduate farmers have maintained statistically significant labour yield both before and after the program implementation. However, comparison in terms of change in labour productivity between the two time periods reveals that the difference actually disappeared. Similarly, although FFS graduate farmers had statistically significant higher difference in terms of technical efficiency before the program implementation, this difference rapidly diminished two years after the program implementation. As a result, we couldn't find any statistical difference in terms of technical efficiency change between the two farmers groups over time period.

In addition, our analysis shows that the FFS graduate farmers have used more fertilizer per hectare and hence incurred more cash cost of production than those of non-FFS graduate farmers while the latter incurred significantly [t=2.0419] higher non-cash cost of production such as family labor, oxen and compost. Furthermore, the higher labour productivity difference in the face of lower productivity difference for the FFS graduate farmers also suggests less labour employment per hectare while the non-FFS graduate farmers increased the use of such input each year. After two years of the FFS training, crop income of the non-FFS graduate farmers both in terms of accounting and economic profits has matched with that of the FFS graduated farmers, although the latter had significantly higher net crop income during the baseline year, 2010. The consequence of FFS on technical efficiency is further investigated below employing more rigorous technique in the following section

3.4 Assessment of Farmer Field School Impacts Using PSM

3.4.1 Propensity Score Estimates

In estimating propensity score matching, the samples of program participants and non-participants were pooled, and then participation equation was estimated on all the observed covariates *X* in the data that are likely to determine participation (World Bank, 2010). We first fitted all data collected on the covariates into logit model and gradually reduced the number of the covariates until we got the desired good match. Finally, we have maintained those influential covariates determining the program participation. These covariates included comprise of different forms of assets such as natural resource (land), financial resource (access to credit), physical asset (infrastructure such as access to roads), social capital (social networks), and human forms of capital (experience and education levels). Table 4 presents the logit estimates of the FFS program participation equation.

Table 4. Estimation of Propensity Score: Dependent Variable (HH Participation in FFS)

						Number of obs =445
						Wald chi2(20)=74.71
						Prob > chi2= 0.0000
Log pseudolikelihood = -190.04376						Pseudo R^2 = 0.1549
Variables	Coef.	Robust St.Er.	z	P>|z|	[95%Conf.interval]	
Household head age	-.0108551	.026434	-0.41	0.681	-.0626648	.0409546
Household head sex (1 male)	.0938002	.3921801	0.24	0.811	-.6748586	.862459
Household education	.0955047	.0697257	1.37	0.171	-.0411551	.2321646
Household literacy (1 yes)	1.139841	.3750863	3.04	0.002	.4046854	1.874997
Farming Experience	.0138987	.025946	0.54	0.592	-.0369545	.064752
None farm income (Birr)	.0000365	.0000438	0.83	0.404	-.0000492	.0001223
Family Size	-.0275738	.0631437	-0.44	0.662	-.1513332	.0961857
Distance from techno centre	-.0086456	.1285851	-0.07	0.946	-.2606677	.2433766
Distance from district town	-.0675697	.0393377	-1.72	0.086	-.1446702	.0095308
Has a pair of oxen	.6056229	.2973728	2.04	0.042	.0227828	1.188463
Has mobile phone	.2386495	.286769	0.83	0.405	-.3234074	.8007064
Estimated asset value	7.35e-06	.0000104	0.71	0.479	-.000013	.0000277
Has land use certificate	.0971948	.3450007	0.28	0.778	-.5789941	.7733838
Head is member of coop.	.453459	.3240438	1.40	0.162	-.1816549	1.088573
Number of DA visit/year	.017125	.0101495	1.69	0.092	-.0027674	.0370178
Head has access to credit	-.524440	.3757721	-1.40	0.163	-1.260941	.2120588
Household land size (ha)	.042385	.1042641	0.41	0.684	-.1619685	.2467394
Maize farm land (ha)	.198122	.1925527	1.03	0.304	-.1792743	.5755184
Constant	-2.9335	.7304996	-4.02	0.000	-4.365277	-1.501771

