IMPACT ASSESSMENT OF INPUT AND OUTPUT MARKET DEVELOPMENT INTERVENTIONS BY IPMS PROJECT: THE CASE OF GOMMA WOREDA, JIMMA ZONE

M.Sc. Thesis

ALEMU TOLEMARIAM

October 2010

Haramaya University

IMPACT ASSESSMENT OF INPUT AND OUTPUT MARKET DEVELOPMENT INTERVENTIONS BY IPMS PROJECT: THE CASE OF GOMMA WOREDA

A Thesis Submitted to the School of Agricultural Economics and Agri. Business Management, School of Graduate Studies
HARAMAYA UNIVERSITY

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF SCIENCE IN AGRICULTURE
(AGRICULTURAL ECONOMICS)

BY
Alemu Tolemariam

October 2010
Haramaya University

SCHOOL OF GRADUATE STUDIES

HARAMAYA UNIVERSITY

As Thesis research advisors, we here by certify that we have read and evaluated this Thesis prepared, under our guidance, by Alemu Tolemariam Entitled "Impact Assessment of Input and Output Market Development Interventions by IPMS Project: The Case of Gomma Woreda". We recommend that it be submitted as fulfilling the thesis requirement.

Moti Jaleta(PhD)	_____	_____
Major Advisor	Signature	Date
Berhanu G/Medhin (PhD)	_____	_____
Co-Advisor	Signature	Date

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

_____	_____	_____
Name of Chairman	Signature	Date
_____	_____	_____
Name of Internal Examiner	Signature	Date
_____	_____	_____
Name of External Examiner	Signature	Date

.

DEDICATION

I dedicate this thesis manuscript to all of my family members.

STATEMENT OF THE AUTHOR

First, I declare that this thesis is my solely work and that all sources of materials used for this thesis have been duly acknowledged. This thesis has been submitted in partial fulfillment of the requirements for M.Sc. degree at the Haramaya University and is deposited at the University Library to be available to borrowers under rules of the library. I seriously declare that this thesis is not submitted to any other institution anywhere for the award of any academic degree, diploma, or certificate.

Brief quotations from this thesis are allowable without special permission provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the School of Graduate Studies when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Name: Alemu Tolemariam Signature: _____

Place: Haramaya University, Haramaya

Date of Submission: _____

BIBLIOGRAPHICAL SKETCH

The author was born in 1983 in Bako district, West Shoa zone from his father Tolemariam Ejeta and his mother Tejitu Midheksa. After completion of secondary school at Ambo Comprehensive Senior Secondary School he joined Haramaya (the then Alemaya) University in 2000/01, and graduated with B.Sc. degree in Agricultural Extension in July 2004.

After his graduation, he joined Oromiya Agricultural Research Institute (OARI) and worked at Adami Tulu Research Center in Socioeconomics Research Division as Junior Researcher. He had been working for the Institute until he joined the School of Graduate Studies at Haramaya University to pursue his M.Sc. degree in October 2008.

ACKNOWLEDGMENTS

First and foremost, many thanks is directed to our Lord Jesus Christ, the Almighty God, the everlasting father, and the prince of peace and love for giving me the endurance and patience in accomplishing this piece of work.

Special thanks go to my supervisors, Dr. Moti Jaleta, major advisor and Dr. Berhanu Gebremedhin, co-advisor, for their invaluable and constructive comments on the manuscript starting from proposal development to the final thesis preparation.

I would like to thank ILRI-IPMS project for granting me the research fund for this study. I also wish to extend my gratitude for IPMS office staff at Gomma namely Ato Yisehak Baredo,Dadi Gelashe, Habtamu Tiruneh and Abas Jemal, for their keen facilitation of data collection when I was there. Particular thanks go to Dadi Gelashe for his support, facilitation and coordination of activities and providing me morale is the one to be exemplary. Zerihun Teshome,Gazali Mohamed,Mustefa Mogga,Endale Ayele and Hayato AbbaBiya deserve my gratitude for collecting data used for the study diligently. I am indebted to Birke Eneyew,IPMS head quarter, for her timely sending the necessary finance whenever I requested her.

I am very grateful to Oromiya Agricultural Research Institute (OARI) for allowing me for the study leave. Special thanks goes to Dr. Asefa Ta'a, Deputy Director General of OARI, for his unreserved support and cooperation he rendered for me to continue the study at Haramaya University.

I owe a lot to my parents Ato Tolemariam Ejeta,Tejitu Midheksa,Beletu Mindaye,Chaltu Woyesa for nursing and bringing me up to this level. Really this is your fruits that you have contributed starting from my childhood. My special thanks goes to my elder brother Dr. Taye and his family for providing me a laptop computer and hosting me at Jimma during data collection. His encouragement, inspiration and advice are the source of my endurance too. My

brothers Abdisa, Dessalegn and Turi Tolemariam deserve my acknowledgements for their encouragement and praying.

Last but not least let me take this opportunity to forward my thanks to my friends Wubshet Chala,Mesay Yami,Belainesh Hailu,Bayisa Gedefa,Hiwot Abayneh,Hiwot Teshome,Bezawit Yilma,Eden Dagnachew,Beshatu Abera, Million Kassaye and Wubshet Tesfaye for having memorable and funny time during my stay in the University. I respect and love you all.

ACRONYMS AND ABBREVIATIONS

ADLI	Agricultural Development Lead-Industrialization
ATE	Average Treatment Effect
ATT	Average Treatment Effect on Treated
CBSN	Commodity Based Safety Net
CG	Consultative Group
CIDA	Canadian International Development Agency
DA	Development Agent
DID	Difference in Difference
FAO	Food and Agriculture Organization
FGD	Focused Group Discussion
GDP	Gross Domestic Product
GPS	Global Positioning System
HH	Household
ILRI	International Livestock Research Institute
IPMS	Improving Productivity and Market Success
JARC	Jimma Agricultural Research Center
MARC	Melkasa Agricultural Research Center
MoARD	Ministry of Agriculture and Rural Development
MoFED	Ministry of Finance and Economic Development
MoPED	Ministry of planning and Economic Development
OoARD	Office of Agriculture and Rural Development
OCSSCO	Oromiya Credit and Saving Share Company
PA	Peasant Association
PASDEP	Plan for Accelerated and Sustained Development to End Poverty
PLW	Pilot Learning Woreda
PSM	Propensity Score Matching
SNNP	Southern Nation and Nationality People
TLU	Total Livestock Unit

TABLE OF CONTENTS

STATEMENT OF THE AUTHOR **v**

BIBLIOGRAPHICAL SKETCH **vi**

ACKNOWLEDGMENTS **vii**

ACRONYMS AND ABBREVIATIONS **ix**

LIST OF TABLES **xiii**

LIST OF FIGURES **xiv**

LIST OF TABLES IN THE APPENDIX **xv**

ABSTRACT **xvi**

1. INTRODUCTION **1**

1.1. Background 1

1.2. Problem Statement 4

1.3. Objectives of the Study 7

1.4. Scope and Limitation of the Study 7

1.5. Significance of the Study 8

1.6. Organization of the Thesis 8

2. REVIEW OF LITERATURE **9**

2.1. Market Development 9

2.2. Market Participation 11

2.3. Commercial Orientation 12

2.3.1. Measurements of commercial (market) orientation 13

2.3.2. Rationale of commercial orientation 14

2.4. Agricultural Market Institutions and its Role in Marketing 15

2.5. Definitions, Types and Approaches of Impact Assessment 17

2.5.1. Types of impact assessment 18

2.5.1.1. People level impact 18

2.5.1.2. Community level 19

2.5.2. Approaches to impact assessment study 21

2.5.2.1. Experimental Method 23

TABLE OF CONTENTS (*Continued*)

2.5.2.2. Quasi-experimental method 24

2.6. Empirical Studies 29

3. RESEARCH METHODOLOGY 31

3.1. Program Description 31

3.2. Description of the Study Area 32

3.3. Sources of Data 35

3.4. Sampling and Data Collection Techniques 35

3.5. Data Analysis Techniques 37

3.5.1. Qualitative analysis 37

3.5.2. Descriptive analysis 37

3.5.3. Propensity score matching 37

3.5.3.1. Estimating propensity score using binary response model 42

3.5.3.2. Choice of matching algorithm 44

3.5.3.3. Checking overlap and common support 46

3.5.3.4. Testing the matching quality 46

3.5.3.5. Estimation of standard error 48

3.5.3.6. Sensitivity analysis 49

3.6. Variable Choice and Its Definitions 53

3.6.1. Choice and definition of explanatory variables 53

3.6.2. Choice, measurements and indicators of the outcome variables 54

4. RESULTS AND DISCUSSION 58

4.1. Description of Sample Households' Characteristics 58

4.2. Changes in the Organizational and Institutional Aspect of Agricultural Market 60

4.2.1. Change in marketing system 61

4.2.2. Innovative credit for marketing interventions 64

4.2.3. Input supply system 64

4.2.4. Community based safety net (kind of insurance) 65

4.3. Empirical Results 66

TABLE OF CONTENTS (*Continued*)

4.3.1. Estimation of propensity scores 66

4.3.2. Matching participant and comparison households 69

4.3.3. Choice of matching algorithm 71

4.3.4. Testing the balance of propensity score and covariates 72

4.3.5. Estimating treatment effect on treated 75

4.3. 5.1. Impact estimate on total household net income 75

4.3.5.2. Impact estimate on intensity of input use and productivity 76

4.3.5.3. Impact estimate on marketed surplus 78

4.3.5.4. Impact estimate on market orientation of the household 79

5. CONCLUSIONS AND RECOMMEDATIONS 82

5.1. Conclusions 82

5.2. Recommendations 84

6. REFERENCES 87

7. APPENDICES 94

LIST OF TABLES

Table **Page**

Table 1. Conceptual framework in marketing strategies of market development 10
Table 2. Sample size by *Kebele* 36
Table 3. Type, definitions and measurement of variables 54
Table 4. Descriptive statistics of sample households (for continuous variables) 59
Table 5. Descriptive statistics of sample households (for dummy variables)..... 60
Table 6. Types and roles of actors along value chain 61
Table 7. Logit results of household program participation..... 67
Table 8. Distribution of estimated propensity scores..... 69
Table 9. Performance of different matching estimator 72
Table 10. Propensity score and covariate balance 74
Table 11. Chi-square test for the joint significance of variables..... 75
Table 12. ATT for total net income from coffee and other commodities of intervention 76
Table 13. ATT for different input use intensity of coffee production 77
Table 14. ATT for input use for improved hives, sheep fattening and fruits..... 77
Table 15. ATT for productivity for improved hives 78
Table 16. ATT for proportion of produce sold for commodities of intervention 78
Table 17. ATT for market orientation indicators for the commodities..... 79
Table 18. Result of sensitivity analysis using Rosenbaum bounding approach 80

LIST OF FIGURES

Figure **Page**

1. Conceptual framework for comprehensive impact assessment 21
2. Location of study area 34
3. Kernel density of propensity scores 68
4. Kernel density of propensity scores of participant households 70
5. Kernel density of propensity scores of non-participant households 70

LIST OF TABLES IN THE APPENDIX

Appendix Table **Page**

1. Multicollinearity test for continuous explanatory variables..... 95
2. Contingency coefficient for discrete variables..... 95
3. Conversion factor used to calculate TLU 95
4. Labor supply conversion factor (person day equivalent)..... 96
5. Joint significance test (likelihood ratio test) 96
6. Histogram of propensity score 97
7. Histogram of Pscore with common (off) support regions..... 97

IMPACT ASSESSMENT OF INPUT AND OUTPUT MARKET DEVELOPMENT INTERVENTIONS BY IPMS PROJECT: THE CASE OF GOMMA WOREDA

ABSTRACT

The study evaluates the ex-post impact of input and output market development interventions on total household net income, intensity of input use and productivity, marketed surplus and market orientation behavior of the households. Moreover, the study has assessed the change in the institutional and organizational aspect of market of the woreda due to market interventions. For quantitative analysis both program participant and non participant respondents were drawn and cross-sectional survey data were collected from 200 households in Gomma woreda. Propensity score matching method was employed to analyze the impact of the project interventions quantitatively. This method was checked for covariate balancing with a standardized bias, t-ratio, and joint significance level tests. Furthermore, sensitivity analysis of the estimated participation effect to unobserved selection bias was checked using the Rosenbaum bounds procedure. Results show that participation in market development interventions has a significant, positive and robust impact on the outcome variables measured using different indicators. However, for some outcome variable indicators such as household income from coffee commodity, input use for apiculture and fruit production, productivity of improved hives, land allocation for coffee and number of hives possessed by the household are positive but statistically insignificant. The sensitivity analysis also shows that the impact result estimates are insensitive to unobserved selection bias. The qualitative assessment shows that the main changes were the private sector (including agro-industry) involvement in supplying inputs by opening alternative village shops as well as the development of community based input supply system, linking the farmers/private traders to the input importer and potential buyers and innovative credit provision specifically meant to enhance input and output marketing interventions. Furthermore, different platforms specifically for the apiculture and sheep fattening has been set by the community which help them to abide by. To hedge against the risk involved in sheep fattening practice, a kind of community based insurance program was established. In order to alleviate access to market, establishing farmers' cooperatives and linking this cooperatives or individual farmers to potential buyers (including exporters) was done. The thesis finally discusses these results in detail and draws some recommendations.

Key words: Propensity score matching, Gomma, market development, impact, IPMS

1. INTRODUCTION

1.1. Background

Ethiopia is still predominantly agricultural economy which is almost rain fed. The sector accounts for 40 percent national GDP 90 percent of export earnings and 85 percent employment opportunity are dominated by agricultural products (World Bank, 2007).

Such characteristics or contribution to the socioeconomic well-being of the population lead that Ethiopia should have to start from agricultural sector as it employs most of the labor force. Despite its importance in the livelihood of the people and its potential, the sector has still remained at subsistence level due to multifaceted problems (Dercon and Zeitlin, 2009).

The overall development strategy of Ethiopia is based on the development of a strong free market economic system. Policies towards the development of the agricultural sector and its role in the Ethiopian economy as a whole are guided by the strategy of Agricultural Development Led Industrialization (ADLI), which has been put forward by the Government of Ethiopia in 1993. ADLI has an aim to bring about a structural transformation in the productivity of the peasant agriculture and to streamline and reconstruct the manufacturing (Industrial) sector, so that it makes extensive use of the country's natural and labor resources (MoPED, 1993). This strategy has driven the introduction of policies to promote: a more supportive macroeconomic framework and development; liberalized markets for agricultural products; and a strong extension- and credit-led push for intensification of food staples production through the use of modern inputs, especially seed and fertilizer.

Since then, this strategy has been developed further and fine-tuned, most recently in the more nuanced PASDEP, Ethiopia's strategic framework for the five-year period 2005/06-2009/10. Commercialization of agriculture and the growth of the non-farm private sector are two main thrusts of the initiative to accelerate growth. PASDEP also recommends specialization both at farm and community level, a shift to high-value crops, promotion of niche high-value export crops, a stronger focus on selected high-potential areas, supporting the development of large-scale commercial agriculture where it is feasible, and facilitating the commercialization of

agriculture, among others, through improved integration of farmers with markets - both local and global (MoFED, 2006). Moreover, increased availability and utilization of appropriate technologies, an effective and efficient service delivery system and, improving institutional competence and performance, integrated and coordinated service delivery, sustained demand for the agricultural outputs are some strategies which are crucial to making market orientation of agricultural sector a reality are the component of this strategy (Puskur and Hagmann, 2006).

In this strategy, markets are expected to lead production, not the other way round as it has been practiced where farmers look for markets after they produce (Berhanu *et al.*, 2006). The policies, strategies and instruments document clearly emphasizes that the development of Ethiopian agriculture should be based on market-oriented production system. Although both the local and international markets are recognized, in the short term emphasis is put on developing the local markets and in the longer term penetrating the international market. To be successful in competing in the international market, continuous improvement in production efficiency at farm level and quality of products has been envisaged.

As an integral part of this overall strategy, improving the efficiency of markets is underlined. In this regard, four areas are especially emphasized. These are establishing a system of labeling and standards, improving the provision of market information, expanding and strengthening cooperatives, and improving and strengthening the participation of private investors in agricultural marketing. The strategy also stipulates that rural banks be established and expanded to provide financial services to farmers. Acceleration of private-sector involvement in agricultural production, marketing and providing different service is the other components (MoFED, 2006).

In order to realize these policy directions and strategies, several options/efforts have been promoted by the government to increase farmers' income from marketing and processing, including the formation of cooperatives and better access to market information. Even though the Government has market orientation as a goal, government policy is not very clear on how the potential benefits of increased smallholder commercialization could be maximized and the

potential damage minimized i.e. in creating an enabling economic environment in which smallholders can take advantage of commercialization opportunities and progressively move away from the widespread subsistence orientation towards a more viable and market-oriented smallholder sector(Samuel and Ludi,2008). In addition to this, Puskur *et al.* (2007), argue that most past development efforts have been geared towards increasing food production but the development of agricultural markets was not sufficiently emphasized. During this period there was a high degree of control by government institutions with limited involvement private sector and other players.

The challenge, therefore, is to develop a knowledge based system which is capacitated and responsive to markets with linkages between different partners in development and improved development processes, including technology introduction, and input/output marketing to facilitate the development of marketable commodities (IPMS, 2005). Recognizing these government initiatives, MoARD embarked on the Improving Productivity and Market Success (IPMS) Ethiopian Farmer project, which is donor-supported and implemented by the International Livestock Research Institute (ILRI) on behalf of the MoARD.

The project follows a value chain development approach, which is made up of several interconnected components. These components include the development and availability of farm inputs and technology, the agricultural production process, harvesting, storage, processing, marketing and distribution which involves different stakeholders along the value chain including the active involvement of private sectors. It aims at making the institutional linkage between producers, processors, marketers and distributors which are very important in sustained agricultural growth (IPMS, 2005).

The project aims to contribute to improved agricultural productivity and production through market-oriented agricultural development, as a means for achieving improved and sustainable livelihoods for the rural population in Ethiopia. To accomplish this goal, the project supported development and research on innovative technologies, processes and institutional arrangements in four focus areas i.e. knowledge management, innovation capacity building of public and private sector partners, farmers and pastoralists, market oriented production

technologies and input/output marketing and financing; contributing to evidence-based policy making to support innovation processes and capacity development and developed strategies, policy, technology and, institutional options from research and lessons learned (IPMS, 2005).

In doing so, the project has been assisting government endeavors by accelerating the introduction of technology and institutional innovations, as well as adding/modifying innovations in collaboration with relevant stakeholders so that the technology adoption and application is enhanced which in turn help the farmers to improve the farm productivity and their market orientation status.

To this end, the project was implemented in Gomma *woreda* in Jimma zone as one of the cash crop growing areas among ten PLWs.

1.2. Problem Statement

Agricultural marketing in Ethiopia is generally weak and inefficient (Puskur *et al*., 2007). Past agricultural development strategies have mainly focused on production and productivity. Farmer organizations, on the other hand, are weak and are not yet business oriented. The involvement of private sectors can potentially improve the delivery efficiency of input for the producers. However, even though there are conducive policy environments their involvement in input/output marketing and investments are weak (Eleni and Goggin, 2006). The Ethiopian government has recognized the situation and is currently paying attention to the improvement of agricultural marketing in order to improve the economic well being of the farming population who depend on agriculture as a source of income and employment.

Development of agricultural markets contributes towards revitalizing the agricultural sector by increasing agricultural production and productivity. In the past, the government has instituted various programs to encourage agricultural production but the development of agricultural markets was not sufficiently emphasized. Government and the various stakeholders currently recognize the important role that a well-developed market can play in

the process of agricultural development and commercialization which further catalyze production growth and boost rural incomes in the country (Eleni *et al.*, 2003).

Program or project evaluation is one of the components of project design matrix whenever any project has been designed and implemented. According to Ponniah *et al*. (1996), there are several reasons to undertake *ex-post* impact evaluation of any project. These include provision of feed-back to the scientists and the system including policy makers, for accountability purposes including establishing the credibility of the public sector research and development, as justification for increased allocation of resources, learning from and adjusting to new challenges.

