

Doctoral Thesis



University of Trento

School of Social Sciences

**Agricultural Advisory Services in Eastern Ethiopia:
Access, Impact, and Willingness to Pay**

By

Muluken Gezahegn Wordofa

A dissertation submitted to the Doctoral School of Local Development and Global
Dynamics in partial fulfillment of the requirements for the Degree of Doctor of
Philosophy (PhD) in *Local Development and Global Dynamics*

January 22, 2015

Trento, Italy

Advisor

Prof. Roberta Raffaelli

Department of Economics and Management, University of Trento

Co-advisor

Prof. Sandra Notaro

Department of Economics and Management, University of Trento

Doctoral Thesis Examination Committee

Prof. Maria Sassi

Department of Economics and Management, University of Pavia, Italy

Prof. Gabriella Berloffia

Department of Economics and Management, University of Trento, Italy

Prof. Geremia Gios

Dipartimento di Economia e Management, Università degli Studi di Trento

Prof. Donato Romano

Dipartimento di Scienze per l'Economia e l'Impresa, Università degli Studi di Firenze

Dott.ssa Elena Pisani

Dipartimento Territorio e Sistemi Agroforestali, Università degli Studi di Padova

Table of Contents

Acronyms and Abbreviations	xii
Acknowledgments.....	xvi
Summary	xix
1 Predictors of Smallholder Farmers’ Intensity of Participation in On-farm Agricultural Advisory Service Provision in Haramaya District, Eastern Ethiopia †	1
1.1 Introduction	1
1.2 Literature Review	5
1.2.1 Smallholder Farmers’ Access to AASs.....	5
<i>i. Smallholder Farmers’ Participation in AASs in Ethiopia</i>	8
1.3 Research Methodology.....	15
1.3.1 Sampling Procedures and Data	15
1.3.2 Empirical Strategy for Data Analysis	17
<i>i. Inspecting the Distribution of the Outcome Variables and Model Fit</i>	20
<i>ii. Selection of Explanatory Variables</i>	28
1.4 Result and Discussion	30
1.4.1 Descriptive Statistics of Sociodemographic Characteristics of Respondents.....	30
1.4.2 Poisson Regression and Negative Binomial Regression Estimation Results	34
<i>i. Extent of Involvement in On-farm Training in FTC and non-FTC Areas</i>	35

ii. <i>A Comparative Discussion on the Determinants of On-farm Training in FTC and non-FTC Areas</i>	39
iii. <i>Extent of Involvement in On-farm Demonstration in FTC and non-FTC Areas</i>	46
iv. <i>Comparing Intensity of Access to On-farm Demonstration in FTC and non-FTC Areas</i>	47
1.5 Conclusion and Recommendation.....	53
2 Improving Smallholder Farmers’ Income through Farmers’ Training Centers: an Impact Evaluation in Haramaya District, Eastern Ethiopia ^{*†}	56
2.1 Introduction	56
2.2 Data and Methods.....	63
2.2.1 Study Area and <i>Kebele</i> /PA Selection	63
2.2.2 Household Head Selection in Treatment and Comparison Areas.....	67
2.2.3 Data Collection	68
2.2.4 Empirical Strategy and Data Analysis	70
i. <i>Matching Methods in Evaluating Program/Treatment Effects</i>	74
ii. <i>Assumptions in Implementing Propensity Score Matching</i>	78
iii. <i>Matching Quality Analysis (Balance Checking)</i>	80
iv. <i>Choice of Variables for Estimating Propensity Score</i>	82
2.3 Results	87
2.3.1 Descriptive Statistics of Baseline Characteristics Determining Participation in Modular Training at Farmers’ Training Centers	87

2.3.2	Propensity Score Estimation	88
2.3.3	Construction of Outcome Variable for Impact Evaluation	90
	<i>i. Income from Crop Production</i>	92
	<i>ii. Income from Livestock Production</i>	96
2.3.4	Estimation of the Average Effect of Modular Training on Farm Income and Other Outcome Variables	100
2.3.5	Matching Quality (Balance Checking) and Common Support	102
2.4	Discussion	106
2.4.1	Determinants of Participation in FTC-Based Modular Training	106
2.4.2	Impact of Modular Training on Farm Income of Rural Households	107
2.5	Conclusion and Recommendation.....	111
3	Smallholder Farmers' Willingness to Pay for Agricultural Advisory Services: Investigating the Effect of Solemn Oath in Mitigating Hypothetical Bias in Best-Worst Discrete Choice Experiments in Haramaya District, Eastern Ethiopia	115
3.1	Introduction	115
3.2	Literature Review	120
3.2.1	The Stated Preference Techniques	120
3.2.2	Application of Choice Experiments in Developing Countries	121
3.2.3	Best-Worst Approach to Preference Elicitation.....	122
3.2.4	The Status Quo Effect and the Use of Hypothetical Baselines in Choice Experiments	124

3.2.5	Hypothetical Bias and Strategies to Circumvent Hypothetical Bias	125
	<i>i. The Cheap Talk Script</i>	127
	<i>ii. The Solemn Oath Script.....</i>	128
3.3	Research Methodology.....	134
3.3.1	Sampling Procedures	134
3.3.2	Survey Design and Methods	135
	<i>i. Identifying Relevant Attributes</i>	135
	<i>ii. Assigning Attribute Levels.....</i>	138
	<i>iii. Describing Scenarios and Choice Cards</i>	141
	<i>iv. Experimental Design and Designing Choice Sets</i>	143
	<i>v. Survey Administration.....</i>	145
3.4	Results and Discussion.....	151
3.4.1	Descriptive Statistics of Sociodemographic Characteristics of Respondents.....	151
3.4.2	Parameter Estimates in the Full Sample	154
3.4.3	Exploring Heterogeneity around Mean Parameter Estimates	159
3.4.4	Parameter Estimates in the Sub-Sample	164
3.4.5	Marginal Willingness to Pay (WTP) Estimates	168
	<i>i. Changes in the WTP in Relation to Socio-demographic Characteristics ('Full Sample')</i>	172
	<i>ii. Changes in the WTP in Relation to Socio-demographic Characteristics ('Sub-Sample')</i>	
	176

3.5 Conclusion and Recommendation.....	180
References.....	185
Appendices.....	199

List of Tables, Figures, and Appendices

List of Tables

Table 1.1 Some Previous Studies on Factors Affecting Access to AASs in Developing Countries	7
Table 1.2 Distribution of On-farm and FTC-based Advice, Training, and Demonstration in the FTC Areas and Non-FTC Areas in 2013	22
Table 1.3 Descriptive Statistics of Sociodemographic Characteristics of Respondents in 2013..	32
Table 1.4 Poisson and Negative Binomial Regression Results (2013).....	37
Table 1.5 Overall Significance of Categorical Predictor Variables.....	40
Table 2.1 Sample Household Selection in Treatment and Comparison Areas	68
Table 2.2 Baseline Information on Selection Criteria by Study Kebele/PA (Measured in 2009)	87
Table 2.3 Participation in Modular Training based on Baseline Covariates Measured in 2009 ..	89
Table 2.4 Average Value by Household of Crop Produced, Sold, Consumed, and Stored, Including the Value of Crop Residues Produced in the Last 12 Months, 2013.....	96
Table 2.5 Average Value by Household of Livestock Owned, Products Sold, Oxen Rented, and Intermediate Expenses Incurred in the Last 12 Months, 2013.....	97
Table 2.6 Contribution of Different Crops to Gross Crop Income in 2013 (Percentage Share)...	99
Table 2.7 Distribution of the Average Values (Birr) by Household of Outcome Variables in the Study Sample in 2013	100
Table 2.8 Estimation of Average Treatment Effect on the Treated (ATT) Using Nearest Neighbors, Kernel, and Radius Matching.....	101
Table 2.9 Matching Quality Analysis: T-Test (Before/After) between Unmatched and Matched Treated and Control Groups.....	103

Table 2.10 Matching Quality Analysis: Standardized Percentage Bias (Before/After Matching) and Mean/Median Bias as Summary Indicators of the Distribution of the Bias	104
Table 3.1 Attributes and Attribute Levels for the CEs in FTC-based Provision of AASs	141
Table 3.2 An Example of a Choice Card Used in the Collection of Prior Data	143
Table 3.3 Experimental Design.....	145
Table 3.4 Summary Statistics of Sociodemographic Characteristics of Respondents (Standard Deviations in Parenthesis).....	153
Table 3.5 MNL Model Estimation Results on the Treatment Groups in the Full Sample (Trained in 2008 & 2009)	157
Table 3.6 RPL Model Estimation Results on the Treatment Groups in the Full Sample (Trained in 2008 & 2009)	157
Table 3.7 RPL Model Estimation Results with Interactions (Full Sample – Trained in 2008 & 2009)	161
Table 3.8 MNL Model Estimation Results on the Treatment Groups in the Sub-Sample Trained in 2009	166
Table 3.9 RPL Model Estimation Results on the Treatment Groups in the Sub-Sample Trained in 2009.....	166
Table 3.10 Marginal Willingness to Pay Estimates (in Birr) for Advice, Training, and Demonstration in the Full Sample and Sub-Sample.....	169
Table 3.11 Changes in the WTP (birr) Based on Some SDCs (Full Sample – Trained in 2008 and 2009)	175
Table 3.12 Changes in the WTP (birr) Based on Some SDCs (Sub-Sample – Trained in 2009).....	178

List of Figures

Figure 1.1 Distribution of Outcome/Dependent Variables in the Non-FTC Sample	24
Figure 1.2 Graphical Inspection of Model Fit in the FTC Areas.....	26
Figure 1.3 Graphical Inspection of Model Fit in the Non-FTC Areas	27
Figure 2.1 Map of East Hararghe Zone of Oromia Region Showing the 18 Districts in the Zone and Map of Haramaya District Showing the 33 Peasant Associations (PAs)/Kebeles in the District.....	66
Figure 2.2 Distribution of Propensity Scores between Comparison (Left Histogram) and Treatment (Right Histogram) Respondents	90
Figure 2.3 Histogram of Estimated Propensity Score between Participants (Treated) and Nonparticipants (Untreated) of FTC-Based Modular Training	105

List of Appendices

Appendix 1 FGDs to Test Attributes, Attribute Levels, Payment Vehicle, and Scenario Description.....	199
Appendix 2 Scenario Description for CEs in AASs in Haramaya District, Ethiopia.....	203
Appendix 3 Illustrative Show Cards Used in the Choice Experiment Survey	208
Appendix 4 Summary Statistics of Sociodemographic Characteristics of Respondents Trained at FTCs in 2009 (Standard Deviations in Parenthesis).....	215
Appendix 5 Choice of Hypothetical Baseline as Best, Second-best, and Worst	217
Appendix 6 Changes in the WTP in Relation to Socio-demographic Characteristics ('Sub- Sample', Signed on the Solemn Oath Form)	218
Appendix 7 Focus Group Discussions (FGDs) in the CE in AASs.....	221

Appendix 8 Household Survey Questionnaire - Treatment (Oath) Group 226

Acronyms and Abbreviations

AASs	Agricultural Advisory Services
AIC	Akaike's Information Criterion
ATT	Average Treatment Effect on the Treated
ATVET	Agricultural Technical and Vocational Education and Training
BIC	Bayesian Information Criterion
BOARD	Bureau of Agriculture and Rural Development
CEs	Choice Experiments
CPPs	Comprehensive Package Programs
CT	Control Treatment
CVM	Contingent Valuation Method
CV	Contingent Valuation
DAs	Development Agents
DID	Difference-in-Difference
EPRDF	Ethiopian People's Revolutionary Democratic Front
ERHS	Ethiopian Rural Household Survey
ETB	Ethiopian Birr (Official Currency)

FFSs	Farmer Field Schools
FGDs	Focus Group Discussions
FTCs	Farmers' Training Centers
GMXLOGIT	Generalized Mixed Logit
Ha	Hectare
IECAMA	Imperial Ethiopian College of Agriculture and Mechanical Arts
IFPRI	International Food Policy Research Institute
IIA	Independence from Irrelevant Alternatives
IID	Independently and Identically Distributed
IPMS	Improving Productivity and Market Success
IV	Instrumental Variables
K-S	Kolmogorov-Smirnov
LR	Likelihood-ratio
MLE	Maximum Likelihood Estimation
MNL	Multinomial Logit
MOARD	Ministry of Agriculture and Rural Development
MPPs	Minimum Package Projects

NAADS	National Agricultural Advisory Services
NGOs	Non-governmental Organizations
NRM	Natural Resource Management
OLS	Ordinary Least Squares
OT	Oath Treatment
PADEP	Peasant Agricultural Development Program
PADETES	Participatory Demonstration and Training Extension System
PAs	Peasant Associations
PLWs	Pilot Learning Woredas
PSM	Propensity Score Matching
RCTs	Randomized Control Trials
RDD	Regression Discontinuity Design
ROL	Rank Ordered Logit
RPL	Random Parameters Logit
Sd	Standard deviation
SDCs	Socio-demographic Characteristics
SMNLOGIT	Scaled Multinomial Logit

SP	Stated Preference
SQ	Status Quo
WTP	Willingness to Pay

Acknowledgments

A number of individuals and organizations contributed – directly and/or indirectly and at different levels – to the successful accomplishment of this PhD study. It is, therefore, with due respect that I recognize their efforts and inputs.

I am also deeply indebted to my major advisor, Prof. Roberta Raffaelli, for her intellectual guidance, encouragement, and support from the very conception of the PhD research to the write up of the dissertation and discussion with the doctoral commission. In particular, her keen attention to detail while at the same time being concise; her technical support with *NLOGIT* software; her constant availability and prompt feedback; and, her overall dedication, concern, and positivity in all matters related to the study are some of her greatest contributions that made this dissertation research a reality. I have also received invaluable comments and suggestions from Prof. Sandra Notaro, co-advisor, especially on the experimental design part of the Choice Experiment chapter, some technical issues related to *NLOGIT* software, and the analysis of the choice experiment data using the software. I sincerely acknowledge her contributions in this regard.

I am also very grateful to Prof. Maria Sassi for her extraordinary advice, kindness, and friendliness throughout the dissertation research. In particular, I am very fortunate to have worked with her on the first two parts of the dissertation, where her critical comments and suggestions were very useful in improving the quality of the dissertation. She is a wonderful advisor, a patient reader, an approachable and enthusiastic person in terms of handling students, and a great person filled with encouragement and positive attitude.

I extend my deepest appreciation to the doctoral thesis examiners for the productive discussion, general comments, and chapter-specific comments, which altogether have helped to substantially improve the quality of the dissertation. I also duly acknowledge the technical inputs from Profs. Ugo Trivellato and Enrico Rettore. Special thanks are also due to Dr. Sara Cardey, Mark Beittel, and Prof. Yiorgos Alexopoulos for their support during the development of the thesis research proposal. I am very grateful to Prof. Gabriella Berloffia for her useful comments on the survey instrument and the dissertation. I also acknowledge the help I received from the ‘impact evaluation team’ at IRVAPP (Istituto per la Ricerca Valutativa sulle Politiche Pubbliche) in Trento.

My sincere thanks also go to the participants of the following conferences: Faculty of Economics and Management, University of Pavia. Pavia, Italy (26 Nov. 2013); Third Italian Association of Agricultural and Applied Economics (AIEAA) conference “Feeding the Planet and Greening Agriculture: Challenges and Opportunities for the Bio-Economy”. Porto Conte Research Center, Alghero (SS), Italy (25-27 June 2014); Società Italiana degli Economisti dello Sviluppo (SITES)/Italian Development Economists Association (IDEAs). SITES-IDEAS First Annual Conference. Florence, Italy (11-12 Sep. 2014); and, Associazione Italiana per gli Studi Africani/Italian Association for African Studies (ASAI). ASAI Third Conference “Africa on the Move” organised in partnership with the Interdisciplinary Centre for African Studies of the University of Macerata. Macerata, Italy (17-20 Sep. 2014).

I am also very thankful to the following organizations for facilitating the field work for the thesis research: Haramaya University; East Hararghe Zone Bureau of Agriculture and Rural Development; and, Haramaya District Bureau of Agriculture and Rural Development. In particular, I would like to thank Ato Yohannes and Ato Fekadu; the field supervisors Ato Nigussie and Ato Sisay Belay; the enumerators W/ro Tejitu Tesfaye, W/ro Kokobe Fentahun, Ato Efrem Solomon, Ato Anteneh Matiwos

(Abushye), Ato Sherif Mohammed, Ato Aman Adem Ali, and Ato Yasin Musa. I am also very grateful to the respondents of the study, the administrators of the PAs and of the FTCs within, and the wider community.

The generous financial support of the University of Trento (for both the dissertation research and the PhD study as a whole) is also specially acknowledged. I would also like to thank Prof. Giuseppe Folloni, Davide Santuari, and Nicole Bertotti (School of Social Sciences); Welcome Office staff; Opera Universitaria Accommodation Office staff; Province of Trento; and, Doctoral Office staff. My heartfelt thanks and appreciation also go to my wife, Yodiye, for shouldering the responsibility of raising our two daughters – Natanim and Liya – and constant encouragement and support throughout the study. Above all, I thank God, ‘the author and perfecter’ of my destiny, around which my entire life revolves and towards which it leaps forward.

Muluken G. Wordofa

Trento, Italy

Summary

In developing countries, public agricultural advisory services (AASs) have been one of the key components of local development in terms of enhancing production and productivity, achieving food security, contributing to poverty reduction, and improving livelihoods in the face of economic, social, and environmental challenges. However, the benefit that farmers obtain out of participating in AASs and the resulting impact depend, to a great extent, by the (intensity of) farmers' direct and indirect access to these services. The issue of intensity of farmers' involvement in AASs is also especially important considering the various stages involved in farming – land preparation to post-harvest management. In this regard, the extent of farmers' participation in such services and the determinants of their intensity of participation are not very well documented in the context of many developing countries. On the other hand, recent studies have shown that the impact of AASs is mixed and that some empirical investigations lack methodological rigor. Furthermore, there is a dearth of context-specific empirical evidence on the impact of new approaches to AAS provision. One such approach is the Farmers' Training Centers (FTCs) implemented by the government of Ethiopia to improve smallholder farming systems.

Although public AASs have been hailed to be a key constituent of agricultural development and poverty reduction strategies in many developing countries, due to declining public budget allocation, alternative ways of providing AASs have been sought and implemented. One way of diversifying AASs is through a fee-for-service arrangement, which proved useful in many developed countries. In the context of developing countries, however, the potential for payment for AASs has yet to be explored in full. This is especially relevant considering the current

economic crisis, large number of farmers relying on agriculture for their livelihoods, inefficiencies of public AASs, and specialized needs of market-oriented producers.

In order to address these issues that relate to intensity of farmers' access to on-farm AASs, impact of FTC-based AASs on outcome indicators, as well as willingness to pay for improved FTC-based AASs, this study is carried out in Haramaya district of eastern Ethiopia. More specifically, a field-based household survey was conducted from May to October 2013 on a total sample of 340 farm households. In relation to the first objective of the investigation, i.e., analyzing the predictors of farmers' intensity of participation in on-farm training and demonstration, due to the nature of the outcome variables (i.e., count outcome data) both the Poisson regression and the negative binomial regression models are employed on data collected from the full sample. The findings of the investigation indicate that a host of factors – relating to human capital, financial capital, physical capital, social capital, and access to infrastructure and services – influence the farmers' differential involvement in these services.

With regard to the impact evaluation, the study made reference to a sub-sample within the full sample (i.e., 90 households trained at FTCs in 2009 and 160 control respondents selected for this purpose). Using the same data collection instrument in both treatment and comparison areas, data were generated specifically for the purposes of impact evaluation. Due to the non-random allocation of FTCs and self-selection issues as well as the cross-sectional nature of data collected, the propensity score matching (PSM) procedure was employed to estimate the causal effect of an FTC-based modular training on farm income. The results of the analysis show a positive and statistically highly significant gain of farm income (excluding *chat*) by the participants of the modular training, which is between Birr 9,557.47 and Birr 10,387.53 per household, on average.

Lastly, in order to analyze whether a payment for improved AAS scheme is a possible way towards diversifying the institutional options for providing AASs, a choice experiment was conducted to estimate the farmers' willingness to pay (WTP) for services describing improved FTC-based provision of AASs. By using the best-worst approach to elicit preference data from 120 households selected from treatment areas, applying the solemn oath procedure to mitigate hypothetical bias, and estimating the random parameters logit (RPL) model, the WTP estimates are derived for advice, training, and demonstration provided by the FTCs. In doing so, the study shows that the largest WTP values are associated with demonstration, followed by training. A small premium price is also associated with advice in the group that is not exposed to the solemn oath (i.e., the control treatment). Moreover, it is found that signing on a solemn oath form reduces the mean WTP values compared to not signing (but agreeing to tell the truth) and the control treatment.

This dissertation consists of three essays: Essay 1 *“Predictors of Smallholder Farmers’ Intensity of Participation in On-farm Agricultural Advisory Service Provision in Haramaya District, Eastern Ethiopia”*; Essay 2 *“Improving Smallholder Farmers’ Income through Farmers’ Training Centers: an Impact Evaluation in Haramaya District, Eastern Ethiopia”*; and, Essay 3 *“Smallholder Farmers’ Willingness to Pay for Agricultural Advisory Services: Investigating the Effect of Solemn Oath in Mitigating Hypothetical Bias in Best-Worst Choice Experiments in Haramaya District, Eastern Ethiopia”*. Each of the three essays begins with an introductory section where the background to the problem and gaps in previous empirical investigations are discussed. In Essays I & III, there is also an additional discussion on this under the section ‘literature review’. The ‘research methodology’ section in each essay presents a detailed account of the sampling procedure, survey design and methods, as well as empirical strategy and

estimation. Immediately following this is the result and discussion section where the main findings of the study are presented and discussed. Finally, the conclusion and recommendation section provides some policy implications of the studies.

Key words: Agricultural advisory services, farmers' training centers, access/participation, farm income, Poisson regression, negative binomial regression, propensity score matching, choice experiments, willingness to pay, best-worst, solemn oath, hypothetical bias, Ethiopia

1 Predictors of Smallholder Farmers' Intensity of Participation in On-farm Agricultural Advisory Service Provision in Haramaya District, Eastern Ethiopia [†]

1.1 Introduction

Small farms and smallholder farming systems are an integral part of agricultural development in many developing countries. This is more so in the context of Ethiopia where smallholder mixed-farming system is the mainstay for the majority of the rural population (McIntosh et al., 2013; Abro et al., 2014; Berhanu & Poulton, 2014). Despite the existence of an on-going debate on the role of small farms in fostering agricultural development in today's increasingly globalizing world (see, for example, Hazell et al., 2010; Wiggins et al., 2010), a number of studies document their importance in sustainable agricultural development and poverty reduction (D'souza & Ikerd, 1996; Devendra, 2007; Rocha et al., 2012). However, only a small body of literature documents the nature of technology adoption in small farms (Rauniyar & Goode, 1992), the challenges facing extension and advisory services in addressing their problems (Ballantyne, 1987), and the importance of reaching small farm owners with advisory services (Marshall, 2012). Hence, the issue of effective agricultural advisory services (AASs) needs special attention in relation to these farms and farming systems, which support the lives of millions of people worldwide.