The result shows that some covariates are significantly associated with FFS program participation. Educational level of the household head measured in terms of years of schooling, household head literacy measured as ability to read and write; possession of household assets such as one or more pair of farming oxen, are strongly related with FFS program participation. Furthermore, possession of mobile phone, total asset values, as well as social network such as participation in farmers cooperative, number of development agents' contact with the household per year, possession of land use certificate, possession of larger farm size were positively associated with FFS program participation. In the contrary, such covariates as age of the household head, family size, distance from centers where farm technologies were distributed and distance from the district town were negatively associated with the FFS program participation. The younger the household head, the more likely she/he is better educated and hence has more chance of being selected into the training program. These findings are consistent with the stated criteria of selecting household heads for FFS program participation as it was designed to train few affluent households, who are supposed to be easily trained and train others. This result also indicates that participation in the FFS program was mainly influenced by observable covariates and hence hidden covariates played very little role which, in turn, implies that the results of program assessment using PSM approach were unbiased and consistent.

As the main purpose of the propensity score estimation was to balance the observed distributions of covariates across two farmer groups, we need to establish that there is sufficient common support region for the two groups of farmers. We also need to be sure of that the differences in the covariates in the matched two groups have been eliminated. These two requirements are the necessary preconditions for the reliability of the subsequent estimations of the program impacts.

The predicted propensity scores range from 0.0365417 to 0.8797614 with mean value of 0.3310722 for the FFS graduates farmers, while it ranges from 0.0185319 to 0.9011666 with mean value of 0.1716005 for those non-FFS graduate farmers. Accordingly, the common support region was satisfied in the range of 0.03654173 to 0.8797614 with only 17 losses of observations (one from those FFS graduates and 16 from those non-FFS graduates farmers). Figure 1 below shows the regions of common support for the two groups of farmers.

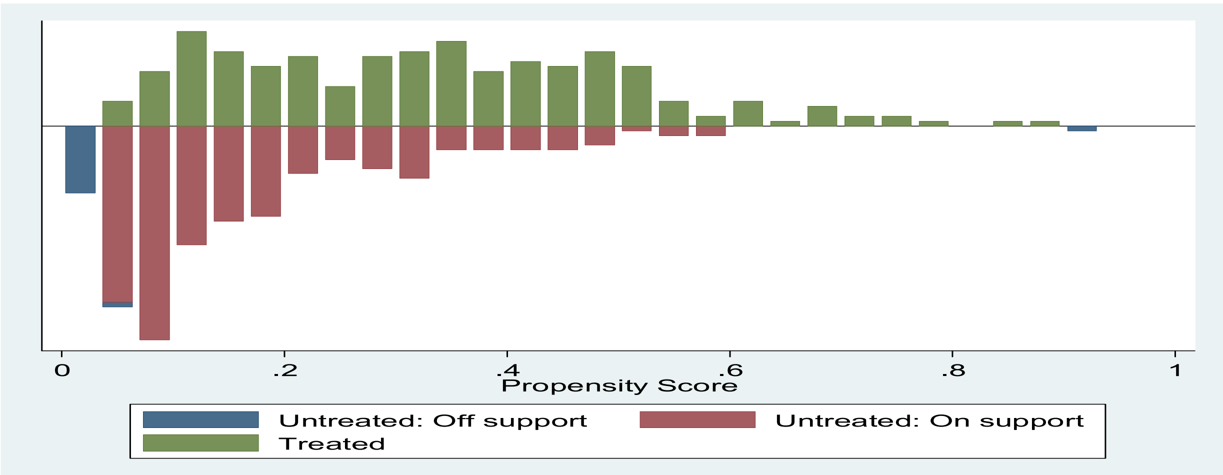

Figure 1.Propensity score distributions and common support for the propensity score estimation.

Note that "untreated off support" indicates those observations in the non-FFS graduates that do not have suitable comparison from the FFS graduates and hence excluded from the analysis while "untreated on support" indicates those observations in the non-FFS graduate that do have suitable comparison from the FFS graduates and used in the analysis. Thus, the graph clearly reveals that there is considerable overlap in the predicted propensity scores of the two groups. To verify whether the differences in the covariates in the matched two groups have been eliminated, we need to test covariate balancing. Accordingly, Table 5 presents results from covariate balancing test before and after matching. Mean standardized bias between the two groups after matching is significantly reduced for all matching algorithms suggesting that there is no systematic difference between the two groups after matching. The standardized mean difference which was around 26 percent for all covariates used in the propensity score before matching is significantly reduced to about five to seven percent after matching[ii], which has substantially reduced total bias to between 73.3 to 82.4 percent depending on which matching algorithm is used.