Similarly, Baker (2000), argues that evaluating impact is particularly critical in developing countries where resources are scarce and every dollar spent should aim to maximize its impact on poverty reduction. If programs are poorly designed, do not reach their intended beneficiaries, or are wasteful, with the right information they can be redesigned, improved, or eliminated if deemed necessary. The knowledge gained from impact evaluation studies will also provide critical input to the appropriate design of future programs and projects.

Most past impact assessment studies have analyzed the impact of the project/program interventions interms of the economic and environmental changes. The consideration of change in institutional and organizational aspect of marketing as outcome variable is limited in most literatures. Moreover, they are providing qualitative insights into processes and do not assess outcomes explicitly which are now widely seen as unsatisfactory (Ravallion, 2005). On the other hand, where the quantitative estimation methods were applied in estimating program impact, parametric estimation methods have been commonly used to capture the impact of the program/project on outcome of interest that has many limitations in attributing the impact to the program.

IPMS project worked on the market development interventions in one of the PLWs, Gomma which is the key for agricultural sector development and poverty reduction. Linking the producers to the potential buyers and input importers, developing/strengthening producers'

cooperatives, establishing alternative input shops, involving private sectors in input and output marketing were some of the interventions done in the market development component of the project for the selected market oriented commodities, namely; apiculture development, sheep fattening, coffee and tropical fruits production (IPMS, 2007).

To the best of our knowledge no work has been done to analyze the impact of input and output market development interventions by IPMS project on institutional and organizational setups of the *woreda* market, marketed surplus, total net income, intensification and productivity and household market orientation behavior for market oriented commodities of interventions in Gomma *woreda*.

Thus, to fill these gaps, the study attempted to analyze the impact of agricultural market development interventions on the mentioned outcome variables by using a blend of both qualitative and quantitative methods.

1.3. Objectives of the Study

The general objective of the study is to generate information on impact of input and output market development interventions by the IPMS project.

Specific Objectives:

In relation to the project's market development interventions the following specific objectives were set.

1. To describe changes in the organizational and institutional aspect of agricultural market in the district.
2. To assess the impact of market intervention on intensification and productivity of commodities of intervention.
3. To assess the impact of market intervention on household total net income from the commodities of intervention.
4. To assess the impact of market intervention on marketed surplus from the commodities of intervention.
5. To assess the impact of market intervention on market orientation of household.

1.4. Scope and Limitation of the Study

The study was undertaken in Gomma *woreda* of Jimma Zone. The main aim is to evaluate the impact of input and output market development interventions on different outcome of interest. Though there were many PAs where project activities have been undertaken in the *woreda* only few of them were included due to time and resource limitations. Moreover, the analysis was limited to the impact of market development interventions for market oriented commodities undertaken in the *woreda* by the project. Data for the empirical study were collected from both households participating and non-participating in the market interventions using the same survey questionnaire at the same time.

The study is constrained by lack of clear and wide range of previous empirical studies on market development and market orientation particularly on establishing market orientation criteria and its clear indicators for its measurements.

1.5. Significance of the Study

As the study focused on the ex-post impact of the intervention, the information provided in this study has much importance for policy makers and scientific community in terms of providing insights and knowledge. It can also potentially contribute for the growing impact evaluation literature in at least identifying casual effect of market development interventions on different outcome variables at household level. Moreover, it is very helpful for the project in providing the feedback information on its effectiveness and in validating the works done on market interventions endeavors.

1.6. Organization of the Thesis

The thesis is organized as follows. The following section describes literature review that includes concepts on market development, market participation, market orientation and their measurements and linkage of institutions and marketing, impact evaluation methods and empirical studies. Section three introduces the methodology which includes description of the project and study area, source and methods data collection and analysis as well. Section four describes the results and discussion of the research outcomes and finally section five present conclusions and recommendations.

2. REVIEW OF LITERATURE

This section tries to discuss the following sub-headings: market development, market participation, commercial orientation and its measurements, market institutions and its roles impact evaluation methods and empirical studies

2.1. Market Development

Input market development: In this case the small and medium size village traders have been supported in capacity strengthening, linking to input importers, provisions of credit and other material and technical support for the market oriented commodities of interventions. Consequently, the input markets for those high value commodities develop over time. Because of a high demand for commercial input like fertilizer by the producers of commodities, traders (input suppliers) can bring in large quantities to rural areas with a low unit cost (Goetz, 1992).

Output market development: This include identifying the products which have potential and demanded in the domestic as well as international market places, linking producers to potential buyers, provision of market information which contributes for reduction in marketing cost, establishing primary cooperatives in order to improve their bargaining power and further reduce the transaction cost in the value chain approach (Mwape, 2009).

It can be said that market is developed, if in addition to the existing markets, new markets are created like niches and linkage to supermarkets which has not been practiced before though marketing practice has been there for long periods, new customers are targeted, different institutions and organizations have get involved in every activities in the value chain development approach (Mwape, 2009). In this regard, improving quality and quantity of the existing product, targeted marketing strategies are paramount important components.

According to Ansoff(19570,Bellmare and Barrett(2006),the main components of market development are; (i) marketing extension and training, (ii) market information and intelligence network, (iii) grading and standardization at producer's level, (iv) improvement in competition and awareness, (v) accessibility of marketing finance and credit, and (vi)promoting the product by targeting different customers.

Similarly, Eleni and Goggin (2006), explain that market development requires an integrated rather than piecemeal approach, in which the key market institutions needed, such as market information, grades and standards, contract enforcement, regulation, and trade and producer groups, involvement of different stakeholders including private sectors which mutually reinforce each other. Moreover, the interaction of these stakeholders and the institutions which are governing them is very important for the best functioning of the activities.

Table 1. Conceptual framework in marketing strategies of market development

	Existing products	**New products**
Existing markets	**Market penetration** Increase sales of products to existing market segments e.g. decrease prices, promotion	**Product development** Identify opportunities for new or modified products e.g. product differentiation through new packaging, brands, additional processing, quality improvement
New markets	**Market development** Expanding into new geographical area, selling to new segments of the population, New product dimensions or packaging etc.	**Diversification** Identify opportunities for new products for new clients or markets

Source: Adopted from Ansoff (1957)

Input and output marketing system play key roles in adoption of agricultural technologies. If farmers do not have efficient input and output markets, they resist investing in new and more productive technologies (Oechmke *et al.*, 1997). Thus, generally it can be said that, market development increases the competitiveness of selected agricultural sub-sectors that target national, sub-regional and international markets thereby contributing to agricultural growth.

2.2. Market Participation

Ana *et al.* (2008), defined market participation in terms of sales as a fraction of total output, for the sum of all agricultural crop production in the household which includes annuals and perennials, locally-processed and industrial crops, fruits and agro-forestry. This sales index would be zero for a household that sells nothing, and could be greater than unity for households that add value to their crop production via further processing and/or storage.

On the other hand, the commonly approach in the literatures is to divide the market-participation decision into two stages. In the first stage, households that produce a particular commodity decide whether to be net buyers, net sellers, or autarkic in the market for that commodity. In the second stage, net buyers and net sellers determine the extent of market participation (Goetz, 1992; Key *et al*., 2000; Holloway *et al*., 2005; Bellmare and Barrett, 2006).

As argued by Reardon and Timmer (2005), market participation is both a cause and a consequence of economic development. Markets offer households the opportunity to specialize according to comparative advantage and thereby enjoy welfare gains from trade. Recognition of the potential of markets as engines of economic development and structural transformation gave rise to a market-led paradigm of agricultural development. He explained further as households' disposable income increases, so does demand for variety in goods and services, thereby inducing increased demand-side market participation, which further increases the demand for cash and thus supply-side market participation.

The poorest people in the world are farmers with low market participation and low agricultural productivity. Increasing either one could help to improve the other, and both could boost living standards: higher market participation could drive productivity by providing incentives, information and cash flow for working capital, while higher productivity could drive market participation since households with higher productivity are more likely to have surpluses above their immediate consumption needs (Ana *et al.*, 2008).

2.3. Commercial Orientation

The definition of commercialization process varies. These definitions have been taking into account different side of markets, types of commodities, decision making power of farmers etc. According to Dawit *et al*.(2006), commercialization can be defined considering three perspective viz. input versus output, sales versus purchases, and the type of commercial activity (cash crops versus other crops).

However, according to Pingali (1997), agricultural commercialization is more than marketing agricultural outputs. He argued that agricultural commercialization is attained when household product choice and input use decisions are made based on the principles of profit maximization. Moreover, according to von Braun *et al.* (1994), commercialization implies increased market transactions to capture the benefits from specialization. Increased market transactions are more easily attained when there are favorable policies and institutional arrangements that promote open domestic and international trade environment and the development of market infrastructure and support services that facilitate access to existing markets and the opening up of new market opportunities under a secured legal system. The review made by Moti *et al.* (2009), argues about the concept of smallholder commercialization, the meaning is not merely about producing significant amount cash commodities and supplying the surplus to the market. Rather it also consider both the input and output sides of production, and the decision-making behavior of farm households in production and marketing simultaneously. In addition to this, commercialization is not only cash crops as traditional food crops are also frequently marketed to a considerable extent.

The prime objective of commercial oriented households are profit maximization and they are targeting markets in their production decisions (based on market signals and comparative advantage) whereas those of subsistence farmers' production decision are based on production feasibility and subsistence requirements, and selling only whatever surplus product is left after household consumption requirements are met (Pingali and Rosegrant, 1995; Berhanu and Dirk, 2009). Their main objective is to fulfill subsistence requirements. Generalizing the various literatures, commercial orientation can be understood as a transition from subsistence-oriented to increasingly market-oriented patterns of production and resource use.

2.3.1. Measurements of commercial (market) orientation

Although the net welfare gain from agricultural commercialization at the household level is universally accepted, there is no common standard for measuring the degree of household commercialization (Moti *et al.,* 2009)

As specified by von Braun *et al.* (1994), there are three types of commercialization indices at household level.

The first index measures proportion of agricultural output sold to the market and input acquired from market to the total value of agricultural production. In the second type, commercialization of the rural economy is defined as the ratio of the value of goods and services acquired through market transactions to total household income. Thirdly, the degree of household integration to the cash economy is measured as the ratio of the value of goods and services acquired by cash transaction to the total household income.

In addition to the above indices, von Braun *et al.* (1994), have measured commercialization in terms of proportion of land allocated by farmers to commercial crops and in terms of the value of input and output sales and purchases weighted by the value of agricultural production.

In measuring household-specific level of commercialization, Govereh *et al.* (1999), and Strasberg *et al.* (1999), used a household commercialization index (HCI), which is a ratio of the gross value of all crop sales per household per year to the gross value of all crop production. This ratio does not incorporate the livestock subsector, which could be more important than crops in some farming systems (Moti *et al.*, 2009). Keister and Nee (2001), have also measured commercialization structure in terms of the degree of allocation of different resources (such as labor and land) to competing agricultural activities and in terms of total sales of agricultural commodities

Recently Dawit *et al.* (2006), used four approaches to measure the level of household Commercialization: sales-to-output and sales-to-income ratios, net and absolute market positions (either as a net buyer, net seller or autarkic/self-sufficient household), and income diversification or level of specialization in agricultural production

Generally, the measurements of commercialization are expressed broadly by higher proportion of agricultural input and crop output that is marketed for cash and resources allocated in increased amount to this commodities. However, as Moti *et al.* (2009), argues, although there is relatively rich body of literature analyzing the extent of commercialization for crop production, the commercialization process and its measurements in the livestock subsector have received little attention.

2.3.2. Rationale of commercial orientation

The recent move towards market reform in developing countries has renewed an interest in the working of agricultural markets as a source of income, employment and food security (Pingali and Rosegrant, 1995; Timmer, 1997). It is increasingly recognized that the commercialization of surplus output from small-scale farming is closely linked to higher productivity, greater specialization, and higher income (Timmer, 1997). Furthermore, in a world of efficient markets, commercialization leads to the separation of households' production decisions from their consumption decisions, supporting food diversity and overall stability. At the macro level, commercialization has also been shown to increase food security and, more generally, to improve allocation efficiency (Timmer, 1997; Fafchamps, 2005).

However, in the face of imperfect markets and high transaction costs, many smallholders are unable to exploit the potential gains from commercialization (de Janvry *et al.*, 1991; Key *et al.*, 2000, Bernard *et al.*, 2010). In the absence of mechanisms to overcome these constraints, smallholders are unlikely to participate in markets or, when they do, to realize the full benefits of participation. These challenges are particularly important in Sub-Saharan Africa, where empirical evidence suggests that the proportion of farmers engaged in subsistence agriculture remains very high. At the same time, those who participate in markets often do so only at the margins because of the high risks and associated costs (Jayne *et al*., 2006).

2.4. Agricultural Market Institutions and its Role in Marketing

Institutions are defined in many different ways. The most widely quoted one is by North (1990), which defines institutions as humanly devised constraints, made up of formal constraints (i.e., rules, laws, constitutions), informal constraints (i.e., norms of behavior, conventions and self-imposed codes of conduct) that structure human interactions, and their enforcement characteristics. These constraints and the technology employed determine the transaction and transformation costs that add up to the production and marketing costs. Following North (1990), Dorward *et al.* (2005) define institutions as "rules of the game" that define the incentives and sanctions affecting people's behavior and distinguish institutional as sets of rules and structures that govern particular contracts, and the context within which the contracts are governed.

The World Bank (2002), offers a working definition of institutions as rules, enforcement mechanisms and organizations that promote market transactions. These definitions indicate that institutions provide multiple, functions to markets; they transmit information mediate transactions, facilitate the transfer and enforcement of property rights and contracts, and manage the degree of competition.

Markets only work because of institutions**.** Market failures are caused by asymmetric information, high transaction costs and imperfectly specified property rights. These market deficiencies are more pronounced in rural areas with underdeveloped road and

communication networks and other market infrastructure. Where supporting market institutions are lacking, rural markets in areas with low market infrastructure tend to be very thin and imperfect. In the absence of institutions that help to coordinate marketing functions or to link producers to markets, the associated high transportation costs and transaction costs undermine the processes of exchange (Kranton 1996; Gabre-Madhin, 2001) and result in limited or localized markets with little rural-urban linkages (Chowdhury *et al*., 2005). In such circumstances, households produce only a limited range of goods and services for their own consumption because social protection for food security is not provided through markets and government interventions (de Janvry *et al*., 1991).

Shocks and vulnerability to production risk (i.e., weather, pests and sickness) and market risk (i.e., price) that seem systemic to agriculture also lead to imperfect markets and transaction failures (Dorward and Kydd, 2004).

When high transaction costs, asymmetric information and incomplete property rights impede the functioning of markets, market players fail to undertake profitable investments (due to the absence of complementary investments) leading to coordination failures that hinder market functions (Dorward *et al.,* 2003; Dorward *et al*., 2005). Thus, coordination failure along the production to consumption value chain may explain constrained agricultural development and the prevalence of a low equilibrium trap, which is a big challenge to policy (Dorward *et al*., 2003). Overcoming the effects of such market imperfections in agricultural input and output markets would therefore require a deliberate attempt to strengthen institutions that promote coordination of market functions, reduce transaction costs and integrate markets to facilitate a continual transition to a higher level equilibrium (World Bank, 2002).

Various private and public sector market-supporting institutions and institutional arrangements have been proposed to bridge market imperfections, reduce transaction costs, enhance opportunities for the poor in markets and to make the market systems more inclusive and integrated (World Bank, 2002). Among the potential market-supporting institutions that can enhance market functions in rural areas are farmer organizations such as Producer Marketing Groups. Their potential in this process lies in enabling contractual links to input and output markets (Coulter *et al.,* 1999); promoting economic coordination in liberalized

markets and in leveraging market functions for smallholder farmers. However, their success in this process depends on their ability in conveying market information; coordinating marketing functions; defining and enforcing property rights and contracts; facilitating smallholder competitiveness in markets (World Bank, 2002) and more critically in mobilizing their members to engage in markets.

2.5. Definitions and Approaches of Impact Assessment

Different definitions have been given to impact assessment by different organizations and scholars. But the commonly used definition of impact assessment as it is given by Omoto (2003) and Rover and Dixon (2007), is that it is a process of systematic and objective identification of the short and long-term effects–positive and negative, direct or indirect effect of intervention on economic, social, institutional and environments. Such effects may be anticipated or unanticipated, and positive or negative, at the level of the individual, household or the organization caused by on-going or completed development activities such as a project or program.

An impact evaluation assesses the extent to which a project has caused desired/undesired changes in the intended users. It is concerned with the net impact of an intervention on individuals, households or institutions, attributable only and exclusively to that interventions (Baker, 2000). Thus, impact evaluation consists of assessing outcomes of research and developmental changes resulting from interventions.

According to FAO (2000), impact assessment is done for several practical reasons: (1) accountability – to evaluate how well we have done in the past, to report to stakeholders on the return to their investment, and to underpin political support for continued investment; (2) improving project design and implementation - to learn lessons from past that can be applied in improving efficiency of research projects; and (3) planning and prioritizing - to assess likely future impacts of institutional actions and investment of resources, with results being used in resource allocation and prioritizing future projects and activities, and designing policies.

Based on the time continuum, there are two types of impact assessment studies. *Ex ante* type is about assessing the impact of the likely future environments and of expected impacts from interventions. It is applied to assist in decisions on approval and funding of any project where as *ex-post* (which this study meant for) evaluates performance, achievements and impacts of the past activities of the project or program (FAO, 2000). The resulting information is used in accounting for the past use of resources, and as a useful input for future planning.

2.5.1. Types of impact assessment

According to Ponniah *et al.* (1996), comprehensive impact evaluation can be undertaken at two levels viz. people (household) and community.

2.5.1.1. People level impact

People level impact refers to the effect of the intervention on the ultimate users or target group for which the technology is developed and adopted. Impact begins to occur when there is a behavioral change among the potential users. The people level impact deals with the actual adoption of the appropriate technologies and subsequent effects on production, income, environment and/or whatever the development objective may be (Omoto, 2003). The people level impact can be economic, socio-economic, socio-cultural, and/or environmental.

Economic impact assessment

Economic impact measures the combined production and income effects associated with a set of research and development activities (Ponniah and Martella, 1999). The economic impact assessment studies range in scope and depth of evaluation from partial impact studies (adoption studies) to comprehensive assessment of economic impacts (FAO, 2000). One popular type of partial impact assessment is adoption studies that look at the effects of new technologies such as the spread of modern crop varieties on farm productivity and farmers' welfare. Economic impact assessments of the more comprehensive types look beyond mere yield and crop intensities to the wider economic effects of the adoption of new technology.

Social impact assessment

Social impacts are important and need to be considered along with the economic and environmental impacts. Social impacts assessment include the effects of intervention of the project on the attitude, beliefs, resource distribution, status of women, income distribution, nutritional implications, institutional implications etc of the community. These can be assessed through socio-economic surveys and careful monitoring. Social impact has the potential to contribute greatly to the planning process of other types of development projects (FAO, 2000). It can assist in the process of evaluation of alternatives, and to help in their understanding and management of the process of social change.

In many impact assessment studies these impact evaluations is rarely applied or overlooked especially in agricultural research and development programs and focusing usually on economic and environmental impacts. Only few economic studies have included social impact analysis through qualitative assessments (FAO, 2000).

Environmental impact assessment

The importance of environment impact assessment is increasing in agricultural research and development interventions due to the growing concerns of land degradation, deforestation and loss of biodiversity around the world. However, there are few countries and research institutions that have formally assessed the environmental impacts associated with agricultural research projects (FAO, 2000).