[†] Preliminary version published as a working paper at Natural Resources, Agricultural Development and Food Security International Research Network (NAF-IRN). NAF International Working Paper Series No. 14/9, Year 2004. <http://economia.unipv.it/naf/>

Agricultural advisory services (AASs) encompass the entire range of rural services designed to foster the access of smallholder farmers to technical knowledge and information, improve their skills through training and demonstration, and facilitate their access to microfinance credit and input/output markets (Birner et al., 2006; Swanson, 2008). In the case of Ethiopia, the provision of formal and organized AASs dates back to the 1950s. Since then, AASs in the country have gone through a number of transformations. Currently, the provision of AASs has been through the Participatory Demonstration and Training Extension System (PADETES), which puts smallholder farmers at the center-stage (for a detailed account of the historical evolution of AASs in Ethiopia, see Belay, 2003; Gebremedhin et al., 2006). In addition to the on-farm/general AASs in the country, the provision of AASs is recently being undertaken through farmers' training centers (FTCs), which are training and information institutions that serve as focal points for agricultural development activities in rural areas.

The importance of AASs in various contexts has been emphasized by some previous studies. Although there is scarcity of empirical evidence on the impact of new approaches to AASs (for example, FTCs) in diverse farming systems, available research suggests that AASs play significant roles in improving farmers' skills (Tripp et al., 2005), knowledge level (Godtland et al., 2004), production, productivity and income (Birkhaeuser et al., 1991; Godtland et al., 2004; Davis et al., 2012), consumption growth and poverty (Dercon et al., 2009), and food security (Larsen & Lilleør, 2014). There are also some studies documenting poor or negligible performance of AASs in some contexts. Nonetheless, the benefit that farmers obtain out of participating in AASs and the resulting impact depends, to a great extent, by the (intensity of) farmers' direct and indirect access to these services.

In relation to farmers' access to on-farm AASs (e.g., training and demonstration), there is a scarcity of empirical evidence. Although some previous studies (such as, Kalinda et al., 1998; Rehman et al., 2013; Anaglo et al., 2014; Baloch & Thapa, 2014) attempted to address the issue, there are a number of concerns. First, almost invariably, all the studies deal with access to agricultural information. However, since agricultural activities are seasonal and have many phases (for example, land preparation, sowing/planting, fertilizer application, harvesting, etc.), they require continuous follow up and support from AAS providers. Therefore, having a one-time mere access to AASs may not have the same effect as getting frequent services. Second, they attempt to provide explanations on the factors affecting access to advisory services. None of them discuss the issue of (intensity of) access to agricultural training and demonstration services that are equally relevant to smallholder farming systems. Finally, there are serious methodological flaws in the studies, including sample size, adequacy and choice of explanatory variables, and empirical strategy to analyze data (see the next section for the details on this issue).

Recognizing these crucial issues, the present study focuses on investigating the determinants of intensity of farmers' participation in on-farm AASs (training and demonstration) in Haramaya district of eastern Ethiopia. To this end, primary data was collected from a sample of 340 farm household heads in the period May to October 2013. The study has at least three unique features. First, by explicitly focusing on two dimensions of AASs (i.e., training and demonstration), it contributes to the literature on access to AASs from a different perspective. Second, by dealing with intensity rather than a simple access to AASs, it provides additional understanding to the way we conceptualize the provision of AASs. Third, by employing methodologically rigorous approaches – count data models with sound choice of explanatory variables – it presents

empirical evidence on the factors that explain the farmers' differential access to on-farm training and demonstration. In summary, at least in these three accounts, the present investigation is innovative and relevant in the current discussion of reinventing AASs. Such an analysis can fill the existing gap in our understanding of AASs and can provide useful insights on future targeting of smallholder farmers with on-farm AASs. More importantly, it can contribute to the current discussion on the principles for responsible investment in agriculture and food systems (FAO, 2014) by highlighting the issue of access to AASs in smallholder farming systems.

The remainder of the paper is structured as follows. In the next section, a brief review of the literature pertaining to the issue of farmers' access to AASs in developing countries (especially focusing on Ethiopia) is provided. What follows this is a discussion on data and methods used in the study, including the choice of empirical strategy. After this is the result and discussion section where the main findings of the study are presented and discussed. Finally, some conclusions and recommendations are provided.

1.2 Literature Review

Agricultural advisory services (AASs)¹ are one of the important vehicles for transformation of livelihoods in rural areas. These services try to adapt and create knowledge by conducting field experiments, disseminating knowledge through demonstrations and seminars, providing on-farm training to farmers, and establishing direct contacts between agricultural advisers² and agricultural growers (Dinar, 1996). This means that AASs not only work towards accelerating the diffusion of innovation and adoption of new technologies, but also improve the managerial ability of farmers and affect the efficient utilization of existing technologies (Birkhaeuser et al., 1991; Dinar, 1996). The main focus of this section is to provide a brief review of the studies on smallholder farmers' participation in AASs in developing countries, especially focusing on Ethiopia.

1.2.1 Smallholder Farmers' Access to AASs

As mentioned in the introductory section, there is a paucity of careful empirical investigation documenting the factors affecting farmers' participation and intensity of participation in on-farm AASs in the context of developing countries. However, a few studies in Zambia (Kalinda et al., 1998), Nigeria (Olajide, 2011), Ghana (Anaglo et al., 2014), and Pakistan (Rehman et al., 2013; Baloch & Thapa, 2014) attempt to provide some insights into sociodemographic characteristics which have significant associations with farmers' participation in AASs. A summary of the

¹ Used interchangeably with the phrase “agricultural extension services”, “rural advisory services” or simply “extension services” (See for example, Anderson, 2008; Swanson, 2008; Davis et al., 2012). Since today's understanding of agricultural extension goes beyond technology transfer to facilitation, beyond training to learning, and includes assisting farmer groups to form, dealing with marketing issues, and partnering with a broad range of service providers and other agencies, many people are now using the phrase, “agricultural advisory services,” instead (Davis, 2008).

² These are agricultural extension agents (also called development agents) who carry out agricultural advisory services.

studies is given in Table 1.1. These studies emphasize the role of human capital and physical capital. However, none of the studies show any significant effect of financial capital (e.g., farm income, off-farm income, credit) and social capital (such as networks and membership in organizations) in determining farmers' involvement in on-farm AASs. Most importantly, however, all the studies mentioned above have severe methodological flaws.

First, they all use a few explanatory variables. By doing so, they leave out other very important characteristics (e.g., socioeconomic, farm level, and institutional) that may likely affect the participation of farmers in AASs. Second, many of the studies base their analyses on a small sample, making generalizability to the wider sampling frame and/or population very hard. Third, and most importantly, the analysis employed in them is substandard and weak. Most of them use descriptive statistics, crosstabulations and chi-square test results of associations between sociodemographic characteristics and access to AASs. Although most of the authors claim to have shown the determinants of access to AASs, in reality, they simply have documented preliminary associations. The study by Baloch & Thapa (2014), although somewhat better than the rest in terms of using appropriate empirical strategy (i.e., logistic regression to investigate the factors affecting participation), also falls short of incorporating relevant covariates. Moreover, it collapses many of the covariates into dummies, instead of using them as they are measured. The same also applies to Rehman et al. (2013), who also group the variables into different categories instead of using them as such. Hence, the findings in each of these studies and the corresponding interpretations have to be approached with great caution.

Table 1.1 Some Previous Studies on Factors Affecting Access to AASs in Developing Countries

Authors	Country	Variables with significant effects	Methodology (sample size and analytical strategy)
Kalinda et al. (1999)	Zambia	age, gender, education, farm size	150 households; descriptive statistics, chi-square
Olajide (2011)	Nigeria	age, gender, education, marital status, information sources	118 farmers; descriptive statistics, chi-square
Anaglo et al. (2014)	Ghana	gender	400 farmers; crosstabulation, chi-square
Rehman et al. (2013)	Pakistan	age, education, land size, farming experience	361 agricultural magazine subscribers; descriptive statistics
Baloch & Thapa (2014)	Pakistan	age, education, farming as main occupation, number of date palm trees (assets), % of dead palm trees	200 households; descriptive and binary logistic regression

Examining the studies in detail, whereas human capital variables (age, gender, education) have no significant association with access to AASs in Zambia (Kalinda et al., 1998), farm size categories exhibit significant effects on farmers' access, where farmers with medium land size (6 – 9 ha) are more likely to be visited by development agents or take part in agricultural training. Unlike the case in Zambia, the study in Nigeria (Olajide, 2011) documents that all the three human capital indicators have positive and significant effects on access to agricultural information. However, marital status has no significant effect on access. In Ghana, Anaglo et al. (2014) indicate that the gender of farmers significantly affects their participation: male farmers have better access than female farmers. The case study in Pakistan (Rehman et al., 2013) suggests mixed effect of human capital – whereas education has a positive effect, age is not significant. In addition, although land size (physical capital) has a positive effect, the influence of farming experience is not significant. Finally, both human capital (age, education) and

physical capital (asset – number of date palm trees) have a positive effect on access to AASs (Baloch & Thapa, 2014). However, the percentage of dead palm trees present on the farm has negative effect.

In general, if one is prepared to ignore the methodological issues mentioned above, these studies suggest that the factors affecting participation in AASs appear to vary from place to place and country to country. However, since none of the studies looked into the determinants of intensity of participation, which matters more than a simple access, it is imperative to consider this issue. Most importantly, there is a heightened need for a methodologically-rigorous analysis of the case.

i. Smallholder Farmers' Participation in AASs in Ethiopia

Ethiopian agriculture is virtually small-scale, subsistence-oriented. More precisely, more than 95% of the country's agricultural output is generated by smallholder subsistence farmers (Belay, 2003). In relation to smallholder farming in the country, AASs contribute a lot in improving production and productivity. These services are the largest and fastest growing in the African continent, which the government uses to achieve economic growth (Berhanu & Poulton, 2014). These programs represent a significant public investment, amounting to over 50 million dollars annually, or almost 2% of the agricultural Gross Domestic Product (GDP) in recent years – a figure that exceeds expenditure in most other developing countries and regions (Mogues et al., 2009).

Over the past many decades, AASs in the country have gone through a number of transformational stages: the land grant system provided by the Imperial Ethiopian College of Agriculture and Mechanical Arts (IECMA) in the 1950s; the Comprehensive Package Programs

(CPPs) in the period 1967 – 1971; the Minimum Package Projects (MPPs) in the period 1971 – 1985; the Peasant Agricultural Development Program (PADEP) in the period 1985 – 1995; and the Participatory Demonstration and Training Extension System (PADETES)³ in the period 1995 to date. The AASs provided by IECAMA in the early 1950s were limited to areas surrounding the experiment stations that were being operated by the college, and hence had limited coverage and impact. Comprehensive package programs were more of rural development approaches than just advisory services and were limited to only few high agricultural potential areas. Minimum package projects had wider coverage compared with the CPPs, but still failed to cover the majority of the country. PADEP was a victim of the ideological doctrine that was being followed by the Marxist military regime and so limited its services to producers' co-operatives. The current advisory service (PADETES – since 1991) appears to give more attention to smallholders compared to its predecessors (Belay, 2003; Gebremedhin et al., 2006).

More importantly, the coming to power of the Ethiopian People's Revolutionary Democratic Front (EPRDF) in 1991 has resulted in significant investments in the agricultural sector and its supporting institutions. Under the *Agricultural Development Led Industrialization (ADLI)* policy and strategy of development, a lot of emphasis has been given to growth and transformation of the agricultural sector and especially to AASs, because it is recognized that most Ethiopian farmers do not use modern agricultural technology, and the agricultural knowledge and information system (agricultural research, education, and advisory service) is poorly integrated (Lemma, 2007). As a bold measure to strengthen AASs in the country, therefore, the government has taken swift action and increased the number of agricultural advisers to more than 60,000 by opening 25 Agricultural Technical and Vocational Education and Training (ATVET) colleges

³ For detailed historical accounts of the Ethiopian AASs, see Belay (2003); Belay and Abebaw (2004); and, Gebremedhin et al. (2006).

(Ethiopia, MOARD, 2009b)⁴. In addition, the government has embarked on the establishment of rural educational centers (i.e., farmers' training centers (FTCs)) at the lowest level of administration in rural areas.

The FTC-based Provision of AASs

Farmers' training centers (FTCs) are training and information institutions that serve as focal points for agricultural development activities at the lowest level of administration (i.e., *kebele/peasant association*) in rural areas. The establishment and functionality of these centers is one of the policy directions followed by the government of Ethiopia in the implementation of the Agricultural Development Led Industrialization (ADLI) development strategy, which aims to transform the age-old legacy of subsistence agriculture to one that is market-oriented and sustainable. More specifically, the setting up of FTCs is aimed at improving intermediate (such as knowledge, skill, and productivity) and final (such as income, food security, and sustainable economic growth) outcomes by promoting market-driven production systems in the country. To this end, FTCs actively involve in the provision of technical information, advice, modular training, and demonstration services related to crop, livestock, and natural resources in the process of producing skillful and knowledgeable farmers in rural areas.

Concerning the FTC-based training, two types of training are being offered to farmers at FTCs. These are short-term and modular training. The short-term training is delivered to 15-20 farmers for a period of 3-20 days on the demonstration plots of the FTCs or on the fields of model farmers. The organization of modular training takes special consideration of the size of participants in order to reduce/avoid adverse effects of large class size. Consequently, a

⁴ Ethiopia, MOARD (Ministry of Agriculture and Rural Development). 2009b. Data on ATVET colleges graduates. Addis Ababa, Ethiopia: Department of Agricultural Technical and Vocational Training and Education, MOARD.

classroom for such training consists of 20-30 trainees. A range of subjects are covered by the modular training – which is 80% practical and 20% theoretical in its content – including agronomy, animal husbandry and health, and natural resource management and conservation. The duration of modular training provided in each subject is 3-6 months. Depending on the type of the subject, two training periods can be arranged each year. Each trainee is expected to attend 2 days/week (6 hours per day, of which 2 hours for theory and 4 hours for practical part). This means, a total of 300 hours of participation in the training is required to graduate with a “Green Certificate.” However, there is flexibility in terms of training hours based on the particular context where the training is organized. The training methodologies employed by the FTCs include: class room lecture, demonstration trial, field practice, exposure visit, and rural radio program.

Hence, in areas where there are established and functional FTCs, AASs are being provided through the FTCs, where the farmers come to the centers to obtain advice, participate in modular training (which lasts for 3 – 6 months), and take part in demonstration of improved technologies or practices. However, since there is also an independent and parallel AAS provision, called the on-farm/general AASs, farmers also get the services on their farm (instead of at the FTCs). In non-FTC areas – i.e., peasant associations (PAs) which do not have any established or functional FTCs – the on-farm AASs are organized in such a way that development agents go to the farmers to provide advisory services.⁵ They also organize on-farm training and demonstration (at

⁵ In both cases, farmers obtain services related to crop, livestock and natural resources. However, the organization and provision of the services is different. In the FTC-based AASs, for example, the farmers receive modular training at FTCs. Likewise, demonstrations are organized at the FTC-managed demonstration plots and farmers go there. In general, the farmers go to the FTCs on a regular basis to obtain these services. In the general/on-farm AASs, however, it is the development agents who come to the farmers and provide on-farm advice, training, and demonstration.

demonstration sites or farmers' plots) so that the farmers participate in short-term training and field days.

In general, although the on-farm AASs and the FTC-based AASs are independent of each other, farmers in the FTC-PAs have access to both of them. Regarding the components of AASs (both on-farm and FTC-based), all the three major elements (i.e., advice, training, and demonstration) are independently provided. There is a stark difference, however, between the on-farm and FTC-based AAS provisions in terms of mode, content, structure, and duration of service delivery. In relation to the mode of AAS provision, for example, training is organized in cycles. A household who participated in one cycle of training will not take part in a subsequent training. The same is true with method and/or result demonstration. This is especially true in the FTC-based provision of AASs. Moreover, whereas one training cycle in the FTC-based AASs may refer to a duration of three to six months, it may mean half-a-day or five days training in the on-farm AASs. Similarly, whereas the AASs in the FTC-based provision are highly structured and organized, they are not very structured in the on-farm AASs. Hence, one cannot sum up, for example, on-farm training and FTC-based training in FTC areas. The same applies to demonstration and advice.

Challenges Faced in On-farm AAS Provision

The history of AASs in the country shows that farmers have differential access to and participation in AASs in general. A handful of reasons were raised in relation to the challenges faced in on-farm AAS provision in the country. First, since the on-farm AASs are supply-led and free of charge, coupled with lack of clearly defined targeting and selection criteria, it is not very clear how farmers are screened and selected to take part in the services. An earlier research in

this regard indicates that agricultural advisers (also called development agents) and local leaders play a significant role in the selection process (Belay, 2003), although farmers' self-selection is also improving recently. For example, a few economically better, progressive, risk-taking and experienced farmers might seek out for AASs. On the other hand, the role of political affiliation in the selection process cannot be undermined (Berhanu & Poulton, 2014). Hence, AAS provision in the country is criticized for the lack of participation of relevant stakeholders in planning, implementation and evaluation (Belay, 2003). Second, although there is a continuous supply of trained development agents from agricultural colleges, the farmer-to-adviser ratio is still too high to address all farmers with timely agricultural information and training. Therefore, some farmers are left out inevitably because of the institutional capacity of AAS providers.

Third, since development agents were expected to demonstrate the impact of new technologies (such as high yielding crop varieties), previous research shows that they tend to work more closely with farmers who have large plots of land and who are willing to allocate a parcel of their land for demonstration purposes. Fourth, due to mobility constraints, advisers also tended to work with farmers located close to all-weather roads, and thus farmers living in remote and inaccessible rural areas may have limited chance of participating in AASs. Fifth, since agricultural development requires complementary inputs other than advisory services, development agents tended to target farmers who have access to formal or informal credit and market. Lastly, since majority of the advisers are males, they tended to work with male farmers for cultural reasons (Belay, 2003; Belay & Abebaw, 2004). However, the post-1991 period has seen an unprecedented investment in the agricultural sector and especially in AASs. Consequently, there are a lot of changes in targeting and provision of on-farm AASs in the

country. However, there are still observable differences in terms of farmers' intensity of participation in AASs.

1.3 Research Methodology

In this section, first, a brief description of the sampling procedure followed during the selection of household heads for the study and the process of primary data collection is presented. Following this is a discussion on the empirical strategy for data analysis, highlighting on the various tests and comparisons on the choice of count data models – Poisson and Negative Binomial.

1.3.1 Sampling Procedures and Data

This study is conducted in Haramaya district of East Hararghe zone, Oromia region, Ethiopia. In the district, there are ten peasant associations (PAs) which have fully functional farmers' training centers (FTCs) and two PAs which do not have any FTCs. From the ten PAs with functional FTCs, three PAs (*Ifa Oromia*, *Adele Waltaha* and *Biftu Geda*) were selected in such a way that they are representative to the rest of the PAs with functional FTCs and comparable to the two non-FTC PAs (*Fendisha Lencha* and *Ifa Bate*) in terms of biophysical (topography, weather, etc.) and socio-economic characteristics. Following this, a list of households trained in the three FTCs was obtained from the district bureau of agriculture and rural development (BoARD). The list contained 450 household heads who were reportedly trained in the three FTC PAs in 2008 and 2009. Although the document from the district BoARD indicated that this is the number of farmers trained in the three FTC-PAs, the actual number of farmers who completed the modular training, according to the development agents, is 30 per FTC per year. This means that there are 90 farmers who completed the training in 2008 and 90 in 2009 in the three FTC PAs. Hence, a total of 180 households who actually completed the FTC-training (i.e., 60 households in each PA) were selected in the PAs with functional FTCs.

The households were selected into modular training following the criteria: educational background, devotion and initiatives, age, and gender (more on this is found in the section ‘choice of variables for estimating propensity score’ in Essay-II). Furthermore, development agents and local leaders are key players in the recruitment and selection of such households into the training. For example, a household that received advice on crop production or livestock management in a given year may not get the same service in the next year. This means that, in the case of on-farm AASs for example, the provision of subsequent advice, training, or demonstration is not contingent up on the first visit, training, or demonstration. In most cases, the decision to target a household over another one is taken by the providers of the on-farm AASs.

In the two non-FTC PAs, the list of households was also obtained from the district bureau of agriculture and rural development office. Accordingly, 824 households live in these areas. However, since it was observed – and also confirmed by key-informants – that in FTC-PAs many relatively better-off, progressive and model household heads were included in the training, this necessitated the development of another list containing households judged to be ‘good’ performers in terms of *devotion* and *initiatives* (as measured by, for example, main occupation, years of experience in farming, and experience with on-farm AASs) in these comparison PAs.⁶ The criteria *devotion* and *initiatives* (discussed in detail in the section ‘choice of variables for estimating propensity score’ in Essay-II) refer to the commitment and self-initiation, hard-working capacity, willingness of trainees to put into practice knowledge and skill learned, capacity of trainees to adopt new technologies, and readiness of the trainees to share such knowledge and skill to other farmers in the area. In addition to these, there were considerations regarding educational background, age, and gender of the farmers. As a result, a total of 495 such

⁶ The new list was developed in consultation with Development Agents, local leaders, district extension personnel, and PA administration.

households were obtained in both PAs (188 in *Ifa Bate* and 307 in *Fendisha Lencha*). Following this, a total of 160 household heads from these areas (i.e., 80 households in each PA) were randomly sampled. Hence, the total sample size for this study is 340 household heads from the five PAs.

Data was gathered from the sample respondents in the period May to October 2013. The process of data collection followed the research guidance for development practitioners in developing countries (e.g., Leones & Rozelle, 1991; Nyariki, 2009). First, a total of seven enumerators from the district and three supervisors from Haramaya University were selected for the purpose of collecting good quality data.⁷ Second, having provided orientation training to the research assistants, a pilot survey on 30 households was carried out to test the questionnaire. Third, following the feedback from the pilot, the questionnaire was translated to *Amharic*, the national language, and also some items in the questionnaire were simplified. Finally, through the main survey, data pertaining to agriculture and allied activities in relation to AASs were gathered. More specifically, data were collected on household and farm level characteristics, participation in general/on-farm and FTC-based advisory services, household assets, livelihood activities (crop and livestock production), and participation in (formal/informal) organizations/associations.