Table 5. Quality of Matching Before and After Matching

Algorithms	Before Matching			After Matching			Total bias reduction (%)
	Pseudo R^2	LR X2 (P-value)	Mean std Bias	Pseudo R^2	LR X2 (P-value)	Mean std Bias	
NNM	0.179	110.28 (p=0.000)	26.2	0.042	23.82 (p=0.250)	5.4	79.4
RBM (0.01)	0.179	110.28 (p=0.000)	26.2	0.037	19.58 (p=0.484)	7	73.3
RBM(0.005)	0.179	110.28 (p=0.000)	26.2	0.029	12.08 (p=0.913)	5.3	79.8
KBM	0.179	110.28 (p=0.000)	26.2	0.01	5.93 (p=0.999)	4.6	82.4

Notes : NNM = Nearest Neighbor Matching with replacements
RBM (0.01) = Radius Based Matching with replacement using caliper of 0.01
RBM (0.005) = Radius Based Matching with replacement using caliper of 0.005
KBM = Kernel Based Matching

In addition, comparisons of the pseudo R^2 and p-values of likelihood ratio test of the joint insignificance of all regressors obtained from the logit estimations before and after matching (Sianesi, 2001) shows that the pseudo R^2 is substantially reduced from about 18 percent before matching to about one percent in the case of kernel matching and to four percent with nearest neighbor matching. The joint significance of covariates was rejected since the p-values of likelihood ratio test are insignificant in all matching cases. In sum, the high total bias reduction, lower pseudo $R^{2,}$ low mean standardized bias and insignificant p-values of the likelihood ratio test after matching suggests that the propensity score equation specification is successful in terms of balancing the distributions of covariates between the two groups of farmers.

Although there are a number of methods to match the sample FFS program participants with the sampled non-FFS program households, the methods used in this analysis are the nearest neighbor matching (attnd), radius matching with two different calipers (attr 0.01 and attr 0.005) and kernel matching (attk), each with two different commands-*Psmatch2*[iii] and *Pscore*[iv].

The nearest neighbor matching method involves choosing a household from the non-FFS graduate or comparison group as a matching partner for a FFS graduate household that is closest in terms of the propensity score. Here, we have used matching with replacement, in which case the same non-FFS graduate household can be used as a match for different FFS graduate households. The problem with such matching is that difference in propensity scores for an FFS graduate household and its closest non-FFS graduate household neighbor may still be very high leading to higher standard errors and hence lower t-test result thereby increasing the chance of accepting the null hypothesis of no program impact. Thus, caliper matching is supposed to supplement the problem with nearest neighbor matching by imposing a threshold or "tolerance" limit on the maximum propensity score distance (*caliper*) between the two groups. Caliper matching, therefore, involves matching with replacement, only among propensity scores within a certain range. Accordingly, we have used two models with two different calipers of 0.01 and 0.005. This is to mean that, in the first case, we have restricted control matches to be those within 0.01 propensity score distance from the FFS graduate farmers while in the second case we restricted matches to 0.005 propensity scores differences. The smaller the caliper chosen, the higher the number of dropped samples is likely, thereby potentially increasing the chance of sampling bias. Lastly, to correct the problem of such loss of data which is common with the caliper matching, we have used the kernel matching method, which matches an FFS graduate household to all non-FFS graduate households weighted in proportion to the closeness between the FFS graduate and the non-FFS graduate households. In this case, non FFS households receive weights based on the distance between their propensity score and the propensity score of the FFS graduate households to which they are being matched. In essence, control households with the lowest propensity score distance receives the highest scores and vice versa.

