Chapter IV

Methodology

This section includes information about the study area, sampling procedure, sample size, types and source of data, method of data collection and data analysis.

4.1 Description of Study Area

4.1.1 Ethiopia-An Overview

Ethiopia is the country situated in East Africa at 9°1.8′N 38°44.4′E. Its total area is about 1,100,000 square kilometers (420,000 sq. miles). It is bordered by Eritrea to the north and northeast, Djibouti and Somalia to the east, Sudan and South Sudan to the west, and Kenya to the south. Its capital city is Addis Ababa. It is one among the fast growing countries in the world. The population of Ethiopia is estimated at 96,506,031 as of July 1, 2014. Ethiopia's population is equivalent to 1.3% of the total world population. Ethiopia ranks number 13 in the list of countries by population. The population density in Ethiopia is 87 people per Km². Eighteen per cent of the population are residing in urban (17,172,948 people in 2014). The median age in Ethiopia is 18.4 years.

The country's population is highly diverse, containing over 80 different ethnic groups. Most people in Ethiopia speak Afro-Asiatic languages, mainly of the Semitic or Cushitic branches. The latter include the Oromo, Amhara, Tigray and Somali, which together make up three-quarters of the population. Nilo-Saharan-speaking Nilotic ethnic minorities also inhabit the southern regions of the country, particularly in areas bordering South Sudan. As many as 90

individual languages are spoken in Ethiopia. Amharic speakers account for 25.7 per cent (as a first language), Afaan Oromo 31.6 per cent, Somali 6.5 per cent, Tigrinya 6.1 per cent, Guragigna 3.5 per cent, Sidamo 3.5 per cent, other local languages. English and arabic are major foreign languages taught in schools. Amharic is the only official national language. Afaan Oromo enjoys almost equal number of native speakers. It is also the language of primary school instruction. English is the most widely spoken foreign language and is taught in all secondary schools.

Ethiopia is a multi-religious country. Most of the Christians live in the highlands, while the Muslims mainly inhabit the lowlands. Adherents of traditional faiths are primarily concentrated in the southern regions. Ethiopian Orthodox accounts 43.5 per cent, Protestant 18.6 per cent (which include P'ent'ay and the Ethiopian Evangelical Church Mekane Yesus), Muslim 33.9 per cent, traditional 2.6 per cent, Catholic 0.7 percent, all others 0.6 per cent. Ethiopian Jewish communities also reside in the northern parts of the country although almost all of them have emigrated to Israel. The country is the site of the first Hijra in Islamic history and the oldest Muslim settlement in Africa at Negash.

Ethiopia is one of the oldest locations of human life known to scientists and is widely considered the region from which Homo sapiens first set out for the Middle East and points beyond. Ethiopia was a monarchy for most of its history. The Kingdom of Aksum was one of the great world powers of the 3rd century. In the 4th century, it was the first major empire in the world to officially adopt Christianity as a state religion.
During the late 19th-century Scramble for Africa, Ethiopia was the only African country to defeat a European colonial power and retain its sovereignty as an independent country. It was the first independent African member of the 20th-century League of Nations and the UN. When other African nations gained their independence following World War II, many of them adopted the colors of Ethiopia's flag. In 1974, at the end of Haile Selassie I's reign, power fell to a communist military junta known as the Derg, backed by the Soviet Union, until it was defeated by the EPRDF, which has ruled since about the time of the collapse of the USSR in 1991.

The origin of the coffee bean, Ethiopia is a land of natural contrasts; with its vast fertile West, jungles & numerous rivers, the World's hottest settlement in its north, Africa's largest continuous mountain ranges and the largest cave in Africa at Sof Omar. Ethiopia has the most UNESCO World Heritage Sites in Africa. Ethiopia's ancient Ge'ez script, also known as Ethiopic, is one of the oldest alphabets still in use in the world. It is the main source of the Nile, the longest river on the earth.

The country’s real GDP growth rate (by 2012) was estimated to be seven per cent. Agriculture contributed 46.4 per cent; service contributed 43 percent while industry’s share was only 10.7 per cent of the GDP.

4.1.2 Oromia at Glance\textsuperscript{5}

\textit{Oromia} is one among the 9 National Regional States and two City Administrations Towns of Ethiopia. Its average estimated area is about 363,375 km\textsuperscript{2}, accounting for about 34.3\% of the country’s total area. \textit{Oromia} is the largest National Regional State in Ethiopia in terms of population size and areal

\textsuperscript{5} Available at http://www.ethiodemographyandhealth.org/Oromia.html date accessed 16/12/2013
coverage. It has a total border length of about 5700 km (about 600 km international borderlines with Kenya (521km) and Sudan (66 km), and about 5100 km with National Regional States (1860 km with SNNPSE, 1410 km with Somali N.R.S., 706 km with Amhara N.R.S., 690 km with Benishangul, 255 km with Gambela N.R.S. and 164 Km with Afar N.R.S.). The Region occupies central position of the country. Oromia shares common boundaries with all national regional states except the National Regional State of Tigray.

