COMMERCIAL OPPORTUNITY FOR INDIGENOUS ALOE VERA PRODUCTS APPEAR TO BE INCREASING AS THERE IS A SUSTAINED DEMAND FROM INTERNATIONAL AS WELL AS AFRICAN MARKET WITH HIGH PARTICIPATION OF WOMEN. HOWEVER, ITS INCOME IMPACT ON THE LIVELIHOOD OF HOUSING IS NOT YET ADDRESSED. THEREFORE, THIS STUDY EVALUATED THE INCOME IMPACT OF PASTORALIST WOMEN PARTICIPATION IN ALOE VERA SOAP PRODUCTION IN YABELLO DISTRICT, BORANA ZONE OF OROMIA, ETHIOPIA. BOTH PRIMARY AND SECONDARY DATA WERE COLLECTED FROM 200 SAMPLE HOUSEHOLDS USING SEMI-STRUCTURED QUESTIONNAIRE. TO ANALYSE THE DATA BOTH DESCRIPTIVE AND INFERENCEAL STATISTICS AND PROPENSITY SCORE MATCHING MODEL WERE APPLIED. THE PROPENSITY SCORE MATCHING WAS APPLIED THE REQUIRED MATCHING PROCESSES, COVARIATE BALANCING AND SENSITIVITY ANALYSIS TESTS. THE RESULT SHOWS THAT PARTICIPATION OF WOMEN IN ALOE VERA SOAP PRODUCTION HAS INSIGNIFICANT RESULT WITH IMPACT ON HOUSEHOLD INCOME. HOWEVER, THE PROPENSITY SCORE MATCHING ALSO INDICATES AVERAGE TREATMENT EFFECT ON TREATED INCOME IS 45.693 BIRR. RESULT OF SENSITIVITY ANALYSIS FURTHER SHOWS THAT THE ESTIMATED EFFECTS ARE INSENSITIVE TO UNOBSERVED SELECTION BIAS WITHIN GAMMA LEVEL USED. THUS, ALOE VERA SOAP PRODUCTION SHOULD BE ENCOURAGED FOR THE PASTORALIST SOCIAL WELLBEING.
significant role in their family economy.

In Ethiopia, there is a participation of women groups in Bio-enterprise developments in sustainable wild harvest and domestication of indigenous Aloe species and gums product of non-forest forest products (NTFPs). Aloe is one of a vast growing plant species in Ethiopia even though the commercialization of this amole resource by residents and transient communities is very low (Demissew and Nordal, 2010).

As Wren and Mamo (2009) currently a few Bioenterprise are established at different parts of Ethiopia to utilize the natural resource in an area. Their goal is to contribute to the poverty reduction, economic empowerment and social wellbeing of pastoralists, particularly women. Example; rural women groups that participate in Aloe Vera soap production came from Tigray and Amhara region to Borana zone for cross visits organized to rural enterprises initiated by women groups for experience sharing.

As presented by Hurst, et al., (2012) in Borana community, women generate income for their family by establishing women’s groups or savings cooperatives. Mostly NGOs such as CARE, Gayo and SOS Sahel provide financial support to women hoping to engage in petty trade businesses. These activities are purchasing sugar, alcohol, coffee, tobacco, butter, milk and tea leaves and then returning to the village to sell these items, thus earning a small profit. Women who live in close proximity to forests or wooded areas (special in Yabello and Arero districts) may use products from the forest to supplement their incomes. Organizations such as SOS Sahel encourage forest management and teach women how to use products from the forest such as aromatic wood product that is used as perfume, collect gum from trees and sell incense/myrrh to generate profit.

The Aloe Vera soap production established by SOS Sahel Ethiopia in Borana pastoral area in 2006 after they did an assessment on natural resource found in that area. Seed money was funded by the European (Milky union, 2017). Before this assessment Aloe plants taken as bush clearing but the result of assessment found that aloe is one of economical and medical wild tree in an area. Their intention was utilization of wild resource like aloes which is found in an area so as to improve livelihood of local community and social wellbeing of pastoralists, particularly women in sustainable manner. Indeed, it enables local community (like women, youth and people at different economic level) to participate in economic activities by starting Aloe Vera soap production opportunity.

Even though in most pastoral community, women roles are reproductive and they mostly work in the house like fetching water, collecting firewood, cooking food, cleaning and child care (Lasage et al., 2010), this project gives them chance to participate in productive activities to their family living. Today this project has becoming to expand into five districts of the zone.

As usually known the livelihood of pastoral and nomadic people are more relied on livestock (live animal and its product). Problem with this system in the study area is that they mostly depend on natural climate condition (rainfall).

However, the nature of rainfall becomes very erratic and there is a high drought expansion throughout the district of Borana zone including the study area. As Tilahun (2015), this drought causes shortage of feed and water for both livestock and human beings. Thus, severe death on livestock, and human hungry (food insure) is an end result of the drought. Therefore, what more important for this pastoral community is diversifying their basis of livelihood through utilization of available wild resource in an area with all participation of their local community without discrimination of sex and minor groups and people in all age. Aloe is one of those trees that identified as elaborated above.