Asymptotically, all the four matching methods with two different command types are supposed to lead to the same conclusion although the specific results may not be necessarily the same. This is to mean that, if the FFS impact on any of the impact indicator is robust, findings from most matching algorithms must lead to the same conclusion. Thus, such use of different matching algorithms with two different command types is used as effective robustness check of the estimated program impacts, which is again, to be confirmed by the impact assessment using DID in the subsequent section 3.5.

3.4.2 Impact estimation using PSM

Our main interest in this section is to see if the FFS training program has increased the crop income of the FFS graduate farmers as compared to non- FFS graduates. To this end, we have evaluated the impact of FFS training program on net crop income by differentiating *accounting income* from *economic income* for the sake of disposition. As most smallholder farmers in the study areas are less educated, they usually consider only explicit cost of production in making cost benefit analysis of their production decisions. Hence, *accounting income* is meant to evaluate the impact of the FFS on net crop income as perceived by the farmers themselves. In contrast, *economic income* measures the net crop income by considering the excess of the estimated crop income over the sum of all the explicit and implicit costs of production. Table 6 provides the estimates of crop income growth rate comparison among the farmer groups.

Table 6. Comparison of Crop Income /Ha Across Farmer Groups

		Observations (N)		Accounting income growth (%)			Economic income growth (%)		
	Algorithms	FFS	Non FFS	ATT	Std.Err	t	ATT	Std.Err	t
Psmatch2	Attnd	204	211	-0.010	0.094	-0.110	0.025	0.146	0.170
	Attr0.01	189	211	0.015	0.085	0.180	0.095	0.131	0.730
	Attr0.005	148	211	-0.055	0.084	-0.650	-0.011	0.122	-0.090
	Attk	204	211	-0.006	0.074	-0.080	0.040	0.120	0.340
Pscore	Attnd	217	94	0.002	0.237	0.009	0.273	0.258	1.058
	Attr0.01	191	212	0.220	0.163	1.348	0.228	0.278	0.818
	Attr0.005	174	199	0.205	0.168	1.217	0.306	0.348	0.880
	Attk	217	212	0.374	0.283	1.321	0.275	0.169	1.632

Source: Own calculation from survey data

The result shows that the estimated coefficients are very small, inconsistent among different matching algorithms and all statistically insignificant implying that the FFS graduate farmers do not seem different from other farmers in terms of their crop income (accounting as well as economic income). The result is also consistent with the implications of descriptive statistics explained above. Our descriptive statistics shows that although the FFS graduate farmers and non-FFS graduate farmers have achieved mean economic income of Birr 5,422.30 and Birr 4,748.7 per hectare respectively during the baseline year of 2010, the figures become 9,133.50 and Birr 8,552.40 during harvesting year of 2012. These show a decreasing trend of income per hectare difference between the FFS graduates and non FFS graduates from Birr 673.60 in the base year to just income difference of Birr 581.20 during 2012 harvesting year (refer to Table 3 above). This, in turn, implies relatively larger income growth rate registered by the non FFS graduates over time.

3.5 Impact Estimation using DID

In this section, net crop income of the two farmer groups was used as dependent variable in the impact estimation function specified by equation 9 above. In addition to the participation dummy of FFS, various household and village characteristics were also included as explanatory variables. However, as most household and village characteristics were almost stable over the three years, most of them were eliminated by differencing operation. As there could be significant differences of performance among farmers in different districts, it is meaningful to include two district dummies Guto Gida and Boneya Boshe to control for the district specific unobserved factors, while Gida Ayana was made implicit in this case.

Since heteroscedasticity may cause problem to the "difference in difference "models (Wooldridge, 2002; Leifeng, 2010; World Bank, 2010; Admassu et al,2015), we have tested for the existence of such problems. We have observed that Breusch-Pagan Tests detected existence of significant heteroscedasticity for estimated functions. Therefore, we have reported the robust standard errors as correction for heteroscedasticity problem. However, since there was only one period left after differencing, there was no need of testing for serial correlation in the model.

Consistent with the PSM estimates reported above, there seems no statistically difference between the two farmer groups in terms of accounting income growth rate. Table 7 provides the estimates of technical efficiency growth rate.