1.3.2 Empirical Strategy for Data Analysis

In the present study, both the Poisson regression and the negative binomial regression are used because of the nature of the outcome variables (i.e., count outcome data) and the distributions of them (equi-dispersion and over-dispersion). Details on the choice of the models are provided in

⁷ The criteria to select the enumerators were: education, work experience, and language competence. Regarding education, the enumerators were required to have some years of training at agricultural colleges (in crop production, animal husbandry, and/or natural resource management). In addition, due attention was also given to select enumerators who are stationed at the study PAs, who have been working with the respective communities, and who are competent in both local and English languages.

the next section. Count outcome variables, such as the number of visits by development agents per year or the number of on-farm training per year, are often modeled using the Poisson regression, provided that each subsequent visit or training is not conditional on the former. In situations where subsequent visits or training are affected by the previous experience/encounter with the service (visit or training in this case), neither the Poisson nor the families of it (such as the negative binomial regression) are appropriate. Hence, in order to use such models, it is a requirement that each subsequent visit or training be independent or not causally related to the former. Otherwise, models that account for such pattern in the count outcome data have to be employed. In the present study, however, since there is a lack of forward causal relationship between the first, second, and so on visits, training, or demonstrations, the choice of the count data model falls in the Poisson or its family.

The basic assumption behind the Poisson model is that the variance of the outcome variable is constrained to equal the mean, which is referred to as equi-dispersion (Greene, 2008; Cameron & Trivedi, 2013). This assumption may fail to hold in many circumstances (Cameron & Trivedi, 1986). For example, in many studies employing discrete outcome variables, the sampling distribution may often result in higher frequency of zero counts than would be expected from a Poisson distribution (Byers et al., 2003). In the present study context, this can be observed if higher number of farmers were not visited by development agents or did not take part in at least one on-farm training or demonstration.

According to Agresti (2007) and others, the Poisson loglinear model can be give as

$$(1) \quad \log(\mu) = \alpha + \beta x$$

where μ is the non-negative count outcome variable (also denoted as $E(Y)$ – the expected value of the outcome variable Y), α is the constant, and x is a vector of explanatory variables with the corresponding coefficient estimates β . The outcome variable μ is in logarithmic term, since a linear model can yield $\mu < 0$, while the possible values for it are $\mu \geq 0$. Hence, the mean satisfying the exponential relationship in (1) can be written as

$$(2) \quad \mu = \exp(\alpha + \beta x) = e^\alpha e^{\beta x}$$

In situations where the outcome variable takes discrete and non-negative values (as in the case of this study), standard ordinary least squares (OLS) regressions are difficult to implement because the distribution is not likely to be normal. For example, implementing the OLS regression will become even inappropriate when the distribution is highly skewed (i.e., when many scores or counts of zero are found for the outcome variable). In such circumstances, it is necessary to use models that accommodate a higher frequency of zero counts and over-dispersion (Byers et al., 2003). Since it is a common encounter that observed outcome data will exhibit a considerable amount of over-dispersion (Agresti, 2007; Green, 2008) and violate the key requirement for Poisson, many scholars seek alternatives to the Poisson model. One such alternative is the Negative Binomial model, which is a generalization of the Poisson regression model that accounts for over-dispersion by including a disturbance/error term (Agresti, 1996, 2007, 2014; Byers et al., 2003; Cameron & Trivedi, 2009).

According to Green (2008), the negative binomial model is by far the most common specification, perhaps because data suitable for a Poisson regression are difficult to find in reality. This model, unlike the Poisson, has an additional parameter such that the variance can exceed the mean (i.e., $E(Y) = \mu$, $\text{Var}(Y) = \mu + D\mu^2$). The non-negative index D is called a dispersion parameter (some scholars refer to it as the over-dispersion parameter *alpha*). When

there is greater heterogeneity in the Poisson mean values, this heterogeneity results in larger values of D . However, as D approaches to 0, $\text{Var}(Y)$ converges to μ and the negative binomial distribution reverts to the Poisson distribution. The farther D falls above 0, the greater the over-dispersion relative to Poisson variability (Hilbe, 2011).

The usual functional form of the negative binomial regression (e.g., Byers et al., 2003) can be given as

$$(3) \quad \log(\mu_i) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \sigma \varepsilon_i$$

where μ_i is the expected value of the outcome variable (i.e., number of training and demonstration per year) for farmer i , x_i are the independent variables with the corresponding regression coefficients β , and $\sigma \varepsilon_i$ is the disturbance term.

i. Inspecting the Distribution of the Outcome Variables and Model Fit

This study takes into account three discrete outcome variables – number of on-farm visits by development agents, number of on-farm training farmers participate, and number of on-farm demonstrations farmers take part in.⁸ It should be noted here that all these components of on-farm AASs are provided to the farmers independent of each other, and that the intensity of each service provided does not depend on initial encounter with the service. Table 1.2 presents the mean and variance of each of these variables in FTC areas (column 1) and non-FTC areas (column 2). In addition, the table contains the mean and variance of FTC-based advice, training, and demonstration in FTC areas, in order to illustrate the general picture of AAS provision in these areas. However, since only the on-farm AAS is common to both areas, and since it is not

⁸ However, since, as shown latter, the distribution of the outcome variable ‘number of on-farm visits’ does not fit either the Poisson or the negative binomial distribution, it is left out of further investigation.

possible to add up the quantities of on-farm and FTC-based services for reasons discussed earlier, it is taken into consideration for further analysis. Moreover, due to the existence of FTC-based AASs only in the FTC areas, an area-wise approach was followed in the analysis.

As can be seen from the last column of the table, there exists significant mean difference between FTC areas and non-FTC areas for all the outcome variables. Likewise, there is a general huge dispersion (comparing mean and variance in each PA) in the sample. However, looking at the FTC areas alone, the dispersion is more pronounced only for the variable on-farm *advice/visit* (column 1a). In the corresponding non-FTC areas, there is huge dispersion for all the three variables (column 2a).

Table 1.2 Distribution of On-farm and FTC-based Advice, Training, and Demonstration in the FTC Areas and Non-FTC Areas in 2013

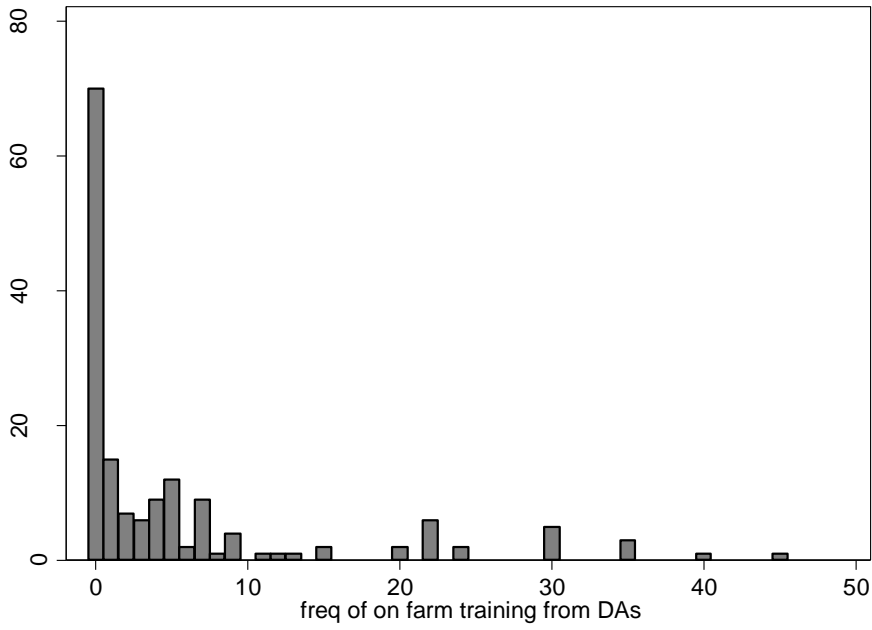
	(1) FTC areas			(2) Non-FTC areas			(3) ttest ^b	
	<i>Ifa</i> <i>Oromia</i>	<i>Adele</i> <i>Waltaha</i>	<i>Biftu</i> <i>Geda</i>	(1a) Total	<i>Fendisha</i> <i>Lencha</i>	<i>Ifa Bate</i> (2a) Total		
On-farm:								
advice								
<i>Mean (Sd)</i>	7.1 (4.8)	3.5 (11.0)	6.7 (12.4)	5.8 (10.0)	3.4 (3.3)	20.5 (19.5)	12.0 (16.4)	4.26 ***
<i>Variance</i>	22.9	121.1	153.4	100.5	10.6	380.6	267.6	
training								
<i>Mean (Sd)</i>	2.2 (1.0)	1.4 (0.9)	0.6 (0.7)	1.4 (1.1)	0.9 (1.7)	10.2 (11.4)	5.6 (9.4)	5.89 ***
<i>Variance</i>	1.0	0.8	0.5	1.2	2.9	130.7	88.4	
demonstration								
<i>Mean (Sd)</i>	1.0 (0.4)	1.3 (1.0)	0.4 (0.7)	0.9 (0.8)	0.3 (0.7)	7.4 (8.4)	3.9 (6.9)	5.72 ***
<i>Variance</i>	0.1	0.9	0.5	0.6	0.5	70.7	48.2	
FTC-based:								
advice								
<i>Mean (Sd)</i>	25.5 (5.9)	2.1 (0.8)	7.2 (6.2)	11.6 (11.2)				
<i>Variance</i>	35.2	0.7	38.4	126.1				
training								
<i>Mean (Sd)</i>	7.4 (2.4)	2.0 (0.5)	2.2 (1.1)	3.8 (2.9)				
<i>Variance</i>	5.6	0.3	1.2	8.5				
demonstration								
<i>Mean (Sd)</i>	2.2 (0.7)	1.9 (0.6)	1.8 (0.8)	1.9 (0.7)				
<i>Variance</i>	0.5	0.4	0.6	0.5				
Sample size	60	60	60	180	80	80	160	

Standard deviations (Sd) in parenthesis.

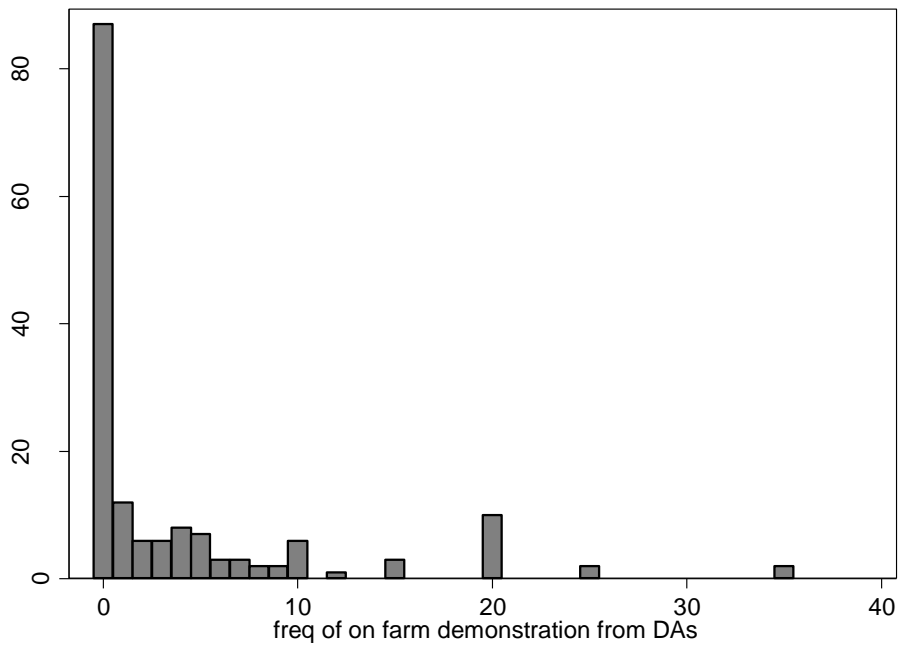
*** denotes significance at 1% level.

^b Absolute value of t-statistic (between FTC areas (1a) and Non-FTC areas (2a)).

On the basis of these observations and the requirements for the count data models (Poisson and Negative Binomial) discussed above, it appears that the negative binomial model can reasonably fit the data in the non-FTC areas (for all the three variables). It also appears that this model can fit the data for *advice* in the FTC areas. However, the data for *training* and *demonstration* in the FTC areas appear to fit more the Poisson than the negative binomial distribution. Nonetheless, in order to confirm these observations, there are some more data inspections required. A graphical inspection of the distributions of the outcome variables (training and demonstration) in the non-FTC areas, for example, shows that the data are highly skewed to the left, suggesting that the distribution is not normal and that either the Poisson or the negative binomial regression is appropriate (Figure 1.1). However, this graphical examination by itself does not clearly inform which of the two models is potentially applicable for advice, training, and demonstration in the FTC and non-FTC areas. In order to address this issue, there is a need to plot each of the variables together with the Poisson and negative binomial distributions.



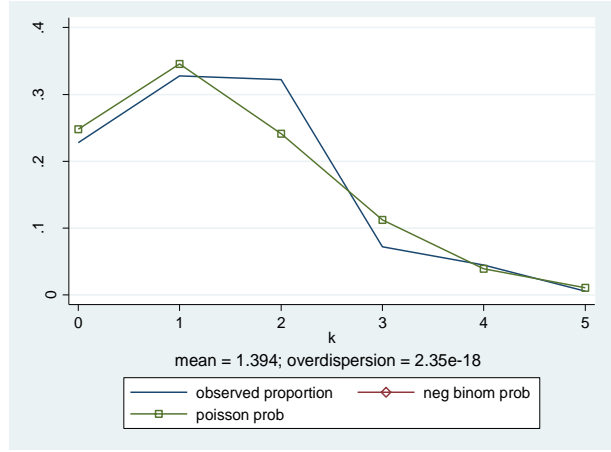
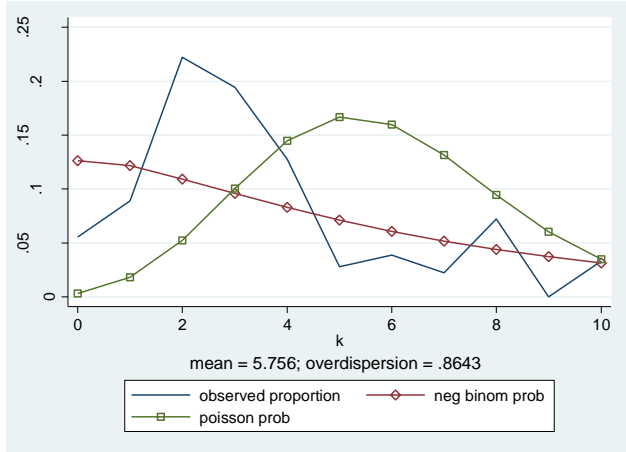
(a) Training



(b) Demonstration

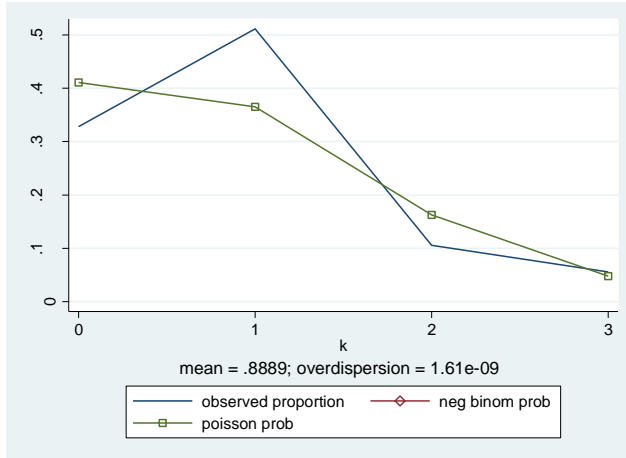
Figure 1.1 Distribution of Outcome/Dependent Variables in the Non-FTC Sample

Figures 1.2 and 1.3 display the observed distribution of the outcome variables in relation to Poisson and negative binomial distributions for the FTC and non-FTC areas, respectively. In all the figures, it is visibly evident that the outcome variable on-farm *advice* does not fit either the Poisson or the negative binomial distribution. However, a logistic regression on participation in on-farm advice/visit – by collapsing the number of advice into dummies (i.e., 1/0, those who took on-farm advice and those who did not, respectively) – was a possibility, but it is not consistent with the very aim of this research, which looks into intensity of participation. Hence, the outcome variable ‘on-farm advice’ is left out of further investigations. The distributions of *training* and *demonstration* fit the Poisson distribution in the FTC areas (Figure 1.2). In the non-FTC areas, however, they fit the negative binomial distribution (Figure 1.3).



(1) *Advice/visit*

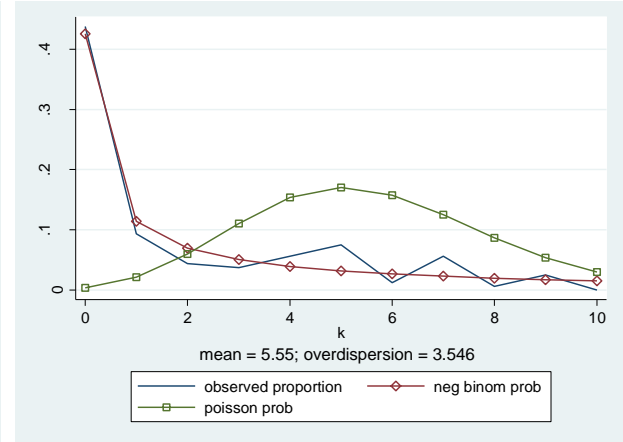
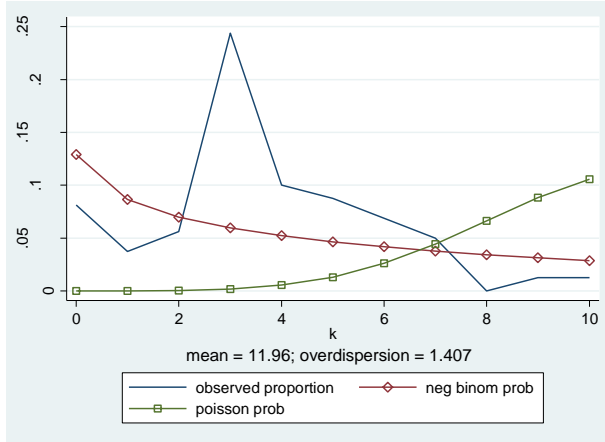
(2) *Training*



(3) *Demonstration*

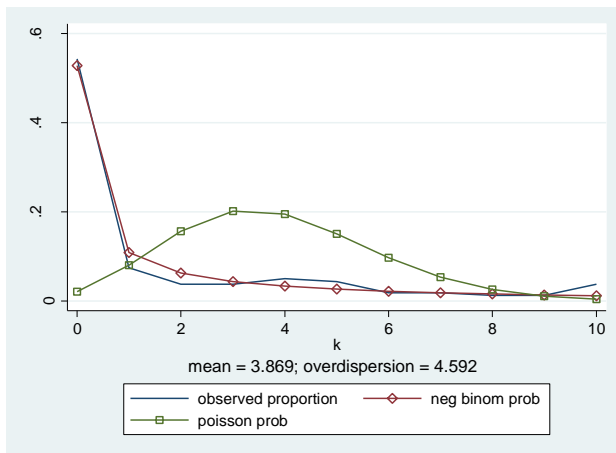
Figure 1.2 Graphical Inspection of Model Fit in the FTC Areas

Note: (1) Observed distribution of *advice* does not fit either the Poisson or Negative Binomial distributions. (2) Observed distribution of *training* reasonably fits the Poisson distribution. (3) Observed distribution of *demonstration* reasonably fits the Poisson distribution.



(1) *Advice/visit*

(2) *Training*



(3) *Demonstration*

Figure 1.3 Graphical Inspection of Model Fit in the Non-FTC Areas

Note: (1) Observed distribution of *advice* does not fit either the Poisson or Negative Binomial distributions. (2) Observed distribution of *training* reasonably fits the Negative Binomial distribution. (3) Observed distribution of *demonstration* reasonably fits the Negative Binomial distribution.

In addition to mean-variance comparisons, checking skewed distributions, and graphical inspection of outcome data, two additional tests of model fit were carried out (i.e., post-estimation tests). The first one is the goodness-of-fit χ^2 test for Poisson regression (second row from below, Table 1.4 in the ‘result and discussion’ section). The test result indicates that the use of the Poisson regression in the FTC areas is appropriate. As a rule-of-thumb, when there are statistically significant and large values of χ^2 , Poisson regression becomes inappropriate. The second test is the likelihood-ratio (LR) test after running the negative binomial regression on the same set of explanatory variables. If the test statistic for α is significantly different from zero, this reinforces the fact that Poisson regression is not appropriate and, hence, the use of the negative binomial distribution is justified (see the third row from below in Table 1.4 in the ‘result and discussion’ section).

In summary, in this study, intensity of participation in the provision of on-farm training/demonstration is operationalized as the number of on-farm training/demonstration organized on crop production, livestock production and management, and/or natural resource management a farmer participates in one year. In the analysis, the negative binomial regression was used for the non-FTC areas (for both *training* and *demonstration*) and the Poisson regression for the FTC areas (for both *training* and *demonstration*).

ii. Selection of Explanatory Variables

The inclusion of explanatory variables into the regression models considers the following aspects. First, a special attention was paid to the aim of on-farm AAS provision in the study area – that is, improving smallholder farmers’ income. Hence, indicators such as farm income and its relationship with off-farm employment are considered important in this regard. Given this, it is

expected, for example, that AAS providers should channel more services to low income farmers (to be consistent with their aim). If, however, AASs target more high income farmers, this may suggest that these services are biased against the low income farmers. Second, the observed significant differences in selected sociodemographic characteristics between the respondents in FTC and non-FTC areas (Table 1.3) also point out to the inclusion of some covariates in explaining intensity of participation in the provision of on-farm training and demonstration. Moreover, correlations among these covariates were tested in the final selection of explanatory variables for the regression models. Lastly, the small body of literature on access to and participation in AASs (e.g., Kalinda et al., 1998; Anaglo et al., 2014 – summarized in Table 1.1 of ‘literature review’ section) also provided some entry point in the selection of the explanatory variables.

Therefore, on the basis of these considerations, rich dataset at hand, and aim of the study, the following explanatory variables were included: human capital (age, education, household size); physical capital (value of household asset, value of livestock possessed, land holding size); economic/financial capital (farm income, off-farm employment, credit); social capital (membership in *iddir*, networks); access to infrastructure and services (animal health care, experience in AASs, FTC-based training, FTC-based demonstration); and, location of the households (i.e., PA dummies to account for the effect of a household’s location on intensity of participation in on-farm AASs). In the final regression, although different specifications (including/excluding some variables, continuous versus categorical variables, etc.) were tried, the ones that resulted in a better model fit were used.

1.4 Result and Discussion

In this section, the main findings of the study are presented and discussed. First, the descriptive statistics of sociodemographic characteristics of respondents in FTC and non-FTC areas, along with the test of equality of means between the two groups, are presented. Then, a discussion is given on the Poisson and negative binomial regression results pertaining to the determinants of the extent of farmers' involvement in on-farm training and demonstration in FTC and non-FTC areas.