Previous studies such as Tesfome (2014) tried to address Aloe soap value chain initiatives and its effect on livelihood diversification strategy, by assessing the impact and determinants of participation in three kebele or peasant associations. Moreover, Tesfaye (2006), Demissew and Nordal (2010), Oda and Erena (2017), and Walker (2017) conducted study on the aloe species and distribution in different parts of Ethiopia including Yabello district. But, to the knowledge of this study the impact of Aloe Vera soap production particularly for women economic empowerment and income of household were not yet addressed. Therefore, considering aloe as one of commercial natural resource, and women as important human capital for Borana pastoral community, this study aims to fill the gaps by analysing the income impact of pastoralist women participation in Aloe Vera Soap production in Yabello district, Borana Zone of Oromia, Southern Ethiopia.

DATA AND METHODS

Description of Study Area

The study was conducted in Yabello district, Borana zone, southern Oromia State in Ethiopia. Yabello is located in the coordinates of 4°25’ – 5°15’ North latitude and 37°50’- 38°50’ East longitude. The ecology of the zone is arid and semi-arid savannah (Beza, 2011), with an altitudinal range of 1000 -1600 masl (McCarthy et al., 2001). Borana zone rangelands are dominated by tropical savannah vegetation with varying proportions of open grasslands and perennial woody vegetation (Homann et al., 2007).

The mean annual daily temperature and a mean annual rainfall are around 19°C and below 600mm respectively. It is frequently exposed to droughts that characterized as hot temperature and erratic rainfall. The erratic rainfall pattern causes vast area of the zone is not suitable for crop production. As (JICA, 2015; Tilahun, 2015), there are two rain season; long rainy ganna (from March to May) and short rainy Hagayya (from September to November) and the other remain months are dry season.

According to Demissew and Nordal (2010), the vegetation type in Borana zone was Acacia-Commiphora woodland and bush land with scattered remnant forests. This vegetation type is particularly rich in Aloes and other lilies including quite a few endemic or near endemic Aloe species. The natural populations taking place in study area is found in Sidamo floristic region.
Research Design

The study employed survey research of cross-sectional design. This cross-sectional survey study collected the data from study population at a single time from March to April of 2018 to examine the relationship between variables of interest. The study target population was both participant and non-participant women who are living in Yabello district of Oromia region, Ethiopia.

The study employed multi-stage sampling techniques that involve a combination of both simple random sampling and purposive sampling techniques.

In the first stage, on the basis of production of Aloe Vera soap out of the 12 kebeles of Yabello district, five sample kebeles were selected using stratification basis on the existence of intervention of Aloe Vera soap production in kebeles or not. The reason for this stratification is to have sample data from both kebeles of intervention and non-intervention of Aloe Vera soap production program. The first stratum was included women living in Aloe Vera soap production kebeles and a second stratum was accounted for non-participant women from the kebeles where production is not introduced yet.

Accordingly, Yabello district has only three kebeles (Dadim, Dida Yabello and Dikale) which produces Aloe Vera soap production; so that all three kebeles were selected purposively based on their production. From the other nine remain kebeles of which production is not yet started; two kebeles (Areri and Hara Awatu), for the case of taking as control group in order to reduce spillover effect for analysis of the impact of women participation in Aloe Vera soap, were selected purposively. This is because compared to others; those two kebeles are relatively more similar to intervention kebeles as they are on the same agro-ecology and on the same livelihoods basis (according to the interview of Yabello Pastoralist Development Office). Spillover is occurred due to interaction between participant and non-participant women live in the same village. Since this interaction is out of the control so taking some of the sample from the people outside the kebeles of Aloe Vera soap production is important.

On the second stage, sample respondents were selected randomly from each stratum on the basis of proportion to their population size. This is due to the homogeneity of population of study area.

Using the population data from (Yabello Pastoralist Development Office, 2017) the sample size was calculated since this data is not available on the CSA 2014 Ethiopian population projection. As in 2015/2016 Borana zone was separated into two zones and then many of district’s and kebele’s data were changed. By using the Yamane (1967) formula for sample size determination (Eq. 1) the study computed a total of representative 200 sample size.

\[
n = \frac{N}{1+\frac{N\sigma^2}{N}}
\]

Method of Analysis

The demographic and determinant characteristic of respondents were analysed using the combination of simple descriptive and inferential statistics. In addition, Propensity Score Matching model applied.

The propensity score matching of non-experimental methods of impact evaluation was used in ordered to assess the income impact of pastoralist women participation in the Aloe Vera soap production in the Yabello district of Borana zone, southern Ethiopia.

PSM builds a statistical comparison group that is based on a model of the probability of participating in the production, using observed characteristics. Participants are then matched on the basis of this probability, or propensity score, to nonparticipants. The average treatment effect of the production on treated is then calculated as the mean difference in outcomes across these two groups. PSM valid if two conditions are satisfied:

(a) Conditional Independence Assumption; (namely, that unobserved factors do not affect participation). It assumes that given a set of observable covariates X which are not affected by treatment, potential outcomes are independent of treatment assignment. This condition was tested using sensitive analysis.

(b) Sizable Common Support or Overlap: is about probability, propensity scores, across the participant and nonparticipant samples (Khandker et al., 2010). Common support assumption is central for this study analysis and both on support as well as off support households are found in the result. But, for all respondent overlap range is 0 <P(D = 1|X) < 1. This ensures that persons with the same X values (observed covariates or characteristic X) have a positive probability of being both participants and non-participants. Consider the outcome of participants after participating in the production as Y1 and the outcome of nonparticipants or control households as Y0. This with-and-without group comparison measures the program’s effect (participating in production) as Y1 - Y0. This difference is called impact of intervention. But this measure is not always give a right estimate of program effect because of the pre intervention situations of treated and control groups, the counterfeit comparison could yield an over- or underestimation of the program’s effect.