Table 7. Estimated Coefficients for Accounting Income Per Hectare

Dependent variable: Log of accounting income per hectare (Birr)						
N= 446			R^2= 0.514	F= 5.16		P= 0.0000
Variables	Coef.	Robust St. Err	t	P>|t|	95% Conf. Interval	
Constant	0.6940	0.0820	8.470	0.0000	0.5331	0.8549
FFS Graduates	-0.0798	0.0851	-0.940	0.3480	-0.2468	0.0872
Plough frequency	0.4494	0.3906	1.150	0.2500	-0.3172	1.2160
Fertilizer	0.0001	0.0025	0.030	0.9730	-0.0048	0.0050
Farm size	-0.0241	0.0608	-0.400	0.6920	-0.1433	0.0952
Family labor	0.0133	0.0034	3.890	0.0000	0.0066	0.0199
Hired labour	0.0032	0.0028	1.140	0.2550	-0.0023	0.0087
Herbicide	-0.0303	0.0288	-1.050	0.2940	-0.0869	0.0263
Tractor use	0.0000	0.0000	-1.060	0.2890	-0.0001	0.0000
Compost	0.0041	0.0060	0.690	0.4930	-0.0077	0.0160
DA visits	-0.0080	0.0068	-1.180	0.2400	-0.0213	0.0053
Guto Gida	-0.3661	0.1141	-3.210	0.0010	-0.5901	-0.1421
Boneya Boshe	-0.4869	0.0990	-4.920	0.0000	-0.6811	-0.2927

The constant term is statistically significant with positive coefficient implying that there is natural growth rate in accounting income owing to partly improvement in productivity and partly improvement in the estimated selling price of maize output. Although statistically insignificant, the model shows FFS graduate farmers are identified with relatively lower growth rate in accounting income. In contrast, the family labour has the expected and statistically significant coefficient which implies that a 10 percent increase in family labour application increases accounting income growth rate by 0.133 percent. Although positive, the coefficient of fertilizer application is almost close to zero and statistically insignificant. This could be because soils in the study areas are actually deficient not only in Phosphorus and Nitrogen supplied by DAP and Urea but also in other micronutrients equally important for enhancing maize crop yield. Hence, given the ever rapidly increasing trend of fertilizer prices which usually exceeds the increase in the productivity and estimated output prices, the contribution of such inputs to net crop income is inevitably insignificant. The coefficients of dummy variables representing Guto Gida and Boneya Boshe districts are found to be negative and statistically significant. These coefficients imply that, on average, sampled farmers in both Guto Gida and Boneya Boshe districts have achieved lower accounting income growth rate than farmers in Gida Ayana. Table 8 provides estimated economic income growth rate over time. The result is consistent with the model estimates for the growth rate of accounting income except some difference in the sizes of the coefficients and degree of the statistical significance for the estimates.

Table 8. Estimated Coefficients For Economic Income Per Hectare

Dependent variable: Log of Economic income per hectare (Birr)						
		N= 446		R^2= 0.620	F= 6.16	P= 0.0000
Variables	Coef.	Robust Std. Err	t	P>|t|	95% Conf. Interval	
Constant	0.7362	0.0867	8.490	0.0000	0.5660	0.9065
FFS Graduates	-0.1071	0.0730	-1.470	0.1430	-0.2503	0.0362
Plough frequency	0.5486	0.3521	1.560	0.1200	-0.1426	1.2398
Fertilizer	0.0013	0.0020	0.630	0.5310	-0.0027	0.0052
Maize farm	0.0207	0.0406	0.510	0.6100	-0.0590	0.1004
Family labor	0.0066	0.0029	2.300	0.0220	0.0010	0.0122
Hired labour	0.0036	0.0023	1.580	0.1150	-0.0009	0.0080
Herbicide	-0.0006	0.0257	-0.020	0.9830	-0.0509	0.0498
Tractor use	0.0000	0.0000	-0.680	0.4940	0.0000	0.0000
Compost	0.0070	0.0044	1.580	0.1150	-0.0017	0.0156
DA visits	-0.0043	0.0046	-0.930	0.3520	-0.0132	0.0047
Guto Gida	-0.2881	0.0906	-3.180	0.0020	-0.4660	-0.1103
Boneya Boshe	-0.5357	0.0938	-5.710	0.0000	-0.7198	-0.3516