1.4.1 Descriptive Statistics of Sociodemographic Characteristics of Respondents

The descriptive statistics of sociodemographic characteristics of the respondents is given in Table 1.3. Together with some empirical evidence from previous studies, this information is used to explain the reasons behind the observed differences in the sign and explanatory power of the various covariates included in the regression models. As can be seen from the table, except for level of education, household size, asset ownership, land certification, value of livestock, farm income, and networks, there are no statistically significant differences between the respondents in the two areas. Regarding education, there is an overall statistically significant difference between the respondents in the two areas ($t = 6.89$, $p = 0.000$), which is attributable to the differences in primary and lower secondary schooling as well as some college/university level education. In general, although there is no difference in terms of primary school coverage in the two areas, many respondents in the non-FTC areas went to school compared to those in FTC areas. Household size exhibits significant differences only in the category *household size > 10 members* because about 3% and 11% of the respondents in FTC and non-FTC areas,

respectively, fall in this category. However, the difference in household size in the two areas is not significant overall ($t = 0.91$, $p = 0.183$).

Table 1.3 Descriptive Statistics of Sociodemographic Characteristics of Respondents in 2013

Variable	(1) FTC areas			(2) Non-FTC areas			(3) ttest ^b (1a) & (2a)	
	<i>Ifa Oromia</i>	<i>Adele Waltaha</i>	<i>Biftu Geda</i>	(1a) Total	<i>Fendisha Lencha</i>	<i>Ifa Bate</i>		(2a) Total
Age (years)	39.02 (3.65)	33.87 (4.88)	41.43 (4.90)	38.11 (5.49)	37.49 (7.87)	38.65 (5.49)	38.07 (6.79)	0.06
Female household head ^a	0.02	0.03	0.0	0.02	0.03	0.05	0.04	1.19
Education ^a								
No formal education	70.00	85.00	81.67	79.89	27.50	37.50	32.5	9.73
Primary	6.67	5.00	8.33	6.7	35.00	30.00	32.5	6.42 ^{***}
Primary & lower secondary	18.33	5.00	10.00	11.11	25.00	28.75	26.88	3.80 ^{***}
Secondary	5.00	5.00	0.00	3.33	10.00	1.25	5.6	1.03
Some college/university	0.00	0.00	0.00	0.00	2.50	2.50	2.5	2.14 ^{**}
Household size ^a								
< 5 members	1.67	30.00	0.00	10.56	10.00	11.25	10.63	0.02
5 – 10 members	98.33	70.00	91.67	86.67	81.25	76.25	78.75	1.94
> 10 members	0.00	0.00	8.33	2.78	8.75	12.50	10.63	2.97 ^{***}
Experience in farming (years)	23.63 (5.15)	25.97 (5.01)	22.42 (4.80)	24.01 (5.18)	23.95 (7.52)	21.29 (8.0)	22.62 (7.85)	1.94
Experience in AASs (years)	16.62 (3.58)	26.4 (6.50)	14.65 (4.36)	19.22 (7.14)	20.91 (6.57)	14.35 (6.78)	17.63 (7.42)	2.01
Household assets (Birr) ^a								
< 5,000	43.33	51.67	36.67	43.89	72.50	46.25	59.38	2.88 ^{***}
5,000 – 10,000	33.33	36.67	10.00	26.67	5.00	27.50	16.25	2.34
10,000 – 15,000	3.33	11.67	26.67	13.89	12.50	15.00	13.75	0.04
15,000 – 20,000	1.67	0.00	18.33	6.67	3.75	5.00	4.38	0.92
> 20,000	18.33	0.00	8.33	8.89	6.25	6.25	6.25	0.91

(continued)

Table 1.3 (Continued)

Variable	(1) FTC areas				(2) Non-FTC areas			(3) ttest ^b
	<i>Ifa Oromia</i>	<i>Adele Waltaha</i>	<i>Biftu Geda</i>	(1a) Total	<i>Fendisha Lencha</i>	<i>Ifa Bate</i>	(2a) Total	(1a) & (2a)
Land size (ha)	0.88 (0.35)	0.49 (0.15)	0.80 (0.26)	0.72 (0.31)	0.42 (0.21)	0.82 (0.33)	0.62 (0.34)	2.87
Certified land ^a	100.00	43.33	1.67	48.33	76.25	75.00	75.63	5.35 ***
Value (Birr) of livestock ^a								
< 10,000	0.00	35.00	0.00	11.67	35.00	16.25	25.63	3.37 ***
10,000 – 30,000	16.67	63.33	6.67	28.89	53.75	42.50	48.13	3.71 ***
30,000 – 60,000	78.33	1.67	63.33	47.78	10.00	38.75	24.38	4.59
> 60,000	5.00	0.00	30.00	11.67	1.25	2.50	1.88	3.57
Household farm income (Birr)	18,185.18 (6,587.74)	14,913.68 (14,133.02)	28,013.48 (9,700.42)	20,370.78 (11,929.76)	20,851.92 (11,259.10)	52,477.76 (45,101.54)	36,664.84 (36,404.45)	5.67 ***
Off-farm employment ^a	56.67	1.67	8.33	22.22	23.75	21.25	22.5	0.06
Microfinance ^a	11.67	8.33	8.33	9.44	1.25	7.50	4.38	1.83
Membership in <i>iddir</i> ^a	98.33	90.00	100.00	96.11	93.75	43.75	68.75	7.22
Network (no. of people)	4.43 (1.64)	5.53 (0.57)	5.7 (1.42)	5.22 (1.40)	6.99 (2.24)	6.35 (4.78)	6.67 (3.73)	4.83 ***
Access to basic facilities ^a								
Daily market	1.67	0.00	73.33	25.0	0.00	7.50	3.75	5.72
Drinking water	85.00	0.00	100.00	61.67	42.50	92.50	67.5	1.12
Human health center	18.33	5.00	0.00	7.78	5.00	10.00	7.5	0.10
Veterinary	100.00	11.67	95.00	68.89	0.00	1.25	0.6	18.36
Sample size	60	60	60	180	80	80	160	

Standard deviations in parenthesis.

*** and ** denote significance at 1% and 5% level, respectively.

^a Proportion of households possessing the specific characteristics.

^b Absolute values of t-test statistic (between FTC areas and non-FTC areas).

In relation to household asset endowment (i.e., value of productive assets, household goods, and consumer durables), there is significant variation between the two groups only in the category *assets < 5000 birr/household*. Overall, however, asset ownership is not significantly different in the two areas ($t = 1.73$, $p = 0.958$), although respondents in the FTC areas are slightly better-off compared to those in the non-FTC areas. With respect to the value of livestock possessed by the respondents, there is significant variation only in the category *value of livestock < 30,000 birr/household*. In general, however, the value of livestock possessed is not significantly different in the two areas ($t = 7.61$, $p = 1.000$), though respondents in the FTC areas are comparatively better-off, on average. Regarding household annual farm income, the non-FTC areas outperform the FTC areas. This difference is due mainly to the fact that farmers in the non-FTC areas produce and sell more high-value cash crops, the dominant of which is Chat (*Catha edulis*).

With this brief discussion on the characteristics which are significantly different in the two areas, in the following section, the regression results pertaining to the determinants of on-farm training and demonstration in FTC and non-FTC areas are presented and discussed.

1.4.2 Poisson Regression and Negative Binomial Regression Estimation Results

The Poisson regression and the negative binomial regression results – estimated on *Stata 11* – for the intensity of involvement in on-farm training and demonstration are given in Table 1.4. The results are presented in two categories for both FTC and non-FTC areas: on-farm training and on-farm demonstration. Whereas the FTC sub-sample refers to the respondents drawn from PAs which have fully functional FTCs, the non-FTC sub-sample consists of respondents who reside in PAs which do not have any fully functional FTCs at the time of the survey. The analysis at the

sub-sample level was necessary in order to compare on-farm AAS provisions in FTC and non-FTC areas.⁹ Note that the overall significance of categorical variables in the regression models, i.e., education, household size, value of livestock, FTC-based training, and FTC-based demonstration, is tested, with the results presented in Table 1.5.

i. Extent of Involvement in On-farm Training in FTC and non-FTC Areas

Here, first the determinants of farmers' intensity of participation in on-farm training in FTC areas are presented, followed by the issue of on-farm training in non-FTC areas, and finally, a comparison between the two.

The intensity of farmers' participation in on-farm training in the FTC areas is estimated using the Poisson regression. The Wald χ^2 test (which is equivalent to the likelihood-ratio (LR) test for the negative binomial regression) shows that the model is statistically significant overall (Wald $\chi^2_{22} = 187.38, p = 0.000$). This statistic is calculated as negative two times the difference of the log likelihood for the null model and the fitted model, where the null model corresponds to the last iteration from fitting constant-only model. From the Poisson regression coefficients, it is apparent that physical capital (household asset endowment), financial/economic variables (farm income, microfinance credit), and access to services (animal health care, experience in AASs, FTC-based training) positively and significantly influence the intensity of farmers' involvement in on-farm training in these areas. However, value of livestock possessed by the respondents (another kind of physical capital) has a negative relationship with the number of on-farm training. Human capital (such as, age, education, and household size) and social capital (network,

⁹ In the FTC areas, the farmers have access to an FTC-based AAS provision (such as FTC training and demonstration) in addition to the on-farm/general AASs. However, in the non-FTC areas, the farmers have access only to the on-farm AASs. Therefore, pooling these samples together was not a feasible strategy since there are the FTC-based AASs only in the FTC areas.

membership in *iddir*) have no significant role in explaining the number of on-farm training in FTC areas. Concerning the geographic distribution of the respondents, whereas residing in *Adele Waltaha* PA increases the expected log count of training, living in *Biftu Geda* PA decreases the log count of training (compared to the reference PA, i.e., *Ifa Oromia*, in both cases).

In the non-FTC areas, the extent of participation in on-farm training is estimated using the negative binomial regression. The LR χ^2 test – of whether all the coefficients of the explanatory variables in the model are simultaneously equal to zero – indicates that the model as a whole is statistically significant (LR $\chi^2_{20} = 126.44$, $p = 0.000$). The coefficient estimates reveal that human capital (education and household size), financial capital (farm income), social capital (membership in *iddir*), and access to services (years of participation in AASs) positively and significantly affect the intensity of farmers' participation in on-farm training. However, physical capital (land holding size) and financial capital (access to microfinance credit) are negatively associated with the number of training. Moreover, the analysis makes it clear that farmers living in *Ifa Bate* PA have increased expected log count of training compared to those residing in the reference PA (i.e., *Fendisha Lencha*).

Table 1.4 Poisson and Negative Binomial Regression Results (2013)

	FTC areas		Non-FTC areas	
	Training ^b	Demonstration ^b	Training ^a	Demonstration ^a
<i>Human capital</i>				
Age (years)	-0.002 (0.01)	-0.06 (0.02) ***	-0.02 (0.02)	-0.03 (0.03)
Education (<i>reference</i> : No formal education)				
Primary (grades 1 – 4)	0.21 (0.19)	-0.002 (0.20)	0.89 (0.28) ***	0.90 (0.32) ***
Primary and lower secondary (5 – 8)	0.05 (0.14)	-0.23 (0.15)	0.74 (0.28) ***	0.91 (0.32) ***
Secondary (9 – 12)	0.17 (0.15)	-0.19 (0.25)	1.66 (0.57) ***	1.09 (0.73)
Some college/university (12+)			0.01 (0.73)	-0.67 (0.97)
Household size (<i>reference</i> : < 5 members)				
5 – 10 members	0.13 (0.22)	0.57 (0.28) **	1.08 (0.37) ***	1.00 (0.44) **
> 10 members	0.55 (0.43)	0.18 (0.80)	1.62 (0.50) ***	0.92 (0.55) *
<i>Physical capital</i>				
Value (Birr) of household assets	0.0000068 (0.0000033) **	0.000011 (0.000005) **	-0.000006 (0.00002)	-0.000002 (0.00002)
Land size (ha)	0.20 (0.18)	0.60 (0.18) ***	-0.64 (0.38) *	-1.15 (0.43) ***
Value of livestock (<i>reference</i> : < 10,000 birr)				
10,000 – 30,000	-0.49 (0.16) ***	-0.16 (0.16)	-0.53 (0.29) *	0.43 (0.36)
30,000 – 60,000	-0.56 (0.20) ***	-0.32 (0.26)	-0.23 (0.35)	0.39 (0.41)
> 60,000	-0.46 (0.26) *	-0.18 (0.36)	-1.21 (0.87)	-0.80 (1.02)
<i>Financial capital</i>				
Farm income (birr) per year	0.0000058 (0.0000030) **	-0.000010 (0.000005) **	0.000007 (0.000004) *	0.000008 (0.000005) *
Off-farm employment	0.03 (0.12)	-0.32 (0.13) **	0.17 (0.27)	0.09 (0.34)
Microfinance	0.30 (0.11) ***	0.08 (0.13)	-1.65 (0.53) ***	0.01 (0.50)
<i>Social capital</i>				
Network	0.01 (0.03)	0.11 (0.04) **	0.005 (0.02)	-0.01 (0.03)
Member of <i>iddir</i>	0.31 (0.21)	0.12 (0.27)	0.98 (0.40) **	0.88 (0.45) **
<i>Access to services</i>				
Access to animal health center	0.75 (0.16) ***	0.34 (0.27)	1.61 (1.11)	2.03 (1.18) *
Experience in AASs (years)	0.02 (0.01) *	0.02 (0.01)	0.07 (0.02) ***	0.04 (0.02) *

(continued)

Table 1.4 (Continued)

	FTC areas		Non-FTC areas	
	Training ^b	Demonstration ^b	Training ^a	Demonstration ^a
FTC-based training ^c (reference: 1 training/year)				
2 – 3 training/year	0.01 (0.18)			
> 3 training/year	0.90 (0.28) ***			
FTC-based demonstration (reference: 1 demons./year)				
2 – 3 demons./year		0.42 (0.17) **		
> 3 demons./year		0.35 (0.25)		
PA/Kebele reference:	<i>Ifa Oromia</i>	<i>Ifa Oromia</i>	<i>Fendisha Lencha</i>	<i>Fendisha Lencha</i>
<i>Adele Waltaha</i>	0.87 (0.32) ***	– 0.05 (0.32)		
<i>Biftu Geda</i>	– 0.76 (0.21) ***	– 0.82 (0.28) ***		
<i>Ifa Bate</i>			3.53 (0.28) ***	4.12 (0.36) ***
_cons	– 1.45 (0.59) **	0.26 (0.74)	– 3.10 (0.89) ***	– 3.30 (1.02) ***
Number of observations	180	180	160	160
Log Likelihood	– 218.04	– 179.58	– 335.57	– 274.54
Pseudo R2	0.17	0.15	0.16	0.19
Alpha (α)	2.35e–18 ^c	1.61e–09 ^c	0.87	0.92
Likelihood-ratio test of $\alpha = 0$	0.00 ^c	0.00 ^c	176.44 ***	111.08 ***
Goodness-of-fit χ^2 test ^d	97.98	95.59		

Standard errors in parenthesis.

***, ** and * denote significance at 1%, 5%, and 10% level, respectively.

^a Negative binomial regression results.

^b Poisson regression results (with robust standard errors).

^c Based on negative binomial regression Likelihood-ratio test of $\alpha = 0$: $\text{chibar2}(01) = 0.00$ $\text{Prob} \geq \text{chibar2} = 1.000$. This reinforces that the Poisson distribution fits the data better than the negative binomial distribution.

^d Since the χ^2 values are not statistically significant, this reinforces the fact that the Poisson regression is appropriate for this set of data.

^e The year of FTC-training was also accounted for (since the sample comes from 2008 and 2009). However, the dummy indicator for year of training shows no statistically significant effect on the number of on-farm training and/or demonstration.

ii. A Comparative Discussion on the Determinants of On-farm Training in FTC and non-FTC Areas

In the above section, the various characteristics that influence the extent of farmers' participation in on-farm training in FTC and non-FTC areas were highlighted. In this section, a detailed discussion on them is provided. The first aspect of comparison relates to the role of physical capital. In this study, two types of physical capital (value of asset and value of livestock) were found to affect participation in training in FTC areas, but have no significant effect in non-FTC areas.¹⁰ Regarding household asset endowment, although its effect is positive and significant only in the FTC areas, it has a negative and insignificant effect in non-FTC areas. This suggests that an increase in the value of asset possessed increases the expected log count of training in FTC areas. This implies that asset-constrained farmers, i.e., the majority of households with low asset portfolios in the present study case (44% in FTC areas), may not frequently take part in on-farm training. Though on the basis of this observation one may deduce that the on-farm training is biased against households with low levels of asset, the importance of productive assets in translating knowledge and skills obtained from training sessions into productivity and income is indispensable. To this end, making a concerted effort to help such households build their asset base should be considered, in addition to giving priority to them in the recruitment process.

Another form of physical capital represented by the value of livestock (e.g., cattle, small ruminants, chicken, etc.) possessed by the households is also an important determinant of the number of on-farm training. Although it is overall a statistically significant predictor in FTC areas only (Table 1.5), all its categories included in the regression indicate that the expected log

¹⁰ Land holding size (another kind of physical capital) plays some role in non-FTC areas. More discussion on this is given in relation to on-farm demonstration.

count decreases as the value of livestock increases. For instance, the expected log count of training for the category *value of livestock 30,000 – 60,000 birr* is 0.56 lower than that for the reference category. This suggests that households with a few livestock possession are encouraged to frequently take part in on-farm training. From the descriptive statistics, however, it is evident that about two-third of the respondents possess livestock whose value is between 10,000 and 60,000 birr. Historically, AASs have been biased against the livestock sector in the country (Belay, 2002; 2003) and more so is with households possessing higher valued livestock. On the other hand, it may be the case that households possessing such livestock may not seek out for on-farm training as they are financially in a more sound position compared to the others.

Table 1.5 Overall Significance of Categorical Predictor Variables

	Training	Demonstration
<i>FTC areas (n = 180)</i>		
Education χ^2_3	2.19	2.62
Household size χ^2_2	1.65	4.35
Value of livestock χ^2_3	10.06 **	1.77
FTC training χ^2_2	14.40 ***	
FTC demonstration χ^2_2		6.38 **
<i>Non-FTC areas (n = 160)</i>		
Education χ^2_4	14.75 ***	12.59 **
Household size χ^2_2	11.24 ***	5.18 *
Value of livestock χ^2_3	4.71	3.36

***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Note: For example, χ^2_4 means the four degrees-of-freedom chi-squared test.

An important aspect related to livestock production and management is animal health care/veterinary service. It is found in this study that having access to veterinary services in FTC areas significantly increases the expected log count of training compared to not having such

access.¹¹ This may be because farmers who interact with veterinary service providers (i.e., development agents at PA level) are more likely to establish good relationships with them and may end up being recruited for several on-farm training. The role of development agents in selecting which farmers to work with is a long documented reality in the country (e.g., Belay, 2003). It has to be noted here that three development agents (trained in plant science, animal science, and natural resource management) reside in each PA (especially in those PAs where there are established and functional FTCs). These agents are responsible to organize and deliver training on various aspects of crop and livestock production. For example, the agents specialized in animal science assist farmers in livestock production and management, including animal health care. There are also units in the district bureau of agriculture and rural development that provide veterinary services. Therefore, making contact with such agents is more likely to increase the probability of a farmer to frequently take part in on-farm training.

The second aspect of comparison concerns the role of economic/financial capital (farm income and microfinance credit) in FTC and non-FTC areas. To start with the common element, it is found that farm income has a positive and significant association with the number of on-farm training in both FTC and non-FTC areas, implying that high income farmers receive preferential treatment. This is clearly not in line with the aim of AASs in the study area. Although it is not the intention of the investigation to suggest that high income farmers should not be involved in AASs, smallholder farming cannot be improved by focusing much on high income earning farmers. Concerning credit, the study shows that whereas access to microfinance credit significantly increases the expected log count of training in FTC areas, it decreases the log count

¹¹ Access to veterinary services is not a significant predictor of training in non-FTC areas probably because a very few proportion of the respondents (6%) have access to such services. The corresponding figure for the FTC areas is 69%.

in non-FTC areas (compared to not having such access). In general, however, since only a few proportion of the respondents (approximately 9% and 4% in FTC and non-FTC areas, respectively) are currently beneficiaries of organized microfinance services, promoting the farmers' access to microfinance credit should be encouraged, given its importance to achieving increased agricultural productivity and poverty reduction in the study area (Brehanu & Fufa, 2008). Microfinance credit is also shown to have a causal effect on increased consumption and housing improvements in the country (Berhane & Gardebroek, 2011). On top of this, since a recent study suggests that many farmers in the study area also save (though in informal ways) irrespective of their level of household income (Teshome et al., 2013), a mechanism that combines efficient credit provision and saving can be highly relevant. This calls for increased coordination between AAS providers and credit and saving institutions operating at various levels.

Human capital (education and household size) and social capital (membership in *iddir*) also play significant roles in predicting the intensity of farmers' participation in training in non-FTC areas but not in FTC areas. Concerning educational attainment of the respondents, the results show that the expected log counts (of on-farm training) for farmers educated up to secondary school level is, in general, higher than that for the reference group (i.e., no formal education). For example, the expected log count for farmers with primary school level is 0.89 higher than that for *no formal education* category. This may be because some contents of the on-farm training require certain level of education from the participants in order to grasp them easily. However, beyond secondary school, the expected log count decreases compared to that for the reference group, probably because the more educated members of the community prefer to work in off-farm employment. Nonetheless, since roughly one-third of the sample comprises farmers with no

formal education, there is a need to explicitly consider this issue in future targeting and recruitment of farmers for on-farm training.¹²

The effect of household size on the number of training is also positive and significant in non-FTC areas, indicating that the more the number of household members, the higher the log counts of on-farm training. For instance, the expected log count for households with *5 – 10 members* is 1.1 higher than that for the reference category. In terms of magnitude, however, it is the category *> 10 members* that has the highest expected log count. Since the majority of households in the study (about 79%) fall in the category *5 – 10 members* and since large family size is a common phenomenon in eastern Ethiopia (e.g., Emanu & Hadera, 2007; Haji, 2007), addressing such households with frequent training is important, given the challenges they face in fulfilling basic needs of the household. On the other hand, since large households can get support from some of their members (e.g., members who earn income from employment elsewhere or educated members who support the household through remittances), giving priority to them always may not be a singularly effective rural development strategy. There should be a balance between size of a household and its capability in terms of generating sufficient and sustained income.

