Journal of Economics and Sustainable Development www.iiste.org
ISSN 2222-1700 (Paper) ISSN 2222-2855 (Online)
Vol.9, No.22, 2018

Households' Welfare Effects of Snow Peas Production by Small Holder Producers in Mt. Kenya Region, Kenya

Naomi Njeri Mukami Jackson Kipngetich Langat
Department of Agricultural Economics and Agribusiness Management, Egerton University

Authors:
Corresponding Author:
Naomi Njeri Mukami, Masters of Science student at Egerton University, present/ affiliation address: P.O. Box 536-20115 Egerton, Kenya;
Co-Author:
Dr. Langat, Jackson Kipngetich[3], Monitoring Learning and Evaluation Officer at Tegemeo Institute of Agricultural Policy and Development, present address: P.O. Box 20498-00200 Nairobi, Kenya

Abstract
This study aimed at determining the households' welfare effects of snow peas production by evaluating the difference in income, assets and expenditure of adopters and non-adopters of snow peas farming in Mt. Kenya region, Kenya. Propensity score matching technique was used to evaluate the impact. Findings reveal that the impact of snow peas farming was significantly different between participants and non-participants in terms of their income, assets and expenditure. The estimation of treated effect on treated showed that participating farmers had relatively higher monthly income, higher total value of assets and higher expenditure compared to non-participating farmers. To enhance participation, farmers are encouraged to form and actively participate in farmer groups through collective productiction and marketing of their produce. Policies that provide extension services to farmers and affordable credit are also important in enhancing farmer's participation in snow peas production.
Keywords: snow peas, welfare effects, propensity score matching

1.0 INTRODUCTION

Kenya produces a wide variety of horticultural commodities, including fruits, vegetables, and cut flowers. Most production is rain-fed, but irrigated vegetable, fruits and flower cultivation can be found especially if production is for export purposes (Authority, E. P. Z. 2005). The horticultural sector is therefore important as a producer of food, source of income, employment, and foreign exchange and this has been recognized by the Kenyan Government. According to the vision 2030, crop cultivation is one of the main pillars of unlocking the potential of Kenya through increased productivity of crops (Nguguna *et al.,* 2009). The sub-sector has grown since 2002 to become a major pillar of economic growth. Currently, the horticulture industry is the fastest growing agricultural sub-sector and is ranked third in terms of foreign exchange earnings from exports after tourism and tea (KNBS, 2013).

Central Kenya offers a favourable environment for the production of horticultural crops like peas. It is mostly carried out by small scale farmers who are contracted by exporters (Chemining'wa *et al., 2004*). Most of them lack basic knowledge of producing them as well as capital for production and export opportunities (Davis, 2006). The industry is therefore characterized by brokers and middlemen who place farmers at a disadvantaged position in terms of prices and other benefits that should be associated in the entire value chain (Odero *et al., 2013).* This could possibly be explained by the nature of peas. Owing to the perishable nature of the peas, brokers take advantage of the farmers by buying the products at low price on realizing they have already harvested. This could therefore be one of the major reasons why the adoption of snow peas has been slow in Kenya. Beside this dilemma, the market for snow peas has also been unstable due the consistent fluctuation in prices. This has discouraged farmers from investing in the venture (Rugenyi, 2011).

Previous studies have shown that production of snow peas in Mt. Kenya region does not only benefit the entire economy in terms of it being a source of foreign exchange earnings, but the crop also creates employment especially for the rural population as well as generating income for farmers. In spite of the benefits, snow peas farmers in Mt. Kenya regionhave been withdrawing from its production perhaps due to low profitability. The study therefore aimed at determining whether there is a significant welfare difference between adopters and non-adopters in terms of assets, income and expenditure.

2.0 METHODOLOGY

2.1 Study area

This study was conducted around the Mt. Kenya region of Kenya. The altitude which is 3,500m above sea level and good climate are quite favourable for agricultural activities. The region lies between latitude $0.50°0'$ to the

North and $0°50'$ to South and between $36.20°\,0'$ East and $36°50'$ West (Regional Center for Mapping of Resource for Development, 2016).

The main economic activities in the area were farming where main crops included: maize, wheat, beans, Irish potatoes, cabbages, carrots, snow peas and garden peas. A large proportion of the farming in the region is dedicated to food crops. The crops are not exclusively meant for subsistence as they also contribute to household income. The area has a high population density, the average size of land per farmer is small (2 hectares) (Njarui *et al*, 2016), and hence small scale farming is the most practiced method of farming. Nevertheless, a few large scale farmers also exist. The cool temperatures, high rainfall and altitude make the area a conducive place for the production of snow peas hence the selection of the area for purposes of this research. The well drained and highly fertile soils with high water retention capacity facilitate production of snow peas and also minimize the use of inputs. Snow peas are cultivated mainly for commercial purpose. Land size under snow peas farming ranges from 0.25 to 1 acre. Most farmers grow snow peas beside other crops. Mostly, farmers depend on brokers to market their produce even though there are informal arrangements with some companies through various groups.

2.2 Sampling procedure

The target population of the study was small-scale farmers. The study applied multiple-stage sampling procedure.

The first step included purposively selecting major snow peas growing areas in Mt. Kenya region followed by stratified sampling of farmers with and without snow peas venture and finally proportionate random sampling was conducted to get the desired sample size. The required sample size was then determined by proportionate to size sampling methodology (Cochran formula in Mutai*,* 2014).

$$n=\left(\frac{z}{m}\right)^2 p(1-p) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 1$$

$$n=\left(\frac{1.96}{0.06}\right)^2 0.5(1-0.5)=267$$

Where: n = Sample size; z = confidence level; p = proportion of the population containing the major interest, and m= allowable error. Since the proportion of the population is not known with certainty, it is normally assumed that p= 0.5, and z = 1.96 and m= 0.06 (error the researcher is willing to accept). This results to a sample of 267 respondents.