Although statistically insignificant, the estimated coefficient for the FFS training with negative sign implies that FFS graduate farmers have achieved 10.7 percent lower economic income growth rate than the non-FFS graduates farmers. This finding is also consistent with earlier study by Feder et al., (2004) who had concluded that the FFS training program in Indonesia did not have significant impacts on the net crop income performance of graduates and their neighbors. Our result also supports the study on the impacts of FFS program in Ethiopia on technical efficiency growth rate by Admassu et al, (2015) and impact of the same program on knowledge and farm technology adoption (Admassu and Workneh, 2016). Furthermore, the result is consistent with the results reported by Praneetvatakul and Waibel (2006) concluding absence of evidence for positive impact of FFS program on net crop income of the participants. However, our result does not support the literature review reported by Braun and Duveskog (2008) indicating evidence for existence of high FFS impacts on yields and farm profits in Vietnam, Ghana, Cote d'Ivoire and Burkina Faso.

4. Summary and Conclusions

This paper assesses the impacts of Farmer Field School (FFS) on crop income of the two farmer groups: FFS graduate farmers and non-FFS graduate farmers. The FFS training program was sponsored by the Ethiopian government in 2010. To see the impact of the program on these two impact indicators (accounting income and economic income), we have employed two impact assessment models: Propensity Score Matching (PSM) with numerous matching algorisms and Difference-in-Difference (DID). The PSM method helps to match program participating farmers and non-participating farmers based on their baseline similarities and clear out those factors to single out only program impacts. The DID approach helps to difference out unobservable factors from the impact analysis process.

Both PSM and DID estimates show that net crop income of the FFS graduate farmers (measured in terms of accounting and economic income) was not statistically different from those of non FFS graduates. Although statistically insignificant, both model estimates show that participation in the FFS training program has reduced the participants' income growth rate compared to those non FFS graduates. Our descriptive statistics of the mean crop income growth rates of the two farmer groups has also confirmed the same conclusion. According to our model, family labour was the main explanatory factor for the net crop income growth rate. However, those FFS graduates were identified with relatively lower family labour application per hectare owing to their heavy involvement in political issues thereby reducing their available labour for their own agricultural activities. Admassu et al, (2015) has explained that "although the FFS graduate farmers have higher knowledge test score than other farmers, they couldn't use their knowledge to increase their technical efficiency for the reasons that the FFS program have put disproportionately higher burdens on the FFS graduates which sharply contradicts with farmers' own production decisions".

Overall, our analysis shows that the training program was implemented in the study areas without thorough understanding of the principles of FFS approach and the context within which it is expected to bring the desired impacts. The program implementation has been suffering from a number of impediments which basically emerges from the nature of the program itself. These include, (1) top-down approach of the program design and its implementation which has failed to consider the rural producers' heterogeneity in terms of production conditions and prioritization of problems, (2) difficulty in supplying the required resources and logistical support for the training program, (3) lack of well trained and technically competent facilitators, and (4) use of FFS program and farmers' networks as political tools to widely diffuse the developmental state ideology of the ruling party. These findings suggest a number of policy implications: farmers' networks and organizations need to be formed by the smallholder farmers own freewill in a way it promote their "human agency"; policy makers need to understand the rural producers' heterogeneity and be able to avoid "one size fits all" approach in development endeavors; the necessary arrangements (competent facilitators, credit facilities, and other logistic supports) need to be in place before launching FFS training programs; government need to clearly separate activities required for agricultural transformation from activities required for political issues.

Notes

[i]Rosenbaum and Rubin (1985) suggested that a standardized mean difference greater than 20 percent should be considered too large and an indicator that the matching process has failed.

[i]i Psmatch2 is Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and Covariate imbalance testing developed by Leuven and Sianesi (2003).

[iii] Pscore was developed by Becker and Ichino (2002) for the estimation of average treatment effect based on propensity score. Although the estimated effects under both commands may differ, both estimates are expected to lead to the same conclusion if the detected impact estimation results are robust enough.