Regarding participation in informal local organizations, the result shows that being a member of an *iddir* in non-FTC areas increases the expected log count of training by 0.98 compared to not being a member. The *iddir* is an indigenous social insurance institution developed to help members during funeral (Pankhurst & Mariam, 2000; Dercon et al., 2006; LeMay-Boucher, 2009). An *iddir* discharges its responsibilities through money obtained from its members (in the

¹² Such considerations are especially relevant in FTC areas where the proportion of farmers with no formal education is much higher compared to that in the non-FTC areas. However, since there is a parallel AAS provision in these areas, i.e., the FTC-based AASs, this requires improved coordination between the two, so as to avoid wastage of scarce resources and inefficiencies.

form of membership fee and a fixed sum of monthly contributions). In an event when a member or a relative/family of a member passes away, the *iddir* authorities, based on the rules and regulations set out during the establishment of the *iddir*, provide a stipulated amount of money to cover expenses associated with funeral. Besides this, an *iddir* is also an important social platform where members come together and share information. As members of the community, development agents also belong to some *iddir* in their locality and interact informally with community members. This encounter, probably, may have helped members of an *iddir* to participate frequently in on-farm training compared to non-members. Although the proportion of households belonging to some form of *iddir* is higher in FTC areas compared to the non-FTC areas (96% and 69%, respectively), membership in *iddir* is not a significant determinant of intensity of participation in training in FTC areas. This may be due to the composition (profile of the members) of the *iddir* in those areas (e.g., no development agents or local leaders being a member of such *iddir*).

Another point of comparison relates to the issue of access to services as characterized by the number of years spent in AASs and participation in FTC-based training. In relation to experience in AASs, although the coefficient has the same sign in both areas, its magnitude and level of significance suggest that a one year increase in AAS experience increases the expected log count of training more in the non-FTC areas than in the FTC areas. Although the difference in experience in AASs is not statistically significant between the two areas (Table 1.3), farmers in the non-FTC areas have a relatively lesser experience in AASs than farmers in FTC areas, on average. This result also suggests that development agents in non-FTC areas prefer to work more with experienced farmers than inexperienced/less-experienced ones. Such a tendency was also documented in earlier studies in the country (e.g., Belay, 2003). Although working with more

experienced farmers is useful in relation to accelerated adoption and diffusion of agricultural innovations, younger and less experienced farmers should not be left out.

In fact, the providers of on-farm AASs should aim to build up the confidence of younger farmers to try out new technologies and practices. One way of doing so can be by providing frequent orientation training to these farmers. Such targeting can retain the rural youth of productive age in rural areas and reduce the migration of them to urban areas in search of employment. The issue of diffusion of agricultural technologies is especially important considering the generally low rates of adoption of modern varieties of staple food crops (e.g., sorghum) in the study area (Cavatassi et al., 2011). This may be associated with the ever increasing expansion of the most lucrative Chat crop in the area. One practical suggestion regarding adoption and diffusion can be enhancing farmers' participatory evaluation of sorghum (as was tried in the *Kile-Bisidimo* plains of eastern Ethiopia) and encouraging farmer-to-farmer exchange of seeds (Mulatu & Belete, 2001). The participatory evaluation can be arranged in a training session where farmers, assisted by experts, evaluate the various traits of modern varieties in relation to the traditional ones.

Finally, in this investigation, the effect of FTC-based training on the intensity of farmers' participation in on-farm training is also considered. This is done in order to account for the effect participation in the FTC-based AASs has on the on-farm/general AASs in FTC areas. The result shows that the more the number of FTC training a farmer receives, the better the involvement in on-farm training. From this, it is apparent that there is a lack of coordination between the FTC-based training and on-farm training in terms of targeting – there is some redundancy in AAS provision in FTC areas.

iii. Extent of Involvement in On-farm Demonstration in FTC and non-FTC Areas

The basic notion of on-farm demonstrations is to help farmers to acquire technical skills and knowledge related to the management of their farm enterprises (through method-demonstration), and to show them the performance of improved technologies and practices (through result-demonstration), so that they may be convinced about the technologies and develop interest to try them on their own farm. The expectation is that having developed interest in the new technology, farmers can go as far as trying and adopting the technology as well as diffusing it to fellow farmers in their vicinity. Demonstrations are also organized in such a way that farmers in one area share their experiences (through field visits) with farmers in another area, thereby encouraging farmer-to-farmer flow of information. In Ethiopia, besides on-farm advice/visit and training, demonstration has been a principal component of on-farm AASs (Belay, 2003; Belay & Abebaw, 2004).

In the present study, the extent of farmers' involvement in on-farm demonstration in FTC and non-FTC areas is modeled using the Poisson regression and the negative binomial regression, respectively. The Wald χ^2 test for the Poisson regression suggests that the model is statistically significant overall (Wald $\chi^2_{22} = 138.42, p = 0.000$). The same is true for the non-FTC areas (LR $\chi^2_{20} = 125.54, p = 0.000$). The coefficient estimates for the FTC areas reveal that physical capital (value of household asset and land holding), social capital (networks), and access to services (FTC-based demonstration) have a positive and significant effect on the intensity of participation in on-farm demonstration; financial capital/economic variables (farm income and off-farm employment) and human capital (age) have negative effects. In addition to these, farmers residing in *Biftu Geda* PA have a lower expected count of on-farm demonstration compared to the reference PA for the FTC areas.

In the non-FTC areas, whereas human capital (education and household size), social capital (membership in *iddir*), financial capital (farm income), and access to services (veterinary and years of participation in AASs) play a positive and significant role, physical capital (land holding) has a negative relationship with the intensity of farmers' involvement in on-farm demonstrations. Furthermore, being in *Ifa Bate* PA is found to increase the expected log count of on-farm demonstration compared to living in the reference PA.

iv. Comparing Intensity of Access to On-farm Demonstration in FTC and non-FTC Areas

The main focus of this section is to discuss the characteristics exhibiting significant effects on farmers' participation in on-farm demonstration but are not discussed in the previous sections. The first aspect of comparison revolves around the relationship between human capital (household size, age, education) and extent of participation in on-farm demonstration. Concerning household size, it is found that household size positively predicts the extent of farmers' participation in on-farm demonstration in both areas, although its effect is overall significant only in non-FTC areas. Similar to the case with on-farm training in these areas, the demonstration services are targeting households with large family size. Since about 90% of the respondents in the non-FTC areas have a family size of more than five, such targeting may be beneficial to these households. Nonetheless, there should be a mechanism to integrate small households in such services.

Regarding the role of respondents' age, it is evident from the regression coefficient that the expected log count of demonstration decreases for every one year increment in age (both in FTC and non-FTC areas, although it is significant only in the former). This suggests that younger farmers are in good position of getting more demonstration services. Since many of the result-

demonstration services are arranged in demonstration plots that are located somewhat far from a farmer's plot, compounded with poor transportation in rural areas, younger farmers can easily walk to the demonstration plots and witness the performance of an improved technology or practice. Education has a positive effect only in the non-FTC areas probably because the respondents drawn from these areas have a comparatively better educational attainment. Hence, in the eyes of the organizers of on-farm demonstration, such farmers might be the preferred ones compared to those with lower or no education.

The second aspect of comparison relates to physical capital (value of household asset and land holding). The coefficient of household asset is positive and significant in FTC areas but it is negative and not significant in non-FTC areas, suggesting that household asset levels matter in participating in demonstrations in FTC areas. A similar observation was also noted in relation to intensity of participation in training in FTC areas. In relation to the area of land cultivated by the farmers, the coefficients indicate that for each one hectare increment in land holding, the expected log count of demonstration increases in the FTC areas, but it decreases in the non-FTC areas.¹³ In the latter, this means that farmers with small land holding have better counts of on-farm demonstration. This is contrary to previous work, which shows that development agents frequently work with farmers who have large land holdings in order to use part of the land for demonstration purposes (Belay, 2003). However, this earlier claim appears to hold in FTC areas.

In general, throughout the country, land is state-owned and farmers have only usufruct rights on land allocated to them by local authorities. However, through some arrangements, farmers can have access to additional cultivable land. For instance, in eastern Ethiopia, this is achieved

¹³ Average land holding size is not significantly different in the two areas (0.72 and 0.62 hectares in FTC and non-FTC areas, respectively).

through renting (e.g., sharecropping arrangements) and land borrowing (Belay & Manig, 2004). Very recently, the government has started issuing land certificate, which aims to strengthen tenure security and investments in the land (Deininger et al., 2008; Holden et al., 2009, 2011). In this study, it is found that 48% and 76% of the respondents' land holdings in FTC and non-FTC areas, respectively, are certified by authorities. Nevertheless, the average land holding in the region has been declining over the years mainly due to increased demographic pressure. This has resulted in individual farm units that are generally too small to be economically viable. Access to cultivated land is also inextricably linked to market integration, availability of credit, and effective AASs (Belay & Manig, 2004).

The third aspect of comparison involves the influence of economic/financial capital (farm income and off-farm employment) on farmers' participation in on-farm demonstration. Regarding the effect of household farm income, whereas an increase in farm income negatively influences the number of demonstrations in FTC areas, it has a positive effect in non-FTC areas. This may suggest that as household income improves, farmers in the FTC areas are unlikely to seek out for more on-farm demonstration services. An alternative explanation can be that development agents in FTC areas choose to work more with those farmers who earn lower amount of farm income, which is consistent with the very aim of AASs in these areas. Note that there exists a statistically significant difference in farm income between the two areas, where farmers in the non-FTC areas earn more, on average, compared to those in the FTC areas (Table 1.3). This difference is attributable to the fact that farmers in the non-FTC areas produce and market high quality high value cash crops, the dominant one being Chat.

Chat (*Catha edulis*) is a perennial crop whose leaves and tender stems are chewed for their stimulating effects. Its production in the study area is usually undertaken by intercropping it with

sorghum, maize and other crops. In spite of the government's premeditated attempt to weaken the production and expansion of this narcotic tree/bush crop, it has been dominating the Hararghe highlands of eastern Ethiopia as a cash crop. Hence, it is not uncommon to observe farmers shifting their limited resources to the production of this crop (Mulatu & Kassa, 2001). The shift, as documented by these authors, increased farm income and promoted crop-livestock integration. The authors go as far as asserting that the production of chat improves household food security and enhances sustainability of smallholder mixed farming systems. In another study analyzing the factors contributing to the expansion of this crop in the Hararghe highlands, Tefera et al. (2003) document the significant power of market incentives that lure farmers into this crop production.

The issue of market access – to sell crop and livestock products as well as buy production inputs associated with the enterprises – is also a relevant aspect in rural development. Even though there is no significant effect of market access on the expected log count of demonstration (or training for that matter) in the present study, only a small proportion of the respondents have such access (25% and 4% in FTC and non-FTC areas, respectively). In general, however, access to market can play an important role in the adoption of improved agricultural technologies and practices. For example, in a study investigating whether adoption of improved sorghum varieties can potentially serve as a risk reduction mechanism in eastern Ethiopia, researchers show that access to market, among others, significantly influences farmers' decision to adopt modern sorghum varieties (Cavatassi et al., 2011). Hence, linking smallholder farmers to potential input/output markets through, for example, improved road and communication infrastructures should be integrated in rural development efforts. Moreover, the dissemination of market information during training and demonstration sessions can be a useful approach in order to keep

farmers up-to-date with not only the prevailing market circumstances, but also the market scenarios forecasted for the future.

In relation to income from off-farm employment, the analysis shows that it significantly predicts the number of demonstrations in FTC areas (although the proportion of households involved in such employment is roughly the same in FTC and non-FTC areas). The coefficient shows that an increased participation in off-farm activities reduces the intensity of involvement in on-farm demonstrations in FTC areas. From the side of the farmer, this may be due to the fact that the farmer may not have enough time to frequently take part in demonstrations since his/her time is spent elsewhere in off-farm employment. From the point of view of development agents, they might deliberately avoid such farmers and support the ones who are devoted full time in their farm. More generally, however, since off-farm employment requires households to adjust their employment portfolio to changing circumstances and challenges (Bezu & Barrett, 2012; Bezu & Holden, 2014), it may pose difficulty to frequently participate in demonstrations. This is more so when a household is involved in off-farm activities very far from his/her farm or homestead and unable to coordinate farm and off-farm tasks.

A fourth aspect of comparison considers the role of social capital (networks and membership in *iddir*). The analysis reveals that networks and membership in *iddir* are important in FTC and non-FTC areas, respectively. Concerning network size, i.e., the number of people/farmers in one's network, there is a significant variation between the two areas, with the non-FTC areas having more network members, on average, compared to the FTC areas (Table 1.3). However, it appears from the regression coefficients that an addition of one network member increases the expected log count of demonstration in FTC areas and not in non-FTC areas. This may suggest that although the respondents in the non-FTC areas have large network size compared to the FTC

ones, large size may not necessarily mean greater influence. For example, a network with a few yet very influential, experienced, and resourceful persons can exert greater influence on its members than a network comprising many individuals with less exposure, information, and external contact. In general, social relationships among network members are important channels of information flow, which can play significant roles in agricultural technology adoption and diffusion (Maertens & Barrett, 2013; Jensen et al., 2014; Krishnan & Patnam, 2014).

Finally, similar to the effect of FTC training on the expected log count of on-farm training, the study shows that participating in FTC-based demonstration positively and significantly influences the expected log count of on-farm demonstration in FTC areas. This implies the existence of a weak coordination between FTC-based services and on-farm AASs in the area, because the same household is targeted by both. The effect of access to other services (animal health care) and experience in AASs on intensity of participation in on-farm demonstration is also positive and significant in non-FTC areas.

1.5 Conclusion and Recommendation

In the present study, an investigation is undertaken on the predictors of smallholder farmers' intensity of participation in on-farm AASs (training and demonstration) in Haramaya district, Ethiopia. By collecting primary data from a sample of 340 farm household heads and employing both the Poisson regression and the negative binomial regression models, the study shows that a host of factors affect farmers' extent of involvement in these services. Specifically, whereas financial capital (farm income, credit), physical capital (value of livestock, value of household asset), and access to services (FTC-based training, veterinary, experience in AASs) are significant predictors of on-farm training in FTC areas, human capital (education, household size), social capital (membership in *iddir*), financial capital (farm income, credit), physical capital (land holding size), and experience in AASs are important determinants of participation in training in non-FTC areas. Concerning demonstration in FTC areas, human capital (age), physical capital (asset, land), financial capital (farm income, off-farm employment), social capital (networks), and access to services (FTC-based demonstration) have a significant effect on intensity of farmers' participation. In non-FTC areas, human capital (education, household size), physical capital (land), financial capital (farm income), social capital (*iddir* membership), and access to services (veterinary, experience in AASs) play significant roles.

Considering the determinants of intensity of participation in on-farm AASs (both training and demonstration) jointly, the following patterns are noted. First, seven out of eight (about 88%) characteristics that significantly affect intensity of farmers' participation in training in non-FTC areas also influence the extent of farmers' involvement in demonstration in these areas. The influence is so remarkable that all these covariates have the same sign (for training and demonstration). This is strong evidence that points to the conclusion that the characteristics

which affect farmers' intensity of participation in training and demonstration in non-FTC areas are largely the same. In FTC areas, however, only 50% of the variables (four out of eight) jointly affect the extent of farmers' participation in both training and demonstration. This suggests that, unlike the case in non-FTC areas, the factors responsible for farmers' differential access to on-farm AASs in FTC areas vary greatly.

Another interesting pattern also emerges when comparing the determinants of training in FTC and non-FTC areas as well as the determinants of demonstration in the two areas. In the case of training, except for financial capital (farm income, microfinance credit) and experience in AASs, a different set of factors affect the farmers' intensity of participation in training in FTC and non-FTC areas. Even then, whereas the effect of credit is positive in FTC areas, it has a negative association with training in non-FTC areas. Concerning demonstration in both areas, except for physical capital (land holding size) and financial capital (farm income), there is no similarity between the factors affecting intensity of participation in demonstration in the two areas. Even in these two cases, the effects are mixed. These comparisons indicate that the determinants of the extent of farmers' participation in training and/or demonstration vary greatly across the two areas.

On the basis of these findings, therefore, the following recommendations are suggested. First, since there exists a huge difference in the types of covariates explaining farmers' differential participation in training and/or demonstration in the two areas, an area-specific targeting and selection criteria should be followed by the implementers of AASs in the study area. Second, if the ultimate aim of AASs in the area is improving smallholder farmers' intermediate and final outcomes, the possibility for a reorientation of service provision should be explored. This can be achieved, for example, by focusing training efforts on asset-constrained and low-income

households (in FTC areas). Furthermore, since the results suggest the existence of redundancy of service provision in FTC areas, there is a need to improve coordination between the FTC-based services and the on-farm AASs. This can be achieved through joint planning, implementation, and evaluation of AAS activities. Likewise, since the findings indicate that farmers residing in *Biftu Geda* and *Fendisha Lencha* PAs have lower expected counts of both training and demonstration, the activities of AASs need to be strengthened in these PAs.

Although the predictors of intensity of participation in on-farm AASs are shown in this study, the relationship between intensity of participation and outcome variables (such as farm income, food security, poverty, consumption, etc.) is not investigated. Even in the wider evaluation literature, the impact of intensity of access to AASs is not very well analyzed. Most of the available studies collapse intensity into 0/1 (non-participants versus participants) due to methodological difficulty or neglect of the issue. Likewise, the current investigation did not undertake an enterprise-level analysis (i.e., considering the three sectors independently – crop, livestock, and natural resource management) of the factors affecting intensity of access to on-farm training and demonstration. Such an analysis can help in setting up a sector-specific targeting and selection criteria for AASs. Finally, since the focus of the present study is on the supply-side of on-farm AAS provision in the study area, future research can look into the demand-side. In general, these are some indications where future research activities in improving AASs can focus on.

2 Improving Smallholder Farmers' Income through Farmers' Training Centers: an Impact Evaluation in Haramaya District, Eastern Ethiopia ^{*†}

2.1 Introduction

About 75 percent of the world's poor people live in rural areas, and most of them depend – directly or indirectly – on agriculture for their livelihoods (Mogues et al., 2009). Although agriculture is the engine for development in these areas, it is often constrained by issues of access to relevant technologies and improved practices, institutional weaknesses, and deep-rooted problems with the organization and management of agricultural education, research, and advisory systems. Consequently, agrarian communities and agricultural systems in many countries operate under severe limitations and face major stumbling blocks to the use of knowledge, management skills, and innovation for development (Asenso-Okyere et al., 2008). To deal with these challenges and to foster agricultural growth especially in rural areas, agricultural advisory services (AASs) are brought to the center of today's international development discourse (Birner et al., 2009).

* Paper presented at: a) Società Italiana degli Economisti dello Sviluppo (SITES)/Italian Development Economists Association (IDEAs). SITES-IDEAS First Annual Conference. 11-12 Sep. 2014, Florence, Italy. b) Associazione Italiana per gli Studi Africani/Italian Association for African Studies (ASAI). ASAI Third Conference “Africa on the Move” organized in partnership with the Interdisciplinary Centre for African Studies of the University of Macerata. 17-20 Sep. 2014, Macerata, Italy. c) Italian Association of Agricultural and Applied Economics (AIEAA). Third AIEAA Conference “Feeding the Planet and Greening Agriculture: Challenges and Opportunities for the Bio-Economy”. 25-27 June 2014, Porto Conte Research Center, Alghero (SS), Italy.

† Preliminary version published as a working paper at Natural Resources, Agricultural Development and Food Security International Research Network (NAF-IRN). NAF International Working Paper Series No. 14/7, Year 2004. <http://economia.unipv.it/naf/>

AASs can be defined as the entire set of organizations involved in supporting and facilitating the access of farmers to technical knowledge, skills, improved technologies, credit, markets, and other rural services (Birner et al., 2006; Swanson, 2008). The renewed interest in AASs is linked to a ‘rediscovery’ of the role of agriculture for pro-poor development, as emphasized in the World Development Report 2008 (World Bank, 2007), and the recognition that these services are a critical means for achieving local development (Birkhaeuser et al., 1991). Nevertheless, the benefits that farmers obtain from AASs – and the resulting impact of the services – depend, to a great extent, on their level of participation, which in turn is influenced by their direct and indirect access to the services.

In the case of Ethiopia, the organized and formal provision of these services dates back to the early 1950s when the then Imperial Ethiopian College of Agriculture and Mechanical Arts (IECAMA), now called Haramaya University, was established. Since then, AASs have passed through a number of stages (for details, see Belay, 2003; Gebremedhin et al., 2006). Nonetheless, smallholder productivity remains low because of scarce, inaccessible, and expensive production inputs, extremely limited market and credit (IFPRI, 2010), and poor provision of advisory services in terms of quality and intensity. In the face of these challenges and in order to provide effective, coordinated, and decentralized AASs at the *woreda* (district)¹⁴ level, the provision of AASs has recently been based on the use of Farmers’ Training Centers (FTCs).

The FTCs are expected to serve as hubs for farmers to receive advisory services and information, training, and demonstrations on improved and sustainable farm management practices (Gebremedhin et al., 2006; Moguees et al., 2009; IFPRI, 2010). Currently, there are about 8,500

¹⁴ AAS provision falls under the agricultural development subsector and is further subdivided into services on crops, livestock, and natural resources management (Gebremedhin et al., 2006). Limited advisory service is conducted by NGOs, usually working through the *woreda*-level Bureaus of Agriculture and Rural Development (BOARD).

FTCs established at the *Kebele*¹⁵ level, with roughly 2,500 of these FTCs reported to be fully functional (MOARD, 2009a¹⁶ cited in IFPRI, 2010). Stationed at each FTC are three Development Agents (DAs) responsible for providing AASs on livestock, crop production, and natural resource management (Spielman et al. 2006; IFPRI, 2010). The agents, at least in the early stages of the establishment and functioning of the FTCs, are also responsible for the recruitment and selection of individuals and groups to take part in the FTC-based services. This is due to the fact that these FTC-based AASs are relatively new to the farming communities and not many farmers are convinced and ready to freely sign up for them. Moreover, the limited institutional capacity of the providers of the FTC-based services also required to limit the number of farmers per training cycle, for instance.

Regarding effectiveness of AASs, many scholars in the field agree that AASs provide important benefits in a range of farming systems and contexts around the globe (Birkhaeuser et al., 1991; Anderson, 2008; Swanson, 2008; Dercon et al., 2009; IFPRI, 2010; Benin et al., 2012). These researchers document the effects of AASs on knowledge, skills, technology adoption and diffusion, productivity, consumption, and poverty reduction. On the other hand, there are also studies showing negative or negligible effect of AASs. Nonetheless, though there is a vast literature on issues related to AASs, there is an acute shortage of rigorous impact evaluations to sufficiently inform policy makers and development practitioners on the causal effects of AASs (Waddington et al., 2010). In particular, in Ethiopian context, there is little systematic and careful empirical investigation regarding the impact of AASs (Dercon et al., 2009).

¹⁵ Kebele/Peasant Association is the lowest administrative division in urban and rural areas, respectively.