Oromia is characterized by diverse relief features, and it is a Region of great geographic diversity with altitude of ranging from less than 500m to over 4300m above sea level. Oromia is also a region of great physiographic diversity, having Mountain ranges and massifs that culminate into mountain Tullu Dimtu that is found in Bale zone. Tullu Dimitu is the highest peak of the Region with 4377m above sea level (a.s.l.). About 50 per cent of the Region's land surface is above 1500m asl (i.e. Highlands), while the lowlands and Rift Valley systems also account for the remaining 50 per cent of the Region's total land surface.

Oromia has diverse climatic condition. Average annual rainfall of the Region varies from 400mm to over 2400mm. The Region has summer, autumn, winter and spring seasons. Average annual temperature ranges from less than 7.5 °C to greater than 27.5 °C in the Region. Oromia is endowed with huge natural resource base. These include resources such as surface water, energy, mineral, wildlife, tourism, arable land, agricultural resource bases and historical and cultural heritages as well.

Oromia has fertile soil, diverse geographical setting and varied agro-climatic zones that are suitable for the cultivation of different food and cash crops. The Region has four climatic zones namely, Tropical zone that accounts
50.8 per cent, Sub-tropical 42.2 per cent, Temperate 7.5 per cent and Wurch 0.4 per cent. It has four growing seasons, namely summer, autumn, winter & spring. Summer and winter are the major seasons of the region.

Out of its total area, cultivable land accounts about 28 per cent of its total area. Natural forest accounts nine per cent; wood, shrub, bush, grassland, and plantation cover about 60 per cent of its total area, while and swamp, water courses, rocks and urban centers account only three per cent of its total area. The Region has 12 major lakes, covering an area of 3135 km\(^2\), of which some of them are currently used for small scale fish production. The Region has an estimated potentially irrigable land of about 1.7million hectares. Oromia is the largest contributor to Ethiopia's immense hydropower potential.

4.1.3 Description of Sample Districts

The study has been carried out in two districts in Ethiopia namely Toke Kutaye and Dendi Districts. Toke Kutaye is located in West Shewa Zone of Oromia Regional state in Ethiopia. Toke Kutaye is bordered on the east by the Ambo Zuria, on the north by Midakegn, on the west by Cheliya. This District is found at about 12 Km from capital town of West Shewa Zone, Ambo. Administratively the District is classified into 31 rural and 4 urban kebeles\(^6\). The town of the District is Guder. The projected total population of the district for the year 2014 was 131,086, of which women account 51 per cent. Besides, about 87 per cent of the population lived in rural area. The District contains three agro ecological zones, namely highland which accounts (27%), midland (55%) and lowland (18%). The minimum rainfall is 800mm whereas the maximum is 1000mm. Three major soil types are dominantly found in the District. They are black or vertisol which account (27 %), red soil (48 %) and mixed soil account (25%) of the total area. With regard to land use and land cover, the majority of

\(^6\) Smallest administrative unit
the area (57.6%) is potentially arable land while 14.7 per cent is grazing land. Forest and woodland cover only 0.37 per cent of the total area of the District. Shrubs account 1.37 per cent whereas 10.86 per cent of the area is allocated for industrial and related use. The remaining 15.10 per cent is allocated for other uses.

Similarly, *Dendi* is one among the Districts of *Oromia* Regional state in Ethiopia. Distance of the District capital (Ginchi) from zonal capital (Ambo) is 35 Km while distance of the District capital from regional capital (Finfinnee) is 77 Km. *Dendi* is bordered on the south by the Debub Mirab (Southwest) Shewa Zone, on the west by Ambo, on the north by Jeldu, and on the east by Ejerie Districts. Administratively the District is divided into 48 peasant associations and 6 urban *kebeles*. Ginchi, Olenkomi, Asgori and Gaba Dilbata are the towns found in the District. The administrative center of this District is Ginchi. The highest point, Mount *Dendi* (3260 meters above sea level), is situated in this District. Notable landscape feature known as Chilimo forest which is a remnant of the dry afro-mountain forest on the Ethiopian Central Plateau with wooded area covering 2400 hectares in size is found in this District. Rainfall pattern is bimodal with the main rainy season from June to September and short rainy season from February to March. On-farm activities are the main economic activities in the District with very little number of farmers working in off-farm activities. *Dendi* has an estimated population density of 165.2 people per square kilometer. The two largest ethnic groups in the District are the Oromo (91.97%) and the Amhara (6.9%); all other ethnic groups constitute 1.12 per cent of the population. Oromiffa is spoken by 93.17 per cent, and 6.37 per cent of the population speak Amharic; the remaining 0.46 per cent speak all other primary languages. The study area is shown in Figure 4.
4.1.4 Description of Sample Cooperatives and Sample Households