References

[i] Kebele is the lowest administrative unit in Ethiopia.

Admassu, T. H., & Workneh, N. (2016). The Impacts of Farmer Field School Training on Knowledge and Farm Technology Adoption: Evidence from Smallholder Maize Farmers in Oromia, Ethiopia. Journal of Economics and Public Finance , 2 (1), 1-23.

Admassu, T. H, Workneh, N., & Sisay, A. (2015). Does Farmer Field School Training Improve Technical Efficieny? Evidence From Smallholder Farmers in Oromia, Ethiopia. *Journal of Economics and Sustainable Development , 6* (11), 49-64.

Admassu, T. H. (2015). The Impact of Farmer Field School Training on Farmers' Technical Efficiency: Evidence from Smallholder Maize Farmers in Oromia, Ethiopia. *European Journal of Training and Development Studies , 2* (3), 1-28.
Anderson, J. R., & Feder, G. (2007). Agricultural Extension. (R. Evenson, & P. Pingali, Eds.) Handbook of Agricultural Economics , Volume 3, 2344-2367.
Anderson, J.R. (2007). Agricultural advisory services (background paper to ―Science and technology for pro-poor growth‖ chapter). In World development report 2008, ed. D. Byerlee. Washington, D.C: World Bank.
Bamlaku .A., E-A Nuppenau and H. Boland (2009). Technical Efficiency of Farming Systems across Agro-ecological Zones in Ethiopia: An Application of Stochastic Frontier Analysis. Agricultural Journal 4(4): 202-207,
Battese, G.E. and T.J. Coelli .(1995). A Model for Technical Inefficiency in a Stochastic Frontier Production Function for Panel Data. Empirical Economics, 20:325-332.
Bereket Kebede and Daniel John Zizzo (2011) Envy and Agricultural Innovation: An Experimental Case Study from Ethiopia, University of East Anglia.
Braun, A. and D. Duveskog (2008) "The Farmer Field School Approach – History, Global Assessment and Success Stories" Background Paper for the IFAD Rural Poverty Report 2011, October 2008
Bravo-Ureta, Boris E., and Robert E. Evenson. (1994). Efficiency in Agricultural Production: The Case of Peasant Farmers in Eastern Paraguay. *Agricultural Economics* 10, no. 1: 27–37.
Bryman, A. 1988. *Quantity and Quality in Social Research,* London: Unwin Hyman
Caliendo, M., and S. Kopeinig. (2005) "Some Practical Guidance for the Implementation of Propensity-score matching". Iza Discussion Paper 1588. Institute for the Study of Labor (IZA).
Douthwaite, B., Kuby, T., van de Fliert, E., & Schulz, S. (2003). Impact pathway evaluation: an approach for achieving and attributing impact in complex systems. Science Direct , 243-265.
Douthwaite, B., Schulz, S., Olanrewaju, A. S., & Ellis-Jones, J. (2007). Impact pathway evaluation of an integrated Striga hermonthica control project in Northern Nigeria. Agricultural Systems 92 , 201-22.
Duflo, Esther, Michael Kremer and Jonathan Robinson, (2006) Understanding Technology Adoption: Fertilizer in Western Kenya: Evidence from Field Experiments
Essama-Nssah, B. (2006). Propensity Score Matching and Policy Impact Analysis: A Demonstration in Eviews. Poverty Reduction Group (PRMPR), The World Bank Washington, D.C.
Farrell, M. J. (1957). The measurement of productive efficiency. Journal of Royal Statistical Society, (Series A), 120:253-281.
FDRE, MoARD.(2010). Ethiopia's Agriculture Sector Policy and Investment Framework: Ten Year Road Map (2010-2020). (Prepared by Demese Chanyalew Berhanu Adenew, and John Mellor) Addis Ababa, Ethiopia, May 3, 2010
FDRE, MoFED .(2014). Growth and Transformation Plan Annual Progress Report for F.Y. 2012/13. Addis Ababa, Ethiopia, February 2014
FDRE, MoFED .(2014). Project Proposal for Blended Fertilizer Plant Establishment at Gibe Didessa Farmers' Cooperative Union. A proposal to build a sustainable fertilizer blending business in Ethiopia and drive local adoption of blended fertilizer, Addis Ababa, Ethiopia, March 2014
Feder, G.(1985), "The Farm Size and Farm Productivity: the Role of Family Labour, Supervision and Credit Constraints", Journal of Development Economics 18: 297-313.
Feder, G., R. Murgai, and J.B. Quizon. (2004b) "The acquisition and diffusion of knowledge: The case of pest management training in farmer field schools, Indonesia". Journal of Agricultural Economics 55 (2): 217–239.
Feder, Gershon, Rinku Murgai, and Jaime Quizon. (2004a). Sending Farmers Back to School: The Impact of Farmer Field Schools in Indonesia. Review of Agricultural Economics 26, no. 1:45–62.
Godtland, Erin, Elisabeth Sadoulet , Alain de Janvry , Rinku Murgai , and Oscar Ortiz. (2004). "The Impact of Farmer Field Schools on Knowledge and Productivity: A Study of Potato Farmers in the Peruvian Andes". Economic Development and Cultural Change 52 (1): 129-158.
Heckman, James J., Ichimura, Hidehiko, and Todd Petra (1998). Matching as an Econometric Evaluation Estimator. Review of Economics Studies 65:261-294.
Heinrich. C, A. Maffilo and G. Vazquez . (2010). A Primer for Applying Propensity-Score Matching. Impact-Evaluation Guidelines. Technical Notes No. IDB-TN-161. August 2010
IDB- Inter-American Development Bank (2010). Designing Impact Evaluations for Agricultural Projects Impact-Evaluation Guidelines Technical Notes No. IDB-TN-198 December 2010. Inter-American Development Bank, Washington, D.C.
IDB -Inter-American Development Bank (2010). Development Effectiveness Overview, Special Topic, Assessing the Effectiveness of Agricultural Interventions. Inter-American Development Bank, Washington, D.C.

