DETERMINANTS OF RURAL FEMALE-HEADED HOUSEHOLD PARTICIPATION DECISION IN INCOME DIVERSIFICATION AND ITS IMPACT ON THEIR LIVELIHOOD IN GUDURU DISTRICT, OROMIA REGIONAL STATE, ETHIOPIA

A Thesis Submitted to the School of Graduate Studies of Jimma University in Partial Fulfillment of the Requirements for the Award of the Degree of Masters of Science (M.Sc.) in Economics (Economic Policy Analysis)

By:

OROMO TADESE OLJIRA

JIMMA UNIVERSITY

COLLEGE OF BUSINESS AND ECONOMICS

M.Sc. PROGRAM

NOVEMBER, 2021

JIMMA, ETHIOPIA

DETERMINANTS OF RURAL FEMALE-HEADED HOUSEHOLD PARTICIPATION DECISION IN INCOME DIVERSIFICATION AND ITS IMPACT ON THEIR LIVELIHOOD IN GUDURU DISTRICT, OROMIA REGIONAL STATE, ETHIOPIA

A Thesis Submitted to the School of Graduate Studies of Jimma University in partial fulfillment of the Requirements for the Award of the Degree of Masters of Science (M.Sc.) in Economics (Economic Policy Analysis)

By:

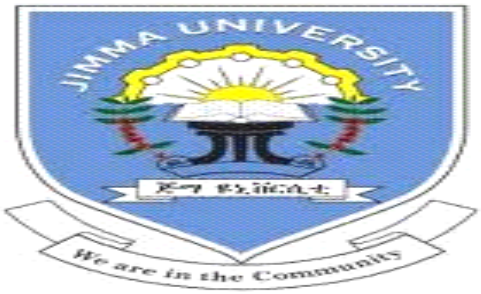
OROMO TADESE OLJIRA

Under the supervision of:

Mr. Sisay Tolla (Assistant professor)

And

Mr. Aleka Jeldu (M.Sc.)



JIMMA UNIVERSITY

COLLEGE OF BUSINESS AND ECONOMICS

M.Sc. PROGRAM

NOVEMBER, 2021

JIMMA, ETHIOPIA

CERTIFICATE

This is to certify that the thesis entitled "**Determinants of Rural Female-Headed Household Participation decision in income diversification and Its Impact on their Livelihood in Guduru district, Oromia regional state, Ethiopia**", submitted to Jimma University for the award of the Degree of Master of Science (MSc) in economic policy analysis and is a record of valuable research work carried out by Mr. Oromo Tadese Oljira, under our guidance and supervision.

Therefore, we hereby declare that no part of this thesis has been submitted to any other university or institution for the award of any degree of diploma.

Main Advisor's Name	Date	signature
Mr. Sisay Tolla (Assistant professor).	_____	_____
Co-Advisor's Name	Date	Signature
Mr. Aleka Jeldu (MSc)	_____	_____

As member of the Board of Examiners of the M.Sc. Thesis Open Defense Examination, We certify that we have read, evaluated the Thesis prepared by Mr. Oromo Tadese and examined the candidate. We recommended that the Thesis be accepted as fulfilling the Thesis requirement for the Degree of Master of Science in Economics (economic policy analysis).

_____	_____	_____
Chairperson	Signature	Date
_____	_____	_____
Internal Examiner	Signature	Date
_____	_____	_____
External Examiner	Signature	Date

DECLARATION

By my signature below, I declared and affirm that this Thesis which entitled "**Determinants of rural female-headed household participation decision in income diversification and Its Impact on their livelihood: in Guduru district**", has been carried out by me under the guidance and supervision of Mr. Sisay Tolla (Assistant professor) and Mr. Aleka Jeldu. I have followed all ethical and technical principles of scholarship in the preparation, data collection, data analysis and compilation of this Thesis. Any scholarly matter that is included in the Thesis has been given recognition through citation.

This Thesis is submitted in partial fulfillment of the requirements for a Master of Science degree at Jimma University. The Thesis is deposited in the Jimma University Library and is made available to borrowers under the rules of the library. I solemnly declare that this Thesis has not been submitted to any other institution anywhere for award of any academic degree, diploma or certificate.

Brief quotations from this Thesis may be made without special permission provided that accurate and complete acknowledgement of source is made. Requests for permission for extended quotations from or reproduction of this Thesis in whole or in part may be granted by the Head of the School or Department when in his or her judgment the proposed use of the material is in the interest of scholarship. In all other instances, however, permission must be obtained from the author of the Thesis.

Researcher's Name	Date	Signature
Oromo Tadese Oljira	_____	_____

BIOGRAPHICAL SKETCH

Oromo Tadese was born from his parents Sena Merera and Tadese Oljira in Fincha town, Abby Commen District, H/G/Zone of Oromia Regional State, Ethiopia in May 17, 1997. He attended his primary school at Fincha primary school and high school at Tullu Habib Secondary School. Then, he joined Fincha Preparatory School and completed his preparatory education in 2015.

After completion of his preparatory school education, he joined Wollega University College of Business and Economics Department of Economics in November 2015 and graduated with a BA in Economics on June 22, 2018. After graduation, he directly employed Wollega University, College of Agriculture Department of Agricultural Economics. The author started his Graduate Studies at Jimma University College of Business and economics in September 2019 to pursue his M.Sc. Degree in economic policy analysis in a regular program.

ACKNOWLEDGEMENT

First and foremost, I would like to extend my unshared thanks to the Almighty Waaqaa for providing me the opportunity and smoothening of all aspects regarding writing this research.

From the initial stages of thesis research proposal development to the completion of the thesis write-up, it gives me great pleasure to express my heartfelt gratitude to my major advisor Mr. Sisay Tolla (Assistant professor) and co-advisor Mr. Aleka Jeldu (MSc) for their keen interest, constant supervision, valuable guidance, kindness, encouragement, and constructive criticisms. “Thank you” is not enough to express my gratitude to my esteemed and honored advisors, but it is a more appropriate expression of gratitude; if it were more than “thank you,” my advisors would be the only ones to receive it.

I would like also to express my sincere gratitude to Wollega University Particularly, the College of Agriculture for giving me postgraduate study by covering my education expense. My deepest thanks to my sister Zoditu Tadese as well as to my co-workers for their commitment and scarifies bring me to this stage and my deepest thanks to my brothers and sisters who have always been there for me whenever I need them; their encouragement keeps me going and their love empowers to me. Equally important, my gratitude goes to enumerators and all respondents to my research questionnaire.

Finally, I would like to thank all Guduru district female headed households, Office of Agriculture, and office of women and child affairs for their patience and hospitality in giving basic information.

.

TABLE OF CONTENT

Contents **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AE	Adult Equivalent
AIC	Akakias Information Criteria
ATE	Average Treatment on Treated
ATT	Average Treatment Effect
ATU	Average Treatment effect on Untreated
CC	Contingency Coefficient
CIA	Conditional Independence Assumption
CM	Caliper Matching
DD	Difference in Difference
DHM	Double- Hurdle Model
ETB	Ethiopian Birr.
FAO	Food and Agricultural Organization
FHH	Female-Headed Household
GDP	Gross Domestic Product
HH	Household
HI	Herfindahl index
IHI	Inverse of Herfindahl Index
IMR	Inverse Mills Ratio
KM	Kernel Matching
KM	Kilometer
MLE	Maximum Likelihood Estimator
NNM	Nearest Neighbor Matching
PSM	Propensity Score Model
RDD	Regression Discontinuity Design
RM	Radius Matching
RNFE	Rural Non-farm Employment
RSM	Randomized Selected Model
SID	Simpson Index of Diversity
SSA	Sub-Saharan African

TLU	Tropical Livestock Unit
USD	United State Dollar
VIF	Variance Inflation Factor

ABSTRACT

The rural economy in developing countries usually characterized by an agrarian economy in which large number of the smallholder farmers are practice farming activities. However, agriculture as a sole livelihood activity is not sufficient to overcome the livelihood constraints in rural area. As a result, income diversification is considered as important strategy for livelihood improvement. Therefore, this study was aimed at analyzing the determinants of female-headed household participation decision in income diversification and its impact on their livelihood in Guduru district, Ethiopia. To do this research, both primary and secondary data were used. The primary data was collected through multi-stage sampling technique from 245 rural female-headed households through structural questionnaires. In addition to this, the secondary data was collected by reviewing relevant sources such as documents of the office of agriculture of the district and other relevant organizations. Descriptive statistics (mean, chi-square test, and t-test), Simpson index of diversity, and econometric models (Double-hurdle and Propensity Scores Matching) were used to analyze the data. The descriptive statistics result showed that the degree of income diversification was 0.27 which is low. This is because of different constraints like lack infrastructure, altitude of the society, adultness, lack of credit, and lack of awareness and training. Beside this, there were opportunities that motivate rural female-headed household to participate in income diversification from two side pull and push factors. The result of the Double-hurdle model showed that, except age, gender-based discrimination, distance from market center and annual average agricultural income other variable like education states, family size, Livestock ownership, and access of credit were positively affect income diversification in the first hurdle. The intensity of income diversification was negatively affected by age of the household, family size, livestock ownership and gender-based discrimination whereas positively affected by frequency of extension contact in the second hurdle. Moreover, Propensity Score Matching result indicates that income diversification was a significant and positive impact on income and saving of income diversifying female headed households. The study concluded that, participation of rural female-headed household into income diversification is one of the key valuable potential solutions to enhance livelihood indicators (income. consumption expenditure, and saving) in the study area. As a result, it is suggested that the government and all other stakeholders should focus on rising income source and further research should be conducted for a more holistic understanding of the significance of income diversification than specialization in rural area.

Keywords: Double-hurdle, Female-Headed Household, Income Diversification, Livelihood, Propensity Score Matching,

CHAPTER ONE

1. INTRODUCTION

This study introduces and discusses the determinants of female-headed household participation income diversification in the rural area and its impact on their livelihoods. The first chapter of this study shows specifically:- the background of the study; statement of the problem; research question; and the overall object of the finding; hypothesis of the study; how the paper is organized; limitation of the study; ethical consideration and description of the terms.

1.1 Background of the Study

Agriculture is major economic activity of human being in the real world, since it provides basic needs such as food, clothing and shelter. It has been confirmed that every 1% rise in agricultural output translates into a 0.6–1.2% reduction in the numbers of extreme poor households in the world (Ismail et al., 2020). Likewise, in real world, economic activities are more dominated by agricultural sector which is the major source of income for rural households. It is one of the large components of national income of world poor country and it reduces the extreme poverty, help as source of income and improves the food security of world poor countries, for those who live in rural areas and work mainly farming (FAO, 2020).

Most less developed county particularly Latin America and Sub Saharan Africa economic activities highly dominated by the agriculture (Abate, 2019). Especially, in Sub-Saharan Africa with approximately two-thirds of the population depending on agriculture for their livelihood (Onadeko et al., 2020). However, the sector has been continually blamed for its failure to guarantee sufficient livelihood for smallholder farm households in the region because of decreasing farm sizes, low productivity and high degree of subsistence farming (Abera et al., 2021). As a result, food insecurity and poverty are widespread in Sub- Saharan Africa (Balense and Debebe, 2020)

Ethiopia is one of the Sub-Saharan Africa countries where majority of the population lives in the rural areas mainly depending on agriculture for their livelihoods (Abera et al., 2021). Therefore, the sector contributes about 34% of total Gross Domestic Product (GDP) and more than 70% of total employment opportunities, 70% of the raw material requirements for local manufacturing industries, and about 70% of total export (Welteji, 2018).

However, agricultural activities in rural Ethiopia are highly characterized by rain fed, low productive, fragmented land size, recurrent drought occurrence compounded with poor usage of improved agricultural inputs, high population growth, subsistence oriented and dominated by smallholders those who cultivating less than 0.5 ha (Abera et al., 2021). As a result, the contribution of the sector for poverty reduction and enhancement of food security is limited for the last long year. Therefore, widening livelihood diversification choices and diversification of income sources supplement the effort of food security and poverty reduction in rural Ethiopia (Balense and Debebe, 2020).

Accordingly, rural households participate in income diversification to resist push and pull factors which rise in rural household economic activities (Abbeam et al., 2020). The rural household's livelihoods are not permanently run by the income derived from seasonal farming, because farming as a principal source of income has failed to achieve the maximum livelihood of rural farmer households due to harsh condition (Fassil and Elias, 2016). As a result, income diversification among rural households in rural area has been grown to become a common phenomenon (Vien, 2017)

Diversification refers to a process in which households voluntarily or involuntarily increase the number of economic activities they are involved in. Thus, we can say diversification is the norm (Teji, 2020). Moreover, it is defined as the process by which rural households construct increasingly diverse livelihood portfolios, making use of increasingly diverse combinations of resources and assets to meet their basic needs, improve their living standards or welfare, and manage. Diversification is becoming an increasingly important livelihood strategy among rural households in SSA (Alobo Loison and Bignebat, 2018)

According to Kassie et al (2017) farm households can diversify their income source into on-farm, off-farm, and non-farm income components. On-farm income is income gained from either farming own-land or land acquired or accessed by cash or share tenancy, and income from livestock production. Off-farm income is income gained from labor wage working from other farms with-in the agriculture sector. Nonfarm income refers to income from non-agricultural sources like non-farm employment, urban-to-rural remittances, rental income, non-farm rural-wage, and international remittances to a farm household. Thus, income diversification between rural households brings strategically distribution of productive assets on the different income-

generating activities (Onadeko et al., 2020). In addition, it is also a tools that rural household used to manage risk and coping strategy meant to mitigate the effects of economic hardship (Debesai, 2020).

Income is one of the essential instruments to analyze the development of human capital especially in less developed countries (Le and Le, 2020). Therefore, the level of income generated in a rural area is highly influenced by human capital related variable (gender and age of household head, number of family size, education level of the household, attitude to risk and), livelihood assets (livestock holding, size of cultivated land, irrigated land, a wealth of households, aggregate total crop product) and infrastructure-related variable (proximity to market, extension service, access to credit and cooperative) (Teji, 2020). These constraints are especially difficult for poor female-headed households to overcome, becomes it is linked to gender. Female-headed households may be constrained because they often are poorer and because they face special constraints due to gender. Generally, they have less education and less access to productive assets and credit, which limit their options of diversification. Female-headed households might also be hindered by norms about female labor force participation (Elin, 2015).

Rural women play a lion share in the rural economies of both developed and developing countries (Omirin and Okpara, 2018). Specially, in Africa women labor participation to rural economic activity is very high (Palacios *et al.*, 2017). As a result, women are the essential pathway of dramatic change in fighting against extreme poverty specifically in rural households. In addition to this, they are a base for the country in general and their family in specific. Thus, women are also considered as an agent of rural economy by playing the important role, like selling and buying input and output, crop collection, food production and overall management of the household at all (FAO, 2020).

Guduru district is one of the districts of Oromia Regional State of Ethiopia, where agriculture is the primary source of income for the rural people. By its nature, agriculture could not cover all livelihoods of rural people. So smallholder rural female households are engaged into diverse their source of income. As a result, the goal of this research is to look at income diversification of female-headed households and how it affects livelihood indicators (income, consumption expenditure, and saving) in Guduru district, H/G/Zone of Oromia regional state, Ethiopia.

1.2 Statement of the problem

In developing countries, agriculture is an important sector for majority of the rural populations livelihood as it is a source of income, employment and foreign exchange (Abera et al., 2021). According to Palacios *et al.*,, (2017) in African agriculture, women labor participation is typically estimated to be between 60 and 80 percent. In five Sub-Saharan African countries, the average female labor involvement in agriculture was estimated to be 40% (Ethiopia, Malawi, Rwanda, Uganda, and the United Republic of Tanzania). Therefore, in this sector women play a critical role.

However, agricultural growth, resilience, and agricultural strategies alone as the primary solution for rural poverty reduction may not be a long-term option, because of agriculture of less developed counties characterized by very small land-holdings, drought, floods, crop loss, poor infrastructure status, and gaps in market access in rural areas. As a result, the monetary living standards of households are very low (consumption levels of USD 2.2/day) Adem et al., (2018). Moreover, multidimensional poverty and food insecurity is exacerbating the conditions of smallholder farmers in rural areas which mean that agriculture alone unable to support all of the rural population. Hence, rural income diversification is equally important for poverty reduction, food security, and wellbeing (Majbauddin et al., 2020).

Therefore in developing countries, particularly in the rural areas, income diversification is getting more consideration in the development economics research than ever before (Loison and Bignebat, 2018). The share of off- farm and non-farm income is increasing in most developing countries as a result of push and pulls factors (Astatike and Gazuma, 2019). This issue attracts the attention of stockholder for two important reasons. First, to increases the real income of the rural household. Second, due to population growth and idiosyncratic and covariate shocks, income constraints and poverty are common in most developing countries (Fentahun and Kemaw, 2019). These are some problems that force scholars to investigate income diversification as pathway to move out of these constraints across regions and countries.

Several recent studies show the relationship between income diversification and some of livelihood indicators like income, food security, and consumption expenditure. But, there is a controversial argument among the researchers concerning the impact of income diversification on the livelihood of rural households who participate in income diversification. Some recent

empirical findings indicate that income diversification has a positive impact on participant households livelihood (Adem et al., 2018; Alemayehu et al., 2018; Derbe, 2020; Teji, 2020). While others argue that income diversification harms participant household livelihood by decreasing agricultural output of the farmer (Kassie et al., 2017; Anteneh and Ganamo, 2019; Salam et al., 2019). In addition to this, the researcher disagrees on the income diversifying capacity of the rich and poor households. Based on this, most of the scholars say the poor farmers are more diversify their income than rich ones (Ababbo, 2015; Diep and Vien, 2017; Ntwalle, 2019). While another researcher argues that, the rich have more power to diversify their income than the poor (Ghosh and Sujan, 2020).

However, to the best of researcher knowledge, Ethiopia has diverse climatic zones and the results obtained from the specific region or area cannot be generalized for the whole country. Therefore, this signifies the need for further investigation on the factors determining income diversification in the study area. Moreover, the researcher couldn't find any study undertaken on the determinant of female-headed household participation in income diversification and its impact on their livelihood in the study area and in Ethiopia also. Because of female-headed households are not the main interest of most previous researcher, the gender of the head household is often included as a control variable.

In addition to the above, most of recent studies in Ethiopia have focused on examining the factors that affect income diversification of households and its impact on the food security (e.g. Adem et al., 2018; Fentahun and Kemaw, 2019; Etea et al., 2019), that attempt to create a research framework that integrate socio-economic and demographic factors; but, significant altitudinal factors receive little attention. Furthermore, there are a number of unobserved factors that largely influence income diversification, like gender based discrimination on FHHs in case of rural labor force participation in rural income generating activity and again there is no clear link between income diversification and disposable income, consumption spent, and saving of HH in case of previous studies. Moreover, little attention was given for identifying the constraints and opportunities of income diversification in rural area specifically for female-headed households. Therefore, this study is initiated to analyze and contribute to the knowledge gap on the determinant of rural female-headed household participation in income diversification and its impact on their livelihood.

1.3. Research Questions

- What is the current level of income diversification of female-headed household in the study area?
- What are factors that affect female-headed household participation decision and intensity of income diversification in the study area?
- What is the impact of income diversification on income, consumption expenditure, and saving of female-headed households in the study area?
- What are the major constraints and opportunities of female-headed household face to participate in income diversification in the study area?

1.4. Objective of the Study

1.4. 1. General objective

The general objective of this study was to identify the factors that affect female-headed household participation in income diversification and evaluate its impact on their livelihood in the study area.

1.4. 2. Specific objectives of the study

- To measure the current level of income diversification FHH in the study area.
- To identify the factors that affecting female-headed households participation decision in income diversification and its intensity in the study area**.**
- To examine the impact of income diversification on female-headed households income, consumption expenditure and saving in the study area.
- To identify major challenges and opportunities in income diversification in study area.

1.5. Hypothesis of the Study

The study formulated the following hypothesis;

The determinants of female-headed household participation decision and intensity of participation in income diversification

H_{01}: Age does not influence the participation decision and intensity of female-headed households in income diversification.

H_{02}: The family size of the household doesn't influence the female-headed household participation decision and intensity participation in income diversification.

H_{03}: The education level of the household doesn't influence the female-headed household participation decision and intensity participation in income diversification.

H_{04}: The dependency ratio of the household doesn't influence the female-headed household participation decision and intensity participation in income diversification.

H_{05}: Distance of household homes from the market does not influence the female-headed household participation decision and intensity participation in income diversification.

H_{06}: Average annual agricultural income the household doesn't influence the female-headed household participation decision and intensity participation in income diversification.

H_{07}: Land size does not influence the female-headed household participation decision and intensity participation in income diversification.

H_{08}: Livestock ownership doesn't influence the female-headed household participation decision and intensity participation in income diversification.

H_{09}: Member of farmer-based organization does not influence the female-headed household participation decision and intensity participation in income diversification.

H_{10}: Frequency of extension contact does not influence the female-headed household participation decision and intensity participation in income diversification.

H_{11}: Availability of shock does not influence the female-headed household participation decision and intensity participation in income diversification.

H_{12}: Access of credit service does not influence the female-headed household participation decision and intensity participation in income diversification.

H_{13}: Gender-based discrimination does not influence the female-headed household participation decision and intensity participation in income diversification.

H_{14}: Access of irrigated land does not influence the female-headed household participation decision and intensity participation in income diversification.

Impact of income diversification on outcome variables (income, consumption expenditure, and saving)

H_{01}: Participation in income diversification doesn't help to increase the income of the female-headed household

H_{02}: Participation in income diversification doesn't help to increase the consumption expenditure of female-headed households.

H_{03}: Participation in income diversification doesn't help to increase the saving of female-headed households.

1.6. Significance of the Study

The purpose of this paper is to analyze determinants of rural female-headed household participation decision in income diversification and its impact on their livelihood in Guduru district, Oromia regional state, Ethiopia. The results generated by this research may be important to progress trainers, researchers, policy makers. Especially, the results of study may be helpful to development practitioners working on gender parity by demonstrating which issue needs to intervene in. It may provide researchers as a reference for those who are interested in doing further study on determinants of income diversification and its impact on livelihood. Besides, it might be very important for policymakers to define, approve and implement appropriate policy based on the current situation of the countries.

1.7. Scope of the Study

This study was delimited both geographically and conceptually. Regarding the geographical scope, even though the issue of income is interesting to conduct throughout the country, because of financial and time constrains the study was delimited to the Guduru district. Conceptually, the subject matter of the study was delimited to the determinant of female-headed household participation in income diversification and its impact on their livelihood in the Guduru district.

1.8. Limitation of the Study

This study focused on the determinants of female-headed rural household participation decision in income diversification and its impact on their livelihood in Guduru district, Oromia regional

state, Ethiopia. However, the result of the study may have limitations as far as the research is conducted. The study employed a cross-sectional data research design, which limits the researcher's ability to generalize the findings into continuous outcomes. Accordingly, the researcher was unable to draw broad generalizations about the country based on data from one location since livelihood indicators (such as income, consumption expenditure and saving) were studied at the household level. Beside this, in case of data collection some respondents were unwillingness to give necessary information; but by creating good relationship with the respondents, all the relevant information was collected.

1.9. Organizations of the Study

This thesis contains five chapters. Chapter one introduces and discusses background of the study, statement of the problem, research questions, general and specific objectives of the study, hypothesis, significance of the study, scope and limitation of the study. Chapters two covered the relevant related review literature and conceptual framework. Chapter three is about description of the study area, research design, sources of data and collection methods, sampling techniques, sample determination, and method of data analysis. Chapter four is about results and discussion. Finally, chapter five includes conclusion, recommendations and suggestion for future research

1.10. Ethical Consideration

Before starting the data collection, the study design was explained to the Officials of the Agriculture office and the Administrative of Guduru district for their permission and support. The nature of the study was fully explained to respondents to obtain consent and because of the questionnaires for only female headed household respondents; awareness was given for populations before actual data collection and finally the administration of actual fieldwork. No false promise. Information was collected after obtaining consent from the study participant. Data obtained from each study participant was kept confidential, and all people who participated in the study were acknowledged.

1.11. The Description of the Terms

Female-headed household: - A female head of household refers to a woman in charge of managing the family as a result of divorce, separation, immigration, widowhood and

occupational states of the husband, spouse death, addiction or disability of husband, increased life expectancy among women, migration, or being abandoned by husband (Davis et al., 2017).

On-farm income: - Is refers to income generated from either farming own-land or land acquired or accessed by cash or share tenancy, and income from livestock production (Onadeko et al., 2020).

Off-farm income: - Refers to the activities either individual rural households are wage employment or self-employment activities which derive income for the household's labour supplies outside their own farm. Wage employment include, farm wage, skilled and unskilled regular wage and casual daily work. In the other hand, self-employment off-farm income comprises from selling firewood, stone mining, grain and livestock trading, petty trading, weaving, mat making, pottery and hand craft (Kassie et al., 2017).

Non-farm income: - It is also one of the sources of income in rural household which rise from outside farm and off-farm source of income. For example, income derived from remittance from relative and friend not present living with the household, pension, gift, renting out asset, and government aid (Kassie et al., 2017).

Participation: - refers the engagement of female-headed households in off-farm and non-farm economic activities or the act of taking part an activity or processes by which women have shared from the involvement in non-farm work. It is viewed as the involvement of female-headed households through allocating their time for off-farm and non-farm economic activities by considering either as a main job or as alternative tasks. Participation often means involvement in externally conceived activities via contributions and benefits (Bayu, 2018)

CHAPTER TWO

2. LITERATURE REVIEW

This chapter comprises theoretical, empirical and conceptual literature relating to income diversification participation, intensity of participation and its impact on the livelihood. The chapter starts by offering some definitional aspects of the terms and concepts of the study, namely: income, and income diversification. The chapter also reviews literature on the determinants of income diversification participation and extent of income diversification participation and followed by presenting empirical and conceptual evidence on the analysis of income diversification participation decision and its impact.

2.1. Theoretical Literature Review

2.1.1 Basic concepts and definitions of income and income diversification

According to Collins Essential English Dictionary (2009) income is defined as the total amount of money derived from work or obtained from other sources over a given period of time. Free online dictionary defines income as the amount of money or its equivalent received during a period of time in exchange for labor or services, from the sale of goods or property, or as profit from financial investments. The same source alternatively describes income as money received by a person or organization because of effort work or from return on investments.

Diversification means addition of livelihood sources of income other than those of farm related ones. It is the most important way of reducing rural poverty and increasing household income (Hengsdijk *et al.,* 2007). According to Brugère *et al*. (2010), diversification is defined as the process by which rural households construct a diverse portfolio of income generating occupations in their struggle for survival and in order to improve their standards of living. The term income diversification is mostly used interchangeable with livelihood diversification simplify analysis and interpretation. However, Ellis, (1999) makes a distinction between the two and defines income diversification as the composition of household income at a given point in time while livelihood diversification is considered as an active social process involving engagement in increasingly complex portfolio of activities overtime.

Therefore, from existing literature the term "income diversification" can be described using five distinct but related concepts. According to Minot *et al*. (2006) income diversification defined from two important point of views. Firstly, income diversification considered as a means of raising the number of income sources or soothing the multiple sources of income for rural households. Secondly, income diversification can also be exists when an individual firm transfer from low-value crop production to high-value crop production.

Thirdly, when the rural households change from small-scale or subsistence crop production to industrial production also constitutes income diversification progress (Delgado and Siamwalla, 1997). Fourthly, according to Reardon (1997) indicates an extension of income source to non-farm economic activities is considered as the third concepts of income diversification. Finally, Teshome and Edriss, (2013); Idris-Adeniyi et al. (2020) were used quantitate number to define the income diversification status of the rural households. Accordingly, an individual rural household is said to be income diversifier if and only if his or her quantitative status of income diversification is greater than 0.05%. Therefore, this study adopted the diversification concept suggested by Teshome and Edriss, (2013); Idris-Adeniyi et al. (2020), in addition to the concept raised by Reardon (1997).

2.1.2. Approaches of diversification analysis

According to Barrett et al. (2001) there are different methods for diversification analysis empirically. Like: - asset based approach, activity based approach and income based approach. Assets are part of wealth that directly or indirectly generate income in the form of cash or in-kind returns. In portfolio theory, on which the diversification literature is based, assets are emphasized as objects of agent's choice for the sake of income maximization, risk minimization or both. Accordingly, assets can be chosen as a means of diversification analysis and a number of authors have used assets to characterize and study diversification (Yizengaw, 2014).

However, asset based approach of diversification has their own drawback. First, a productive asset cannot always be allocated to a particular activity instead of being used across activities, so it is relatively hard to sum up assets in a single activity. Second, calculating the true value of some assets is difficult due to insufficient development of asset markets in developing countries.

Accordingly, it is argued that in order to study asset diversification, assets must be treated as a vector of physical quantities rather than a single, money-metric aggregate. This is, in turn, unable to fully capture patterns of diversification of households across all income-generating and unearned income options (Barrett et al., 2001).

As an alternative, activity can be used to study income diversification but it also has some drawbacks, according to Barrett and Reardon (2001). First, although we can identify which assets are used in each activity, they cannot be fully valued as mentioned above. Therefore, as in the case of assets, activities cannot be aggregated in monetary value and hence cannot be used to examine diversification patterns. Second, if based on activities, unearned income sources are completely ignored. This may lead to an incomplete understanding of the relationship between diversification and poverty reduction.

However, if researchers purposively ignore unearned income sources and define diversification as participation in income-generating activities, activity diversification can be adopted as a suitable measure. Thus time allocated to, or income earned from each activity may be used to analyze diversification. Unfortunately, another weak point is that reported employment share of nonfarm activities is believed to be understated (Yizengaw, 2014). Due to non-farm activities are widely recognized to offer supplementary work during slack periods of the agricultural cycle, real working time allocated to those activities is often unintentionally added to the total account of agricultural employment, the primary source of income of farmers. Therefore, this result for an underestimate of the actual proportion of labor time that is allocated to non-farm activities (Barrett and Reardon, 2000)

To overcome the limitation of asset and activity based approaches, income approach was employed in empirical work on diversification analysis. Using income may offer several merit. First, since the two main motives of income diversification are maximization of income and stabilization of income, or both, discussing diversification in terms of income diversification appears to be a natural candidate (Barrett et al., 2001). Second, income is the end outcome of income-generating activities, to which both productive and non-productive assets are allocated, and of unearned income options, for example transfers, as well. It is also easier to convert in-kind payments into a money-metric due to higher development of goods market compared with asset market (Barrett and Reardon, 2000). Third, income is closely related to the concept of

absolute poverty as it is more or less used to define the poverty line and measure household wealth (Yizengaw, 2014). Because of these reasons I seems that, defining diversification in terms of income may be the most suitable approach (for a review of various empirical studies that used income approach). As a result, this paper employee income based approach to analysis diversification

2.1.3. Reasons for income diversifications and classification of HH income sources

For poor farm households, participating on non-farm/off farm income diversification activities are important than specialization (produce one product or services efficiently) income generating activities (Adem et al., 2018). This is due to poor farm house households motivated by different factors.