The study has been made in two Cooperatives namely Olonkomy and Mutulu Agricultural Multipurpose Cooperatives. Olonkomy is established in 1969 while Mutulu is established in 1968. Olonkomy has a total member of 2405, while Mutulu has 1882 members. The two Cooperatives constitute a total of 4287 members. They account 12 per cent of the total members and 6.7 per cent of total rural households in the two sampled districts (Annex 11).
Chapter IV: Methodology

The majority of the sampled households are smallholder farmers with an average farm size of 2.14 hectare. The sampled households allocated on average 1.69 hectare of land for crop production. All are engaged in both crop production and livestock rearing. The average production per annum in the sampled area is 18.37 quintals. The mean consumption expenditure in the two sample cooperatives is 2,334 ETB. Regarding level of education, 77 per cent of household respondents are literate. On average 5 members in a family are dependent. The sampled households in the two Cooperatives are homogeneous in terms of ethnic group, socio-economic condition, culture and language. Agro-ecologically they are also similar (Annex 14).

4.2 Sampling Design

4.2.1 Sampling Procedure

The study has been carried out in Oromia Regional State in Ethiopia. Oromia Regional State has been selected purposively from among the nine Regional States and the two administrative cities in Ethiopia since it registers larger number of Cooperatives (29%) with larger membership size (33%) and huge capital share as compared to other regions (Annex 9). Agricultural Cooperatives have been chosen as a study unit because they account 26 per cent of cooperatives in Ethiopia, they distribute 90 per cent of agricultural inputs, commercialize 10 per cent of marketable surplus, and they cover 50 per cent of the kebeles in Ethiopia.

The two Districts namely Toke Kutaye and Dendi have been selected as a sample purposively on the basis of accessibility, availability of reliable data, population size, and experience. Similarly, the two Agricultural Cooperatives have been selected purposively as sample Cooperatives from the two Districts on
the basis of the same criteria. Then, sample households have been disaggregated into members and non-members in the respective kebeles.

4.2.2 Sampling Frame and Techniques

The sampling frame has been the records of the lists of names of households in the kebele that are members and non-members of Cooperatives. The respective sample has been taken from each study area on the basis of proportion to size probability sampling technique.

Finally, sample respondents or households from both members and non-members have been chosen from the sample frame using simple random sampling technique and a random number table as a tool.

4.2.3 Sample Size

The study made by Bernard et al. (2009) indicated a rather limited coverage of Cooperatives. As of 2005, only 9 per cent of all households have been members of such Cooperatives. Gradually, the size has risen. However, in recent period, even where Cooperatives do exist, only 17 per cent of the local households are, in fact, members.

Several reasons may be proposed for a household’s reluctance to participate in a Cooperative in Ethiopia. Under the previous regime, Cooperatives have been used to extend strong government control to the local level and promote socialist ideology through compulsory participation. Field observations suggest that present-day Cooperatives must go through a slow process of trust recovery to overcome persisting suspicion and wariness on the

---

7 Is determined assuming the following information. Maximum total population size in a kebele has been taken as 1200 households in the Districts. The two sampled in the two Districts yield maximum N =2450. The proportion of population becoming members of Cooperatives out of the total population of the highest kebele has been 20per cent only. Hence p=0.2, and q=0.8

---

64
part of their potential members. Another class of explanation is based on an analysis of a household’s economic incentives to participate in a Cooperative.

As a result, the highest participation of households in a Cooperative has not exceeded 20 per cent. This is the probability of success. The remaining 80 per cent are found to be non-members. Then, based on this estimate, the sample size is determined using a formula given by Kothari (2004).

\[ n = \frac{z^2 pqN}{e^2(N-1) + z^2 pq} \text{ ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... }\]

Where, \( N^8 \) = population size

\( n \) = sample size

\( Z = 1.96 \), at 95% confidence interval under normal curve

\( e = \) acceptable error term (0.05) and

p and q are estimates of the proportion of population to be sampled. Hence \( p = 0.20 \) and \( q = 0.8 \). Therefore the sample size has been determined to be a total of 224. However, the sample size, \( n \), has been taken as 240 with 7% contingency. Out of this, 0.33\( n \) = 80 households and 0.67\( n \) = 160 households have been randomly assigned as a treatment group and as a control group respectively to ensure robust match. Equal proportion of sample households have been selected from each sampled Cooperatives of the respective sampled Districts and Cooperatives.