IFPRI. (2010). Fertilizer and Soil Fertility Potential in Ethiopia: Constraints and opportunities for enhancing the system, July, 2010. http://www.ata.gov.et/projects/ethiopian-soil-information-system-ethiosis/

Imbens, Guido W. (2004). Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. The Review of Economics and Statistics **86**(1): 4-29.

Lifeng, W. (2010). Farmer Field School and Bt Cotton in China: An Economic Analysis. Doctoral Study. University of Hannover. Pesticide Policy Project Publication Series Special Issue No. 15, Murch 2010.

Praneetvatakul, S. and H. Waibel. (2006) “Impact Assessment of Farmer Field School using a Multi-Period Panel Data Model”. Paper presented at International Association of Agricultural Economics conference, 12-18 August 2006, Gold Coast, Australia.

Ravallion, M. (2008). Evaluating Anti-poverty Programs. In Handbook of Development Economics, vol. 4, ed. T. Paul Schultz and John Strauss, 3787 846. Amsterdam: North-Holland.

Rola, A.C., S. Jamias and J.B. 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(1): 65-76.

Sianesi, Barbara. (2001). Implementing Propensity Score Matching Estimators with STATA. Presentation at the UK Stata Users Group, VII Meeting. London (May).

Soniia, D., & Christopher, A. (2011). Farmer Knowledge As An Early Indicator Of Ipm Adoption: A Case Study From Cocoa. Journal of Sustainable Development in Africa , 13 (4), 213-224.

Timer, C.P. (1988)’ The Agricultural Transformation’, In: H. Cheneo and T.N. Srinivasan, eds., Handbook of Development Economics, Volume I, Elsevier Science Publishers B.V., 1988y: 276-328

Todo, Y. and R. Takahashi. (2011). Impact of Farmer Field Schools on Agricultural Income and Skills: Evidence from an Aid-Funded Project in Rural Ethiopia. Impact Evaluation Analyses for the JICA Projects. JICA-RI Working Paper. N0.30, May 2011.

World Bank, (2010). Handbook on impact evaluation: quantitative methods and practices. Public Disclosure Authorized, 52099, Washington, DC