Does Farmer Field School Training Improve Technical Efficiency? Evidence from Smallholder Maize Farmers in Oromia, Ethiopia

Admassu Tesso, Wollega University
Workneh Negatu, Addis Ababa University
Sisay Asefa, Western Michigan University
Adugna Lemi, University of Massachusetts
Bamlaku Alemirew, Addis Ababa University

Abstract

This study carries out the impact evaluation of Farmer Field School (FFS) training program on the technical efficiency of smallholder farmers. The FFS program was sponsored by the Ethiopian government and launched in 2010 to scale-up best agricultural practices in the country. The study aims to compare changes in the technical efficiency of those FFS graduate and non-FFS graduate maize producing farmers in Ethiopia, Oromia. For this, panel data were collected in two rounds from 446 randomly selected households from 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 the 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 to estimate the program impact. In the third stage, we used Difference-in-Difference as robustness check in detecting causality between program intervention and the technical efficiency. The combined uses of these alternative estimation techniques indicate that the program has a negative impact on the technical efficiency of the FFS graduates. Numerous plausible explanations for this outcome are discussed, and recommendations for improvements are suggested accordingly.

Key words: *impact evaluation, technical efficiency, propensity score matching.*

The agricultural sector has always been an important component of the Ethiopian economy. During 2012/13, agriculture accounted for 42.7 percent of the gross national product (GDP), 80 percent of employment and over 70 percent of total national foreign exchange earnings. In contrast, industry and service sector accounted for 12.3 and 45 percent of GDP, respectively, during the same period (MoFED, 2014).

A unique feature of Ethiopian agriculture is the role of smallholder farms in the total national output production and labour employment. For example, of the total production of 251 million quintals in 2012/13, about 96 percent (241 million quintal) was produced by the smallholder farmers and the rest 4 percent (10 million quintal) was produced by commercial farms. On the average, land holding share of 83 percent by smallholders farming setup is less than 2 hectares and the average size of the small farms is about 1.25 hectare in Ethiopia. These data clearly denote that small farms are the main sources of the production and employment generation in Ethiopia. Evidence also suggest that small farms provide a more equitable distribution of income and an effective demand structure for other sectors of the economy (Bravo-Ureta and Evenson 1994). Thus, the current strategic focus on increasing the productivity and production of smallholder farmers in socio-economic development of the country is justified.

Accordingly, the Ethiopian government has issued agricultural policy and investment frame-work (PIF). PIF provides a clear statement of the goals and development objectives of the country spanning the roughly ten years between 2010 and 2020. The development objective, as stated in the policy document, aims to sustainably increase rural incomes and national food security through increased production and productivity. To this end, farmer field school training is considered as the best strategy to scale up the best practices used by the model smallholder farmers whose productivity is more than two times higher than the average (FDRE, 2010).

The aim of FFS is to give special training to some purposively selected 'model farmers,' who, in turn, are supposed to transfer the knowledge to the rest through their farmers' networks that are administratively organized rather than using the existing social relationship. Accordingly, the selection of the 'model farmers' into the training program was made by the district level government officials in collaboration with the Kebele level (the lowest administrative unit in Ethiopia) development agents. Although there was no as such transparent criterion guiding the selections of the model farmers, the past performance of the farmers with adoption of technological packages, agricultural production outputs, accessibility of the farmers in terms of geographical location and educational level were mainly considered as selection criteria. Ultimately, those who were administratively sampled attended all 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 network called "sub-development team" so as to facilitate the diffusion of knowledge and best practices from the FFS participant model farmers from now onwards, referred to as "FFS graduates" to non FFS participants. The desired outcome of FFS was to increase technology adoption and technical efficiency of the smallholder farmers as means to increase their production and productivity. In effect, policymakers have focused their attention on increasing the adoption of new technologies and improving their technical efficiency as means to increase smallholder farmers' productivity and crop income.

However, the prices of new technologies are increasing in the face of capricious output prices and declining farm holding sizes which discourage such technology adoptions. Furthermore, presence of possible technical inefficiency means that output can be increased without the need for new technology. If there appears significant inefficiency among the smallholder farmers, then, the agricultural policy should gear towards training them on how to increase their efficiency with the existing technology. Increasing the adoption of more expensive agricultural technology may result in liquidating the existing meager assets of the rural producers with very little productivity gain. This calls for increasing productivity and production through optimum and efficient uses of existing technologies (Bravo-Ureta and Pinheiro 1993). However, studies that systematically analyze the impacts of FFS on technical efficiency of the smallholder farmers are lacking. Therefore, this study aims to fill this knowledge gap. To this end, the paper aims to empirically examine the impact of FFS on the technical efficiency of the two farmer group. 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 later approach (DID) help to difference out unobservable factors from the impact analysis process.

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 plays in their socio-economic developments. In these areas, cultivation of maize crop occupies an important place in the crop production plan of the farmers. For this study, 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: (a) FFS graduate farmers who were selected for the FFS training program, and; (b), 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 since 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.