The basic evaluation problem comparing outcomes Y across participant and non-participant individuals i; is given by Eq. 2. Income is an outcome variable.

\[
Y_i = \alpha X_i + \beta T_i + \epsilon_i
\]

Where:

T is a dummy equal to 1 for those who participate and 0 for those who do not participate, \( X_i \) is a set of other observed characteristics of the individual and perhaps (maybe) of her household and local environment and \( \epsilon \) is an error term reflecting unobserved characteristics that also affect Y.

The average treatment effect (ATE) of the program might be represented by Eq. 3.

\[
D = E(Y_i(1) | T_i = 1) - E(Y_i(0) | T_i = 0)
\]

Where: D is representing an impact of program.

The problem is that the participant and non-participant groups may not be the same prior to the intervention, so the expected difference between those groups may not be due entirely to program intervention. If, in equation 2, one then adds and subtracts the expected
outcome for nonparticipants had they participated in the program $E(Y(t)(0) \mid Ti = 1)$, or another way to specify the counterfactual one gets Eq. 4 - Eq. 6.

$$D = E(Y(t)(1) \mid Ti = 1) - E(Y(t)(0) \mid Ti = 1) + E(Y(t)(0) \mid Ti = 1) - E(Y(t)(0) \mid Ti = 0) \quad (4)$$

$$D = ATE + [E(Y(t)(0) \mid Ti = 1) - E(Y(t)(0) \mid Ti = 0)] \quad (5)$$

$$D = ATE + B \text{ (selection bias)} \quad (6)$$

In Equations 4-6, ATE is $E(Y(t)(1) \mid Ti = 1) - E(Y(t)(0) \mid Ti = 1)$, that is, the average gain in outcomes of participants relative to nonparticipants, as if nonparticipating households were also treated. The ATE corresponds to a situation in which a randomly chosen household from the population is assigned to participate in the program, so participating and nonparticipating households have an equal probability of receiving the treatment T. The term B, $E(Y(t)(0) \mid Ti = 1) - E(Y(t)(0) \mid Ti = 0)$, is the extent of selection bias that crops up in using D as an estimate of the ATE. Hence $E(Y(t)(0) \mid Ti = 1)$, is unknown, the calculation of magnitude of selection bias is become difficult. As a consequence, it may impossible to know the exact difference in outcomes between the treated and the control groups. Therefore the basic objective of a sound impact assessment is then to find ways to get rid of selection bias (B = 0) or to find ways to account for it. The solution for this problem is conditional independence assumption or unconfoundeness assumption. It means assuming that whether or not households or individuals receive treatment (conditional on a set of covariates, X) were independent of the outcomes that they have. So B = 0 (selection bias is disappeared).

$$E(Y(t)(1), Y(t)(0)) \perp Ti \mid Xi \quad (7)$$

Generally in independence assumption, the participant and non-participant groups must be the same in at least three ways. The average characteristics of the participant and non-participant groups must be identical in the absence of the program, the participant should not affect the non-participant group either directly or indirectly this is called (no spillovers) and the outcomes of units in the non-participant group should change the same way as outcomes in the participant group, if both groups were given the program (or not) Khandker et al., (2010).

RESULTS AND DISCUSSION

Descriptive Statistic Results

The study uses a total of 200 sample respondent data for analysis which was collected from Yabello district. Comprising 57 participants and 143 non-participants women in Aloe Vera soap production, as indicated in Table 2.

Sex of the household head: According to the survey result obtained in the study area, 42 households are headed by female and the remains 158 households are male headed. Result of chi²-test found that the difference is statistical insignificant. This means participation is not different by which sex is headship of the family (Table 3).

Trade experience of respondent women: according to survey result obtained in the study area, 23.5 percent of respondent women had an experience of trading in different small business sectors. According to their response, the livelihood of study area community is depends on livestock, crops and livestock products whereas the crops and livestock products such as milk and butter are sold by women. From the comparison views between participant and non-participant women, there is higher frequency of trader in non-participant (29) than participant women (18). This difference is statistical positive and significant at 10% level.

<table>
<thead>
<tr>
<th>Table 1: Definition of hypotheses variables and expected sign/s</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable code</strong></td>
</tr>
<tr>
<td>COOP</td>
</tr>
<tr>
<td>LAND</td>
</tr>
<tr>
<td>Sexhhh</td>
</tr>
<tr>
<td>EDURW</td>
</tr>
<tr>
<td>AGERW</td>
</tr>
<tr>
<td>FAMILYSIZE</td>
</tr>
<tr>
<td>Training</td>
</tr>
<tr>
<td>TRW</td>
</tr>
<tr>
<td>CREDIT</td>
</tr>
<tr>
<td>EXTN</td>
</tr>
<tr>
<td>TLU</td>
</tr>
<tr>
<td>DROUGHT</td>
</tr>
<tr>
<td>AGEMARRIED</td>
</tr>
<tr>
<td>Labourforce</td>
</tr>
<tr>
<td>DISTANCE</td>
</tr>
<tr>
<td>FASS:</td>
</tr>
</tbody>
</table>
Table 2: Total sample of respondent

<table>
<thead>
<tr>
<th>Variables</th>
<th>Freq.</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-participant</td>
<td>143</td>
<td>71.50</td>
</tr>
<tr>
<td>Participant</td>
<td>57</td>
<td>28.50</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Source: Own survey data, 2018

As Table (3) indicated, 47 percent of respondent women had extension contact per month and 0.5 percent of them had extension contact when asked. According to respondent reaction, agricultural development agents are advising them on agricultural production whereas the issue about resource management and how to work in cooperatives are not yet reached most of respondent by GOs. But, some NGOs are initiated to give them training on this concern. Majority of respondent had also credit services from different organization starting from their own women cooperative groups.