To obtain impact estimates that are generalizable to the population of interest, it is necessary for the pool of comparison units to have a sufficient number of observations with characteristics corresponding to those of the treated units (Heinrich *et al.,* 2010). Therefore, higher sample size for untreated (60%) was used.

S

The objective aimed at comparing the level of consumption, expenditure and total annual income of households growing snow peas with households sharing the same social-economic characteristics but not involved in snow peas venture. Following Heckman et *al.* (1970);Smith and Todd (2001), estimating the effects of household participation in snow peas production would involve various steps. Let Y_1 be the mean of the outcome conditional that a farmer participates in snow peas production and let Y_0 be the mean outcome of the control group. The impact of snow peas production on income, expenditure or assets is the change in the mean outcome caused by farmers participating in snow peas production. It's given by:

$T_I = Y_i\,(D_i = 1) - Y_i(D_i = 0)$ ……………………………………………………………2

Where: T_i is the notation for the effect of the crop for a given households, Y_I is the outcome on household, D_i is whether household i is a participant or not. Since the two outcomes cannot be observed for the same household simultaneously, the problem of missing data arises (Gebrehiwot and van der Veen, 2015;Tolemariam, 2010). This means that the method gives biased estimates hence the need to introduce the sample average for the impacts of the treated group rather than the individual. Two treatment effects are the most common in empirical studies. The first one is the Average Treatment Effect (ATE), which is the difference of the expected outcome after participation and non-participation given by;

$\Delta Y_{ATE} = E(Y_1) - E(Y_0)$. ……………………………………………………………………3

The measure shows the effect if households in the population were randomly assigned to treatment. Heckman *et al*. (1997) noted that the measure would not be appropriate for policy makers since it includes the effects for which the intervention was not intended leading us to the second measure of treatment effects. The Average impact of the treatment on the treated (ATT) concentrates solely on participants pointing out the realized impacts of participation. It aims at determining how much the households participating benefited in the program compared to what they would have achieved without the intervention. It is given by:

$T_{ATT} = E(T|D = 1) = E(Y_1|D = 1) - E(Y_0)|D = 1$ ………………………………………4

Where: D, is a factor that indicate whether a household i received treatment or not. That $D_i = 1$ if the farmer was involved in snow peas production and 0 otherwise. Data on $E(Y_1|D = 1)$ is derived from the participants. The only problem here is to find $E(Y_0)|D = 1$. As a result, the difference between $E(Y_1|D = 1) - E(Y_0)|D = 1$ cannot be observed for the same household. This problem creates a need to use a better substitute to estimate ATT. The solution involves the use of mean outcome of the comparison individuals as a substitute control mean for the participants:

The ATT can only be identified when there is no self-section bias but if selection bias is present, the estimates are biased too (Gilligan and Hoddinott, 2007). The study area was purposively determined and this might have caused selection bias. To overcome this challenge and achieve this objective, Propensity score matching technique was applied. To enhance the validity of PSM, the treated and the control group were derived from similar agro-ecology and socioeconomic conditions (Tolemariam, 2010).

Propensity score matching method is preferred to the normal regression methods because it utilizes only comparable observations without imposing a functional form. This helps in overcoming the problems of multicollinearity and heteroscedasticity. Furthermore, the matching technique emphasizes the issue of common support thus avoiding the bias due to the extrapolation to non-data region. In addition, the results from the matching technique are easy to explain to policy makers, since the idea of comparison of similar group is intuitive. PSM depends on conditional independence and common support region assumptions. The implementation process involves five steps that include selection of variables to estimate the propensity scores, estimating the propensity scores, choosing the matching technique, assessment of overlap and common support, evaluation of the matching quality and calculation of sensitivity analysis (Caliendo and Kopeining, 2008).

Selection of variables to estimate the propensity scores

The variables that were included in the probit model consisted of both continuous and discrete variables and included, occupation, age, number of years of formal education, size of the household, marital status, presence of extension visit, local group membership and the gender of the farmer. In selecting these variables to estimate propensity scores, Conditional Independence Assumption (CIA), was utilized. It was assumed potential outcomes were independent of treatment status. This assumption was supported by the fact that the covariates selected determined the selection process, there were no unobserved confounders and there were high degree of post-match balance across the covariates. From the table, occupation, total land acreage and the presence of extension visit were statistically significant but after the post-match balance, none of the covariate was significant in influencing participation.

2.3 Data analysis

Estimation of the propensity scores

Estimation of propensity score was first accomplished using probit model following (Owuor, 2008) as shown:

$$pr\ ob(y_i = j) = \frac{\exp(\beta_j x_i)}{s(j=0-j)\exp(\beta_j x_i)}, j = 0{,}1 \quad \ldots\ldots 5$$

Where: the left hand side represent the probability of participation in snow peas farming for jth household and 'x_i' variables are characteristics of the observed household, which are the same across all outcomes. These include farmer's age in years (*Age*), education level given as the number of years of formal education (*Education*), gender of the household either male or female (*Gender*), household size in numbers (*Hhsize*), occupation of the farmer either pure farmer or non-pure farmer (*Occupation*), marital status of the farmer either married or not (*Martstatus*), extension visit to the farmer either a farmer was given extension service or not (*Extcame)* membership in any local organization as yes or no (*Localmemb*).

In linear form, equation is reduced to:

$$D(0{,}1) = \beta_0 + \beta_{ij} x_{ij} + \varepsilon, pscore(mypscore), [blockid(myblock), com\sup \quad \ldots\ldots 6$$

Where: *D* is the indicator for participation, whereby *D*=1, if a household is a participant in snow peas production and 0 otherwise. *Xi* represents a vector of participation covariates of the household which are common across all farmers. This is then followed by options commands that generate propensity score index '*pscore'*, for the program. Specification of the outcome ***Y*** (income, expenditure or assets) is also specified in the command. The option (Y) for common support generates a dummy variable, which identifies households that meet the matching condition. The common support variable attaches numerical '1' corresponding to the subjects that meet the matching condition and '0' to those that do not meet the condition.