Estimation of Technical Efficiency

Stochastic frontier production function is widely proposed efficiency measures for the analysis of farm-level data (Farrel, 1957; Battese, 1995; Bamlaku et al., 2009). Thus, we have used the technical efficiency model specified by (Battese and Coelli, 1995) which allows a stochastic frontier production function for panel data with farmer effects that can vary systematically over time and are assumed to be distributed as truncated normal random variables. The model can be specified as:

$$Y_{it} = X_{it}\beta + (V_{it} - U_{it}) \quad i = 1, 2, ...N, \quad t = 1, 2, ...T \text{ ....(1)}$$

Where, Y_{it} is the logarithm of the production of the *i*-th household in the *t*-th time period, X_{it} is vector of values of known functions of inputs of production and other explanatory variables associated with the i-th firm at the *t-th* observation; and β is a vector of unknown parameters. Here, the error term comprises two separate parts, V_{it} are random variables outside the control of the households which affects the productivity of the households and assumed to be identically and independently distributed (iid) $N(0, \sigma_v^2)$ and independent from U_{it} ; U_{it} represents factors contributing towards technical inefficiency but which are supposed to be within the control of the households.

The measure of technical efficiency is equivalent to the ratio of the production of the *i*-th household in the *t*-th time period to the corresponding production value of the frontier household whose U_i is zero. Thus, it follows that given the specifications of the stochastic frontier production function defined by equation (1), the technical efficiency of the *i*-th household in the *t*-th time period can be defined by:

$$TE_{it} = (X_{it}\beta + v_{it} - U_{it}) / (X_{it}\beta + v_{it}) = (X_{it}\beta - U_{it}) / (X_{it}\beta) = -U_{it} \text{ ....(2)}$$

Where U_{it} and $X_{it}\beta$ are defined by the specifications of the model in equation (1). In this study, Cobb-Douglas stochastic frontier production function, which is the most commonly used model, is considered to be the appropriate model for the analysis of the technical efficiency of the farmers. On the basis of panel data, equation (1) above can be expressed in the following form:

$$Y_{it} = A L_{it}^{b_L} K_{it}^{b_K} e^{V_{it}} e^{-U_{it}} \text{ ....}(3)$$

Where V_{it} follows $N(0, \sigma_v^2)$ and U_{it} follows a half or truncated normal distribution at zero. Taking natural log on both sides of equation (3), the following equation is obtained:

$$\ln Y_{it} = \ln A + \beta_L \ln L_{it} + \beta_K \ln K_{it} + (V_{it} - U_{it}) \text{.........................}(4)$$

Finally, the following equation was estimated by the computer programme FRONTIER 4.1 developed by Coelli (1994) that computes the parameters estimates by iteratively maximizing a nonlinear function of the unknown parameters in the model subject to the constraints.

$$\ln(\mathrm{Yit}) = \beta_0\mathrm{it} + \beta_{1it}\ln\mathrm{X}_1\mathrm{it} + \beta_{2it}\ln\mathrm{X}_2\mathrm{it} + \beta_{3it}\mathrm{X}_3\mathrm{i} + \beta_{4it}\ln\mathrm{X}_4\mathrm{it} + \beta_{5it}\mathrm{X}_5\mathrm{it} + \beta_{6it}\mathrm{X}_6\mathrm{it} + \beta_{7it}\mathrm{X}_7\mathrm{it} + \beta_{8it}\mathrm{X}_8\mathrm{it} + \mathrm{Vit} - \mathrm{Uit} \text{...................}(5)$$

Where, βi's are parameters to be estimated (coefficients) of inputs to be estimated by maximum likelihood estimation method (MLE). Here, the βi's refer to output elasticity. Ln is natural logarithm, Yit is denotes the production in (kg) at the t-th observation (t = 1,2,..., T) for the i-th farmer (i= 1,2,..., *N);* X_{1it} is maize farm size (ha), X_{2it} is human labor (man-days), X_{3it} is oxen labour used (oxen-days), X_{4it} is DAP fertilizer used (kg), X_{5it} is urea fertilizer used (kg), X_{6it} is improved seed used (kg), X_{7it} is compost used (quintal), X_{8it} represents year of observation; V are assumed to be iid $N(0, \delta_v^2)$ random errors, independently distributed of the $\mu_{it}s$, U_{it} represents technical inefficiency effects independent of Vi, and have half normal distribution with mean zero and constant variance while i shows households during the time t year. Battese and Coelli, (1995) noted that the *year* variable in the stochastic frontier accounts for Hicksian neutral technological change.

Following Battese and Coelli (1995) model, the mean of farm-specific technical inefficiency Ui, is also defined as:

$$u_{it} = \delta_0 + \delta_{1it}z_{1it} + \delta_{2it}z_{2it} + \delta_{3it}z_{3it} + \delta_{4it}z_{4it} + \delta_{5it}z_{5it} + \delta_{6it}z_{6it} + \delta_{7it}z_{7it} + \delta_{8it}z_{8it} + \delta_{9it}z_{9it} + \delta_{10it}z_{10it} + \delta_{11it}z_{11it} + \delta_{12it}z_{12it} + \delta_{13it}z_{13it} + \delta_{14it}z_{14it} + \delta_{15it}z_{15it} + \delta_{16it}z_{16it} + \delta_{17it}z_{17it} + w_{it} \text{.....}(6)$$