Moreover, the chi²-test result from Table (3) shows extension contact and membership in other cooperative of respondent women is positive and significant at 5% and 1% level respectively. But, rest of the variables (credit, drought and Father’s asset) are with chi²-test of insignificant result even though there were some figurative differences among participant and non-participant women are seen.

Table (4) shows minimum, maximum, mean and std. deviation of age of respondent’s women are 17, 80, 39 and 14.083 years respectively. This table also reflected that mean age of participant women are 44.87 years and that of non-participants are 36.65 years. Which is also statistical negative and significant at 1% level as a witness of age difference between participant and non-participant. This result also tallied with the study hypothesis which stated that age is one of factor which increases women participation.

The minimum, maximum and mean age at which respondent women got marriage is 5, 30 and 16-17 years respectively. As one can see from Table (4), there is no that much difference between this two groups on their age of getting married both are on average at their 16th years. Statistical t-test for this variable also shows no significant difference.

The educational level: Results of education show that on average respondent women are below grade one which means high illiteracy rate. Statistical t-test is also revealed insignificant result.

The minimum, maximum, mean and std. deviation of family size are 2, 12, 6-7 and 2-3 persons per household respectively. Mean family sizes of participant women (7.94) are greater than that of non-participant women (6.24). This difference is statistical negative and significant at 1% level. This result is also tallied with the study hypothesis which is stated that child is one of a driven force that pushes women towards productive works for their family wellbeing.
According to the Table (4), the minimum, maximum and mean of labour force person per household is 0, 6, and 2 respectively. The mean labour force of both family of participant and non-participant women are approximately 2 person which is too much less than mean family size (6-7). Statistical t-test for labour force is also revealed negative and significant result at 1% level. This result also tallied with the study hypothesis.

According to UN (2017) on average family size in Ethiopia is 4.6 person per household. They also said that “Small average household sizes, fewer than three persons per household are concentrated in Europe and Northern America, whereas large average household sizes, five or more persons per household are observed across much of Africa and the Middle East”. Thus, the family size of the respondent in both cases is that of Ethiopian national level and within the range of most African and Middle East countries.

Table (4) also shows the minimum, maximum, mean and str. deviation of average estimated household income per month are 200, 5000, 961.50 and 650.3 Birr respectively. But, mean monthly income of participants and non-participants women household is 1019.3 and 938.46 Birr respectively. It show that there is some what difference on mean monthly income earning of participants and non-participant in which the participant’s mean income is greater than that of non-participants’ households even though this difference is statistical insignificant.

The possible reason for this difference is that; according to the respondents’ responses of participant women they were benefited for being participant of this production, they are always gaining the acting of learning and exercising experience from different organization such as local, national and international NGOs and GOs. Beside this, the participant themselves formed different form of cooperatives like butter, mirt stove, milk cooperative and also joining into other formed cooperatives. This all training and learning process are gradually developed among participant as they come together as co-workers. In monetary benefit, participants also had a lot of saved deposited money on account and some capital asset in form of cooperative.

But, according to participant women responses their problem was: there is no dividend sharing to members that cause them disincentive to work. And another problem that they were raised as hindering factor for this production was problem of input supplied especially caustic soda and cooking oil. Many women said that, if they were getting these ingredients (inputs) individual, they will produce soap as individual business.

The minimum, maximum and mean total tropical livestock units are 0.3, 121.05 and 14.25 respectively. The difference in this characteristic is also similar to household monthly income. There is gap between minimum and maximum which explained in std. deviation (15.09) of mean (14.25) total tropical livestock unit. As result indicated the mean and standard deviation of participant women household livestock holding is smaller than that of non-participant women household. This result is statistical positive and significant at 10% level. This result also supports the study hypothesis.

The minimum, maximum and mean distribution of land is 0.0, 5 and 1.1087 hectares respectively. The mean land holding of participant (1.09 hectares) is below that of non-participant (1.115 hectares) but this difference is statistical insignificant. The mean land holding of total respondent is below that of national level. As national level land use survey shows the average household farm size in Ethiopia is 1.37 hectares, but its varies by place of residence and the gender of the household head (CSA, 2013).