\(^8\) Total population; \( N = 2450 \)
4.3 Method of Data Collection

4.3.1 Types and Source of Data

In this study both primary and secondary data are used. Primary data have been gathered from sampled respondents, while secondary data have been collected from published and unpublished materials available in the Cooperatives, promotion offices and other sources.

4.3.2 Data Collection Techniques and Tools

Both quantitative and qualitative research approach has been employed. Semi-structured interview schedule has been prepared and used for primary data collection. Survey method has been employed to gather quantitative and qualitative primary data. The secondary data has been gathered from published and unpublished sources. Besides, qualitative data has been gathered through focus group discussion and key informant interview for the purpose of triangulation and to substantiate the quantitative information. The interview-schedule has been pre-tested to check its reliability. Cronbach’s Alpha has been used to verify the extent to which the same answer is obtained when the same question is asked repeatedly or how closely answers agree with each other. Accordingly, results from reliability analysis shows that the value of cronbach’s Alpha is 0.865 indicating that information generated so far is reliable.

4.4 Method of Data Analysis

4.4.1 Descriptive Statistics and PSM Methods

Descriptive analysis has been used to produce evidence on the livelihood impact of Cooperatives, which, latter has also been used for triangulation purpose. Moreover, inferential statistical tests like student t-test and an independent-sample t-test procedure has also been used to examine whether
there is a significant relationship between variables and to compare means for the two groups in terms of selected livelihood and poverty impact indicators.

4.4.2 Propensity Score Matching

Propensity score matching method has also been applied to analyze the impact of Cooperatives on household livelihood and poverty. Various matching estimators have been suggested in the literature such as the nearest neighbor matching (NNM), radius (caliper) matching (RM), stratification matching (SM) and kernel matching (KM) estimators (Becker and Ichino 2002). In this study, nearest neighbor matching and radius (caliper) matching estimators have been applied.

Welfare status of the participant and non-participant households have been analyzed by using major poverty indices such as headcount index, poverty gap index and squared poverty gap index.

On top of this, the study also employed econometric model (PSM) to measure the impact of Cooperatives on poverty or household’s welfare, using per capita consumption expenditure as proxy indicator. The rationale for using PSM for this study include: First, due to its non-parametric approach to the balancing of covariates between, the “treatment group” (member) and the “control group” (non-member households), it improves the ability of regression to generate accurate casual estimates and hence removes bias due to unobservable variables. Second, the conventional approach of assessing the impact of an intervention using with and without method has been hampered by a problem of missing data. As a result, the impact of an intervention program cannot be accurately estimated by simply comparing the outcomes of the treatment and the control groups. Third, the PSM builds matched pairs of
comparable households from the participated and the non-participated individuals that show similarity in terms of their observable characteristics. Hence, the use of the PSM method to assess the impact of a discrete treatment on an outcome is considered as one best alternative (Rosenbaum and Rubin 1983).

Using nearest neighbor matching, each treated household “$i^{th}$ household” has been matched with one control “$j^{th}$ household” that has the closest propensity score. The method has been applied with replacement, to ensure that a particular control unit can be a best match for several treatment units. The NNM faces the risk of bad matches, if the closest neighbour is far away. As Smith and Todd (2005) note, a possible drawback of caliper matching is that it is difficult to know a priori what choice for the tolerance level is reasonable. A benefit of this approach is that it uses only as many comparison units as available within the caliper and therefore allows for usage of extra (fewer) units when good matches are (not) available (Dehejia and Wahba 2002).

Besides, the common support restriction has been used to improve the quality of the matches. With common support, control households have been included only for those whose propensity scores fall within the range of propensity scores for the treated households. In this study, the treated group represents member households called participants and the control group represents non-member households named as non-participants.

---

9 The risk of bad match is avoided by imposing a tolerance level on the maximum propensity score distance (caliper). Imposing a caliper works in the same direction as allowing for replacement. Bad matches are avoided and hence the matching quality rises. However, if fewer matches can be performed, the variance of the estimates increases. Applying caliper matching means that those individual from the comparison group is chosen as a matching partner for a treated individual that lies within the caliper (“propensity range”) and is closest in terms of propensity score.
10 Further Dehejia and Wahba suggest a variant of caliper matching which is called radius matching. The basic idea of this variant is to use not only the nearest neighbour within each caliper but also all of the comparison members within the caliper.
The goal is to estimate the change in livelihood and welfare level between members and non-member households in terms of selected livelihood and poverty impact indicators. The comparison techniques account for the effects of exogenous factors that influence the assignment of households into one of these two groups. There are n households, indexed by i = 1⋯n=240.