2.5.1.2. Community level

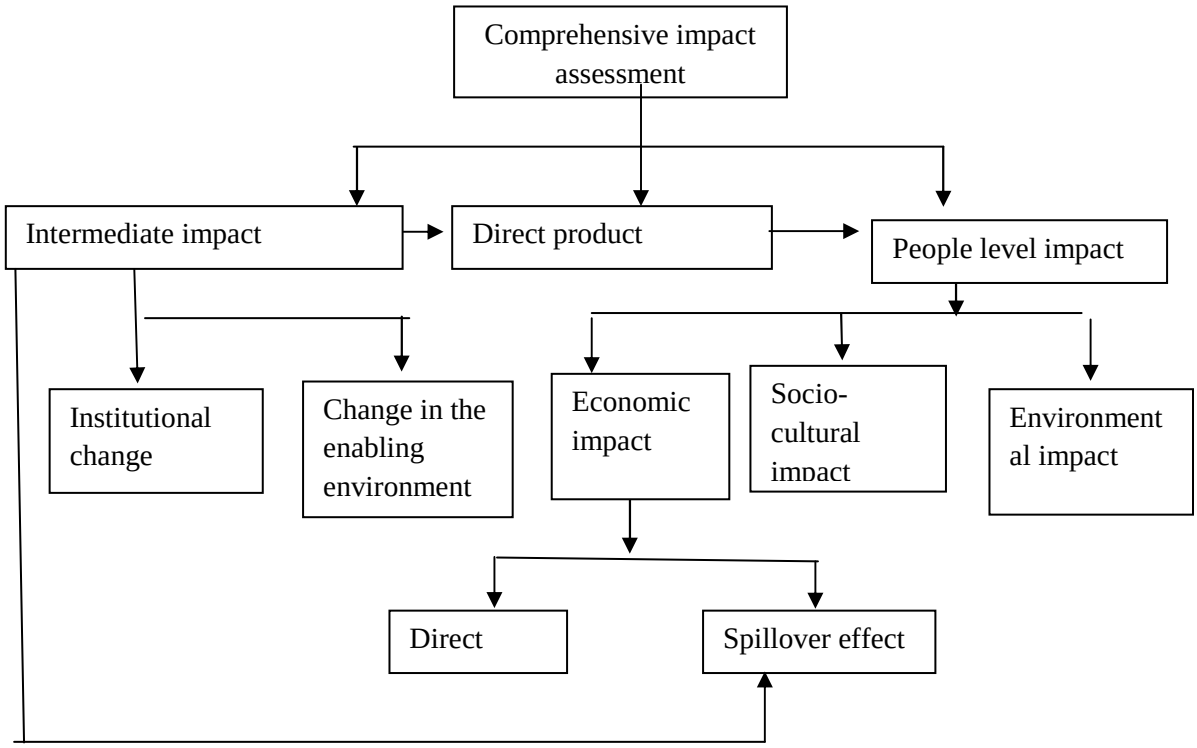
Institutional impact Assessment

According North (1996), institutions are rules of game and organizations and their entrepreneurs are players. Increasing agricultural productivity, whilst strengthening local

institutions, has long been an important goal of agricultural research and development. Organizations play an important role in meeting this goal by improving technologies and knowledge base of the biological, social, economic and political factors that govern the performance of an agricultural system, and by strengthening local institutions' capacity and performance. Most impact evaluation studies are often subjected to rigorous appraisals from economic and environmental perspectives, without giving due attention to the institutional aspect of the interventions (Ponniah *et al.*, 1996).

While economic ,environmental and social impact focuses on the impact evaluation of the technological outputs of research and development organizations in the form of new techniques, methods, information and practices of agricultural systems, institutional impact assessment involves the evaluation of the performance of an intervention in non-technical activities such as training, networking, facilitation, development of methodologies, and advisory services in the areas of research and other policies, organization and management. The concrete results and impacts of institutional development can be difficult to see and may take time to emerge. However, information, generated from institutional impact assessment has the great potential to lead to better, more effective actions and institutional performance of a research and development system (FAO, 2000). Institutional and organizational impact is measured in terms of changes in policy, institutional structure, networking, arrangements and achievements in human capacity buildings (Omoto, 2003).

Figure 1. Conceptual framework for comprehensive impact assessment



Source**:** Ponniah *et al.* (1996)

2.5.2. Approaches to impact assessment study

If one could observe the same individual at the same point in time, with and without the project, this would effectively account for any observed or unobserved intervening factors or contemporaneous events and the problem of endogeneity do not arise (Ravallion,2005; Gilligan *et al.*,2008). Since this is not happening in practice, something similar is done by identifying non-participating comparator groups identical in every way to the group that receives the intervention, except that comparator groups do not receive the intervention.

To know the effect of a project on a participating individual, we must compare the observed outcome with the outcome that would have resulted had that individual not participated in the project. However, as stated earlier two outcomes cannot be observed for the same individual. In other words, only the factual outcome can be observed. Thus, the fundamental problem in

any social project evaluation is the missing data problem (Bryson *et al.*, 2002; Ravallion, 2005).

Estimating the impact of a project requires separating its effect from intervening factors which may be correlated with the outcomes, but not caused by the project. To ensure methodological rigor, an impact evaluation must estimate the counterfactual, that is, what would have happened had the project never taken place (Baker, 2000).

This task of "netting out" the effect of the project from other factors is facilitated if control groups are introduced or constructed form non-beneficiaries. Control groups consist of a comparator group of individuals or households who did not receive the treatment, but have similar characteristics as those receiving the intervention called the treatment groups, the only difference between groups is being project participation. The comparison group should be identical to the treatment group except that the treated group receives the intervention and the non treated ones do not. They make it possible to control for other factors that affect the outcome(confounding factor).Identifying these groups correctly is a key to identifying what would have occurred in the absence of the intervention (Ezemenari *et al.*, 1999; Gilligan *et al.*, 2008).

However, this is difficult to achieve for two reasons. First, beneficiaries of the intervention may be selected on the basis of certain characteristics (purposive targeting) based. If these characteristics are observed then a comparison group with the same characteristics can be selected. But if they are unobserved then in principle only a randomized approach can eliminate selection bias. Second, the comparison group may be contaminated either by spillover effects from the intervention or a similar intervention being undertaken in the comparison area by another agency.

If these differences that could arise from the non-random placement of the program and/or from the voluntary nature of participation in program (self-selection) is not properly accounted for, comparison of outcomes between program participants and non-participants is likely to yield biased estimates of program impact(Gilligan *et al.*,2008).

In theory, evaluators could follow two main quantitative methods in establishing control and treatment groups namely randomization/pure experimental design and non-experimental/ quasi-experimental design. In practice, in the social sciences, the choice of a particular approach depends, among other things, on data availability, cost, and ethics to experiment (Yibeltal, 2008).

2.5.2.1. Experimental Method

Experimental designs, also known as randomization, are generally considered as the most robust of the evaluation methodologies (Baker, 2000). By randomly allocating the intervention among eligible beneficiaries, the assignment process itself creates comparable treatment and control groups that are statistically equivalent to one another, given appropriate sample sizes.

In a randomized experiment, the treatment and control samples are randomly drawn from the same population. In other words, in a randomized experiment, individuals are randomly placed into two groups, namely, those that receive treatment and those that do not. In this case observable and unobservable characteristics get uncorrelated thus no selection bias problem arises. This allows the researcher to determine project impact by comparing means of outcome variable for the two groups which yields an unbiased estimate of impact (Nssah, 2006).

According to Ezemenari *et al*. (1999), a random assignment of individuals to treatment and non-treatment groups ensures that on average any difference in outcomes of the two groups after the intervention can be attributed to the intervention (i.e. both observed and unobserved characteristics is the same for both the treated and the control group).

Random assignment ensures the two groups are statistically similar (drawn from same distribution) in both observable and unobservable characteristics, thus avoiding program placement and self-selection biases (Bernard *et al.,* 2010). If implemented appropriately, this design ensures that potential confounders are balanced across program (intervention) and

control units and therefore any differences in the outcomes between the two can be attributed to the program.

Although experimental designs are considered the optimum approach to estimate project/program impact, in practice, there are several problems. It is not feasible in demand-driven programs in which participants make their own decisions of whether to participate and about the kind of activities to do in the learning process (Ravillion, 2005; Bernard *et al.*, 2010). Baker (2000), also argues that individuals in control groups may change certain identifying characteristics during the experiment that could invalidate or contaminate the results. Moreover, experimental designs can be expensive and time consuming in certain situations, particularly in the collection of new or raw data

2.5.2.2. Quasi-experimental method

Quasi-experimental (nonrandom) methods can be used to carry out an evaluation when it is not possible to construct treatment and comparison groups through experimental design. For projects that are often setup intentionally, it is common to only have access to a single cross-sectional survey done after the project is introduced (Jalan and Ravallion, 2003). These techniques generate comparison groups that resemble the treatment group, at least in observed characteristics, through econometric methodologies which include matching methods, double difference methods, and reflexive comparisons.

When these techniques are used, the treatment and comparison groups are usually selected after the intervention by using nonrandom methods. In some cases a comparison group is also chosen before treatment, though the selection is not randomized. Therefore, statistical controls must be applied to address differences between the treatment and comparison groups and sophisticated matching techniques must be used to construct a comparison group that is as similar as possible to the treatment group (Gilligan *et al.*, 2008).

A quasi-experimental method is the only alternative when neither a baseline survey nor randomizations are feasible options (Jalan and Ravallion, 2003). The main benefit of quasi-

experimental designs are that they can draw on existing data sources and are thus often quicker and cheaper to implement, and they can be performed after a project has been implemented, given sufficient existing data. The principal disadvantages of quasi-experimental techniques are that (a) the reliability of the results is often reduced as the methodology is less robust statistically; (b) the methods can be statistically complex and data demanding; and (c) there is a problem of selection bias. This study employed this method as there is no base line data and as the project placement is not random.

The central methodological challenge in non-experimental evaluation method is that examining outcome response of an intervention involves distilling the effect of intervention per se from that of the factors that affect individuals (Foster, 2003). There are different econometric approach that has been used to avoid or reduce this problem.

Double difference or difference-in-differences (DID): Method in which one compares a treatment and comparison group (first difference) before and after a project (second difference). Comparators should be dropped when propensity scores are used and if they have scores outside the range observed for the treatment group. In this case potential participants are identified and data are collected from them. However, only a random sub-sample of these individuals is actually allowed to participate in the project. The identified participants who do not actually participate in the project form the counterfactual (Jalan and Ravallion, 1999; Baker, 2000).

With this method program impacts are estimated by calculating the difference in outcomes between treatment and control groups after program implementation minus the difference in outcomes between treatment and control groups prior to the implementation. Often, we refer to this double difference or this simple comparison-in-means as the difference-indifference (DID) estimator.

The strength of the panel-based-DID estimator comes from its intuitive appeal and simplicity. It can derive an estimate of the impact by comparing the treatment and control groups using the post treatment data (second difference), after we use the pre-treatment data to equate

treatment and control groups (first difference). In addition to this, DID estimates are known to be less subject to selection bias because they remove the effect of any unobserved time-invariant differences between the treatment and comparison groups.

However, there are at least two disadvantages that relate to the very simplicity of such a panel based impact assessment. First, constructing panel data sets can be expensive, time consuming, and logistically challenging particularly because we need to collect baseline and follow-up data that straddle the implementation of a program. Second, the design assumes that the potential selection bias (i.e., due to administrative targeting or volunteering) is linear and time invariant such that it can be subtracted off in the first differencing (Jalan and Ravallion, 1999). However, these assumptions might be violated if the time period between two panel data sets is long enough so that the unobservable variables of subjects are altered. In addition, the unobservable variables can be changed as the subjects participate in the program which leads the estimate to be biased.

A reflexive comparison: Methods in which a baseline survey of participants is done before the intervention and a follow-up survey, is done after. Here, Participants who receive the intervention are compared to themselves before and after receiving the intervention. The counterfactual group is the set of participating individuals themselves (Jalan and Ravallion, 1999; Baker, 2000).

Propensity Score Matching: Among quasi-experimental design techniques, matched-comparison techniques are generally considered a second-best alternative to experimental design (Baker, 2000). Intuitively, PSM tries to create the observational analogue of an experiment in which everyone has the same probability of participation. The difference is that in PSM it is the conditional probability (P(X)) that is intended to be uniform between participants and matched comparators, while randomization assures that the participant and comparison groups are identical in terms of the distribution of all characteristics whether observed or not. Hence there are always concerns about remaining selection bias in PSM estimates (Ravallion, 2005).

On the other hand researchers also usually applied ordinary regression in adjusting preexisting differences treated and comparison households. Although common, such an approach has some limitations. The first is that regression generally assumes a set of linear relationships between the covariates and the outcome of interest. A second, more subtle problem involves the so called common support or distribution of the covariates (Ravallion, 2005). Not only high- and low-treatment groups differ in terms of the means of those variables, but the distribution of those variables could overlap relatively little. In that case, regression essentially projects the behavior of individuals in one group outside the observed range to form a comparison for the other at common values of the covariate. Such projections can be highly sensitive to functional form (Foster, 2003).

An alternative to econometric regression is statistical matching methods. With this method meaningful counterfactual (control) group is selected among a large group of non-participants, which is identical to the participating group (Bryson *et al.,* 2002; Caliendo and Kopeinig, 2008) to match the characteristics of the project population (causality of potential outcomes) as closely as possible. It matches control groups to treatment groups on the basis of observed characteristics or by a propensity (to participate) score; the closer this score, the better the match. A good control group is from the same economic environment and is asked the same questions by similar interviewers as the treatment group. In recent years there have been substantial advances in PSM technique applications (Rosenbaum and Rubin, 1985; Jalan and Ravallion, 1999).

Unlike econometric regression methods, PSM compares only comparable observations and does not rely on parametric assumptions to identify the impacts of projects and it does not impose a functional form of the outcome, thereby avoiding assumptions on functional form and error term distributions, e.g., linearity imposition, multicollinearity and heteroscedasticity issues. In addition, the matching method emphasizes the problem of common support, thereby avoiding the bias due to extrapolation to non-data region. Results from the matching method are easy to explain to policy makers, since the idea of comparison of similar group is quite intuitive.

Matching the treated and the control subjects becomes difficult when there is a multi-dimensional vector of characteristics (Rosenbaum and Rubin, 1983). The PSM solves this type of problem by summarizing the pre-treatment characteristics of each subject into a single index variable, and then using the propensity score (PS) to match similar individuals. This constitutes the probability of assignment to treatment conditional on pre-treatment variables (Rosenbaum and Rubin, 1983).

Matching estimates is more reliable if: (i) participants and controls have the same distribution of unobserved characteristics; (ii) they have the same distribution of observed characteristics; (iii) the same questionnaire is administered to both groups; and (iv) treated and control households are from the same economic environment. In the absence of these features, the difference between the mean impact of the participants and the matched non-participants is biased estimate of the mean impact of the project (Jalan and Ravallion, 1999).

PSM is not without its potentially problematic assumptions and implementation challenges. First, PSM requires large amounts of data both on the universe of variables that could potentially confound the relationship between outcome and intervention, and on large numbers of observations to maximize efficiency (Bernard *et al.*, 2010) .Second, related to the previous point we can never be entirely sure that we have actually included all relevant covariates in the first stage of the matching model and effectively satisfied the conditional independence assumption (CIA). Furthermore, PSM is non-parametric: we do not make any functional form assumptions regarding the average differences in the outcome. Although the first stage involves specification choices - e.g., functional form like logit and probit, empirical analyses tend to find impact estimates that are reasonably robust to different functional forms. Moreover, if unobservable characteristics also affect the outcomes, PSM approach is unable to address this bias (Ravallion, 2005).

Irrespective of its shortcomings, PSM is extensively used in the recent literature on economic impact evaluation (Jalan and Ravallion 2003). It is very appealing to evaluators with time constraints and working without the benefit of baseline data given that it can be used with a single cross-section of data, where this study envisaged to employ.

2.6. Empirical Studies

Due to dearth of available information on effect of market development intervention studies, only application of the model used by different researcher is discussed.

Studies on application of PSM methods

A number of researchers have applied this semi parametric model to evaluate social programs both in Ethiopia and elsewhere in the world. Below are some of the recent studies who have applied PSM in program evaluations particularly in Ethiopia.

Fitsum *et al,* (2006), used PSM in order to analyze the impact of small scale water harvesting on household poverty in Tigray .The main objective here was to assess whether households with ponds and wells are better off compared to those without. Results show that households with ponds and wells are not significantly better off compared to households without, even though they are comparable in essential household characteristics.

Yebeltal (2008), applied the model to assess the impact of Integrated Food Security Program in Ibant district of Amahara region. The study found that the program has increased participating households' calorie intake by 30% (i.e., 698 calories) compared to that of non-participating households.

In assessing the impact of the Productive Safety Net Program (PSNP) in Ethiopia on livestock and tree holdings of rural households, Andersson *et al.* (2009), have applied PSM model. They found that there was no indication that participation in PSNP leads households to disinvest in livestock or trees. In fact, the number of trees increased for households that participated in the program. It could be the case that participation in PSNP (where tree planting and subsequent forest management work on public lands are usual activities) leads to households becoming more skilled in forestry, and that they switch to increased forest planting as a result.

In analyzing the impact of social protection on food security and coping mechanisms in Ethiopia's productive safety nets program, Gilligan *et al.* (2008), used PSM methods and they found that participation in the public works component of the PSNP (defined as receipt of at least 100 Birr) in payments over the first five months has modest effects. It improves food security by 0.40 months and increases growth in livestock holdings by 0.28 Tropical Livestock Units (TLU). It leads to an increase of 4.4 percentage points in the likelihood that a household is forced to make a distress asset sale.

Bernard *et al*. (2010), applied PSM in assessing the impact of cooperatives on smallholders' commercialization of staple crops using the output price offered and proportion of output sold as indicators. They found that cooperatives deliver, on average, 7 percent price premium for their members' output, relative to what these farmers would have received had they decided to market their output individually. On the other hand, the quantity of grain coming to market from this smallholder farmer is less than it would be without the cooperative's services.

3. RESEARCH METHODOLOGY

3.1. Program Description

IPMS project, funded by the Canadian International Development Agency (CIDA), was implemented since 2006/7 to assist the MoARD in the transformation of smallholder farmers from a predominantly subsistence oriented agriculture to a more market oriented (commercial) oriented agriculture. It has been implemented by ILRI Ethiopia on behalf of MOARD in four major regions (viz. Oromia, Amhara, Tigray and SNNP). The project adopted a participatory market oriented commodity value chain approach which is based on innovation systems and value chain concepts. Crucial elements in the approach are the value chain instead of a production focus, the linking and capacitating of value chain partners and the assessment, synthesis and sharing of knowledge among the partners, participatory commodity development. The project introduced this approach in 10 Pilot Learning *Woredas* (PLWS) in Ethiopia with the objective of testing/adopting the approach so that the respective PLW best practice can be scaled up/out nationwide.

At the PLW level, the program was implemented by existing public and private institutions, including extension/advisory services, agricultural input/service suppliers, credit institutions, cooperatives and private traders. An important aspect of the project was providing these institutions with new ideas and best practices from Consultative Groups (CG) centers and other institutions outside of Ethiopia in addition to the existing institutions in the country. The project's role in the PLW is to facilitate access to agricultural innovations – technologies, policies and processes as well as strengthening the capacity of institutions to better serve farmers and communities (IPMS, 2007).

In Gomma PLW, the project integrates coffee, fruit, apiculture development and sheep fattening as market oriented commodities of intervention. Selection of households into the program involved local consultation (experts and administrators) and a non-random placement. In the first place, *kebeles* were identified in the district based on certain criteria like their accessibility to road and availability of agricultural extension services and

willingness of the farmers to participate and the opportunity and potential of the *kebele* for specific commodity of intervention. Upon selection, households participate in one or more program activities, which include coffee and fruit production, sheep fattening and apiculture development. Households who have been involved in different project's component since 2006/07 were considered as participants.

3.2. Description of the Study Area

Gomma *woreda* is one of the 13 *woredas* in Jimma zone known for predominantly growing coffee. It is located 390 km south west of Addis Ababa and about 50 km west of the Jimma town. One of the coffee biodiversity centers in Ethiopia is found in this *woreda*. There are 36 peasant associations and 3 towns. The number of agricultural households in the district is 45,567 of which 35,533 are male headed and 10,034, female headed (IPMS, 2007). The total population of the district was 21662 of which 110,448 are males and 106,174 females (CSA, 2007). Gomma is the second most densely populated district in the zone with the total area 96,361.72 ha (96.4 km^2) including the two coffee state farms which cover an area of 2704 ha.