Where Z_1 is age of the farmers (years) during the year, Z_2 is gender of household head [1 male, 0 otherwise] , Z_3 is marital status of household [1 married, 0 otherwise], Z_4 represents that the household head can read and write [1 yes, 0 otherwise], Z_5 is educational level of household head (years of schooling), Z_6 farming experience of household head (years), Z_7 is family size, Z_8 is average annual non farm income (Birr), Z_9 is household head has radio [1 yes, 0 otherwise], Z_{10} shows that the household has land use certificate [1, yes; 0 otherwise], Z_{11} is total land size of the household (hectare), Z_{12} is distance of household residence from the technology distribution center (hours), Z_{13} is average annual development agents visit to the house hold (number), Z_{14} is plough frequency of maize land (number), Z_{15} represents Guto Gida district [1 Guto Gida, 0 otherwise], Z_{16} represents Gida Ayana district [1 Gida Ayana, 0 otherwise], Z_{17} represents year of observations, Wit, is defined by the truncation of the normal distribution with zero mean and variance, δ^2, and δs are parameters to be estimated. Here, the *year* variable in the inefficiency model (6) specifies that the inefficiency effects may change linearly with respect to time. This is because "the distributional assumptions on the inefficiency effects permit the effects of technical change and time-varying behavior of the inefficiency effects to be identified in addition to the intercept parameters, $\beta_0$it and δ_0, in the stochastic frontier and the inefficiency model" (Battese and Coelli, 1995).

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 level technical efficiency 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 later 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 (PSM).

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 the agricultural productivity and technical efficiency of the farmers. 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 training 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:

$$P(x) = pr(T = 1 \mid x) \quad \text{...(7)}$$

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 thesis.

However, in order to determine if matching is likely to effectively reduce selection bias, it is crucial 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.

This 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:

$$(Y_1, Y_0) \perp T_i \mid X_i \quad \text{...(8)}$$

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, 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 un-confoundedness 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).

$$Y_0 \perp T_i \mid X_i \quad \text{...(9)}$$

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.

This 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:

$$0 < \Pr(T = 1|x) < 1 \text{..........} (10)$$

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:

$$P(x) = \Pr(T = 1|x) < 1 \text{..........} (11)$$

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 score 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 thesis: 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).

$$B_{before}(x) = 100.\frac{\overline{X}_T - \overline{X}_C}{\sqrt{\frac{[V_{T(X)} - V_{C(X)}]}{2}}}, B_{after}(x) = 100.\frac{\overline{X}_{TM} - \overline{X}_{CM}}{\sqrt{\frac{[V_{T(X)} - V_{C(X)}]}{2}}} \ldots(12)$$

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:

$$ATT_{PSM} = E_{p(x)|T=1,}\left\{E\left[Y_1|\mathrm{T}=1,\ \mathrm{p(x)}\right] - E\left[Y_0|\mathrm{T}=0,\ \mathrm{p(x)}\right]\ \right\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots(13)$$

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)

Unlike the propensity score matching, DID assume 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). 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.

$$\ln Y_{ijt} = \alpha_0 + \alpha D_t + \beta FFS_{ijt} + \mu x_{ijt} \gamma z_{ijt} + \lambda_i + \eta_j + \varepsilon_{ijt} \ldots\ldots\ldots\ldots\ldots(14)$$

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:

$$\Delta \ln Y_{ijt} = \alpha + \beta FFS_{ijt} + \mu \Delta \chi_{ijt} \gamma \Delta z_{ijt} + \varepsilon_{ijt} \ldots\ldots\ldots\ldots\ldots (15)$$

The symbol (Δ) in equation 15 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.

Significance of the study

This study has enormous academic contributions. It has unique contributions in that it employs propensity score matching (PSM) and difference in difference (DID) impact assessment methods attempting to supplement the limitation of the first method by the later. To the best of our knowledge, this study is the first to combine *Psmatch2* and *Pscore* stata commands with four different matching algorithms attempting to ensure the robustness of the estimated program impacts.

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 so as to verify the similarities of the samples. 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 before and after the program was implemented by farmer groups on the basis of some selected performance indicators followed by section four presenting FFS impact assessments by farmer groups using PSM. Section five extends the impact assessment further by using DID method.

Table1

Household and farm characteristics during 2010 (by farmer groups)

	Mean		t-test	
Variables	FS 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
Oxen labour (Oxen day/Ha)	13.528	10.43	**3.680**	**0.000**
Total maize farm (Ha)	1.4463	1.1012	**3.620**	**0.000**
Percent of maize land to total		86.4000		0.691
89.600			0.398	

Note. Source: Own calculation from survey data of June to July, 2012.

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 had the highest scores in terms of educational levels, non-farm income, family sizes, estimates of asset values, total land size as well as percent of farm size covered by maize. Significant differences were also observed in the proportions of household owning mobile cell phone, radio ownership, participation in farmers training center, participation in farmers cooperatives, as well as number of contacts with the Kebele level development agent as those FFS graduate farmers had the highest scores than those non-FFS graduate farmers in all cases. In a sharp contrast with the FFS graduate farmers, non-FFS farmers are found at more distance from such important locations as centers for farm technology distributions and from their respective district offices.

Such significant difference between the farmers groups was not just the result of non-random selection of the farmers into the FFS training program. Rather, it was the result of the intended principles of selection criteria followed by the government. As the result, 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 the 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.

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 of June to July, 2012.

Cost and Returns of Maize Production by Farmer Groups

Table 2 presents cost and returns of maize production by farmers' groups. Comparison of costs and returns among the two farmers groups shows that FFS graduate farmers had significant differences from their counterpart, non-FFS graduate farmers, specifically in terms of total maize obtained, technical efficiency, and 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.