The treatment is a binary variable, with D=1 for member households and D=0 for non-members households. Whereas \( Y^T_i \) is change in livelihood outcomes of member households (treated group), and \( Y^C_i \) is change in livelihood outcomes of non- member households (control group). Each household also has a vector of characteristics or covariates\(^{11}\) \( x_i \). The livelihood outcome \( Y_i \) and treatment \( D \) can be formally expressed as:

\[
Y_i = \begin{cases} 
Y^C_i, & \text{if } D=0 \\
Y^T_i, & \text{if } D=1 
\end{cases}
\]  
\[ \text{......... (1)} \]

The propensity score for subject \( i (i = 1, \ldots, N) \), is the conditional probability being assigned to treatment \( D_i = 1 \) Vs control \( D_i = 0 \) given a vector \( x_i \) of observed covariates: Meaning, the estimated propensity score, for subject \( i, (i = I, \ldots, N) \) is the conditional probability of being assigned to a particular treatment given a vector of observed covariates \( xi \) (Rosenbaum and Rubin 1983). It is given as:

\[
P(x_i) = Pr(D = 1| x_i) = E(D| x_i)
\]  
\[ \text{......... (2)} \]

\(^{11}\) These characteristics represent variables such as economic, social, environmental and demographic factors that are likely to influence the livelihood of each household.
First the conditional probability of becoming a member in a Cooperative given observed household characteristics has been estimated using a logit model prior to estimation of livelihood impacts, where membership\textsuperscript{12} status in a Cooperative is considered as dependent variable and covariates determining membership as explanatory variables. After the propensity score is calculated, it has been used for matching treated and control units (member households with non-member households) in order to later estimate differences in the livelihood outcome indicators, also known as the Average Treatment Effect on the Treated (\(ATT\)) which is given as:

\[
ATT = E(Y_T^T - Y_T^C | D = 1) = E(E(Y_T^T - Y_T^C | D = 1, P(x_i))
\]

\[
ATT = E[E(Y_T^T | D = 1, P(x_i)) - E(Y_T^C | D = 0, P(x_i)) | D_i]
\] ........ (3)

To measure the impact or the average treatment effect on treated (\(ATT\)), propensity score matching (PSM\textsuperscript{14}) the technique which is proved to be an appropriate, efficient and effective tool for evaluation of programme impact in social research has been used (Rosenbaum and Rubin 1983; Bernard et al. 2007; Francesconi and Ruben 2007). Because, when applied correctly, the consensus among researchers is that non-experimental methods produce the most accurate estimate of program impact (Burtless 1995; Fraker and Maynard 1987; Friedlander and Robbins 1995; Lalonde and Maynard 1987; Michalopoulos et al. 2004; as cited in

\textsuperscript{12} Membership in this study denote participation

\textsuperscript{13} \(ATT\) is the mean difference in livelihood indicators between member and non-member households had they been non-members of Cooperatives. Because the second term is a counterfactual situation that is not observable, it needs to be estimated. The treated and control units have been matched based on their propensity scores before the effects/outcomes are compared.

\textsuperscript{14} PSM is one of a family of statistical techniques that attempts to estimate the effect of a treatment, policy, or other intervention by accounting for the covariates that predict receiving the treatment. It employs a predicted probability of group membership e.g., treatment vs. control group—based on observed predictors, usually obtained from logistic regression to create a counterfactual group.
Methodology

Diaz and Handa 2005). Unlike experimental studies, the assignment of units into treated and control groups are not random for observational studies.

Therefore, the estimation of the effect of treatment may be biased due to the existence of unobservable or confounding factors. However, the propensity score matching method reduces the bias in comparisons between the treated and control groups by comparing outcomes for treated and control units that have similar propensity score. Hence, in this study too, this technique has been used to analyze the livelihood and poverty impact Cooperatives by using propensity score (p-score\textsuperscript{15}). Accordingly, to calculate the p-score, a step by step process has been followed: First, logit model has been estimated in order to determine significant observable covariates.

Second, the propensity score has been calculated. Two choices have been made in estimating the p\textsuperscript{16}-score. The first choice has been concerning the model to be used for the estimation. The second choice involved the variables to be included in the model. In principle, any discrete choice model could be used for such purpose. However, there is a strong preference for logistic or probit models in p-score estimation. For estimating the probability of participation versus non-participation; logistic and probit models usually yield similar result. Hence, logistic\textsuperscript{17} model has been used in this study.