Estimation of average effect of participation in the programme follows commands in stata, namely '*attnd'* for nearest neighbor matching, '*attr*' for radius matching, *'attk*' for kernel matching and *'atts'* for stratified matching methods. The general formulation of the empirical model is as follows:

$$command: Z = \beta_0 + \beta D + \beta_i\ X_i + \beta_i + \in, pscore(mypscore), comsup, logit \ldots\ldots 7$$

Where command stands for either one of the matching estimation above (*attns, attr, atts, attk*), 'Z' is the income, expenditure or assets. X_i is a vector of participation covariates, followed by the propensity score option, then the common support option. The two options are important since they sense the average effect of participation (AEP). It is computed from propensity score index.

Choice of matching algorithm

Estimation of propensity scores is not sufficient to estimate ATT. This is because propensity score is a continuous variable and the probability of observing more than one unit with similar propensity score is zero. To overcome this problem, various matching techniques have been proposed in the literature. Commonly used matching algorithms are nearest neighbor (NN), kernel matching and caliper maching (Tolemariam, 2010)

Nearest neighbor matching method is straightest forward. An individual from the control group is chosen as matching partner for the treated household that is closest in terms of propensity score (Caliendo and Kopeinig, 2008). It can be done with and without replacement. Matching with replacement increases the quality of matches but decreases degree of precision of estimates. On the other hand matching with replacement increase precision but it's liable to biasness (Dehejia and Wahba, 2002). Caliper matching on the other hand involves getting matching partner within a given range of propensity score and the closest partner in terms of propensity score (Caliendo and Kopeinig, 2008). The main problem with the technique is that it is difficult to know the choice for the tolerance level which is reasonable (Tolemariam, 2010). In kernel matching, all the treated individuals are matched with a weighted average of all the controls with weights which are proportional to the distance between the propensity scores of treated and control (Becker and Ichino, 2002). The drawback of this method is that it's possible to get bad matches as estimator hence it's important to impose the common support condition for kernel matching technique. Again, it will not be obvious how to set tolerance. However, kernel matching with 0.25 band width is mostly used (Mendola, 2007). The choice of matching method depends on the data in question (Bryson *et al.,* 2002). When there is considerable overlap in the distribution of propensity score between the control and the treated groups, most of the matching techniques yield similar results (Dehijia and Wahba, 2002).

Assessing of region of common support and overlap

Common support is also a mandatory option to ensure matching is done only on controls that are similar to participants (Bryson *et al.,* 2002). The common support region is the area which contains the minimum and maximum propensity scores of the treated and control households respectively. This ensures that only the subset of the comparison group that is comparable to the treated group should be applied in the analysis (Tolemariam, 2010). The basic way of achieving this is to delete all observations whose propensity score is less than the minimum and greater than the maximum in the opposite group (Caliendo and Kopeinig, 2008). This is because there is no match that can be made to estimate the average effects on the ATT parameter when overlap exists between the treated and control groups.

Testing the matching quality

The matching procedure should be able to balance the distribution of the relevant variables in both the treated and non-treated groups. This is because the conditioning is not done on all the covariates but on the propensity score. While differences in the variables are expected before matching, it should be avoided after matching. The idea behind balancing tests is to determine whether the propensity score is balanced well or to check if at each value of propensity score, a given characteristic has equal distribution for the control and treated groups (Tolemariam, 2010). The idea is to compare the condition before and after matching to examine if there is any differences after conditioning on propensity score (Caliendo and Kopeinig, 2008).

Sensitivity analysis

Checking the sensitivity of the estimated outcomes is increasing its important in applied evaluation literatures (Caliendo and Copeining, 2008). Matching technique is based on the assumption that all variables affecting participation decision and outcome variables are tested simultaneously. It is hard to test the assumption (unconfoundedness assumption) because the data are uninformative about the distribution of the controlled outcome for the treated units (Becker and Caliendo, 2007). Where the assumption does not hold, it means there are unorbservable covariates which influence the assignment into treated and the results simultaneously. This results in a hidden bias (Rosenbaum, 2002). This translated in biased estimation of ATT. The magnitude of bias depends on the strength of the correlation between the observable covariates and the treated outcomes (Tolemariam, 2010).

Sensitivity analysis therefore involves the testing of the robustness of the outcome deviation from the assumption. The main concern is whether the treated effects may be affected by unobserved factors. This may be tested using Rosenbaum bounding approach (Rosenbaum, 2002). The approach does not test the unconfoundedness assumption rather it provides evidence on the magnitude to which any significance outcome is dependent on this untestable assumption. The approach involves calculating upper and lower bounds, using the Wilcoxon signed rank test. The rank tests the null hypothesis of control effect for various hypothesized values of unobserved selection bias. In case the results are sensitive, the researcher might have to consider about the validity of the identifying assumption and think of other estimation methods (Tolemariam, 2010).

Choice and definition of explanatory variables

When estimating the propensity score, the interest is not in the effects of covariates on the propensity score because the aim of the work is to determine the impact of snow peas growing on the outcome variables. Omitting important variables can increase the bias it the outcome (Heckman *et al.,* 1997). In this particular case, covariates that determine households' decisions to participate in snow peas production could affect the outcome variables in question. Pre-intervention characteristics, that brings differences in outcome of the interest among snow peas growers and non-growers were used. There is no general criterion for which variables to include in the model (Anderson et *al.,* 2009). However, the choice of variables is guided by the economic theory and empirical studies to know which observable independent variables to include in the model (Bryson *et al.,* 2002). The covariates used are identified in Table 1:

2.4 Definitions of explanatory variables

Table 1: Explanatory variables definition and measurements

Variable	Types and definition	Measurements
Occupation	Dummy, pure or non-pure farmer	1 if pure farmer, 0 otherwise
Age	Continuous age of the household head	In years
Extension came	Dummy, yes or no	1 if yes, 0 otherwise
Local membership	Dummy, yes or no	1 if yes, 0 otherwise
Household size	Continuous, total family size	Number of household
Years of education	Continuous, number of years of formal education	in years
Land size	Continuous, size of land owned	In hectares
Marital status	Dummy, married or otherwise	1 if married, 0 otherwise

Choice, measurements and indicators of the outcome variables

Income

It is one of the outcome variables as a result of household's participation in snow peas production. It was measured in Kenya shillings per year and it is calculated as the total revenue from all the income generating sources of the farmer.