Given the fact that FFS graduate farmers own larger farm sizes than those non-FFS graduate farmers, profit margin diminishes as we look at their per hectare contributions although they were still significant at 10 percent. FFS graduate farmers had modest difference from non-FFS graduate farmers in terms of total cost per hectare and cash cost per hectare they incurred. 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.

In general, fertilizer application per hectare of the smallholder farmers in the study areas is low compared to both the African and world standards. However, we it was revealed that the basic factor underlying poor technology adoption in the study areas is neither lack of awareness as government claims nor lack of desire for success by the smallholder farmers. The major constraints identified as limiting technology adoption by the smallholder farmers are the escalating price of the technology themselves, lack of credit arrangement for such input purchase, inconsistent supply of the technologies, poor quality of the technologies supplied by the unions and their cooperative as well as fear of risks associated with adopting such technologies in the face of rapidly changing environmental factors owing to global warming. These findings suggest the need to create and sustain a number of institutions whose functions are the base for the desired agricultural transformation in the Ethiopian context.

Performance Indicators by Farmer Groups "Before and After"

Table 3 presents comparisons of various input and output performance indicators between the two farmer groups before and after the FFS program intervention. A statistical comparison in the table 3 reveals some seemingly 'illogical' and surprising results. 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 the result, we couldn't find any statistical difference in terms of technical efficiency change between the two farmers groups over time periods.

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

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 later uses 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. Thus, more technological adoption may not automatically result in productivity enhancement without proper agronomic practices such as timely field preparation, timely planting, and timely applications of agronomic chemicals.

Eventually, after two years of 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 later farmer group had significantly higher net crop income during the baseline year of 2010. With this understanding, more sophisticated assessment of FFS on technical efficiency of the farmers is presented in the next sections.

Assessment of Farmer Field School Impact Using PSM

In this section, we have employed PSM which doesn't require distributional assumptions to identify casual effects of the program (Kassie, Shiferaw and Murich, 2010). Although there are a number of matching methods to match the FFS program participant sampled households with the sampled non-FFS program households, in this study, we have used 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*[i] and *Pscore*[ii].

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 impact, which is again, to be confirmed by the impact assessment using DID in the subsequent section.

Estimation of the Propensity Scores

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). In estimating the propensity sores, we first tried by fitting all data collected on the covariates into logit model and gradually reduced the number of the covariates until we get the desired good match. Finally, we have maintained those influential covariates determining the program participation. The covariates included comprises 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.Err.	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 of 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

Note. Source: Own calculation from survey data of June to July 2010

It shows that some covariates are statistically 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, mobile phone, total asset values, social network participation, such as participation in farmers' cooperative, number of development agents contact with the household per year, possession of land use certificate, and 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 educated and hence the more chance of selection into the training program was. 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. These findings also indicate 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.

The main purpose of the propensity score estimation was to balance the observed distributions of covariates across two farmer groups. We need to ascertain that there is sufficient common support region for the two groups of farmers. The differences in the covariates in the matched two groups have been eliminated. These two issues 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.

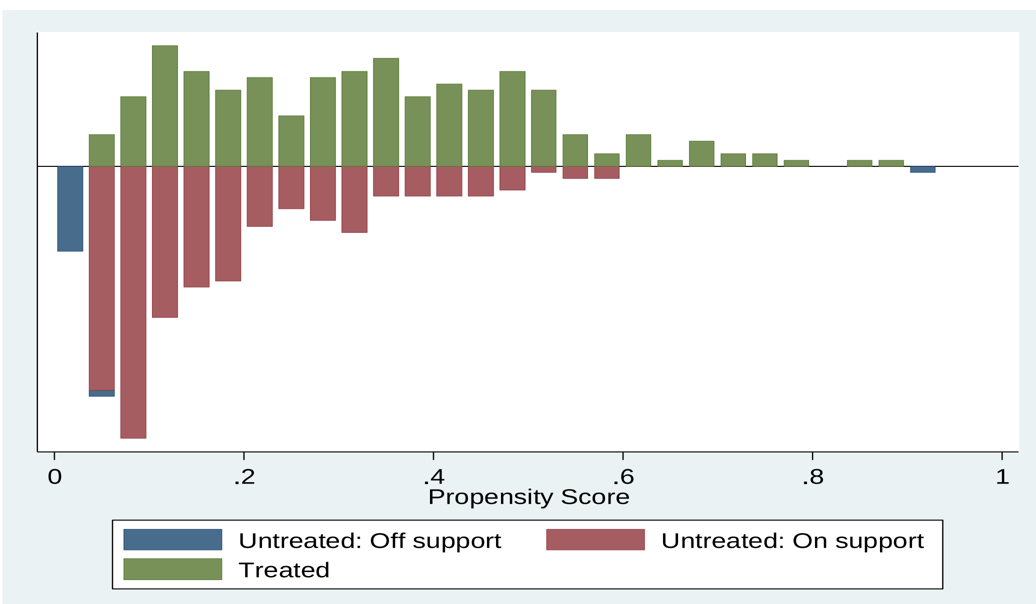

Figure 1. Propensity score distributions and common support for the propensity score estimation.
Source: own calculation from survey data

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 make test of covariate balancing. Accordingly, Table 5 presents results from covariate balancing test before and after matching. Mean standardized bias between the two groups after matching has been significantly reduced for all matching algorithms. This suggests 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[iii], which has substantially reduced total bias between 73.3 to 82.2 percent depending on which matching algorithm is used.