Third, a balancing propensity test has been made using t-test in order to check whether there is a significant difference between the treatment and the

\textsuperscript{15} Propensity score is the probability of a unit (household) being assigned to a particular treatment given a set of observed covariates. Propensity scores are used to reduce selection bias by equating groups based on these covariates.

\textsuperscript{16} Obtain propensity score: predicted probability (p) or log \( \frac{p}{1-p} \). \( B_0 + B_1X_1 + B_2X_2 + \ldots + B_nX_n \)

\textsuperscript{17} Dependent variable: \( Y_i = 1 \), if participate; \( Y_i = 0 \), otherwise. Choose appropriate conditioning (instrumental) variables.
control group. Fourth, matching has been carried out based on a single-index variable, called a propensity score, which summarizes a multi-dimensional vector of characteristics for the treated and control units, and help to compare the differences in outcomes between the two groups in the later steps. Finally, ATT has been calculated from the average p-score differences.

### 4.4.3 Specification of Model

A Logistic regression model, specified hereunder has been applied to estimate the likelihood of a household to participate or not in a Cooperative (P\(_S\)^{18}):

\[
\ln \left( \frac{P_S}{1 - P_S} \right) = \log \left( \frac{e^{(x_i)}}{1 - e^{(x_i)}} \right) = B_0 + B_1X_1 + B_2X_2 + \cdots + B_nX_n \quad \cdots (1)
\]

\[
P_S = \frac{\exp(B_0 + B_1x_1 + B_2x_2 + \cdots + B_nx_n)}{1 + \exp(B_0 + B_1x_1 + B_2x_2 + \cdots + B_nx_n)} \quad \cdots (2)
\]

Where, P(X) = P\(_S\) = e\(^{(x_i)}\) = propensity score; \(x_1, x_2, \ldots x_n\) are the predictor variables determining household participation, and \(P_S\) denotes the probability that the household participates in Cooperatives. Stepwise logistic regression method has been used to determine automatically which variables to add into or drop from the model.

### 4.4.4 Poverty Threshold Level Identification

Measurement of poverty involves identifying or defining a poverty line that helps to determine the status of the member and the non-member households with respect to this line. Despite the availability of numerous approaches for estimating poverty line, the Direct Calorie Intake (DCI); the Food Energy Intake (FEI) and the Cost of Basic Needs (CBN) approach are the most

---

18 Ps is the propensity score
commonly used ones. These three approaches are different on the basis they use to define the threshold or poverty line.

The DCI uses a minimum requirement of 2200 kilocalories per day per person to be healthy as a reference to define the poverty line (WHO 1985). So any household that meets this criterion can be considered as not in poverty while those below this line are called households in poverty.

The FEI defines the poverty line using the level of per capita consumption at which people are expected to meet the 2200 kilocalories nutritional requirement. In other words, this method defines the threshold line by estimating the value of the per capita consumption at which every household is expected to fulfill their calorie requirement. The CBN accommodates estimating cost of direct calorie intake (food) and other basic non-food requirements.

Among these, the CBN approach is used in this study due to its consideration of basic non-food requirements while defining the threshold level. One major challenge of poverty study results is sensitivity to the choice of the threshold or poverty line and poverty measure. The national level food and total poverty line set by the government of Ethiopia was Birr 1985 and 3781 respectively (MoFED 2012). Since the choice of approaches to define the line are at the disposal of the researcher, the results obtained are likely not robust. However in this study the agreed-upon international poverty line (US$1.25/capita/day) adjusted for purchasing power parity have been employed.
4.4.5 The FGT Poverty Index

To compare the poverty levels among treated and control groups, the three measures of poverty index developed by (Foster and Shorrock 1988) has been used to get exact poverty levels. This helps to identify the number of households below and above the estimated poverty line. The poverty index has been given by and defined as follows:

\[
P_a = \frac{1}{N} \sum_{i=1}^{m} \left( \frac{z - y_i}{Z} \right)^\alpha
\]

….. (1)

where; \( Z \) -is the agreed-upon poverty line (US$1.25/capita/day) adjusted at purchasing power parity; \( y_i \) - is consumption expenditure per capita for the \( i^{th} \) person; \( N \) - is the number of people in the sample population; and \( m \) - is the number of poor (those with consumption expenditure per capita at or below \( z \)); and \( \alpha \) - is a poverty aversion (sensitivity) parameter.

When \( \alpha = 0 \), the above equation gives the incidence of poverty that is also called the Head Count Ratio or simply Head Count Index. It is defined as the percentage of people falling below the poverty line. When \( \alpha = 1 \), the above equation gives the depth of the poverty called Poverty Gap Index. When \( \alpha = 2 \), the equation shows a measure called the severity of poverty index or squared poverty gap.