¹⁶ Ethiopia, MOARD (Ministry of Agriculture and Rural Development). 2009a. DAs and FTC data at national level. Addis Ababa, Ethiopia: MOARD.

Most of the available impact studies in the country focus on intermediate outcomes, such as knowledge (Efa et al., 2005), agricultural productivity (Ayele et al., 2005¹⁷ cited in Davis, 2008), and productive and technical efficiency (Alene & Hassan, 2003; Ayele et al., 2006; Thangata & Mequaninte, 2011; Elias et al., 2014). The only rigorous impact evaluation of AASs on final outcomes (i.e., poverty and consumption growth) in the country is that of Dercon et al. (2009), where the researchers show that “receiving at least one visit by agricultural adviser reduces headcount poverty by 9.8 percentage points and increases consumption growth by 7.1 percentage points” (p. 1008). However, the study examined only one of the components of AASs (namely advice/visit) and did not take into account either modular or short-term training. In general, most of the impact evaluations do not cover all parts of the country – including different agro-ecological zones and the various farming systems – and hence it is difficult to make definitive conclusions regarding the effectiveness of AASs in the country.

Most importantly, despite its crucial implication to policy making, to the best of my knowledge, there is not any careful empirical study conducted to systematically document the effectiveness of the recent FTC-based approach to AAS provision. So far, the only available study is that of Lemma et al. (2011), who investigated the extent and efficiency of utilization of physical, human, and other resources as well as training programs conducted in Improving Productivity and Market Success (IPMS) supported FTCs. However, the study concentrates only on IPMS Pilot Learning Woredas (PLWs) and does not cover the wider government-supported FTCs. In addition, and this cannot be stressed enough, the study falls short of identifying and explaining the causal links between FTC-based AASs and outcome variables, such as (agricultural) income

¹⁷ Ayele, G., Alemu, D., & Kelemework, F. (2005). *The provisions of rural services in Ethiopia: Characterization, impacts, and farmers' priorities*. Unpublished manuscript, International Food Policy Research Institute, Washington, DC.

and poverty. As a matter of fact, the FTC-based AAS provision is a somewhat recent approach, and there is very shallow literature on it (IFPRI, 2010). Therefore, recognizing the gaps in empirical investigations, scarcity of evidence of impact of FTC-based AASs (especially training), suggestions of Gebremedhin et al. (2006) and Dercon et al. (2009) on the importance of further inquiries, the present study aims at investigating the causal effect of modular training organized at FTCs on farm income of participating farmers in Haramaya district of eastern Ethiopia.

The present analysis relies on the use of a non-experimental design implemented through a field-based household survey¹⁸ conducted from May to October 2013. A total of 250 household heads from three treatment and two comparison *Kebeles*/peasant associations (PAs) of Haramaya district responded to the survey. To ensure the quality of data gathered, a number of procedures as outlined by Leones & Rozelle (1991), Nyariki (2009), and others were implemented. In particular, specific attention was paid to the selection and training of research assistants, translation of the questionnaire into local language, and pre-testing of the survey instrument. The evaluation strategy used in the analysis is propensity score matching (PSM). This is preferred to the use of randomized control trials (RCTs) or other non-experimental techniques (such as difference in difference, regression discontinuity design) primarily because of the nature of the program, data collected, and other considerations discussed in section 2.2.4.

An important issue arises in the use of PSM – the choice of variables to estimate propensity score. Therefore, in section 2.2.4, detailed discussion on this point is presented. Likewise, from the range of PSM techniques (see, for example, Khandker et al., 2010), Nearest Neighbor

¹⁸ In order for the survey to be relevant for our purpose, we constructed the survey tool based on documents obtained (with permission) from the International Food Policy Research Institute (IFPRI).

Matching, Kernel Matching (both Gaussian and Epanechnikov), and Radius Matching estimators were used following the suggestions of Bryson et al. (2002). Section 2.2.4 presents details on the choice of these matching algorithms. In addition to this, since PSM is based on the assumption of ‘*strong ignorability*’ (Rosenbaum and Rubin, 1983), this is discussed in section 2.2.4. Furthermore, since the results obtained using PSM estimators are judged reliable and robust based on balancing tests (e.g., Lechner, 1999; Sianesi, 2004; Caliendo and Kopeinig, 2008), these issues are also discussed in section 2.2.4.

The construction of the outcome variable is also an important issue addressed in this study. In reference to the policy aim of setting up FTCs (MOARD, 2009) and in accordance with the literature dealing with agriculture and allied activities in developing countries (e.g., Cunguara and Darnhofer, 2011; Davis et al., 2012; Benin et al., 2012b), farm income is adopted as an outcome variable (see section 2.3.3 for details). The definition of farm income in this study encompasses not only incomes from dominant crops (sold, consumed and/or stored) and livestock (products sold and oxen rented out), but also revenues from the otherwise neglected components in many previous studies, i.e., incomes from less-dominant crops and by-products of crop production. Moreover, each crop and livestock product is evaluated using actual prices received by the household rather than a single price (such as average price). This procedure is followed to emphasize the negotiation capacity of the respective households in getting better prices for their products (among other things discussed in section 2.3.3).

Following the above mentioned procedures, the average effect of modular training on farm income (without *chat*) is estimated to be positive and statistically highly significant, ranging from 9,557 birr/year (*Nearest Neighbor*) to 10,388 birr/year (*Kernel – Gaussian*). Moreover, by disaggregating farm income into its constituents, the study shows the effect of the modular

training on crop income and livestock income separately (as can be seen in Table 2.8). In addition, a discussion is provided on the specific role of income from *chat* on farm income and crop income in sections 2.3.3 and 2.3.4. Finally, all the matching quality checking mechanisms (section 2.3.5) indicate that the results are reasonably robust to changes in estimation methods.

2.2 Data and Methods

2.2.1 Study Area and *Kebele*/PA Selection

This study employs a non-experimental design implemented through a field-based household survey. Using a pre-tested questionnaire as a survey instrument, primary data for the purpose of impact evaluation were collected from both treatment and comparison areas of Haramaya district from May to October 2013. Whereas the ‘treatment’ areas refer to the peasant associations (PAs) with functional farmer training centers (FTCs), the ‘comparison’ areas refer to those PAs which do not have any FTC at the time of the survey. The survey questionnaire was constructed based on the research project documentations obtained (with permission) from the International Food Policy Research Institute (IFPRI). These documents and datasets are: Ethiopian Rural Household Survey (ERHS); Food and Water Security under Global Change – Developing Adaptive Capacity with a Focus on Rural Africa; Policies for Sustainable Land Management in the Ethiopian Highlands Dataset 1998-2000; and, Assessing the Potential of Farmer Field Schools (FFSs) to Fight Poverty and Foster Innovation in East Africa. In the context of impact evaluation, the survey tool used in the data collection process presented a number of questions related to household and farm level characteristics, participation in general and FTC-based advisory services, household assets, livelihood activities (crop and livestock production), and participation in (formal/informal) organizations/associations (see Appendix 8 for the details).

Regarding sampling procedures, a multi-stage sampling technique was employed to select study district, *Kebeles*/PAs, and household heads. From the 18 districts in East Hararghe Zone of Oromia Region (Figure 2.1), Haramaya district was purposively selected based on its representativeness to the major farming systems and agro-ecological zones, its proximity to the

collaborating institutions (i.e., Haramaya University and East Hararghe Zone Bureau of Agriculture and Rural Development), availability of established and functional FTCs, and personal experience in the area.

According to the Haramaya District Extension Department, there are 33 PAs in the district, out of which 31 have *established* FTCs. The establishment of the FTCs is not done at random since ‘randomness’ is not in the best interest of the implementers of the program, similar to other development interventions elsewhere (e.g., Davis & Nkonya, 2008; Cavatassi et al., 2011). In the PAs with established FTCs, only 14 have *operational* FTCs which provide some sort of AASs. At the time of the survey, ten¹⁹ of these FTCs were actually providing modular training to participating farmers. The remaining four²⁰ were involved in the provision of demonstration services only. From the ten *fully functional* FTCs, three PAs – *Ifa Oromia*, *Adele Waltaha*, and *Biftu Geda* – were selected based on representativeness to the rest of the PAs with functional FTCs as well as comparability to the PAs without FTCs (which are to be selected in the comparison areas).

With regard to PA selection in the comparison areas, there are only two PAs – *Ifa Bate* and *Fendisha Lencha* – which purely did not have any established or functional FTC at the time of the survey. In these comparison PAs, there were neither FTC modular training nor FTC-based demonstration services being offered to the farmers. Moreover, the PAs were relatively comparable to the sampled FTC-PAs in terms of biophysical (topography, weather, etc.) and socio-economic characteristics (as confirmed by some experts from the district). Hence, these two PAs were included in the study as comparison areas. Therefore, the total number of PAs

¹⁹ These PAs are: *Becheke*, *Kuro Jalala*, *Ido Belina*, *Damota*, *Melka Gamachu*, *Biftu Geda*, *Ifa Oromia*, *Adele Waltaha*, *Karo Tarkanfi*, and *Arwadi*.

²⁰ These are: *Ugaz Lencha*, *Qerensa Derba*, *Qorke*, and *Qarsa Kejima*.

covered by the study is five (i.e., three from treatment and two from comparison areas) (Figure 2.1).

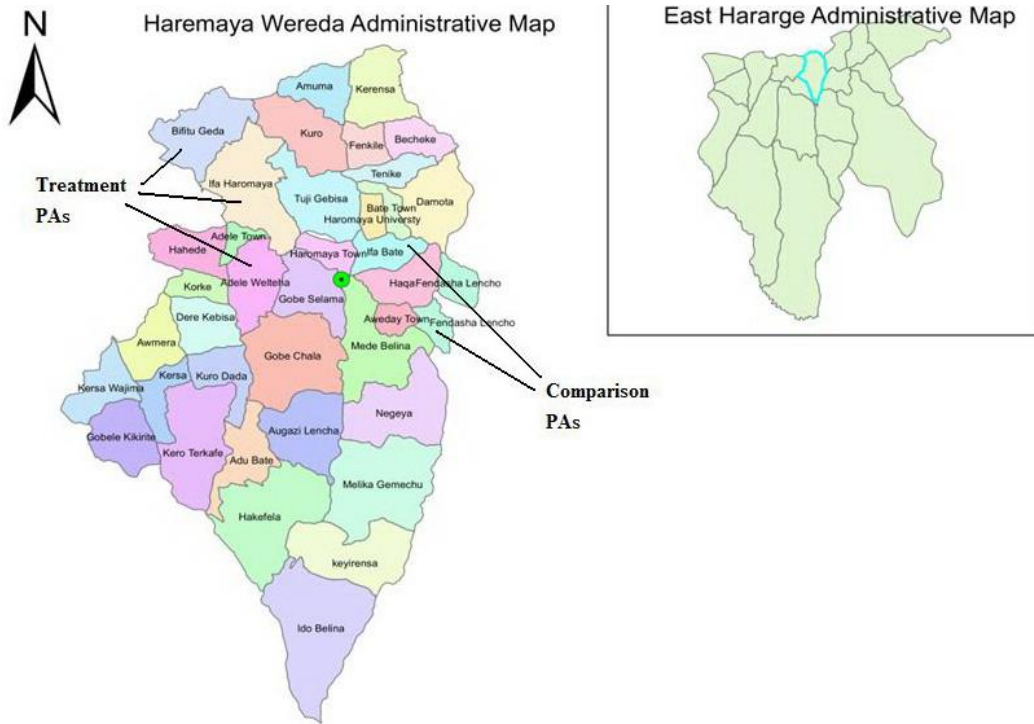
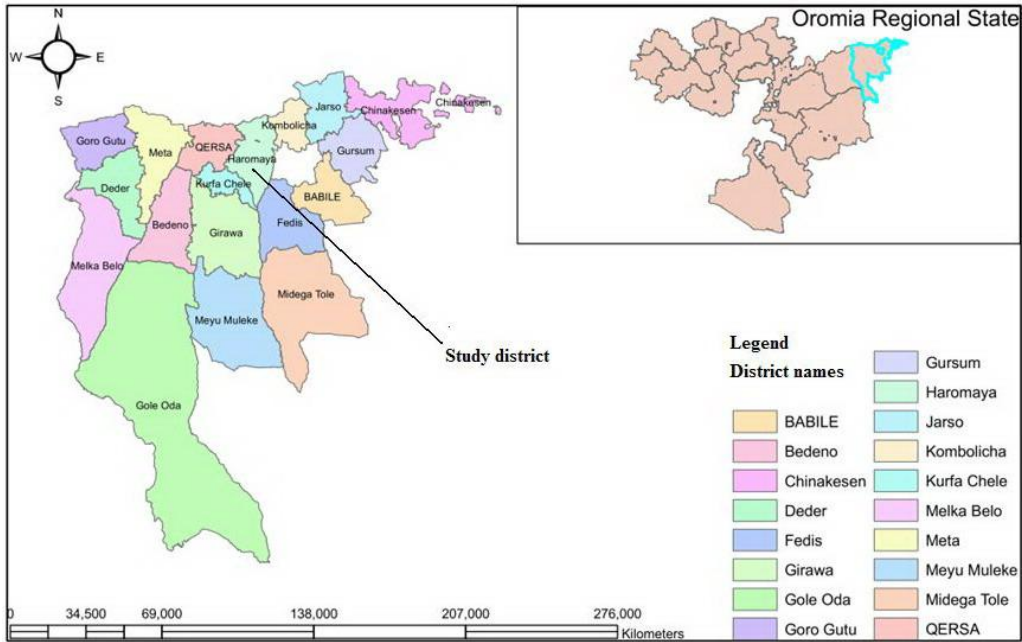


Figure 2.1 Map of East Hararge Zone of Oromia Region Showing the 18 Districts in the Zone and Map of Haramaya District Showing the 33 Peasant Associations (PAs)/Kebeles in the District

2.2.2 Household Head Selection in Treatment and Comparison Areas

On the basis of the Haramaya District BoARD, there are a total of 13,916 households in the district, of which 2,449 live in the five PAs selected for this study. Sample household selection was undertaken as follows. In treatment PAs, a list of household heads who participated in modular training in 2009 (2001 Ethiopian calendar) was obtained from the district bureau of agriculture and rural development office. Accordingly, there were a total of 754 household heads trained in the ten *fully functional* FTCs, of which 225 were reportedly from the three sampled treatment PAs. Nevertheless, at the time of the survey, Development Agents – who provide the training to the farmers – indicated that the actual intake was smaller than what was reported and that only 30 household heads successfully completed the training in each training cycle in the respective PAs. Hence, the actual number of trained household heads in the treatment PAs in 2009 was 90, which forms the treatment group.²¹

In the two comparison PAs, the list of households was also obtained from the district BoARD office. Accordingly, 824 households live in these areas. However, since it is observed – and some key-informants confirmed – that in treatment PAs many relatively better-off, progressive and model household heads were included in the training, another list containing households judged to be ‘good’ performers in these comparison PAs was developed (more on this is also available in section iv). Consequently, it was found out that there are a total of 495 such households in both PAs (188 in *Ifa Bate* and 307 in *Fendisha Lencha*). Since the aim of this exercise was to select a reasonably large number of comparison farmers who did not have any access to modular training but who are relatively close to the respondents in the treatment

²¹ The criteria used to select households into the modular training are discussed in section ‘iv. Choice of Variables for Estimating Propensity Score’

category in terms of observable characteristics, 160 such household heads were randomly sampled from these non-FTC areas. The number of comparison respondents was higher compared to the treatment ones because of the requirements of the non-experimental design used. However, it is not high enough to allow for the evaluation of spillover effects of the FTC-based AASs in the area. Table 2.1 summarizes the sample selected for the study.

Table 2.1 Sample Household Selection in Treatment and Comparison Areas

Area	PA/Kebele	No. of households ^a	No. of hhs trained at FTCs in 2009	No. of sampled hh heads
Treatment	<i>Ifa Oromia</i>	696	75	30
Treatment	<i>Adele Waltaha</i>	566	75	30
Treatment	<i>Biftu Geda</i>	363	75	30
Comparison	<i>Ifa Bate</i>	294	<i>na</i>	80
Comparison	<i>Fendisha Lencha</i>	530	<i>na</i>	80
Total		2,449	225	250

^a Obtained from the Haramaya District BoARD and Haramaya District Revenue Office (2013)

na = not applicable (since these are non-FTC PAs).

2.2.3 Data Collection

Other than the selection of study district, PAs and household heads, selection and training of research assistants as well as translation of the questionnaire to local language was carried out. The selection and training of research assistants was made on the basis of research guidance available to development practitioners in developing countries (e.g., Leones & Rozelle, 1991; Nyariki, 2009) in order to ensure the collection of high quality data. Consequently, based on information from key-informants, research assistants who were trained (at agricultural colleges) in crop production, livestock production and management, or natural resource management were selected. Moreover, it was a requirement that the enumerators are stationed at the study PAs and that they have been working with the respective communities. In addition to this, local language

ability was also a key criterion in the recruitment process. As a result, a total of seven enumerators were selected for the purpose of data collection. Furthermore, in order to ensure that quality data be collected for the study, three supervisors from Haramaya University were employed (in order to closely supervise the process of data collection by the enumerators and provide real-time feedback whenever necessary).

Both the enumerators and supervisors were given orientation training on the overall process of data collection. At the beginning of the training, the purpose of the research was stated to them. Following this, an English copy of the survey tool was distributed to them. Each enumerator was requested to take the copy home, read it, and come back the next day for discussion. This was done to check the level of understandability of the questionnaire as well as the competence levels of the enumerators themselves. This exercise helped to detect some ambiguities with a few of the enumerators regarding some items in the questionnaire. Following this, an extensive explanation about all the items contained in the questionnaire was provided to all the enumerators.

Although the above exercise was fruitful in equipping the enumerators with the necessary preparation, it was judged that it would be challenging to use a questionnaire written in English in a community where *Afan Oromo* is dominantly spoken. Hence, this issue was discussed with the enumerators, supervisors, and other collaborators. Then, the questionnaire was translated to *Amharic*, the national language. In the translation process, the level of complexity was kept to the minimum so that the enumerators and supervisors understand exactly what an item in the questionnaire refers to. However, an effort was also made to balance simplicity with accuracy.

In terms of the research process, before embarking on the actual data collection, the questionnaire was tested on a randomly sampled 30 households in *Ifa Oromia* PA. This was an

intensively supervised testing of the questionnaire to make sure that the data collection instrument was fully understood by the respondents. Based on the results of the pre-test, additional orientations were given to the enumerators and supervisors. Pre-tests were also conducted in the other PAs and issues were rectified accordingly. In relation to the main data collection, each supervisor followed his respective enumerator and provided timely feedbacks. Accordingly, each supervisor received the filled in questionnaires, checked them for missing or incompatible information, and returned them to the respective enumerator, so that the enumerator goes back to the respondent to re-fill the missing data. Although this created an additional burden on the enumerators, given the length of the questionnaire, this procedure was undertaken to once again ensure quality data to be gathered. At the beginning of the survey, many questionnaires were returned to the enumerators to re-fill. However, as the survey progressed, the enumerators get used to the issues and a few questionnaires were returned to be corrected. On average, each interview lasted for 84 minutes in treatment PAs (Number of observations=90), and 104 minutes in comparison PAs (Number of observations=160).

2.2.4 Empirical Strategy and Data Analysis

Investigating the causal effect of an intervention on outcome indicators depends on the type of data available and empirical methods employed in the analysis (Winters et al., 2011). Over the past two decades, methodological developments in impact evaluation have focused on issues of attribution of impact to interventions (Imbens & Wooldridge, 2008) and determining the counterfactual outcome of participants (Caliendo & Kopeinig, 2008; Winters et al., 2011; Benin et al., 2012). These two problems are important in interventions in the agricultural sector because any outcome that an intervention aims to improve has many other factors that could affect it (IFPRI, 2010). Hence, standard impact evaluations are required to consider the problem of

placement bias (in which the location or target population of the intervention is not randomly selected), selection bias (in which households choose whether or not to participate), and spillovers (intended or unintended contamination of treatment and comparison groups).

The process of evaluating the impact of a program on an outcome indicator of interest inevitably requires answering a daunting question: ‘what would have happened to the participants of the program had they not participated in it?’ Referred in the evaluation literature as ‘*the fundamental problem of causal inference*’ (Holland, 1986) or ‘*fundamental evaluation problem*’ (Heckman et al., 1999), this is a serious question to answer since an individual can only be observed at one state in a given time – either participating or not participating in the program (Cavatassi et al., 2011; Winters et al., 2011).

The ideal way to deal with the problem of counterfactuals is to employ randomized control trials (RCTs) following the *potential outcome approach* or *Roy–Rubin model* (Roy, 1951; Rubin, 1974), which entails the random assignment of eligible individuals to a treatment and comparison group. In doing so, any difference in an indicator of impact can be attributed to the intervention. The treatment effect following this approach can be given as

$$T_i = Y_{i1} - Y_{i0} \tag{1}$$

where T_i is the program/treatment effect for an individual i , Y_{i1} and Y_{i0} are the potential outcomes with and without the program, respectively. However, as stated above, only one of these outcomes can be observed for an individual at a given time. The outcome that is not observed is called the counterfactual outcome. The potential outcome approach using RCTs is not viable in the present study setting due to concerns about placement/targeting and self-selection. Moreover, RCTs entail careful planning and implementation to ensure compliance to

assignment, a lot of human and financial resources, ethical considerations, and issues of internal and external validity. Random assignment is also not in the best interest of implementers of many development programs aimed at improving the livelihoods of rural dwellers (Davis & Nkonya, 2008; Cavatassi et al., 2011).

The alternative to the experimental approach is the use of quasi-experimental approaches, which seek to create, using empirical methods, a comparable control group that can serve as a reasonable counterfactual (Dillon, 2011; Winters et al., 2011). These approaches try to estimate the impact of an intervention when individuals are not randomly assigned to treatment and control groups.²² In the presence of spillover effects, an added difficulty is finding control groups for both groups – those who directly participate in the intervention and those who do not participate directly but benefit from their indirect involvement. Nevertheless, a credible evaluation of the direct and indirect effects of an intervention can be conducted as long as the evaluation design has an in-built strategy to account for them (Dillon, 2011). Some common quasi-experimental approaches include propensity score matching (PSM), double difference or difference-in-difference (DID), regression discontinuity design (RDD), and instrumental variables (IV) estimation.

Difference-in-Difference (also called Double Difference) is an approach to impact evaluation which allows for comparisons of outcome indicators before and after an intervention between participants and non-participants. According to Ravallion & Chen (2005), the DID approach can account for observable and unobservable factors which may bias the estimates of impact, provided that data are available to make pre-post comparisons between the two groups.

²² Non-random assignment of interventions is a widespread problem in developing countries (Duflo et al., 2007). In such situations, the main challenge is finding a reliable comparison group which can be compared to the treatment group.