The *woreda* has two farming system namely shaded coffee/livestock farming system and cereal/livestock. Thirty-two of the 36 PAs belong to the coffee/livestock farming system. More than 92% of the people in the *woreda* live in this farming system. On the other hand, cereal/livestock farming system consists of four PAs among the total PAs.

The average annual rainfall of the district is 1524 mm with low variability. It is bimodaly distributed in which the small rains are from March to April and the main rainy season lasts from June to October. Hence, crop and livestock production is not constrained by the amount and distribution of rainfall.

Agro ecologically, Gomma *woreda* is classified as 96% Wet *Weina Dega* (Wet Midland) and and 4% Kolla (lowland). Altitude in Gomma ranges from 1387 to 2870 metres above sea level (masl). Most parts of the *woreda* lie between 1387 and 2067 masl.

The three dominant soil types are Eutric Vertisols, Humic alisols and Humic Nitosols. Among these soil types, Nitosols is the most abundant covering about 90% of the *woreda*. These soils are young soils and are generally acidic soils. However, farmers grow crops that are acid tolerant. The pH of the soils in Gomma ranges between 4.5 and 5.5. However, the commonly observed problem related to aluminum and magnesium toxicity as a result of low pH is minimal. There are about 5 rivers in the *woreda*. Even though available land and water resources offer high potential for irrigation development in Gomma, the present utilization level is very poor (IPMS, 2007).

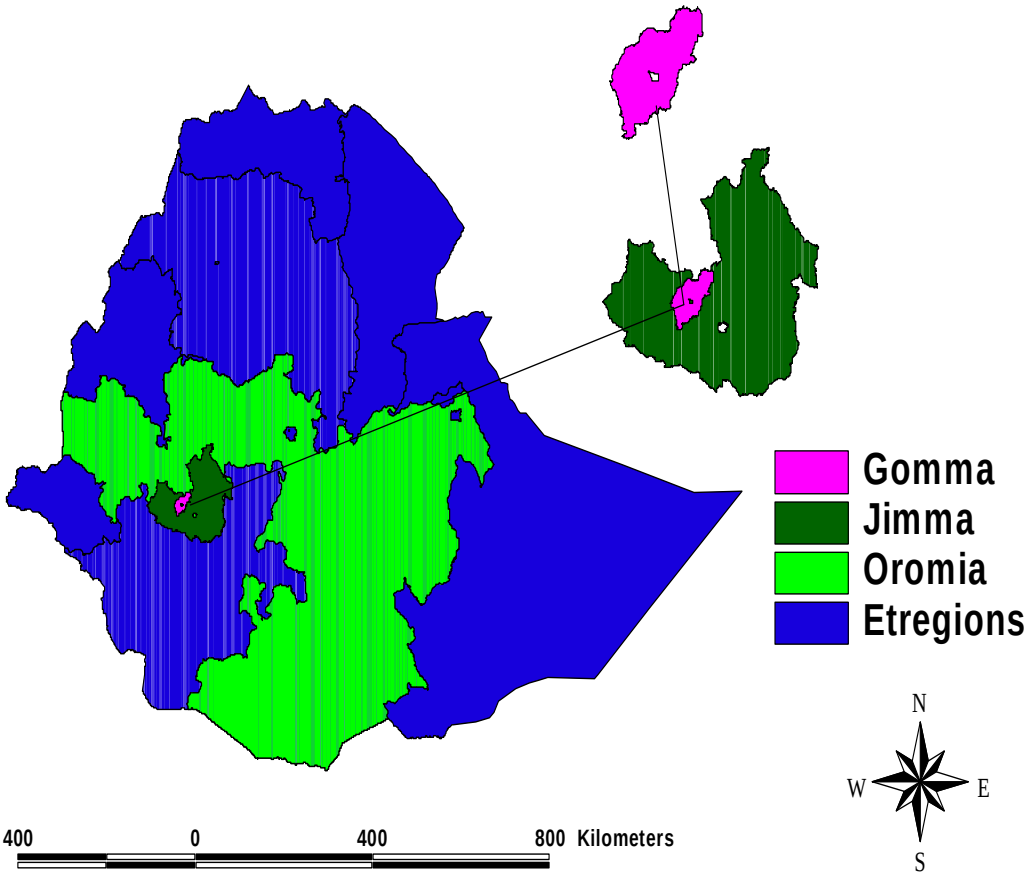


Figure 2. Location of study area

3.3. Sources of Data

Both qualitative and quantitative data were collected from secondary and primary sources. Secondary data relevant to this study were collected from Gomma office of agriculture and rural development, Gomma based PLW and other relevant organizations. Primary data were elicited from the respondents using formal and informal survey.

3.4. Sampling and Data Collection Techniques

Both primary and secondary data sources were used for this study. The primary data needed for the study were obtained from randomly selected farm households. A two stage sampling technique was adopted to generate the required primary data. First 5PAs were selected purposively from 20 project target PAs. Then households were stratified as program participant and non-participants in the selected PAs. Finally, probability proportional to the size was employed to select 100 households from participants and 100 households from non-participants which totally constitute the size of the sample to 200.

Then after, a structured household questionnaire was administered to 200 sampled households of participant and non-participant households in the selected *kebeles*[1]. In doing so, training was given to enumerators about the questionnaire and follow up was made to ensure that the process of data collection was smooth. The survey questionnaire was pre-tested before full scale data collection in order to clarify issues in the questionnaire if any. Finally, the survey was conducted from December, 2009 to January, 2010.

The survey questionnaire was designed to elicit information from a variety of topics including on household resource endowments, access to markets, agricultural services and demographic characteristics of the respondents both at the time of the survey as well as before-interventions of the program using recall methods. Respondents were asked to recall information on easily remembered household characteristics.

1 . In this thesis, program households, treatment households are used interchangeably. On the other hand, non-program households, control households, comparison households are used interchangeably.

In fact, while some household characteristics such as sex are time-invariant, one can easily trace-down pre-intervention characteristics on age, and education given current information questionnaire was designed to elicit information from a variety of topics including on household resource endowments.

Table 2. Sample size by *Kebele*

Sample *Kebeles*	Participants HHs	%	N	Non-participants HHs	%	N	Total
Kilole Kirkir	244	46	46	1400	23	23	69
Behsasha	48	9	9	1200	19	19	28
Bulbulo	86	16	16	1110	18	18	34
Yachi Urache	72	14	14	1140	18	18	32
Omo Gurude	80	15	15	1350	22	22	37
Total	530	100	100	6200	100	100	200

Source: OoARD, 2010

In addition to formal survey, informal survey was undertaken using PRA tools in different villages of the *woreda*. Accordingly, community discussions at four different *kebeles* were undertaken in order to understand the overall community situations and insights about the project activities, its performance, limitations and strength. This information also helps to provide critical insights into beneficiaries' perspectives, the value of projects to beneficiaries, the processes that may have affected outcomes, and a deeper interpretation of results observed in quantitative analysis (Baker, 2000).

Furthermore, interview with experts working in collaboration with the project like *woreda* level agriculture office working on different commodities of intervention were made to broaden the qualitative data base of the study and to enrich the interpretations of the result of quantitative result .

3.5. Data Analysis Techniques

3.5.1. Qualitative analysis

Information on changes in organizational and institutional aspect of agricultural market in the *woreda* were collected from the community using FGD, interviewing for experts in different organizations in the district and reference made to secondary sources which were described and explained qualitatively. This information also used to augment the quantitative analysis results.

3.5.2. Descriptive analysis

Descriptive statistics such as mean, standard deviation, percentages, graphs and cross tabulations were used in analyzing the data.

3.5.3. Propensity score matching

One of the critical problems in non experimental methods is the presence of selection bias which could arise mainly from nonrandom location of the project and the nonrandom selection of participant households that makes evaluation problematic (Heckman *et al.*, 1998). According to Bernard *et al.* (2010), there are three potential source of bias. The first one is that participant households may significantly differ from nonparticipants in community as well as household level due to observable characteristics((such as geographic remoteness, or a household's physical and human capital stock) that may have a direct effect on outcome of interest. Secondly, the difference arises due to unobservable community level characteristic. For instance, the existence of a project may be in part driven by particularly dynamic local leaders at community level. At the household level, a household's expected benefits, its entrepreneurial spirit, or its relationship with other program/project may significantly influence behavior. Thirdly, externalities (spillover effect) exerted by project on nonparticipants.

As a result of the above problems, differences between participants and non-participants may, either totally or partially, reflect initial differences between the two groups rather than the effects of participating in the market interventions.

PSM controls for the households' observable characteristics by comparing the outcomes of program participants with those of matched non-participants, based on similarity in observed characteristics which minimizes the first bias. If not feasible to control for these characteristics, PSM estimation become biased. Having control households from the same communities as program beneficiaries helps to reduce the risks of such bias. However, removing unobservable characteristic remains the main problem of this method.
As Ravallion (2005), argues contamination of the control group can be hard to avoid due to the responses of markets and governments. For instance, Bernard *et al*. (2010), minimize the effect of spillover effect on comparison group by comparing cooperative members to similar households located in other *kebeles* where there are no cooperatives. Nevertheless, as argued by Heckman *et al.* (1998), treatment and comparison households should operate in the same markets and should have come from similar agro-ecology (from sufficiently close locations) and socioeconomic conditions in order to ensure the validity of PSM method.

In order to achieve objectives 2-5, PSM non-experimental method was employed to know the impact of market development interventions made by IPMS on different outcome variables. It is chosen among other non experimental methods because it does not require baseline data, the treatment assignment is not random and considered as second-best alternative to experimental design in minimizing selection biases mentioned above (Baker, 2000).

Mathematical specifications of PSM method

In our case estimating the effect of household's participation in the markets developed by IPMS interventions on a given outcome (Y) is specified as:

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

Where τ_i is treatment effect (effect due to participation in the specific market), Y_i is the outcome on household i, D_i is whether household i has got the treatment or not (i.e., whether a household participated in the market development interventions facilitated by IPMS or not).

However, one should notice that $Y_i(D_i = 1)$ and $Y_i(D_i = 0)$ cannot be observed for the same household at the same time. Depending on the position of the household in the treatment (participation in market development), either $Y_i(D_i = 1)$ or $Y_i(D_i = 0)$ is unobserved outcome (called counterfactual outcome).Due to this fact, estimating individual treatment effect τ_i is not possible and one has to shift to estimate the average treatment effects of the population than the individual one. Two treatment effects are most frequently estimated in empirical studies. The first one is the (population) Average Treatment Effect (ATE), which is simply the difference of the expected outcomes after participation and non-participation:

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

This measure answers the question what would be the effect if households in the population were randomly assigned to treatment. But, Heckman *et al.* (1997), note, that this estimate might not be of importance to policy makers because it includes the effect for whom the intervention was never intended. Therefore, the most important evaluation parameter is the so called Average Treatment Effect on the Treated (ATT), which concentrates solely on the effects on those for whom the program/interventions are actually introduced. In the sense that this parameter focuses directly on those households who participated, it determines the realized impact from the program and helping to decide whether the program is successful or not. It is given by:

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

This answers the question, how much did households participating in the program benefit compared to what they would have experienced without participating in the program. Data on $E(Y_1/D = 1)$are available from the program participants. An evaluator's classic problem is to find $E(Y_0/D = 1)$. So the difference between $E(Y_1/D = 1) - E(Y_0/D = 1)$ cannot be observed for the same household. Due to this problem, one has to choose a proper substitute

for it in order to estimate ATT. The possible solution for this is to use the mean outcome of the comparison individuals, $E(Y_0/D = 0)$, as a substitute to the counterfactual mean for those being treated, $E(y_o/D = 1)$after correcting the difference between treated and untreated households arising from selection effect.

Thus, by rearranging, and subtracting $E(Y_0/D = 0)$ from both sides of equation (3), one can get the following specification for ATT.

$$E(Y_1/D = 1) - E(y_0/D = 0) = \tau_{ATT} + E(Y_0/D = 1) - E(Y_0/D = 0) \quad (4)$$

Both terms in the left hand side are observables and ATT can be identified, if and only if only if $E(Y_0/D = 1) - E(Y_0/D = 0) = 0$. i.e., when there is no self-selection bias. This condition can be ensured only in social experiments where treatments are assigned to units randomly (i.e., when there is no self-selection bias). In non-experimental studies one has to introduce some identifying assumptions to solve the selection problem. The following are two assumptions to solve the selection problem.

I. Conditional Independence Assumption (CIA)

Conditional Independence Assumption is given as

$$Y_0 \perp D/X \quad (5)$$

Where $\perp$ indicates independence, X -is a set of observable characteristics, Y_0 – non-participants. Given a set of observable covariates (X) which are not affected by treatment (in our case, participating in market development), potential outcomes (input use intensity, level of productivity, income, etc) are independent of treatment assignment (independent of how the market participation decision is made by the household). This assumption implies that the selection is solely based on observable characteristics (X), and variables that influence treatment assignment (market participation decision is made by the household) and potential outcomes (input use intensity, productivity level, income) are simultaneously observed

(Bryson *et al*., 2002; Caliendo and Kopeinig, 2008). Hence, after adjusting for observable differences, the mean of the potential outcome is the same for D = 1 and D = 0 and $\mathrm{E}(\mathrm{Y}_0/\mathrm{D} = 1, \mathrm{X}) = \mathrm{E}(\mathrm{Y}_0/\mathrm{D} = 0, \mathrm{X})$.

Instead of conditioning on X, Rosenbaum and Rubin (1983), suggest conditioning on a propensity score (propensity score matching). The propensity score is defined as the probability of participation for household *i* given a set X which is households characteristics $\mathrm{P}(\mathrm{X}) = \mathrm{pr}(\mathrm{D} = 1/\mathrm{X})$. Propensity scores is derived from discrete choice model, and then used to construct the comparison groups. Matching the probability of participation, given covariates solves the problem of selection bias using PSM (Liebenehm *et al.,* 2009). The distribution of observables X is the same for both participants and non-participants given that the propensity score is balancing score (Liebenehm *et al.,* 2009). If outcomes without the intervention are independent of participation given X , then they are also independent of participation given (X) . This reduces a multidimensional matching problem to a single dimensional problem. Due to this, differences between the two groups are reduced to only the attribute of treatment assignment, and unbiased impact estimate can be produced (Rosenbaum and Rubin, 1983).

II. Common support region assumption

The common support is the region where the balancing score has positive density for both treatment and comparison units. This assumption rules out perfect predictability of D given X That is:

$$0 < \mathrm{pr}(\mathrm{D} = 1/\mathrm{X}) < 1 \qquad (6)$$

This assumption improves the quality of the matches as it excludes the tails of the distribution of (X) , though this is done at the cost that sample may be considerably reduced. Yet, non-parametric matching methods can only be meaningfully applied over regions of overlapping support .No matches can be formed to estimate the parameters when there is no overlap between the treatment and comparison groups. It also guarantees an individual with identical

observable characteristics to have a positive probability of belonging both to the participants and control group (Rosenbaum and Rubin, 1983).

Given the above assumptions, the PSM estimator of ATT can be written as:

$$\tau_{ATT} = E[Y_1 - Y_0/D = 0, p(x)] = E[Y_1/D = 1, p(x)] - E[Y_0/D = 0, p(x)] \quad (7)$$

Where $p(x)$ is the propensity score computed on the covariates X Equation (7) is explained as; the PSM estimator is the mean difference in outcomes over the common support, appropriately weighted by the propensity score distribution of participants.

According to Caliendo and Kopeinig (2008), there are steps in implementing PSM. These are estimation of the propensity scores using binary model, choosing a matching algorism, checking on common support condition, testing the matching quality.

3.5.3.1. Estimating propensity score using binary response model

First the propensity score was obtained using either logit or probit models to predict the probability of participation of household. According to Gujarati (1999), both provide similar results. Thus, for comparative computational simplicity logit model was used to estimate propensity scores using households pre-intervention characteristics (Rosenbaum and Robin, 1983) and matching is then performed using propensity scores of each observable characteristics, which must be unaffected by the intervention. These characteristics include covariates variables that influence the participation decisions and the outcome of interest. The coefficients are used to calculate a propensity score, and participants matched with non-participants based on having similar propensity scores.

In estimating the logit model, the dependent variable is market development interventions by IPMS, which takes the value of 1 if a household participated in the specific market and 0 otherwise. The mathematical formulation of logit model is as follows:

$$P_i = \frac{e^{z_i}}{1 + e^{z_i}} \quad (8)$$

Where, P_i is the probability of participation for the i^{th} household and it ranges from 0-1

Z_i: is a function of N-explanatory variables which is also expressed as:

$$Z_i = \beta_0 + \sum \beta_i x_i + U_i \quad (9)$$

Where,

i = 1, 2, 3, - --, n

β_0 = intercept

β_i=regression coefficients to be estimated or logit parameter

U_i= a disturbance term, and

x_i=pre-intervention characteristics.

The probability that a household belongs to non participant is:

$$1 - P_i = \frac{1}{1 + e^{z_i}} \quad (10)$$

Therefore, the odds ratio can be written as:

$$\frac{P_i}{1 - p_i} = \frac{1 + e^{z_i}}{1 + e^{-z_i}} = e^{z_i} \quad (11)$$

Now $\frac{Pi}{1-pi}$ is simply the odds ratio in favor of participating in market development interventions. It is the ratio of the probability that the household would participate in the market development interventions to the probability that he/she would not participate in the market. Finally, by taking the natural log of equation (11) the log of odds ratio can be written as:

$$L_i = \text{Ln}\left(\frac{\text{Pi}}{1-\text{pi}}\right) = \text{Ln}\left(e^{\beta_0 + \sum_{j=1}^{n} \beta_j X_{ji}}\right) = Z_i = \beta_0 + \sum_{j=1}^{n} \beta_j X_{ji} \quad (12)$$

Where Li is log of the odds ratio in favor of participation in the market development, which is not only linear in X_j, but also linear in the parameters.

3.5.3.2. Choice of matching algorithm

Estimation of the propensity score per se is not enough to estimate the ATT of interest. This is due to the fact that propensity score is a continuous variable and the probability of observing two units with exactly the same propensity score is, in principle, zero. Various matching algorithms have been proposed in the literature to overcome this problem. The methods differ from each other with respect to the way they select the control units that are matched to the treated, and with respect to the weights they attribute to the selected controls when estimating the counterfactual outcome of the treated. However, they all provide consistent estimates of the ATT under the CIA and the overlap condition (Caliendo and Kopeinig, 2008). Below, only the most commonly applied matching estimators are described.

Nearest Neighbor (NN) Matching: It is the most straightforward matching estimator. In NN matching, an individual from a comparison group is chosen as a matching partner for a treated individual that is closest in terms of propensity score (Caliendo and Kopeinig, 2008). NN matching can be done with or without replacement options. In the case of the NN matching with replacement, a comparison individual can be matched to more than one treatment individuals, which would result in increased quality of matches and decreased precision of estimates. On the other hand, in the case of NN matching without replacement, a comparison individual can be used only once. Matching without replacement increases bias but it could improve the precision of the estimates. In cases where the treatment and comparison units are very different, finding a satisfactory match by matching without replacement can be very problematic (Dehejia and Wahba, 2002). It means that by matching without replacement, when there are few comparison units similar to the treated units, we may be forced to match

treated units to comparison units that are quite different in terms of the estimated propensity score.

Caliper Matching: The above discussion tells that NN matching faces the risk of bad matches, if the closest neighbor is far away. To overcome this problem researchers use the second alternative matching algorism called caliper matching. 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). If the dimension of the neighborhood is set to be very small, it is possible that some treated units are not matched because the neighborhood does not contain a control unit. One problem in caliper matching is that it is difficult to know a priori what choice for the tolerance level is reasonable.

