Assessment of the Impact of Adoption of Improved Maize Variety on Farm Productivity Using Propensity Score Matching (PSM): the Case of Southern Ethiopia

*1Dibaba Debela Geleta & 2N Subba Rao
1Research Scholar, Department of Economics, Andhra University, Andhra Pradesh, Vishakhapatnam (India)
2Associate Professor, Department of Economics, Andhra University, Andhra Pradesh, Vishakhapatnam (India)

Abstract: This study has been carried out to investigate whether or not technology, as applied to improved seed, boosts agricultural production and productivity. If so under what instrument of impact assessment can one most prudently conduct an impact assessment of this sort while duly paying attention to the problem of endogeneity and self-selection? The fact that a certain change occurred on the question at hand should not necessarily imply that it is just because of the factor under study. In light of this the study opted for Propensity Score Matching as one of the most credible tool of impact assessment and justified its plausibility in the context of our study. The result of the Propensity Score Matching under various matching algorithms revealed that there is a quantitatively large and statistically significant impact of improved maize variety on productivity. And hence, question that really matters, the question of proper attribution of the amount of an impact together with its statistical test of the impact causing factor alone, is well addressed.

***Index Terms* – Impact, Propensity Score Matching, Productivity, Improved Seed, Agriculture, Ethiopia.**

1. INTRODUCTION

Agriculture is the major contributor of the Ethiopian economy since time immemorial and heather to. Both the backward and forward contribution of the agricultural sector made the sector to have been the hub of the Ethiopian economy. It is in light of this that many have dubbed Agriculture in Ethiopia, just like other sub Saharan countries, as the foundation of the country's economy. Be that as it may, however, agriculture has remained underdeveloped, unproductive, inefficient and unable to answer the food security question of the rural households primarily on account of the fact that these resources are not used to their full socio economic potential. The under development, unproductivity and in efficiency of the sector can be attributed to the commonest problems like drought, poor economic base, weak infrastructure, and low level of technology. A vast majority of rural small holders depend on traditional and subsistence farming, the characteristic features of which are, among others, low productivity. [EEA/EEPRI, 2015; EEA, 2007]

The use of improved technologies helps to increase productivity and shift the Production function upwards thereby playing a great role in the nation's age old attempt to attain food security and combat poverty as it has helped several countries all over the world. This is especially more so in a situation where both labor and land productivities are at a very low level. This very intuitive fact of classical microeconomics theory of production, should, however, be empirically justified with in the rubric of our study topic. It is with this in mind that the study endeavored to assess the impact of agricultural technology on farm productivity while paying special emphasis to maize farmers of Boricha Woreda.

2. IMPACT ASSESSMENT TOOL: PROPENSITY SCORE MATCHING METHOD

A frequent question of noble worth in impact assessment is whether or not the impact is only attributable to the treatment variable. The fact that a certain change occurred on the question at hand should not necessarily imply that it is just because of the factor under study. There could be other unobserved factors affecting the change and may even be the cause of the change.

One of the critical problems in non-experimental methods is the presence of selection bias which could arise mainly from non-randomness of the treatment variable, or non-random location of the project and the nonrandom selection of participant households that makes impact assessment problematic (Heckman et al, 1998).

There are several potential source of bias. The first one is that participant households may significantly differ from nonparticipants in demographic, institutional, socioeconomic characteristics due to observables that may have a direct effect on outcome of interest. Secondly, the difference arises due to unobservable household characteristic. Thirdly, externalities like the spillover effect exerted by the treatment variable on the non-treated may exaggerate or undermine the net impact of the treatment variable. Fourth the missing data problem coupled with endogeneity and non-randomness (Bryson et al., 2002; Ravallion, 2005)

As a due solution, something fairly similar to randomized experimental approach is done by identifying non-participating comparable groups identical in every way to the group that receives the intervention or the treatment variable, with the exception that control groups do not receive the intervention or the treatment factor.