Table 5

Quality of Matching before and after matching

Algorithms	Before Matching			After Matching			
	Pseudo R^2	LR X2 (P-value)	Mean std Bias	Pseudo R^2	LR X2 (P-value)	Mean std Bias	Total bias reduction (%)
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

Note. 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

Source: own calculation from the survey data

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, 2004) shows that the pseudo R^2 has 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 summary, 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 very much successful in terms of balancing the distributions of covariates between the two groups of farmers.

Table 6

Estimates of stochastic frontier production function for maize farmers (panel data)

Input variables	**coefficient**	**St. .error**	**t-ratio**
Constant	6.8057	0.0972	70.0289
Maize land (hectare)	1.1688	0.0558	20.9301
labour used (man-days)	0.1008	0.0172	5.8682
Oxen labour (oxen days)	0.0765	0.0204	3.7537
DAP applied (kg)	0.1106	0.1357	0.8150
Urea applied (kg)	-0.0270	0.1406	-0.1924
Seed used (kg)	0.0511	0.0310	1.6468
compost used (qt)	0.0311	0.0114	2.7150
Year of observation	0.0190	0.0299	0.6352
Inefficiency variables:			
Constant	9.7654	1.2695	7.6921
Age of HH head	0.0557	0.0162	3.4400
Gender of HH head	-1.1409	0.5093	-2.2401
Marital status of HH head	-0.6680	0.2593	-2.5766
HH head can read and write	-0.0168	0.2527	-0.0666
Educational level of the HH head	0.0503	0.0404	1.2457
Farming experience of HH head in years	0.0002	0.0160	0.0138
HH family size (number)	-0.0777	0.0321	-2.4190
Average annual non farm income	0.0001	0.0000	3.7598
HH head has radio	-0.6236	0.1492	-4.1795
HH has land use certificate	-0.3815	0.2086	-1.8286
HH land [hectare]	-0.4320	0.0559	-7.7333
Distance from technology center[hrs]	0.9483	0.1028	9.2283
Average DA contacts	-0.0380	0.0091	-4.1556
Plough frequency	-2.7585	0.1878	-14.6874
Guto Gida District	-14.5711	0.7942	-18.3459
Gida Ayana District	-7.4981	0.5075	-14.7759
Time (year)	0.0865	0.0954	0.9071
Sigma-square ($\delta^2 = \delta u^2 + \delta v^2$)	4.4100	0.3866	11.4071
Gamma ($\gamma = \delta u^2 / \delta^2$)	0.9911	0.0016	611.9111
eta (η)	-0.0622	0.0251	-2.4812
ln (Likelihood) LR test	-1060.14		
Mean Technical Efficiency	0.59		

Note. Source: own calculation from the survey data

Estimation of Farmers' Technical Efficiency

Following Battese and Coelli (1995) model, the mean of farm-specific technical inefficiency Ui, was estimated using equation 5 above. Table 6 presents the estimates of stochastic frontier production function for maize farmers using pooled data of three years both for FFS graduate and non-FFS graduate farmers. Before proceeding to the analysis of impact of FFS on the technical on the technical efficiency of the farmers, it was necessary to assess the presence of inefficiency in the production data for the sampled households. Given the specifications of the stochastic frontier production function defined by equation (5), the null hypothesis that technical inefficiency is not present in the model is expressed by $H_o : \gamma = 0$, where γ the variance ratio is explaining the total variation in output from the frontier level of output attributed to technical efficiencies and defined by $\gamma = \frac{\sigma_u^2}{\sigma_v^2 + \sigma_u^2}$. The parameter γ must lie between 0 and 1; the closer the value of γ to zero, indicates that the inefficiency effects are insignificant and vice versa.

Accordingly, generalized likelihood-ratio tests of null hypotheses of the estimated parameters are presented in Table 7 below.

Table 7

Likelihood-ratio tests of hypotheses for parameters of the stochastic frontier production function

Null Hypothesis	Log likelihood	$\lambda*$	Critical value	Decision
Given Model	1060.14			
$\gamma = 0$	1493.86	867.43	6.63	Reject the null hypothesis
$\mu = \gamma = \eta = 0$	1337.10	553.92	11.34	Reject the null hypothesis
$\mu = \eta = 0$	1337.10	553.92	9.21	Reject the null hypothesis
$\mu = 0$	1336.96	553.63	6.63	Reject the null hypothesis
$\eta = 0$	1276.27	432.26	6.63	Reject the null hypothesis

Note. Source: own calculation from survey data

The first null hypothesis tested states that technical inefficiency is not present in the model, $H_o : \gamma = 0$ was strongly rejected. Similarly, the null hypotheses states that technical inefficiency effects are time invariant and that they have half normal distribution defined by H0: η = 0 and H0: μ = 0 were also strongly rejected. As the estimated parameter η was found to be significantly negative, which was -0.0622 at [t=2.5], it means that the technical efficiency of the sampled farmers decreases over time. It was also proved that the inefficiency effects in the stochastic frontier are clearly stochastic and are not unrelated to the household and farm specific variables and year of observation included in the model. The fact that the null hypothesis stating that parameter μ is zero was rejected implies that truncated-normal distributional assumption of one sided error term is more appropriate for the farmers in the study area than half-normal distributional assumption.