Kernel Matching: This is another matching method whereby all treated units are matched with a weighted average of all controls with weights which are inversely proportional to the distance between the propensity scores of treated and controls (Becker and Ichino, 2002). Kernel weights the contribution of each comparison group member so that more importance is attached to those comparators providing a better match. The difference from caliper matching, however, is that those who are included are weighted according to their proximity with respect to the propensity score. The most common approach is to use the normal distribution (with a mean of zero) as a kernel, where the weight attached to a particular comparator is proportional to the frequency of the distribution for the difference in scores observed (Bryson *et al*., 2002).

According to Caliendo and Kopeinig (2008), a drawback of this method is that possibly bad matches are used as the estimator includes comparator observations for all treatment observation. Hence, the proper imposition of the common support condition is of major importance for kernel matching method. A practical objection to its use is that it will often not be obvious how to set the tolerance. However, according to Mendola (2007), kernel matching with 0.25 band width is most commonly used.

.

The question remains on how and which method to select. Clearly, there is no single answer to this question. The choice of a given matching estimator depends on the nature of the available data set (Bryson *et al*., 2002). In other words, it should be clear that there is no `winner' for all situations and that the choice of a matching estimator crucially depends on the situation at hand. The choice of a specific method depends on the data in question, and in particular on the degree of overlap between the treatment and comparison groups in terms of the propensity score. When there is substantial overlap in the distribution of the propensity score between the comparison and treatment groups, most of the matching algorithms yield similar results (Dehejia and Wahba, 2002).

3.5.3.3. Checking overlap and common support

Imposing a common support condition ensures that any combination of characteristics observed in the treatment group can also be observed among the control group (Bryson *et al*., 2002). The common support region is the area which contains the minimum and maximum propensity scores of treatment and control group households, respectively. Comparing the incomparable must be avoided, i.e. only the subset of the comparison group that is comparable to the treatment group should be used in the analysis. Hence, an important step is to check the overlap and the region of common support between treatment and comparison group. One means to determine the region of common support more precisely is by comparing the minima and maxima of the propensity score in both groups. The basic criterion of this approach is to delete all observations whose propensity score is smaller than the minimum and larger than the maximum in the opposite group. Observations which lie outside this region are discarded from analysis (Caliendo and Kopeinig, 2008). No matches can be made to estimate the average treatment effects on the ATT parameter when there is no overlap between the treatment and non-treatment groups.

3.5.3.4. Testing the matching quality

Since we do not condition 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

control and treatment group. The main 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 should be avoided after matching. The primary purpose of the PSM is that it serves as a balancing method for covariates between the two groups. Consequently, the idea behind balancing tests is to check whether the propensity score is adequately balanced. In other words, a balancing test seeks to examine if at each value of the propensity score, a given characteristic has the same distribution for the treated and comparison groups. 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), Dehejia and Wahba (2002), emphasized that the crucial issue is to ensure whether the balancing condition is satisfied or not because it reduces the influence of confounding variables.

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.

Standard 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. For each variable and propensity score, the standardized bias is computed before and after matching as:

$$\mathrm{SB(X)} = 100.\frac{\overline{\mathrm{X_1}} - \overline{\mathrm{X_0}}}{\sqrt{0.5.\left(\mathrm{v_1(X)} + \mathrm{v_0(X)}\right)}} \qquad (13)$$

Where $\overline{X_1}$ and $\overline{X_0}$ are the sample means for the treatment and control groups, and ($v_1(X)$ $v_1(X)$ *and* $v_0(X)$ and $v_0(X)$ are the corresponding variance (Caliendo and Kopeining, 2008). The bias reduction (BR) can be computed as:

$$\mathrm{BR} = 100(1 - \frac{\mathrm{B(X)_{after}}}{\mathrm{B(X)_{before}}}) \quad (14)$$

One possible problem with the SB approach is that one does not have a clear indication for the success of the matching procedure.

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 in both 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

Sianesi (2004), suggests re-estimating the propensity score on the matched sample, i.e. only on participants and matched nonparticipants, and comparing the pseudo-R^2s before and after matching. The pseudo-R^2 indicates how well the regressors X explain the participation probability. After matching there should be no systematic differences in the distribution of covariates between both groups and therefore the pseudo-R^2 should be fairly low. Furthermore, one can also perform a likelihood ratio test on the joint significance of all covariates in the probit or logit model. The test should not be rejected before, and should be rejected after, matching.

In our case, in order to test the matching quality of matching estimators the combinations of the above procedures were applied.

3.5.3.5. Estimation of standard error

Testing the statistical significance of treatment effects and computing their standard errors is not a straightforward thing to do. The problem is that the estimated variance of the treatment

effect should also include the variance due to the estimation of the propensity score, the imputation of the common support, and possibly also the order in which treated individuals are matched. These estimation steps add variation beyond the normal sampling variation (Heckman *et al.,* 1998). For example, in the case of NN matching with one nearest neighbor, treating the matched observations as given understate the standard errors.

Bootstrapping: Standard errors in psmatch2 are invalid, since they do not take into account the estimation uncertainty involved in the probit/logit regressions (pscore). One way to deal with this problem is to use bootstrapping as suggested by Lechner (2002). This method is a popular way to estimate standard errors in case analytical estimates are biased or unavailable. Recently it has been widely applied in most of economic literatures in impact estimation procedures. Each bootstrap draw includes the re-estimation of the results, including the first steps of the estimation (propensity score, common support). Bootstrap standard errors attempted to incorporate all sources of error that could influence the estimates.

Abadie and Imbens (2006), argue that using the bootstrap after nearest neighbor matching, until recently a common approach to estimating standard errors in evaluation studies, does not yield valid estimates .In other words, bootstrapping estimate of standard errors is invalid for nearest neighbor matching selection. Thus, calculating analytical standard error is applicable here. Bootstrapping standard errors for kernel matching estimators is not subject to this criticism because the number of observations used in the match increases with the sample size.

The distribution of these means approximate the sampling distribution and thus the standard error of the population mean. Clearly, one practical problem arises because bootstrapping is very time-consuming, computationally expensive and might therefore not be feasible in some cases (Caliendo and Kopeinig, 2008).

3.5.3.6. Sensitivity analysis

Recently checking the sensitivity of the estimated results becomes an increasingly important topic in the applied evaluation literatures (Caliendo and Kopeining, 2008).

Matching method is based on the conditional independence or unconfoundedness assumption, which states that evaluator, should observe all variables simultaneously influencing the participation decision and outcome variables. This assumption is intrinsically non-testable because the data are uninformative about the distribution of the untreated outcome for treated units and *vice versal* (Becker and Caliendo, 2007). As outlined in equation (5) that the estimation of treatment effects with matching estimators is based on the unconfoundedness or selection on observables assumption. However, if there are unobserved variables which affect assignment into treatment and the outcome variable simultaneously, a 'hidden bias' might arise (Rosenbaum, 2002). In other word, if treatment and outcomes are also influenced by unobservable characteristics, then CIA fails and the estimation of ATTs are biased. The size of the bias depends on the strength of the correlation between the unobservable factors, on the one hand, and treatment and outcomes, on the other.

It should be clear that matching estimators are not robust against this 'hidden biases. Different researchers become increasingly aware that it is important to test the robustness of results to departures from the identifying assumption. Since it is not possible to estimate the magnitude of selection bias with non-experimental data, the problem can be addressed by sensitivity analysis.

Rosenbaum (2002), proposes using Rosenbaum bounding approach in order to check the sensitivity of the estimated ATT with respect to deviation from the CIA. The basic question to be answered here is whether inference about treatment effects may be altered by unobserved factors. In other words, one wants to determine how strongly an unmeasured variable must influence the selection process in order to undermine the implications of matching analysis.

The bounding approach does not test the unconfoundedness assumption itself, because this would amount to test that there are no (unobserved) variables that influence the selection into treatment. Instead, Rosenbaum bounds provide evidence on the degree to which any significance results hinge on this untestable assumption. If the results turn out to be sensitive,

the evaluator might have to think about the validity of his identifying assumption and consider other estimation strategies.

As noted above, it is not possible to estimate the magnitude of selection bias using observational data, instead the sensitivity analysis using the bounding approach that involves calculating upper and lower bounds, using the Wilcoxon signed rank test. This rank tests the null hypothesis of no-treatment effect for different hypothesized values of unobserved selection bias.

The central assumption of the analysis is that treatment assignment is not unconfounded given the set of covariates X, i.e., that equation (5) no longer holds. In addition, it is assumed that the CIA holds given X and an unobserved binary variable U: In other words the probability of participation *F(·)* needs to be complemented by a vector U containing all unobservable variables and their effects on the probability of participation captured by γ.

$$P(X,U) = pr(D = 1/X,U) = F(X\beta + U\gamma) = e^{X\beta+U\gamma} \qquad (15)$$

Where γ is the effect of U on the probability of participation in the program. Assuming that *F* follows logistic distribution, the odds ratio of two matched individuals (let say m and n), who are identical in observable characteristics, receiving the treatment written as:

$$\frac{P(X,u_m)}{P(X,u_n)} X \frac{(1-P(X,u_n)}{(1-P(X,u_m)} = \frac{e^{\beta_m X_m+\gamma_m u_m}}{e^{\beta_n X_n+\gamma_n u_n}} = e^{[\gamma(U_n-U_m)]} \qquad (16)$$

Equation (16) states that two units with the same X differ in their odds of receiving the treatment by a factor that involves the parameter γ and the difference in their unobserved covariates U. As long as the there is no difference in U between the two individuals or if the unobserved covariates have no influence on the probability of participation (γ = 0). This happens if the probability of participation will only be determined by the X vector and the selection process is random. γ>0 implies that two individuals with the same observed

characteristics have different chances of participating in the program due unobserved selection bias. In our sensitivity analysis, we examined how strong the influence of γ or ($U_m - U_n$ $U_m - U_n$) on the participation process needs in order to attenuate the impact of market development on potential outcomes.

Following Rosenbaum (2002), equation (16) can be rewritten as:

$$\frac{1}{e^{\gamma}} \leq \frac{P(X, U_m)(1 - P(X, U_n)}{P(X, U_n)\ (1 - P(X, U_m)} \leq e^{\gamma} \qquad (17)$$

Both matched individuals have the same probability of participating only if e^{γ}=1 provided that they are identical in$\mathbf{X}$. Consequently there will be no selection bias on unobservable covariates. If e^{γ}=2, one of the matched individuals may be twice as likely to participate as the other agent (Rosenbaum, 2002). If e^{γ} is close to one and changes the inference about the treatment effect, the impact of participation on potential outcomes, the estimated effect is said to be sensitive to hidden bias. In contrast, insensitive treatment effects would be obtained if a large value e^{γ} does not alter the inference about treatment effects. In this sense, e^{γ} can be interpreted as a measure of the degree of departure from a study that is free of unobservable selection bias (Rosenbaum, 2002). Several values of e^{γ} bounds are calculated on the significance level, and hence, the null hypothesis of no effect of treatment on potential outcomes, is then tested.

Eventually, using predicted probabilities of participation in the program (i.e. propensity score) match pairs are constructed using alternative methods of matching estimators. Then the impact estimation is the difference between simple mean of outcome variable of interest for participant and non participant households.

The difference involvement in market development by IPMS project between treatment and matched control households is then computed. The ATT is obtained by averaging these

differences in market development outcomes (Y_i) across the *k* matched pairs of households as follows:

$$\mathrm{ATT} = \frac{1}{\mathrm{K}}\sum_{\mathrm{i}=1}^{\mathrm{K}}\left[\mathrm{Y}_{\mathrm{i}}^{\mathrm{i}\in\mathrm{D}=1} - \mathrm{Y}_{\mathrm{i}}^{\mathrm{i}\in\mathrm{D}=0}\right] \qquad (18)$$

A positive (negative) value of ATT suggests that households who have participated in market development have higher (lower) of outcome variable Y_i non-participants.

3.6. Variable Choice and Its Definitions

3.6.1. Choice and definition of explanatory variables

In the estimation of the propensity score, we are not interested in the effects of covariates on the propensity score because the purpose of our work is to assess the impact of market development interventions by IPMS project on outcome variables. However, the choice of covariates to be included in the first step (propensity score estimation) is an issue. Heckman et *al.* (1997) argue that omitting important variables can increase the bias in the resulting estimation. In our particular case, variables that determine households' decision to participate in the markets developed by the IPMS project could also affect the outcome variable mentioned above. Here, pre-intervention characteristics, which bring variation in outcomes of interest among program participants and non-participants, were used. In other word, variables which are not affected by being participate in the program or not or those explanatory variables which are fixed throughout are assumed to be used as explanatory variables.

There are no general rules for which variables to include in the model (Anderson *et al.*, 2009). However, the evaluator is guided by economic theory and empirical studies to know which observables (explanatory variables) affect both participation and the outcomes of interest (Bryson *et al.*, 2002).

Accordingly, different socioeconomic, demographic, institutional and location factors were identified below.

Table 3. Type, definitions and measurement of variables

Variable	Types and definition	Measurements
Dependent variables		
Treatment	Dummy, participation in market development of IPMS program	1 if yes,0 otherwise
Covariates		
AGEHH	Continuous, Age of the household head	in years
SEXHH	Dummy, Sex of households	1 if male,0 otherwise
EDULHH	Dummy, Education of head of household	1 if can read and write 0 otherwise
TOTALFMZ	Continuous, Total family size	number of household
EXPFRMG	Continuous, experience of head of households in farming	in years
SZEONLD	Continuous, Size of owned land	in hectare
TLU	Continuous, Livestock holding size	tropical livestock unit
DISNMKT	Continuous, Distance to the nearest market	in kilometers
DISEXTO	Continuous, Distance to extension office	in kilometers
DPCRTO	Continuous, Dependence ratio	number of dependents

Source: Own definitions

3.6.2. Choice, measurements and indicators of the outcome variables

Intensification and productivity: This is one of the outcome variables which is measured by the quantity of inputs used for the market oriented commodities of interventions. There are various inputs type supplied by the private traders and sold to the participant households which are necessary for the production, maintenance and management of commodities. In addition to this, extra labor is required in addition to the family labor especially for coffee commodity especially during peak season (picking, harvesting and processing red cherry). Moreover labor is required for land clearing, preparation and transplanting coffee plants. Thus, intensity of labor use is also measured using person days.

As input use is increasing the productivity of the commodity usually increases. For food crops especially, and other commodities the effect of an increment in input use is immediately reflected in improvement in the productivity which is usually after one year. Nevertheless, for the perennial crops like coffee and fruits the resulting productivity might not immediately be

realized (i.e. it needs relatively long periods). Due to these reasons, our focus here is to analyze the effect of market development interventions on productivity of apiculture commodity only using amount of honey production from improved hives as indicator of productivity.

Marketed surplus: It is the quantity actually sold without meeting farmers' consumptions and utilization requirements (Wolday, 1994).). It is expected that market interventions for each commodities of interventions improves the amount of product taken to the market. The effect of market development interventions on marketed surplus is measured in percentage increase in volume (proportion of sold) of each market oriented crop and apiculture commodities of intervention. However, for the sheep fattening commodity of intervention, the number of fattened sheep taken to the market among the total sheep purchased (kept from own stock or purchased) for fattening purposes was considered as an indicator of marketed surplus.

Household net income: It is one of the outcome variables as a result of the household's participation in specific market development intervention which is measured in birr. Household net income is calculated as the difference between the total revenue generated from sale and total cost incurred by households for the commodities of intervention.

Market orientation behavior: It is the other outcome variable which is used to determine whether the household is market oriented or not due to participation in the market development interventions by the project.

Different researchers have used different market orientation indicators or measurements which are not uniform. According to Bernard *et al*. (2010), farmers' involvement in producing cash crops such as coffee, fruits, or *khat* rather than staple crops such as cereals may by itself reflect the commercial orientation of the household. They have used proportion of the cereal production that was sold and average price received by members of cooperatives as indicators in analyzing the impact of smallholders' commercialization through cooperatives in Ethiopia.

On the other hand, Berhanu and Dirk (2008), used the proportion of households producing market oriented crop and the proportion of land allocated to them as indicators of level of market orientation in studying market orientation of smallholders in selected grains in Ethiopia.

In our study we used the proportion of area under coffee commodity as market oriented indicator. The proportion of sample household possessing improved hives (Kenya Top bar and modern hives) was the other market orientation indicators for the apiculture development intervention. Similarly, number of sheep allocated for fattening either from their own or purchase is an indicator of market orientation for sheep fattening commodity.

Before proceeding to estimate the data using logit model, different tests were undertaken. One of the tests is checking the existence of multicolinrarity between explanatory variables. The presence of multicollinearity among the variables seriously affects the parameter estimates of any regression model. The Variance Inflation Factor (VIF) technique was employed to detect the problem of multicollinearity for the continuous variables (Gujarati, 2004). VIF can be defined as;

$$\mathrm{VIF}(\mathrm{X_i}) = \frac{1}{1 - \mathrm{R_i^2}} \qquad (19)$$

Where R_i^2 is the squared multiple correlation coefficient between X_i and other explanatory variables. The larger the value of VIF, the more troublesome it is. As a rule of thumb if a VIF of a variable exceeeds10, the variable is said to be highly collinear.

Similarly, for dummy variables contingency coefficients test were employed using the following formula

$$\mathrm{C} = \sqrt{\frac{\chi^2}{\mathrm{n} + \chi^2}} \qquad (20)$$

Where C is contingency coefficient, χ^2 is the chi-square value and n=total sample size. For dummy variables if the value of contingency coefficients is greater than 0.75 the variable is said to be collinear.

Heteroscedasticity exists when the variances of all observations are not the same, leading to consistent but inefficient parameter estimates. More importantly, the biases in estimated standard error may lead to invalid inferences (White, 1980). Heteroscedasticity was detected by using Breusch- Pagen test (hettest) in STATA.

Finally, the impact of market development interventions on the outcomes were estimated using STATA 10.0 software using the propensity scores matching algorithm (psmatch2) developed by Leuven and Sianesi (2003).

4. RESULTS AND DISCUSSION

This section consists of three sub-sections. The first one is description of sample households' characteristics. The second subsection is description of the institutional and organizational change of agricultural marketing in the district. The third sub-section is estimation results which include propensity score matching, treatment effect and sensitivity analysis results.

4.1. Description of Sample Households' Characteristics

Both continuous and discrete variables were used in order to describe the sample households included in this study. As already discussed above, pre-intervention type of variables have been used to describe both program participants and non participants. Table 4 shows, the mean differences between the participants and non-participants were significantly differ in size of owned land, distance to extension agents office ,distance to the nearest market, total family size and dependence ratio. On average, participant households have larger size of land; smaller dependence ratio and family size. Compared to non-participants, participant households are living nearer to the office of extension agent and market place. However, dummy variables described in table 5 are statistically insignificant ($p>0.1$) between participant and non-participant households.

Table 4. Descriptive statistics of sample households (for continuous variables)

Pre-intervention variables	Sample Households (N=200)		Participant (N=100)		Nonparticipant(N=100)		Difference in means		T-Value
	Mean	STD	Mean	STD	Mean	STD	Mean	STD[a]	
AGEHH	40.020	8.096	39.911	8.675	40.130	7.508	-0.219	1.144	-0.191
TOTALFMZ	4.816	2.221	4.238	2.035	5.400	2.256	-1.162	0.303	-3.834***
EXPFRMG	18.781	6.867	18.723	6.894	18.840	6.875	-0.117	0.971	-0.121
SZEONLD	1.125	0.814	1.217	0.947	1.031	0.646	0.186	0.114	1.658*
TLU	3.723	2.805	3.578	2.498	3.869	3.090	-0.291	0.397	-0.733
DISNMKT	3.866	1.921	3.381	1.646	4.355	2.058	-0.974	0.263	-3.702***
DISEXTO	2.882	1.626	2.298	1.292	3.472	1.719	-1.174	0.215	-5.469***
DPCRTO	1.805	1.585	1.495	1.361	2.118	1.733	-0.622	0.220	-2.830***

Source: Own survey result, 2010
*** and** means significant at the 1%, 10% probability levels, respectively

[a] *STD for mean difference* = $\sqrt{\frac{STD_1^2}{N_1} + \frac{STD_2^2}{N_2}}$

Table 5. Descriptive statistics of sample households (for dummy variables)

Pre-intervention variables	Category	Participant (N=100)		Non participant (N=100)		Total		$\chi 2$
		N	%	N	%	N	%	
SEXHH	Male	90	90	83	83	173	86.5	2.178
	Female	10	10	17	17	27	13.5	
EDULVLHH	Can read and write	95	95	86	86	181	90.5	3.643
	Cannot read and write	5	5	14	14	19	9.5	

Source: Own survey data, 2010

4.2. Changes in the Organizational and Institutional Aspect of Agricultural Market in the District

After the implementation of the project in the *woreda* changes different organizational and institutional aspects of agricultural market change was observed for different commodities of intervention. There are different individuals/organizations that came together for the accomplishment of different activities and discharging their roles in the value chain approach of the project. Though these actors have been there in the *woreda*, after the start of the project, the synergy and linkage among these actors appeared to be strong and improved.