Thus when one carries out a statistical analysis of impact assessment on cross sectional data by so using propensity score matching (PSM), a due endeavor should be made to estimate the impact of a treatment variable (policy, program, intervention, training, technology etc.) by taking into consideration the detrimental factors affecting the receipt of treatment. This tool, PSM, reduces the possible bias that may arise from estimation of the impact of a treatment by simply comparing the two groups which are commonly called treated and non-treated (Paul Rosenbaum and Donald Rubin, 1983).

Ideally propensity score matching (PSM) implements a sort of artificial randomization experiment where by a group of sample that received the treatment variable are made to be comparable with respect to all observed covariates to another group of sample who did not receive the treatment.

The intuition behind propensity score matching, PSM, method is an attempt to create the observational equivalent of an experiment in which everyone has the same or comparable probability of adoption. The difference, in our context of impact of adoption studies, is that in PSM it is the conditional probability (P(X)) that is intended to be uniform between adopters and non-adopters.

3. MATHEMATICAL SPECIFICATION OF PSM

When we come to the mathematical specification of the PSM method, estimating the impact of farm households' adoption of improved maize decision on outcome variable of Productivity (Y) can be specified as:

$$\tau_i = Y_i(D_i = 1) - Y_i(D_i = 0) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where $\boldsymbol{\tau}$ is treatment effect or impact that is attributable to adoption decision, Yi is the outcome variable (output productivity or yield) on farm household, D_i is whether household i^{th} has got the treatment or not (i.e., whether a household adopted improved maize variety or not). It should be noted, however, that adopters and non-adopters,$Y_i(D_i = 1)$ and $Y_i(D_i = 0)$ cannot be observed for the same farm household at the same time. This is because, now that we are dealing with cross sectional data, farm households cannot be adopters and non-adopters at the same time. On condition that the rural farm household is either an adopter or non-adopter. Either, the adopter $Y_i(D_i = 1)$ or the non-adopter $Y_i(D_i = 0)$ is unobserved outcome (called counterfactual outcome). This makes the possibility of impact assessment a difficult research endeavor as estimating individual treatment effect is an impossibility.

As a solution to this predicament uncovering the average treatment effect is worthwhile. There are two types of average treatment effect. One is the Average treatment effect (ATE) which is can be, mathematically, computed as the difference between the expected outcomes of adopters and non-adopters, the second one is, which is the most plausible, Average Treatment Effect on the Treated (ATT), that tries to single out the effect of adoption only on the adopters.

The mathematical formulation of the average treatment effect (ATE) can be depicted as

$$\Delta Y_{ATE} = E(\tau_i) = E(\Delta Y) = E(Y_1) - E(Y_0) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2)$$

The average treatment effect (ATE) tries to measure the expectation of the impact of adoption across all rural household farmers. Using the above equation as a tool of estimating impact assessment is problematic. The basic source of the problem is our data is not experimental data. The non-experimental cross sectional data that we have reveals only adopters or non-adopter at a given point in time. We, thus, need to resort to Average Treatment effect on the Treated (ATT) which is mathematically depicted as:

$$ATT = \tau_{ATT} = E(\tau/D = 1) = E(Y_1/D = 1) - E(Y_0/D = 1) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (3)$$

Equation (3) above answers the question as to how adoption of improved maize variety enhanced the Productivity of adopters compared with non-adopters. Obviously outcome data on adopters $E(Y_1/D = 1) = 1$ is available but outcome data on non-adopters $E(Y_0/D = 1)$ is not available for the same farm household.

This necessitates the choice of a reasonably acceptable substitute. Due to this problem, one has to choose a proper substitute, hence the question of matching pops in so much so that predicting ATT is possible.

In order to do that we employ the average outcome of the outcome variable (maize yield) of non-adopters as a proxy for the mean of the counterfactual group of adopters having, first, accounting for the problem of selection bias.