Expenditure

It is another outcome variable used to determine the welfare of individuals. It was measured in Kenya shillings and calculated by adding consumption expenditure and all other expenses incurred by a particular household per month.

Assets

It is the value of all the items a household owns. It includes the value of land, furniture, tools and equipment, livestock and anything else that can be disposed to generate income. They were valued at the current market price minus the depreciation cost for the assets that depreciates.

3.0 RESULTS AND DISCUSSION

3.1 Differences in income, assets and expenditure of adopters and non-adopters of snow peas enterprise

The overall mean age of farmers was found to be 38.94 years and the figure reflects the age of individual participants and non-participants. This indicates that there is no significant difference between ages of the two categories of farmers. Age is therefore not a determinant of whether a farmer participates in snow peas production or not.

There is no significant difference between the household sizes of the farmers in the two different categories. The average household size of all the farmers was 5 members. Household size is therefore not a determinant of participation in snow peas production. The levels of education of farmers do not influence the decision to participate in snow peas production. There is no difference in education level of both snow peas farmers and the control group where the overall average is 9.50 years of education. In addition; the size of land under the ownership of the farmers does not affect the participation decision of the farmer. The results also show that there is no statistically significant difference between the size of the land of the treated and the control group.

There is a significant difference in the outcome variables where participants had higher values compared to non-participants. On average participants had Kes 406,111.20, Kes 413,114.40 and Kes 24,974.49of income, assests and expenditure respectively; while non-participants had Kes 22,8630.30, Kes 232221.30 and Kes 11810.76 of income, assests and expenditure respectively (Table 2). This indicates higher household gains for partucipants which could be attributed to snow peas farming. On the other hand, result depicts an association between participation in snow peas farming and occupation of the farmer, extension service as well as group membership (Table 3).

Table 2: Descriptive statistics of sample households

Variable	Overall Mean	Sd	Participant Mean	Sd	Non-participant Mean	Sd	t-test
Age	38.94	6.52	38.94	6.52	39.94	8.19	1.058
Income	406111.20	132388.80	406111.20	132388.80	228630.30	63854.55	-14.61***
Assets	413114.40	162210.60	413114.40	162210.60	232221.30	178825.80	-8.40***
Expenditure	24974.49	26656.63	24974.49	26656.63	11810.76	9685.43	-5.71***
Household size	5.08	3.03	5.08	3.03	4.85	2.87	1.57
Years of education	9.50	3.45	9.50	3.45	9.50	2.52	0.01
Total land size	2.71	1.89	2.71	1.89	3.10	2.09	1.57

Note: **** represents significance level at 1%*

Table 3: Descriptive statistics of household dummies and their association with participation

	snow peas grower No	%	Yes	%	Total	%	chi
Farmers occupation							6.3518**
Non- pure farmer	37	14	40	15	77	29	
pure farmer	123	46	67	25	190	71	
Marital status							5.0474
Married	144	54	95	36	239	89	
Never married	3	1	7	3	10	4	
Widow	4	1	1	0	5	2	
Divorced	9	3	4	1	13	5	
Extension							4.7581**
No	133	50	77	29	210	79	
Yes	27	10	30	11	57	21	
Group membership							72.6473***
No	133	50	60	22	193	72	
Yes	0	0	47	18	47	18	

Note: ***, *** represents significance level at 5% and 1% respectively*

3.2 Factors influencing participation in snow peas production

The logistic regression model was applied to estimate the propensity score matching for participant and non-participants households. The mean average of the generated propensity scores was 0.04 with the range of 0.12 to 0.82. This means, the probability of any randomly selected farmer being a participant is 40%.

Considering the estimated coefficients (Table 4), the outcome indicates that snow peas' growing is significantly influenced by five explanatory variables. That is, occupation, total land size, extension service, marital status and gender of the farmer. From the results, Farmers who are also involved in other income generating activities or careers are likely to participate in snow peas growing. Being a pure farmer decreases the likelihood of participation. Occupation was found significant at 1% significant level. This could be explained by their desires to diversify risk by trying more than one business ventures.

Farmers with large land size are less likely to participate in snow peas production in that, an increase in land size by 1 acre decreases the likelihood of participation in snow peas production by 0.0307 at 10% significant level. The entire process of snow peas production is intensive and requires a lot capital, time and most management practices require attention. Most farmers would prefer to dedicate the production in small portions of land that are manageable. This probably explains why farmers with small pieces of land are more likely to be snow peas growers.

The positive and significant extension coefficient depicts that farmers who received extension service are likely to participate in snow peas production than their counterparts. The results were significant at 10% confidence level. An extra extension visit increases the likelihood of participation in snow peas production. Extension agents play a vital role in introducing new crops and technologies to farmers. Probably, in the course of their visits to the farmers, the extension agents influenced farmers to try snow peas production. This explains why those farmers who were visited by the agents have a high possibility of participating in snow peas growing.

The results depict that being singe and never married increases the likelihood of participation in snow peas compared to being married by 1% significance level. People who are single make independent decision. This could explain why they are more likely to participate in snow peas production. Again, being a widow decreases the likelihood of participation compared to being married at 5 % significance level. Most widows are aged and lack motivation and incentives to participate in snow peas production. This could be used to support the results Being male increased the probability of participation in snow peas production at 1% level of significance. In most African societies, male are the head of the family and consequently the decision makers. This could be the reason why in the households they are the head are likely to be snow peas growers.