The signs of the coefficients of the stochastic frontier are as expected, with the exception of the negative estimate of the urea applied. The estimated elasticities of mean output with respect to land, human labour, and oxen labor are 1.1688, 0.1008, and 0.0765 respectively. This means that for a 10 percent increase in area cultivated to maize, maize output will increase by 11.68 percent. This shows the importance of farm size for maize production. This could be related to achievement of economies of scale. This clearly indicates the rejection of the strongly held view of the Ethiopian government who assumes the smallholder farmers as more efficient than the larger farm size operators. Similarly, a 10 percent increase in the amount of human labour increases maize output by 1.01 percent, again indicating the significance of human labor for routine maize farm management. A 10 percent increase in oxen labour increases maize output by about 0.8 percent. The estimated elasticity for compost is 0.0311 implying that a 10 percent increase in its application increase maize output by 0.31 percent and this result is statistically significant at 1 percent.

As the estimated coefficients in the inefficiency model are more relevant for this study. It is reasonable to discuss these results in a more detail. As expected, the *age* coefficient is positive, which indicates that the older farmers are more inefficient than the younger ones. This could be because the elders lack the required capacity to deal with routine agricultural work and or because they lack literacy. The negative estimate for gender implies that the males are more efficient than females. This is actually true in the reality of the study areas as females are usually preoccupied with in-house activities including child caring while the agricultural activities which demand more labour are customarily considered as the responsibility of males. Similarly the negative sign for literacy implies that farmers who can read and write tend to be more efficient. The coefficient of family size is negative implying the importance of labour for maize production. Those farmers who have more non-farm income tend to be more inefficient and this is statistically significant even at 1 percent. Other variables such as having a radio, land use certification, and size of land owned, have negative signs. This shows that the individuals who have radio acess may acquire updated information and hence tend to be more efficient. Also, having a land use certificate will increase their tenure security and hence make the farmers more efficienct. The negative sign for land size is consistent with the importance of larger farm land for achievement of scale of economies. Plough frequency has the expected large negative signs with statistical significance showing that if maize land is ploughed many times before planting, the more the efficient the farmer will be. Dummy variables for the districts show negative sign implying that sampled farmers in Guto Gida and Gida Ayana are more efficient than farmers in the Boneya Boshe district. The positive coefficient for year variable in the inefficiency model although statistically insignificant, suggests that the inefficiencies of the maize farmers tended to increase throughout the year. This is also confirmed by the decreasing mean technical efficiency of the farmers which was 0.60 during the year before the FFS training and reduced to 0.59 during the subsequent two years after the training. The estimate for the variance parameter, y, is 0.9911 which is close to one, indicating that the inefficiency effects are likely to be highly significant in the analysis of the value of output of the farmers. Furthermore, the estimates for parameters of the time varying inefficiencies model indicate that the technical inefficiency effects tend to increase over time since the parameter η is estimated to be negative (–0.0622) which is statistically significant at 5 percent.

Impact Estimation Using PSM

Our main interest in this section is to see if the FFS training program has brought any desirable change in the technical efficiency of the FFS graduate farmers as compared to non- FFS graduates. For this, the estimated technical efficiency for each farmer in the sample from the equation 5 was used as dependent variable in the models specified by equation 13 above so as to examine the technical efficiency difference between the two farmer groups. Accordingly, comparison of technical efficiency across farmers groups is presented by Table 8.

Table 8

Comparison of technical efficiency across farmer groups

Command	Algorithms	FFS Graduate (N)	Non FFS (N)	ATT	Std.Err	t
	Attnd	217	228	-0.0178	0.0336	-0.53000
	attr 0.01	202	228	-0.0011	0.0310	-0.04000
Psmatch2	attr 0.005	177	228	0.0028	0.0320	0.09000
	Attk	217	228	0.0094	0.0285	0.33000
	Attnd	217	94	0.027	0.038	0.72900
	attr 0.01	191	212	0.022	0.024	0.90000
Pscore	attr 0.005	174	199	0.025	0.025	0.98000
	Attk	217	212	0.023	0.03	0.77300

Note. Source: own calculation from survey data

The result shows that the estimated coefficients are very small and inconsistent among different matching algorithms. Since all are statistically insignificant, this implys that the FFS graduate farmers do not seem different from other farmers in terms of their technical efficiency. The result is also consistent with the implications of descriptive statistics explained above.

Impact Estimation using DID

In this section, household technical efficiency index estimated by equation 5 was used as dependent variable in the impact estimation function specified by equation 15 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 periods, 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.

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

Consistent with the technical efficiency estimates reported above, all variables included in the estimates of technical efficiency growth rate are found with expected signs.