Table 6. Types and roles of actors along value chain

Stakeholders	Number	Role and responsibilities
Alternative village Input shop	5	Supplying and marketing inputs like hive accessories, harvesting materials, farm implements for coffee production
Local honey traders	NA	Buy honey from beekeepers and sell to other middlemen
Local carpenters	5	Construct Kenya top bar hives and sell to beekeepers
Foundation sheet sellers	1	Produce foundation sheet through buying wax from local beekeepers, prepare the foundation sheet and sell to beekeepers having improved hives
Primary Cooperatives	11	Selling the product of members with better price Searching for better markets Provision of different inputs demanded by farmers
OCSSCO	1	Provision and collection of credit
Apiculture input private traders	NA	Provision and selling of inputs such as hive accessories

Source: OoARD and IPMS-GPW, 2010

NA-not available

4.2.1. Change in marketing system

i. Clustering of individual producer and linking to potential buyers

One of the critical problems for the apiculture development commodity in the *woreda,* which was identified during diagnostic survey, was access to market for both honey produce and inputs (like accessories) which are used to purify the crude honey. There is relatively more production of honey during the harvesting period and during this season prices are drastically lower compared to other times. In order to tackle the marketing problem in the district the project facilitated the establishment of *Wojjin Guddana* primary cooperatives which is meant for increasing the bargaining power, and access to market information thereby increasing

beekeepers' potential for earning better income. This cooperative has developed and ratified bylaws and gets the credential from the *woreda* cooperative promotion desk.

Besides helping the beekeepers to form cooperatives, the project has also tried to link *woreda* honey producers to potential buyers and honey processors like Beza and Alem Mar. The delegate of these organizations came to the *woreda* and took the sample of honey from the beekeepers and tested for its quality and then agreed to buy from the producers.

The project has also tried to link participant coffee producers to private coffee exporters. Two known exporters from Addis Ababa came to *woreda* and observed the process of coffee drying using raised bed made from mesh wire. They witnessed that farmer's drying methods are up to the standard and agreed that product could be exported. Auction based coffee selling was also facilitated/arranged by the project where these exporters came to the *woreda* and bid for buying the special sun dried coffee even though, eventually, the premium offered by the exporter was rejected by the farmers. Similarly, the project has already finalized the process of linking the coffee producers' to Oromiya Coffee Producers Union for better market access. The traceability of the *woreda's* coffee producers who are drying coffee cherry using raised beds constructed from Mesh wire was prepared using GPS and the Union is currently ready to buy the sun dried coffee from participant farmers.

ii. Market information service

The absence of reliable and updated market information service is one of critical problems in the *woreda*, as elsewhere in the country. Availability of market information helps to make informed decisions by market participants if effectively and timely conveyed. In order to minimize this gap, the project tried to provide market information through posting weekly or daily market prices on boards where the majority of producers could see and read. But due to high irregularities and spontaneous change in prices, the efforts have not been as such successful and continued though the initiation was very encouraging.

iii. Improved product quality, storage and processing

The project attempted to improve coffee quality and storage in the *woreda* by backstopping the producers technically, starting from red coffee cherry picking to selling. Different inputs such as mesh wire for raised drying red cherry coffee, jute sacks for dried coffee storage were supplied which enhance the quality of coffee production.

At each village where the project made intervention on coffee quality improvement, coffee quality control committee has been established to monitor and control the respective village starting from coffee picking to drying which is one of the best innovative activities not practiced before.

In addition to this, to improve the quality and quantity of apiculture development in the *woreda,* different hive accessories and other inputs marketing was arranged and some private traders started to bring and sell those inputs to beekeepers. The project facilitated the linkage of these traders to beekeepers and provides credit through OCSSCO. Moreover, since access to improved hives is important in order to keep honey quality and quantity, efforts were made to link beekeepers to carpenters who had experience in manufacturing Kenya Top bar hives which are relatively affordable to the beekeepers than the framed hive.

iv. Business oriented production system

Different experience sharing tours were arranged for selected farmers to acquaint them with production system of market oriented commodities. This has helped them to develop small business running sentiments such as advertising the unique nature of the new variety and overall performance of the varieties of the commodities to other farmers. Participant farmers have already started targeting different holidays, festivals to fatten and sell sheep to fetch better income from sale. Furthermore, farmers have already started multiplying avocado seedlings on their backyard and selling to other farmers and generated additional income.

4.2.2. Innovative credit for marketing interventions

In most developing countries including Ethiopia, financial services and that of market development are largely separated. This finance gap is a key barrier to small-scale farmers' participation in modern markets. Developing innovative financial services that cater for the needs of the small-scale farmer and rural entrepreneur is crucially important to bridge the gap. Prior to IPMS intervention, though there has been a micro finance institution in the *woreda*, the services rendered to farmers were limited and loan conditions were restricted (IPMS, 2007).

Facilitation of credit services for participant farmers by the project which has been channeled through OCSSCO is one of the innovative ways especially to serve the purpose of input marketing. The project channeled the credit to the *woreda* level micro finance branch office and it is this office that disburses the credit and collects the loan as per their rules and regulations. The risk involved during the credit collection was borne by the project and OCSSCO. However, during discussion made with community, farmers were repeatedly complaining the shortage of the repayment period and the large group size that the company requires as peer collateral to provide the loan.

Similarly, the project has supplied credit for private traders who do have capital limitations to supply inputs demanded by the producers for production of the commodities. Moreover, the project managed to link these traders to the input importers in Addis Ababa so as to purchase the required quality and quantity of inputs and then supply then to farmers at a reasonable price.

4.2.3. Input supply system

During the project's diagnosis study, supply of inputs was identified as a major bottleneck for a market oriented commodities in the PLW. To improve this situation, the project focused its attention on strengthening the private sector including community in supplying inputs used for market oriented commodities and the supply of all other inputs and services. Accordingly,

private food oil industry supply cotton meal input which is used as supplementary feed for sheep fattening and with this agro industry linkage was made which helps the fatteners to purchase the required quantity in their nearby town. Moreover, the involvement of private actors such as traders and entrepreneurs in supplying inputs by opening alternative village shops for commodities of interventions especially in supplying input for coffee and apiculture were very encouraging which were entirely the role of the public sector in the previous periods. These private traders were linked to the input importers in the Addis Ababa. The role changes are consistent with the strategy document set by MoARD, which deals with input and output marketing and implementation mechanisms. The document clearly states the need for increased privatization of input supply, while recognizing the role of the government.

Lack of seedling of improved fruits, especially avocado, was one of the critical problems in the district. There has been an occasional distribution of seedling through OoARD and JARC though it was not adequate. But after the intervention by the project model farmers were selected, trained and purchased the mother seedling from MARC with the help of the project. After grafting, farmers have been selling improved seedlings to other interested farmers. Due to this encouragement observed on the participants' farmers, currently, there is a great effort in establishing and scaling out the community based improved seedling production, marketing and exchange system which further reduce the shortage of the availability of improved seedling of fruits in the *woreda* at large.

4.2.4. Community based safety net (kind of insurance)

Community Based Safety-Net (CBSN) is a sort of small-scale insurance type established and owned by the sheep fattening groups in the *woreda* through the facilitation and support of the project. The main purpose of insurance type was to pay compensation to farmers for lost or dead sheep. In order to accomplish the activities the fatteners developed bylaw for its smooth operation.

The scheme consists of two committee groups elected by the fattening group in the community. One of the committee is the sub-committee which has a chairman and secretary.

The main role of this committee is; to receive claims from fattening group members', to verify the claim using its own mechanism; to collect a written justification from all fattening group members be it for or against the loss cover, to review it thoroughly and passes it to the executive committee with its firm comment. The final report of the sub-committee assessment is completed and submitted to the executive committee only 10 days after the claim is reported.

The second type is the executive committee. It consists of chairman, secretary and casher. The executive committee is accountable to the general assembly and undertakes the following activities; reviewing the claim report submitted by each sub-committee and verifying it using its own mechanisms if need be; decide to cover or not to cover the loss; draw the cash saved in the saving wing of OCSSCO and effect payment to the claimer. Payment is effected within 5 days after the sub-committee submits the cases to the executive committee. The scheme covers 80% of the cost of the fattened sheep lost or dead to the claimers.

In order to run the scheme each fattener was contributing 10 birr per sheep and this amount of money is deposited in the saving account wing of OCSSCO.

4.3. Empirical Results

This part explains the estimation of propensity score, matching methods, common support region, balancing test and eventually sensitivity analysis.

4.3.1. Estimation of propensity scores

The logistic regression model was used to estimate propensity score matching for participant and non-participants households. As, indicated earlier, the dependent variable is binary that indicate households' participation decision in the market development interventions.

Before proceeding to impact estimation, Variance Inflation Factor (VIF) was applied to test for the presence of strong multicollinearity problem among the continuous explanatory

variables (see Appendix 1). Moreover, by using contingence coefficients(C) multicollinerty between discrete variables were checked (Appendix 2). There was no explanatory variable dropped from the estimated model since no serious problem of multicollinearity was detected from the VIF results. Similarly, heteroscedasticity was tested by using Breusch-Pagen test. This test resulted in rejection of the existence of heteroscedasticity hypothesis as (p= 0.507) and there was no need to make the standard error robust.

Results presented in Table 7 show the estimated model appears to perform well for the intended matching exercise. The pseudo-R^2 value is 0.20. A low R^2 value shows that program households do not have much distinct characteristics overall and as such finding a good match between program and non-program households becomes easier.

Table 7. Logit results of household program participation

Covariates	Coefficients.	Std. Err.	Z
AGEHH	0.0051	0.0364	0.14
SEXHH	0.3348	0.4821	0.69
EDULHH	0.6408	0.6327	1.01
TOTALFMZ	-0.2571	0.1013	-2.54**
EXPFRMG	0.0050	0.0408	0.12
SZEONLD	0.4348	0.2387	1.82*
TLU	0.0094	0.0679	0.14
DISNMKT	-0.2155	0.0919	-2.35**
DISEXTO	-0.4762	0.1148	-4.15***
DPCRTO	-0.1521	0.1287	-1.18
_cons	2.0286	1.3515	1.5
N	200		
LR chi2(10)	57.08		
Prob > chi2	0.0000		
Log likelihood	-110.77		
Pseudo R2	0.2049		

***, ** and * means significant at the 1%, 5% and 10% probability levels, respectively.

Looking into the estimated coefficients (Table7), the results indicate that program participation is significantly influenced by four explanatory variables. Access to the market, living nearer to office of extension agent, size of owned land and family size are significant variables which affect the participation of the household to the program. Households nearer to market and office of the extension agents are more likely to be included in the program than those living far from the market and extension office. Similarly, households who do have largest family size are less likely to participate in the market development interventions developed by IPMS project than households having family sizes. By contrast, size of owned land has a strong and positive effect on household program participation.

Figure 3 below portrays the distribution of the household with respect to the estimated propensity scores. In case of treatment households, most of them are found in partly the middle and partly in the right side of the distribution. On the other hand, most of the control households are partly found in the center and partly in the left side of the distribution.

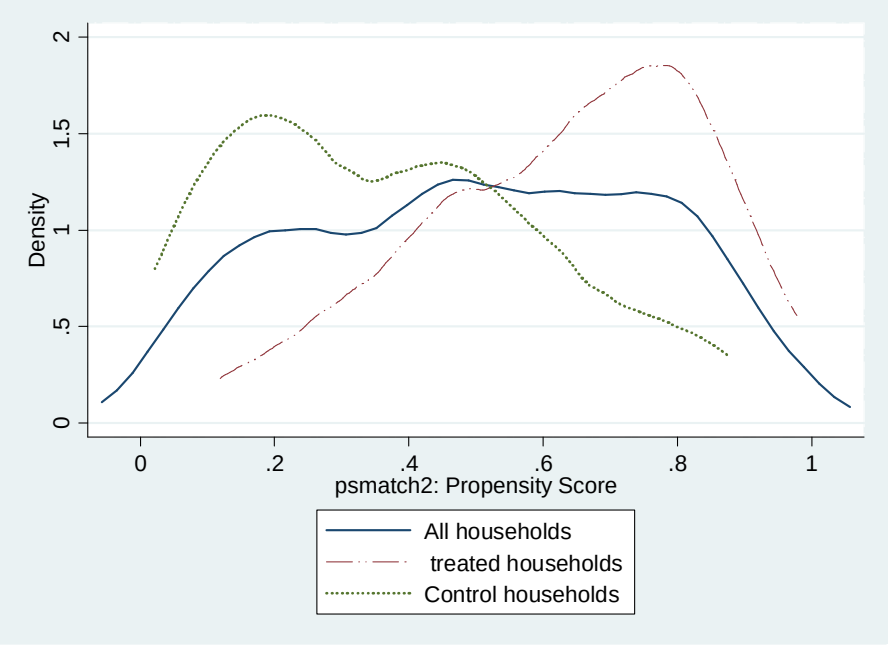


Figure 3. Kernel density of propensity score distribution

4.3.2. Matching participant and comparison households

As stated before, four main tasks should be accomplished before one launches the matching task itself. First, predicted values of program participation (propensity scores) should be estimated for all households in the program and outside the program.

Second, a common support condition should be imposed on the propensity score distributions of household with and without the program. Third, discard observations whose predicted propensity scores fall outside the range of the common support region. And finally sensitivity analysis should be done in order to check the robustness of the estimation (whether the hidden bias affects the estimated ATT or not).

As shown in Table 8, the estimated propensity scores vary between 0.12 and 0.98 (mean =0.63) for program or treatment households and between 0.02 and 0.87 (mean = 0.37) for non program (control) households. The common support region would then lie between 0.12 and 0.87. In other words, households whose estimated propensity scores are less than 0.12 and larger than 0.87 are not considered for the matching exercise. As a result of this restriction, 23 households (10 program and 13 control households) were discarded from the analysis.

Table 8. Distribution of estimated propensity scores

Group	Obs	Mean	STD	Min	Max
Total households	200	0.50	0.25	0.02	0.98
Treatment households	100	0.63	0.21	0.12	0.98
Control households	100	0.37	0.23	0.02	0.87

Source: Own survey result

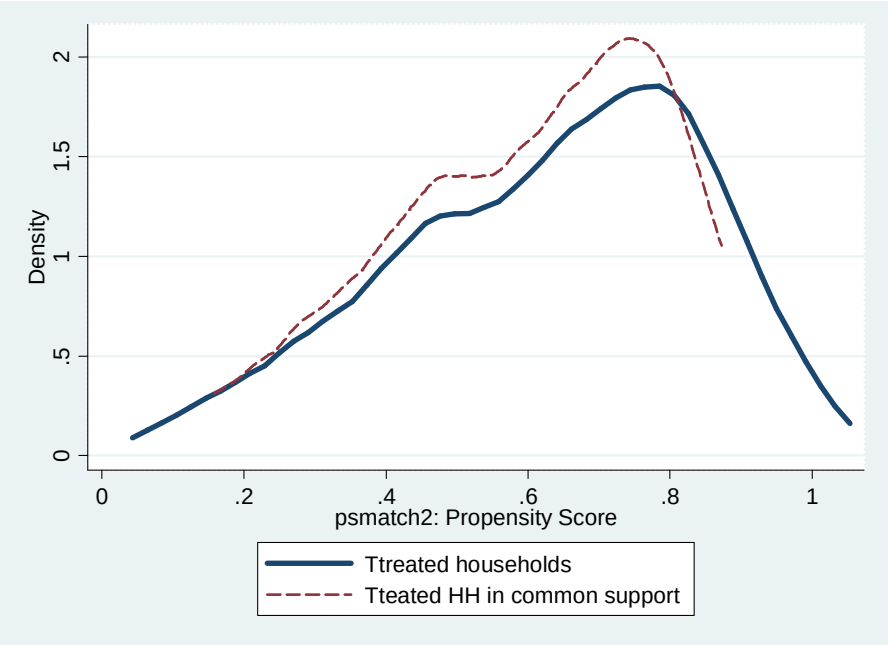


Figure 4. Kernel density of propensity scores of participant households

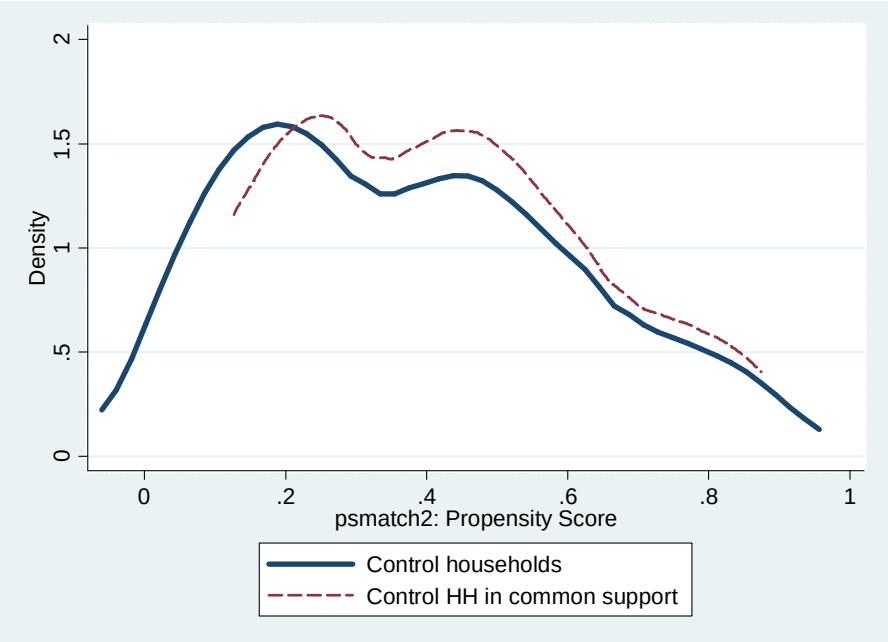


Figure 5. Kernel density of propensity scores of non-participant households

4.3.3. Choice of matching algorithm

Alternative matching estimators were tried in matching the treatment and control households in the common support region. The final choice of a matching estimator was guided by different criteria such as equal means test referred to as the balancing test (Dehejia and Wahba, 2002), pseudo-R^2 and matched sample size. Specifically, a matching estimator which balances all explanatory variables (i.e., results in insignificant mean differences between the two groups), bears a low R^2 value and results in large matched sample size is preferable.

Table 9 shows the estimated results of tests of matching quality based on the above mentioned performance criteria. After looking into the results, it has been found that kernel matching with a band width of 0.25 is the best estimator for the data at hand. As such, in what follows estimation results and discussion are the direct outcomes of the kernel matching algorithm based on a band width of 0.25.

Kernel matching associates the outcome of the treated household with the matched outcome that is given by a kernel-weighted average of all control groups for market development. Since the weighted averages of all market development interventions in the control group are used to construct the counterfactual outcome, kernel matching has an advantage of lower variance because more information is used (Heckman *et al.*, 1998).

Table 9. Performance of different matching estimator

	Performance criteria		
Matching estimator	Balancing test*	pseoudo-R^2	matched sample size
NN			
NN(1)	6	0.182	177
NN(2)	6	0.142	177
NN(3)	6	0.132	177
NN(4)	7	0.134	177
NN(5)	6	0.128	177
Radius caliper			
0.01	8	0.096	151
0.25	10	0.022	175
0.50	9	0.048	177
Kernel			
band width 0.1	6	0.112	177
band width 0.25	10	0.084	177
band width 0.5	8	0.0006	177

Source: own calculation result

* Number of explanatory variables with no statistically significant mean differences between the matched groups of program and non-program households.