Table 4: Factors influencing participation in snow peas production

Variable	dy/dx	Standard Error	P-value
Occupation	-0.2492***	0.0750	0.0010
Age	0.0023	0.0048	0.6280
Years education	-0.0183	0.0121	0.1280
Total land size	-0.0307*	0.0168	0.0680
House hold size	0.0013	0.0111	0.9060
Single never married#	0.3861***	0.1423	0.0070
Widow#	-0.2952**	0.1155	0.0110
Divorced#	-0.1200	0.1422	0.3980
Visited by extension officer	0.1525*	0.0812	0.0600
Local group membership	-0.0450	0.0747	0.5470
Gender	0.2515***	0.0827	0.0020

Note: **, **, *** represents significance level at 10%, 5% and 1% respectively; # dummy for marital status*

3.3 Matching participants and comparison households

Before the actual matching task, four steps are involved: predicting the propensity score, examining the common support condition, discarding observations whose predicted propensity scores lies outside the range of common support and conducting a sensitivity analysis.

The estimated propensity scores ranges between 0.12 and 0.78 with an average of 0.37 for the treated households and between 0.16 and 0.82 with an average of 0.46 for control households. From the results, the common support region would lie between 0.16 and 0.78. To satisfy the common support condition, households whose estimated propensity scores are less than 0.16 and greater than 0.78 were dropped and not considered for the matching exercise. As a result, 2 participating households were discarded from the analysis.