According to Khan et al. (2019) income diversification can be driven by different motivations. Development economics literature has identified two main factors that drive diversification into non/off farm activities among farm households in developing countries. These factors are broadly classified into pull factors and push factors. Pull factors will attract households to the nonfarm sector when the nonfarm activities offer higher returns compared to farming. Reasons why a farm household can be pulled into the non-farm/off-farm sector include expectation of higher returns and the less risky nature of investment in the non/off farm sector (Kilic et al., 2009). The desire to increase income in order to become more food-secure, upgrade housing, educate children, accumulate assets or otherwise improve the household's standard of living are also the pull factors.

In the literature there were many different systems in classifying sources of income. Following (Davis et al., 2009; Benjamin et el., 2013) income sources are allocated into five basic categories: (1) agriculture; (2) informal employment, (3) formal employment; (4) self-employment; (5) remittances. The five categories of income are aggregated into higher level groupings depending on the type of analysis (i.e., on-farm, off-farm and non- farm). On-farm income is refers to income generated from either farming own-land or land acquired or accessed by cash or share tenancy, and income from livestock production (Onadeko et al., 2020).

Off-farm income refers to the activities either individual rural households are wage employment or self-employment activities which derive income for the household's labour supplies outside

their own farm. Wage employment include, farm wage, skilled and unskilled regular wage and casual daily work. In the other hand, self-employment off-farm income comprises from selling firewood, stone mining, grain and livestock trading, petty trading, weaving, mat making, pottery and hand craft. While Non-farm income also one of the sources of income in rural household which rise from outside farm and off-farm source of income. For example, income derived from remittance from relative and friend not present living with the household, pension, gift, renting out asset, and government aid (Kassie et al., 2017). Therefore, in these study three types of income categories was discussed: on-farm, non-farm and off-farm income based on supported literature. Generally classifications of household source of income are explained by the following diagram.

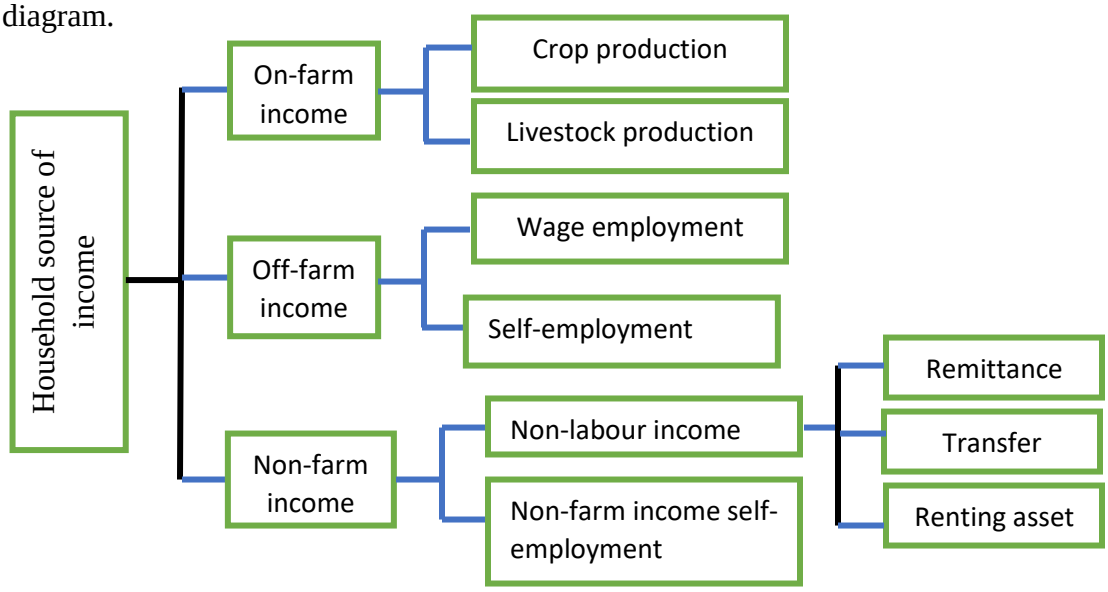


Figure 2.1 Classifications of Household Income Sources

2.1. 4. Relationship between off-farm and on-farm activities

From panel data on Slovenian farmers, (Bojnec and Fert, 2011) examined the impact of off farm income on farm income. Their result revealed that off farm income increases production of farmers over time, showing spillover effects of income from off farm activities on farm activities. Similarly, Babatunde and Leliveld, (2013) also studied the link between on farm works and off farm works in rural Nigeria. In his study, he examined the impact of off farm income on farm level output, purchased inputs and production of farmers. The finding of his research confirmed positive associations between off-farm income and output and purchased inputs.

In contrary to this, typology of non-farm and off-farm income generating activities that were exerting positive impact on farm production was not identified. Rural nonfarm income is likely to have a positive impact on farm activities in cases where the rural markets do not function properly. This is in line with the arguments given by Kilic et al. (2009) who suggest that "non-farm earnings may help households overcome credit and insurance market constraints by providing liquidity that can be utilized for productivity enhancing input purchase and long-term investments in agriculture." Participating in non-farm activities could increase overall cash income. Several articles show a positive effect of off farm income on the use of purchased inputs, for instance: Davis *et al*. (2009) from Kenya; Maertens (2009) from Senegal. Hence, in most developing countries farm households highly reliant on off farm income and that can have good implications to be considered by agricultural research and extension Participation in Non-farm Activities

2.1.5. Importance of the non-farm activities

In Ethiopia, even if agriculture is the dominant sector where many farm households make a living, rural nonfarm and off farm activities also play significant role in employment creation, income generation and enhancing farm production activities (Beyene, 2015). In developed countries like the United State that nonfarm work has risen steadily and has become the most important component of farm household income, far more important than farm income. However, the impact of nonfarm income is not that straightforward in developing countries. It is assumed that farm activities remain important in rural households as they provide the main source of employment and income in rural areas of developing countries Kilic et al. (2009) review empirical evidence on the RNFE in a number of developing countries.

One of the major reasons for a rural household to diversify into the rural nonfarm economy is to minimize risk: under imperfect insurance market, nonfarm and off farm income has a great role to minimize the risk of farm activities. The agricultural investment effect of nonfarm/off farm income diversification is particularly important for poor farm households. This is because lack of liquidity and poor access to credit are the most pressing constraints to improve agricultural productivity among farm households in developing countries Davis et al. (2009).

2.1.6. Female-headed households and livelihoods diversification

Gender is an integral and inseparable part of rural livelihoods. Men and women have different assets, access to resources, and opportunities (Ellis, 1999). The female-headed household's income diversification options would be determined by the specific constraints linked to gender norm, labor, land, human capital. Because female-headed households are generally smaller, they sometimes lack sufficient labor to diversify their income sources (Elin, 2015). In addition to labor, female-headed households, and women in general, have less access to productive resources. In his seminal paper about gender differentials in farm productivity, (Udry, 1996) point out that the productivity of plots farmed by women is proportionally lower than that of plots farmed by men due to lower efficiency of female. Furthermore, while women are also less likely to own land, those who do often own plots that are smaller and lower quality than those owned by men. In 9 of 14 countries (in the RIGA database), landholdings were smaller among female-headed households than among male-headed households (FAO, 2020)

In addition to lack of access to productive resources, female-headed households sometimes lack access to credit. This is especially important if the household wants to diversify into self-employment. In the above-mentioned survey, rural, female-headed households were less likely to use credit in seven of nine countries (FAO, 2020). (Reardon, 1997) concludes that female households are sometimes constrained to businesses with low start-up costs.

Furthermore, access to high-return off-farm wage employment requires a certain level of human capital (Woldenhanna and Oskam, 2001; Bigsten and Tengstam, 2011) . In addition, even if the gender gap in education was decreased in most less developed countries, female-head households generally have lower levels of education than their male counterparts (FAO, 2020). They might also lack the connections needed to access these forms of employment, and sometimes, social norms regarding female labor force participation hinder women from entering the labor market (Mammen and Paxson, 2000) . In Ethiopia, the gender gap in education has narrowed, and in 2019, there were as many girls as boy in primary education. However, in secondary education, only 93 girls were enrolled for every 100 boys (WB, 2019).

Because of the special constraints faced by women, we expect that female-headed households are less likely to report off-farm wage work or to operate their own business than male-headed

households. As a result, we believe that female-headed households face constraints that reduce the probability of diversifying into off-farm activities; we expect them to be more reliant on the agricultural sector, farming full time and more of their time consumed by homework. However, female-headed households are a heterogeneous group with different reasons for becoming female-headed. In households where there is a man present, he may be able to help the household to overcome some of the constraints listed above. Even if the man does not live in the household (for example a husband living elsewhere because of work), he might help the family with access to productive resources such as land and credit (Elin, 2015).

Previous research has also found that women and men differ in degree of risk aversion, which could influence their optimal level of diversification (Eckel and Grossman, 2008) and (Croson and Gneezy, 2009) summarize a large amount of research and conclude that evidence from field studies generally show that women are more risk averse than men. This finding is supported by evidence from experiment-studies, while the evidence from contextual environment is more mixed. However, most of these studies are conducted in developed countries. If this holds also in our setting, we would expect female-headed households to demand higher levels of diversification than male-headed households.

2.1.7. Measuring income diversification

Most finding employee either univariate or multivariate regression models to estimate the determinants of income diversification with mainly income diversification as the dependent variable and is regressed against a set of independent variables. Accordingly, the extent of household livelihood diversification in the literature is commonly quantified using income diversification (Yizengaw, 2014). There are two approaches used to measure income diversification in different literature. First, using the vector of income shares associated with different income sources (e.g. Barrett et al., 2001). Secondly, using different diversification index, like (i) Index of maximum proportion, (ii) Herfindahl Index, (iii) SID, (iv) Ogive Index, (v) Entropy Index, (vi) Modified Entropy Index, and (vii) Composite Entropy Index (Debesai, 2020).

In this paper, to measure income diversification, I was used on-farm, off-farm and non-farm income share in total income of female-headed households, to gather with SID. Since, SID has

the following advantages over the other methods. SID has the advantage of estimating both the number of household's income source and the contribution of each income source to total household's income (Balense and Debebe, 2020).

2.1.8. Impact of income diversification

Income diversification has both positive and negative impacts on farm households livelihood (Yizengaw, 2014). These effects are discussed in detail below.

Positive impact

Seasonality: - Income diversification can contribute to reducing the adverse effects, by utilizing labor and generating alternative sources of income in off-peak periods. It would be misleading to see the growth in rural nonfarm income in isolation from agriculture as both form part of complex livelihood strategies adopted by rural households

Increase income: - Diversification promotes making better use of available resources and skills (as in seasonality above), and taking advantage of spatially dispersed income earning opportunities.

Minimize risk: - Diversification enables spreading of risk across different activities whereby factors that create risk for one income source are not the same as those that create risk for another.

Gender benefits: - Where activities are equally or better accessed by women, it is possible for diversification to improve the independent income-generating capabilities of women and in so doing, also improve the care and nutritional status of children.

Environmental benefits: - Diversification can potentially provide environmental benefits by providing options that make time spent in exploiting natural resources.

Asset improvement: - Cash resources obtained from diversification may be used to invest in or improve the quality of household assets.

Negative impact

There are some demerits of income diversification examined in empirical studies. Those are:-

Reduced agricultural output:- Some types of diversification may result in stagnation on the home farm especially when there is lucrative distant labor markets for male labor, resulting in depletion of the labor force required to undertake peak farm production (Ellis, 1999)

Adverse gender effects: - Where it is male labor that is predominantly able to take advantage of diversification opportunities, then women may be even more relegated to the domestic sphere and subsistence food production. (Baiphethi et al., 2008) suggests that "one of the major impacts of livelihood diversification is feminization of agriculture, as men frequently pursue migratory labor opportunities." Consequently, women remain home to tend to home gardens and other agricultural tasks to ensure food production for the household. The empowerment of women may yield positive results as women are more likely to invest the additional income in children and family (Ellis, 1999)

Income inequality: - Diversification can be related with broadening the difference between the incomes of the rural poor and the rich. This happens if the better- off are able to diversify in more advantageous labor markets than the poor.

2.1.9. Impact assessment methods

The adoption of several methodologies ranging from simple to more advanced approaches has been expected for impact evaluation. In theory, evaluators could develop control and treatment groups using one of three methods: randomization/pure experimental design, non-experimental design, or quasi-experimental design. In practice, in the social sciences, the choice of a particular approach is influenced by a variety of factors, including data availability, cost, and the willingness to experiment ethically. Propensity score matching (PSM), differences in differences (DD), regression discontinuity design (RDD), and instrumental variables (IV) are the most commonly used non-experimental/quasi-experimental design methodologies for evaluating development programs (Hap *et al.*, 2006).

Difference-In-differences (DID): The method of difference-in-differences contrasts a treatment and evaluation group (first distinction) before and after a mission (second distinction). When propensity scores are utilized, comparators must be discarded if their scores fall outside the treatment group's range. However, only a random subset of these people is allowed to participate in the task. The counterfactual is made up of diagnosed people who do not participate in the project (Jalan and Ravallion, 2003).

Regression discontinuity design: Designing a regression discontinuity when program participation is determined by an explicitly special exogenous rule, a method can be used. The

approach is based on the assumption that persons on both sides of the eligibility cut-off point are comparable, and it uses those on the other side of the cut-off point as the counterfactual. The RDD method's main technical flaw is that it only evaluates the program's marginal impact at the eligibility cut-off point; therefore nothing can be said about people who are further away from it. Furthermore, for the RDD estimate to be valid, a threshold must be applied in practice, and persons must be unable to manage the choice score to be eligible (Hap *et al.*, 2006)

Instrumental variable: The instrumental variable method employs one or more variables that influence involvement. However, rather than identifying the exogenous variant in outcomes as a result of the program recognizing that its placement isn't always random but purposeful, this identifies the exogenous variant in outcomes as a result of the program recognizing that its placement isn't always random but purposeful. After using the instrumental variables to forecast program participation, the final result's indicator is compared to the expected values (Baker, 1999). Ordinary least square (OLS) techniques limit the model to a linear purposeful form, suggesting that the coefficients at the manipulating variables for the treatment and control groups are similar (Ali and Abdulai, 2010).

Propensity score matching: In the absence of baseline survey numbers for impact assessment, the propensity score matching method has been widely used as an impact evaluation tool. The method is automatically appealing because it makes evaluating the observable consequences of the counterfactual control institution's impact easier (Heckman *et al.*, 1998). In the lack of experimental statistics, propensity score matching was appropriate. As a result, this study used propensity score matching to examine the influence of income diversification on female-headed household livelihood indicators (income, consumption expenditure, and saving) in the study area.

2.1.10. The relationship between income, consumption and saving

Income is a monetary value that human generates in business world either to consume or save. Accordingly, many author set the linkage between income, consumption and saving using identical formula:-

$$Y = C + S \quad \text{_____} \quad 2.1$$

Where Y = disposable income, C = consumption spent, and S = saving (Takala, 1995). Moreover, in case of simple economy (there is no government and external sector) using the

above identical equation, the private sectors receive disposable income which is a whole of income used for consumption and saving (Dornbusch and Fischer, 1977).

In addition to above, the theories of consumption, Fisher's intertemporal model briefly discussed the relationship income, consumption spent and saving. The consumption function introduced by Keynes relates current consumption to current income. This relationship, however, is incomplete at best. When people decide how much to consume and how much to save, they consider both the present and the future. The more consumption they enjoy today, the less they will be able to enjoy tomorrow. In making this tradeoff, households must look ahead to the income they expect to receive in the future and to the consumption of goods and services they hope to be able to afford (Santos, 2013).

According to (Mendershausen, 1939) there is positive relationship between average income of the household and net saving. If income raises further, the families show net savings; the percentage of income saved increases with income, but the rate of increase becomes smaller and smaller, first rapidly, then more slowly. Since savings as a percent of income are increasing over the whole income range considered, the absolute amount of savings must rise too.

Moreover, according to the consumption theories and the risk theories, income diversification can affect consumption in two ways. Firstly, diversification can increase the peasants' income, and then encourage them to consume more. For rural households, income diversification means they can gain access to more income sources than farming, and the aggregate income may be enhanced more or less (Xu, 2017). Empirically, there is also a consensus on the potential positive effect of income diversification in raising the rural households' income (Barrett *et al.,* 2001). Secondly, income diversification can mitigate the rural households' income risk, and stabilize their expectation and promote their consumption (Alderman and Paxson, 1992).

Existing theoretical finding summarized that multiplying income sources can effectively protect the farmers from being negatively influenced by income shocks, which has virtually performed the function of risk management with portfolio of occupations (Alderman and Paxson, 1992). The empirical evidence has also been found that income diversification can complement the peasants' income and, hence, mitigate income risk (Xu, 2017). On the other hand, the consumption theories state that the households' risk attitudes are related to their income

(Dohmen *et al*., 2011). The households with higher income will definitely have greater abilities to deal with risk from those with lower income (McKenzie and Woodruff, 2006). Thus, the role of income diversification in consumption depends on the household's income, as well.

2. 2. Empirical Literature Review

2.2.1. Empirical literature on determinants of income diversification decision and its intensity

The study conducted by Idris-Adeniyi et al. (2020) on determinants of income diversification among arable crop farmers in Osun state, Nigeria by employing Tobit regression model to analysis data. Beside this, the researchers used diversity index (Herfindahl-Hirschman diversity index) to measure degree of income diversification in study area. Therefore, their study point out that the mean income diversification index in the study area was 0.46 which implies medium level of income diversification of farmers). Only 6.67% of the farmers had income diversity index of 0 meaning that most of the respondents adopted multiple income generating activities while crop farming remained their dominant income source. Factors like: credit, household size, frequency of extension visits are positively affect the income diversification, whereas age of the household affect income diversification negatively. In the face of climate change and its attendant risks including total crop failure, farmers should be exposed to other viable farm and off-farm income generating activities, while they are provided with credit facilities to harness such opportunities.

Debesai (2020) also employed Tobit regression model in his study to analysis rural household income diversification in developing countries: A Case study of Eritrea. Beside this, the researcher adopts Simpson Index of Diversification to calculate the diversity status both at household and regional level. Thus, Tobit regression model indicated that level of education, ethnicity, household size, gross income, income per capita, and access to credit had a positive relationship with income diversity, while age of the household heads, dependency ratio, and size of land ownership had a negative relationship. In addition to this, the study implies income diversification was pervasive in all regions and households irrespective of income levels. Although income diversification was more in high income groups, it was also substantial with the "extreme" and "low income" groups. The difference was that low income groups diversified

in less risky ventures as a necessity, whereas the high income groups diversify even in more risky ventures as a choice

The study conducted by Balense and Debebe (2020) in Ethiopia employed Multivariate Probit, and Two-limit Tobit models to analysis income diversification and livelihood strategies among pastoral and agro-pastoral households' in southern Ethiopia. Accordingly, the result of model reveals that factors influencing the use of livelihood strategies are age, sex, family size, educational level, farm size, market distance to the main market, livestock holding size, cooperative membership, use of credit and access to transport were as sex, family size, educational level, livestock holding size, crop failure, a distance of nearest main market income from farm and share of non & off-farm income influenced income diversification. All of the variable except crop failure, main market distance and share of non-farm and off-farm income positively affect income diversification and its intensity.

Khan et al. (2019) conducted the study on determinants of income diversification of farm households in Uttar Pradesh, India. They employed logit regression model to identify the determinant of income diversification in the study area. Age, education, social category, family, land size, access to market, access to finance, information sources, access to organized input supply and perceive weather risk were explanatory variables included in the model. Therefore, the econometric result revealed that education, family size, land size, proper infrastructure for livestock, adequate production technology, information sources, access to market, and climatic risk are positively and significant variables which affect income diversification.

Alobo and Bignebat (2018) used Tobit model in order to identify the determinants of Income diversification in rural Senegal and Kenya using data on 1,747 farm households collected in 2007-2008 from six regions in rural Senegal and Kenya. The empirical finding point out that the regional difference in income diversification does not follow any clear patterns, with push and pull determinants acting concurrently within and between regions. More generally, income diversification is significantly related with household asset endowments, demographic factors, accessibility to rural towns, migration opportunities, and perceptions on food security.

According to the study conducted by Gebiso et al. (2019) on the rural livelihood diversification status and determinant factors in Arsi, Ethiopia using negative binomial regression model and

double-hurdle model to identify determinants of non-farm activities and to identify factors affecting participation in non-farm and amount of earnings respectively. In addition to this, Simpson diversification index was used to estimate the diversification status. According to the result of negative binomial regression model, there demographic variables which are age, gender, and education level were significant. Age and education household were positively affecting income diversification while being maleness affects income diversification negatively. In the other hand, from socio-economic variable land holding and market distance were negatively income diversification. In addition to this, in case of double-hurdle model age and education level of the household positively affect income participation probability while Gender (Being femaleness affect the participation probability negatively. Similarly, the amount of income gained from non-farm income was affected by different variables. Size of cultivated land, educational level and crop income were affect amount income gained from non-farm positively and significantly and negatively by gender.

Most studies in the area of off farm/nonfarm income indicated that, farm characteristics of the household are considered as main factors determining the decision of participation in off/nonfarm activities. For example, using data on 200 households selected from 40 villages of Southeast Nigeria, (Babatunde and Leliveld, 2017) examined factors determining nonfarm income. Their findings show that age of the household, education level and farm size are the most significant variables determining both farm and off farm income.

Ababbo, (2015) studied factors determining the decisions to participate in off farm work in western Ethiopia. The finding of the study shows that variables on access for credit and size of farm land are major determinants of decisions to participate in off farm activities. There is widespread agreement that smallholder farmers require improved access to agricultural markets to raise their farm productivity and living standards (Majbauddin et al., 2020). The causes and consequences of diversification are differentiated by location. Better infrastructure such as roads is linked to higher opportunities for farm and nonfarm employment and to increased agricultural production (Salam et al., 2019)

2.2.2. Empirical literature on impact of income diversification on the livelihoods

The study conducted by Bojnec and Knific (2021) on farm household income diversification as a survival strategy in selected hilly and mountainous areas in Slovenia before and after the

accession to the European Union, by employing descriptive statistics like t-test and F-test. Therefore, the result of the study point out that diversification of income from self-employment is important for more than one-third of households that maintain agro-food production for the market. Income from self-employment is an important source of income for household well-being and for investment in agricultural production to improve incomes from farming activities. Expansion of self-employment impacts the lack of time, business risks, and lack of interest of households to expand the business by renting external sources.

The stud by Ehiakpor *et al.*(2020) on the impact of rural non-farm income diversification on households' welfare and adoption of Zai-technology (a proxy for agricultural technology adoption) using cross sectional data collected from agricultural households in the Upper East region of Ghana. They employed Propensity Score Matching (PSM) and Inverse-Probability weighted Regression Adjustment (IPWRA) model to estimate welfare and Zai-technology effect of non-farm income diversification. After controlling for differences in covariates, the results show that non-farm income diversification increases the likelihood of Zai-technology adoption and contributes to significant household welfare gains. We therefore suggest that the activities of agricultural extension services and farmer based organizations be enhanced as they facilitate the diversification of non-farm incomes, thereby increasing investment in productivity-enhancing technologies (Zai) and household welfare.

Ghosh and Sujan (2020) also conducted study on changing pattern of rural income diversification and its impacts on household economy: A study on a selected village in Jashore District, Bangladesh. They employed descriptive statistics to analysis the data collected from the respondent. Therefore, the result shows that the average rural household income (BDT 16,983 per month) has significantly increased over the time. In 2017, more than 61% of sampled households had been earned above the national rural income level of BDT 13,353 per month. Increases in diversification of income sources were result for enhancement of rural household income.

Moreover, some empirical findings indicate that income diversification has a positive effect on household livelihood while others argue that income diversification has a negative impact on participant household livelihood. For instance, Salam et al. (2019) also carried out study on the Impact of income diversification on rural livelihood in some selected areas of Bangladesh, using

the Two Stage Least Square (2SLS) methods with instrumental variable were applied to estimate impact of the strategies on household welfare. Their study shows that involving in any type of non-farm activities jointly with farming has a significantly positive effect on the household's welfare.

Similarly, Vien (2017) had evaluated the determinants of income diversification and its effects on rural household income in Vietnam by using the Poisson and Tobit regression methods to analysis the data. The regression results point out that socio-economic factors have strong influence on household income diversification in the rural areas, and, in turn, income diversification has positive impact on household income growth. It implied that income diversification is an important strategy to improve households.

On the contrary, despite its popularity, other studies on income diversification show that income diversification has a limited and negative impact on the life of poor people. They argue that income diversification has a negative impact and harm the client instead of improving their livelihood. For instance, Omotesho et al. (2020) studied the effect of income diversification on the rural farming household in Kwara state, Nigeria using Descriptive statistics and Pearson's Product Moment Correlation were used for data analyses with respect to factors such as age, sex, education level, average income, farm size, extension contact, household sizes and primary occupation of the household In Nigeria. Therefore, the study concluded that the more diverse farmers' income, the lower their livelihood status. It is recommended that the number of their economic activities is kept at such that can be effectively managed for positive contributions to the livelihood of farming households.

In Ethiopia also, different studies have been conducted by different researchers on the impact of income diversification on the livelihood of the households and their determinant through different livelihood indicators get different findings. For example, according to Adem et al. (2018) study conducted on income diversification and food insecurity in the farm households of Ethiopia. They found that income diversification strategies have significant influence on food insecurity. Income diversification strategies of the households involved in a combinations of crop production with livestock rearing, crop production with off-farm activities, off farm activities with livestock enterprises and crop production only, at 60%, 10%, 8%, and 22% of households respectively.

In the same way, the study conducted by Astatike and Gazuma (2019) used Heckman two-step model to analysis the impact of participation in off-farm activities on household income level in Wolaita zone, southern Ethiopia. Accordingly, the result of Heckman's two-step model indicated that participation in off-farm activities has a significant impact on annual income of the households. Moreover, the study identified the factors that affect participation of households in off -farm activities. Consequently, the probit model result revealed that age of the household head, sex of the household head, marital status, distance to the nearest market, agro-ecological zone, credit access, livestock ownership, the amount of farm income, and having mobile phone were key factors that influence participation of households in off-farm activities. With regard to results of Heckman's two-step model, the participation in off-farm activities, education level of the household head, and agro-ecological zone were found to be the main factors that affect households' total income level.

According to study conducted by Etea et al., (2019) they used binary logistic regression and ordinary least squares (OLS) estimation methods were employed to determine the effect of household income diversity on food security in the Ambo district, Ethiopia. In this study, the instrumental variable (IV) method was employed to over-come an endogeneity bias. Simpson's index of diversity (SID) and daily calorie consumption were used to measure the level of household income diversity and food security respectively. In this study the following explanatory variables were included to analysis the effect of income diversification on the food security. Those are; sex, age, education, family size, dependency ratio, credit services, distance to market, training or special skill, farm land size and total income. Therefore, the result of the study showed that, the degree of household income diversification was minimal, and the majority of households were food insecure in the study area.

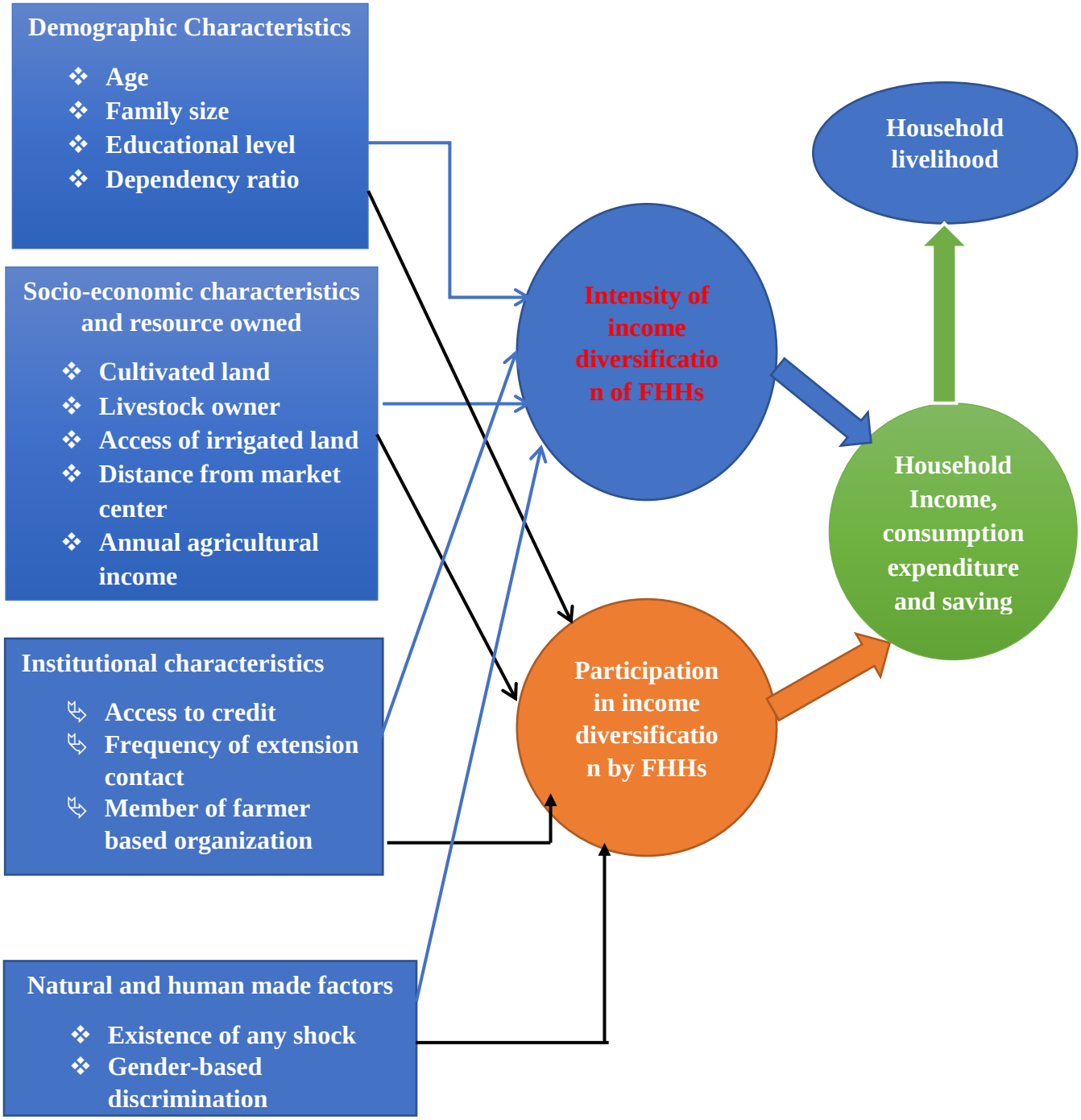
Similarly, (Teji, 2020) provides empirical evidence from rural Ethiopia that poor households rely more on off-farm activities while the rich earn more from agriculture. His model result on determinants of participating on off-farm activities however confirms that household size and level of education has a significant and positive influence on income diversification. Additionally, study in Southern Ethiopia (Adem et al., 2018), again in West (Gebiso et al., 2019). These studies explain that educational status represent household human capital endowment; an increase of which will strengthen the ability of engaging in other livelihood options.

Generally, the above empirical evidence reviews show that the determinate, degree, and impact of income diversification. All of the studies were focused on household heads which lead by either male or female rather than separately examine. Hence, this study was conducted with detailed information about income diversification, specifically on female headed household income diversification, and its impact on household income per capita, consumption expenditure per capita, and saving per capita in the study area. Thus, this study intended to fill this gap.