4.3.4. Testing the balance of propensity score and covariates

After choosing the best performing matching algorithm the next task is to check the balancing of propensity score and covariate using different procedures by applying the selected matching algorithm(in our case kernel matching). As indicated earlier, the main purpose of the propensity score estimation is not to obtain a precise prediction of selection into treatment, but rather to balance the distributions of relevant variables in both groups. The balancing powers of the estimations are ascertained by considering 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.

The mean standardized bias before and after matching are shown in the fifth columns of Table 10, while column six reports the total bias reduction obtained by the matching

procedure. In the present matching models, the standardized difference in X before matching is in the range of 1.7% and 77.2% in absolute value. After matching, the remaining standardized difference of X for almost all covariates lie between 0.8% and 26.2%, which is below the critical level of 20% suggested by Rosenbaum and Rubin (1985). In all cases, it is evident that sample differences in the unmatched data significantly exceed those in the samples of matched cases. The process of matching thus creates a high degree of covariate balance between the treatment and control samples that are ready to use in the estimation procedure

Similarly, t-values in Tables 10 show that before matching half of chosen variables exhibited statistically significant differences while after matching all of the covariates are balanced.

Table 10. Propensity score and covariate balance

Variable	Sample	Mean		% bias	% reduction bias	T-test	
		Treated	Control			T	P>/t/
	Unmatched	0.629	0.375	115.5		8.19	0.00
PSCORE	Matched	0.593	0.494	45.1	60.9	1.14	0.258
	Unmatched	39.911	40.13	-2.7		-0.19	0.848
AGEHH	Matched	39.622	39.525	1.2	55.4	0.01	0.990
	Unmatched	0.901	0.830	20.8		1.48	0.141
SEXHH	Matched	0.900	0.859	12.1	42.0	0.29	0.768
	Unmatched	0.941	0.860	27.0		1.92**	0.057
EDULVLHH	Matched	0.933	0.928	1.7	93.8	0.01	0.996
	Unmatched	4.238	5.400	-54.1		-3.84***	0.000
TOTLFMSZ	Matched	4.500	4.887	-18.0	66.7	-0.32	0.749
	Unmatched	18.723	18.840	-1.7		-0.12	0.904
EXPFRMG	Matched	18.511	18.454	0.8	51.0	0.04	0.971
	Unmatched	1.217	1.031	22.9		1.63	0.106
SZOWLN	Matched	1.130	1.065	8.0	65.0	0.08	0.937
	Unmatched	3.578	3.869	-10.3		-0.73	0.464
TLU	Matched	3.619	3.778	-5.7	45.2	-0.27	0.790
	Unmatched	3.381	4.355	-52.3		-3.71***	0.000
DISTMKTN	Matched	3.528	3.905	-20.2	61.3	-0.55	0.586
	Unmatched	2.298	3.472	-77.2		-5.48***	0.000
DISTEXTO	Matched	2.457	2.856	-26.2	66.0	-0.71	0.476
	Unmatched	1.495	2.118	-39.9		-2.83***	0.005
DPCERATI	Matched	1.567	1.787	-14.1	64.6	-0.52	0.604

Source: Own estimation result

*** and ** means significant at the 1%, and 5% probability levels, respectively.

The low pseudo-R^2 and the insignificant likelihood ratio tests support the hypothesis that both groups have the same distribution in covariates X after matching (see Table 11). These results clearly show that the matching procedure is able to balance the characteristics in the treated and the matched comparison groups. We, therefore, used these results to evaluate the effect of market development interventions by IPMS project among groups of households having similar observed characteristics. This allowed us to compare observed outcomes for

participants with those of a comparison groups sharing a common support. For detail of Chi-square test for joint significance for the three different matching algorithms (see Appendix 5).

Table 11. Chi-square test for the joint significance of variables

Sample	Pseudo R^2	LR chi2	p>chi2
Unmatched	0.207	57.62	0.000
Matched	0.008	1.90	0.999

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

4.3.5. Estimating treatment effect on treated

In order to attain the above stated objectives 2-5, the following impact indicators of the treatment effect have been performed using the already mentioned PSM model.

4.3. 5.1. Impact estimate on total household net income

Table 12 shows that our impact estimate does not yield an impact significantly different from zero for total household net income from coffee production. The explanations for this is the absence of better market in paying better price for quality sun dried coffee cherry. Even though the project facilitated auction selling base, on one hand the number of exporters who appeared for the auction were few which make them to set the price they want than the competitive price. On the other hand, though one of the exporters won the bid and agreed to pay additional 2 birr/kg compared to conventional dried coffee, most farmers were not satisfied and finally rejected the premium. Then most of them have decided to sell through the usually process (to collectors). Similarly, during informal survey participant farmers said that though the linkage made to the coffee exporter is good, as it promotes their quality coffee, they have been highly dissatisfied to the price the exporters set for their special sundry coffee.

Most of them sold the product to the private traders without significant price difference with conventionally dried coffee.

Table 12**.** ATT for total net income from coffee and other commodities of intervention

	Treated	Control	Difference	SE[1]	t-value
Household income from coffee(birr)	5324.29	4972.75	351.54	321.13	1.09
Household income from other(birr)	2443.26	1897.75	545.51	281.15	1.94*

1. The bootstrapped SE is obtained after 100 replications
*Significant at 10% probability level

The above Table also shows that on average, participant household get 22% more income than the control one from other commodities of intervention (apiculture, sheep fattening and fruits). This result is also statistically significant at 10%.

4.3.5.2. Impact estimate on intensity of input use and productivity

a) Impact estimate on intensity of input use for coffee

Labor use is one of the limiting factors for coffee production. There is an extra-household labor demand during peak seasons (during harvesting, processing and selling red coffee cherry). Labor is also required during coffee land preparation and its management.

As indicated in the Table13, the result reveals that the intensity of labor use is relatively higher for both treated (144PD) and control (118PD). Program households use 18% higher labor than comparison households which is statistically significant at 10% probability level. This result is consistent with the findings of a study by Samuel and Ludi(2008) ,which was done in the same area, who found more labor demand for coffee production for different a level of commercialized households during peak season

Inputs such as mesh wires which are used to construct raised beds were purchased by participating households for drying coffee that further helps in improving its quality.

Machetes and *Zapas* are also other input types which are widely purchased and used by the farmers for coffee production and drying purposes in the *woreda*. We also estimated the impact of the program and found that number of beds constructed from mesh wire use for program household increased by nearly 90% compared to non participants. Similarly, on average, the number of Machetes and *Zapas* purchased and used by participant households are 2.31 more than non participant households.

Table 13. ATT for different input use intensity of coffee production

	Treated	Control	Difference	SE[1]	t-value
Intensity of labor use (PD[b])	144.182	118.249	25.934	15.168	1.710*
Beds used(no.)	1.878	0.197	1.680	0.186	9.030***
Machets and *Zapa* use(no.)	3.333	1.026	2.307	0.214	10.798***

*** and *Significant at 1% and 10% probability levels

1. The bootstrapped SE is obtained after 100 replications

[b] person days

b) Impact estimate on intensity of input use for apiculture, sheep fattening and fruits

After controlling for observable confounding factors, we found statistically significant program effect for quantity of cotton meal used as feed supplements for sheep fattening commodities between treated and control households. However, our estimate does not yield statistically significant effect for the values of input use intensity for apiculture and fruit seedling production

Table 14. ATT for input use for improved hives, sheep fattening and fruits

	Treated	Control	Difference	S.E[1].	t-value
Value of apiculture input use(birr)	111.256	80.683	30.573	39.650	0.770
Quantity of cotton meal use(kg)	92.135	0.000	92.135	8.712	10.580***
Value of fruit input use(birr)	90.250	55.683	34.567	32.650	1.059

1. The bootstrapped SE is obtained after 100 replications

*** Significant at 1% probability level

Impact estimate on productivity of apiculture

The empirical analysis for the productivity of honey indicates that the difference between the two groups does not yield statistically significant effect (P>0.1).

Table 15. ATT for productivity for improved hives

	Treated	Control	Difference	S.E[1].	t-value
Productivity of improved hives(kg/hive)	28.432	25.407	3.025	2.275	1.33

1. The bootstrapped SE is obtained after 100 replications

4.3.5.3. Impact estimate on marketed surplus

Our findings in Table 16 indicate that the proportion of coffee sale is relatively high for both treated (82%) and control (78%) households which further indicate that coffee is the major commercialized commodity in the district. Treated households' sell coffee 5% higher compared to control households which is statistically significant at 5% probability level.

Similarly, our impact estimate for the proportion of sale for honey, sheep fattening and fruit seedling shows 19%, 21% and 46% higher respectively for participants which is significantly different from zero.

Table 16. ATT for proportion of produce sold for commodities of intervention

	Treated	Control	Difference	S.E[1].	t-value
Sale of coffee (%)	0.822	0.782	0.040	0.017	2.353**
Sale of honey (%)	0.711	0.577	0.134	0.066	2.031**
Sale of fattened sheep (%)	0.937	0.741	0.196	0.065	3.020***
Sale of fruit seedling (%)	0.832	0.451	0.381	0.093	4.10***

*** and **Significant at 1% and 5% probability levels

1. The bootstrapped SE is obtained after 100 replications

4.3.5.4. Impact estimate on market orientation of the household

As Table 17 shows, the effect of the program on the proportion of households' allocation of land for coffee commodities do not yield significant difference between treated and comparison households. Stated in other words, using land allocation to coffee commodities as indicator of market orientation, our impact estimate does not show significant difference between two groups in their market orientation of coffee production. This result corroborates with the findings of Samuel and Ludi (2008), who found that the proportion of land allocated for coffee did not show significant difference between market participants and non-participants in Gomma district.

Similarly, our impact estimate for the proportion households possessing improved hives does not yield an effect significantly different from zero. On the other hand, our result of measuring the market orientation for sheep fattening commodity shows that the project has larger and significant impact on number of sheep fattened.

Table 17. ATT for market orientation indicators for the commodities

	Treated	Control	Difference	S.E[1]	t-value
Land allocated for coffee (%)	0.680	0.610	0.070	0.056	1.25
Possessions of improved hives (%)	0.361	0.302	0.059	0.043	1.370
Sheep allocated for fattening(no.)	5.151	2.958	2.193	0.717	3.059***

1. The bootstrapped SE is obtained after 100 replications

*** Significant at 1% probability level

In order to control for unobservable biases, Table 18 below shows the result of sensitivity of market development effects on different outcome variables.

Table 18. Result of sensitivity analysis using Rosenbaum bounding approach

No.	Outcomes	$e^{\gamma} = e^{\gamma} = 1$	e^{γ} =1.25	e^{γ} =1.5	e^{γ} =1.75	e^{γ} =2	e^{γ} =2.25	e^{γ} =2.5	e^{γ} =2.75	e^{γ} =3
1	THHIOC	P<0.000	P<0.000	0.000	0.000	0.000	1.10E-16	3.80E-15	6.20E-14	6.40E-13
2	LBUSECFE	P<0.000	P<0.000	0.000	0.000	0.000	1.10E-16	4.70E-15	7.50E-14	7.60E-13
3	TNSHPF	P<2.40E-15	P<1.20E-12	8.10E-11	1.60E-09	1.50E-08	8.90E-08	3.70E-07	1.20E-06	3.10E-06
4	PROFSHPS	P<4.40E-16	P<3.30E-13	2.70E-11	6.30E-10	6.70E-09	4.20E-08	1.90E-07	6.30E-07	1.80E-06
5	PROHNEYS	P<8.20E-14	P<2.10E-11	8.60E-10	1.20E-08	9.20E-08	4.40E-07	1.50E-06	4.30E-06	0.00001
6	QTCOTTML	P<1.70E-14	P<6.00E-12	3.00E-10	5.00E-09	4.10E-08	2.20E-07	8.10E-07	2.40E-06	6.00E-06
7	NOMAZ	P<0.000	P<0.000	0.000	0.000	0.000	1.10E-16	4.90E-15	7.80E-14	7.90E-13
8	NOBEDS	P<0.000	P<0.000	0.000	3.10E-15	1.50E-13	3.00E-12	3.40E-11	2.40E-10	1.30E-09
9	PROCOFES	P<0.000	P<0.000	0.000	0.000	2.20E-16	8.30E-15	1.70E-13	1.90E-12	1.50E-11
10	PROFRUS	P<0.000	P<0.000	0.000	0.000	0.000	2.20E-16	5.60E-15	8.80E-14	8.80E-13

Source: Own estimation

e^{γ} (Gamma)=log odds of differential due to unobserved factors where Wilcoxon significance level for each significant outcome variable is calculated

Table 18 presents the critical level of e^{γ} (first row), at which the causal inference of significant market development effect has to be questioned. As noted by Hujer *et al.* (2004), sensitivity analysis for insignificant effects is not meaningful and is therefore not considered here. Given that the estimated market development effect is positive for the significant outcomes, the lower bounds under the assumption that the true treatment effect has been underestimated were less interesting (Becker and Caliendo, 2007) and therefore not reported in this study. Rosenbaum bounds were calculated for market development effects that are positive and significantly different from zero. The first column of the table shows those outcome variables which bears statistical difference between treated and control households in our impact estimate above. The rest of the values which corresponds to each row of the significant outcome variables are p-critical values (or the upper bound of Wilcoxon significance level -Sig^{+}) at different critical value of e^{γ}.

Result show that the inference for the effect of the market development interventions is not changing though the participants and non participant households has been allowed to differ in their odds of being treated up to 200% ($e^{\gamma} = 3$**)** in terms of unobserved covariates**.** That means for all outcome variables estimated, at various level of critical value of e^{γ} , the p- critical values are significant which further indicate that we have considered important covariates that affected both participation and outcome variables. We couldn't get the critical value e^{γ} where the estimated ATT is questioned even if we have set e^{γ} largely up to 3, which is larger value compared to the value set in different literatures which is usually 2 (100%).Thus, we can conclude that our impact estimates (ATT) are insensitive to unobserved selection bias and are a pure effect of market development interventions by IPMS project.

5. CONCLUSIONS AND RECOMMEDATIONS

5.1. Conclusions

In this study the impact of input and output market development interventions by IPMS project in Gomma *woreda* has been assessed using cross sectional data collected for the same purpose. The primary data for this study were collected from 200 households from both program and non-program households in Gomma *woreda* using a structured questionnaire. The main research question of the study was what would have happened to an outcome of interest had the program not been in place. Answering this question requires observing outcomes with-and-without the program for the same household. However, it is impossible to observe the same object in two states simultaneously. While the program evaluator observes the factual for an object, it is impossible to observe the counter-factual for the same object.

In a randomized experimental design, the impact of a program can be estimated by a simple difference in means between treatment and control outcomes. However, in non experimental design, since the program placement creates selection effect, simple with-and-without comparison of means for program and non-program households would make the biased estimates.

Hence, the study has applied a propensity score matching technique which has become the most widely applied non-experimental tool for impact evaluation of social programs. It is used to extract comparable pair of treatment-comparison households in a non-random program setup and in the absence of baseline data. Moreover, it can adjust for (but not totally solve the problem of) selection bias and in estimating the counterfactual effects.

As expected, participation in the program was determined by a combination of factors. Program participation is significantly influenced by four explanatory variables. The variables distance to market, distance to office of DA, size of owned land and family size are the significant variables which affect the participation of the household in the program. Households nearer to market and office of the DAs are more likely to be included in the

program than household living far from the market and office. Similarly, households who have large family size are less likely to participate in the market development interventions developed by IPMS project than those who do have small family size. By contrast, size of owned land has a strong and positive effect on household program participation.

Finding a reliable estimate of the project impact necessitates controlling for all such confounding factors adequately. In doing so, propensity score matching has resulted in 90 participant households to be matched with 87 non-participant households after discarding households whose values were out of common support region. In other words, matched comparisons of different outcome of interest were performed on these households who shared similar pre-intervention characteristics except the program participation effect. The resulting matches passed on many process of matching quality tests such as t-test, reduction in standard bias and chi-square test. Moreover, the computed parametric standard error was bootstrapped in order to capture all sources of errors in the estimates and finally sensitivity analysis was made.

The impact estimation results then indicate that there are significant differences in market development outcomes between treatment and comparison households, which could be attributable to the participation in input and output market interventions. The effect of the program on total household net income for the commodities of intervention is higher for the participant households which are statistically significant except for coffee commodity. Moreover, the program effect on intensity of input use for most of the commodities of interventions is higher for treated households which are measured using different indicators. However, the project has no impact on input use intensity for apiculture and fruits commodities. Similarly, our estimate also reveals that the productivity of improved beehives is not significant between participant and non participant households.

The impact of the project on the proportion of produce sold to the market is significant and robust for the commodities considered. Treated households sold significantly large proportion of coffee, fattened sheep, honey and fruit seedlings compared to the comparison ones.

The project impact on market orientation behavior of the household also indicates that the treated households who participated in sheep fattening commodities were more market oriented than non fattening comparison households that was measured using number of sheep meant for fattening. But in contrast to this, our estimate result shows that there is no significant difference between treated and control households' land allocation and improved hives possessions for coffee and apiculture commodities respectively.

Due to the market interventions of the project, various changes in institutional and organizational aspect of market have been observed in the district that include establishing/strengthening cooperatives primarily meant to access better market, better linkage with potential traders, business orientation of producing commodities, innovative credit provision meant for input and output marketing and community based fruit seedling multiplication and supply/marketing system, establishment of community based insurance system were changes observed . There is relatively better synergy and linkage of various stakeholders found along the value chain in discharging their specific roles and responsibilities. Moreover, private sector involvement, including the agro industries and exporters, in input supply and output marketing become an important change observed which could substitute the public sector role in the long run.

The result of Rosenbaum bounding procedure to check the hidden bias due to unobservable selection shows that all estimated ATTs for all significant outcome variables are insensitive which clearly indicate its robustness.

5.2. Recommendations

Based on the empirical findings reported in this thesis, the following recommendations are forwarded

Though the project has facilitated quality coffee sale on auction base, bidders were very few in number which in turn affects in setting competitive premium price for coffee. During the informal survey farmers were complaining to the price the bidders (exporters) set per

kilogram was not as they expected and was that not fair price compared to the cost they incurred. The linkage made to the coffee exporters to purchase sundried coffee from the producers was not as such strong too. Thus, there is a need for interventions in the coffee market towards managing high price fluctuations and developing institutional mechanisms that can help coffee growers to better deal with market risks. In this regard, strengthening the linkage between cooperatives and Oromiya Coffee Producer Union is very important.

Specialized market requires investing in certification program. In this market, consumers are willing to pay premium price for specialty coffees, organic and environmentally friendly coffee. However, obtaining certification is not easy for poor/individual farmers to meet the cost. Thus, helping the primary cooperatives in obtaining certification for the special sun dried coffee is very important as it has rewarding impacts in the long-term for producers.

Apiculture accessories such as honey extractor are becoming very expensive and the majority of smallholder farmers might not afford to buy and use from the private traders. Thus, capacitating (in financial and business management) the established honey producer cooperatives in supplying these accessories and other inputs to members is of paramount important which further enhances the production and productivity as well as the quality of honey.

Farmers who have taken credit from OCSSCO meant for innovative market interventions were complaining about large group size as peer collateral and short repayment period. Revising and searching for appropriate group size and loan payment period could help the beneficiaries in maximizing the return from loans.

The impact of the participation in the market development interventions might not be homogenous among participating households. Identifying the factors which contribute to this difference was not covered in this study. Thus, working on identifying the important factors for the variation needs further research. Moreover, evaluating the overall project performance by incorporating other components like environment and gender mainstreaming by having

relatively large sample size and coverage, as matching is a “data hungry” estimation strategy, is the other research gap that ought to be addressed in the future.

6. REFERENCES

Abadie, A. and Imbens, G.W., 2006. Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica*, 74:235- 67.