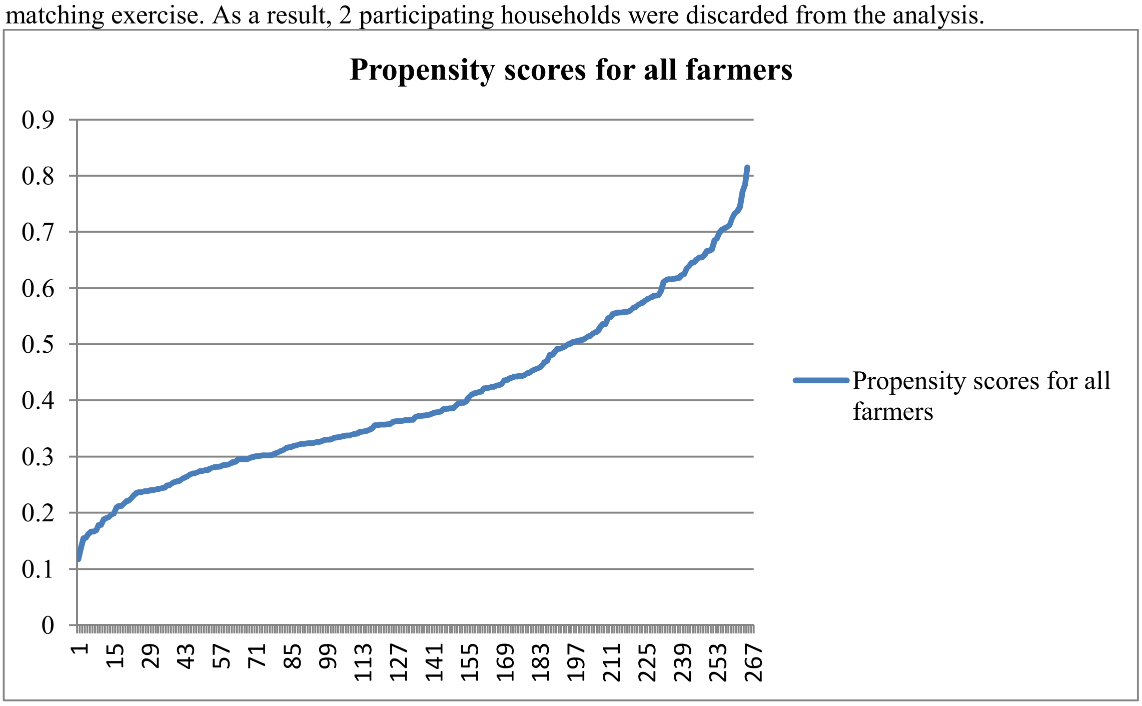

Figure 1: Propensity scores line graph for all the farmers

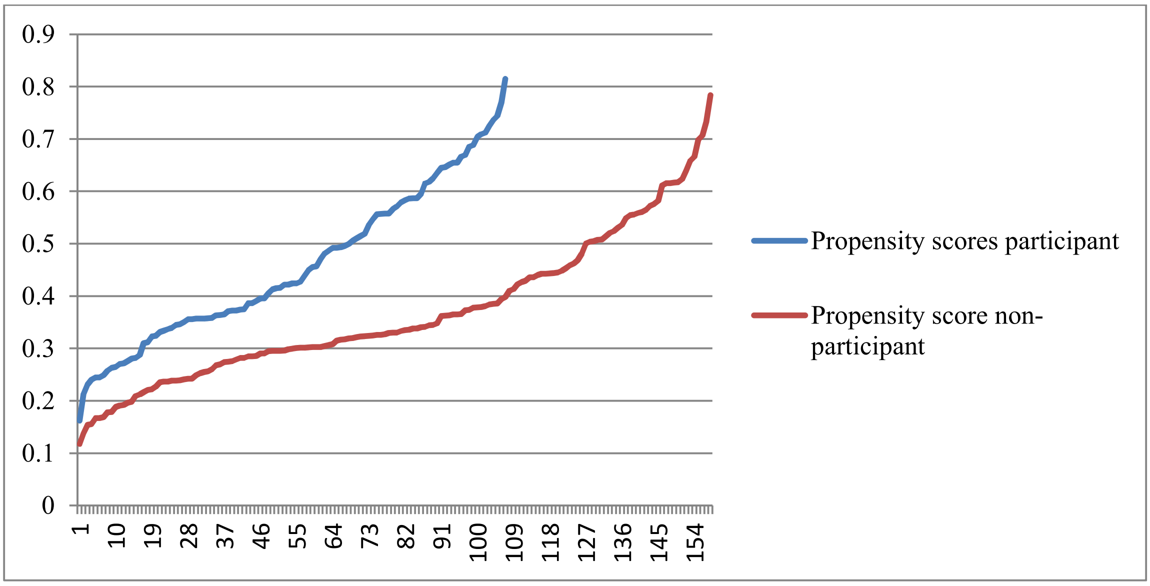

Figure 2: Propensity score graph for participants and non-participants

Choice of matching algorithm
Various matching estimators were considered in matching the treated and control group in the region of common support. The tests showed that the choice of matching algorithm chosen is suited for the data set at hand. Thus, we can progress to estimate ATT for households.

Testing the balance of propensity score and covariates
After deciding on the best matching technique, the next step is to check the balancing of propensity score and variables using different procedures by selecting matching procedure table 5. The main purpose of estimating propensity scores is to balance the distribution of relevant covariates in to both the treated and control groups. In order to ascertain the balancing powers of the estimates, different test methods such as the reduction in the mean standardized bias between the treated and the control groups, equality of means though t-test and chi-square for joint significant tests of the covariates in question are used.

The table shows the mean standard bias before and after matching. The first column presents all the chosen variables in the model. The second and the third column indicate the mean of the treated and control before matching respectively. The t-test of the treated before matching is highlighted in the fourth column. The fifth and the sixth column indicate the mean of the treated and control after matching respectively. The last column shows the t-test after matching.

The table also shows that there are statistically significant differences between the t-tests of the chosen variables before matching. After matching, all the variables are balanced.

Table 5: Test for balance of propensity score and covariates

	BEFORE MATCHING	**(N=267)**		**AFTER MATCHING(N=265)**		
	Mean			**Mean**		
Variable	**Treated**	**Control**	**t-test**	**Treated**	**Control**	**t-test**
Occupation	0.6262	0.7785	-2.73***	0.6262	0.6320	-0.09
Age	38.9440	39.9370	-0.09	38.9440	39.1440	-0.19
Extension service	0.2804	0.1646	2.28**	0.2804	0.2774	0.05
Local membership	0.7103	0.7405	-0.54	0.7103	0.6852	0.40
Household size	5.0841	4.8354	0.68	5.0841	5.0144	0.18
Years education	9.4953	9.4810	0.04	9.4953	9.4309	0.16
Total land size	2.7056	3.0968	-1.55***	2.7056	2.7522	-0.18
Marital status	1.1963	1.2215	-0.29	1.1963	1.2203	-0.26

Note: ***, *** represents significance level at 5% and 1% respectively*

A low pseudo-R^2 and the insignificant likelihood ratio tests shows that both the treated and control groups have the same distribution of covariates X after matching (table 6)

Table 6: Chi-square test for joint significant of variable

Sample	Ps R2	LR chi2	p>chi2
Unmatched	0.052	18.46	0.018
Matched	0.001	0.41	0.99

Table 7: Impact estimate on total household net income

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Income	Unmatched	406111.21	228519.30	177591.91	12215.93	14.54
	ATT	406111.21	230292.78	175818.42	14192.80	12.39

The results in table 7 show that there is a statistically significant difference in income between snow peas farmers and those who do not grow the crop. The estimates are significant at 10% level. Most snow peas farmers pointed out that they harvest snow peas crop two days per week. This means that they have a stream of income especially when the markets are good. As opposed to their counterparts who only depend on other types of income source, snow peas farmers have an additional source of income. This explains why they probably have more monthly net income as compared to non-snow pea's growers.

The results are contrary to the findings of Tolemariam (2010) who found that households' participation in market development intervention by coffee producers did not have statistically significant impact on their income.

Table 8: Impact estimate on expenditure

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Expenditure	Unmatched	24974.49	11815.96	13158.52	2318.86	5.67
	ATT	24974.49	11175.20	13799.29	2740.23	5.04

The results depict a statistically significant between the expenditure of participants and non-participants. Those farmers who are involved in snow peas production are depicted to spend more amount of money on various consumption expenditures. This could be explained by the fact that snow peas could be giving them more income compared to their counterparts. They are therefore able to afford all the basic commodities in satisfying amounts. They are able to afford good education for their children by taking them to good schools. They are able to afford booth fresh and non-fresh staples more frequently than non-participants. They are also able to afford being members of saving societies and organizations. For the non-participants, their income comes only from other investments which participants also have. This makes it hard for them to purchase commodities less frequently and in fewer amounts. This explains the statistically significant difference.

Table 9: Impact estimate on assets

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Assets	Unmatched	413114.39	232374.68	180739.71	21662.35	8.34
	ATT	413114.39	230942.49	182171.90	23283.25	7.82

Assets was measured in terms of the value of all the durable commodities a farmer has ranging from livestock, furniture, electronics, land, tools and equipment. Their value was estimated using current market prices. The results depict a significant difference between the assets of participants and non-participants. Snow peas farmers are likely to use part of their income from snow peas production to purchase assets and this explains why they probably have more assets than their control counterparts.

3.4 The sensitivity of the results

Stata Mhbounds was applied to compute Mantel-Haenszel bounds to check sensitivity of estimated average treatment effects and critical hidden bias. Appendix Table 10, 11 and 12 contains the test results. $\Gamma = 1$ indicates an absence of unobserved factors. The bounds were increased slightly by 0.5 and the various levels of bounds tells us at which degree of unobserved positive or negative selection the effect would become significant. From the results the Q_mh+ and Q_mh- test statistic gave a similar result across all bound of odds assigned due to unobserved factors. The negative values of Q_mh+ therefore indicated negative selection bias where snow peas farmers tend to have low annual income, expenditure and value of assets even without participation in production of snow peas. The bias was however not significant at different bound levels in the case of overestimation and underestimation of the treated effect as indicated by P_mh + and P_mh- values. Result on the tables' further show that the study was insensitive to bias that will double or triple the odds of change in the level of income, assets and expenditure as a result of participation in snow peas production.