2.3. Conceptual Framework of the Study

The conceptual framework of the study is based on the empirical analysis. It links factors which affect income diversification participation decision and intensity of income diversification of the household. Choice of rural farm households to engage in farm, nonfarm and off farm activities depend on demographic, socio-economic, institutional and other factors which allow farm households to engage in income diversification. A demographic factor consists of age, education, family size, and dependency ratio, while Socio-economic characteristics and resource owned are cultivated land size, livestock owner, access of irrigated land, distance from market center, and annual agricultural income. The institutional characteristics of the household are access to credit, frequency of extension contact, and member of farmer based organization were included. Lastly, natural and human made factors are included like existence of any shock and gender-based discrimination as shown in below figure 2.1.

In addition, the ideas in income, consumption and saving theory are used to define the relationship between income diversification and livelihood indicators, as shown in Figure 2.1. Female-headed household income, consumption expenditure, and saving are considered as an outcome variable that affected by the income diversification and other independent variables.



Source: Modified from Alemayehu (2020)

Figure2.1. conceptual framework of the study

CHAPTER THREE

3. RESEARCH METHODOLOGY

This chapter presents the methodology employed to address the objectives of the study. It consists of six sections. The first section describes the study area in which the study conducted. The second section describes the data types and source of the data. The third section describes the mothed of the data collection. The fourth section describes sampling procedure and sample size determination. The fifth section describes methods of the data analysis (descriptive statistics, inferential statistics, and econometric model) and the last section of this chapter presents hypothesis and variable definition.

3.1 Description of the Study Area

This study was conducted in Guduru district of Horo Guduru Wollega Zone in Oromia region. Guduru district is one of the districts of Horo Guduru Wollega Zone and the administrative center of the district is Kombolcha town. It is located at about 331 kilometers distance from Addis Ababa, the capital city of Ethiopia, to the western direction of the country. Today, it is sub-divided in to 25 farmers associations. Guduru district is bordered on the south by Jimma Rare, on the southwest by Jimma Horo, on the west by Fincha river, on the northwest by Abay Chomen, on the north by the Abay River which separates it from the Amhara Region, and on the east by the Guder River which separates it from the Mirab Shewa Zone.

This district is divided in to three distinct geographical areas with different proportions; namely the high land 0.26 percent which is very small part of the district, midland 46.74 percent and the low land 53 percent. The district is situated at an altitude of 1350 to 2450 meters above sea level. The dominant climatic condition is a sub- tropical type. As a result, this area is experiencing mean annual temperature of slightly greater than 15℃ and mean annual rainfall of 1600 mm to 2000 mm (DADO, 2020)

Clay loam is among the soil types found in the district i.e., it covers 16.33% of the total land of the district, Sandy soil covers 55,734.60 hectares of land which is about 23.06% of the total land of the district. The other soil type exists in the district is loam soil, dominantly found in the district, which is good potentiality for agriculture and covers 42.80 % of the total land of the district (District Agricultural Development Office, 2020).

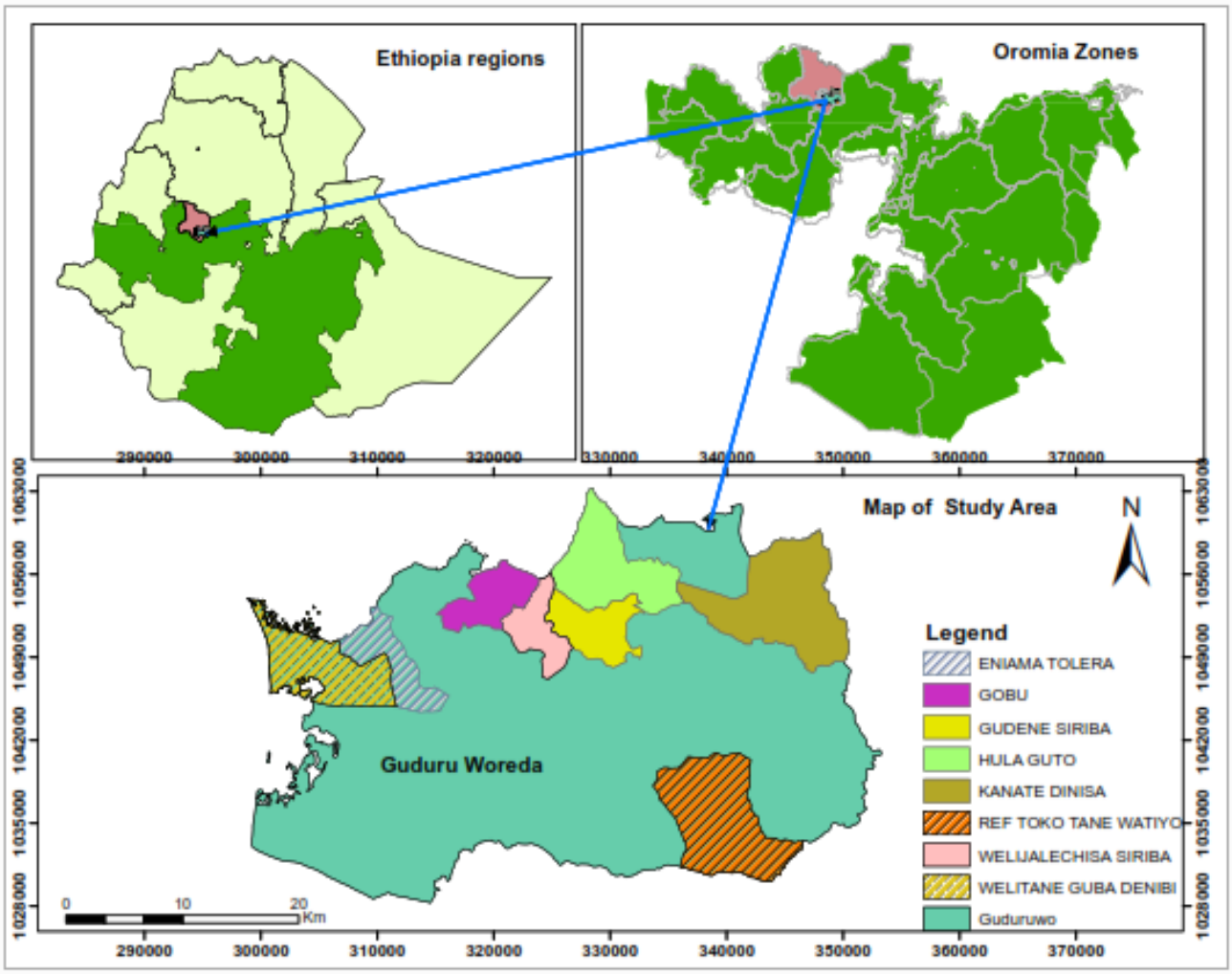


Figure 3.1: Location map of the study area

The 2007 population and housing census result is the base of population projection all over the country. Based on this census result, the population of Guduru district is projected to be 106,798 of whom 53,714 are male and 53,084 are females. The percentage of urban population for the district doesn't include. The majority of the inhabitants follow Protestantism, with 53.11% reporting that as their religion, while 30.16% observed Ethiopian Orthodox Christianity.

3.2. Research Design

In order to conduct a research on the study area, the researcher adopted a cross-sectional survey. The targeted population was female-headed households those who are participants (treated) and non-participant (control) of income diversification, in the Guduru district and unit analysis was female-headed household. For sampling techniques, multi-stage sampling technique was used to

select representative sample for the study. Data was analyzed using descriptive statistics and econometric model. Under econometric, double-hurdle model (probit and truncated) and PSM were used. In general, this section focuses on a discussion of the research process, the selection of the data sites, sampling methods, type of data, data collecting and analyzing mechanism (analytical and statistical techniques used in analyzing the data for the study).

3.3. Source and Type of Data

Both quantitative and qualitative types of data were collected from both primary and secondary data sources for the purpose of study. Primary data was collected using interview questionnaire specifically structural questionnaire on a variety of respondent demographic characteristics and socio-economic variables. The survey schedule was designed in such a way to capture the necessary information on household level of livelihood indicators, demographic and socio-economic variables based on the objective of the study. The study also supplemented by secondary data source like reports of CSA, FAO, different journals, unpublished materials and data from agricultural development office were used in this study.

3.4. Methods of Data Collection

The data used for this study was collected in 2021. Since agricultural output was not sold at all during the collection of the data, the study used the agricultural output income of 2020 year of production. Prior to data collection, the enumerators were trained regarding on how to collect data from sample households using structural questionnaires through interview. The household questionnaire was first pre-tested to select female headed households, and appropriate modifications were made based on pre-tested on results before actual data collected. Then, the enumerators under supervision of researcher collect the data from the selected sample of rural female-headed households.

3.5. Sampling Technique and Sample Size

The target population in this study was rural female-headed household in Guduru district. Multi-stage sampling technique was used, so as to reach the selection of a sample of female headed rural household participants and non-participants of income diversification in the study area. In the first stage, from the total of 13 districts of Horo Guduru Wollega Zone, Guduru district was purposively selected based on the fact that it is one of the most populous and the area is far from

research conducted. Beside this, it has relatively populous and opportunity to diversify income source is high than other district in the zone. Therefore, it was feasible to conduct this thesis in this district. In Guduru district, there are 25 rural kebeles and 2 urban kebeles. In all kebeles, income source is not only farm alone and 2 urban kebeles were not included in this study because they are not rural.

In the second stage, among those kebeles, eight rural kebeles were selected by simple random sampling method. Those kebeles are Eniama Tolera, Gobu, Gudane Sirba, Hula Guto, Kenate Dinisa, Ref Toko Tane Watiyo, Weljalechisa Sirba and Walitane. In the third stage, 245 samples of female headed households were selected by simple random sampling method from 634 total populations. Since kebeles differ in terms of the total number of sample households they encompass, probability proportional to sample size-sampling technique was employed to determine the number of households from each kebeles.

The sample size was determined by the following simplified formula provided by Yamane (1967). This formula was used to calculate the sample size from a given population at 95% confidence level and 5% precision level. Accordingly, sample size is estimated as follows:

$$n = \frac{N}{1+N(e)^2} \qquad (3.1)$$

Where n = sample size, N = population size, and e = the level of precision. The total number of female-headed household in the selected kebeles is about 634. Therefore, sample size was obtained by using the above formula is

$$n = \frac{634}{1+634(0.05)^2} = 245$$

Consequently, 245 female-headed households were used as a representative sample for the study, and this is considered as third stage sampling. To decide the sample size from each Kebeles, the sample size determination formula of (Israel, 1992) was used. That is ns= (Nh/Ns)*n, where: ns = sample size from each stratum, N h = total population in each stratum, Ns = Total population of the sum of strata for study, and n = Total sample size of population in the study, then determining the sample size of each selected kebeles proportionately as follows:

Table 3 .1 Determination of sample size from each stratum by applying the formula

Kebeles	Sample frame	ns	Proportion sample size from each stratum
Gobu	101	$\frac{101}{634} * 245$	39
Enama Tolera	96	$\frac{96}{634} * 245$	37
Walitane	56	$\frac{56}{634} * 245$	22
Guddanne Sirba	71	$\frac{71}{634} * 245$	27
Ref toko Tane	89	$\frac{89}{634} * 245$	34
Kenate Dinsa	67	$\frac{67}{634} * 245$	26
Weljalechisa	93	$\frac{93}{634} * 245$	36
Hula Guto	61	$\frac{61}{634} * 245$	24
Total			245

Accordingly, the total sample households selected from each kebele were as follow: From Gobu, Enama Tolera, Walitane, Guddanne Sirba, Ref toko Tane, Qanate Dinsa, Weljalechisa, and Hula Guto are 39, 37, 22, 27, 36, 34, 26 and 24 respectively. Finally, the sample households were stratified in to two stratums based on their degree of income diversification which is measure by SID. As a result, female-headed households those who have SID greater than 0.05 were diversified or treated group, while those who have less than 0.05 SID were non-diversified or control group. Based on this, 108 and 137 FHHs were selected under diversified and non-diversified group respectively.

3.5.1. Ensuring validity and reliability data

Ensuring the quality of data was the main objectives of the researchers. Hence to determine the validity of instruments, the researcher used face to face, and format questionnaires. Irrelevant items were modified as comments were given for instruments from advisors. Then English versions of questionnaires language translated into Afaan Oromoo local language orally to minimize confusion and communication barriers in data collections.

3.6. Methods of Data Analysis

The collected data was examined using statistical tools such as Statistical package for Social Sciences (SPSS) version 20 for data entry and STATA version 16 for data analysis. Beside this, Microsoft Excel was used to compute income diversification index as well as for the conversion of some variables into their appropriate figure. Moreover, two types of method of data analysis, namely descriptive statistics inferential statistics and econometric models were used to analysis the data collected from the sample household.

3.6.1. Descriptive statistics

The descriptive analysis is important tools to present research output clearly and concisely. Therefore, descriptive analysis was applied to examine demographic characteristics and socio-economic profiles of the female-headed households and performed using descriptive indicators such as frequency, mean, and percentages. In this study, descriptive statistics was calculated and arranged in the way that permits one to quickly understand their meanings. Furthermore, inferential statistics like t-test and chi-square were employed along with descriptive statistical tools in order to compare statistical significance of the mean difference between income diversified and non-diversified with regard to dummy and continuous independent variables.

Degree of income diversification

There are different methods that were used to measure diversification. Basically, to measure degree of income diversification, the following indexes are most frequently applied. These are Herfindahl Index (HI), Inverse Herfindahl Index (IHI), Ogive Index, Entropy Index, Modified Entropy Index, and Composite Entropy Index and many other uses Simpson Index of Diversity (SID). Therefore, this study prefers SID than other approach used to estimate the degree of income diversification among rural female-headed households in the Guduru district. According to Adem and Tesafa (2020) Simpson Index of Diversity (SID) is preferable than other approach due to it takes into consideration both the number of income sources as well as how evenly the distributions of the income between the different sources.

The value of SID ranges between zero (0) and one (1). Consequently, if the value of SID is zero indicates specialization of income source (only one source of income, where Pi = 1) and one the

extremity of diversification. Therefore, income diversification index called Simpson income diversification index can be calculated as follows:

$$SID = 1 - \sum_{i=1}^{n} P^2 \quad (3.2)$$

Where n = number of income sources, Pi = Proportion of income coming from the source ith source to the total household income obtained from all sources, and i = 1, 2 … n.

For this study, SID was defined as follow:

$$SID = 1- \left\{\left(\frac{CR}{ATI}\right)^2 + \left(\frac{LI}{ATI}\right)^2 + \left(\frac{LSAI}{ATI}\right)^2 + \left(\frac{SGI}{ATI}\right)^2 + \left(\frac{RNI}{ATI}\right)^2 + \left(\frac{RMI}{ATI}\right)^2 + \left(\frac{SEI}{ATI}\right)^2\right\} \quad (3.3)$$

Where CI = crop income, LI = Livestock income LSAI = Labour sale to agricultural income

SGI = Service giving income, RNI = Rental income, RMI = Remittance income, SEI = Self-employment income.

3.6.2. Econometric methods

Econometric analysis was used to examine the causal relationship between the dependent and independent variables.

3.6.2.1. Econometric model selection and specification

To analysis the second and third objective of the study, two econometric models, namely double hurdle and Propensity score matching model were employed. Double hurdle was employed to identify the factors that affect female-headed household participation decision in income diversification and its intensity, and propensity score matching model was used to estimate impact of income diversification on female-headed household livelihood (income, consumption expenditure, and saving) in rural areas of the selected kebeles in Guduru district.

Model specification for Double hurdle

Basically, Tobit, Heckman two steps, Heckman maximum likelihood, and Double hurdle model are alternatively employed when dealing with limited dependent variable model. Therefore, in order to analysis income diversification participation decision and intensity of diversification, the

econometric models mentioned above are possible alternatives. Accordingly, essential tests that verify which econometric model to use for the analysis were undertaken.

Tobit model was a statistical model proposed by James Tobin to describe the relationship between non-negative dependent variable and independent variable. The Tobit model is sufficient to accommodate the zero observed figures alongside other positive values if simultaneity of decision is assumed (Tobin, 1958). Tobit specification has its own limitation; first it is actually applied in cases where the dependent variable is not observed for some sample households due to censoring and not due to individual decision. This means Tobit specification can assume negative values, but will actually take zero for some censored observations. Therefore, all non-diversified are interpreted as corner solutions. Second Tobit specification is based on a restrictive assumption that both the decision to income diversify and level (amount) at which they diversify given that decision are determined by the same set of variables which implies that a variable that increases the likelihood of household to diversify income source will also increase the extent of income diversification (Burke, 2016 and Lidi, 2017). Therefore, double hurdle model is used as better alternative over Tobit specification. Despite its restrictive assumption, Tobit model has been used in the analysis of problem which comprise two steps (i.e., decision and intensity) as it was used by Ademe (2017) and Getahun (2018) upon comparing with the other models such as Heckman sample selection, Heckman maximum likelihood and Double hurdle model through performing statistical test.

In sample selection problems, some part of the dependent variable is not observed as a result of the outcome of another variable. Inverse Mills Ratio (IMR) computed from the first step probit selection model is included as explanatory variable in the second step regression to correct for self-selection and obtain unbiased, consistent, and efficient estimator (Wooldridge, 2002). The Heckman two stage sample selection model was used to determine factors affecting income diversification participation decision and intensity. Heckman has developed a two-step estimation procedures model that corrects for sample selectivity bias and participation and level of participation might be affected by different factors.

In order to overcome the limitation of above model, Double hurdle was appropriate model incase were two hurdle analyzed at the same time (Lidi, 2017). Double hurdle model also known as two tier models and it is corner solution model in which zero values associated with non-

participation are assumed to be outcome of rational choice (Wooldridge, 2002). It was initially proposed by Cragg (1971) and it is two-tiered process which incorporates relevance of the participation decision to the Tobit model with the probability of participation and the intensity of participation being determined by separate processes. As opposed to double hurdle model, Heckman (1979) model considers the zero observation to arise mainly from respondents' self-selection.

Practically, the difference between the Double hurdle and Tobit model should be evident in the data (Greene, 2012) which implies that the idea behind double hurdle is that looking at an event that may or may not occur. Occurrence of the event is associated with a continuous positive random variable. Whereas the random variables take value of zero if the event does not occur. The double hurdle model is an extension of Tobit model which relaxes the restriction imposed by Tobit model allowing different mechanism to determine the discrete probability of participation and volume of transaction, condition on participation.

A hypothesis test for the double hurdle against the Tobit model can be made. The test could be made by estimating three regressions (Tobit model, the truncated regression and the probit models) separately and use a log-likelihood ratio (LR) test. The significant value of likelihood ratio implies the rejection of the null hypothesis that Tobit model is nested in the double hurdle model. Following Greene (2012) the likelihood ratio test of Tobit restriction can be performed as follows;

$$\text{LRstatic} = -2[lnL_{T-}(lnL_{TR} + lnL_{TR})] \quad (3.4)$$

Where LT represents the likelihood of Tobit, LP represents likelihood of probit and LTR represents likelihood of truncated regression.

This test statistic has LRstatic a chi-square distribution with degrees of freedom equal to the number of parameter restrictions made to get the Tobit model.

Information criterions develop by Akaike (1974) to estimate the kullback-Leibler information for selection of the two competing models was also along with likelihood ratio test. It is the relationship between the maximum likelihood which is used in many statistical analyses and the kullback-Leibler information. Similarity, According to Ademe (2017) joint decision criteria of log likelihood test and Akaike's information criteria (AIC) was employed to determine the

rejection or acceptance Tobit model. Thus, he defined Akaike's information criterion (AIC) as follows:

AIC= -2(log-likelihood) +2k_____ (3.5)

Where, K is the number of estimated parameters included in the model, the log-likelihood of the model is derived by separately estimating each competing models readily available in statistical output. Decision rule for this statistical test suggests that the model with the lowest AIC taken as the best model among competing models.

Therefore specification of double hurdle model is as follow:

Double hurdle specification requires two latent variables; Y* related with binary choice model determining decision to income diversification (which is probit model) and Y_i referring to the level (intensity) to which income diversified) that is a truncated regression in nature. These latent variables are expressed as linear functions of the first and second hurdle regressors, zi and xi, respectively, where zi represents the regressors used to explain the decision to income diversify and xi shows those variables used to explain the decision regarding the intensity of income diversification.

1. Participation equation

$$Y^* = Z_i\delta + u_{i,}\ u_i \approx N(0\ 1) \quad (3.6)$$

$$p_i = \begin{cases} 1 \text{ if} z_i\delta + u_i > 0 \\ 0 \text{ if} z_i\delta + u_i \leq 0 \end{cases}$$

2. Intensity equation

$$Y_i = X_i\beta + \varepsilon_{i,}\ \varepsilon_i \approx N(0, \delta) \quad (3.7)$$

$$Y_i = \begin{cases} Y^*\ if Y^* > 0\ and\ p_i = 1 \\ 0, \qquad otherwise \end{cases}$$

Where: corr (u_i, ε_i) = p refers to unobserved factors affecting participation May or not affect intensity of participation

Y_i* = refers to latent variable describing the household's decision to participate in income diversification.

Z_i = refers to vector of explanatory variables explaining the participation decision

δ = refers to unknown parameter to be estimated in the mode

Y_i = refers to dependent variable that describe the degree at which household after participate in the income diversification,

β = refers to unknown parameter to be estimated in the model, and ε_i and u_i refers to are respective error assumed to be independent and distributed

The maximum likelihood estimator of double hurdle model (MLE) produces first and second hurdle results. The hurdle can be obtained from probit estimator. Then, the maximum likelihood estimator (MLE) for the second hurdle can be estimated truncated normal function, the estimation results was identified whether estimations made simultaneously or one regression at a time. On other hand, while using Craggit makes estimation more coherent, use of probit and truncated regression for Double hurdle model would not change results (Burke, 2016). Furthermore, under the assumption of independent, homoscedasticity, and normally distributed between two error terms (vi and εi), the Log-likelihood function of the double-hurdle is the summation log-likelihood of probit model and truncated regression model(McDowell, 2003). Therefore, the log-likelihood function for the double-hurdle model that nests a univariate probit model and a truncated regression model is given following Cragg (1997) by:

$$\ln L_i = \sum 0 \ln\left[1 - \Phi(\varphi i'\alpha)\left\langle\frac{Xi'\beta}{\sigma}\right\rangle\right] + \sum + \ln\left[\Phi(\varphi i'\alpha)\frac{1}{\sigma}\phi\left\langle\frac{yi - Xi'\beta}{\sigma}\right\rangle\right] \quad (3.8)$$

Where,

‘‘0’’ indicates summation over the zero observation in the sample;

‘‘+’’ indicates summation over positive observations

Φ (.) and ϕ (.) Represent standard normal probability and density functions respectively,

Φi And X_i represent independent variables for the probit model and truncated regression model respectively, β, σ, and δ are parameters estimated from each model.

The first term the right-hand side denotes the summation over the zero observation in the sample. Whereas the second term on the right-hand side indicates summation over the positive observations.

Model specification for propensity score matching model

In this study, PSM model was employed to see the impact of income diversification on outcome variables (income, consumption expenditure and saving). The reason for the adoption of these model is, the study lacks baseline data or longitudinal data and thus depends on cross-sectional data for which PSM model is more appropriate. Therefore, to evaluate the impact of income diversification on the livelihood of female-headed household, PSM model was applied.

The PSM technique enable us to extract information from the sample of treated (participant) female-headed household and a set of matching (non-participant) female-headed household that look like the participant (treated) female-headed household in all relevant pre-intervention characteristics. The objective of PSM is to find the closest comparison group from a sample of control female- headed household and treated female-headed household. Closest was measured in terms of observable characteristics. FHH with the same propensity scores is paired and the average treatment effect on the treated (ATT) will then be estimated by the difference in outcome between the treated and control/comparison group (Greene, 2012).

In this study, the main pillars to apply Propensity score matching (PSM) are existence of two group (income diversified and non-diversified), female- headed household and potential outcome variables (income, consumption expenditure and saving). The idea is to match FHH that participates in income source multiplication with that of non-participants in income diversification sharing full observable characteristics. Therefore, the average effect of participation in income diversification was measured as the average difference in outcome variables (income, consumption expenditure and saving) between the participants/treated group and non-participants /control group in income source multiplication. That means the impact is the change in income, consumption expenditure and saving as an outcome variable.

The use of propensity score matching model is to answer the question “what would be the income, consumption expenditure and saving of female-headed household who participated in income source diversification had than rural women headed household not participated in income diversification?” Participants (treated) and non-participants (control group) of FHH’s in income diversification are related on some characteristics (age, educational level, family size, dependency ratio, land size, livestock ownership, frequency of extension contact, member of farmer based organization, access of credit service, gender-based discrimination, Pre received

shock, access to irrigated land, distance from market center, and average annual agricultural income). These variables are important to identify comparison group.

According to Rosenbaum and Rubin (1983), PSM can be explained as the conditional probability of taking a treatment given pretreatment characteristics. The propensity score model is defined as:

$$P(X) = Pr\ (D=1|Xi) = E\ (D/X_i) \quad (3.9)$$

Let, Y_i^T −is outcome of treated (income, consumption expenditure and saving) in birr of i^{th} female-headed household, when she is treated, Y_i^C: is outcome of control (income, consumption expenditure and saving) in birr of i^{th} women headed household, when she is controlled and ΔI: change in outcome between the treated and control group. Therefore, the difference in outcome between the treated and control group can be calculated from the following mathematical equation:

$$\Delta I = Y_i^T - Y_i^C \quad (3.10)$$

Let the above equation can be determined in causal effect notational form and in this study 'D' represent participation in income diversification which is a dummy variable such that D =1 if women headed household is participant in income diversification and D = 0 otherwise. Then, the formula for average treatment effect on the treated (ATT) can be seen as follow:

$$ATT = E(Y_i^T - Y_i^C | D_i = 1) = E\ (Y_i^T | D_i = 1) - E\ (Y_i^C | D_i = 1) \quad (3.11)$$

Where:

$E\ (Y_i^T | D_i = 1)$ =Average outcome for participant women headed household if they participated in income diversification.

$E\ (Y_i^C | D_i = 1)$ =Average outcome for participant women headed household if they were not participated

ATT=Average treatment effect on the treated for the sample.

$$ATT = E\ (Y_i^T - Y_i^C | D_i = 1) = E\ (Y_i^T | D_i = 1) - E(Y_i^C | D_i = 1) \quad (3.12)$$

The main problem in the evaluation of impact is difficult to observe a person's outcome for with and without treatment at the same time. The post-intervention outcome $(Y_i^T|D_i = 1)$ can be observable, however, the counter factual outcome of the i^{th} female headed household when she does not a treated, the treatment is not observable in the data. Therefore, an alternative counterfactual has to be constructed through the formation of control groups that resembles to the observed outcomes of participants or the treatment group. Therefore, ATT used to estimate the true impact as follow as;

$$\text{ATT= E } (Y_i^T - Y_i^C|D_i = 1) = \text{E } (Y_i^T|D_i = 1) - E(Y_i^C|D_i = 0) \quad (3.13)$$

There are two important assumptions that need to be satisfied for the PSM model to correctly estimate the impact of participation in income diversification on outcome variable (income, consumption expenditure, and saving). These are the conditional independence assumption and the common support condition.

Conditional Independence Assumption (CIA): It indicates that outcomes are independent of treatment and conditional on (Xi). This assumption shows that the selection is only depend on observable variables that affect participation decision of households and outcome variables simultaneously (Caliendo and Kopeinig, 2008).

$$(Y_i^T, Y_i^C) \perp \text{D|X} \quad (3.14)$$

Common Support assumption: A further assumptions besides conditional independence (CIA) is the common support or overlap condition. The assumption is that P(x) (probabilities) lies between 0 and 1.This restriction implies that the balancing propensity is performed only on the observations whose propensity score falls in the common support region of treated and control groups (Becker and Ichino, 2010). Individuals that lie outside the common support region would be discarded in the estimation of treatment effect. That is;

$$0<\text{P } (D=1) |X<1 \quad (3.15)$$

3.6.3. Matching algorithm

There are four commonly used matching algorithms. Those are nearest neighbor matching (NNM), radius matching (RM), caliper matching (CM) and kernel matching (KM) was used to assess the impact of income diversification outcome variable (income, consumption expenditure, and saving) of female-headed household. Let us discuss one by one:-

Nearest neighbor matching: It is the straight forward matching estimator. In a nearest neighbor matching, female-headed household from the comparison group is chosen as a match for a treated woman headed household in terms of the closest propensity score or similarity in terms of observed characteristics. Women headed household from the controlled group are chosen as a matching partner for a treated woman that is closest in terms of propensity scores. For each treated woman headed household i, a control woman headed household j that has the closet scores in terms of the observable characteristics is selected. A propensity score that minimizes the distance between the treated and untreated defines the nearest neighbor matching algorism.

Caliper and Radius Matching: NN matching faces the risk of bad matches, if the closest neighbor is far away. This can be avoided by imposing a tolerance level on the maximum propensity score distance (caliper). To overcome this problem the caliper matching algorithm is another alternative. Caliper matching means that an individual from the comparison group is chosen as a matching partner for a treated individual that lies within a given caliper (propensity score range) and is closest in terms of propensity score (Caliendo and Kopeinig, 2008). 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.

Kernel matching: It is non-parametric matching estimator that compare the outcome of each treated women-headed household to a weighted average of the outcomes of all the untreated women headed household with the highest weight placed on those with scores close to the treated rural female-headed household. Caliendo and Kopeinig (2008) argue that Kernel matching uses weighted average of all rural female-headed household in a comparison group to construct the counterfactual outcome. The assignment of weights depends on the distance between each rural FHH from the comparison group and treated group for which the counterfactual is estimated. Therefore, more weight is assigned to comparison rural female-

headed households whose propensity score is closer to that of the treated group. Each rural woman-headed household from the treated group is thus matched with several control rural women-headed household with weights inversely proportional to the distance between treated and control group.

3.6.4. Testing the matching quality

Since we do not depend on all covariates but on the propensity score, it has to be checked if the matching procedure is able to balance the distribution of the relevant variables in both the comparison and treated group. The purpose of the propensity score matching is not to perfectly predict selection into treatment but to balance all covariates. While differences in covariates are expected before matching, these would be avoided after matching. The main purpose of the PSM is that it serves to balance covariates between the two groups. Consequently, the idea behind balancing tests is to check whether the propensity score is adequately balanced or not.

The basic idea of all approaches is to compare the situation before and after matching and check if there remain any differences after conditioning on the propensity score (Caliendo and Kopeinig, 2008). Rosenbaum and Rubin (1983) emphasized that the crucial issue is to check whether the balancing condition is satisfied or not.

There are different approaches in applying the method of covariate balancing (i.e., the equality of the means on the scores and all the covariates) between treated and non-treated individuals. Among different procedures the most commonly applied ones are described below.