To do that equation (3) can be re formulated as

$$ATT = \{[\tau_{ATT} + E(Y_0/D = 1) - E(Y_0/D = 0)] = [E(Y_1/D = 1) - E(Y_0/D = 0)]\} \ldots \ldots \ldots (4)$$

In the above equation both terms in the left hand side and right hand side are observables. Furthermore in the equation (4) above ATT identifiable and can be estimated provided that there is no self-selection bias i.e. $E(Y_0/D = 1) - E(Y_0/D = 0) = 0$. For this to hold true, two prominent assumptions should be explicitly stated and duly satisfied in order to insure that self-section problems are accounted for.

The first assumption is the conditional Independence Assumption (CIA).Conditional Independence Assumption can be expressed as:

$$Y_0 \perp D/X$$

Where $\perp$ refers to independence and X is a set of observable characteristics, and Y_0 is non-adopters. Conditional independence assumption refers to the equation that for a given set of explanatory variables or covariates (X_i) are assumed to be not affected by the treatment variable (i.e. Adoption of improved maize variety). This means the outcome variable (output) should be independent of treatment variable i.e. outputs are not affected by treatment. The implication of this assumption is adoption decision is solely based on observable covariates (X_i) and determinants of adoption and farm output are observable simultaneously. (Caliendo and Kopeinig, 2008).

Now that the observable differences are adjusted the average outcome for both groups of adopters and non-adopters is the same or comparable, i.e. $E(Y_0/D = 1, X) = E(Y_0/D = 0, X)$.

Rosenbaum and Robin(1983) have made a clearer definition of the propensity score wherein they said, contextualizing it to our case study, it is the probability of participation in an adoption decision for a given farm household under certain characteristics that determines adoption decision. Propensity score has its derivation from limited dependent variable models (LEMDEP Models), where in these discrete choice models are used to create comparable or match able groups. It is in this way that PSM, given covariates, solves the problem of selection bias. The distribution of observables, at the end of the day, will be similar for both groups of adopters and non-adopters now that the propensity score is a balancing score

The second assumption relates to the common support or overlap condition. This refers to the law of probability being satisfied for both groups. The common support region rules out the perfect predictability of D for a given values of$X$. Mathematically, it is expressed as

$$0 < P(D = 1/X) < 1$$

This condition guarantees that farmers with the same covariates of X values have a positive probability of being both an adopter and a non-adopter. This makes Average Treatment Effect (ATE) is only identified with in the rubric of the common support. Farm households who are not with in the common support region are not included in the estimation of the ATE.

On condition that the two assumptions are satisfied, the propensity Score Matching (PSM) estimator of the Average Treatment effect on the Treated (ATT) can be expressed as

$$\tau_{ATT}^{PSM} = E_{P(X)/D=1}\{E[Y(1)/D = 1, P(X)] - E[Y(0)/D = 0, P(X)]\} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (5)$$

Thus according to the above equation (5), one can safely estimate the impact of adoption of improved maize variety on farm Productivity of maize by simply taking the difference of the average maize output of the two groups.

4. DATA AND SAMPLING TECHNIQUE

The main source of the data for this study was a primary data. A cross sectional survey was conducted on a sample of 223 farmers from Boricha Woreda. Boricha is one of the woredas in Sidama Zone of the Southern Nations, Nationalities, and Peoples' Region of Ethiopia. The study applied multi stage sampling technique where by credible sampling technique is employed in each and every stage.

5. RESULT AND DISCUSSION

The table below is the result of the propensity score matching analysis by using different matching algorism and band width. It has very vividly unveiled the disparity in terms of farm productivity of maize output while also justifying it with statistical significance.

Table 1: Propensity Score Matching Result

Source: own survey and calculation 2018.