4.0 CONCLUSION AND POLICY IMPLICATIONS

The aim of the study was to determine the households' effects of snow peas production in Kenya. The study achieved this by comparing the annual income, expenditure and total assets of participants and non-participants using propensity scores matching method. The study showed that the impact of snow peas farming was significantly different between participants and non-participants in terms of their income, assets and expenditure. The estimation of treated effect on treated showed that participating farmers had relatively higher monthly income, higher total value of assets and higher expenditure compared to non-participating farmers. To enhance participation of farmers in the production of snow peas, farmers should be provided with up to date information concerning snow peas farming by extension service providers. Farmers should be educated on snow peas varieties, seed rate, agro-chemicals, spacing and any other technology that can improve snow peas production. Farmers are encouraged to form groups and organizations to will help them in lowering the entire cost of production through bulk buying of inputs, bulk marketing of products, increased bargaining power, eliminating middlemen and access to credit. The government should provide farmers with credit facilities and support the entire value chain of snow peas because it's a profitable venture.

REFERENCES

Anderson, S. R., Auquier, A., Hauck, W. W., Oakes, D., Vandaele, W., and Weisberg, H. I. (2009). *Statistical methods for comparative studies: techniques for bias reduction* (Vol. 170). John Wiley & Sons.

Authority, E. P. Z. (2005). Horticulture Industry in Kenya 2005. Nairobi, Kenya.

Becker, S. O., and Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The stata journal*, *2*(4), 358-377.

Bernard, T., Spielman, D. J., Seyoum Taffesse, A., and Gabre-Madhin, E. Z. (2010). *Cooperatives for staple crop marketing: evidence from Ethiopia* (Vol. 164). Intl Food Policy Res Inst.

Bryson, A., Dorsett, R., and Purdon, S. (2002). The use of propensity score matching in the evaluation of active labour market policies.

Caliendo, M., and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, *22*(1), 31-72.

Chemining'wa George, N., Ngeno, J., Muthomi, J. W., and Shibairo, S. I. (2012). Effectiveness of indigenous pea rhizobia (Rhizobium leguminosarum bv. viciae) in cultivated soils of central Kenya. *Journal of Applied Biosciences*, *57*, 4177-4185.

Davis, J.R. (2006). How Can the Poor Benefit from the Growing Markets for High Value Agricultural Products? Kent: Natural Resources Institute.

Dehejia, R. H., and Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, *84*(1), 151-161.

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., and Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health services research*, *49*(5), 1701-1720.

Gilligan, D., Hoddinott, J., Taffesse, A. S., Dejene, S., Tefera, N., and Yohannes, Y. (2007). Ethiopia Food Security Program: Report on 2006 Baseline Survey. *International Food Policy Research Institute, Washington, DC Photocopy*.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, *64*(4), 605-654.

Heckman, J., Ichimura, H., Smith, J., and Todd, P. (1998). *Characterizing selection bias using experimental data* (No. w6699). National bureau of economic research.

Heinrich, C., Maffioli, A. and Vázquez, G. (2010). A Primer for Applying Propensity-Score Matching: impact-Evaluation Guidelines. Inter-American Development Bank.

Jalan, J., and Ravallion, M. (2003). Estimating the benefit incidence of an antipoverty program by propensity-score matching. *Journal of Business & Economic Statistics*, *21*(1), 19-30.

James, K. (2007). Export Oriented Crop Production: Snow Peas and French Beans in Mount Kenya Region. Retrived on April, 2016 from http://www.google.com/url?sa= t&rct= j&q=&esrc=s&source=web&cd=9&ved=0ahUKEwjz1pOHvYbMAhWJSRoKHQPMBJYQFghIMAg&url =http%3A%2F%2Fwww.ravinvest.biz%2Fdownloads%2FStories%2FStoryKanyeMOPAExports.pdf&usg =AFQjCNE_Vdu7BGWKL0hnVHJ2KtYB0h2Jjg

Kimiti, J. M., Odee, D. W., and Vanlauwe, B. (2009). Area under grain legumes cultivation and problems faced by smallholder farmers in legume production in the semi-arid eastern Kenya.

KNBS, K. (2013). Economic Survey 2013 Highlights.

Lamborn, C. (2001). Origin, history, and genetic improvement of the snap pea (Pisum sativum L.). *Plant Breeding Reviews*, *21*, 93.

Mark, G. (2016). Edible Pod Production in California. University of Carlifornia Cooperative Extension Farm Advosor, Santa Barbara County.

Mendola, M. (2007). Agricultural technology adoption and poverty reduction: A propensity-score matching analysis for rural Bangladesh. *Food policy*, *32*(3), 372-393.

Mutai, B.K. (2014). *How To Write a Quality Research Proposal: A Complete Simplified Recipe* (Rev. 2nd ed.). Pune, Maharashtra, India: Dream Web Solutions.

Nguguna N., Kamau T. and Owino, O. Unlocking the Future Potential for Kenya: The Vision 2030 chapter 4 (preliminary draft).

Njarui, D. M. G., Gichangi, E. M., Gatheru, M., Nyambati, E. M., Ondiko, C. N., Njunie, M. N., ... & Ayako, W. (2016). A comparative analysis of livestock farming in smallholder mixed crop-livestock systems in Kenya: 1. Livestock inventory and management. *Livestock Research for Rural Development*, *28*, 66.

Odero, D. O., Mburu, J., Ackello-Ogutu, C. and Nderitu, J. H. (2013). 832 Value Chain Analysis Of Smallholder Snap Bean Production In Kirinyaga County, Kenya.