The mean distance of respondent from the nearest market is 2hrs on foot. This result is also positive and

---

**Table 4: Descriptive Statistics for continuous Variables**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Participant Mean (std. d.)</th>
<th>Nonparticipant Mean (std. d.)</th>
<th>Total Mean (std. d.)</th>
<th>Min</th>
<th>Max</th>
<th>Mean (std. d.)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGERW</td>
<td>(11.51)</td>
<td>(1.35)</td>
<td>17</td>
<td>80</td>
<td></td>
<td>39.00</td>
<td>-3.85***</td>
</tr>
<tr>
<td>AGEMARRIED</td>
<td>(2.42)</td>
<td>(0.04)</td>
<td>5</td>
<td>30</td>
<td></td>
<td>16.41</td>
<td>-1.43</td>
</tr>
<tr>
<td>EDURW</td>
<td>(0.47)</td>
<td>(0.16)</td>
<td>0</td>
<td>5</td>
<td></td>
<td>0.14</td>
<td>0.64</td>
</tr>
<tr>
<td>TLU</td>
<td>(9.85)</td>
<td>(16.60)</td>
<td>.3</td>
<td>121.05</td>
<td></td>
<td>14.25</td>
<td>1.8*</td>
</tr>
<tr>
<td>Family size</td>
<td>(2.22)</td>
<td>(2.51)</td>
<td>2</td>
<td>12</td>
<td></td>
<td>6.73</td>
<td>-4.48***</td>
</tr>
<tr>
<td>Labourforce</td>
<td>(1.02)</td>
<td>(0.56)</td>
<td>0</td>
<td>6</td>
<td></td>
<td>2.00</td>
<td>-3.03***</td>
</tr>
<tr>
<td>Monthly income</td>
<td>(1019.3)</td>
<td>(645.45)</td>
<td>200</td>
<td>5000</td>
<td></td>
<td>961.50</td>
<td>-0.79</td>
</tr>
<tr>
<td>Land</td>
<td>(1.09)</td>
<td>(1.115)</td>
<td>0</td>
<td>5</td>
<td></td>
<td>1.108</td>
<td>0.20</td>
</tr>
<tr>
<td>Distance</td>
<td>(.94)</td>
<td>(.80)</td>
<td>0</td>
<td>4</td>
<td></td>
<td>2.007</td>
<td>4.1***</td>
</tr>
</tbody>
</table>

Source: own survey data, 2018
significant at 1% level where on average it take less hours for participants women to go a nearest market. But standard deviation result revealed that there is somewhat higher variation within participant (.94) than non-participant (.80). This means, compared to non-participant, not all participants were nearby market. It is tallied with the study hypothesis.

**Propensity Score Matching Model Results**

The “sum myscore” command was used to check/summarize the propensity score (Table 5).

As propensity score is a probability, it has to be in the interval of [0:1]. Hence, the average probability of respondent women to participate in the Aloe Vera soap production for all the individuals is 28.5%.

**Check Range of Common Support**

Psgraph test for the common support was used to check the extent to which distributions of propensity scores in treatment (participants) and comparison (non-participants) groups are overlapped from logistic regression of propensity score matching model.

Note: the common support option has been selected. The region of common support is [.00411911, 0.9775877]. This assumption of common support was also checked graphical as following; if an assumption of common support holds, there must be an overlap of the propensity scores of the participants and non-participants.

That’s why the Figure (1) depicted as three colour of blue, red and green. The blue and red colours are on common support region, but green colour indicated the propensity score out of common support regions. In each class of the propensity score there is a certain number of non-treated and treated individuals as well.

**Check Balancing Property**

Step 1: Identification of the optimal number of blocks. The final number of blocks is 5. This number of blocks ensures that the mean propensity score is different for treated (participants) and controls (non-participants) in each blocks. The Stata was used t-tests to determine whether each covariate is balanced within each block.

Step 2: Test of balancing property of the propensity score across groups. It should be needed to check balancing before trusting the ATT estimation. This was the test of whether mean of propensity score is equal in treatment and comparison groups within each quintile. Therefore, the Stata result of both steps revealed that the balancing property is satisfied in each blocks for each covariates.

Table (6) shows the inferior bound, the number of treated and the number of controls for each block under the assumption of common support. And in this case, total 190 respondents are on common support of which 133 are non-participant women and 57 are participant women.

Figure (2) shows the distribution of the all household with respect to their estimated propensity score. Accordingly, most of the treatment households are found partially in the middle and partially in the right side of the distribution. On the other hand, most of the control households are partially found in the center and partially in the left side of the distribution.

Of course, two conditions were identified and as well satisfied, the success of the matching for each of the independent variables was also tested from the matching algorithm of propensity score after the choice of matching algorithm estimator.

**Choice of Matching Algorithm**

Note: There is no universal best strategy among matching algorithm of propensity score matching but the focus should be given on the trade-off between bias and variance/efficiency.

For instance, According to Caliendo and Kopeinig (2005) on their review of “practical guidance for the implementation of propensity score matching”, one way of assessing the quality of matching is t-test of standard bias for all covariates as all covariates should be balanced. In this case of study on hand, as indicated in Tables (7 and 8), the test revealed that all covariates are balanced after matching. Again, Caliendo and Kopeinig (2005) offered that the joint significance test was found another method to assess matching quality. And they elaborated this joint significant as this; the best matching result was found in such away as before matching there might be significant difference in all covariates, and should be the insignificant difference result existed after matching or for matched sample case.