Nevertheless, the approach may produce biased estimates in situations where an intervention results in a differential impact on individuals having different observable characteristics (Ravallion & Chen, 2005). Moreover, the fact that DID requires at least two-period panel data²³, which is difficult to find for programs such as AASs in most developing countries, makes it less preferable compared to PSM.

Another quasi-experimental approach which is applicable under less stringent assumptions and requires less data (compared to PSM) is regression discontinuity design (RDD). The idea behind RDD is that in a situation where interventions are assigned on the basis of a score or index and there is a cutoff/threshold point below which units (individuals or localities) are not eligible for the intervention and above which they are, it can be assumed that the units do not differ greatly in terms of observable or unobservable characteristics nearby the threshold. Then the impact of the intervention can be evaluated by comparing the outcomes of those units closer to the threshold from both sides (Dillon, 2011).

Although the small data requirements and the ability to estimate the impact of an intervention controlling for both observable and unobservable characteristics are the main strengths of RDD, the approach has a strong internal validity but weak external validity. RDD is a powerful approach to estimating the direct and indirect impact of an intervention around the cutoff point but the results cannot be extended to units far from the threshold. This makes sense if there are concerns about representativeness of areas around the cutoff point when compared to the rest of the population (Dillon, 2011).

²³ Such data can be constructed using a backward recount, but one may end up gathering a very poor quality data on relevant variables (especially when recall periods are more than one year).

The instrumental variable approach accounts for the bias resulting from non-random assignment to an intervention by means of another variable (an instrument, Z) which is correlated with the assignment to treatment variable (T), but the only effect it has on the relevant outcome variable is through its effect on T (Dillon, 2011). This approach helps to control for differences in many observed characteristics, but in practice the assignment to treatment would be based on criteria that are not necessarily observable, such as budgetary considerations, logistics, convenience, or political priority. Moreover, finding a valid instrumental variable is a great practical challenge. As Dillon (2011) points out, the validity of instrumental variables can convincingly be challenged.

In the present study, the propensity score matching procedure is implemented primarily due to the nature of data available. Although PSM is based on a quite strong assumption (see section 2.4.2), requires a large amount of data from participants and non-participants, and fails to account for selection bias due to unobservable characteristics,²⁴ it remains an influential quasi-experimental approach to estimating the impact of an intervention in a fairly straightforward manner (Dillon, 2011).

i. Matching Methods in Evaluating Program/Treatment Effects

The use of matching methods in evaluating program/treatment effects has grown in popularity across a wide range of disciplines (Caliendo and Kopeinig, 2008). The fundamental notion behind matching is to construct a comparable (sub) group of individuals – who are similar to the treatment individuals/groups in all relevant pre-treatment characteristics X – from a sample of

²⁴ It is recognized that although treatment and control groups are matched in terms of observable characteristics, they may differ in unobservable characteristics (Heckman et al., 1998).

untreated ones. Then, having created this comparable group and performed matching under some identifying assumptions, any observed difference in outcome between the two groups can be attributed to the program/treatment. However, since matching on all observed characteristics X is problematic – referred to as the ‘curse of dimensionality’ (Rosenbaum & Rubin, 1983) – the technique mostly relies on the use of balancing scores, the commonest of them being propensity scores, i.e., the probability of participating in a treatment/program conditional on observable characteristics X . Matching techniques based on this score are referred to as propensity score matching. In practice, a model (probit or logit for binary treatment) is estimated in which participation in a treatment/program is explained by several pre-treatment characteristics and then predictions of this estimation are used to create the propensity score which ranges from 0 to 1. Having done this, one can compare the units (individuals or groups) which are made "close" to each other in terms of the propensity score (Dillon, 2011).

There are different approaches of implementing propensity score matching. Khandker et al. (2010) discuss a range of these techniques, including nearest neighbor (NN) matching, caliper or radius matching, stratification or interval matching, and kernel and local linear matching. There are two issues worth considering in making distinctions among these matching techniques. First, the way in which the neighborhood for each individual in the treatment group is defined as well as how the overlap condition is handled. Second, the way in which the weights are assigned to these neighbors (Caliendo & Kopeinig, 2008).

The most straightforward and one of the most frequently used matching estimator among the abovementioned techniques is the nearest neighbor matching. The basic notion behind this estimator is that an individual from a treatment group is matched to an individual with the closest propensity score in the comparison group. Variants of this estimator are nearest neighbor

matching with replacement and without replacement. Whereas the former refers to the use of an individual from a comparison group as a match for individuals in the treatment group more than once, the latter refers to the use of an individual from the comparison group as a match to an individual in the treatment group only once. In addition to these, one can choose n nearest neighbors (e.g., $n = 2, 3, 5$, etc.) and perform matching. A problem with nearest neighbor matching is that the difference in the propensity scores of a treatment and its closest comparison neighbor may still be very high, which may result in poor matches.

One way out of such a problem is imposing a tolerance level on the maximum propensity score distance (caliper). Hence, caliper matching handles this problem well. It involves matching treated and untreated individuals which are close to each other within a given propensity score range (given by caliper). A potential downside of caliper matching is the difficulty to know *a priori* the tolerance level that is reasonable (Smith & Todd, 2005). A variant of caliper matching called radius matching is suggested by (Dehejia & Wahba, 2002). The central idea of this estimator is to utilize not only the nearest neighbors within the given caliper, but also all the comparison individuals within that caliper. An advantage of this method is that it uses only as many comparison units as are available within the caliper and therefore allows for usage of extra (fewer) units when good matches are (not) available (Caliendo and Kopeinig, 2008).

The matching techniques discussed up to now have in common that only a few subjects from the comparison group are used to create the counterfactual outcome of a treated individual. Kernel matching (KM) and local linear matching (LLM) are nonparametric matching estimators that use a weighted average of all – depending on the choice of the kernel function – individuals in the comparison group to construct the counterfactual outcome. Hence, a key benefit of these methods is the use of more information which leads to lower variance. However, some of the

subjects might be poor matches. This is why it is important to properly set the common support condition in kernel matching and local linear matching. For kernel matching, the choice of the kernel type (e.g., Gaussian, Epanechnikov, etc) and the bandwidth are also relevant issues.

Although the different matching techniques are discussed in brief, a practical question remains to be answered – how to choose among the matching methods. Evidently, all the matching estimators should give the same result, since as sample size grows, argues Smith (2000), they all become closer to comparing only exact matches. The choice of a specific matching procedure matters, however, in small samples (Heckman et al., 1997), where there is a usual trade-off between bias and variance (see Table 1 of Caliendo and Kopeinig (2008) for comparison among the matching methods in terms of bias and variance).

Since the choice of a specific matching method can affect the estimated impact parameter through, for example, the weights assigned (Khandker et al., 2010), it seems that there is no hard-and-fast rule on the choice of matching methods. There is also no ‘magic’ matching technique capable of being used in all situations. Each matching method can perform reasonably well in a specific situation given rich dataset. Hence, it is suggested to try as many matching methods as possible (Bryson et al., 2002). Should the methods yield the same result or converge to a closer result, the choice of a matching method would not be an important issue. Should the results differ from method to method, a careful analysis has to be made concerning the difference and the choice of an estimator. Moreover, implementing different matching methods can also reveal the underlying existence or lack of consistency, which can be used as a way of assuring sensitivity of the results to changes in matching methods. Therefore, in this investigation, the Nearest Neighbor Matching, Kernel Matching (both Gaussian and Epanechnikov), and Radius Matching are implemented.

ii. Assumptions in Implementing Propensity Score Matching

Although the choice of a matching estimator is a crucial aspect in implementing PSM, there are other equally vital issues in this procedure. In particular, there are two assumptions to be made and verified somehow. The first one is a quite strong assumption, referred to as *unconfoundedness* (Rosenbaum and Rubin, 1983), *selection on observables* (Heckman & Robb, 1985), or *conditional independence assumption (CIA)* (Lechner, 1999). According to this assumption, the treatment needs to fulfill the criterion of being exogenous, implying that any systematic difference in outcomes between the treatment and comparison groups with the same values for characteristics X can be attributed to the treatment. This assumption can be noted as

$$Y_1, Y_0 \perp D|X \quad (2)$$

where \perp denotes independence, $D|X$ is assignment to treatment (D equals 1 for treated and 0 otherwise) conditional on characteristics X , Y_1 and Y_0 are the outcomes for the treatment and comparison groups, respectively. This means that given a set of observable characteristics X which are unaffected by treatment (i.e., not contaminated by treatment; not affected by treatment or the expectation of it), potential outcomes are independent of assignment to treatment. In order to credibly justify this assumption, therefore, one practical requirement is that researchers must observe all the covariates that simultaneously affect assignment to treatment and potential outcomes (Caliendo and Kopeinig, 2008).

The second assumption, called *common support* or *overlap*, ensures that individuals/groups with the same values for characteristics X have a positive probability of being both participants and non-participants of a program/treatment (Heckman et al., 1999). This assumption is noted as

$$0 < P(D=1|X) < 1 \quad (3)$$

This means that any combination of covariates X observed in the treatment group can also be observed in the corresponding comparison group (Bryson et al., 2002). The *overlap* condition enables to compare comparable units. Since violation of this condition results in a major bias (Heckman et al., 1997), it is recommended that the comparison units that are made comparable to the treatment ones should be used in the analysis as counterfactuals (Dehejia & Wahba, 1999). Hence, it is a requirement to check the goodness-of-overlap of the propensity score distributions and the common support region between the treatment and comparison groups.

The two aforementioned assumptions are jointly referred as ‘*strong ignorability*’ (Rosenbaum and Rubin, 1983). According to Caliendo and Kopeinig (2008), under ‘*strong ignorability*’, mean independence is a satisfactory condition to define the average treatment effect on the treated (ATT) for all values of covariates X ; in which case, the assumptions can be simplified as $Y_0 \perp D|X$ (*unconfoundedness for comparison groups*) and $P(D=1|X) < 1$ (*weak overlap*). These assumptions are sufficient for the identification of ATT, because the moments of the distribution of Y_1 for the individuals in treatment group are directly estimable.

Nevertheless, in order to deal with the ‘curse of dimensionality’ problem, Rosenbaum and Rubin (1983) show that if the potential outcomes Y_1 and Y_0 are independent of treatment allocation conditional on covariates X , then they are also independent of treatment conditional on the propensity score (i.e., $P(D=1|X) = P(X)$). It follows, therefore, that if the *unconfoundedness* assumption holds, all biases associated with observable covariates can be removed by conditioning on the propensity score (Imbens, 2004). The *unconfoundedness* assumption given the propensity score is noted as $Y_1, Y_0 \perp D|P(X)$.

Generalizing the above issues, assuming that the *unconfoundedness* assumption holds and there is sufficient *overlap* between the treatment and comparison groups, the PSM estimator for ATT conditional on the propensity score can be written as

$$ATT = E_{P(X|D=1)} \{E[Y_1|D=1, P(X)] - E[Y_0|D=0, P(X)]\} \quad (4)$$

This means, the PSM estimator is simply the mean difference in outcomes over the common support region, appropriately weighted by the propensity score distribution of treated participants (Caliendo and Kopeinig, 2008).

iii. Matching Quality Analysis (Balance Checking)

After having performed matching, it is a recommended practice to check whether balancing of the relevant covariates in the two groups is achieved through the matching procedure. The main idea behind analysis of matching quality is the comparison of situations before and after matching and checking if there remains any difference between the two groups after conditioning on the propensity score (Caliendo and Kopeinig, 2008). A number of techniques are available to check balancing, including mean comparisons between treatment and comparison groups (before and after matching), standardized bias, and overall measures of covariate imbalance. In terms of mean comparisons, a two-sample t-test (before and after matching) can be used to check the existence or lack of significant differences in covariate means between the treated and comparison groups (Rosenbaum & Rubin, 1985). As a rule-of-thumb, there should not be any significant difference in means after matching. However, to what extent the different matching procedures reduced/eliminated original bias cannot be visible from t-test results.

Standardized bias is a commonly used approach in a number of evaluation studies (e.g., Lechner, 1999; Sianesi, 2004; Arpino & Aassve, 2013; Lee, 2013). Rosenbaum and Rubin (1985) define the absolute standardized bias (for each covariate X) as the absolute difference in sample means between the matched treatment and comparison samples as a percentage of the square root of the average sample variance in the two groups. As pointed out by Caliendo and Kopeinig (2008), one possible drawback of the standardized bias approach is the lack of a clear indication about the level to indicate matching success, although a standardized bias below 5% after matching is seen as sufficient in most empirical studies. The standardized bias before matching can be written as

$$\text{Standardized bias}_{\text{before}} = 100 * \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{0.5 * (V_1(X) + V_0(X))}} \quad (5)$$

The standardized bias after matching can be written as

$$\text{Standardized bias}_{\text{after}} = 100 * \frac{\bar{X}_{1M} - \bar{X}_{0M}}{\sqrt{0.5 * (V_{1M}(X) + V_{0M}(X))}} \quad (6)$$

where \bar{X}_1 (V_1) is the mean (variance) in the treatment group before matching and \bar{X}_0 (V_0) the corresponding values for the comparison group. \bar{X}_{1M} (V_{1M}) and \bar{X}_{0M} (V_{0M}) are the mean (variance) values for the matched samples.

In addition to mean comparisons and standardized bias, there is a measure of overall covariate imbalance. Sianesi (2004) suggests the comparison of Pseudo- R^2 before and after matching as a method to check balance. The Pseudo- R^2 indicates how well the covariates X explain the probability of participating in the treatment. The Pseudo- R^2 has to be very low after matching to indicate success of the matching process. Moreover, the likelihood ratio test on the joint

significance of all covariates in the (logit) model should not be rejected before matching, but should be rejected afterwards (Caliendo and Kopeinig, 2008). In the present study, balancing of covariates is tested using all of the techniques discussed in this sub-section.

iv. Choice of Variables for Estimating Propensity Score

A very crucial aspect in the estimation of propensity scores is the criteria to include/exclude variables into/from the model. The most widely followed practice in selecting variables into the model estimating propensity score is to include all the variables which simultaneously affect both participation in treatment/program and outcome variable of interest, but are not affected by the treatment or the expectation of it (Lechner, 2002; Caliendo and Kopeinig, 2008). This implies that a variable that affects outcome but not participation should not be included. Likewise, any variable which does not affect either participation or outcome should not be included in the model (Cuong, 2013). Although there is a debate for and against the widely adopted tradition, the bottom line in this regard seems to have a sound economic theory and a detailed understanding of the program-/treatment-specific considerations (Caliendo and Kopeinig, 2008).

In the choice of variables for the estimation of propensity score for this study, therefore, both theoretical and empirical sources were consulted. Most importantly, previous works in various fields employing propensity score matching were reviewed, paying special attention to the studies which – directly or indirectly – deal with impact evaluation in agricultural or related disciplines. In order for the reviews to be more relevant to the specific study context, empirical sources from Africa (i.e., Ethiopia, Mozambique, Tanzania, Kenya, Mali, Burkina Faso, Uganda, Burundi, DRC, and Rwanda) were focused at. However, valuable insights and suggestions were

also obtained from other studies conducted elsewhere (Philippines, Germany, and Ecuador). Many theoretical sources, including Caliendo and Kopeinig (2008) and Cuong (2013), were also consulted.

Directly relevant to the current investigation, previous empirical works were reviewed on broad aspects of *agricultural extension* (Dercon et al., 2009; Gebregziabher et al., 2013; Thangata and Mequaninte, 2011; Elias et al., 2014), *Farmer Field Schools* (Davis et al., 2012; Todo & Takahashi, 2013), and *adoption of technologies* (Cunguara & Darnhofer, 2011; Amare et al., 2012; Deschamps-Laporte, 2013; Mutuc et al., 2013; Shiferaw et al., 2013). Some of the outcome variables analyzed by the impact evaluations in the aforementioned studies include: household consumption expenditure; agricultural technology adoption, input use, and agricultural productivity; household food security, poverty and consumption growth; economic impact, farm household technical efficiency, and agricultural income; crop production, productivity, and vulnerability.

Regarding the methodologies followed in selecting the variables for the estimation of propensity scores, the reviewed studies include or exclude variables based on some underlying theory, as highlighted by Caliendo and Kopeinig (2008), and program-specific considerations. Nonetheless, there are common elements in almost all the studies – demographic and household characteristics, farm level attributes, human assets, financial capital, social capital, input/output markets, and (formal/informal) institutions. Depending on the specific nature of the study, there are also considerations regarding shocks, participation in community life, and off-/non-farm employment.

Another relevant aspect in the choice of the variables is the consideration regarding the criteria used to select farmers into modular training. According to the Farmers' Training Centers Operational Manual/Guideline/ (October 2009),²⁵ the selection of farmers into modular training is done by the *Kebele extension unit*²⁶, tasked with duties, such as identifying eligible farmers from the *Kebele* register, following the process of training and its accomplishment, serving as a bridge between the local community and FTCs, reporting the activities of the FTCs to the *Kebele* Cabinet, and initiating the community to contribute their share to the success of the training program. The operational manual states that “the selection of farmers for the training should focus on identification of committed ones who have the desire and initiative to modernize the existing traditional agricultural practices of the area” (p. 17). Accordingly, for the selection of such farmers – irrespective of religious, economic or other background – four important criteria are used. These are: *educational background, devotion and initiatives, age, and gender.*

Educational background. In terms of educational attainment, both uneducated and educated farmers are eligible to the modular training. However, the uneducated farmers are expected to be those who are widely accepted by the community as role models in their devotion to agricultural activities and ability to learn and adopt new/improved technologies. In the three FTC-PAs of this study, 70% of the respondents are uneducated, indicating that greater proportion of the sample comprises trainees who do not have any formal school background. The remaining 30% had some years of primary and/or secondary schooling. Concerning the educated participants, the list includes educated youth or youth who dropped out of school, whether they manage their own

²⁵ Farmers' Training Centers Operational Manual/Guideline/ (October 2009). Agricultural Extension Department, Ministry of Agriculture and Rural Development, Federal Democratic Republic of Ethiopia. [http://www.oerafrica.org/ftpfolder/website%20materials/Agriculture/haramaya/Perspective_Agricultural_Extension/AICM%20Module%20Final%20Jan%2031,%202012/FTC_guideline\[1\].doc](http://www.oerafrica.org/ftpfolder/website%20materials/Agriculture/haramaya/Perspective_Agricultural_Extension/AICM%20Module%20Final%20Jan%2031,%202012/FTC_guideline[1].doc) [accessed: October 2013]

²⁶ Chaired by Head, Development Agents and consisting of Development Agents in the *Kebele*; Head, Localities Development Committee; Head, Women's Representatives; and Head, Youth Representatives.

land or live with their parents. The distribution of educational level at baseline (i.e., year 2009) across the treatment PAs and the comparison PAs (Table 2.2 in ‘results’ section) indicates that the comparison respondents have comparably better educational attainment than the treatment ones. Although there is no statistically significant difference ($t=1.45$, $p=0.15$), it is actually observed in the data that the comparison respondents are located, on average, 1.1 km (Sd 0.8) away from primary schools, whereas the corresponding figure for the treatment ones is 1.2 km (Sd 0.9).²⁷

Devotion and initiatives. The inclusion of this criterion is justified by the FTC operational manual in that the commitment and self-initiation of farmers in the agricultural activity of a certain locality greatly contributes to the achievement of development objectives of the area. Consequently, it is recommended to consider initiation and hard-working capacity of farmers for candidacy. More specifically, the complete willingness of trainees to put into practice the theoretical knowledge they are going to acquire from training, the capacity of trainees to adopt and implement new/improved technologies, and the readiness of trainees to share knowledge to friends and neighbors are key issues in the selection process.

In order to measure the level of *devotion* among the sampled treatment respondents, three indicator variables – that are considered by the *Kebele extension unit* – were used: main occupation, experience in farming, and years of participation in general/on-farm AASs. In terms of main occupation of the household head, 100% of the treatment respondents were engaged in farming/agricultural activity as their main-stay. On average, these respondents spent 18 years in

²⁷ On the basis of educational background and other relevant differences among trainees, separate classrooms are sometimes arranged because of the realization that all participants do not learn at the same pace.

farming, suggesting their rich experience in agricultural activities (prior to joining the modular training). Similarly, these respondents participated, on average, for 15 years in general AASs, once again showing their deep involvement in AASs aiming to bring about changes in their living standards. Concerning the willingness to share knowledge obtained from training, it is actually observed that about 96% of the participants of the training shared the contents of the training with three neighbors/friends (min. 0, max. 10), on average. Moreover, since the willingness and capacity of respondents to adopt technologies and practices depends on, to some extent, the resources available at their disposal and institutional affiliation, size of land holding and membership in organizations were included as additional proxies in order to better account for the criterion *devotion and initiatives*.

Age. The ideal age required to be eligible for the modular training is 18 years. However, farmers aged above 18 years can also be accepted provided that the accommodation capacity of the FTCs is adequate.²⁸ In the treatment group, the mean age of the respondents is 33 years.

Gender. Both genders are eligible to the training. However, there is a provision to encourage women's participation. Regarding gender distribution among the treatment group, only two of the 90 respondents are female household heads. This should not be taken as an indication of program-biasness against women, since the study considers the household heads, which by default are males, unless divorced or widowed community members are encountered.

On the basis of these considerations, eligibility and selection criteria, outcome variable of interest, and policy direction of setting up *Farmer Training Centers* in Ethiopia, variables were included in the logit model to estimate the propensity score as can be seen in Table 2.2.

²⁸ There is also a recommendation in the operational manual that the age structure of the participants has to be taken into consideration while grouping the participants in classrooms.

2.3 Results

2.3.1 Descriptive Statistics of Baseline Characteristics Determining Participation in Modular Training at Farmers' Training Centers

This section presents the descriptive statistics of covariates influencing the participation of farmers in modular training provided at FTCs. Table 2.2 depicts the statistics on the covariates measured at the baseline year (i.e., 2009). It is evident from the table that there exists a statistically significant difference between treatment and comparison respondents regarding level of education, household size, asset ownership, experience in advisory services, organizational membership, and number of educated members of the household.

Table 2.2 Baseline Information on Selection Criteria by Study Kebele/PA (Measured in 2009)

	Treatment	Comparison	T-test
Age of the household head (=1 if >18 years)	0.99	0.99	- 0.41
Gender (=1 if male)	0.98	0.96	0.66
Education (=1 if attended some formal school)	0.30	0.68	- 6.10 ***
Household size (number of people in the house)	3.30	3.99	- 2.47 **
Assets (number of productive assets)	3.01	2.61	3.22 ***
Farming experience (number of years)	18.29	18.62	- 0.36
Experience in AASs (number of years)	15.24	13.63	1.69 *
Land holding (=1 if ≥ 1 ha) ^a	0.32	0.36	- 0.54
Organization (=1 if a member of at least one organization/association)	0.99	0.86	3.38 ***
Educated members (=1 if number of educated members of the household is ≥ 1)	0.89	0.95	- 1.80 *
Number of observations	90	160	

***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

^a This classification of land is used because the minimum land size required to set up a demonstration plot in ones land holding is 0.5 ha. Hence, to be able to allocate 0.5 ha of land for such demonstration plots while at the same time be able to manage own farming business, one might need at least one hectare of land. Data on comparison PAs is provided alongside for ease of comparison.