Standardized bias

One suitable indicator to assess the distance in marginal distributions of the X variables is the standardized bias (SB) suggested by (Rosenbaum and Rubin, 1985).It is used to quantify the bias between treated and control groups. The standardized bias before matching is given by;

$$SB_{before} = 100.\frac{\overline{X_1}-\overline{X_0}}{\sqrt{0.5(V_1(X)+V_0(X))}} \quad (3.16)$$

The standardized bias after matching is given by;

$$SB_{after} = 100.\frac{\overline{X_{1M}}-\overline{X_{0M}}}{\sqrt{0.5(V_{1M}(X)+V_{OM}(X))}} \quad (3.17)$$

Where $\overline{X_1}$ and $\overline{X_0}$ are the sample means for the treated and control group respectively

Where X (V_1) and X (V_0) are the mean (variance) in the treatment and control group before matching respectively, X_1M (V_1M) and X_0M (V_0M) are the corresponding values for the matched samples.

T-test

A two-sample *t*-test to check if there are significant differences in covariate means for both groups (Rosenbaum and Rubin, 1985). Before matching differences are expected, but after matching the covariates should be balanced between the two groups and hence no significant differences should be found. The *t*-test might be preferred if the evaluator is concerned with the statistical significance of the results. The shortcoming here is that the bias reduction before and after matching is not clearly visible.

Joint Significance and Pseudo-R^2

Additionally, Sianesi (2001) suggests re-estimating the propensity score on the matched sample that is only on participants and matched non-participants and compare the pseudo R^2 before and after matching. The pseudo-R^2 indicates how well the explanatory variables explain the participation probability. After matching there should be no significant differences in the covariates between the two groups and the pseudo-R^2 should be fairly low.

3.7. Diagnostic Test.

3.7.1 Multicollinearity test

Under this, the researcher was tested whether the both explanatory variables are correlated or not in the model separately. Accordingly, the existence of multicollinearity among the continuous variables was tested. There are different methods to see the existence of multicollinearity problem between the included explanatory variables in the model. Among these methods, variance - inflating factor (VIF) is commonly used and it also employed in the present study to detect multicollinearity problem among continuous explanatory variables (Gujarati, 2004). While doing this test, all independent variables were orthogonal or uncorrelated with each other.

In the case of orthogonal regressors both the tolerance and variance inflation factor (VIF) are 1, if not orthogonal the tolerance becomes closer to 0 and VIF very large. Hence, the closer the

tolerance and VIF to 1, the less severe the problem of multicollinearity, and the reverse is true. The rule of thumb recommends, it is necessary to consider the severity of the multicollinearity problem when the VIF is 10 or greater and tolerance is 0.1 and less. Therefore, to deal with the severe correlation among regresses, we use the VIF result and identify the source of multicollinearity from the test (Greene, 2002).

Mathematically, VIF for individual explanatory variable (Xi) can be computed as:

$$VIF(Xi) = \frac{1}{1 - R_2} \quad \text{(3.18)}$$

Where, R^2 is the coefficient of correlation among explanatory variables.

In other way, multicollinearity among dummy variable were also computed and the value of the CC were less than 0.75 which indicates absence of multicollinearity problem (see Appendix: Table 3).

$$CC = \sqrt{\frac{\chi^2}{n + \chi^2}} \quad \text{(3.19)}$$

Where, CC is contingency coefficient and χ^2 = chi-square value

3.7.2. Heteroscedasticity test

Under classical regression model once at the basic assumption that the probability distribution term remains same overall observations of explanatory Var (U_i) = $S^2_=$ constant variance. The condition of non-constant variance or non-homoscedasticity of variance is known as heteroscedasticity. This type of test is used to examine the pattern of the error terms variance are constant or to test the assumption of homoscedasticity. Heteroscedasticity is present if the variance of the error term is not constant or different variance for different segments of the population or sample size (Gujarati, 2004). To see the problem, the researcher was run the '*hettest'* command in software Stata and in response to see the p-value of Breusch pagan test to decide the issue.

3.7.3. Model specification test

The model specification refers to the description of the process by which the dependent variables is generated independent variable. Thus, it encompasses the choice of dependent and independent variables, as well as the functional form connecting the independent variable to the dependent. The specification of a regression model should be based primarily on theoretical considerations rather than empirical or methodological ones. Testing the model specification is very important to check out whether one or more relevant variables are omitted from the model or one or more irrelevant variables are included in the model. There are different methods to check the specification error of the model. The Ramsey reset test for omitted variables is commonly used in the test.

3.7.4. Sensitivity test

Recently checking the sensitivity of the estimated results becomes an increasingly important topic in the applied evaluation literatures (Caliendo and Kopeinig, 2008). In observational studies, treatments are not randomly assigned to experiments units, so that the randomization tests and their associated interval are not generally applicable. In attempt to compensate for lack of randomization, treated and control units are often matched on the basis of observed covariates.

To confirm the robustness of the finding of ATT; the post estimation analysis of sensitivity test was checked. Sensitivity analyses examine how strong the influence ϒ (unobserved) on the participation process needs to be. If there are unobserved variables that affect participation decision and the outcome variable simultaneously, a hidden bias might arise to which the average treatment effect are not robust (Rosenbaum and Rubin, 1985).

In participation probability is given by;

$$P_i = p(xi, ui) = P(Di=1|xi, ui) = F(\beta xi + \Upsilon ui) \quad (3.19)$$

Where Xi is the observed characteristics for an individual, u_i is the unobserved variables, and ϒ is the effect of ui on participation decision. If the analysis is free of hidden bias ϒ is zero and the participation decision was fixed only by Xi. In case of hidden bias both group with the same observed covariates x has different chances of receiving treatment.

Sensitivity test evaluates how program effect is affected by change in ϒ. The following bounds on the odds ratio of the participation probability of both individuals are applied.

$$\frac{1}{e^{\Upsilon}} \leq \frac{Pi(1-Pj)}{Pj(1-Pi)} \leq e^{\Upsilon} \quad (3.20)$$

3.8. Variable Description and Hypothesis

For the purpose of this study, different variables were selected based on economic theory and previous empirical findings from the existing literatures on similar studies. In these study variables which affect participation decision in to income diversification are selected depending on observable characteristics of respondents in the study area. Thus, in impact evaluation study, variable choice must be, those variables which affect both participants and non-participants (both treated and control groups share characteristics of X covariates). According Heckman et al., (1997) only variables which affect both participation and outcome should be included in the estimation of propensity score.

3.8.1. Variable description for income diversification participation

Dependent variable

Income diversification participation decision: It is dummy variable measured through engagement into income diversification through SID, which taking a value of one (1) if female-headed household participate in income diversification or share of income from off-farm, and non-farm is greater than 5% and zero (0) if not participant or the share of income from diversification into off-farm and non-farm is less than 5% (Teshome and Edriss, 2013; Idris-Adeniyi et al., 2020). This is regressed in the first hurdle of double hurdle model by using probit model.

Intensity of income diversification: It is continuous variable measured in terms of income diversification index (Simpson income diversification index) since income diversification is defined in terms of both the number of income sources as well as how evenly the distributions of the income between the different sources (Adem and Tesafa, 2020). This is the second hurdle of the double hurdle model which produces truncated regression.

Independent variables

Age of the household head (AGE): This is continuous variable which is defined as the number of years after birth of the household. It is expected to have negative relationship with participation in income diversification. Gebru *et al.* (2018) found that in Ethiopia, age household is negatively related with income diversification and its intensity of diversification. The young are more likely to participate in income diversification because of their ambition to invest and accumulate wealth during their working age. The old on the other hand are less likely to participate in income diversification activities. The old are more likely to depend on past saving and accumulated wealth for their consumption. Therefore, based on these arguments, age of the household head was hypothesized to have a negative relationship with participation in income diversification.

Education states of the house hold heads (EDL): This is continuous variable representing the female households head's education level measured by year of schooling. Household heads with more years of formal education were expected to have a higher ability to accept new ideas and innovations, get access of information than female headed households with no or low education and therefore would be more willing to diversify their income source with higher intensity to protect risk in the rural area. Hence, this variable is hypothesized to have positive and negative effect on the probability of income diversification participation decision. The result of Adem and Tesafa (2020) in Ethiopia, farmers income diversification participation decision and intensity of diversification positively influenced by education level. However, Demissie (2013) by arguing that educated persons specialized one income generating activity rather than diversification.

Family size (FSIZE): This is the total number of family members in the household and it was continuous variable expressed by total number of family in terms of adult equivalent. Accordingly, household who have larger families are more likely to diversify their income source compared *to* smaller ones as they are more likely to have a higher dependency ratio with limited and marginalized land agricultural income alone could not meet food security/livelihood and hence farmers might tend to involve in different non-farm and off-farm activities that bring additional income to survive livelihood shock. Adem and Tesafa (2020); Khan *et al.* (2019) reveal that family size has positive effects on both income diversification and its intensity. In this

study therefore, family size is expected to have a positive relationship with income diversification participation decision and the extent of participation in study area.

Dependency ratio (DRATIO): It is a continuous variable which obtained by dividing economically inactive labor force by economically active labor force**.** The age limit to include a family member as economically dependent differs from researcher to researcher in developing country; some take children aged less than 15 years. Others take the age when children involve in economic activity. For example, in case of Ethiopia, Beyene (2008) takes this to be less than five years and above sixty five as dependent. While Mekonnen (2011) uses age limit less than seven years and greater than and equals to sixty five years as economically inactive household family member. Gebru *et al.* (2018) and Batool *et al.*(2017) use age limit less than 15 and above 65 as economically inactive.

Therefore this study used the economically inactive labor force interpreted as, those members of farm households whose age is less than 15 years or younger and 65 years or above as economically dependent household members. Since international law put age limitation for children involve in any economic activity. When there is a dependent ratio, household has to raise more money and the marginal value of leisure was declined. Anang (2017) found positive relationship between dependency ratio and participation in income diversification. Whereas, Owusu et al.(2011) in Ghana found income diversification is negatively influenced by dependency ratio. Therefore, the number of economically inactive within the households is expected to have positive and negative impact on income diversification participation decision and intensity of participation.

Cultivated land size (LSIZE): Land size is in continuous variable measured hectares. It is one of the most important asset which generate income through multi-dimensional in rural area. The larger the land holding less like to diversify income source due to the household's full time employed on their farms. Batool *et al.* (2017) found that having large sizes of land affect income diversification negatively due to rural farmers specialize in to on-farm activities. Therefore, this variable was hypothesized to have a negative and positive relationship with the extent of income diversification.

Total livestock ownership (LOWNED): This is a continuous variable represented by the total number of livestock owned by rural household and measured in tropical livestock unit (TLU). A household livestock size in TLU is calculated by multiplying the number of each type of animal by an appropriate conversion factor and then summing together. Livestock is considered as an asset which is liquid. Balense and Debebe (2020) found positive effect of livestock ownership on the income diversification participation decision. Therefore, in this study, total livestock ownership was expected to have positively relationship with income diversification.

Frequency of extension contact service (FECONTACT): This variable is coded as a continuous variable to proxy for number of contacts between farmers with public extension workers. This may be due to the fact that the role of rural development agent in farming communities had helped farm households engage in certain income-generating activities, assisted them how to apply selective and emerging technology of crops and livestock production to boost income of the rural household. Fekadu et al. (2021) found that service of agricultural development agent had positively affect income diversification participation decision and its intensity. In this study, therefore, number of contact with extension workers is expected to have a positive relationship with income diversification participation decision and the extent of participation.

Member of any farmer based organization (MFBO): is dummy variable which takes the value 1 if the household head is a member and 0 otherwise. It is conceivable that, member of any farmer based organization have a number of contributions for smallholder rural farmers in developing countries. For example, it helps the society as a source of income for their member, necessary inputs, market information, buy their produce at better prices. Danso *et al*. (2020) found that member of farmer based organization was positive effect on income diversification participation decision. Thus it is expected those farmers who are members of any farmer based organization would be more likely to participate in income diversification and in its intensity in the study area.

Access of credit service (ACREDIT): It is a dummy variable which take 1 if the respondent is credit user and 0 otherwise. This variable is expected to influence income diversification decision positively on the assumption that credit use improves the financial capacity of rural house to buy more improved production inputs and to start up a small business, thereby

existence of credit service important for farmer in rural. The variable is hypothesized to have positive effect on income diversification participation decision. Adem and Tesafa (2020) found that access of credit service positively affect both income diversification participation decision and level of income diversification.

Gender-based discrimination (GENDIS): Gender-based discrimination is one of variable which affect income diversification of female headed households. In less developed country, female has no equal opportunity with male due to gender norm and culture. According to Shannon (2019), gender based discrimination refers to the social construct that expressed psychologically from the society (i.e. identifying as male and female). Accordingly, in rural area there is income generated activity which is delivered by man alone; for example, long distance trade, ploughing land, etc. In addition to this, female headed households also considered incapable in human capital to participate in farm based organization in rural area. Therefore, in this study gender based discrimination is coded as a dummy variable if households are victim of gender-based discrimination one (1) and zero otherwise. Moreover, it was expected to influence negatively income diversification participation decision and intensity of income diversification.

Pre-received shock (PRSHOCK): It is a dummy variable, which takes 1 if a household face flooding and illness of the household head member and 0, otherwise. Households experiencing health problems or crop failure are expected to link positively with income diversification strategies. According to Yizengaw (2014) there are two type of shock variable broadly in the rural area. Those are:-

Covariant Shocks: It is a shock that are occurred in the in the community such as occurrence of flooding/water logging, too early availability of rain at the beginning of summer, lack of enough rainfall at growing season of crops, to early or too late stop of rain at the for their farm, which affect income generated from on-farm income specifically.

Idiosyncratic Shocks: It is also other type of shock, which is peculiar to the households like damage of crops by insects; livestock's and birds, damage due to weed infestation, illness of household members. Therefore, shock variable might be hinder or push rural household heads diversify their income source to survive the problem.

Distance of household home from market (DMRK): It is a continuous variable measured by hour taken to travel from their residence to market place. According to Balense and Debebe (2020); Adem and Tesafa (2020) were found negative effect of distance of households home from market on both income diversification and level of diversification. Hence, being farther location from marketplace decreases the probability of income diversification access. It was expected that being far from market place limit being participant in income diversification. Therefore, the expected relation would be negative.

Average annual agricultural income (LnAAI): It is the income that households receive by participating in agricultural activities. It is continuous variable which is measured in monetary value (birr). Teshome and Edriss, (2013) and Balense and Debebe, (2020) found that positive relationship between income from agriculture and income diversification participation decision and its intensity. Therefore, in this study also it is hypothesized that average annual agricultural income will have positive and negative significant effect on income diversification participation decision and intensity of diversification.

Outcome variables (impact indicators): In this study, three outcome variables namely average yearly household income, consumption expenditure and saving were used as an indicator of the impact of multiplication of income source on household livelihood. Average yearly household income is a continuous variable which determines household livelihood. It has a positive expected sign. And also, consumption expenditure and saving is also expected positively related with household livelihood.

Table 3.2: Summary of hypothesized explanatory variables used in the econometric model

Variable's code	Definition	Measurement	Type	Expected sign
Dependent variable	Female headed household participation income diversification	1for participant and 0 for not participant	Dummy	
	Intensity of income diversification	SID	Continuous	
Outcome variable	Annual income, consumption expenditure, and saving amount of female headed household	Birr	Continuous	+
AGE	Age of household head(female)	Year	Continuous	-
FSIZE	Family size of household	Adult equivalent	Continuous	+
LnAAI	Annual average agricultural income	ETB	Continuous	-/+
EDU	Education states of the household	Year of schooling	Continuous	+
LSAZE	Cultivated land size of household	Hectares	Continuous	-/+
LOWNED	Livestock owned	TLU	Continuous	-
DRATIO	Dependency ratio	Family members < 15 > 64 old age in adult equivalent	continuous	+
ACREDIT	Access to credit service	1 =Yes 0 = No	Dummy	+
FECONTAC	Frequency of extension contact	Frequency	Continuous	+
GENDIS	Gender-based discrimination	1= Yes 0 = No	Dummy	+
AILAND	Access of irrigated land	1 = Yes 0 = No	Dummy	+
PRSH	Pre received shock	1 = Yes 0 = No	Dummy	+
MFBO	Member of any farmer-based organization	1 = Yes 0 = No	Dummy	+
DMRK	Distance from the market center	hour	continuous	-

Source: own survey, 2021

CHAPTER FOUR

4. RESULTS AND DISCUSSION

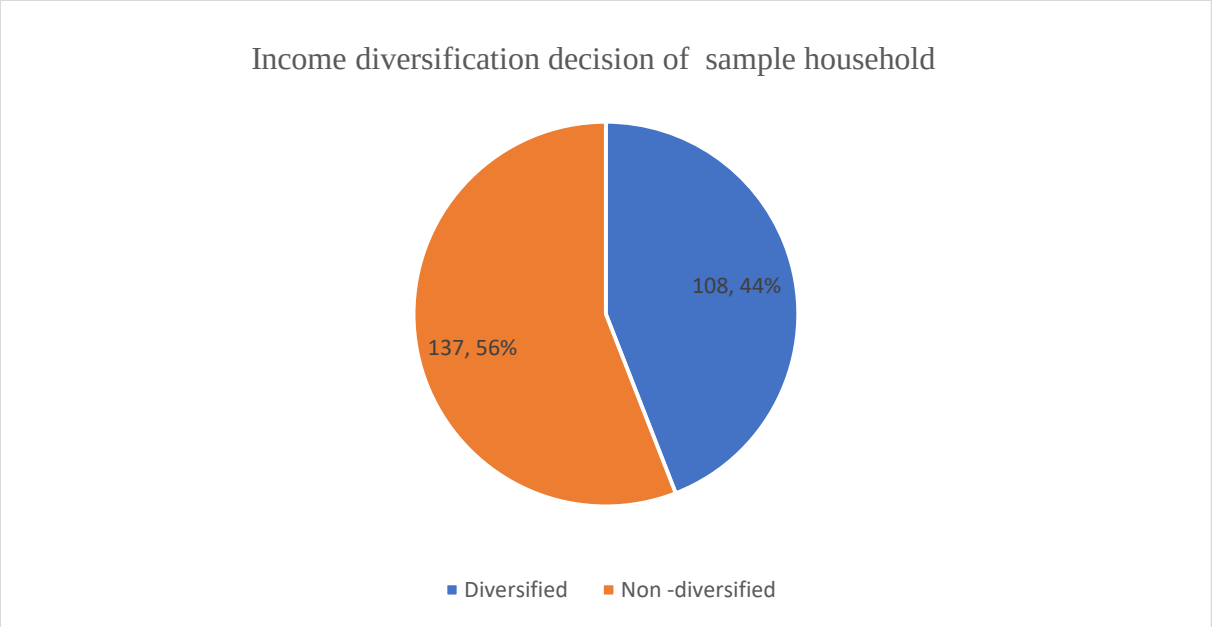
This Chapter presents the profile of female headed household, extent and determinants of female headed household participation in income diversification and empirical assessment on the impact of income diversification on female headed household livelihood through livelihood indicators. The livelihood assets used to measure household livelihood in this study were income per capita, consumption expenditure per capita and saving per capita. Data were collected through structural interview questionnaire from 245 female headed households of whom 108 comprise treated (diversified) and 137 represent comparison (non- diversified) households.

In addition, this chapter has two major sections. The first section describes the socio-economic and demographic characteristics of respondents using descriptive statistics based on the concepts drawn from the conceptual framework. The second component of the chapter is about econometric model, mainly double–hurdle model and propensity score matching models. The double- hurdle model (DHM) was used to identify the factors that determine female headed household participation in the income diversification and its intensity. Impact study was assessed using the PSM model. Both descriptive and econometric analysis was under taken by using EXCEL, SPSS version 25 and STATA software version 16.

4.1. Descriptive Statistics

This section discusses and reports information about the socio-demographic characteristics of respondents using descriptive statistics.

The results of the descriptive statistics relied on the data collected from randomly selected income diversified and non-diversified rural household. As indicated in the figure below, the total of 245 respondents were surveyed in the study area. Of these 108 (44.08%) respondents were multiplying their source of income and represent the treated group whereas the rest 137(55.92%) respondents were specializing their source of income and were classified as the control or comparison group based on their degree of income diversification measured by SID.



Source: Author's computation from own survey data, 2021
Figure 4.1: Income diversification decision of the sample household

Table 4.1 of below shows the profile of the respondents in terms of the socio-economic and demographic factors to evaluate the values using statistical tests such as the t-test, χ^2 test, mean, frequency and percentage. The t-test was employed to test the statistical significance mean difference between continuous variables while, χ^2 test was used to test the association between dummy variables, with the comparison of treated (diversified) and control (non-diversified) female headed households. Therefore, descriptive analyses of both continuous and categorical variables indicated that there were statistical significant differences between the income diversified and non-diversified female-headed household in terms of all variable included in the analysis except, cultivated land size, livestock owned and average agricultural income.

4.1.1. Demographic characteristics of sample households

Age of household head (female): as shown on table 4.1, the average age of the sample female headed households were found to be 43.55102 years where the minimum is 25 and the maximum is 68. On the other hand, the average age of the diversified female-headed households was 41.40741years and the corresponding figure for non-diversified female-headed households was 45.24088. From the statistical analysis performed, it was found that the mean age difference between diversified and non- diversified female headed households in income diversification is 3.833469 years and the mean comparison test shows there is a statistically significant difference

in the distribution of household head age between the diversified and non-diversified female-headed households at 1% significance level

The educational level of female-headed household: It is expressed in terms of year of schooling. The survey result indicates that, the average of education level for the female headed household was 3.253061. The maximum and minimum grade completed in formal schooling was 12 and 0 grade, respectively. With regarding to the distribution according to their education level, 45.31% of sampled female headed households were never attended any formal education, while 54.79% of households were literate at different levels of year schooling. As shown in table 4.1, the statistical analysis from independent sample t-test indicate that, income diversified female headed household had more year of schooling (5.314815) than nan-diversified (1.627737) on average and the result of statistics show that income diversified female-headed households had more year of schooling than non-diversified, and it was statistically significant at 1% of significance level

Total family size: As it was showed below table, total family size of the respondents was converted into adult equivalent in order to consider the effects of age and sex of each family member. Accordingly, the family size of the sample female headed households which were measured in adult equivalent ranged from 1 to 12.5 with an average of 6.269388. For the diversified (treated group) and non- diversified (control group), the average family size is about 7.046296 and 5.656934 respectively. When we compare the average household size between diversified and non-diversified, the study result revealed that female headed households that participate in income diversification have more household size than non- diversified households. The mean difference in household size between the two groups is -1.389362. Therefore, the variable is statically significant at 1% significance level.

Dependency ratio: Regarding the dependence ratio of the female headed households, the average dependency of all sample respondents is about 0.2489378 per economically active people. On the other hand, the average dependence ratio for diversified and non-diversified is .3066608 and .2034336 per economically active people respectively. This result shows that income diversified female households have more economically inactive dependent people than non-diversified female headed households. The mean difference in dependence ratio between

diversified and non-diversified is about -.1032272. Therefore, the result was statistically significant at 5% significance level

Table 4.1: The demographic characteristics of sample household

Variables	Total sample	Diversified	Non-diversified	Mean difference			p-value	t-value
	Mean (Std. Dev.)	Mean (Std. Dev.)	Mean (Std. Dev.)		Max	Min		
AGE	43.55102 (7.510815)	41.40741 (6.934995)	45.24088 (7.539626)	3.833469	68	25	0.0001	4.0924
EDU	3.253061 (3.439352)	5.314815 (3.381224)	1.627737 2.488229	-3.687078	13	0	0.0000	-9.8284
FSIZE	6.269388 (2.272628)	7.046296 (2.24393)	5.656934 (2.108937)	-1.389362	12.5	1	0.0000	-4.9769
DRATIO	0.115102 (0.140355)	0.1427778 (0.138495)	0.0932847 (0.138437)	-.0494931	0,5	0	0.0059	-2.7778

Source: computation based on survey data (2021)

4.1.2 Socio- economic characteristics and resource of sample household heads

Cultivated land size: As shown in table 4.2 below, the average land size for the sample household is 2.613265 hectare with minimum and maximum of 1 and 5.5 hectares, respectively. The average cultivated land size for the diversified and the corresponding figure for non-diversified is 2.511574 hectare and 2.693431 hectare respectively. The mean difference is .1818566 hectare. This implies that participant female-headed households have less cultivated land size than non-diversified households. Therefore, the mean comparison test revealed that the means difference between the two groups regarding landholding size is statistically insignificant at any chosen significance level.

Livestock owned: As shown in Table 4.2 below, numbers of livestock owned have no a statistically significant difference between diversified and non-diversified. On average livestock owned by the sample household is 10.97702 in TLU. The average livestock owned by diversified and non-diversified female headed household is 11.10083 and 10.87942 in TLU respectively. The mean difference is -.2214173. The result of these statistical analyses indicates that

diversified households have more livestock population than non- diversified household and the computed mean difference is statistically insignificant at chosen significance level.

Average agricultural income: There are different important source of on-farm and off-farm income in the study area. However, the major part of sample households were participate on-farm (i.e., crop production and livestock rearing) income source whereas, some of sample households were participate in off-farm income. As shown in table 4.2 the average annual agricultural income for sample household is 9.817789. And out of this the average annual agricultural income of income diversified and non-diversified were 9.74068 and 9.878575 birr per a year. Accordingly, the mean comparison t-test revealed that the means difference between the two groups regarding average agricultural income is statistically insignificant at chosen significance level.

Table 4.2: Socio- economic characteristics and resource of sample household heads

Characteristics	Total sample	Diversified	Non-diversified	Mean difference			t-value	P-value
Continuous variable	Mean (Std. Dev.)	Mean (Std. Dev.)	Mean (Std. Dev.)		Max	Min		
LSIZE	2.613265 (.9851237)	2.511574 (1.111949)	2.693431 (.8681416)	.1818566	5.5	1	1.4377	0.1518
LOWNED	10.97702 (5.500332)	11.10083 (5.41878)	10.87942 (5.58166)	-.2214173	25.35	1.29	-0.3123	0.7551
lnAAI	9.817789 (1.485621)	9.74068 (1.432092)	9.878575 (1.528965)	.1378948	10.757903	.0001	0.7206	0.4718

Characteristics		Diversified		Non-diversified		Total sample		Pearson chi2	
Dummy variable		N	%	N	%	N	%		
AILAND	Yes	45	41.66	32	23.35	77	31.42	9.3938	0.002
	No	63	58.33	105	76.64	168	68.57		

Source: computation based on survey data (2021)

Access to irrigated land: Irrigation opportunities make multiple cropping possible which will create agricultural surplus. This surplus generates the financial capability for doing non-farm activities for the rural household head, particularly self-employment activities. Therefore, as

shown in table 4.2 above survey result point out that, among total respondent 31.42% of them were have access of irrigated land while the rest 68.68% were not. From the total of 108, income diversified households 41.66% and from the total of 137 non-diversified 23.35% reported that they had access of irrigation land. Therefore, the chi square test of the two groups was run and the result showed that, there is statistical relationship between participation in income diversification and access of irrigated land. And the variable is statistically significant at 1% level of significance.

4.1.3. Institutional and infrastructural service of sample household

Frequency of extension service: The agricultural extension service providers in the district are office of agriculture experts and development agents. The main objective of the rural development agent in rural is working to diversify the income sources of the household heads in addition to increase agricultural production and productivity which enhance agricultural income. Specifically, they induced and guide rural household to diversification income at the household level and to increase the crop production by using modern agricultural technologies like chemical fertilizer, irrigation etc. Therefore, based on table 4.3 the survey result indicated that, the maximum and minimum of extension contact in the study area are 4 and 0 respectively per a year. The average rural development contact of all female headed household is about 1.036735 per a year. On the other hand, the average of extension contact for diversifier and non-diversifier is 1.287037 per a year and .8394161 per a year respectively. Thus, the mean comparison test shows there is a statistically significant difference due to frequency of extension service between the diversifier and non-diversifier female-headed households at 1% significance level.

Distance of household home from market: As shown table 4.3, the distance from home to the nearest market place where farmers sold their agricultural output was on average of 6.504082 km. The minimum and maximum distance that the sample household travels to the nearest market center was 1 and 16 km, respectively. The mean difference between distance to the nearest market among diversifier (treated) and non- diversifier (control) are 4.717593 and 7.912409 respectively. The mean comparison test result of the two groups concerning distance of FHH home from market is statistically significant at a 1% significance level.

Access of credit service: Credit is assumed to play crucial role in increase production and enhance rural household to diversify their income source to off-farm and non-farm activities by

reducing financial constraint. In the study area, the main source of credit for the sampled household was micro-finance institution. However, the credit received was not only used for production but also for trading purpose. As shown in table 4.3 above, about 55.10% of the total sampled households had gone credit service and the rest 44.89% were not. Among the income diversifier 86.12% and from non-diversified 30.65% had access to credit service. The chi-square test revealed that there was significant association between income diversified and non-diversified with respect to the credit received at 1% significance level.

Table 4.3 Institutional and infrastructural service of sample household

Character Eristics	Total sample	Diversified	Non-diversified	Mean difference			t-value	p-value
Continuous variable	Mean (Std. Dev.)	Mean (Std. Dev.)	Mean (Std. Dev.)		Max	Min		
FECONTACT	1.036735 (1.009523)	1.287037 (1.041664)	.8394161 (.9437826)	-.447621	4.00	.00	-3.5255	0.0005
DMRK	6.504082 (3.416448)	4.717593 (2.481024)	7.912409 (3.400853)	3.194816	16	1	8.1928	0.0000

Characteristics		Diversified		Non-diversified		Total sample		Pearson chi2	
Dummy variable		N	%	N	%	N	%		
ACREDIT	Yes	52	48.15	45	32.85	97	55.10	5.9121	0.015
	No	56	51.85	92	67.15	148	44.89		
MFBO	Yes	36	33.34	61	44.52	171	69.79	11.9300	0.001
	No	72	66.66	76	55.48	74	30.21		

Source: computation based on survey data (2021)

Membership to any farmer-based organization: is an important social capital that promotes sharing of knowledge, information, experience and source of income. In addition, being a member of a group (cooperatives) opens a means of gaining off-farm employment opportunities and open access to loan for the household head to engage them in to non-farm activities like petty trade. Accordingly, in this survey study, among the sample households 69.79% of respondents are member of farmer-based organization while the rest 30.21% are not member of nay farmer-based organization in the study area. Beside this, from the total of 108 income diversified households 75% and from the total of 137 non-diversified households 65.69% were

member of farmer-based organization. This variable found to have ($\chi 2$= 11.9300) which portrays presence of statistically significant association between the two groups in terms of this variable at less 1% of significance level.

4.1.4 Natural and human related characteristics of sample household

Pre-received shock: is important variable in the analysis of rural livelihood since shock happen in the life of rural population determine the production and productivity of the households. Accordingly, table 4.4 shows, among the sample households 55.52% of the respondents were the victim of shock avail in the rural area. While 44.48 % of the sample households were not. Beside this, from the total of 108 income diversified households 22.23% and from the total of 137 non-diversified households 81.75% were victim of both idiosyncratic and covariant shock which hinder the female headed households to diversify their income. The chi-square test result shows that there was statistically significant association or difference between the diversified and non-diversified female headed households regarding to shock availed at 1% of significance level.