### Table 5: Summary of pscore

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total household</td>
<td>200</td>
<td>.285</td>
<td>.3229254</td>
<td>3.36e-06</td>
<td>.9857269</td>
</tr>
<tr>
<td>non-participant</td>
<td>143</td>
<td>.1371315</td>
<td>.2047364</td>
<td>3.36e-06</td>
<td>.9857269</td>
</tr>
<tr>
<td>Participant</td>
<td>57</td>
<td>.6559683</td>
<td>.2614782</td>
<td>.0041191</td>
<td>.9777588</td>
</tr>
</tbody>
</table>

Source: own survey result, 2018

### Table 6: Inferior block of propensity score

<table>
<thead>
<tr>
<th>Inferior of block of pscore</th>
<th>PARTALS</th>
<th>Participant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-participant</td>
<td>104</td>
<td>4</td>
<td>108</td>
</tr>
<tr>
<td>.2</td>
<td>15</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>.4</td>
<td>8</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>.6</td>
<td>2</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>.8</td>
<td>4</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td>133</td>
<td>57</td>
<td>190</td>
</tr>
</tbody>
</table>

Source: Model result, 2018
In line with this connotation, the joint significant test of different matching algorithms for this study was taken as this and the result was found, in Table (7), that all used algorithm satisfied to some extent the test condition even though the difference was found in the level of percentage significance. Accordingly, the result of kernel band width (0.25) has the lowest Pseudo-$R^2$ and all covariates are balance after matching for matched sample among all algorithm. Similarly, (Dehejia and Wahba, 2002) noted that, the choice of a matching estimator was based on different criteria like equal means test (balancing test of covariates), lowest pseudo-$R^2$ and largest numbers of matched sample size.

Therefore, the matching algorithm with highest sample size matched, insignificant matched sample or equal means and with the lowest Pseudo-$R^2$ value for matched sample, among the other used matching algorithm for this study is kernel band width (0.25) as indicated in Table (7).

Covariate Imbalance Tests (before and after matching) and Graphing

According to the summary of (Leuven and Sianesi, 2003) on psmatch2 module; the main focus of this test was; for each variable t-test the equality of means before and after matching, standardized % bias before and after, and achieving % reduction in |bias| is found to be important. And as long as overall significance concerned; pseudo $R^2$ from logistic of treatment on covariates before matching and on matched samples, p-values of the likelihood-ratio test of joint insignificance of covariate before and after matching and summary indicators of the distribution of |bias| before and after are critically issues. Accordingly, for these point illustrations, ptest was applied and discussed for this study as follows:

The Table (8) show that a t-test on the hypothesis that the mean value of each variable is the same in the treatment group and the non-treatment group. It was done before and after matching.

Moreover, a bias before and after matching was calculated for each variable and the change/reduction in this bias was indicated. In this table, one can see the difference of the values of the exogenous variables between the two groups before matching.

With matching, all covariates are shown us, no significance mean difference after matching as it indicated by the p-value of test in Table (8). Caliendo and
**Kopeinig (2005)** “before matching differences are expected, but after matching the covariates should be balanced in both groups and hence no significant differences should be found”.

Also standard bias reduction result in Table (8) revealed that the differences between treatment group and non-treatment group are reduced for many variables where exception was found in the variables such as training, sexhhh (sex of household head); family size and AGEMARRIED (age at first married). Even though in the case of these variables the difference between two groups were not reduced, but these variables are statistical significant effect on participation as it revealed.

Therefore, pstest indicated all variables are satisfied the insignificance test of after matching which means there is no mean difference after matching for each variables are balanced.

The Table (9) description was about the joint significance, taking together all predictors variable, that explained by Pseudo-$R^2$ and $p$>chi2. The pseudo-$R^2$ indicates how well the regressors $X$s explain the participation probability. As it was explained, after matching there should be no systematic divergences in the distribution of covariates between both groups. This means for the joint significance; $p$>chi2 value before matching might be statistical significant, and $p$>chi2 value after matching (for matched sample) should be statistical insignificant.

Accordingly, test in Table (9) was also illustrated that ($p$>chi2 is 0.000 for unmatched sample) so there is not to be rejection before matching, and ($p$>chi2 is 0.937 for matched) there is to be highly rejection after matching which was an expected result of this regression.

Figure (3) portrays graphically the distribution of mean score of each of explanatory variable for the participant and non-participants of unmatched and matched. And it shows the standardized % bias across covariates.

**Table 7: Comparison of different matching algorithm estimators**

<table>
<thead>
<tr>
<th>Matching estimator</th>
<th>Nearest neighbour</th>
<th>Balancing test*</th>
<th>PS $R^2$</th>
<th>LR $\chi^2$</th>
<th>$\text{Prob}&gt;\chi^2$</th>
<th>Matched N</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN(1)</td>
<td>16</td>
<td>0.101</td>
<td>15.92</td>
<td>0.459</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>NN(2)</td>
<td>16</td>
<td>0.069</td>
<td>10.98</td>
<td>0.811</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>NN(3)</td>
<td>15</td>
<td>0.070</td>
<td>11.04</td>
<td>0.807</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>NN(4)</td>
<td>16</td>
<td>0.074</td>
<td>11.62</td>
<td>0.770</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>NN(5)</td>
<td>16</td>
<td>0.068</td>
<td>10.73</td>
<td>0.826</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Radius caliper</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>16</td>
<td>0.055</td>
<td>8.69</td>
<td>0.925</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>16</td>
<td>0.102</td>
<td>8.52</td>
<td>0.932</td>
<td>173</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>16</td>
<td>0.061</td>
<td>9.67</td>
<td>0.883</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>16</td>
<td>0.110</td>
<td>15.27</td>
<td>0.644</td>
<td>193</td>
<td></td>
</tr>
<tr>
<td>Kernel matching</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band width 0.1</td>
<td>16</td>
<td>0.068</td>
<td>10.80</td>
<td>0.822</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Band width 0.01</td>
<td>16</td>
<td>0.085</td>
<td>7.09</td>
<td>0.972</td>
<td>173</td>
<td></td>
</tr>
<tr>
<td>Band width 0.25</td>
<td>16</td>
<td>0.053</td>
<td>8.38</td>
<td>0.937</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Band width 0.5</td>
<td>15</td>
<td>0.083</td>
<td>13.08</td>
<td>0.667</td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