Ana, R.,William R.,Masters A. and Shively,G.E., 2008. Linkages between market participation and productivity: results from a multi-country farm household sample**.** Prepared for presentation at the American Agricultural Economics Association Annual Meeting, Orlando, FL, July 27-29, 2008

Andersson, C., Alemu Mekonen and Stage S., 2009. Impacts of the productive safety net program in Ethiopia on livestock and tree holdings of rural households. Discussion Paper Series. Environment for Development (EfD) Initiative. EfD DP 09-05.

Ansoff, I. H., 1957. Strategies for diversification, *Harvard Business Review*, 35(2):113-124.

Baker, J.L., 2000. Evaluating the impact of development projects on poverty: A handbook for Practitioners. Washington D.C. World Bank.

Becker, S. O. and Ichino, A., 2002. Estimation of average treatment effects based on Propensity Scores: *The Stata Journal*, 2(4): 1-19.

Becker, S.O. and M. Caliendo, 2007. Sensitivity Analysis for Average Treatment Effects. *Stata Journal*, 7(1): 71–83.

Bellemre, M.F. and Barrett, C.B., 2006. An Ordered Tobit Model of Market Participation: Evidence from Kenya and Ethiopia. *American Journal of Agricultural Economics* 88(2): 324-337.

Berhanu Gebremedhin, Dirk, H. andAzage Tegegne, 2006. Commercialization of Ethiopian agriculture: Extension service from input supplier to knowledge broker and facilitator. IPMS (Improving Productivity and Market Success) of Ethiopian Farmers Project Working Paper 1. ILRI (International Livestock Research Institute), Nairobi, Kenya. 36 pp.

Berhanu Gebremedhin and Dirk,H., 2008. Market orientation of smallholders in selected grains in Ethiopia: Implications for enhancing commercial transformation of subsistence agriculture. IPMS (Improving Productivity and Market Success) of Ethiopian Farmers Project Working Paper 11. ILRI (International Livestock Research Institute), Nairobi, Kenya. 44 pp.

Berhanu Gebremedhin and Dirk, H., 2009. Staple food crops turning into commercial crops: Case studies of Teff, Wheat and Rice in Ethiopia. Revision submitted to the Ethiopian Journal of Economics. International Livestock Research Institute (ILRI). Addis Ababa, Ethiopia.

Bernard,T.,Spileman,D.J., Alemayehu Seyoum and Eleni Gabre-Madhin, 2010. Cooperatives for staple crop marketing: evidence from Ethiopia. IFPRI research monograph; 164.

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

Caliendo, M., and Kopeinig, S., 2008. Some practical guidance for the implementation of propensity score matching. IZA Discussion Paper No. 1588, University of Cologne.

Chowdhury ,S., Asfaw Negessa and Torero M., 2005. Market institutions: Enhancing the value of rural-urban links. FCND Discussion Paper 195/MTID Discussion Paper 89. Washington DC: International Food Policy Research Institute.

Coulter, J., Millns, J., Tallontire, A .and Stringfellow R., 1999. Marrying farmer cooperation and contract farming for agricultural service provision in liberalizing economies in sub-Saharan Africa. ODI Natural Resources Perspectives, No. 48, London: Overseas Development Institute.

CSA(Centeral Statistics Agency),2007. Summary and statistical report of population and housing census,Addis Ababa,Ethiopia

Dawit Alemu, Eleni Gabre-Madhin and Samson Dejene, 2006. From farmer to market and market to farmer: Characterizing smallholder commercialization in Ethiopia. Paper submitted for ESSP Policy Conference on Bridging, Balancing, and Scaling up: Advancing the rural growth agenda in Ethiopia. Addis Ababa, Ethiopia.

Dehejia, R. H. and Wahba, S., 2002. Propensity score matching methods for non-experimental causal studies. *The Review of Economics and Statistics,* 84(1):151-161.

de Janvry, A., Fafchamps, M. and Sadoulet, E., 1991. Peasant household behavior with missing markets: Some paradoxes explained. *Economic Journal,* 101:1400-1417.

Dercon, S. and Zeitlin, A., 2009. Rethinking agriculture and growth in Ethiopia: A conceptual discussion. Paper prepared as part of a study on agriculture and growth in Ethiopia.

Dorward, A., Poole, N., Morrison J., Kydd, J. and Uray, I., 2003. Markets, institutions and technology: Missing links in livelihoods analysis. *Development Policy Review,* 21(3): 319-332.

Dorward, A. and Kydd, J., 2004. The Malawi 2002 food crisis: The rural development challenge. *Journal of Modern African Studies,* 42(3): 343-361.

Dorward, A., Kydd, J., Morrison, J. and Poulton, C., 2005. Institutions, markets and economic development: Linking development policy to theory and praxis. *Development and Change,* 36(1): 1-25.

Ezemenari, K., Rudqvist, A. and Subbarao, K., 1999. Impact evaluation: A note on concepts and methods, poverty reduction and economic management network, World Bank.

Fafchamps, M., 2005. Market institutions in Sub-Saharan Africa. Cambridge, Mass, U.S.A.: MIT Press.

FAO (Food Agriculture Organization), 2000. Impact assessment of agricultural research: Context and state of the art. Revised version of a paper prepared by the Impact Assessment and Evaluation Group (IAEG) of the Consultative Group on International Agricultural Research (CGIAR) for the ASARECA / ECART / CTA. Workshop on impact assessment of agricultural research in Eastern and Central Africa. Rome, Italy.

Fitsum Hagos, Gideon Kruseman, Zenebe Abreha, Vincent L., Afeworki Mulugeta and Girmay G/Samuele, 2006. Impact of small scale water harvesting on household poverty: evidence from Northern Ethiopia. Poverty Reduction and Environmental Management (PREM), 07/01.

Foster, E., 2003. Propensity score matching. An illustrative analysis of dose response. Lippincott Williams & Wilkins, Inc. *Journal of Medical Care*, 41(10):1183-1192.

Eleni Gabre-Madhin, 2001. Market institutions, transaction costs, and social capital in the Ethiopian grain market. IFPRI Research Report 124. Washington DC: International Food Policy Research Institute.

Eleni Gabre-Madhin, Wolday A., Tafara, E., Schluter, J., Teshome, T. and Kilkile, G., 2003. Getting markets right in Ethiopia: An institutional and legal analysis of grain and coffee marketing. Final report. IFPRI (International Food Policy Research Institute), Washington, DC, USA.

Eleni Gabre-Madhin, and Goggin,I., 2006. Does Ethiopia need a commodity exchange? An integrated approach to market development. Paper presented at the Ethiopian Strategy Support Program 2006 Policy Conference, Addis Ababa.

Gilligan, D.,Hoddinott,J. and Alemayehu Seyoum,2008. The Impact of Ethiopia's Productive Safety Net Programme and its linkages. IFPRI Discussion Paper 00839.

Goetz, S.J., 1992. A Selectivity model of household food marketing behavior in Sub-Saharan Africa. *American Journal of Agricultural Economics* 74(2): 444-452.

Govereh, J., Jayne, T.S. and Nyoro, J., 1999. Smallholder commercialization, interlinked markets and food crop productivity: Cross-country evidence in eastern and southern Africa

Gujarati, D.N., 1999. Essentials of Econometrics. 2nd edition. Mc-Graw-Hill Companies

Gujarati, D.N., 2004. Essentials of Econometrics. 4th edition. Mc-Graw-Hill Companies

Heckman, J.J., Ichimura, H. and Todd, P., 1997. Matching as an econometric evaluation estimator: evidence from evaluating a job training programme, *The Review of Economic Studies*, 64:605-654.

Heckman, J., H. Ichimura, J. Smith, and Todd, P., 1998. Characterizing selection bias using experimental data. *Econometrica,* 66: 1017–1098.

Holloway, G., Barrett, C.B. and Ehui, S., 2005. The Double-Hurdle model in the presence of fixed costs. *Journal of International Agricultural Trade and Development*, 1: 17-28.

Hujer, R., Caliendo, M., Thomson, S.L., 2004. New evidence on the effects of job creation schemes in Germany. A matching approach with threefold heterogeneity. *Res. Econ*. 58(4), 257–302.

IPMS (Improving Productivity and Market Success) of Ethiopian Farmers, 2005. Project implementation plan. www.ipms-ethiopia.org.

IPMS (Improving Productivity and Market Success) of Ethiopian Farmers, 2007. Gomma Pilot Learning Woreda. Diagnosis and program design. www.ipms-ethiopia.org.

Jalan, J. and Ravallion, M., 1999. Income gains to the poor from workfare:Estimates for Argentina's Trabajar Program. Indian Statistical Institute and World Bank. Policy Research Working Paper 2149. Washington, DC.

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

Jayne, T., Zulu, B., and Nijhoff, J., 2006. Stabilizing food markets in Eastern and Southern Africa. *Food Policy* 31(4): 328–341.

Keister, L.A. and V.G. Nee, 2001. The rational peasant in China. Flexible adaptation, risk diversification and opportunity. Rationality and Society. 13 (1).

Key, N. E. Sadoulet, and A. DeJanvry, 2000. Transactions costs and agricultural household supply response. *American Journal of Agricultural Economics*, 82(2): 245-259.

Kranton, R.E., 1996. Reciprocal exchange: A self-sustaining system. *American Economic Review*, 86(4): 830-851.

Lechner, M., 2002. Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies. *The Review of Economics and Statistics*, 84(2): 205-220.

Leuven, E. and Sianesi, B., 2003. PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing and covariate imbalance testing. http://ideas.repec.org/c/boc/bocode/s432001.html.Accessed on April, 2010

Liebenehm, S., Affognon, H. and Waibel, H., 2009. Impact assessment of agricultural research in West Africa: An application of the propensity score matching methodology. Paper

presented at the International Association of Agricultural Economics Conference, Beijing, China.

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

MoFED (Ministry of Finance and Economic Development), 2006. Building on progress: A Plan for Accelerated and Sustained Development to End Poverty (PASDEP). Addis Ababa, Ethiopia.

MoPED (Ministry Planning and Economic Development), 1993. An economic development strategy for Ethiopia: A comprehensive guidance and development strategy for the future. Addis Ababa, Ethiopia.

Moti Jaleta, Berhanu Gebremedhin and Dirk,H., 2009. Smallholder commercialization: Processes, determinants and impact. Discussion Paper No. 18. Improving Productivity and Market Success (IPMS) of Ethiopian Farmers Project, ILRI (International Livestock Research Institute), Nairobi, Kenya. 55 pp.

Mwape, F., 2009. Paper presented at the opening speech of Towards Priority Actions for Market Development for African Farmers, ILRI Conference, Nairobi, Kenya

North, D.C., 1990. Institutions, institutional change and economic performance. Cambridge: Cambridge University Press.

North, D.C., 1996. Epilogue. Economic performance through time.In: L.J.Aston,T.Eggersson, and D.C. North(eds.). Empirical studies on institutional change. Cambridge. Cambridge University press

Nssah, B., 2006. Propensity score matching and policy impact analysis. Poverty reduction group.World Bank.Wasihngton,D.C.

Omoto, W., 2003. Impact assessment of urban agriculture research and development in Nairobi, Kenya.

Oechmke, J.F., Anandajayasekeram, P., Masters, W.A., 1997. Agricultural Technology Development and Transfer in Africa Impacts Achieved and Lessons Learned. USAID, Office of Sustainable Development, Bureau for Africa, SD Publication Series, Technical Paper No. 77, Washington, D. C.

Pingali, P. and Rosegrant, M., 1995. Agricultural commercialization and diversification: Processes and policies. *Food Policy*, 20(3): 171–185.

Pingali, P., 1997. From subsistence to commercial production systems: The transformation of Asian agriculture. *American Journal of Agricultural Economics*, 79(2): 628–634.

Ponniah, A., Martella, D. R. and Rukuni, M., 1996. A training manual on R & D evaluation and impact assessment of investments in agricultural and natural resources research, SACCAR, Gaborone, Botswana.

Ponniah, A. and Martella, D. R., 1999. Evaluation of agricultural research in Eastern, Central, and Southern Africa. *Knowledge, Technology and Policy: Winter,* 11(4):13-41.

Puskur, R. and J. Hagmann, 2006. Synthesis report. Workshop on alternative service delivery systems, Unpublished Report, IPMS, Addis Ababa

Puskur, R., Ponniah,P., Kahsay Berhe,Hoekstra,D.,2007. Partnerships for enhancing market-led Innovation processes – Experiences and lessons from IPMS Ethiopia. Paper presented at the Innovation Africa Symposium, Kampala

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

Reardon, T. and C.P. Timmer, 2005. Transformation of Markets for Agricultural Output in Developing Countries Since 1950: How Has Thinking Changed? In R.E. Evenson, P. Pingali and T.P. Schultz (Eds). Handbook of Agricultural Economics Volume 3. Agricultural Development: Farmers, Farm Production and Farm Markets. Amsterdam: Elsevier.

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

Rosenbaum, P. R. and Rubin, D. B., 1985.Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician,* 39(1): 35-39

Rosenbaum, P. R., 2002. Observational Studies. 2nd ed. New York: Springer

Rover, T. and Dixon, R., 2007. Operational guidelines for assessing the impact of agricultural research on livelihoods. Good practices from CIMMYT. Impacts Targeting and Assessment (ITA) Unit. Mexico.

Samuel Gebreselassie and Ludi ,E., 2008. Agricultural commercialization in coffee growing areas of Ethiopia. A paper for the Future Agricultures Consortium, Institute of Development Studies.www.future-agricultures.org.

Sianesi, B., 2004. An evaluation of the active labor market programs in Sweden. *The Review of Economics and Statistics*, 186(1):133-155.

Strasberg, P.J., T.S. Jayne, T. Yamano, J., Nyoro, D. Karanja and J. Strauss, 1999. Effects of agricultural commercialization on food cop input use and productivity in Kenya. MSU International Department of Agricultural Economics Development, Department of Economics. Working paper no. 71.

Storck, H., BezabihEmana, BerhanuAdnew, Borowiecki, A.AandShimelisW/Hawaryat, 1991.Farming system and farm management practices of smallholders in the Hararghe highlands. Faming system and resource Economics in the Topics.11: Wissenschfts Varlag Vauk Kiel KG, Germany.

Timmer, C. P., 1997. Farmers and markets: The political economy of new paradigms. *American Journal of Agricultural Economics*, 79(2):621–627.

Von Braun, J., Bouis, H. and Kennedy, E., 1994. Conceptual framework. In: Von Braun J and Kennedy E. (eds.). Agricultural commercialization, economic development and nutrition. John Hopkins University Press, Baltimore, USA

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48: 817-838.

Wolday Amha, 1994. Food grain marketing development in Ethiopia after the market reform 1990. A case study of Alaba Siraro district. PhD.Dissertation. Veriog Koster. Berlin,Germany

World Bank. 2002. World development report 2002: Building institutions for markets. New York: Oxford University Press.

World Bank, 2007. Ethiopia: Accelerating equitable growth: Country Economic Memorandum, Part II: Thematic chapters. ReportN0.38662-ET.Washington, DC: World Bank. World Bank/Poverty reduction and Economic management Unit, Africa Region.

Yibeltal Fentie, 2008. The Impact of Ibnat-Belessa integrated food security program on household food poverty. An M.Sc. Thesis Presented to School of Graduate Studies of Haramaya University. 94p.

7. APPENDICES

Appendix 1. Multicollinearity test for continuous explanatory variables

Covariates	VIF
AGEHH	3.332
TOTALFMZ	1.792
EXPFRMG	3.051
SZEONLD	1.276
TLU	1.334
DISNMKT	1.096
DISEXTO	1.088
DPCRTO	1.458

Appendix 2. Contingency coefficient for discrete variables

Variable	Value of C
SEXHH	0.104
EDULHH	0.133

Appendix 3. Conversion factor used to calculate TLU

Livestock Category	TLU
Calf	0.34
Heifer	0.75
Cow and Ox	1.0
Horse	1.1
Donkey	0.7
Sheep and goat(adult)	0.13
Chicken	0.013

Source: Storck *et al.*, 1991

Appendix 4. Labor supply conversion factor (person day equivalent)

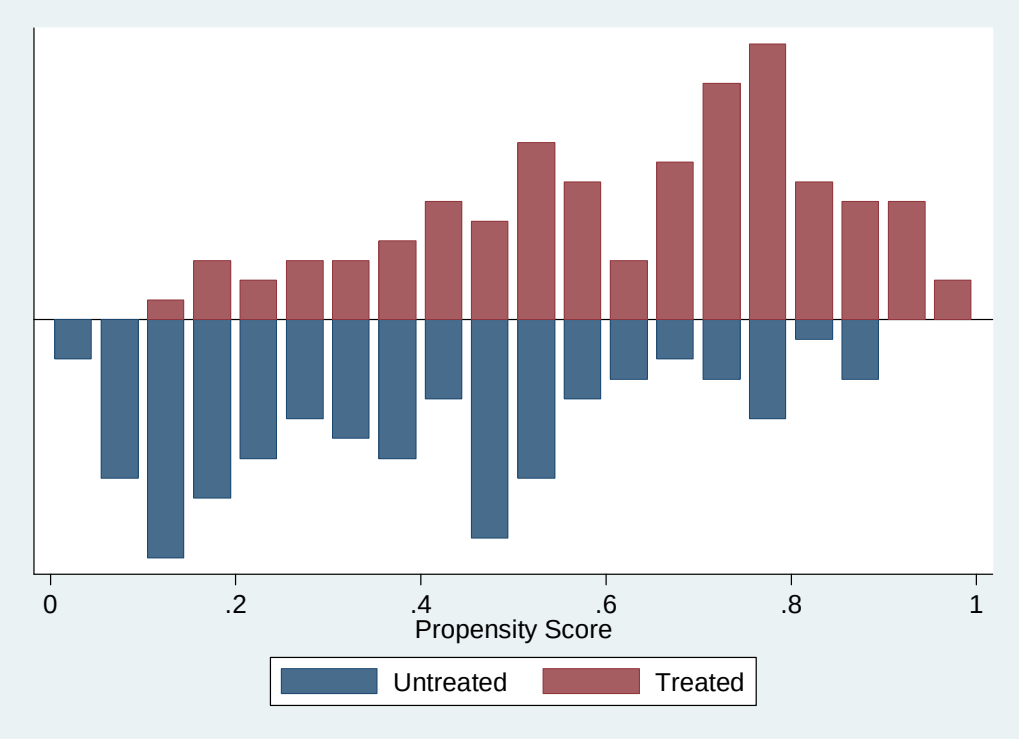
Age group in years	Male	Female
<10	0.0	0.0
10-13	0.35	0.35
15-50	1.0	0.80
>50	0.55	0.5

Source: Storck *et al*., 1991

Appendix 5. Joint significance test (likelihood ratio test)

Matching algorithms	Sample	Pseudo R^2	LRchi^2	P>chi^2
	Unmatched	0.207	57.62	0.000
NN(1)	Matched	0.182	44.69	0.000
	Unmatched	0.207	57.62	0.000
NN(2)	Matched	0.142	34.77	0.000
	Unmatched	0.207	57.62	0.000
NN(3)	Matched	0.132	32.27	0.001
	Unmatched	0.207	57.62	0.000
NN(4)	Matched	0.134	32.79	0.001
	Unmatched	0.207	57.62	0.000
NN(5)	Matched	0.128	31.33	0.001
	Unmatched	0.207	57.62	0.000
Caliper(0.01)	Matched	0.096	20.00	0.045
	Unmatched	0.207	57.62	0.000
Caliper(0.25)	Matched	0.084	22.33	.0152
	Unmatched	0.207	57.62	0.000
Caliper(0.5)	Matched	0.048	11.68	0.389
	Unmatched	0.207	57.62	0.000
Kernel(0.1)	Matched	0.112	27.46	0.004
	Unmatched	0.207	57.62	0.000
Kernel(0.25)	Matched	0.008	1.90	0.999
	Unmatched	0.207	57.62	0.000
Kernel(0.5)	Matched	0.0006	1.43	1.000

Appendix 6. Histogram of propensity score



Appendix 7. Histogram of Pscore with common (off) support regions