Gender-based discrimination: As indicated in table 4.4 below, among the total sampled households, 69.79% were discriminated from labour market participation in the study area, and remain 30.21% were not. Among the diversified female headed household, 75.93% were discriminated and 24.08% were didn't discriminated while in non-diversified female headed households 31.38% were discriminated and the remaining 68.62% were didn't equally participate in labour market with their counterpart. The result shows that there is a statistically significant difference between the diversified and non-diversified female headed households regarding participation labor market with male headed households 1% of significance level.

Table 4.4 Natural and human related characteristics of sample households.

Characteristics		Diversified		Non-diversified		Total sample		Pearson chi2	p-value
Dummy variable		N	%	N	%	N	%		
PRSHOCK	Yes	50	46.29	26	18.98	76	31.03	21.0627	0.000
	No	58	53.71	111	81.02	169	68.97		
GENDIS	Yes	18	16.66	86	62.78	104	42.44	52.5524	0.000
	No	90	83.34	51	37.22	141	57.56		

Source: computation based on survey data (2021)

4.1.5. Income, consumption expenditure, and saving of respondents

As shown in table 4.5, the maximum and the minimum average income received by respondents were 10.84934 and 8.517193 birr. The average annual income of the sampled household heads were 10.28032 birr and the corresponding figure for diversified and non-diversified female-headed households were 10.49238birrs and 10.11314 birrs respectively. This implies that the average annual income of the treated female-headed household was greater than the control group. The mean difference is -.3792439 and it was statistically significant at a 1 percent probability level.

Table 4.5 Descriptive statistics of income and saving of diversified and non-diversified female headed household.

Variable	Total sample	Diversified	Non-diversified	Mean difference	T- value
	Mean Std.Dev.	Mean Std.Dev.	Mean Std.Dev.		
lnAHINCOME	10.28032 (.3703859)	10.49238 (.3336861)	10.11314 (.3076637)	-.3792439	-9.2277***
lnCONEX	10.18705 (.385347)	10.37105 (.3800292)	10.04199 (.3235791)	-.3290608	-7.3155***
lnSAVING	6.801232 (2.625608)	7.716715 (1.696671)	6.079537 (2.985504)	-1.637178	-5.0867***

***, shows significance at 1% probability level

Source: computation based on survey data (2021)

Similarly, the average annual consumption expenditure of the total sample households were 10.18705 birr per a year with in the maximum and minimum of average consumption expenditure 10.84934 and 8.517193 birr respectively. Moreover, the average annual consumption expenditure for treated and control group were 10.37105 and 10.04199 birr respectively with the mean difference of -.3290608 birr and the variable is statistically significant at 1% significant level. The average annual saving of the total sample is 6.801232 birr per a year. The average annual saving of income diversified participant household is 7.716715 birr and 6.079537 birr per a year for non- diversified female-headed households. This shows that the

average annual saving of participant household is more than the non-participant. The mean difference -1.637178and it was statistically significant at one percent significance level.

4.1.6. Types of income generating activities practiced in study area

In the study area farm households were engaged in different types of activities that are practiced to overcome the livelihood problems. Accordingly, off-farm and non-farm activities are among the major activities that supplement and complement farm income. The table 4.6 shows the proportion of rural household's participation income diversification in terms of farm, off-farm and non-farm and mixture of all livelihood strategies in eight selected rural kebeles. From the total respondents, majority of female-headed households in the study area participate on farm livelihood strategy. This on farm income includes income from crop production and livestock rearing which account (44.48%). Non-farm income activities exercised in the study area: small trade; self-employment income; rental income; remittance income and service giving income were practiced. Out of total sample (8.16%) of them participate on it.

Table 4.6 Types of economic activity practiced in the study area

Activity	N	%
On farm	109	44.48%
Non-farm	20	8.16%
Off-farm	8	3.26%
Mixture strategy	108	44.08%

Source: computed from own survey data, (2021)

While the remaining part were participated into off-farm (Sale of labor for agricultural/non-agricultural work (Daily wage work) and mixture of all livelihood strategies are 3.26% and 44.08% respectively.

4.1.7. Level of income diversification

As stated under literature review part of the study, there is no common standard for measuring degree of income diversification in different previous studies. However, proportion of source of income was used as a proxy for measuring degree of income diversification among farmer FHHs in the study area. Accordingly, female headed households in the study area mainly earn their

income from farming, renting farm land, petty trading, wage labour, service giving etc. However, most of the people in the study area are involved in agriculture.

For this survey study, household heads were classified into six categories of source of income based on how they obtain their living. Consequently, six income diversification sources were identified among the households, namely, farming, labour selling to agricultural income, service giving income, self-employment income, remittance income and rent income share. Majority of the household's members derived their livelihood by farming. The SID was used in this study to estimate the degree of income diversification among farm households in the Guduru district

Table 4.7 Levels of income diversification among female-headed households.

Diversification level	N	%
Non- diversifier	137	55.92
Low diversifier	2	0.8
Medium diversifier	36	14.7
High diversifier	70	28.58
Average		0.27
Total	245	100

Source: Author's calculation from own survey data, 2021

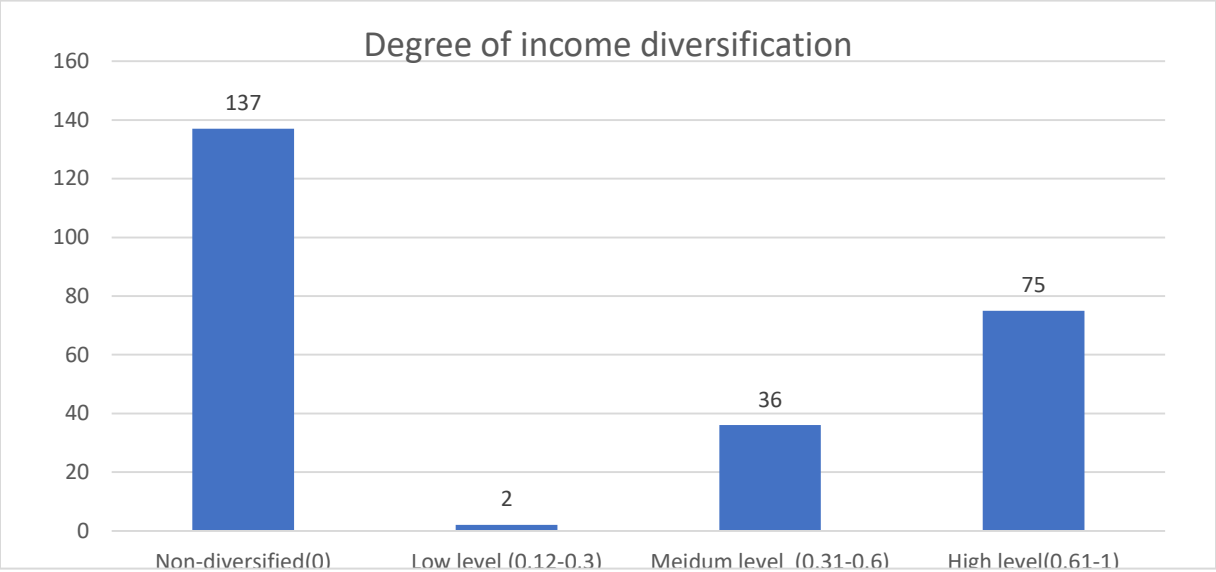
According to Adem and Tesafa (2020) the level of income diversification from 0 up to 0.3, 0.31 up to 0.6, and 0.61 up to 1 was low, medium and high level of income diversification index respectively. The results in figure below point out that the degree of income diversification among the rural FHH in the study area. The sample respondents with the most diversified income sources had the largest index and those with the small income sources had the smallest index.

As indicated in below figure, the survey result from the total sample household heads 137 (55.92%) had a diversity index of 0. Based on the result more than half of the sample household does not diversify their income sources. This implies that those farmers did not multiplying their income sources or did not participating in to more than one economic activity. Sample household heads whose Simpson diversity index equal to zero means household participate in to one types of economic activity it may be agriculture, formal employment, petty trade only etc. They specialize in to one economic activities or income- providing activity. About 2(0.8%) sample

household heads had diversity index between 0.12 up to 0.3, 36 (14.7) households had between 0.31 up to 0.6 and about 70 (28.58) households had diversity index between 0.61 and 0.79.

Therefore, the average income diversification index in the study area is 0.27 which is low. This shows that degree of diversification of 0.27 is lower than previous study conducted in different country. For example, the study conducted by Babatunde and Qaim (2009) of 0.479 in Nigeria, the finding of Agyeman et al. (2014) in Ghana 0.338 and higher than the finding of Tithy Dev et al. (2016) in Bangladesh with the SID value of 0.25. According to the survey result, the main cause behind this is that most of the people living in the rural area are vulnerable as they depend only on agriculture related activities for their livelihood do to different types of limitation like gender, adultness, social altitude, drought, scarcity of irrigation water and land, access of infrastructure and finance etc. Although most of off-farm and non-farm economic activities are the new sources of income emerged to the rural households, these activities are mostly run by the educated and rich farmers. This and other factors are the reason why low degree of income diversification observed among female headed household in the Guduru district.

Figure 4.2 Degree of income diversification in the study area.



Source: own survey data, 2021

4.2. Econometric Analysis

Under this, the study tried to analyze the factors that influence female headed household participation decision in income diversification with its intensity and the impact of participation

in the income diversification on their livelihood through income, consumption spent, and saving using Double-hurdle model and propensity score matching model respectively.

4.2.1. Data cleaning and management

In any data there may be a problem of multicollinearity, heteroscedasticity and sensitivity, which result for the outcome should be biased and inconsistent in econometric analysis especially cross-sectional data. Therefore, the data should be checked before it used for the analysis purpose. Accordingly, for this study, the problem of multicollinearity is detected by looking variance inflation factor (VIF) for continuous independent variables and contingency coefficients (CC) for the discrete independent variables.

Multicollinearity

Prior to running the double hurdle model, the presence or absence of multicollinearity has to be checked. There are two measures that are often suggested to test the existence of multicollinearity. These are: Variance Inflation Factor (VIF) for association among the continuous explanatory variables and Contingency Coefficients (CC) for dummy variables. The larger the value of VIF, the more "troublesome" or col-linear the variables Xi is. As a general rule, if the VIF of a variable exceeds 10, there is series multicollinearity. According to (Gujarati, 2004), to overcome the serious problems of multicollinearity, it is quite essential to omit one of the variables which have VIF greater than 10 and transforming the data are the most important remedial measure to overcome the problem. Thus, the Variance Inflation Factor (VIF) was employed to test the extent of multicollinearity between the continuous variables. The values of the VIF for ten continuous variables were found to be small (less than 10). This result point out that, the data have no serious problem of multicollinearity, (see Table 2 in the appendix). Therefore, all the ten continuous independent variables were retained and entered into the analysis.

Similarly, Contingency Coefficient was figured from survey data to check the existence of high degree of relationship among dummy independent variables. The contingency coefficient value lays 0 and 1. Based on the limited value of contingency coefficient, there are decision rule to identify the association between dummy independent variable. Accordingly, the rule states that,

when its value approaches 1, there is a collinearity problem between the dummy explanatory variables. While, if it's the value approach to zero there is no association between the variables.

Heteroscedasticity

This type of test is used to examine the pattern of the error term variance is constant or to test the assumption of homoscedasticity. Heteroscedasticity is present if the variance of the error term is not constant or different variance for different segments of the target population or sample size. Heteroscedasticity is more likely to exist in cross-sectional than time-series data. Therefore, to check the possible existence of heteroscedasticity problems a Breusch-Pagan test was applied with the following decision rule. If the p-value of the Breusch-Pagan test is less than any of the chosen significance levels (i.e., 10%, 5%, and 1%) which indicates that there is a probable problem of heteroscedasticity. Accordingly, the result shows that existence of the problem (see table: 3 Appendix). Therefore, the remedial measure to overcome such problem is robust the standard error estimation. Accordingly, the researcher was robust the data to overcomes the problem.

Model specification test

Testing the model specification is very important to check out whether one or more relevant variables are omitted from the model or one or more irrelevant variables are included in the model. There are different methods to check the specification error of the model. The Ramsey reset test for omitted variables is commonly used. Therefore, the Ramsey reset test for omitted variables was tested. The result depicted in (Appendix 7.2) the null hypothesis that Ho model has no omitted variables, as a decision rule according to Ramsey reset test, a model specification is fit for regression analysis if the p-values stated in P>F is greater than the chosen level of significances (i.e., 1%, 5% and 10%). According to this taste indicates the model has no relevant omitted variable since the test was failed to reject the null hypothesis. Prop> F is 1% which is greater than any of the significance levels of the specified model of the study.

Since the result of all tests shows that, absence of the problem list above. Therefore, all selected variables were decided to be included in the model analysis.

4.2.2. Determinants of income diversification participation decision and its intensity

The possible econometric model expected to be employed in the analysis of income diversification participation decision and the intensity of income diversification were Topit model, Heckman two-step, Heckman maximum likelihood model, and Double-hurdle model. So, it is very important to identify which econometric model is important when dealing with such kinds of investigation. Therefore, there are some criteria or tests that are used to choose the best fit model among alternative. Those are significance of inverse mill ratio, likelihood ratio test of the independent equation, akakias information criteria and likelihood ratio,

Accordingly, the AIC also shows that the double-hurdle model is preferred than Tobit model since the value of the test statistic from Double -hurdle (134.01) is lower than that of Tobit model (227.52). This result implies that Tobit model was rejected in favor of Double-hurdle model for analyzing factors that affecting income diversification participation decision and intensity participation.

Moreover, Heckman's two-step model is an econometric model developed to correct for sample selection bias. In this study, the result from the Heckman two-step showed that the inverse mills ratio (IMR) or Lambda value (0.816) was found statistically insignificant. The result implies no sample selection bias in the data. Therefore, no need to use the Heckman two-step model (see appendix table 4). In addition to this, the result of Heckman's maximum likelihood model output showed that the two equations are independent because the null that the participation decision in income diversification and intensity of income diversification are independent is accepted (Appendix table 4). The independence of the two equations suggests permissibility of analyzing the two equations separately using probit and truncated regression model which is double-hurdle model. Therefore, the double-hurdle model was employed in this study.

Therefore, the Stata command ‘Craggit’ was used in Stata version 16 for the estimation of the double hurdle model to identify factor affecting income diversification participation decision of the female headed household with its intensity. In this case, since double hurdle model is the combination of probit and truncated regression model, the income diversification participation decision and its intensity (extent of diversification) was separately estimated and the model output estimation using ‘Craggit’ command was placed under (appendix Table 5). The summation of log pseudo likelihood generated from the separate probit and truncated regression

is equal with that generated by craggit command. This implies that, no matter whether double-hurdle model is estimated by 'Craggit' command or separately using probit and truncated regression model.

Double -hurdle model was used to identify factors affecting income diversification participation decision and its intensity in the study area. The overall significance and goodness of fit of the model was checked with the value of Wald chi square value of 130.65 at 14 degree of freedom is statistically significant at less than 1% significance level. The log pseudo likelihood value of 36.007 which indicates that assumption of null hypothesis that all explanatory variables in the regression model is simultaneously equal to zero is rejected at less than 1% level of significance (see appendix table 5)

4.2.2.1. Determinants of income diversification participation decision

The first hurdle of double hurdle model result point out that, out of the 14-variable hypothesized to affect household heads in income diversification participation decision, eight (8) explanatory variables namely: age of household heads, education status of the household heads, family size, livestock owned, access of credit service, gender-based discrimination, distance from the nearest market and agricultural income were significantly determine household's decision to participate in income diversification.

Age of the Female-headed households (AGE): The double hurdle model results of the first hurdle shows that, age of the female-headed household was found statistically significant and negatively related to the probability of participation in income diversification at 1% significance level. The marginal effect result revealed that, a year increased in the age of female headed household results in 2.5% decrease in the probability of participation in income diversification on average, ceteris paribus. This might be due to the fact that even if experience increases with age of household head, at the same time the capacity of households to do multiple sources of income decreases. Moreover, old age induce household heads not to participate in different income generating activities rather to consume more leisure. The result consistent with the finding of (Idris-Adeniyi et al. 2020; Ehiakpor *et al.*2020; Debesai, 2020) who found that as age of farmers increase, the probability of participation in non-farm income activities decrease, which show negative relationship between age of the household and income diversification participation decision. And also the result disagree with the report of previous finding of Gebiso

et al. (2019) which shows positive relationship between age and income diversification decision participation.

Education level of female-headed household (EDU): Education is a powerful tool that increases the productivity of human capital by making people aware of the various opportunities for the multiplying their source of income. Many empirical studies, like Abbeam *et al*.(2020) apply level of education. While, Dagunga *et al.*(2018) have operationalized education as the number of years in school. Accordingly, this study used year of schooling for the analysis due to educational states of the respondent is not leveled in such away.

Therefore, education level of the female-headed household was hypothesized to affect household decision to participate in income diversification since it is assumed that increments in education level of household head increase the probability of participation in income diversification. As it was hypothesized, the econometric result shows that there is positive and significant relationship between the educational status of female-headed households and the decision to participation in income diversification at 1% significant level. The marginal effects indicate that as education level of female-headed households increase by 1 year of schooling, the probability of participating in income diversification could increase by 7.4% on average, keeping other things remain constant. This may be due to the fact that female headed household with more educational level have better to locate their time on different source of income since the agricultural income is seasonal. Moreover, it also evident that educated farmer's tendency to accept different agricultural technology which result for increasing on-farm and off-farm income. This result is consistent with the previous finding conducted by Abbeam *et al*.(2020) and Khan *et al.* (2019) which show education increase the probability of participating in to income diversification.

Family size (FSIZE): Family size measured in adult equivalent was hypothesized to have either positive or negative effect on the probability of income diversification participation decision of the female-headed household. As it was hypothesized earlier, the econometric result from the double-hurdle model point out positive and significant influence of family size on the probability of female headed households participating income diversification at 1% significance level. Accordingly, the marginal effect shows that as family member increase by one adult equivalent, income diversification increased by 6.7%. This result is found as expected because households

with more family size need additional income rather than on-farm income to overcome the variability (shock) of agriculture output. Therefore, this finding is similar with the previous finding of Khan et al. (2019); Idris-Adeniyi et al. (2020); Adem and Tesafa. (2020);Teji, (2020) which showed that increments in the number of family size increase the probabilities of household heads participate in income diversification. The result also contradict with the finding of Onadeko et al. (2020) which point out that existence of negative relationship between family size and income diversification participation decision.

Total livestock ownership of the households (LOWNER): Livestock is important major assets for household in the rural area. They are used as Food, income, for organic fertilizer and domestic fuel supply. Therefore, livestock ownership measured in the tropical livestock unit to have positive or negative effect on the income diversification of household heads. As hypothesized above, it was found to have positive and significant effect on the probability of multiplying income source of the female-headed households at 5% significance level. The marginal effect indicates that, a one TLU increase in livestock holding increases households' participation in income diversification by 1.9%, keeping other variables constant. This result is agree with previous study conducted by Gebru et al. (2018); Balense and Debebe, (2020)that found as the livestock holding of the household head increased the participation in income diversification decrease. The possible reason is that household with relatively more livestock make use of the income obtained from livestock for expanding non-farm income activities.

Access to credit service (ACREDIT): The coefficient of dummy access credit service "yes" of income diversifier was positive and significantly affects income diversification at 10% significance level. As it was hypothesized earlier access to formal or informal credit service induces income diversification and there was strong positive relationship between access to credit and income diversification participation decision. Therefore, the marginal effect shows that as the credit use of female-headed household's increase, the probability of participation in income diversification increase by 15.58% on average, keeping other factors constant. The accessed credit service was highlighted as an important to increase the probability of increasing household's income diversification participation decision. The result implies access to credit service improves the financial capacity of income diversifier to start up new business and expand the existing one. This finding is similar with the previous findings conducted by Debesai, (2020)

found that access to credit from both formal and informal sources had a positive and significant effect on income diversification participation decision. Whereas, the result contradict with previous finding conducted by Diep and Vien, (2017) which shows there is negative relationship between access of credit service and income diversification participation decision.

Gender-based discrimination (GENDISC): Gender-based discrimination is one of the explanatory variables hypothesized to affect participation decision in to income diversification of female-headed household. As it was hypothesized so far, the econometric result from first hurdle shows negative and significant effect of gender-based discrimination on the participation decision female-headed households in income diversification at 1% of significance level. The marginal effect shows that, female-headed households those more victim of discrimination have approximately 26.63% less probability of participating in income diversification compared to non-victim of discrimination. In other way, the households those victims of discrimination, decrease their probability of participating by 26.63% on average, keeping other things remain constant. This implies that, cultural and social point of view of society about female-headed household's participation in agricultural and non- agricultural lobur force in the rural area is very low.

Distance from the nearest market (DMRK): The distance from the nearest market place was assumed to affect the households participate in income diversification. As it was hypothesized earlier, the econometric result shows that there was negative and statistically significant relationship between distance from the nearest market and the female-headed households' decision to participate in income diversification at 1% significant level. This is may be due to female headed households who far ways from market center face greater challenges like lack infrastructure, fear of robber, and transaction cost. Therefore, the marginal effects shows that as the distance from female-headed households house from the nearest center of the market increase by one km, the probability of participate in income diversification decrease by 6.97% on average, keeping other things constant. This implies that being far from the market center reduce the chance of in income source multiplication. This result is consistent with result reported by Ababbo, (2015); Kassie et al. (2017); Balense and Debebe (2020); Abbeam *et al*.(2020) which shows that distance from the market center is negatively related with probability of participating in to income diversification

Table 4.8 Factors affecting income diversification participation decision of the household (first hurdle/ probit model)

PHID	Coef.	Robust Std. Err	z	P>z	Marginal effect
Age of household	-.0666405***	.0162069	-4.11	0.000	-.0258076
Education level of household	.1933173***	.0373476	5.18	0.000	.0748652
Family size	.1747055***	.0504816	3.46	0.001	.0676575
Dependency ratio	.3394871	.822412	0.41	0.680	.1314718
Land size	-.0405711	.1242216	-0.33	0.744	-.0157118
Livestock owned	.0501399**	.0209196	2.40	0.017	.0194175
Frequency of extension contact	.0210757	.112018	0.19	0.851	.0081619
Member of farmer-based organization	.0643465	.2402942	0.27	0.789	.0248983
Access of credit	.4011732 *	.2258163	1.78	0.076	.1558104
Gender based discrimination	-.7112686***	.234146	-3.04	0.002	-.2663482
Pre received shock	.2740129	.2345968	1.17	0.243	.1070018
Access of irrigated land	.1778446	.2450437	0.73	0.468	.0693047
Distance from the market center	-.1800057 ***	.0377932	-4.76	0.000	-.0697101
Annual average agricultural income	-.1259735**	.0586057	-2.15	0.032	-.0487852
_cons	2.803885**	1.131306	2.48	0.013	
Log pseudo likelihood = -82.529353					
Wald chi2(14) = 130.65					
Prob > chi2 = 0.0000					
Pseudo R2 = 0.5090					

Source: Own computation (2021)
Note: ***, **, * shows 1%, 5%, and 10% significance level
The dependent variable (income diversification participation decision) is a binary variable that takes the value 1 if the household head had participated in income diversification, 0 otherwise

Annual average agricultural income (LnAAI): In line with hypothesis annual average agricultural income has negative and significantly affected the income diversification participation decision at 5% level of significance. Therefore, the negative relationship between farm income and income diversification participation decision shows that, the female household heads who generate relatively large income from both crop and livestock production less diversify income than those who get lower agricultural income. The marginal effect result shows

that each additional one birr from agricultural income decrease the probability of female-headed household participation in income diversification by 4.8% on average, keeping other things remain constant. This might be because, availability of finance increase the probability of participation in income diversification but at the same time in the rural area the time of female household was consumed by home practiced work. This result is inconsistent with previous findings conducted by Ababbo, (2015); Balense and Debebe (2020) which show a positive relationship between agricultural income and income diversification participation decisions.

4.2.2.2. Determinants of intensity of participation in income diversification

The second hurdle of the double hurdle model point out that, the intensity of the female-headed households participation in income diversification as a fraction of extent of income diversification which measured by Simpson income diversification index was presented in the following table 4.10. The result showed that out of 14 explanatory variables included in the model, five (5) of them were found statistically significant determinants of the intensity of income diversification. Among those five variables, age of the female-headed household, family size, and gender-based discrimination were negatively and significantly affect intensity of income diversification. While land size and frequency of the extension contact positively and significantly affect intensity of income diversification.

Age of the household (AGE): The age of female-headed households were assumed to affect the level of participation in income diversification. As it was hypothesized age of the female- headed household was found to have negative and significant effect on the intensity of income diversification participation at 10% significance level. The econometric model results show that, as the age of the female-headed households increase by one year, the intensity of income diversification decreases by 0.004 on average, other things remain constant. The implication is that the elder rural female-headed household being well established, more experienced in agricultural production, allocate more time for homework and more resistance to new ideas and information that hinder them from diversifying their income source. Whereas, younger rural female-headed households cannot get enough farmland to support their families, and thus they get into diversified venture activities. This result agrees with the findings of Adem and Tesafa, (2020) that found negative relationship between extent of income diversification and age of rural farmer households. Nevertheless, the result disagree with the finding of Diep and Vien, (2017)

which shows positive relationship between age of rural household and extent of income diversification.

Family size (FSIZE): As it was hypothesized so far, household size of female-headed household was found to have negative and significant effect on degree of income diversification at 10% significance level. The econometric result revealed that as the household member increase by one-man equivalent unit, the female-headed household degree of income diversification participation decision decrease by 0.012 on average, keeping other factor constant. This is may be due to the female household heads have more responsible in-home management than income-earning activity, culturally define roles, social mobility limitations, and differential ownership of or access to assets. The finding is inconsistent with the findings of Teshome and Edriss, (2013) which shows positive relationship between family size and intensity of income diversification of the rural household.

Land size (LSIZE): Land size holding of the female-headed household was one of the variables that hypothesized to have a positive and significant influence on level/degree of income diversification. As it was hypothesized earlier, land size holding had positive and significant effect on the degree of income diversification at 5% significance level. The econometric result shows that, having one more extra hectare of cultivated land size could increase the level of participating in income diversification by 0.028 on average; by keeping other factors remain constant. This implies that, when we compare those households own less cultivated land size with more own land size will participate income diversification. Because of those household who has more cultivated land size rent some part of it to generate non-farm income and increase source of their income. This agree with finding of Gebiso *et al.* (2019) who found positive relationship between land size and intensity of income diversification.

Frequency of extension contact (FECONTACT): As it was hypothesized, frequency of extension contact had positive and statistically significant effect on female-headed household intensity of participation in income diversification at 5% level of significance. Thus, the econometric model result from second hurdle point out that availability of extension contact increment by one day leads to increase the intensity or level of participating in income diversification by 0.034 on average, keeping other thing constant.

Table 4.9: Factor affecting extent of income diversification (Second hurdle/ truncated regression)

SID	Coef.	Robust Std. Err	z	P>z
Age of household	-.0041532*	.0023956	-1.80	0.072
Education level of household	.0070344	.0053221	1.34	0.179
Family size	-.0126545*	.0071539	-1.77	0.077
Dependency ratio	.0856053	.1250178	0.68	0.494
Land size	.0286801**	.0140315	2.04	0.041
Livestock owned	.0001524	.0031408	0.05	0.961
Frequency of extension contact	.0341251**	.0147913	2.31	0.021
Member of farmer-based organization	.0216135	.0147528	0.65	0.513
Access of credit service	.0148424	.0330701	0.49	0.628
Gender based discrimination	-.1297951 ***	.0306027	-2.83	0.005
Pre received shock	.0029711	.045926	0.08	0.933
Access of irrigated land	-.0070482	.0354324	-0.21	0.832
Distance from the market center	.0094182	.033184	1.42	0.154
Annual average agricultural income	.0062774	.0066136	0.74	0.461
_cons	.5311942***	.0085201	3.26	0.001
/Sigma	.1585477***	.0117808	13.46	0.000
Number of obs = 108 Wald chi2(14) = 56.48 Prob > chi2 = 0.0000 Log pseudo likelihood = 46.521812				

Source: Own computation (2021)
Note: ***, **, * shows 1%, 5%, and 10% significance level respectively.
The dependent variable (Simpson income diversification index); SID refers to income diversification index which is continuous.

This implies that, existence of rural development agent in the rural area helps the rural society by increasing awareness about merit of income .The result is consistent with the previous finding by Yishak , (2017); Gebiso *et al.* (2019); Teji,(2020) which shows existence of positive relationship between frequency of extension contact and intensity of income diversification.

Gender based discrimination (GENDIS): Gender-based discrimination was assumed to affect the level of FHH participation in income diversification. As it was hypothesized earlier, the

econometric result from second hurdle shows that there was negative and statistically significant relationship between gender-based discrimination on female-headed household and intensity of participation in income diversification at 1% of significance level. Therefore, the regression coefficient result shows that being victim of gender-based discrimination decrease the intensity or level participation in income diversification of female-headed household by 0.129 on average, keeping other factor constant. The implication is that female-headed household discriminated from rural labor participation into off-farm and non-farm income generating activities duo to culturally defined roles, social mobility limitations, and differential ownership of or access to assets.

4.2.3 Impact of income diversification on the rural female-headed household livelihoods

This section presents the entire process of implementation of propensity score matching model (PSM), to analysis the impact of the income diversification on the livelihood of rural sample households. Specifically, it discusses the estimation of the propensity score, common support region, matching algorithm and balancing test. Lastly, it provides the income diversification effect among the treated Female-headed household.

4.2.3.1 Estimation of the propensity scores

Under this sub section, to address the third objective of the study, the propensity score matching (PSM) method was applied. Therefore, to estimate the propensity score matching the appropriate econometric models were logit and probit model. Since the outcome of both models is the same, the researcher wants to use logit model to estimate propensity score matching for diversified (treated) and non-diversified (control) group of female head households.

Therefore, the estimation of propensity score matching is important for the following purpose in the impact analysis using PSM. Firstly, it is used to identify and define the common support region for the control and treated group. Secondly, it’s used to estimate the average treatment effect on the treated for the sample. Finally, propensity score matching allow the researcher to directly address the question of what can be earned from diversifying income source and the loss of being specializing income source.

As explained below table 4.13, 14 matching variables have been used in the model as explanatory variables like age, education level of household, family size, dependency ratio, land

size, livestock owned, frequency of extension contact, member of farmer-based organization, access of credit service, gender based discrimination, pre received shock, access of irrigated land, distance from the market center, and Annual average agricultural income were taken to estimate.

4.2.3.2 The Common support region

After estimating propensity score for household heads who diversify their income source and specialize their source of income. The next step would be identifying common support region, depending on the propensity score distribution of income diversified household heads and non-diversified household head. Accordingly, the common support region was estimated. As shown on the table 4.10 below, the estimated propensity scores lie between 0.1154501 and 0.9853427 (mean = 0.5588739) for the diversified (treated group) of female head households and 0.0060424 and 0.8818751 (mean = 0.3485615) for the non-diversified (control group) of female headed households. Therefore, the common support region would lie between 0.1154501 and 0.8818751. Households whose estimated propensity scores are less than 0.1154501 and larger than 0.8818751 are not considered for the matching exercise.