The Foster, Greer, and Thorbecke (FGT) poverty index has been computed, especially the headcount ratio, and other indices being generated using stata commands, which is very powerful if one wants to decompose poverty indices. The study used the international poverty line of US$1.25/capita/day. Khonje M. et al. (2015) has also applied same approach to study poverty in eastern Zambia. The consumption expenditure data has been
used because they give a better poverty measurement than income (Christiaensen, Scott, & Wodon 2002). It is worthwhile to mention that, when there is no exact poverty line the first order of stochastic dominance is preferable at this time (Kiriti-Ngangaand Mburu 2007).

4.4.6 Operationalization of Variables

4.4.6.1 Variables Used to Generate Propensity Score

The propensity score has been generated using different covariates that determine household membership status. The logit model has been applied to generate the propensity score that can be used for matching households and ultimately for livelihood outcome estimation.

**Dependent Variable:** In this study, the dependent variable is membership status (participation) which is dummy in nature that assumes a value of 1 if household participates (registered as member) and 0 if otherwise. The dependent variable has been used for generating propensity score that has in turn been used for matching purpose and estimation of ATT.

**Independent Variables:** The major endogenous variables hypothesized to influence both participation (treatment) and the outcome variable (livelihood and poverty) include variables such as:

a. **Family size:** It is a continuous variable. Family size affects participation. Households with average or less family size are assumed to participate in to Cooperatives. The relationship between these two variables is assumed to be inverse. Hence, households with average or less family size are likely to have better livelihoods and welfare.
b. **Household head literacy:** It is a categorical variable. Household head literacy and participation is assumed to have a direct relationship. The literate the household head, the more likely the hhh to participate in to Cooperatives and are likely to exert maximum effort to improve and sustain their livelihood and thereby have better welfare. Households that are literate assume to take up a value 1, and 0 if otherwise.

c. **Age of household head:** It is a continuous variable. Household heads with matured age are likely to participate and are assumed to have less difficulty to come out of poverty and improve their livelihood than those that are not matured. The relationship between these two variables is assumed to be direct. Hence, hhh whose age is above 18 years are assumed to have better livelihood.

d. **Input cost:** Input cost is defined as categorical variable. One of the primary objectives of Cooperatives is to minimize input costs through scale effects (such as low unit transaction cost) and better negotiation power. Low Input cost is assumed to have increased the likelihood of household participation and to have a positive impact on hh livelihood and welfare. It assumes a value 1 if hh input cost has reduced and 0 if otherwise.

e. **Transaction cost:** Transaction cost is defined as categorical variable. Costs incurred by buyer seller search, information search, information processing and decision making, costs of negotiating the terms of the exchange, and cost of contract enforcement activities are assumed to have increased the likelihood of household participation and thereby impacts hh livelihood and poverty. Transaction cost and household livelihood and welfare are assumed to have an inverse relationship. Thus, households who witnessed that the
Chapter IV:  

Cooperative has minimized their transaction cost are assigned a value 1, and 0 if otherwise.

f. **Marketing costs:** It is defined as categorical variable. The cost of transferring title and moving goods to the customer is assumed to have increased the likelihood of household participation and at the same time has a positive impact on hh livelihood and welfare. High market costs reduce producers' returns through lower net farm income. Cooperatives are assumed to improve hh livelihood by lowering marketing costs through improved operational efficiencies\(^\text{19}\). Therefore, a Cooperative that is able to minimize the marketing cost assumes a value 1, and 0 if otherwise.

g. **Market information:** Market information is a categorical variable. Information is power that can be used to take advantage of market situations. Hence, Cooperatives that provide up-to-date market information to hh are assumed to have increased the likelihood of household participation and to have a positive impact on hh livelihood and welfare, because production and pricing decision is made based on market information which in turn will impact hh livelihood and welfare. Thus, households that have got better access to up to date market information will assume a value 1, and 0 if otherwise.

h. **Bargaining power**\(^\text{20}\): It is a categorical variable. Cooperatives are assumed to enable households gain a better balance of power among the market parties and are assumed to have increased the likelihood of household participation. This enables hh to gain advantage of prices for their produce

---

\(^{19}\) Application of less expensive method of operating Cooperative activities such as handling and storing of grains

\(^{20}\) Bargaining power is a relative strength of buyers and sellers in influencing the terms of exchange in a transaction.
which will have an imperative impact on household livelihood and welfare. Bargaining power and hh livelihood is assumed to have direct relationship. If bargaining power of hh has improved then their livelihood is assumed to improve. Thus, Households that have better bargaining power assumes a value 1 and 0 if otherwise.