Okello, J.J., Narrod, C. and Roy, D. (2007). Institutional Innovations for Smallholder Compliance with International Food Safety Standards: Experiences from Kenya, Ethiopian and Zambian Green Bean Growers. *African Association of Agricultural Economists (AAAE) Conference Proceedings held in Ghana 2007(483-487).*

Rosenbaum, P. and Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Casual Effects, *Biometrika*, 70, 41-55.

Rugenyi, A.W (2011). An Evaluation of Smallholder Snow Peas Market Organization and performance in Buuri and Laikipia East Districts, Kenya.

Sekhon, M.K., Amrit, K.M., Manjeet, K. and Sidhu, M.S. (2010). Technical Efficiency in Crop Production: A Region Wise Analysis.

Smith, J. A., and Todd, P. E. (2001). Reconciling conflicting evidence on the performance of propensity-score matching methods. *The American Economic Review*, *91*(2), 112-118.

Tolemariam, A. (2010). *Impact assessment of input and output market development interventions by IPMS Project: The case of Gomma Woreda* (Doctoral dissertation, Haramaya University).

Weinberger, K., and Lumpkin, T. A. (2007). Diversification into horticulture and poverty reduction: a research agenda. *World Development*, *35*(8), 1464-1480.

ACKNOWLEDGMENT

The authors would like to acknowledge African Economic Research Consortium (AERC) for financial support in funding the project. Special regards to the Department of Agricultural Economics and agribusiness management of Egerton University for the support through the staff. We present our sincere gratitude to the enumerators for the good work. Final regards goes the households representatives who took their time to respond to the interview.

APPENDIX

Table 10: Mantel-Haenszel (1959) Bounds for income

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1.00	.	.	.	.
1.05	-0.0818	-0.0818	0.532597	0.532597
1.10	.	-0.0818	.	0.532597
1.15	-0.0818	-0.0818	0.532597	0.532597
1.20	-0.0818	-0.0818	0.532597	0.532597
1.25	-0.0818	-0.0818	0.532597	0.532597
1.30	-0.0818	.	0.532597	.
1.35	-0.0818	-0.0818	0.532597	0.532597
1.40	-0.0818	.	0.532597	.
1.45	.	-0.0818	.	0.532597
1.50	-0.0818	-0.0818	0.532597	0.532597
1.55	-0.0818	-0.0818	0.532597	0.532597
1.60	-0.0818	-0.0818	0.532597	0.532597
1.65	-0.0818	-0.0818	0.532597	0.532597
1.70	-0.0818	-0.0818	0.532597	0.532597
1.75	-0.0818	-0.0818	0.532597	0.532597
1.80	-0.0818	-0.0818	0.532597	0.532597
1.85	-0.0818	.	0.532597	.
1.90	-0.0818	-0.0818	0.532597	0.532597
1.95	-0.0818	-0.0818	0.532597	0.532597
2.00	-0.0818	-0.0818	0.532597	0.532597

Table 11: Mantel-Haenszel (1959) Bounds for expenditure

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1.00	.	.	.	.
1.05	-0.0818	-0.0818	0.532597	0.532597
1.10	.	-0.0818	.	0.532597
1.15	-0.0818	-0.0818	0.532597	0.532597
1.20	-0.0818	-0.0818	0.532597	0.532597
1.25	-0.0818	-0.0818	0.532597	0.532597
1.30	-0.0818	.	0.532597	.
1.35	-0.0818	-0.0818	0.532597	0.532597
1.40	-0.0818	.	0.532597	.
1.45	.	-0.0818	.	0.532597
1.50	-0.0818	-0.0818	0.532597	0.532597
1.55	-0.0818	-0.0818	0.532597	0.532597
1.60	-0.0818	-0.0818	0.532597	0.532597
1.65	-0.0818	-0.0818	0.532597	0.532597
1.70	-0.0818	-0.0818	0.532597	0.532597
1.75	-0.0818	-0.0818	0.532597	0.532597
1.80	-0.0818	-0.0818	0.532597	0.532597
1.85	-0.0818	.	0.532597	.
1.90	-0.0818	-0.0818	0.532597	0.532597
1.95	-0.0818	-0.0818	0.532597	0.532597
2.00	-0.0818	-0.0818	0.532597	0.532597

Table 12: Mantel-Haenszel (1959) Bounds for assets

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1.00	.	.	.	.
1.05	-0.0818	-0.0818	0.532597	0.532597
1.10	.	-0.0818	.	0.532597
1.15	-0.0818	-0.0818	0.532597	0.532597
1.20	-0.0818	-0.0818	0.532597	0.532597
1.25	-0.0818	-0.0818	0.532597	0.532597
1.30	-0.0818	.	0.532597	.
1.35	-0.0818	-0.0818	0.532597	0.532597
1.40	-0.0818	.	0.532597	.
1.45	.	-0.0818	.	0.532597
1.50	-0.0818	-0.0818	0.532597	0.532597
1.55	-0.0818	-0.0818	0.532597	0.532597
1.60	-0.0818	-0.0818	0.532597	0.532597
1.65	-0.0818	-0.0818	0.532597	0.532597
1.70	-0.0818	-0.0818	0.532597	0.532597
1.75	-0.0818	-0.0818	0.532597	0.532597
1.80	-0.0818	-0.0818	0.532597	0.532597
1.85	-0.0818	.	0.532597	.
1.90	-0.0818	-0.0818	0.532597	0.532597
1.95	-0.0818	-0.0818	0.532597	0.532597
2.00	-0.0818	-0.0818	0.532597	0.532597

Gamma : odds of differential assignment due to unobserved factors
Q_mh+ : Mantel-Haenszel statistic (assumption: overestimation of treatment effect)
Q_mh- : Mantel-Haenszel statistic (assumption: underestimation of treatment effect)
p_mh+ : significance level (assumption: overestimation of treatment effect)
p_mh- : significance level (assumption: underestimation of treatment effect)