Table 4.10: Distribution of estimated propensity score

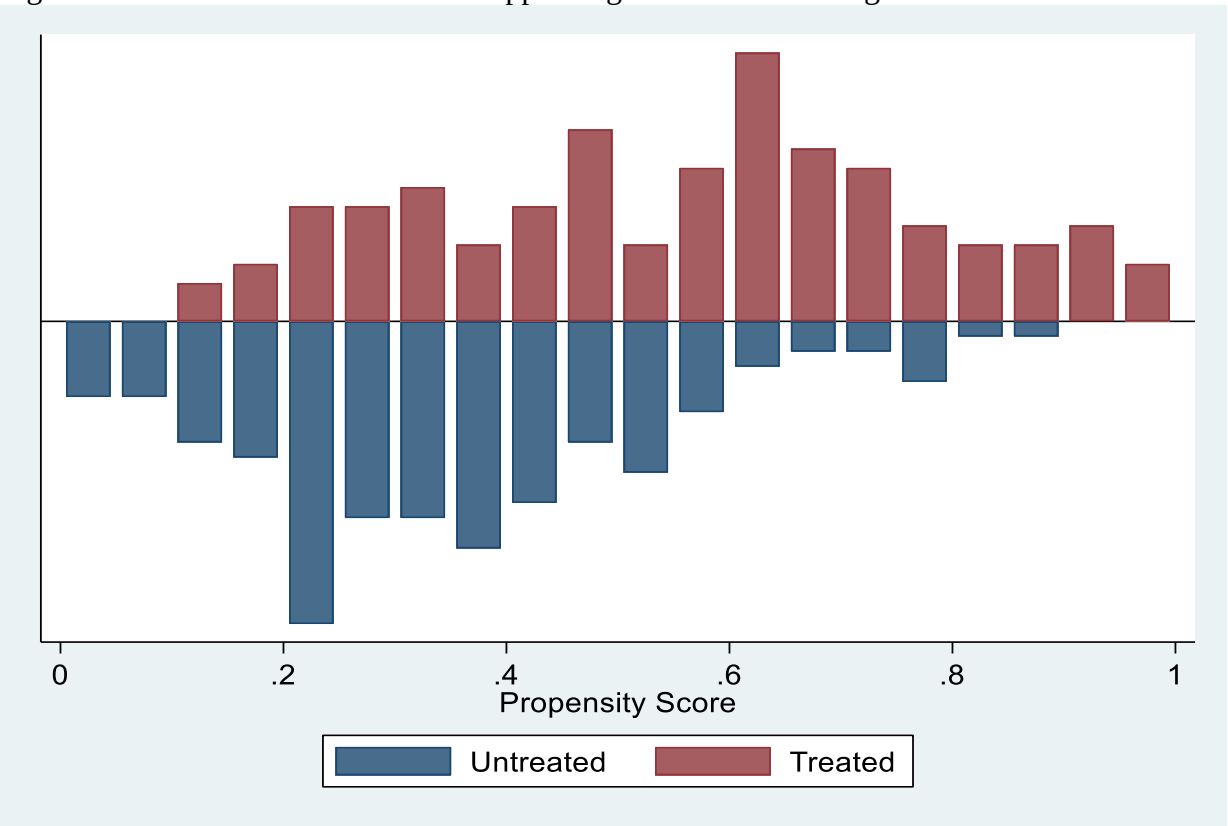
Group	Observation	Mean	Std.Dev.	Minimum	Maximum
Total household	245	.4412706	.2276207	.0060424	.9853427
Diversified	108	.5588739	.2201783	**.1154501**	.9853427
Non-diversified	137	.3485615	.1875445	.0060424	**.8818751**

Source: computation based on survey data (2020)

Figure 4.5: Below, describes the common support region of the untreated and treated group for the outcome variables. Accordingly, 21 or 8.57% observations (9 from treated and 12 from untreated) were discarded from the study in the impact assessment procedure. From the total sampled household 224 observations (91.43%) are on support from both the treated and control groups. This result show 99 diversified and 125 non-diversified groups behave similar characteristics and matched. The decisions for the off and on support observations are based on the summarized p-scores in the treatment and untreated groups and count how many units are off support.

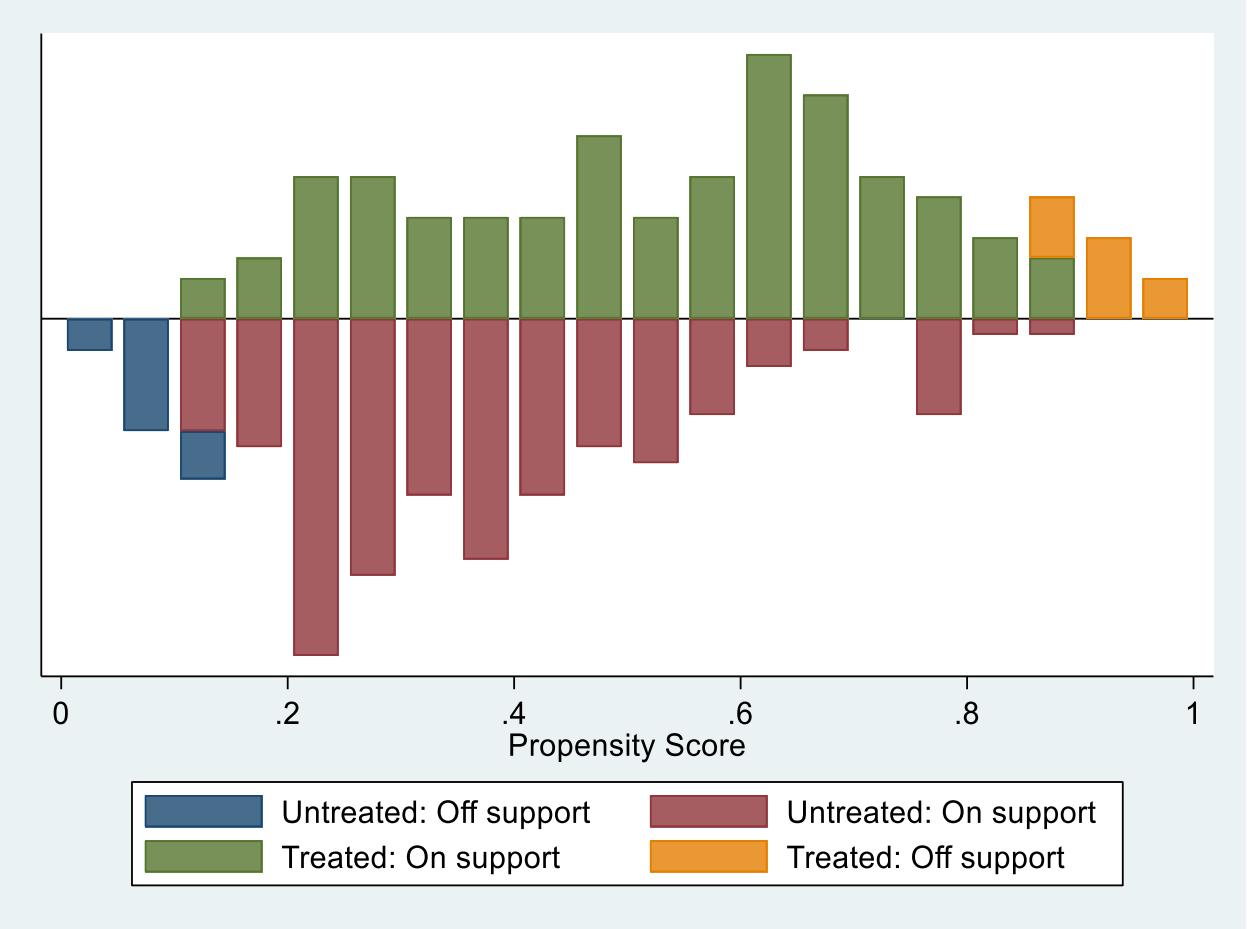
Using color codes, the diagram below distinguishes between the regions of common support. The upper green color shows the treated on support observations. The upper yellow color one indicates the treated off support, the lower blue color indicates the untreated off support and the lower red color represents untreated on support. As a result, the majority of the observations are concentrated in the common support region. This verified that there is enough overlap between the characteristics of treated and untreated components to find enough matches, according to Heckman (1997). As a result, the common support assumption is satisfied.

Figure 4.3 : Presentation of common support region before matching



Source: computation based on survey data (2021)

Figure 4.4 : Presentation of common support region after matching



Source: computation based on survey data (2021)

4.2.3.3 Choice of matching algorithm

After identifying common support region, different matching algorisms were used in matching income diversified (treated) with non-diversified (control) female headed households in the common support region. According to Mulatu (2017) to estimate the effect of treatment on treated, PSM with different matching algorithms were applied. Those are; nearest neighbor matching (NNM), radius matching (RM), caliper matching, and kernel matching (KM) were most frequently used. The final choice of matching algorism is based on three criteria: namely equal mean test (balancing test), pseudo R2 and size of matched sample. Generally, the benefits and drawbacks of each algorithm are clear in theoretical and simulated research, but there is always confusion in practice regarding which method is the best for matching (Khandker et al., 2010). Matching algorism which balances all explanatory variables (result in insignificant mean differences between the two groups), bear low pseudo R2 value and results in large sample size

is preferable (Dehejia and Wahba, 2002). Thus, based on these criteria, kernel (band width 0.1) was selected for this study in which the mean difference of the two groups explanatory variables were insignificant, pseudo R2 was the lowest compared to other matching estimators and finally balance 224 sample sizes. Therefore, impact analysis procedure was followed and discussed using kernel (band width 0.1).

Table 4.11: Performance of matching estimators

Matching estimators	Balance test	Pseudo R2	Matched sample size
Nearest neighbor			
NN_1	11	0.047	224
NN_2	13	0.044	224
NN_3	13	0.030	224
NN_4	14	0.029	224
NN_5	13	0.032	224
Caliper			
0.01	13	0.045	197
0.1	11	0.047	224
0.25	11	0.047	224
0.50	11	0.047	224
Radius			
0.01	11	0.122	224
0.1	11	0.122	224
0.25	11	0.122	224
0.50	11	0.122	224
Kernel			
Band width 0.01	13	0.039	197
Band width 0.1	**14**	**0.017**	**224**
Band width 0.25	14	0.020	224
Band width 0.50	13	0.073	224

Source: Own computation based on survey data (2021)

4.2.3.4. Testing the balance of propensity score and covariates

After choosing the best matching estimator, the next activity is to check the balancing of propensity score and covariates using various techniques by the chosen matching estimator. The primary purpose of the propensity score estimation is not to obtain a precise prediction of selection into treatment, but rather to balance the distribution of covariates in both groups. The balancing powers of the estimation is ensured using different test methods such as the reduction in the mean standardized bias between the matched and unmatched households, equality of means using t-test and chi-square test for joint significance for the variables used and pseudo R2.

The mean standardized bias before and after matching, as well as the total bias reduction achieved by the matching technique, is shown in table 4.13. The standardized bias difference in covariates before matching was lie between 2% and 51.1% in absolute value. However, after matching the standardized bias difference for all covariates was lie between 0.8% and 12% in absolute value, which is less than the critical level of 20% as explained by (Rosenbaum & Rubin, 1985). In every scenario, it is obvious that sample differences in the unmatched data significantly exceed those in the samples of matched cases. As a result of the matching process, the treatment and control samples have a high degree of covariate balance that is ready to utilize in the estimate phase. Similarly, the t-test demonstrated that after matching, all covariates became insignificant. As a result, the propensity score's matching quality or balance for all covariates has been recognized.

Table 4.12: Propensity score and covariate balance test

Variable	Unmatched Matched	Mean		Standard bias %	Reduction bias %	t-test	p>|t|
		treated	control				
AGE	Unmatched	41.731	43.898	-31.0		-2.40	0.017
	Matched	42.303	41.851	6.5	79.2	0.50	0.618
EDL	Unmatched	3.8241	2.2774	51.1		4.02	0.000
	Matched	3.4545	3.5058	-1.7	96.7	-0.12	0.906
FSIZE	Unmatched	6.6111	5.8905	33.9		2.65	0.009
	Matched	6.4545	6.2363	10.3	69.7	0.72	0.471
DRATIO	Unmatched	0.11491	0.09934	11.5		0.89	0.374
	Matched	0.11273	0.10347	6.8	40.5	0.47	0.640
LSIZE	Unmatched	2.6204	2.6015	2.0		0.16	0.876
	Matched	2.6187	2.6258	-0.8	62.2	-0.05	0.960
LOWNED	Unmatched	10.505	10.991	-8.8		-0.68	0.496
	Matched	10.529	11.041	-9.2	-5.3	-0.64	0.522
FECONTACT	Unmatched	1.1204	.94891	17.0		1.32	0.188
	Matched	1.0606	.92524	13.4	21.1	0.96	0.336
MFBO	Unmatched	0.55556	0.44526	22.1		1.72	0.087
	Matched	0.53535	0.59504	-12.0	45.9	45.9	0.400
ACREDIT	Unmatched	0.40741	0.39416	2.7		0.21	0.21
	Matched	0.41414	0.43865	-5.0	-85.0	-0.35	0.729
GENDERD	Unmatched	0.37037	0.45985	-18.2		-1.41	0.160
	Matched	0.39394	0.4211	-5.5	69.6	-0.39	0.699
RSH	Unmatched	0.36111	0.25547	22.9		1.79	0.074
	Matched	0.34259	0.30053	9.3	59.4	0.64	0.521
AILAND	Unmatched	0.34259	0.32117	4.5		0.35	0.725
	Matched	0.34343	0.36564	-4.7	-3.6	-0.32	0.746
DMRK	Unmatched	5.5139	7.0657	-48.9		-3.75	0.000
	Matched	5.6717	5.7263	-1.7	96.5	-0.13	0.898
LnAAI	Unmatched	9.8368	9.7887	3.3		0.25	0.803
	Matched	9.7459	9.9606	-9.4	-184.0	-0.76	0.450

The standardized mean difference of overall variables utilized in the propensity score (25.1 % before matching) is decreased to about (6.6 % after matching), as shown in Table 1.14 below. Furthermore, the p-values of the likelihood ratio tests show that the joint significance of the covariates was always rejected after matching whereas it was never rejected before matching. The low pseudo R2, low standardized bias, high total bias reduction, and the insignificant p-values of the likelihood ratio test after matching suggest that the specification of the propensity is successful in terms of balancing the distribution of covariates between the treated and control groups. Therefore, the result used for analysis the effect of income diversification participation on rural female headed household's income, consumption spent and saving for those having similar observed characteristics. In addition, it allows comparing observed outcomes for diversified (treated) with those of non-diversified (control) groups found in a common support region.

Table 4:1 3 Chi2 test for the joint significance of variables

Sample	Pseudo R2	LR chi2	p>chi2	Mean Bias	Med Bias	B	R	%Var
Unmatched	0.168	56.64	0.000	25.4	18.2	102.0*	1.28	30
Matched	0.065	11.54	0.714	6.6	6.5	30.6*	1.28	10

Source: own computation based on survey data (2021)

All of the above tests suggest that the matching algorism chosen is relatively best with the data at hand. Thus, we can proceed to estimate ATT for the sample households.

4.2.3. 5. Estimating the average treatment effect on the treated (ATT)

The impact of treatment variable (income diversification) on the outcome variables (income, consumption expenditure, and saving) per capita of female headed household which measured by Ethiopian Birr (ETB) was analyzed. The purpose of these all process was to see whether the diversified female headed households have significant difference in livelihood indicator compared to non-diversified female headed households or not. To identify this, there are three parameters; ATE, ATU and ATT (Heinrich et al., 2016). But, ATU measures the impact that the program would have on those who did not participate and ATE also measure the treatment impact without excluding the discord individual from the matching (below the minimum of

diversified and above maximum of non-diversified). As a result, both ATE and ATU are not reveal the true impact of diversification and might not be of relevance to policy makers since it does not consider into account the common support assumption. This implies households who were highly motivated and the households who had extremely low motivated to diversify included in treatment effect (ATE) are included in the analysis (Caliendo and Kopeinig, 2008). Therefore, the average treatment effect on the treated (ATT) was computed to evaluate explicitly the impact on those for whom at least the probability to diversify was approximated. According to table 4.14 below shows the ATT from kernel (band width 0.1) matching results, annual household income, consumption expenditure, and saving amount were considered as indicator of livelihood.

Accordingly, the average treatment effect of the treated (ATT) on household annual average income for female headed household who participation in income diversification was 11122.89 ETB per a year. That means female head households who had participated in income diversification had increased their total income on average by 11122.89 ETB per a year than non-diversified (control group) household heads and it was statistically positive relationship between income diversification and income. The positive and significant impact of the diversification of income estimated in this study is confirm with the results of recent studies (Diiro and Sam, 2016; Ghosh and Sujan 2020).

Beside this, the result also show the consumption expenditure of the female headed households who were diversify their source of income was match greater with 90449.92 Ethiopian birr than non-diversified (control) group. This is may be due to the agricultural income of less developed counties influenced by many factors like flowed, heavy rain, inefficient technology and etc. As a result, rural female headed households diversify their income to overcome livelihood constraints. The result also agree with earlier studies (Salam et al., 2019)which found that positive relationship between income source diversification and consumption expenditure.

The results also show that the ATT on household amount of annual average saving for female headed household who participate in income diversification was 3745.95 ETB per a year. This result indicates that, the female head households who had participated in income diversification had increased their annual average saving by 2037.59 ETB per capita as compared with non-

diversified (control group) female headed household and there was statistically positive relationship between income diversification and saving.

Table 4.14: The result of average treatment effect on the treated

Outcome variable	Sample	Treated	Control	Difference	S. E	t-stat
lnINCOME	ATT	10.4862193	10.1367831	.349436155	.056432021	6.19
lnCONEX	ATT	10.3695869	10.0625425	.307044328	.061439526	5.00
lnSAVING	ATT	7.65602528	6.15147938	1.5045459	.441287554	3.41

Source: computation based on survey data (2021)

4.2.3.5. Sensitivity analysis

In order to check for unobservable biases, sensitivity analysis was performed on the computed outcome variables using Rosenbaum Bounding approach with respect to deviation from the conditional independence assumption. Therefore, this sensitivity analysis method designed to check how the ATT result deviate if the unobserved covariates have been allowed to differ among treated and controlled female headed households. The results show, the impact of income diversification does not change even though the diversified and non-diversified households were allowed to differ in their odds of being treated up to 300% ($e^{\gamma} = 3$) in terms of unobserved covariates**.** This implies that, for outcome variable computed at different level of critical value of gamma, the p-critical values (upper bound significance level or Sig+) were statistically significant. The researcher couldn't get the critical value gamma where the estimated ATT is questioned even if the researcher has set gamma largely up to 3 (see appendix table 7). Thus, it can be concluded that the impact estimate (ATT) is insensitive to the hidden bias and the result is pure effect of participation in income diversification on households' livelihoods.

4.3. Opportunities and Challenges of Income Diversification

4.3.1. Major reasons for participation into income diversification

Depending on the response of the respondents; Decision of rural female-headed households concerning involvement in off farm and non-farm activities depend on two major factors: these are push and pull factors. Accordingly, the survey tried to identify major opportunities that motivate the female-headed households to diversify their source of income. Consequentially, the sample respondents were asked to rank the opportunities they face according to their concern

based on two categories (push and pull) factors. The results of the subjective assessment of the sample female-headed households are summarized in table 4.15 below.

According to the descriptive statistics indicated that 23%, 15%, 11%, 10% and 6% of the sample female-headed household farmers reported that scarcity of land, large family size, seasonality of agriculture, inadequacy of farm income, and illness of family member were the first and the most important opportunities from pushed factors in the study area, respectively. Similarly, the result also point out that, 15.3%, 9%, and 8% of sample female-headed household farmers reported that proximity to urban, availability of non-farm and off-farm income, and near to the road were ranked as the first motivated factors respectively from pull side of opportunities of income diversification in study area.

In addition, the result of the descriptive statistics showed that 19.4%, 15.7%, 14%, and 1.8% of sample households report that large family size, seasonality of agriculture, scarcity of land, and inadequacy of farm income were the second most pushed factors to female-headed households participate in income diversification respectively. Similarly, the result also show that 26.8%, 9.2%, and 5.5% of sample female-headed households priorities as second most pull factor opportunities like availability of non-farm and off-farm income, near to the road, and proximity to urban respectively.

Likewise, the result of the assessment on opportunity that motivate female headed-households to participate in income diversification shows that 36.1%, 14.8%, 9.2%, 6.48%, and 3.7% of sample farmers report that illness of family member, large family size, scarcity of land, Inadequacy of farm income, and seasonality of agriculture were considered as the top opportunities from push factors ranked as the third important motivation in their priority setting, respectively. Moreover, 20.3%, and 10.1% of sample households were pull to income diversification due to near to the road and availability of non-farm and off-farm income, respectively and they ranked as the third important motivation form pull side.

Generally, interview was undertaken to collect an information concerning factors enforcing female-headed household to diversify their income source. Accordingly, those factors are broadly categorized in to two, push and pull factors. Based on this, the summary result of the survey shows that the availability of non-farm and off-farm income, large family size, illness of family member, near to the road and scarcity of land were the five most important factors that

motivates female-headed households at 15.73%, 15.14%, 14.19%, 12.93%, and 12.61% as reported by interviewed female-headed households in their order of priority respectively in table 4.15 summary column.

Table 4.15: Reasons for participating in off-farm and non-farm activities

Opportunity	Rank						Summary	
	1st		2nd		3rd			
	No	%	No	%	No	%	No	%
Push factors								
Scarcity of land	25	23%	15	14%	10	9.2%	40	12.61%
Inadequacy of farm income	10	10%	2	1.8%	7	6.48%	17	5.36%
Large family size	16	15%	17	15.7%	15	14.8%	48	15.14%
Seasonality of agriculture	12	11%	21	19.4%	4	3.7%	37	11.67%
Illness of family member	6	6%	0	0%	39	36.1%	45	14.19%
Pull factors								
Availability of non-farm and off-farm income	10	9%	29	26.8%	11	10.1%	50	15.73%
Near to the road	9	8%	10	9.2%	22	20.3%	41	12.93%
Proximity to urban	17	15.3%	6	5.5%	0	0%	23	7.25%
Other factors	3	2.7%	8	7.4%	5	4.6%	16	5.04%
Total sum	108	100%	108	100%	108	100%	317	100%

Source: computed from own survey data, (2021)

4.3.2. Barriers to participate in income diversification in the study area.

Sample households in the study area face a number of constraints *to* diversify their income sources. As their responses majority of them had an interest to participate and some of them did not participate; however, they constrained by different factors. The table 4.16 below shows those factors being barriers for female-headed households in the study area. Accordingly, female-headed households were constrained to undertake off-farm and non-farm income generating activities mostly due to:

1. Lack infrastructure: Infrastructure is one of essential element in fatting extreme poverty in rural area. In less developed countries, infrastructure like, road, market channel, electricity, bank, etc. are very low when we compare with developed countries. Accordingly, the survey result point out that from 245 sample respondents about (77.95%) of respondents reported that lack of infrastructure constraint to diversify their income source. As a result most rural female-headed households lost their time in non-income generating activities and the rural livelihood supplement only by farm income in the study area.

2. Altitude of the society: The presented survey result revealed that 67.34% of rural female-headed reported altitude of the society has been a constraint to diversification their income source. However, 32.66% of respondents were reported that altitude of society not as such problem for income source multiplication. This may be due to culturally defined roles, and women considered as low in preforming rural income generating activities.

3. **Adultness:** It is one of the challenges of income diversification in the study areas. Gebru *et al.*, (2018) found that in Ethiopia, age of the household is negatively related with income diversification. The young are more likely to participate in income diversification because of their ambition to invest and accumulate wealth during their working age. The old on the other hand are less likely to participate in income diversification activities. Accordingly, the survey result indicated that 60.40% sample household is adult and they are unable to participate in rural labor force income generating activities and other non-farm activities. As a result, adultness is major barriers to participation in income diversification, especially for female-headed households in the study area.

4. Lack of credit: From a survey of a sample of 245 rural female-headed households in the study area in the year of 2021 more than half of total rural household income came from farming while small proportion got from other sources. Because of the lack of collateral and/or credit history, most farmers are bypassed not only by commercial and national development banks, but also by formal micro-credit institutions. In addition to own sources, farmers thus rely on incomes of friends and relatives, and informal money lenders are there. The presented survey result indicated that 69.38% of respondent reported that the probable constraint of income diversification was lack of finance to invest on income source multiplication, however, the remaining 30.62% reported not as such problem. This clearly indicated that lack of finance is a

major problem in income diversification and this problem may be one of the reasons for low degree of income diversification in the study area.

5. Lack of awareness and training: The survey result indicated that 38.77% of the sample households reported that lack awareness and training as the constraint of income diversification in district. However, the remaining 70.80% of respondents reported that lack awareness and training are not as such challenge for income diversification. But, lack of awareness and training has high contribution for low level of income diversification in the study area, due to the agricultural development agent in rural area concentrate on how production and productivity of agricultural output increase rather than on how rural female-headed household generate additional source of income from agriculture and non-agriculture sectors to supplement the on-farm income.

Table 4.16 Challenges of participation in income diversification

Type of constraint	Response	Proportion	Percentage
Lack of credit	Yes	170	**69.38**
	No	75	30.62
Lack of awareness and training	Yes	74	**30.20**
	No	171	70.80
Altitude of the society	Yes	165	**67.34**
	No	80	32.66
Adultness	Yes	148	**60.40**
	No	97	39.60
Lack infrastructure	Yes	191	**77.95**
	No	54	22.05

Source: computed from own survey data, (2021)

CHAPTER FIVE

5. SUMMARY, CONCLUSION AND RECOMMENDATION

This chapter comprises overall summary of the study, conclusion of the finding, suggestion for future research, and recommendation or policy implication and suggestion for future research derived from the finding.

5.1 Summary

The study was conducted to identify factor that affect income diversification participation decision of female-headed household and its impact on their livelihood in the Guduru district of Horo Guduru Zone, Oromia Regional state, Ethiopia. The study was undertaken with the specific objectives of measuring degree of income diversification, identifying factors affecting income diversification participation decision and intensity of income diversification, to examine the impact of income diversification on livelihood indicators (income, consumption expenditure, and saving), and to identify major challenges and opportunities face the female-headed households in income diversification in Guduru district in the study area.

Data used for the study was collected from female household heads drawn from Guduru district. A multistage sampling method was used to select the households. In the first stage, the district was selected purposively. Secondly, from 25 kebeles of the district 8 kebeles were selected by using simple random sampling technique. Thirdly, 245 sample households were randomly selected based on probability of proportionate to size of female household heads in the eight selected kebeles. The primary data obtained from the sampled household was collected by using structural questionnaire through interview by enumerators. Besides the questionnaires, personal observation and key informal interview were also employed to supplement household survey data. The secondary data used in the study was collected from different journal, reports like CSA, FAO, and report from Guduru district agricultural office.

Descriptive statistics, inferential statistics and econometric models were used to analysis the data using statistical tools such as excel, SPSS and Stata version 16. Therefore, descriptive statistics depict that more of the sample household heads were not participate in income diversification. Simpson index of diversity (SID) was computed to address the first objective of the study. Whereas, the second and third objectives of the study were addressed by using Double-hurdle

and propensity score models respectively. Double-hurdle produces first and second hurdle results. The first hurdle uses probit model by taking female-headed households decision to participate in income diversification as dependent variable and the intensity of participation was used as dependent variable in second hurdle result along with different set of hypothesized explanatory variables

Descriptive statistics revealed that from the total sample households, 55.91% and 44.09% were income diversified and non-diversified respectively. The t-test was computed for continuous variable and the result showed that statistically significant mean difference between income diversified and non-diversified regarding age, education level, family size, dependency ratio, frequency of extension contact, and distance from market center. The chi-square test for dummy variable result indicated that there was statistically significant association between the two groups regarding access of irrigated land, access of credit service, member of farmer based organization, pre received shock, and gender-based discrimination. Result from SID indicates that, degree of income diversification in the study area was 0.27 which show as low level.

In addition to the above, it also revealed that existence of different constraints that hinder FHHs to participate in income diversification like: lack of infrastructure, altitude of the society, adultness, lack of credit, and lack of awareness and training. Beside this, there were opportunities that motivate rural FHHs to participate in income diversification from two side pull factors (availability of non-farm and off-farm income, near to the road, and proximity to urban) and push factors (scarcity of land, inadequacy of farm income, large family size, seasonality of agriculture, and illness of family member).

The econometric result from Double-hurdle model indicate that, eight of the 14 explanatory variables included in the analysis were found to have significant effect on female-headed households decision to participate in income diversification in the first hurdle. Those are; education level, family size, livestock owned, and access to credit were positively affected probability of participation decision in income diversification. Whereas, age, gender-based discrimination, distance from market center, and annual average agricultural income had negative effect. The second hurdle result indicated that the intensity of income diversification was positively and significantly affected by land holding size, and frequency of extension

contact. Whereas, negatively and significantly affected by age, family size, and gender-based discrimination.

The result from Propensity Score Matching model revealed that income diversified female-headed households had higher income, consumption expenditure, and saving per a year than those who did not. The average treatment effect on the treated (ATT) from income diversified households were 11122.89, 9049.92, and 2037.59 for income, consumption expenditure, and saving respectively. The study concluded that income diversification is one key valuable potential solution to enhance livelihood indicators in the study area.

5.2. Conclusion

In this study, efforts were made to analyze the determinants of rural female headed household participation decision in income diversification with its intensity and impact on the livelihoods. Accordingly, SID was computed to measure the degree of income diversification in the study area. Double-hurdle model was used to investigate factors that influence participation decision and intensity of participation in income diversification using first and second hurdle, respectively. Moreover, PSM model was used to investigate the impact of income diversification on livelihoods of rural FHHs in study area.

The survey result show that, from the total sample household heads 137 (55.91%) had a diversity index of 0, about 2(0.8%) sample household heads had diversity index between 0.12 up to 0.3, 36 (14.7%) sample households had between 0.31 up to 0.6 and about 70 (28.58) sample households had diversity index between 0.61 and 0.79. Generally, SID indicates that, degree of income diversification in the study area was 0.27 which show as low level.

From the study it was found that different demographic, socio-economic and institutional factors inhibited most of the rural female-headed households from participation in income diversification. Income diversification participation decision was significantly and positively affected by education level, family size, livestock owned, and access to credit. Whereas, age, gender-based discrimination, distance from market center, and annual average agricultural income had negative and significantly. Participation level was affected by land holding size, and frequency of extension contact positively and significantly. Whereas, negatively and significantly affected by age, family size, and gender-based discrimination. According to the

findings of the impact evaluation, it is learned that income diversification has had a considerable beneficial influence on rural FHHs livelihoods (income, consumption expenditure, and saving)

Opportunities and challenges of income diversification in the study area were clearly identified. As it was reported that, major factors that motivate rural FHH's to diversify their income were like: availability of non-farm and off-farm income, near to the road, and proximity to urban from side of pull factors. Whereas, scarcity of land, inadequacy of farm income, large family size, seasonality of agriculture, and illness of family member were from push factors. Beside this, there were major challenges that hinder rural FHHs to participate in income diversification like lack infrastructure, altitude of the society, adultness, lack of credit, and lack of awareness and training are also main challenges factors reported in income diversification.