Source: Model result, 2018

* Number of explanatory variables with no statistically significant mean differences between the matched samples.
### Table 8: ptest table

| Variables        | Unmatched M | Treated | Control | %reduct | %bias | t-test | p>|t| | V(T)/ V(C) |
|------------------|-------------|---------|---------|---------|-------|--------|-----|---------|
| Sexhhh           | U           | .73684  | .81119  | -17.7   | -1.16 | 0.246  | .   |         |
|                  | M           | .73684  | .71208  | 5.9     | 0.29  | 0.770  | .   |         |
| AGERW            | U           | 44.877  | 36.65   | 63.2    | 3.86  | 0.000  | 0.64|         |
|                  | M           | 44.877  | 42.39   | 19.1    | 69.8  | 1.23   | 1.34|         |
| TRAINING         | U           | 3       | 1.8112  | 90.0    | 5.97  | 0.000  | 1.40|         |
|                  | M           | 3       | 2.8041  | 14.8    | 8.35  | 0.80   | 1.46|         |
| AGEMARRIED       | U           | 16.877  | 16.231  | 23.5    | 1.43  | 0.153  | 0.63|         |
|                  | M           | 16.877  | 16.63   | 9.0     | 61.8  | 0.53   | 0.599| 0.87    |
| FAMILYSIZE       | U           | 7.9474  | 6.2378  | 72.0    | 4.48  | 0.000  | 0.78|         |
|                  | M           | 7.9474  | 8.0991  | -6.4    | 91.1  | -0.33  | 0.741| 0.71    |
| TLU              | U           | 11.204  | 15.47   | -31.2   | -1.81 | 0.071  | 0.35*|         |
|                  | M           | 11.204  | 11.856  | -4.8    | 84.7  | -0.36  | 0.721| 1.06    |
| LAND             | U           | 1.0921  | 1.1554  | -3.3    | -0.21 | 0.835  | 0.94 |         |
|                  | M           | 1.0921  | 0.7823  | 16.1    | -389.2| 0.99   | 0.326| 1.75*   |
| Labourforce      | U           | 2.2456  | 1.9021  | 41.7    | 3.04  | 0.003  | 3.33*|         |
|                  | M           | 2.2456  | 2.1187  | 15.4    | 63.1  | 0.85   | 0.397| 4.69*   |
| TRW              | U           | 1.6842  | 1.7972  | -25.8   | -1.71 | 0.090  | 1.35 |         |
|                  | M           | 1.6842  | 1.7318  | -10.9   | 57.9  | -0.55  | 0.580| 1.10    |
| EXTN             | U           | 2.2456  | 2.6573  | -40.5   | -2.65 | 0.009  | 1.26 |         |
|                  | M           | 2.2456  | 2.3087  | -6.2    | 84.7  | -0.32  | 0.752| 1.04    |
| CREDIT           | U           | 1.6667  | 1.6783  | -2.5    | -0.16 | 0.875  | 1.03 |         |
|                  | M           | 1.6667  | 1.6497  | 3.6    | -45.2 | 0.19   | 0.851| 0.98    |
| DROUGHT          | U           | 2.8947  | 2.9161  | -7.3    | -0.47 | 0.634  | 1.24 |         |
|                  | M           | 2.8947  | 2.8285  | 22.5    | -210.1| 1.02   | 0.310| 0.66    |
| COOP             | U           | 1.193   | 1.5524  | -79.6   | -4.86 | 0.000  | 0.64 |         |
|                  | M           | 1.193   | 1.2341  | -9.1    | 88.6  | -0.53  | 0.596| 0.87    |
| DISTANCE         | U           | 1.6132  | 2.164   | -63.0   | -4.17 | 0.000  | 1.38 |         |
|                  | M           | 1.6132  | 1.6205  | 0.8    | 98.7  | -0.04  | 0.967| 1.00    |
| EDUCRW           | U           | 0.08772 | 0.10684 | -1.10   | -0.64 | 0.524  | 0.34*|         |
|                  | M           | 0.08772 | 0.04868 | 5.9    | 46.6  | 0.48   | 0.635| 1.40    |
| FASS             | U           | 2.0175  | 1.9371  | 12.6    | 0.79  | 0.430  | 0.85 |         |
|                  | M           | 2.0175  | 2.0235  | -0.9    | 92.6  | -0.05  | 0.961| 0.75    |

Source: Model result, 2018

* If variance ratio outside [0.59; 1.70] for U and [0.59; 1.70] for M

### Table 9: Joint significance test of psmatch2

| Sample   | Ps | R2 | LR chi2 | p>|ch2 | MeanBias | MedBias |
|----------|----|----|---------|------|----------|---------|
| Unmatched| 0.448 | 107.10 | 0.000 | 36.6 | 28.5 |
| Matched  | 0.053 | 8.38 | 0.937 | 9.5  | 7.7   |