Table 9

Estimated impact on FFS graduate technical efficiency using DID

Dependent variable: Technical efficiency						
N=446	F= 7.1700		R^2= 0.5400			F = 0.0000
Variables	Coef.	St. Err	t	P>|t|	95% Conf. Interval	
Constant	-0.0390	0.0093	-4.2000	0.0000	-0.0572	-0.0208
FFS Graduates	**-0.0257**	**0.0096**	**-2.6700**	**0.0080**	**-0.0445**	**-0.0068**
Plough frequency	**0.0435**	**0.0187**	**2.3300**	**0.0200**	**0.0069**	**0.0801**
Fertilizer used	0.0000	0.0001	-0.0700	0.9470	-0.0002	0.0002
Maize farm	0.0064	0.0097	0.6600	0.5080	-0.0126	0.0254
Family labor	**0.0008**	**0.0002**	**3.5000**	**0.0000**	**0.0004**	**0.0013**
Hired labour	-0.0001	0.0003	-0.4500	0.6540	-0.0008	0.0005
Herbicide	-0.0024	0.0051	-0.4800	0.6310	-0.0124	0.0075
Tractor use	0.0000	0.0000	0.1900	0.8480	0.0000	0.0000
Compost	0.0008	0.0009	0.9200	0.3570	-0.0009	0.0026
DA visits	-0.0006	0.0007	-0.8700	0.3870	-0.0021	0.0008
Guto Gida	**0.0463**	**0.0126**	**3.6600**	**0.0000**	**0.0214**	**0.0711**
Boneya Boshe	**0.0862**	**0.0117**	**7.4000**	**0.0000**	**0.0633**	**0.1091**

*Note.*Source: own calculation from the survey data

Consistent with the descriptive analysis discussed above, FFS graduate farmers are identified with statistically significant lower technical efficiency growth rate. The model estimate shows that participation in the FFS training program has reduced their technical efficiency growth rate by about 0.3 percent and this difference is statistically significant at 1 percent. The Farmers have reported that shortage of time to deal with their routine agricultural practices become the major hindrance for their production and productivity enhancement. They have stated that they are overloaded by the frequency of meetings and short term trainings of various types, rural road construction, and natural resource conservation practices which usually coincide with their farm field preparation seasons tend to make farmers' less efficient than before the FFS training program. The farmers have actually reported that their efficiency declines over time, not because of lack of the required technical skills, but mainly because of lack of time and financial resources to undertake the required agricultural practices at right time. Furthermore, the model farmers have bitterly expressed their concern over the natural resources conservation and rural road construction practices that they are required to do for a minimum of 30 days each year. Such practices not only consume their agricultural time but also severely curtail their efficiency as they are more frequently injured while doing such heavy tasks as digging holes, rolling of rocks and carrying of heavy woods.

Other variables such as plough frequency, application of family labor and dummy variable representing the districts have all expected and statistically significant coefficients. The sign of the estimated coefficient for family labor has statistically significant positive value implying the importance of such labor for efficiency gain while the coefficient for the hired labour is negative. Such finding is also consistent with microeconomic theory which states that in the absence of strict supervision and monitoring, hired laborers fail to increase efficiency owing to their morale hazard problem. Dummy variables for Guto Gida and Boneya Boshe districts have large, positive and statistically significant coefficients implying significant differences in the technical efficiency growth rate among farmers living in different districts. The significantly positive coefficient for Guto Gida and Boneya Boshe districts imply that, on an average, farmers in both districts have higher technical efficiency growth rate than farmers in the Gida Ayana district.

Conclusions and Policy Recommendations

Conclusions

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. Specifically, the establishment of the farmers' networks in the form of a top-down approach are factors negatively affecting both the quality of training and its diffusion effects and hence reducing the program impacts. Our result shows that the farmers' networks are not organized in the way the farmers can take steps for dealing with challenges and obstacles facing them through collective action.

As the FFS graduate farmers allocate most of their time for numerous mandatory meetings, trainings, community mobilization, and their heavy involvement in political canvassing, they tended to use more paid labour than maximizing their own labour for the routine agricultural practices. In addition, most FFS graduate farmers substituted applications of herbicide chemicals in lieu of manual weeding and their cash cost of maize production increased over time, while their technical efficiency declines owing to lack of time to monitor those paid laborers.

The major constraints identified are : limited technology adoption by the smallholder farmers, escalating price of the technology themselves, lack of credit arrangement for such input purchase, inconsistent supply of the technologies, poor quality of the technologies supplied by the unions and their cooperative, and fear of risks associated with adopting such technologies in the face of rapidly changing environmental factors owing to global warming.

In the end, we have employed a number of sophisticated econometric models appropriate for impact evaluation design. There are words of caution with regard to our conclusions. Firstly, given the fact that FFS training program was the national agenda operating in all regions of the country all at the same time, selection of representative districts and households were a real challenge given the very limited research funding and time available. As a result, the data for this study were collected only from three purposively selected maize producing districts and from each district only one Kebele from where households were randomly selected. Thus, this approach has enabled us to positively contribute to impact assessment literature and agricultural policy makers, but it might have come at some expense of representativeness. Secondly, the chosen locations are representative for maize producers in the region, we are not sure how well those locations represent the average conditions under which the FFS training program was implemented in the country and its impacts on other agricultural crops. Thirdly, it is true that the lessons learned from FFS program would be forgotten if not used to practice shortly, by assessing program impact just two years after the program intervention, we may be capturing the only medium term impacts that may or may not last over time. In essence, the estimated impact shows impacts after two years of program implementation, and does not show any possible dynamisms of the impact in the long run. Finally, this study has only considered the impact on the technical efficiency of maize producer farmers, no claim is made with regard to program impacts on other aspects such as general socio-economic development, environmental conservations, health, and political sustainability that the program might have impacted.