5.3 Recommendation

The findings of this study have a wide range of recommendations to enhance the income diversification participation of the rural FHHs in the country in general and Guduru district in particular. Therefore, based on the findings of the study, the following recommendations that can help to design appropriate intervention mechanisms that boast income diversification in the study area were drawn as follows:

In the study area about 56% FHHs were not diversify their income and income diversification level is low compared to other countries. This is may be due to constraint they face in income diversification like: - lack of infrastructure, lack of credit, altitude of the society, and adultness. As a result, more work is needed to engage FHHs to participate in income diversification in various ways by converting exist challenges and push factors into opportunity. Moreover the study's findings may inform policymakers and stockholders who work in this area should design inclusive policy which increase FHHs income source since women are backbone of economy of less developed counties.

The study showed that, age of the female-headed households had negative and significant effect on income diversification participation decision and level of income diversification. This implies that older female-headed households are less in income diversification participation than younger one. This could be due to the fact that aged female-headed households have less power to participate in rural labor market than their counterpart even if they are experienced. Therefore,

Guduru district Women Affairs collaborate with Rural Development Agent of the district should be focused on giving training and awareness for older female-headed households how they diversify their source of income and the relevance it has for them and society as a whole.

The finding shows that, there was a positive and significant relationship between education level and income diversification participation decision of the female household heads. The educated people diversify their income through opting for salaried jobs, self-employment activities, etc., illiterate persons engage themselves in agricultural activities which are mostly valuable to weather and other shocks. So that, society should give priority to the education of the girl child with the hope that this will form the basis for how rural FHHs get alternative income rather than agriculture income in the future. Furthermore, the Ministry of Education and the Ministry of Women and Children's Affairs should implement formal and informal education programs which benefit rural women and rural female-headed household in long and short run respectively.

The study reports that, family size had positive and negative significant effect on the income diversification participation decision and intensity of participation of rural FHHs, respectively. This shows that, female-headed households who have large size of family with in limited land size in rural area may be diversifying their income due to shortage of income from agricultural to survive livelihood constraints. While the levels of diversification is limited since more of their time consumed by home-work. As a result, Guduru district office of job creation should establish more continuous and regular non-farm/off-farm operations, as well as train and support local female-headed households in job creation with their home-work and raise income from those activities.

Livestock ownership had positive and significant impact on the income diversification participation decision female-headed household. Therefore, Guduru district agriculture office and other research institute should be encouraged the female-headed farmers to engage in livestock husbandry through providing with improved livestock production technologies (health service, improved breeds and feeds) to improve production and productivity of the sector, this will ultimately increase the female-headed households income.

Access of the credit service was found to have positive and significant effect on the probability of income diversification participation decision of FHHs. Since financial capital is one of

important precondition to start any income generating activity. Therefore, Government and micro-finance of Guduru district should be design appropriate policy to provide adequate and effective credit service to the rural female-headed household.

The study found that, distance from the center of market affect income diversification participation decision of female-headed households negatively and significantly. This shows that market is rarely accessible to the farm household village and limit income generating capacity of them. Therefore, the intervention of governmental and non-governmental organizations in improving rural infrastructure particularly roads and other access like market center nearby either in form of establishing new or strengthening those already started ones to access the range of opportunities participation both on-farm and outside the farm to improve the livelihoods of rural female-headed households.

Gender-based discrimination was found to have negative and significant effect on both income diversification participation decision and intensity of participation of female-headed households. The result implies that, existence of stereotype among female and male regarding to participation in rural labor income generating activity from the point of view of social and cultural norm. Therefore, the government, office of women and child affairs with other stakeholder should be design appropriate policy that encourages female headed households to participate into rural off and non-farm activity with male counterpart by giving awareness about gender based discrimination in the study area.

Similarly, the result of the PSM indicated that income diversification has a significant and positive impact on annual average income, consumption expenditure, and saving of participant female-headed household which motivates the non- participant household to participate and earn more income, consumption expenditure and save more. Therefore, government and other stakeholders should create awareness, and give training to those households about the advantage of income source multiplication in case of rural economy.

5.4. Suggestion for Future Research

Although there are notable contributions from this study, there are certain limitations. The study incorporated different demographic characteristics and socio-economic factors that may affect female-headed household participation in income diversification. However, there may be

additional socio-economic and demographic factors that can affect female-headed household participation in income diversification. Again, the study considered economic variables indicators of livelihood like income, consumption expenditure, and saving. But livelihood incorporates additional economic variables. Therefore, future research should explore additional socio-economic and demographic factors that may affect female headed household participation in income diversification and additional livelihood indicator of economic variables. Moreover, due to regional disparities in terms of resource endowment, population density, and agro-climate condition the results obtained from the specific region or area cannot be generalized for the whole country, further studies can be conducted to study income diversification and its impact on livelihoods at the national level.

REFERENCES

Ababbo, Y. E. (2015). Determinants of Income Diversification among Rural Farm Households: The Case of Leemo District, Hadiya Zone, South Nation Nationalities People Region, Ethiopia. *International Journal of Science and Research (IJSR) ISSN*, *5*(11), 1162–1172. https://doi.org/10.21275/ART20163308.

Abate, D. (2019). *Review on Market Chain Analysis of Wheat in Ethiopia Review on Market Chain Analysis of Wheat in Ethiopia*. *December*.

Abbeam, G. G., Dagunga, G., & Ehiakpor, D. S. (2020). Rural non-farm income diversification: implications on smallholder farmers' welfare and agricultural technology adoption in Ghana. *Heliyon*, *6*(11), e05393. https://doi.org/10.1016/j.heliyon.2020.e05393

Abera, A., Yirgu, T., & Uncha, A. (2021). Determinants of rural livelihood diversification strategies among Chewaka resettlers ' communities of southwestern. *Agriculture & Food Security*, 1–19. https://doi.org/10.1186/s40066-021-00305.

Adem, M., Tadele, E., Mossie, H., Ayenalem, M., & Yildiz, F. (2018). Income diversification and food security situation in Ethiopia: A review study. *Cogent Food & Agriculture*, *4*(1), 1513354. https://doi.org/10.1080/23311932.2018.1513354.

Adem, M., & Tesafa, F. (2020). Intensity of income diversification among small- holder farmers in Asayita Woreda , Afar Region , Ethiopia Intensity of income diversification among small- holder farmers in Asayita Woreda , Afar Region ,. *Cogent Economics & Finance*, *00*(00). https://doi.org/10.1080/23322039.2020.1759394.

Alemayehu, T. (2020). *Determinants of Rural Households ' Participation in Microfinance Program : The Case of Omo Microfinance Institution , Sodo Woreda , Southern Nations Nationalities , and Peoples Regional*. *11*(1), 23–41.

Alobo Loison, S., & Bignebat, C. (2018). Patterns and Determinants of Household Income Diversification in Rural Senegal and Kenya. *Journal of Poverty Alleviation and International Development*, *8*(1).

Anang, B. T. (2017). *WIDER Working Paper 2017/38 Effects of non-farm work on agricultural productivity: Empirical evidence from northern Ghana*.

Astatike, A. A., & Gazuma, E. G. (2019). The Impact of Off-farm Activities on Rural Household Income in Wolaita Zone, Southern Ethiopia. *Journal of World Economic Research*, *8*(1), 8-16. https://doi.org/10.11648/j.jwer.20190801.12.

Babatunde, R. O., & Leliveld, A. (2013). *On-farm and off-farm works: Evidence from rural Nigeria*.

Baiphethi, M. N., Viljoen, M., & Kundhlande, G. (2008). Rural women and rainwater harvesting and conservation practices: Anecdotal evidence from the Free State and Eastern Cape provinces. *Agenda*, *22*(78), 163–171. http://www.tandfonline.com/doi/abs/10.1080/10130950.2008.9674994.

Balense, A., & Debebe, S. (2020). Analysis of Income Diversification and Livelihood Strategies Among Pastoral and Agro-pastoral Households ' in Southern Ethiopia. *Journal of Investment and Management*, *9*(3), 72–79. https://doi.org/10.11648/j.jim.20200903.12.

Barrett, C B, Reardon, T., & Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa : concepts , dynamics , and policy implications. *Food Policy*, *26*, 315–331.

Barrett, Christopher B, & Reardon, T. (2000). *Asset , Activity , and Income Diversification Among African Agriculturalists : Some Practical Issues Asset , Activity , and Income Diversification Among African Agriculturalists : Some Practical Issues*.

Batool, S., Babar, A., Nasir, F., & Iqbal, Z. S. (2017). *Income Diversification of Rural Households in Pakistan International Journal of Economics &*. *6*(6). https://doi.org/ 10.4172/2162-6359.1000466.

Bayu, E. K. (2018). *Gender and Livelihood Diversification in Rural Ethiopia Women ' s Participation in Non - farm Activities : The Case of Shebel Berenta Woreda , East Gojjam Zone , Amhara National Regional State*.

Becker, S. O., & Ichino, A. (2010). Estimation of average treatment effects based on parametric propensity score model. *And Inference*, *140*(3), 806–816. https://doi.org/10.1016/j.jspi. 2009.09.009.

Beyene, A. D. (2008). *Determinants of off-farm participation decision of farm households in Ethiopia*. *47*(1), 140–161.

Bigsten, A., & Tengstam, S. (2011). Smallholder diversification and income growth in Zambia. In *Journal of African Economies* (Vol. 20, Issue 5). https://doi.org/10.1093/jae/ejr017

Bojnec, Š., & Fert, I. (2011). *85th Annual Conference will be held in Warwick , from 18th – 20th Impact of Off-farm Income on Farm Efficiency in Slovenia*. *April*.

Bojnec, S., & Knific, K. (2021). Farm Household Income Diversification as a Survival Strategy. *Sustainability*, *13*, *6341*, 1–16. https://doi.org/https://doi.org/10.3390/su13116341.

Burke, W. J., Newton, H. J., Baum, C. F., Cameron, A. C., Epstein, D., Hardin, J., Jenkins, S., Kohler, U., & Lumley, T. (2016). Fitting and Interpreting Cragg's Tobit Alternative using Stata. *The Stata Journal*, *January*, 583–592. https://doi.org/10.1177/1536867X0900900405

Caliendo, M., & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *University of Cologne*, *1588*.

Croson, R., & Gneezy, U. (2009). *Gender Differences in Preferences*. 448–474.

Danso-Abbeam, G., Dagunga, G., & Ehiakpor, D. S. (2020). Rural non-farm income diversification: implications on smallholder farmers' welfare and agricultural technology adoption in Ghana. *Heliyon*, *6*(11), e05393. https://doi.org/10.1016/j.heliyon.2020.e05393.

Davis, B., Di Giuseppe, S., & Zezza, A. (2017). Are African households (not) leaving agriculture? Patterns of households' income sources in rural Sub-Saharan Africa. *Food Policy*, *67*, 153–174. https://doi.org/10.1016/j.foodpol.2016.09.018.

Davis, B., Winters, P., Reardon, T., & Stamoulis, K. (2009). Rural nonfarm employment and farming: Household-level linkages. *Agricultural Economics*, *40*(2), 119–123. https://doi.org/10.1111/j. 1574-0862.2009.00374.

Debesai, M. (2020). *Rural household income diversification in developing countries : A Case study of Eritrea*. *13014*.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. In *Review of Economics and Statistics*. https://doi.org/10.1162/003465302317331982.

Demissie, A. (2013). Determinants of income diversification among rural households: The case of smallholder farmers in Fedis district, Eastern hararghe zone, Ethiopia. *Journal of Development and Agricultural Economics*, *5*(3), 120–128. https://doi.org/10.5897/jdae12.104.

Derbe, T. (2020). Determinants of Smallholder Farmers' Livelihood Strategies in Wogera District of Ethiopia. *Erjssh*, *7*(1), 1–11.

Diep, H. T. N., & Vien, H. T. (2017). Determinants of income diversification and its effects on rural household income in Vietnam. *Can Tho University Journal of Science*, *06*(September), 153. https://doi.org/10.22144/ctu.jen.2017.039.

Diiro, G. M., & Sam, A. G. (2016). Agricultural technology adoption and nonfarm earnings in Uganda: a semiparametric analysis.Journal of developing areas, *49*.

Dornbusch, R., & Fischer, S. (1977). *Macro economics, sixth edition* (pp. 43–46).

Eckel, C. C., & Grossman, P. J. (2008). Chapter 113 Men, Women and Risk Aversion: Experimental Evidence. *Handbook of Experimental Economics Results*, *1*(C), 1061–1073. https://doi.org/10.1016/S1574-0722(07)00113-8.

Elin, V. (2015). *Income diversification among female-headed farming households*. 0–41.

Ellis, F. (1999). *Rural livelihood diversity in developing countries : evidence and policy implications*. *40*.

Etea, B. G., Zhou, D., Abebe, K. A., & Sedebo, D. A. (2019). Household income diversification and food security: Evidence from rural and semi-urban areas in Ethiopia. *Sustainability*, *11*(12). https://doi.org/10.3390/SU11123232.

FAO. (2020). Climate change and food security. *Agriculture and Forestry*, *66*(3), 197–210. https://doi.org/10.17707/AgricultForest.66.3.16.

Fassil, E., & Elias, M. (2016). Determinants of off farm income diversification and its effect on rural household poverty in Gamo Gofa Zone, Southern Ethiopia. *Journal of Development and Agricultural Economics*, *8*(10), 215–227. https://doi.org/10.5897/jdae2016-0736.

Fekadu, G., Tebarek, A., Megento, L., & Gurmessa, F. (2021). The extent of livelihood diversification on the determinants of livelihood diversification in Assosa Wereda , Western. *GeoJournal*, *5*(1992). https://doi.org/10.1007/s10708-021-10379-5.

Fentahun, G., & Kemaw, B. (2019). The Role of Income Diversification on Risk Management and Rural Household Food Security in Ethiopia The Role of Income Diversification on Risk Management and Rural Household Food Security in Ethiopia. *Journal of Economics and Sustainable Development*, *9*(August).

Gebiso, C. T., Ashebir, T. M., Aman, N. T., & Dawud, I. (2019). Rural Livelihood Diversification Status and Determinant Factors in Arsi, Ethiopia. *International Journal of Business and Economics Research*, *8*(1), 23. https://doi.org/10.11648/j.ijber.20190801.14.

Gebru, G. W., Ichoku, H. E., Ogbonnia, P., & Eze, P. (2018). Determinants of livelihood

diversification strategies in Eastern Tigray Region of Ethiopia. *Agriculture & Food Security*, 1–9. https://doi.org/10.1186/s40066-018-0214-0.

Ghosh, A. K., & Sujan, M. H. K. (2020). Changing pattern of rural income diversification and its impacts on household economy: A study on a selected village in Jashore District, Bangladesh. *Journal of Agriculture, Food and Environment*, *01*(04), 186–190. https://doi.org/10.47440/jafe.2020.1428.

Greene, W. H. (2002). *Econometric Analysis, New work University, Prentice-Hall International, Inc., 3rd ed., U.S.A.*

Greene, W. H. (2012). *Econometric Analysis, New work University, Prentice-Hall International, Inc., 4rd ed., U.S.A.*

Gujarati, (2004). *Basic Econometrics fourth Edition TATA McGraw –Hill, Companies*.

Hap, N., Seng, L., & Chuenpagdee, R. (2006). Livelihood Importance and Values of Tonle Sap Lake Fisheries. *Fisheries Bethesda, Phase 2*, 24.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*. https://doi.org/10.2307/2971733

Heinrich, C., Maffioli, A., & Vázquez, G. (2016). A Primer for Applying Propensity-Score Matching Impact-Evaluation Guidelines A Primer for Applying Propensity-Score Matching. *Inter-American Development Bank, September*.

Idris-Adeniyi, K. M., Busari, A. O., & Adedokun, S. A. (2020). Determinants of Income Diversification among Arable Crop Farmers in Osun State, Nigeria. *Journal of Agricultural Extension, Vol. 24 (2*, 23–30.

Israel, G. D. (1992). Sampling The Evidence Of Extension Program Impact. *Program Evaluation and Organizational Development, Gainesville, FL: University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS.*, 1–9.

James Tobin. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica*, *26*(1), 24–36.

Kassie, G. W., Kim, S., & Fellizar, F. P. (2017). Determinant factors of livelihood diversification: Evidence from Ethiopia. *Cogent Social Sciences*, *3*(1), 1–16. https://doi.org/ 10.1080/23311886.2017.1369490.

Khan, W., Jamshed, M., Fatima, S., & Dhamija, A. (2019). Forum for Social Economics Determinants of Income Diversification of Farm Households ' in Uttar Pradesh , India *Forum for Social Economics*, *0*(0), 1–19. https://doi.org/10.1080/07360932.2019.1666728.

Khandker, S., Gayatri, S., & Hussain, K. (2010). Handbook on Impact evaluation:Quantitative methods and practices. World Bank. In *Learning* (Vol. 1, Issue 1). https://doi.org/8028-4df/520990PUB0EPI1101Official0Use0Only1.

Kilic, T., Carletto, C., Miluka, J., & Savastano, S. (2009). Rural nonfarm income and its impact on agriculture: Evidence from Albania. *Agricultural Economics*, *40*(2), 139–160. https://doi.org/ 10.1111/j.1574-0862.2009.00366.

Le, L. H., & Le, T. N. (2020). Determinants of income diversification among rural households in

the mekong river delta: The economic transition period. *Journal of Asian Finance, Economics and Business*, *7*(5), 291–304. https://doi.org/10.13106/JAFEB.2020.VOL7.NO5.291.

Lidi, Bogale Yadeta, Amsalu Bedemo, M. B. (2017). Determinants of Saving Behavior of Farm Households in Rural Ethiopia : The Double Hurdle Approach. *Developing Country Studies*, *7*(12), 17–26.

Majbauddin, A., Otani, S., Tsunekawa, A., & Haregeweyn, N. (2020). The Influence of Income and Livelihood Diversification on Health-Related Quality of Life in Rural Ethiopia. *International Journal Environmental Research and Public Health*, 1–12.

Mammen, K., & Paxson, C. (2000). *Women's Work and Economic Development*. *14*(4), 141 164.

Mendershausen, H. (1939). The Relationship Between Income and Savings of American Metropolitan Families;The American Economic Review , Sep ., 1939 , Vol . 29 , No . 3 (Sep ., 1939), pp . 521. *American Economic Association*, *29*(3), 521–537.

Negese Tamirat Mulatu. (2017). *Household Livelihood : - A Case Study of Duna Woreda in MSc* program *Adoption and Impact of Row Planting of Wheat Crop on Household Livelihood : - A Case Study of Duna Woreda in Hadiya Zone , Ethiopia* .

Norsuzila Ya'acob1, Mardina Abdullah1, 2 and Mahamod Ismail1, 2, Medina, M., Talarico, T. L., Casas, I. A., Chung, T. C., Dobrogosz, W. J., Axelsson, L., Lindgren, S. E., Dobrogosz, W. J., Kerkeni, L., Ruano, P., Delgado, L. L., Picco, S., Villegas, L., Tonelli, F., Merlo, M., Rigau, J., Diaz, D., & Masuelli, M. (2020). Title Factors Affecting Yield of Crops. *Intech*, *32*, 137–144.

Ntwalle, J. A. (2019). *Determinants of Tanzania Rural Households' Income Diversification and its Impact on Food security*. http://stud.epsilon.slu.se

Omirin, O. J., & Okpara, C. (2018). *Off-farm activities as income strategies among rural women in ido local government area, Ibadan, Nigeria*. *21*, 224–239.

Omotesho, K. F., Akinrinde, A. F., Ogunlade, I., & Egbugo, J. K. (2020). Effect of income diversification on the livelihood of rural farming households in Kwara State, Nigeria. *Kasetsart Journal of Social Sciences*, *41*(2), 308–314. https://doi.org/ 10.34044/j.kjss.2020.41.2.07.

Onadeko, S., Nelson, I., Shotuyo, A., & Ityavyar, J. (2020). Determinants of Income Diversification among Support Zone Communities of Nigeria National Parks. *Economic and Environmental Studie*, *20*(1), 9–23.

Palacios-Lopez, A., Christiaensen, L., & Kilic, T. (2017). How much of the labor in African agriculture is provided by women? *Food Policy*, *67*(June), 52–63. https://doi.org/ 10.1016/j.foodpol. 2016.09.017.

Reardon, T. (1997). Using Evidence of Household Income Diversification to Inform Study of the Rural Nonfarm Labor Market in Africa. *World Development*, *25*(5), 735–747. https://doi.org/10.1016/S0305-750X(96)00137-4.

Rosenbaum, & Rubin. (1985). *The Central Role of the Propensity Score in Observational Studies for Causal Effects*. *70*(1), 41–55.

Salam, S., Bauer, S., & Palash, M. S. (2019). Impact of income diversification on rural livelihood in some selected areas of Bangladesh. *Journal of the Bangladesh Agricultural University*, *17*(1), 73–79. https://doi.org/10.3329/jbau.v17i1.40666.

Santos, R. (2013). Keynes ' Absolute Income Hypothesis and Kuznets Paradox, Department of Economics, Adekunle Ajasin University Akunbga Akoko, Ondo State Niger. *Munich Personal RePEc Archive*, *49310*.

Shannon, J. (2019). Gender Differences or Gendered Differences: Understanding the Power of Language in Training and Research in Supervision. *International Journal for the Advancement of Counselling*, *41*(4), 598–608. https://doi.org/10.1007/s10447-019-09380-y

Sianesi, B. (2001). Implementing Propensity Score Matching Estimators with STATA. *Fiscal Studies*, *May*, 1–15.

Takala, K. (1995). Permanent Income Hypothesis and Saving in Finland. In *Bank of Finland*.

Teji, M. H. (2020). *Determinants of Income Diversification among Rural Farm Household in Soddo District , Gurage Zone of Southern Ethiopia*. *10*(10), 7–18. https://doi.org/ 10.7176/DCS/10-10-02.

Teshome, B., & Edriss, A.-K. (2013). Determinants and Patterns of Income Diversification among Smallholder Farmers in Akaki District, Ethiopia 1. *Journal of Research in Economics and International Finance (JREIF*, *2*(4), 68–78.

Udry, C. (1996). Gender, agricultural production, and the theory of the household. *Journal of Political Economy*, *104*(5), 1010–1046. https://doi.org/10.1086/262050.

Vien, H. T. (2017a). *Determinants of income diversification and its effects on rural household income in Vietnam*. *September*. https://doi.org/10.22144/ctu.jen.2017.039.

Vien, H. T. (2017b). *Determinants of income diversification and its effects on rural household income in Vietnam*. *January*. https://doi.org/10.22144/ctu.jen.2017.039.

WB. (2020). Ethiopia gender diagnostic report.

Welteji, D. (2018). A critical review of rural development policy of Ethiopia : access , utilization and coverage. *Agriculture & Food Security*, 1–6. https://doi.org/10.1186/s40066-018-0208-

Wiggins, S. (2009). Transforming the Rural Nonfarm Economy. Opportunities and Threats in the Developing World - Edited by Steven Haggblade, Peter B.R. Hazell and Thomas Reardon. *Journal of Agrarian Change*, *9*(4), 595–598. https://doi.org/10.1111/j.1471-0366.2009.00239.

Woldenhanna, T., & Oskam, A. (2001). Income diversification and entry barriers: Evidence from the Tigray region of Northern Ethiopia. *Food Policy*, *26*(4), 351–365. https://doi.org/ 10.1016/S0306-9192(01)00009-4.

Xu, T. (2017). Income Diversification and Rural Consumption — Evidence from Chinese Provincial Panel Data. *Sustainability*. https://doi.org/10.3390/su9061014.

Yizengaw, B. (2014). *Determinants of Household Income Diversification and Its Effect on Food Security Status in Rural Ethiopia: Evidence from Ethiopia Longitudinal Rural Household Survey*.

APPENDIX

Appendix Table 1: *Conversion factor of tropical livestock unit (TLU)*

Livestock Category	TLU
Ox and Cow	1
Heifer	0.75
Young Bull (jibicha)	0.34
Calf	0.25
Sheep and Goat (Young)	0.06
Sheep and Goat (Adult)	0.13
Hen	0.013
Donkey (Young)	0.35
Donkey(Adult)	0.70
Horse and Mule	1.1

Source: Strock *et al (1991)*

Appendix Table 2: *Conversion factor used to compute household size in adult equivalent (AE)*

Age group	Male	Female
< 10	0.6	0.6
10 - 13	0.9	0.8
14 - 16	1	0.75
17 - 50	1	0.75
> 50	1	0.75

Source: strock *et al (1991)*

Appendix Table 3: Over all test of the model

Multicollinarity problem test for continouos expalnatory variable.

```
. vif

    Variable |       VIF       1/VIF
-------------+----------------------
    AHINCOME |      1.35    0.738914
         EDL |      1.28    0.783241
        DMRK |      1.23    0.810112
     LOWNED |      1.22    0.818747
       LSIZE |      1.18    0.848075
       FSIZE |      1.11    0.900065
         AGE |      1.11    0.903546
  FECONTACT |      1.09    0.917660
       LnAAI |      1.08    0.922798
      DRATIO |      1.08    0.926560
      SAVIND |      1.07    0.935558
-------------+----------------------
    Mean VIF |      1.16
```

Multicollinarity problem test for discrite expalnatory variable.

Variable	No observation	χ^2	Contingency coefficient
AILAND	245	9.3938	0.0369
ACREDIT	245	5.9121	0.0235
MFBO	245	11.9300	0.0464
PRSHOCK	245	21.0627	0.0761
GENDIS	245	52.5524	0.1766

```
. hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of PHID

         chi2(1)      =     0.33
         Prob > chi2  =   0.5665
. hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of SID

         chi2(1)      =    12.54
         Prob > chi2  =   0.0004

. ovtest

Ramsey RESET test using powers of the fitted values of PHID
       Ho:  model has no omitted variables
                F(3, 226) =      8.95
                 Prob > F =      0.0000
```

Appendix Table 4: Heckman sample selection model output for income diversification and its intensity.

Heckman selection model -- two-step estimates (regression model with sample selection)

Number of obs = 245
Selected = 108
Nonselected = 137

Wald chi2(14) = 23.62
Prob > chi2 = 0.0508

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
SID						
AGE	-.0039765	.003374	-1.18	0.239	-.0105895	.0026365
EDL	.0052474	.0093111	0.56	0.573	-.0130019	.0234968
FSIZE	-.0148693	.0101574	-1.46	0.143	-.0347774	.0050387
DRATIO	.0953572	.1240731	0.77	0.442	-.1478215	.3385359
LSIZE	.0284097	.0149164	1.90	0.057	-.0008259	.0576454
LOWNED	.0005799	.0035886	0.16	0.872	-.0064535	.0076134
FECONTACT	.0344454	.0168943	2.04	0.041	.0013331	.0675577
MFBO	.0178907	.0363588	0.49	0.623	-.0533712	.0891525
ACREDIT	.0173556	.0360462	0.48	0.630	-.0532937	.0880048
GENDISC	-.114482	.0638838	-1.79	0.073	-.239692	.010728
PRSH	-.0090397	.0397307	-0.23	0.820	-.0869105	.0688311
AILAND	-.0114608	.0352659	-0.32	0.745	-.0805808	.0576592
DMRK	.0115877	.0100468	1.15	0.249	-.0081037	.0312791
LnAAI	.007168	.0131705	0.54	0.586	-.0186457	.0329817
_cons	.5324332	.1816047	2.93	0.003	.1764945	.8883718
PHID						
AGE	-.0666405	.0184425	-3.61	0.000	-.1027872	-.0304938
EDL	.1933173	.0391022	4.94	0.000	.1166783	.2699562
FSIZE	.1747055	.0556144	3.14	0.002	.0657032	.2837078
DRATIO	.3394871	.8228117	0.41	0.680	-1.273194	1.952168
LSIZE	-.0405711	.112458	-0.36	0.718	-.2609847	.1798424
LOWNED	.0501399	.0227278	2.21	0.027	.0055942	.0946856
FECONTACT	.0210757	.1172356	0.18	0.857	-.2087017	.2508532
MFBO	.0643465	.2344694	0.27	0.784	-.395205	.523898
ACREDIT	.4011732	.2308084	1.74	0.082	-.051203	.8535494
GENDISC	-.7112687	.2351931	-3.02	0.002	-1.172239	-.2502987
PRSH	.274013	.2558183	1.07	0.284	-.2273817	.7754076
AILAND	.1778446	.246582	0.72	0.471	-.3054471	.6611364
DMRK	-.1800057	.0405055	-4.44	0.000	-.2593951	-.1006164
LnAAI	-.1259735	.0793054	-1.59	0.112	-.2814093	.0294623
_cons	2.803885	1.258789	2.23	0.026	.3367043	5.271066
/mills						
lambda	-.0246588	.1061197	-0.23	0.816	-.2326496	.1833319
rho	-0.14907					
sigma	.16541918					

Appendix Table 4: Tobit model for income diversification participation decision and its intensity

```
Tobit regression                                Number of obs     =        245
                                                  Uncensored      =        108
Limits: lower = 0                                 Left-censored   =        137
        upper = +inf                              Right-censored  =          0

                                                LR chi2(14)       =     180.77
                                                Prob > chi2       =     0.0000
Log likelihood = -96.061779                     Pseudo R2         =     0.4848
```

SID	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
AGE	-.0165788	.0041462	-4.00	0.000	-.024748	-.0084097
EDL	.0495911	.0089412	5.55	0.000	.0319744	.0672078
FSIZE	.0313385	.012745	2.46	0.015	.0062272	.0564498
DRATIO	.1895944	.2040768	0.93	0.354	-.2124953	.5916842
LSIZE	.0127626	.0270745	0.47	0.638	-.040582	.0661072
LOWNED	.0090673	.0054232	1.67	0.096	-.0016179	.0197525
FECONTACT	.036407	.0277755	1.31	0.191	-.0183188	.0911327
MFBO	.024291	.0593237	0.41	0.683	-.0925937	.1411756
ACREDIT	.0944971	.0561135	1.68	0.094	-.0160626	.2050568
GENDISC	-.2810424	.0670501	-4.19	0.000	-.4131502	-.1489345
PRSH	.0787279	.0602884	1.31	0.193	-.0400574	.1975133
AILAND	.0429579	.0604937	0.71	0.478	-.0762321	.1621479
DMRK	-.0414914	.0102289	-4.06	0.000	-.0616452	-.0213375
LnAAI	-.0209801	.019433	-1.08	0.281	-.0592686	.0173084
_cons	.6656103	.298964	2.23	0.027	.0765655	1.254655
var(e.SID)	.1226218	.0184701			.0911337	.1649896