Source: Model result, 2018

### Figure 3: ptest graph

Source: own survey result, 2018
Table 10: ATT of participation of the women on income and CEI

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample</th>
<th>Treated</th>
<th>Controls</th>
<th>Difference</th>
<th>T-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>MONTHLYINCOME</td>
<td>Unmatched</td>
<td>1019.3</td>
<td>938.462</td>
<td>80.837</td>
<td>0.79</td>
</tr>
<tr>
<td>ATT</td>
<td>1019.3</td>
<td>973.605</td>
<td>45.693</td>
<td></td>
<td>0.27</td>
</tr>
</tbody>
</table>

Note: S.E. does not take into account that the propensity score is estimated. Source: model result, 2018

Table 11: Bootstrap standard error

| Observed Coef. | Bootstrap Std. Err. | Z  | P>|z| | Normal-based [95% Conf. Interval] |
|----------------|---------------------|----|------|-----------------------------|
| bs_1           | -10.61338           | 139.9015 | -0.08 | 0.940 | -284.8153 | 263.5886 |

Source: model result, 2018

Table 12: Rosenbaum bounds for outcome variable (N = 200 matched pairs)

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>Gamma</th>
<th>sig+</th>
<th>sig-</th>
<th>t-hat+</th>
<th>t-hat-</th>
<th>CI+</th>
<th>CI-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly income</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>850</td>
<td>850</td>
<td>800</td>
<td>900</td>
</tr>
<tr>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>800</td>
<td>900</td>
<td>800</td>
<td>900</td>
</tr>
<tr>
<td>1.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>800</td>
<td>900</td>
<td>800</td>
<td>900</td>
</tr>
<tr>
<td>1.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>800</td>
<td>900</td>
<td>800</td>
<td>900</td>
</tr>
<tr>
<td>1.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>800</td>
<td>900</td>
<td>750</td>
<td>1000</td>
</tr>
<tr>
<td>1.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>800</td>
<td>900</td>
<td>700</td>
<td>1000</td>
</tr>
</tbody>
</table>

* gamma - log odds of differential assignment due to unobserved factors
  sig+ - upper bound significance level
  sig- - lower bound significance level
  t-hat+ - upper bound Hodges-Lehmann point estimate
  t-hat- - lower bound Hodges-Lehmann point estimate
  CI+ - upper bound confidence interval (a= .95)
  CI- - lower bound confidence interval (a= .95)

Robust test

The S.E. does not take into account that the propensity score is estimated (Table 10). This is because the estimated variance of the treatment effect also include the variance due to the estimation of the propensity score, the imputation of the common support, and possibly also the order in which treated individuals are matched. Caliendo and Kopeinig (2005), thus, one way to deal with this problem is to use bootstrapping. According to Schmidheiny (2016) the bootstrap takes the sample (the values of the independent and dependent variables) as the population and the estimates of the sample as true values.

Sensitivity Analysis of Psmore Matching

As long the PSM method are concerned for the impact analysis, it might needed to analysis sensitivity of ATT estimation to any unobserved covariates that might introduce the hidden /endogeneity bias. Hence, as the Table (12) indicated the Rosenbaum bounds test was applied to evaluate the sensitivity of how the changing values of a parameter gamma, $\Gamma$, would influence the significance of the results obtained from the matching analysis. According to many literature like If $\Gamma = 1$, then, hidden bias is zero.

The result in Table (12) revealed that for all chosen gamma level (1, 1.1, 1.2, 1.3, 1.4 and 1.5), the upper and lower bound significance level, upper and lower bound of Hodges-Lehmann point estimate and upper and lower bound of confidence interval for outcome variable. The result of upper and lower bound significance level is significant for outcome variable. The result of upper bound Hodges-Lehmann point and confidence interval is decreasing and the lower bound in both cases is increasing.

This witnessed that, the computed ATTs are relatively insensitive to unobservable covariates that might introduce as hidden bias.

CONCLUSION AND RECOMMENDATIONS

The study concluding and recommending it’s finding as follows; the first major problem in production of Aloe Vera soap was input (caustic soda and cooking oil) availability since production is impossible without these inputs. First, caustic soda is only found at national level market which is also just by order, and there is highly problem with market supply and market price of cooking oil. In order to skip this problem of inputs so far aloe soap producers have been organized under milk union and then it facilitate inputs supply and product market of those primary cooperative. District trade office also sometime with very little attention gives them a few litre of oil to those cooperative which is not that much interesting and initiating them for higher level production. Therefore, based on the income impact of this production, the NGOs, government office such as cooperative, trade office, and management at zone and district level should aware of this opportunity and facilitate the input supply (caustic soda and cooking oil) for those primary cooperative, and also for any woman demanding individual in production, and letting this production to go beyond cooperative to individual level.

Overall according to the impact result of propensity score matching of kernel band width (0.25) result revealed, participation of women in Aloe Vera soap has a positive impacts on the household income even though the result is
statistical insignificant. Therefore, non-participant women should be encouraged to have an opportunity to participate more in such productive activities. But, the insignificant result might be due to the study estimation method and data, the study also recommend other researcher to conduct further research taking this study as a base line. Remind, from theory of population growth, Esther Boserup was quoted as “more people there are, the more hands there are to work”. Therefore, so is women role for total wellbeing of community at large.

REFERENCES


LEUVEN, E. & SIANESI, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing and covariate imbalance testing.