4.4.6.2 Outcome Variables

The outcome variables employed in this study includes broader variables of livelihood (measures such as economic, social, environmental and demographic characteristics of a household) and poverty impact indicators such as incidence, depth and severity indices. The outcome variables assume either continuous or binary values. Key economic indicators that have been incorporated under this study include such variables as income, asset (land, livestock and house), saving, employment, access to water, access to health and access to education; while social indicators include variables like sense of openness, self-confidence and discrimination. Similarly environmental indicators include: household participation on natural resources conservation and management, household capacity to cope up with adverse climate change, application of harmful chemicals, awareness on early warning, number of trees planted and grown by households, quantity of soil bund constructed. Above all demographic impact indicators variables such as family size, household head literacy and age of household head has also been considered in this study. The detailed operational definitions have been mentioned as follows.

a. Household income is a continuous variable. Aggregate income by household generated either in the form of dividend from Cooperatives or from
other sources is assumed to have a positive association with household livelihood and welfare.

b. Household asset is a continuous variable. Land and livestock holdings are found to have a positive effect on the likelihood of farmers to participate in agricultural Cooperatives (Ito, Bao and Su 2012) which has a positive bearing on livelihood and welfare.

c. Household saving is a continuous variable. Saving is assumed to have positive relationship with household livelihood since improvement in the level of household saving is expected to increase investments on modern agricultural technologies to improve household livelihood and reduce poverty. The more the level of household savings the more will be the hh livelihood and welfare improvement.

d. Employment is a continuous variable measured in terms of jobs offered to the community. Cooperatives are assumed to offer jobs to the surrounding community. Employment is assumed to have a positive relationship with household livelihood and welfare.

e. Value of land is a continuous variable. Value of household land and livelihood and welfare is assumed to have direct relationship. The higher the value of household land, the better will be their livelihood and welfare.

f. Value of house: Value of household house is a continuous variable. Value of household house and livelihood and welfare is assumed to have a positive relationship. The higher the value of a house of a household, the better will be the livelihood and welfare of the household.
g. *Value of livestock* is a continuous variable. Value of household livestock and household livelihood and welfare is assumed to have a positive relationship. Households who possess livestock with high value are assumed to have a better livelihood and welfare.

h. *Sense of openness* is a categorical variable. Cooperatives are assumed to promote and develop sense of openness among people to help households to democratically, openly and honestly express their views and claim unmet demands that would fulfill their livelihood. Well established and developed sense of openness is assumed to have positive association with household livelihood. Hence, hhs who witnessed their sense of openness are well developed assumes a value 1, and 0 if otherwise.

i. *Self-confidence* is a categorical variable. Cooperatives are assumed to develop individuals’ self-confidence. Households who have self confidence in them are assumed to have better livelihood. Self-confidence is assumed to have positive relationship with household livelihood. Hence, hhs with self-confidence assumes a value 1, and 0 if otherwise.

j. *Discrimination* is a categorical variable. Cooperatives are assumed to promote and develop non-discriminatory approach. It values people are all equal and that equal opportunity is given to all to become a member and participate in the affairs of Cooperatives without discrimination. Absence of discrimination strengthens solidarity and unity of the community and would help them support each other, hence, is assumed to have positive relationship with household livelihood. Household who are not discriminated are assumed to have better livelihood. Hence, households who are not discriminated assume a value 1, and 0 if otherwise.
k. HH participation on natural resources conservation and management is a categorical variable. Cooperatives are assumed to contribute positively to environment through its varied activities. Participation of hh on natural resources conservation and management activities is assumed to have a direct relationship with household livelihood and welfare which will favor activities required to secure the necessities of life. This variable assumes a value 1 if hh participates on natural resources conservation and management activities, and 0 if otherwise.

l. Household capacity to cope with adverse climate change is a categorical variable. Household capacity to cope with adverse climate change has a positive relationship with livelihood and welfare. The higher the coping capacity of the household, the better will be their livelihood and welfare.

m. Number of trees planted and grown by household is a continuous variable. Number of trees planted and grown by a household is assumed to have a positive relationship with livelihood and welfare. A household who planted and grow as many trees as possible has better capacity to conserve the environment which will have a direct positive relationship with livelihood and welfare.

n. Quantity of soil bund constructed by household is a continuous variable. Quantity of soil bund constructed by a household is assumed to have a positive relationship with livelihood and welfare. A household who constructed as many kilometers of soil bund as possible has better capacity to conserve the environment which will have a direct positive bearing on livelihood and welfare.