Policy Recommendations

The first policy recommendation is to contextualize the FFS training curriculum and its timing of implementation to the specific situations of rural producers. It is important to avoid blanket technology recommendation using FFS approach as the use and success rate of a technology is usually location specific. The FFS training program should target at farmers' identified problems and the farmers should decide on the special topics on which they need discovery learning rather than the current top-down approach of FFS curriculum design. In essence, FFS training program needs to be "people centered" in which case the farmers will freely and autonomously participate in problem identification and its prioritization, curriculum design, setting criteria for participant selection, and forming farmers' networks with their own free choice. On the other hand, the role of government has to be limited to assisting the farmers in the form of assigning technically competent FFS facilitators who are conversant with the specific location where the program is implemented; the supply of adequate material and logistic supports needed for the training; and making uninterrupted follow up with the view to create incentives for farmers to continue sharing experiences of technical changes even after the program is closed.

Our second recommendation to the policy makers is to clearly separate activities required for agricultural transformation from activities required for political canvassing. Although the government has claimed “to bring agricultural transformation” as the driving objective for scaling up of FFS training program as the national agenda, in practice, however, both the model farmers’ selection criteria into the program, as well as, their role in society after graduation from the program are found to be popularizing the political doctrines of the ruling party rather than catalyzing agricultural transformation of the country. It is really a temptation to think that the one who is a model in agricultural activities can be the best in politics too. Thus, it is really important for the government not to use both FFS training program and model farmers for political purposes for which these two are not necessarily the best instruments.

For the third policy recommendation, we suggest that the government create and sustain a number of more responsive rural institutions and the related institutional frameworks for the desired agricultural transformation, which include:

i. Government should allow participation of private agricultural technology input suppliers, whose success will depend on providing inputs to the producers when and where needed and hence could be more responsive to shift in weather, cropping patterns and new technology supplies on competitive base than unions and their cooperatives. Our study showed that government and its parastatals such as the unions and cooperatives are almost never in the right place, at the right time, with the right product in the allocations of industrial products and seeds to the rural producers. Given their susceptibility to predatory behaviors such as corruption, rent-seeking, abuse of public resources, and a basic lack of accountability. These parastatals have never been successful in addressing the smallholder farmers’ real interests. Thus, government ought to reduce excessive reliance on the unions and their cooperatives for the distributions of agricultural technologies to the smallholder farmers.
ii. There should be credit arrangements for the poor farmers who are unable to finance the required technologies.
iii. The government has to promote and design incentive structures for private firms to invest in agricultural crop insurance scheme to build up the farmers’ trust in agricultural technologies in the face of volatile output prices and rapidly changing environmental impacts.
iv. The government needs to consider implementation of a forward contract market. A forward contract market refers to a futures market in which both the buyers and the sellers make an agreement stipulating the amount to be exchanged and the exchange price and date of the exchange before crop production while the actual physical exchange of outputs are made at a later date after crop harvest. In this case, the government can use its parastatals such as Ethiopian grain trade enterprise, Oromia Agricultural product market enterprise, Ethiopian Commodity Exchange (ECX) enterprise as well as unions and cooperatives to inter into forward agreement with the smallholder farmers before their production decisions so that the farmers can rationally make cost-benefit analysis of their productions. This system will enable the smallholder farmers to make informed production and marketing decisions simultaneously as they are supposed to know not only the input prices required

for the production but also prices that their resulting output will bring in return as well. Eventually, this system is expected to alleviate the fear of risk of product market failure and hence encourage the smallholder farmers to use full technology packages so as to maximize their income.

v. Finally, there should be farmers' networks and organizations which are formed by the smallholder farmers own freewill and which can promote their "human agency" rather than the one being used as instruments for government political canvassing.

End Notes

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).

ii. 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.

iii. 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.

References

Bamlaku .A. E-A Nuppenau & 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. & T.J. Coelli. (1995). A Model for technical inefficiency in a stochastic frontier production function for panel data. *Empirical Economics*, 20, 325-332.

Bravo-Ureta, Boris E., & Robert E. Evenson. (1994). Efficiency in agricultural production: The case of peasant farmers in Eastern Paraguay. *Agricultural Economics (*10)1, 27–37.

Bryman, A. (1988). *Quantity and quality in social research.* London: Unwin Hyman

Essama-Nssah, B. (2006). Propensity score matching and policy impact analysis: A demonstration in views. *Poverty Reduction Group* (PRMPR). The World Bank Washington, D.C.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of Royal Statistical Society*, (A)120, 253-281.

FDRE, MoA (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

FDRE, MoARD. (2010). *Ethiopia's agriculture sector policy and investment framework: Ten year road map (2010-2020).* Addis Ababa, Ethiopia,

FDRE, MoFED. (2014). *Growth and transformation plan*. Annual progress report for F.Y. 2012/13. Addis Ababa, Ethiopia, Feder, G., R. Murgai, & 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, & Jaime Quizon. (2004a). Sending farmers back to school: The impact of farmer field schools in Indonesia. *Review of Agricultural Economics*, 26(1), 45–62.

Heckman, James J., Ichimura, Hidehiko, & Todd Petra. (1998). Matching as an econometric evaluation estimator. *Review of Economics Studies* (65), 261-294.

Heinrich. C, A. Maffilo & 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, 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. 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 Thesis. University of Hannover. Pesticide Policy Project Publication Series Special Issue No. 15,

Ravallion, M. (2008). *Evaluating anti-poverty programs*. In handbook of development economics, vol. 4, ed. Paul Schultz and John Strauss, 3787 846. Amsterdam: North-Holland.

World Bank, (2010). Handbook on impact evaluation: Quantitative methods and practice