***, **, *** significant at 1%, 5% and 10 % respectively**

The table above revealed the impact of improve maize seed on farm yield using different matching algorithm and band width per matching algorithm. The estimation of average treatment effect analysis as revealed in the table above was based on the three

Outcome (dependent) variable : Maize productivity Treatment (independent) variable: adoption of improved maize variety							
Algorithm	Common Support				ATT	SE	t- value
	off support		on support				
	Adoptors	Non Adoptors	Adoptors	Non Adoptors			
Nearest neighbor (1)	0	0	129	74	425	1.61	2.64*
Nearest neighbor (5)	0	0	129	74	269	0.90	2.98*
Radius matching (0.057)	106	0	23	74	218	1.03	2.12**
Radius matching (0.087)	82	0	47	74	298	1.06	2.83*
Kernel(0.03)	118	64	11	10	278	1.88	1.48
Kernel(0.057)	94	7	35	67	295	1.07	2.75*

alternative matching methods namely: nearest neighbor, radius and kernel matching methods. The outcome variable is farm output productivity which is a yield response of rural households measured in kilograms. The Average Treatment effect on the Treated (ATT) is estimated using equation (5) as above. Accordingly the results unveiled that adoption of improved maize varieties positively and significantly affect output in favor of adopters. This is true for all types of matching algorithm and band width except for kernel matching of 0.03 band width. ATT ranges from 425 kg in the case of nearest neighbor matching of band width (1) to 218 of Radius matching of band width (0.057) and both of them are significant at one percent level of significance. In all of the matching algorithm, adopters are better off than non-adopters. In the case of nearest neighbor matching of (1) band width adopters are better off than their counter parts by 425 kg and the difference is significant at one percent level of significance. This reduces to 269 kg for the same matching algorithm of (5) band width and the difference is significant at one percent level of significance. When we come to Radius matching algorithm of (0.057) band width adopters are better off than non-adopters by 218 kg and the difference is significant at five percent level of significance. The difference goes a bit higher to the level of 298 kg in favor of adopters for Radius matching of (0.087) band width and this difference is significant at one percent level of significance. The above table revealed that the different matching algorithms with different band width consistently confirmed that adoption of improved maize seed has a significant impact on the outcome variable. This implies that, according to equation (5) above and under the condition that there is no selection bias due to unobservable farm household head characteristics, the outcome variable of productivity for adopters is significantly higher than the non-adopters. This result implies that adoption of improved maize variety has qualitatively large and statistically significant impact on farm output and productivity and hence the need for agricultural technology in boosting the agricultural sector is justified.

6. CONCLUSION

The implication that can be made is that application of yield enhancing technologies paves the way for addressing, once productivity is enhanced, the diverse needs of the rural household and thus concerned authorities are advised to pay due attention to the matter.

7. ACKNOWLEDGEMENT

I am very much cognizant and appreciative of the immense support, advice, encouragement and guidance I received from my research Guide Dr. N, Subba Rao. I would like also to thank the ministry of education of Ethiopia for funding my research expense. Last, but not least, I would like to express my deepest gratitude to my darling wife, Welela Zerihun, for single handedly shouldering the responsibility of catering for the needs of our family while graciously letting me focus on my research undertaking.

8. REFERENCES

.

[1] Ethiopian Economic Association/Ethiopian Economic Policy Research Institute :(2007) evaluation of the Ethiopian Agricultural Extension with particular emphasis on the Participatory Demonstration and Training Extension System (PADETES) Addis Ababa, Ethiopia.

[2] Ethiopian Economic Association/Ethiopian Economic Policy Research Institute (2015): Report on the Ethiopian Economy Volume I. Addis Ababa, Ethiopia.

[3] J. Heckman, Hidehiko Ichimura and Petra Todd (1998): Matching As An Econometric Evaluation Estimator; Review of Economic Studies (1998) 65, 261-294

[4] Bryson, A., R. Dorsett and S. Purdon (2002): the use of propensity score matching in the evaluation of labor market policies. Working Paper No. 4. Department for Work and Pensions, London.

[5] Ravallion, M. (2005): Evaluating anti-poverty programs: Policy research working paper 3625, World Bank, Washington D.C.

[6] Rosenbaum, P.R. and Rubin, D.B. (1983): The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1): 41-55.

[7] Marco Caliendo and Sabine Kopeinig (2008): Some Practical Guidance for The Implementation of Propensity Score Matching .Journal Of Economic Surveys (2008) Vol. 22(1)