Appendix Table 5: Double-hurdle model output for income diversification and its intensity.

```
                                                Number of obs     =        245
                                                Wald chi2(14)     =     130.65
Log pseudolikelihood = -36.007541               Prob > chi2       =     0.0000
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Tier1						
AGE	-.0666405	.0162069	-4.11	0.000	-.0984055	-.0348755
EDL	.1933173	.0373476	5.18	0.000	.1201173	.2665173
FSIZE	.1747055	.0504816	3.46	0.001	.0757634	.2736476
DRATIO	.3394871	.8224121	0.41	0.680	-1.272411	1.951385
LSIZE	-.0405711	.1242216	-0.33	0.744	-.284041	.2028987
LOWNED	.0501399	.0209196	2.40	0.017	.0091382	.0911416
FECONTACT	.0210757	.112018	0.19	0.851	-.1984756	.2406271
MFBO	.0643465	.2402942	0.27	0.789	-.4066216	.5353145
ACREDIT	.4011732	.2258164	1.78	0.076	-.0414187	.8437652
GENDISC	-.7112687	.2341461	-3.04	0.002	-1.170186	-.2523508
PRSH	.274013	.2345968	1.17	0.243	-.1857883	.7338143
AILAND	.1778446	.2450437	0.73	0.468	-.3024322	.6581215
DMRK	-.1800057	.0377932	-4.76	0.000	-.2540791	-.1059323
LnAAI	-.1259735	.0586057	-2.15	0.032	-.2408386	-.0111084
_cons	2.803885	1.131306	2.48	0.013	.5865662	5.021204
Tier2						
AGE	-.0043064	.0023956	-1.80	0.072	-.0090017	.0003889
EDL	.007148	.0053221	1.34	0.179	-.0032832	.0175792
FSIZE	-.0126545	.0071539	-1.77	0.077	-.026676	.001367
DRATIO	.0856053	.1250178	0.68	0.494	-.1594251	.3306358
LSIZE	.0286801	.0140315	2.04	0.041	.0011789	.0561813
LOWNED	.0001524	.0031408	0.05	0.961	-.0060035	.0063083
FECONTACT	.0341251	.0147528	2.31	0.021	.0052101	.0630401
MFBO	.0216135	.0330701	0.65	0.513	-.0432028	.0864298
ACREDIT	.0148424	.0306027	0.49	0.628	-.0451378	.0748225
GENDISC	-.1297951	.045926	-2.83	0.005	-.2198083	-.0397818
PRSH	.0029711	.0354324	0.08	0.933	-.0664752	.0724175
AILAND	-.0070482	.033184	-0.21	0.832	-.0720877	.0579913
DMRK	.0094182	.0066136	1.42	0.154	-.0035442	.0223805
LnAAI	.0062774	.0085201	0.74	0.461	-.0104218	.0229765
_cons	.5311942	.1628985	3.26	0.001	.2119189	.8504695
sigma						
_cons	.1585477	.0117808	13.46	0.000	.1354577	.1816376

Probit regression for marginal effect

```
Probit regression                               Number of obs     =        245
                                                Wald chi2(14)     =     130.65
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -82.529353               Pseudo R2         =     0.5090
```

PHID	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
AGE	-.0666405	.0162069	-4.11	0.000	-.0984055	-.0348755
EDL	.1933173	.0373476	5.18	0.000	.1201173	.2665172
FSIZE	.1747055	.0504816	3.46	0.001	.0757634	.2736476
DRATIO	.3394871	.822412	0.41	0.680	-1.272411	1.951385
LSIZE	-.0405711	.1242216	-0.33	0.744	-.284041	.2028987
LOWNED	.0501399	.0209196	2.40	0.017	.0091382	.0911416
FECONTACT	.0210757	.112018	0.19	0.851	-.1984756	.2406271
MFBO	.0643465	.2402942	0.27	0.789	-.4066216	.5353145
ACREDIT	.4011732	.2258163	1.78	0.076	-.0414187	.8437651
GENDISC	-.7112686	.234146	-3.04	0.002	-1.170186	-.2523508
PRSH	.2740129	.2345968	1.17	0.243	-.1857883	.7338142
AILAND	.1778446	.2450437	0.73	0.468	-.3024322	.6581214
DMRK	-.1800057	.0377932	-4.76	0.000	-.2540791	-.1059323
LnAAI	-.1259735	.0586057	-2.15	0.032	-.2408386	-.0111084
_cons	2.803885	1.131306	2.48	0.013	.5865661	5.021204

Appendix Table 6: Summary of propensity score

```
. sum  _pscore
```

Variable	Obs	Mean	Std. Dev.	Min	Max
_pscore	245	.4412706	.2276207	.0060424	.9853427

```
. sum  _pscore if   PHID ==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
_pscore	108	.5588739	.2201783	.1154501	.9853427

```
. sum  _pscore if   PHID ==0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
_pscore	137	.3485615	.1875445	.0060424	.8818751

```
. sum _pscore,detail

                    psmatch2: Propensity Score
------------------------------------------------------------------------------
      Percentiles      Smallest
 1%     .0403345       .0060424
 5%     .1154501       .0213269
10%     .1594224       .0403345       Obs                 245
25%      .268058       .0427847       Sum of Wgt.         245

50%     .4112833                      Mean           .4412706
                       Largest        Std. Dev.      .2276207
75%     .6205081        .94455
90%     .7547525       .9577564       Variance       .0518112
95%     .8556765       .9751476       Skewness         .34432
99%     .9577564       .9853427       Kurtosis       2.305233
```

Result of ATT propensity score using propensity score matching

```
. psmatch2  (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI ), outcome( lnAHINCOME
> lnCONEX lnSAVING ) neighbor(1) common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1601631	.326056163	.065411328	4.98
	ATU	10.1055842	10.4626106	.357026352	.	.
	ATE			.343338634	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0975727	.272014156	.069958626	3.89
	ATU	10.0338666	10.3210409	.287174301	.	.
	ATE			.280474058	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	5.76662936	1.88939592	.613170109	3.08
	ATU	6.06430652	7.50185493	1.43754842	.	.
	ATE			1.63724887	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		
	Off suppo	On suppor	Total
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2  (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI ), outcome(lnAHINCOME l
> nCONEX lnSAVING) neighbor(2) common  logit  ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1678477	.318371556	.060837075	5.23
	ATU	10.1055842	10.466075	.360490715	.	.
	ATE			.341875551	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.1017515	.267835381	.065349515	4.10
	ATU	10.0338666	10.3318627	.297996082	.	.
	ATE			.28466613	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.01792026	1.63810502	.513413287	3.19
	ATU	6.06430652	7.50129892	1.4369924	.	.
	ATE			1.52587699	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2  (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI ), outcome(lnAHINCOME l
> nCONEX lnSAVING) neighbor(3) common  logit  ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1383049	.347914404	.058763662	5.92
	ATU	10.1055842	10.4728302	.367245969	.	.
	ATE			.358702108	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0655885	.303998321	.063611786	4.78
	ATU	10.0338666	10.3361841	.302317561	.	.
	ATE			.303060397	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.2235518	1.43247348	.482104553	2.97
	ATU	6.06430652	7.57646533	1.51215881	.	.
	ATE			1.47694074	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2  (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), outcome( lnAHINCOME l
> nCONEX lnSAVING) neighbor(4) common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1394086	.346810724	.057284612	6.05
	ATU	10.1055842	10.4840526	.378468327	.	.
	ATE			.364476797	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0637871	.305799766	.062244052	4.91
	ATU	10.0338666	10.3426557	.308789095	.	.
	ATE			.307467918	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.28110837	1.37491691	.450694426	3.05
	ATU	6.06430652	7.6355756	1.57126908	.	.
	ATE			1.48448843	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2  (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), outcome( lnAHINCOME l
> nCONEX lnSAVING ) neighbor(5) common  logit  ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1367831	.349436155	.056432021	6.19
	ATU	10.1055842	10.4784308	.372846545	.	.
	ATE			.362499989	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0625425	.307044328	.061439526	5.00
	ATU	10.0338666	10.3416389	.307772281	.	.
	ATE			.307450552	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.15147938	1.5045459	.441287554	3.41
	ATU	6.06430652	7.64857424	1.58426772	.	.
	ATE			1.54903353	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), caliper(0.01) outcome(
> lnAHINCOME lnCONEX lnSAVING) common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.5053453	10.1251437	.380201581	.062489459	6.08
	ATU	10.095421	10.4582146	.362793613	.	.
	ATE			.370127934	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3875899	10.0582927	.329297227	.068225145	4.83
	ATU	10.0229021	10.3101127	.287210615	.	.
	ATE			.304942538	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.71244287	5.90885573	1.80358714	.575396735	3.13
	ATU	6.00101224	7.51170976	1.51069752	.	.
	ATE			1.63409772	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	23	114	137
Treated	25	83	108
Total	48	197	245

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), caliper(0.1) outcome (
> lnAHINCOME lnCONEX lnSAVING) common  logit  ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1601631	.326056163	.065411328	4.98
	ATU	10.1055842	10.4626106	.357026352	.	.
	ATE			.343338634	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0975727	.272014156	.069958626	3.89
	ATU	10.0338666	10.3210409	.287174301	.	.
	ATE			.280474058	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	5.76662936	1.88939592	.613170109	3.08
	ATU	6.06430652	7.50185493	1.43754842	.	.
	ATE			1.63724887	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), caliper(0.25) outcome(
> lnAHINCOME lnCONEX lnSAVING) common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1601631	.326056163	.065411328	4.98
	ATU	10.1055842	10.4626106	.357026352	.	.
	ATE			.343338634	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0975727	.272014156	.069958626	3.89
	ATU	10.0338666	10.3210409	.287174301	.	.
	ATE			.280474058	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	5.76662936	1.88939592	.613170109	3.08
	ATU	6.06430652	7.50185493	1.43754842	.	.
	ATE			1.63724887	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), caliper(0.5) outcome(l
> nAHINCOME lnCONEX lnSAVING) common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1601631	.326056163	.065411328	4.98
	ATU	10.1055842	10.4626106	.357026352	.	.
	ATE			.343338634	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0975727	.272014156	.069958626	3.89
	ATU	10.0338666	10.3210409	.287174301	.	.
	ATE			.280474058	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	5.76662936	1.88939592	.613170109	3.08
	ATU	6.06430652	7.50185493	1.43754842	.	.
	ATE			1.63724887	.	.

Note: S.E. does not take into account that the propensity score is estimated.

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), radius bw(0.01) outcom
> e(lnAHINCOME lnCONEX lnSAVING) common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1055842	.380635056	.047498282	8.01
	ATU	10.1055842	10.4862193	.380635056	.	.
	ATE			.380635056	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0338666	.335720288	.052140627	6.44
	ATU	10.0338666	10.3695869	.335720288	.	.
	ATE			.335720288	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.06430652	1.59171876	.35589459	4.47
	ATU	6.06430652	7.65602528	1.59171876	.	.
	ATE			1.59171876	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), radius bw(0.1) outcome
> (lnAHINCOME lnCONEX lnSAVING) common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1055842	.380635056	.047498282	8.01
	ATU	10.1055842	10.4862193	.380635056	.	.
	ATE			.380635056	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0338666	.335720288	.052140627	6.44
	ATU	10.0338666	10.3695869	.335720288	.	.
	ATE			.335720288	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.06430652	1.59171876	.35589459	4.47
	ATU	6.06430652	7.65602528	1.59171876	.	.
	ATE			1.59171876	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), radius bw(0.25) outcom
> e(lnAHINCOME lnCONEX lnSAVING) common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1055842	.380635056	.047498282	8.01
	ATU	10.1055842	10.4862193	.380635056	.	.
	ATE			.380635056	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0338666	.335720288	.052140627	6.44
	ATU	10.0338666	10.3695869	.335720288	.	.
	ATE			.335720288	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.06430652	1.59171876	.35589459	4.47
	ATU	6.06430652	7.65602528	1.59171876	.	.
	ATE			1.59171876	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), radius bw(0.5) outcome
> (lnAHINCOME lnCONEX lnSAVING) common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1055842	.380635056	.047498282	8.01
	ATU	10.1055842	10.4862193	.380635056	.	.
	ATE			.380635056	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0338666	.335720288	.052140627	6.44
	ATU	10.0338666	10.3695869	.335720288	.	.
	ATE			.335720288	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.06430652	1.59171876	.35589459	4.47
	ATU	6.06430652	7.65602528	1.59171876	.	.
	ATE			1.59171876	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), kernel outcome(lnAHINC
> OME lnCONEX lnSAVING)bwidth(0.01)common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.5053453	10.117641	.387704269	.060122337	6.45
	ATU	10.095421	10.4664851	.37106411	.	.
	ATE			.378074938	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3875899	10.0451301	.342459812	.066179651	5.17
	ATU	10.0229021	10.3263153	.303413276	.	.
	ATE			.319864354	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.71244287	6.05725933	1.65518354	.493817789	3.35
	ATU	6.00101224	7.60899804	1.6079858	.	.
	ATE			1.62787114	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	23	114	137
Treated	25	83	108
Total	48	197	245

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1357787	.350440626	.05321502	6.59
	ATU	10.1055842	10.4917859	.386201691	.	.
	ATE			.370396578	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0626448	.306942091	.057953845	5.30
	ATU	10.0338666	10.3648293	.330962693	.	.
	ATE			.320346445	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.02482159	1.63120369	.422262772	3.86
	ATU	6.06430652	7.66921785	1.60491133	.	.
	ATE			1.61653162	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. pstest   _pscore AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI ,both sum
```

Variable	Unmatched / Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
_pscore	U	.55927	.34743	103.3		8.10	0.000	1.39
	M	.52595	.51951	3.1	97.0	0.23	0.820	1.02
AGE	U	41.731	43.898	-31.0		-2.40	0.017	0.89
	M	42.303	41.851	6.5	79.2	0.50	0.618	1.04
EDL	U	3.8241	2.2774	51.1		4.02	0.000	1.51*
	M	3.4545	3.5058	-1.7	96.7	-0.12	0.906	1.16
FSIZE	U	6.6111	5.8905	33.9		2.65	0.009	1.25
	M	6.4545	6.2363	10.3	69.7	0.72	0.471	1.23
DRATIO	U	.11491	.09934	11.5		0.89	0.374	0.86
	M	.11273	.10347	6.8	40.5	0.47	0.640	0.78
LSIZE	U	2.6204	2.6015	2.0		0.16	0.876	1.36
	M	2.6187	2.6258	-0.8	62.2	-0.05	0.960	1.11
LOWNED	U	10.505	10.991	-8.8		-0.68	0.496	0.99
	M	10.529	11.041	-9.2	-5.3	-0.64	0.522	0.92
FECONTACT	U	1.1204	.94891	17.0		1.32	0.188	0.91
	M	1.0606	.92524	13.4	21.1	0.96	0.336	0.96
MFBO	U	.55556	.44526	22.1		1.72	0.087	.
	M	.53535	.59504	-12.0	45.9	-0.84	0.400	.
ACREDIT	U	.40741	.39416	2.7		0.21	0.834	.
	M	.41414	.43865	-5.0	-85.0	-0.35	0.729	.
GENDISC	U	.37037	.45985	-18.2		-1.41	0.160	.
	M	.39394	.4211	-5.5	69.6	-0.39	0.699	.
PRSH	U	.36111	.25547	22.9		1.79	0.074	.
	M	.34343	.30053	9.3	59.4	0.64	0.521	.
AILAND	U	.34259	.32117	4.5		0.35	0.725	.
	M	.34343	.36564	-4.7	-3.6	-0.32	0.746	.
DMRK	U	5.5139	7.0657	-48.9		-3.75	0.000	0.66*
	M	5.6717	5.7263	-1.7	96.5	-0.13	0.898	0.83
LnAAI	U	9.8368	9.7887	3.3		0.25	0.803	0.38*
	M	9.8241	9.9606	-9.4	-184.0	-0.76	0.450	0.62*

* if variance ratio outside [0.68; 1.46] for U and [0.67; 1.49] for M

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.168	56.64	0.000	25.4	18.2	102.0*	1.32	30
Matched	0.017	4.61	0.995	6.6	6.5	30.6*	1.28	10

* if B>25%, R outside [0.5; 2]

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), kernel outcome(lnAHINC
> OME lnCONEX lnSAVING)bwidth(0.25)common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1305505	.355668817	.048969386	7.26
	ATU	10.1055842	10.4975407	.391956493	.	.
	ATE			.375918636	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0573447	.312242137	.053632042	5.82
	ATU	10.0338666	10.3706464	.336779782	.	.
	ATE			.325935019	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.03524266	1.62078261	.373350553	4.34
	ATU	6.06430652	7.68115194	1.61684542	.	.
	ATE			1.61858552	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

```
. psmatch2 (PHID AGE EDL FSIZE DRATIO LSIZE LOWNED FECONTACT MFBO ACREDIT GENDIS PRSH AILAND DMRK LnAAI), kernel outcome(lnAHINC
> OME lnCONEX lnSAVING)bwidth(0.5)common  logit ate
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnAHINCOME	Unmatched	10.4923836	10.1131397	.37924391	.041098247	9.23
	ATT	10.4862193	10.1096416	.376577679	.045485856	8.28
	ATU	10.1055842	10.4987162	.393131928	.	.
	ATE			.385815541	.	.
lnCONEX	Unmatched	10.3710509	10.0419901	.329060773	.044981466	7.32
	ATT	10.3695869	10.0374098	.332177042	.050106277	6.63
	ATU	10.0338666	10.375468	.341601414	.	.
	ATE			.337436178	.	.
lnSAVING	Unmatched	7.71671515	6.07953731	1.63717784	.3218555	5.09
	ATT	7.65602528	6.0265878	1.62943748	.331475456	4.92
	ATU	6.06430652	7.68918902	1.6248825	.	.
	ATE			1.62689564	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support		Total
	Off suppo	On suppor	
Untreated	12	125	137
Treated	9	99	108
Total	21	224	245

Appendix Table 7: Result of Sensitivity analysis

```
. rbounds  lnAHINCOME , gamma(1(0.25)3)

Rosenbaum bounds for lnAHINCOME (N = 245 matched pairs)

Gamma            sig+      sig-     t-hat+     t-hat-        CI+        CI-
-------------------------------------------------------------------------
    1               0         0     10.309     10.309    10.2532    10.3515
 1.25               0         0    10.2697    10.3412    10.2178     10.386
  1.5               0         0    10.2374    10.3691    10.1847    10.4183
 1.75               0         0    10.2124    10.3927    10.1576    10.4401
    2               0         0    10.1854    10.4138    10.1298     10.459
 2.25               0         0    10.1688    10.4302    10.1062    10.4764
  2.5               0         0    10.1462    10.4427    10.0858    10.4909
 2.75         1.1e-16         0    10.1306     10.459    10.0636    10.5057
    3         2.3e-15         0    10.1087    10.4692    10.0535    10.5172

* 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)

. rbounds _lnAHINCOME , gamma(1(0.25)3)

Rosenbaum bounds for _lnAHINCOME (N = 224 matched pairs)

Gamma            sig+      sig-     t-hat+     t-hat-        CI+        CI-
-------------------------------------------------------------------------
    1               0         0    10.3249    10.3249    10.3011    10.3481
 1.25               0         0    10.3068    10.3417    10.2859    10.3751
  1.5               0         0    10.2936    10.3566     10.277    10.4115
 1.75               0         0    10.2847    10.3802    10.2704    10.4393
    2               0         0    10.2786    10.4046    10.2647    10.4482
 2.25               0         0    10.2734    10.4315    10.2589     10.455
  2.5         1.1e-16         0    10.2695     10.441    10.2523    10.4625
 2.75         2.6e-15         0    10.2656    10.4467    10.2414    10.4695
    3         3.4e-14         0    10.2619    10.4519     10.222    10.4758

* 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)
```

```
. rbounds _lnSAVING , gamma(1(0.25)3)

Rosenbaum bounds for _lnSAVING (N = 224 matched pairs)

Gamma            sig+      sig-     t-hat+     t-hat-        CI+        CI-
-------------------------------------------------------------------------
    1               0         0    6.94205    6.94205     6.8475    7.05866
 1.25               0         0    6.87463     7.0218    6.77641    7.14937
  1.5               0         0    6.81449    7.09967    6.71611    7.24784
 1.75               0         0    6.77014     7.1577      6.626    7.41683
    2               0         0    6.72695    7.22201    6.54478    7.47398
 2.25               0         0    6.67592    7.36413    6.48908    7.52521
  2.5         1.1e-16         0    6.60977    7.43087    6.43754     7.5798
 2.75         2.6e-15         0    6.55353    7.46299    6.38669    7.62416
    3         3.4e-14         0    6.51781    7.49848    6.33671    7.65084

* 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)
```

APPENDIX

JIMMA UNIVERSITY

COLLEGE OF BUSINESS AND ECONOMICS

DEPARTMENT OF ECONOMICS

Dear respondent; this questionnaire is prepared with the intention to gather information on the determinant of female headed household participation decision in income diversification and its impact on their livelihood: A case study of Guduru district. The information you provide will be valuable for the success of the research project. Please be honest and objective while filling the questionnaire. Your genuine response to the following questions would have a crucial importance to the results of the study. The information you provide is only used for academic consumption and will kept confidential. Therefore, you are kindly requested to give accurate information. Thank you for your cooperation!

Instructions to Enumerators

1. Make clear introduction to the respondent before starting the interview (greet them, tell your name, get her/his name, and make clear the purpose and objective of the study that you are undertaking).
2. Please ask the question clearly and patiently until the respondent understands.
3. During the process put the answers of each respondent both in the space provided and encircle the choice or tick mark as requiring

Instruction for respondent: Circle, tick your answer among the given alternative accordingly and fill your own idea in the blank space.

PART I: GENERAL INFORMATION

Name of Interviewer: _____ Signature: _____

Type of respondent: _____ kebeles_____

Part II. Information on Demographic Characteristics and Socio-Economic Variables

1. Age: _____
2. Family size of the household head:

Male _____ Female _____ Total _____

3. Please indicate the number of economically inactive from your family member?

Age	below 5	6-10	11-14	65-70	71-75	Above 75	Total
Number of dependents							

4. Educational background: schooling level of female-headed household

1) Grade _____ 2) Other/specify _____

5. What is the distance of your dwelling in hour from the market center? _____.

5.1. What type of transportation did you use from your home to market center?

1) Head/back load [ ] 2) Using donkey [ ] 3) Using Trucks/Vehicle [ ] 4) Using Carts [ ] 5) Using horse [ ] 6) Others_____

5.2. How you describe the road network in your locality?

1) Very Good [ ] 2) Good [ ] 3) Poor [ ]

5.3. Do you feel there is a problem in transportation of your products to the nearest market?

1) Yes [ ] 0) No [ ]

5.4. If yes, what do you think is the problem?

1) Long distance to the nearby market [ ] 2) No favorable road [ ]

3) No transportation services [ ]

6. Do you have a plot of land on which you cultivate crops? 1. Yes [ ] 2. No [ ]

6.1. If your response to question **No**6 is 'yes' how much hectares of farm land does your household own now? _____

6.2. If 'No' to question 6 how do you get plot of land to cultivate?

1) Share cropping [ ] 2) Rental [ ] 3) Gift from parents [ ] 4) Others (Specify) _____

7. Do you own livestock? 1) Yes [ ] 0) No [ ]

7.1. If your response is 'yes" to question **No7**, indicate the number of livestock owned currently.

Types of livestock		Total owned
Cattles	Oxen	
	Cow	
	Total	
Sheep and goats	Goat	
	Sheep	
	Chicken	
	Total	
Marines	Mules	
	Horse	
	Donkey	
	Total	

8. Did you get any credit from credit lending institution? 1) Yes [ ] 0) No [ ]

8.1 If your response to question **No 8** above is 'Yes', for what purpose(s)?

1) Purchase of farm inputs [ ] 2) Hiring labor [ ] 3) Household expenses in food, clothing and other supplies [ ] 4) For loan repayment [ ] 5) Purchase of jewelry [ ] 6) To start nonfarm business 7) Payment for hired labour [ ] 8) Others specify_____

8.2 If your response to question **No 8** is 'No', what was the reason for not taking loan from formal credit institution?

1) The institution is not available [ ] 2) The institution is far from the kebeles [ ] 3) It requires asset for collateral [ ] 4) The interest rate is too high [ ] 5) The loan payment time is not appropriate [ ] 6) Not allowed by religion [ ] 7) I use my own cash [ ] 8) I do not want [ ]

9) Other (Specify)_____

8.3 Indicate sources of credit

1) Oromia credit and saving company [ ] 2) Cooperatives [ ] 3) NGOs [ ] 4) Banks [ ]

5) Private money lenders [ ] 6) Iddirs [ ] 7) Neighbors [ ] 8) Iqubs [ ]

9) Others specify_____

9. Did you get agricultural extension service? 1) Yes [ ] 2) No [ ]

9.1 If your response to question **No9** above is 'Yes'. How much monthly? _____

9.2 If your response to question **No9** is 'No'. Why?

1) The service is general/not specific to income diversification [ ] 2) The advice is not related to multiplication of income source [ ] 3) Do not have time to get the service [ ] 4) No service provider [ ] 5) Others (specify) _____

10. Is their discrimination among male and female headed households to participate in income generating activities from the point of view of society?

1) Yes [ ] 0. No [ ]

10.1. If your response for question **No11** is 'Yes'. Which one?

1) Cultural [ ] 2) Psychological [ ] 3) Economical [ ] 4) Other_____

11. Are you member of any farmer based institution? 1) Yes [ ] 0) No [ ]

11.1 If response for question **No11** above is yes which institution?

1) Iddirs [ ] 2.Equb [ ] 3.Cooperative [ ] 4. Other (specify)_____

11.2 What are services provided from those institutions?

1) Credit [ ] 2) Training [ ] 3) Input provision [ ] 4) Source of income [ ]

5) Find market for output [ ] 6) other service (specify) _____

12) Did you face any shock? 1. Yes [ ] 0. No [ ]

12.1 If your response to **Q No12** above is 'Yes' what types of shock did you face?

1) Covariant shocks [ ] 2) Idiosyncratic shocks [ ] 3) Others (Specify) _____

Part III: Question for female headed rural households about their income with related to outcome variable (Income, saving, and Consumption)

a. Information about Income of female headed households

1. What is your family source of income?

1) Crop production income [ ] 2) Livestock income [ ] 3) Small trade [ ] 3) Firewood and charcoal selling [ ] 4) Local drink selling [ ] 5) Service giving [ ] 6) Remittance [ ]

7) Rental income [ ] 8) other source_____

1.2. Please State your average yearly income in 2013 E.C from your source of income?

Q. No.	Household's source of income	Amount in birr
1	**On Farm**	

	(Agriculture)	
1.1	Crop production (wheat, barley, teff, maize, lentil, pea, bean etc.)	
1.2	Horticultural production (red pepper, garlic, onion, cabbage, potato etc.)	
1.3	Animal sale (ox, cow, calf, heifer, horse, mule, donkey, goat, sheep, chicken)	
1.4	Animal products sale (milk, eggs cheese, butter, hide and skin)	
1.5	Tree planting	
1.6	Sales of grass and crop residues	
1.7	Sharecropping	
2	**Nonfarm Income**	
2.1	**Self-Employment**	
2.1.1	Shop keeping	
2.1.2	Petty trade (grain, livestock, coffee, spices, salt, *etc*.)	
2.1.3	Food processing for sale-local drink like (alcohol, oil)	
2.1.4	Fuel wood and/or charcoal sale	
2.1.5	rural crafts (pottery, carpentry, blacksmiths, weaving)	
2.1.6	fruits sales	
2.1.7	Services (repair of shoes, barber, grain milling, tailor, traditional healing, *etc*.)	
2.2	**Formal Employment**	
2.2.1	Employment in private enterprises	
2.2.2	Employment in government offices and enterprises	
2.2.3	Employment in non-government organizations	
2.2.4	Local election position (paid), kebele chairman	
2.3	**Remittance/transfer**	
2.3.1	Transfer from relatives	
2.3.2	Transfer from friends	
2.3.3	Other transfer (specify)	
2.3.4	Gifts from others	

2.4	**Rent income**	
2.4.1	Rent out house or room	
2.4.2	Rent of animals (oxen, donkey, horse, mule)	
2.4.3	Rent of land	
2.4.4	Others (Specify)	
3	**Off Farm Income**	
3.1	**Informal Employment**	
3.1.1	Housemaid	
3.1.2	Sale of labor for agricultural/non-agricultural work (Daily wage work)	
3.1.3	Cattle herder	
	Total	

1.3. Did your personal income states during income source multiplied was ___?

1) Increased [ ] 2) Decreased [ ] 3) Constant [ ] 4) If other specify_____

1.4. If your income is increased at all, why? Because of _____

1) Adequate market [ ] 2) Under taken new business [ ] 3) Good agricultural season [ ]

4) Profitable of the business [ ] 5) If any mention_____

1.5. If your income is decreased at all, why? Because of_____

1) Poor agricultural season 2) Poor market/sales 3) The business was not profitable

4) I or my family member has been sick 5) if other (specify) _____

1.6. How do you evaluate the impact income diversification to increase your source and level of income?

1) Very high [ ] 2) High [ ] 3) Medium [ ] 4) Low [ ] 5) Very low [ ]

6.No impact [ ] 7) negatively affect [ ] 8) If any other mention _____.

1.7 If negative, please explain it; _____.

2. Are there any factors motivates you to multiple your source of income? 1) Yes [ ] 2) No [ ]

2.1 If your answer is “Yes” for question **No2** abov**e**, please rank your opportunities according to their priority?

Opportunity	Rank		
	1st	2nd	3rd
Push factors			
Scarcity of land			
Inadequacy of farm income			
Large family size			
Seasonality of agriculture			
Illness of family member			
Pull factors			
Availability of non-farm and off-farm income			
Near to the road			
Proximity to urban			
Other factors			

3. If an individual respondent select only one source of income from question **No1** above, why?

1) Lack of information [ ] 2) Lack of credit (collateral) [ ] 3) Adultness [ ]

4) Discrimination of male and female in labour force participation to agriculture and non-agriculture activities [ ] 5) Absence of diversified work [ ] 6) Lack of infrastructures [ ]

7) other factors (specify) _____.

5. What is your last year agricultural income on average in 2012 E.C? _____.

b. Information about consumption expenditure of female headed households.

1. What is your total expenditure from your total income in year 2013 E.C? _____

1.1 For what purpose you spent?

1) To purchase agricultural inputs [ ] 2) To purchase clothes to the children [ ] 3) To purchase food [ ] 4) To pay loan taken from others [ ] 5) To buy livestock [ ] 6) Payment for hired

labour [ ] 7) To start non-farm business [ ] 8) Health expenses [ ] 9) Social ceremonies [ ] 10) Others (Specify) _____

c) **Information about saving of the female-headed households.**

1. Did you save you part of income? 1) Yes [ ] 2) No [ ]

1.1 If yes for question **No1**, how much did you save per a year in birr? _____

1.2 Where do you save?

1) OCSSCO branch [ ] 2) With relatives/friends [ ] 3) In a bank [ ]

4) At home [ ] 5) other development programs [ ] 6) at village money lenders [ ]

7) others/specify_____

1.3 If no to Q1, what is the reason?

1) Lack of money [ ] 2) Lack of awareness [ ] 3) No nearby institution to save [ ]

4) If other specify_____

2. During the last 12 months, your personal cash saving was_____?

1) Decreased Greatly [ ] 2. Decreased [ ] 3) Constant [ ] 4) Increased [ ]

5) Increased greatly [ ] 6) don't Know [ ]

3. What is your source of money for saving? _____.

4. Why you are saving?

1) To finance emergence cases [ ] 2) For loan repayment [ ]

3) To cover household expenses [ ] 4) for safety of my cash [ ] 5) others (specify) _____

d. Information about tax of the female-headed households

1. Did you pay a tax this year? 1) Yes [ ] 2. No [ ]

1.1. If your response is "Yes" for question No1 above. How much you pay in this year? _____

Thank you!