2.3.2 Propensity Score Estimation

It is to be recalled that the selection of covariates to be included in the model estimating the propensity score is based on knowledge about the modular program (MOARD, 2009), selection and screening criteria, previous empirical studies, and theoretical as well as practical considerations (e.g., Caliendo and Kopeinig, 2008; Cuong, 2013). The propensity score is estimated using logistic regression, where the dependent variable, participation in modular training, takes the value of 1 for participants and 0 otherwise. The results concerning the individual covariates are found in Table 2.3. Accordingly, education of the household head, household size, asset ownership, farming experience, and membership in organizations are significant predictors of participation in modular training.

Table 2.3 Participation in Modular Training based on Baseline Covariates Measured in 2009

Dependent variable:		
Participation in modular training (=1 if trained in 2009)	Coef.	Z
Age of the household head (=1 if >18 years)	- 0.53 (1.47)	- 0.36
Gender (=1 if male)	1.24 (0.92)	1.36
Education (=1 if attended some formal school)	- 1.93 (0.35)	- 5.48 ***
Household size (number of people in the house)	- 0.16 (0.08)	- 1.95 **
Assets (number of productive assets)	0.42 (0.18)	2.38 **
Farming experience (number of years)	- 0.08 (0.04)	- 2.19 **
Experience in AASs (number of years)	0.05 (0.03)	1.38
Land holding (=1 if \geq 1 ha)	- 0.15 (0.36)	- 0.42
Organization (=1 if a member of an organization)	2.47 (1.06)	2.32 **
Educated members of the household (=1 if \geq 1)	0.53 (0.60)	0.90
_cons	- 2.93 (2.17)	- 1.35
Number of observations	250	
Pseudo R ²	0.21	
LR chi ² (10)	67.95 ***	

*** and ** denote significance at 1% and 5% level, respectively.
Standard errors in parenthesis

Regarding the distribution of the propensity scores of the treatment and comparison respondents, the histograms in Figure 2.2 indicate a significant overall overlap between the two groups.

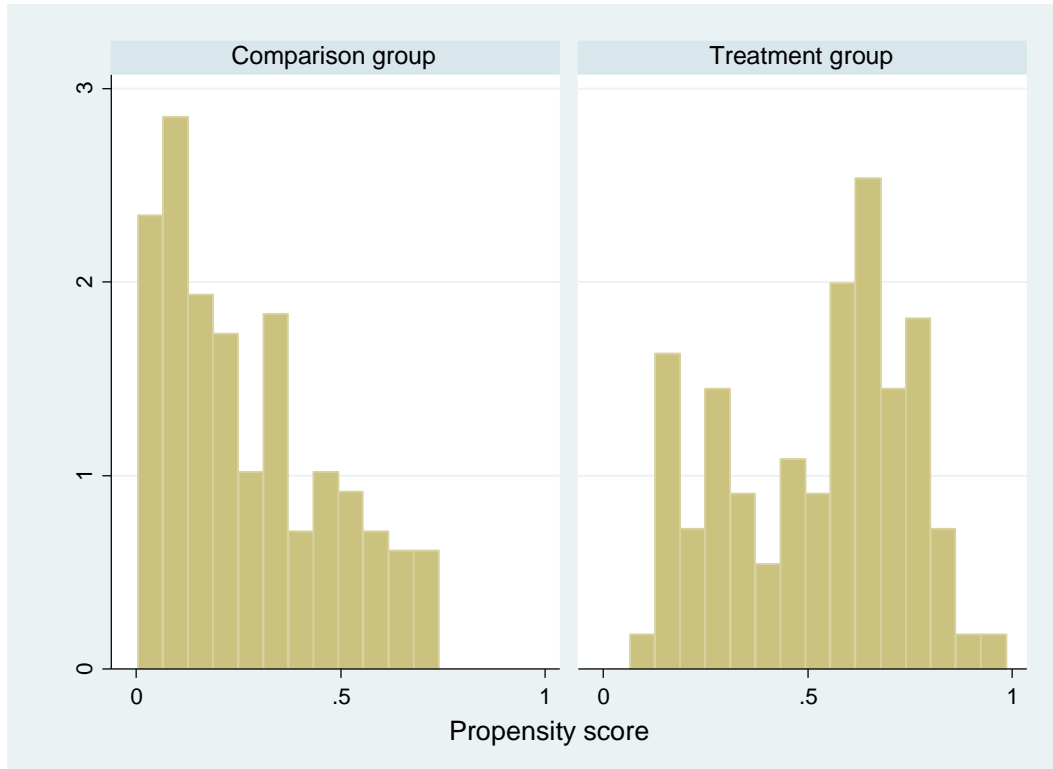


Figure 2.2 Distribution of Propensity Scores between Comparison (Left Histogram) and Treatment (Right Histogram) Respondents

2.3.3 Construction of Outcome Variable for Impact Evaluation

The selection of the outcome variable for impact evaluation adopted in this study has made reference to the policy aim of setting up FTCs in the study area. According to the FTC establishment document (MOARD, 2009), FTCs are constructed to increase income through agricultural production and productivity improvement. This is intended to be achieved by providing technical information and advice, training, and demonstration related to crop, livestock, and natural resources, with crop and livestock being the two most important components where FTCs are actively engaged in currently. In the light of these considerations, the present study takes the agricultural (or farm) income as dependent/outcome variable.

A second choice in the analysis has been the selection of an appropriate concept and sectoral composition of farm income to refer to. On this point, the literature dealing with the relationship between agricultural income, agricultural technologies, practices, and institutions both in developed and developing countries considers various perspectives (Reardon, 1997; Severini & Tantari, 2013; Becerril & Abdulai, 2010; Cunguara & Darnhofer, 2011; Benni & Finger, 2013; Harttgen & Vollmer, 2013; Noltze, Schwarze, & Qaim, 2013; Davis et al., 2012; Benin et al., 2012b; Salvati & Carlucci, 2013; Todo & Takahashi, 2013; Shiferaw et al., 2013).

Following Cunguara and Darnhofer (2011), this study adopted the total farm household factor income from agricultural activities. However, in accordance to Davis et al. (2012) and considering the main focus of the FTCs in the investigated area, this concept was restricted to the dominant agricultural sectors, namely crop and livestock activities. Aggregating these two major sources of income, *farm income* is operationalized as the total value of income from crop and livestock production net of the cost of intermediate inputs associated with each of the activities. Some studies use per-capita household income by dividing the total value (i.e., farm income) by the number of household members (e.g., Harttgen and Vollmer, 2013). Others adopt an equivalence scale to appropriately weigh the household members (e.g., Benin et al., 2012). The use of an equivalence scale is an important approach particularly when one deals with the individuals comprising the household as units of analysis. This issue is not taken into consideration because the current investigation looks at the household as a unit of analysis – collecting data from the household head – thus equivalence scale is of little importance. In the following, the calculation procedure is explained for crops and livestock separately.

i. Income from Crop Production

Crop income by household has been calculated as the value of crop sold, used for own consumption and stored, including the value of crop residues net of the intermediate inputs associated with crop production at household level (Table 2.4). Unlike many previous studies in rural areas, which focus on a few dominant crops grown by farmers, a concerted effort was made to reliably capture the total value of all or almost all the crops managed by the respondents. To this purpose, each of the respondents is asked to list the types of crops planted on their land during the last 12 months²⁹. Although the survey instrument contained a list of 60 types of crops, the interviewed household heads were encouraged to list the types of crops they planted without probing them from this list. However, at times when it is realized that a respondent did not mention – purposely, to understate income, or unintentionally – a crop type which is commonly grown in the area, he/she is reminded from the list contained in the questionnaire. This procedure is followed consistently for all the respondents in order to obtain a complete list of the crops grown by them. Nonetheless, since the respondents managed different types of crops in a given plot of land, which is a typical arrangement in a mixed crop farming system dominantly practiced in the study area, it was not possible to obtain the area under each crop cultivated by the respondents. The existence of such a challenge prevented, for example, the estimation of productivity.

Data on the quantity (in kg) produced by crop in the last 12 months is collected as control variable; production sold, home consumed and stored should not exceed or be lower than this

²⁹ The most important crops produced include *chat*, sorghum, maize, potato, onion, beet root, lettuce, lentils, fenugreek, spinach, carrot, haricot beans, chick peas, tomato, cabbage, groundnuts, pumpkin, cow peas, garlic, and sweet potato.

amount. The declared quantity of production sold by commodity has been converted in terms of Birr (which is the official Ethiopian currency) according to the price declared by each respondent for each commodity (i.e., nominal price). This estimation method has been preferred to the use of a single price, such as *Peasant Association*-level average price or local town market average price, for all commodities due to the fact that respondents in the treatment and comparison group sold their products on different markets. In this way, the different households' negotiation capacity underlined by the variability of price for the same commodity across the respondents has been considered.

The other issue in the choice of nominal price over average price in deriving the value of crops sold, consumed and stored is related to maintaining consistency in the valuation of household revenues and costs associated with crop production and management. In particular, this refers to how the expenditures listed under 'intermediate inputs for crop production' were gathered. Since these item-wise expenditures were declared by each household head and latter summed up to obtain the household level expenses, this has not allowed for the derivation of average price for each item in the expenditure list (e.g., average price for fertilizer, for seed and seedlings, for insecticides, etc.). Hence, the evaluation of both revenues and costs in this study is based on the use of nominal prices.

Summing up the revenues from each crop type sold, the total gross income (Birr) realized by each household from the crops sold is attained. The t-test shows the existence of a statistically highly significant difference between the two groups in terms of mean gross crop revenue. During the field work, it is actually observed that the respondents in the comparison group sold a high-quality high-value crop – the dominant one being *chat* – compared to those in the treatment group, and this might have contributed to the observed difference.

The second component of farm income is represented by a household's own present consumption and food stored for future use. From the interviews, the quantity (kg) of each crop produced and used for own consumption and stored for future consumption or sale is obtained. Although the mean household size is the same in both cases (i.e., seven, including the household head), it appears that the comparison respondents consumed/stored more food crops than the treatment ones, though the difference is not statistically significant. However, since the value of crop sold is much higher for the comparison respondents compared to the treatment group, it indicates that most of the high-value crop produced is sold in these areas.

During the questionnaire design and its validation, the importance of crop residues or by-products in income formation emerged. This is a component generally neglected in empirical investigations. However, it should deserve a specific attention particularly for its potential role in farm income integration. During the investigation, the likely income from this source was estimated at more than 8 percent of crop revenues.³⁰ In the definition of crop income, revenues from crop residues sold and the value of crop residues paid as rent are also considered, the value of which are computed on the basis of the declared amounts obtained by each respondent.

The submitted questionnaire also provided details on household expenses related to intermediate inputs for crop production— fertilizer, pesticides, seeds and seedlings, transport, and other items—in the last 12 months. The result shows the existence of a highly significant difference in terms of expenses for crop production, with the comparison respondents spending, on average, more than (almost twice as large as) the treatment ones. This may be because the physical existence of

³⁰ In relation to the by-products of crop production, the respondents were asked what they did with crop residues. In most cases, they sold part of the crop residues or used it for their own purposes (i.e., feeding their livestock, using it as firewood (e.g., the stalks of sorghum), making compost out of it, or using it to construct fence around the homestead). Although some respondents reported that they paid part of the crop residue as rent, this is not a common practice among the study population.

FTCs in an area as well as partaking in the FTC activities helped improve the access of participants of these FTCs to production inputs.³¹ These expenses have been deducted from total revenues in order to determine farm income from crops. Although more number of respondents from treatment category (66%)³² reported that crop production increased compared to last year's, from the analysis, it is evident that the comparison groups obtained relatively more income from crop production – but with high variability.

³¹ It appears that the FTCs help to significantly reduce production costs associated with crop and livestock production and management (as can be seen in Tables 2.4 & 2.5).

³² The corresponding figure for the comparison group is 50%.

Table 2.4 Average Value by Household of Crop Produced, Sold, Consumed, and Stored, Including the Value of Crop Residues Produced in the Last 12 Months, 2013

	Treatment	Comparison	t-test
Revenues (a to c):			
a. Total value of crop sold, consumed and stored (Birr)	30,216 (11,545)	59,341 (78,748)	3.48 ***
Value of crop sold (Birr)	19,842 (9,398)	39,380 (37,802)	4.82 ***
Value of crop consumed/stored (Birr)	10,374 (5,195)	19,961 (58,865)	1.54
b. Total value of crop residue produced (Birr)	1,645 (1,834)	1,205 (1,732)	1.89 *
Value of crop residue sold (Birr)	143 (188)	83 (771)	0.72
Value of crop residue paid as rent (Birr)	2.22 (21.08)	0	1.34
Value of crop residues consumed (Birr)	1,499 (1,909)	1,121 (1,377)	1.81 *
c. Total value of crop production (a + b) (Birr)	31,861 (12,437)	60,546 (78,811)	3.43 ***
d. Crop production expenses (Birr)	2,440 (978)	4,858 (6,276)	3.63 ***
e. Crop income (c – d) (Birr)	29,421 (12,458)	55,688 (77,442)	3.19 ***
f. Farm income (crop + livestock)	31,499 (12,276)	58,421 (78,025)	3.25 ***
Number of observations	90	160	

*** and * denote significance at 1% and 10% level, respectively.
Standard deviation in parenthesis

ii. Income from Livestock Production

Income from livestock production by household has been computed as the value of livestock products sold and rent from oxen obtained net of the cost of intermediate inputs associated with livestock production at household level (Table 2.5). Information in value terms has been directly provided by the respondents. As in the case of crops, livestock products have been investigated by asking the household head interviewed to provide details on each of them sold in the last 12

months.³³ Despite renting out oxen being an uncommon practice among the study population, this source of income, which interests only a few respondents in the comparison PAs, is also considered. Although the mean number of livestock owned is higher in treatment PAs compared to the comparison PAs, the analysis shows the existence of significant difference among the two study groups, with comparison PAs obtaining higher income compared to the treatment ones.

Table 2.5 Average Value by Household of Livestock Owned, Products Sold, Oxen Rented, and Intermediate Expenses Incurred in the Last 12 Months, 2013

	Treatment	Comparison	t-test
Number of own livestock at home	19 (9.45)	12 (8.88)	5.99 ^{***}
Total value of livestock owned (Birr)	36,297 (20,793)	20,442 (14,355)	7.10 ^{***}
Revenues (a):			
a. Total value of livestock products sold and oxen rented out (Birr)	2,876 (1,639)	4,422 (6,444)	2.23 ^{**}
Value of livestock products sold (Birr)	2,876 (1,639)	4,403 (6,411)	2.22 ^{**}
Value of rented out oxen (Birr)	0	19 (237)	0.75
b. Livestock production expenses (Birr)	798 (791)	1,689 (2,207)	3.69 ^{***}
c. Livestock income (a – b) (Birr)	2,078 (1,849)	2,733 (6,692)	0.91
Own livestock at home (%)	100.00	98.13	
Income from renting out oxen (%)	0.00	0.64	
Compared to last year, is your livestock output improved? (%)	75.00	76.16	
Number of observations	90	160	

^{***} and ^{**} denote significance at 1% and 5% level, respectively.
Standard deviation in parenthesis

The respondents also provided details on intermediate costs in livestock production, among which are feed, water, and veterinary services. These have been aggregated to compute the total

³³ Livestock products sold include meat, hides and skins, butter/cheese, milk, dungcakes, eggs, etc.

value of expenses. Like the case with income from livestock products sold, there is significant variation between the two groups in terms of expenses for livestock production, with the comparison groups reportedly spending more compared to the treatment ones. Similar to the case with lower crop production expenses in treatment PAs, it appears that livestock production expenses are also lower in the FTC-PAs probably due to the role of the FTCs in facilitating the access of smallholder farmers to production inputs and information.

Having computed crop and livestock incomes separately, agricultural (farm) income is derived by combining these two. Accordingly, the average farm income (Birr) realized by the treatment and comparison respondents is 31,499 and 58,421, respectively. The comparison group – despite the high within group variability – obtained higher income overall compared to the treatment ones, which is statistically significant ($t=3.25$, $p=0.001$).

As indicated above, the primary outcome variable on which treatment effect was computed is farm income. However, due to recognition of the special role of *chat* income on gross crop income (Table 2.6) – and thereby on farm income – other outcome variables, such as farm income (with and without *chat*), crop income (with and without *chat*), and livestock income were also used. Table 2.6 also indicates the level and composition of crop production in the two areas.

Table 2.6 Contribution of Different Crops to Gross Crop Income in 2013 (Percentage Share)

	Gross crop income		Total
	Treatment	Comparison	
Cereals	41.23	23.41	29.83
Vegetables	0.88	0.11	0.39
Root crops	11.78	7.22	8.86
Chat	39.97	63.64	55.12
Crop residues	4.76	4.33	4.48
Others	1.38	1.29	1.32
Total	100.00	100.00	100.00

Note: crop production expenses not included.

The distribution of the average values of these outcome variables in the two groups is given in Table 2.7. Accordingly, except for income from livestock production, there are statistically significant variations between the two groups. Moreover, whenever income from *chat* – which accounts for 40% and 64% of the gross crop incomes of treatment and comparison groups, respectively (Table 2.6) – is excluded, the treatment group performs better than the comparison group, on average.

Table 2.7 Distribution of the Average Values (Birr) by Household of Outcome Variables in the Study Sample in 2013

Outcome variable	Treatment	Comparison	Difference	t-test
1. farm income	31,498.63 (1,293.96)	58,420.96 (6,168.42)	26,922.33 (8,288.33)	3.25***
2. farm income (excluding <i>chat</i>)	20,135.42 (1,219.91)	11,653.44 (1,077.61)	- 8,481.97 (1,703.68)	- 4.98***
3. crop income	29,420.65 (1,313.19)	55,687.50 (6,122.31)	26,266.85 (8,229.24)	3.19***
4. crop income (excluding <i>chat</i>)	18,057.44 (1,248.76)	8,919.99 (831.15)	- 9,137.45 (1,450.76)	- 6.30***
5. livestock income	2,077.98 (194.86)	2,733.46 (529.04)	655.48 (720.93)	0.91
Number of observations	90	160		

*** denote significance at 1% level.
Standard errors in parenthesis

2.3.4 Estimation of the Average Effect of Modular Training on Farm Income and Other Outcome Variables

Having estimated the propensity score and constructed the outcome variables, the next step is the computation of the average effect of modular training on farm income of trained farmers (also referred to as the average treatment effect on the treated or ATT). The estimation of the average effect of training was performed on Stata using the command *psmatch2*, which implements a range of propensity score matching techniques (Leuven & Sianesi, 2003). The results are presented in Table 2.8. The average effect of modular training on farm income (including *chat*) of trained farmers is highly significant but negative. This means that untrained farmers performed better compared to the trained ones, with average farm income (including *chat*) ranging from about 24,854 birr/year (*Nearest Neighbor Matching*) to 28,263 birr/year (*Kernel – Epanechnikov Matching*). However, when *chat* income is excluded from farm income – due to

the fact that there is no extension package and FTC training for promoting *chat* – the impact of training on farm income becomes positive and highly significant, ranging from 9,557 birr/year (*Nearest Neighbor*) to 10,388 birr/year (*Kernel – Gaussian*), on average.

Table 2.8 Estimation of Average Treatment Effect on the Treated (ATT) Using Nearest Neighbors, Kernel, and Radius Matching

Outcome variable	(1)	(2)		(3)
		(2a)	(2b)	
1. Farm income per household (Birr)	– 24,854.46 (9,105.19) ***	– 27,344.03 (7,597.10) ***	– 28,262.78 (8,554.61) ***	– 27,536.20 (8,205.94) ***
2. Farm income per household (Birr) excluding <i>chat</i>	9,557.47 (2,561.70) ***	10,387.53 (2,129.27) ***	9,846.20 (2,685.35) ***	10,233.93 (2,509.94) ***
3. Crop income per household (Birr)	– 24,511.34 (8,932.87) ***	– 27,045.93 (7,404.51) ***	– 28,003.22 (8,353.19) ***	– 27,316.92 (7,994.48) ***
4. Crop income per household (Birr) excluding <i>chat</i>	9,900.59 (2,487.14) ***	10,685.63 (2,165.25) ***	10,105.76 (2,713.51) ***	10,453.21 (2,571.27) ***
5. Livestock income per household (Birr)	– 343.12 (837.34)	– 298.10 (565.04)	– 259.56 (666.38)	– 219.28 (647.25)

*** denote significance at 1% level (significance test based on z statistics).

Bootstrap standard errors with 50 replications in parenthesis

Note: 73 treated and 137 untreated individuals on common support (for the purpose of drawing the *psgraph*). (1) Nearest Neighbors matching; (2) Kernel matching; (2a) Normal (Gaussian) Kernel; (2b) Epanechnikov Kernel; (3) Radius matching.

By disaggregating farm income into its components crop income and livestock income, the ATT is also estimated. Similar to the effect of training on farm income (including *chat*), the causal effect of training on crop income (including *chat*) is significant but negative. However, when income from *chat* is excluded from crop income, the effect becomes positive and significant, ranging from 9,901 birr/year (*nearest neighbor*) to 10,686 birr/year (*Kernel - Gaussian*), on average. Although not significant, the impact of modular training on livestock income appears to be negative by all the three matching techniques implemented. Though the dominant contributor

of farm income in the study area is crop production, the livestock sector also has about 8% and 5% share of total farm income in treatment and comparison groups, respectively.

2.3.5 Matching Quality (Balance Checking) and Common Support

In this study, various procedures were implemented to check the quality of matching in terms of balancing the covariates used during the matching process. These are: mean comparisons between treatment and comparison groups (before and after matching), standardized bias, and overall measures of covariate imbalance (Rosenbaum and Rubin, 1985; Caliendo and Kopeinig, 2008). The mean comparisons between the two groups before matching indicate the existence of significant differences in terms of education, household size, assets, organizational membership, experience in advisory services, and number of educated members of the household. However, the post-matching t-test results indicate that none of these variables show significant differences, suggesting that the different matching procedures enabled balancing of the covariates in the two groups (Table 2.9). However, to what extent the different matching procedures reduced/eliminated original bias cannot be visible from t-test results only.

Looking at the absolute standardized percentage bias in the unmatched sample (Table 2.10), the variables education, household size, assets, experience in AASs, organizational membership, and number of educated members of the household show huge values. However, the post-matching results show that there is a significant reduction in the standardized percentage bias for all of these covariates. In general, most of the covariates used for matching have a standardized percentage bias of < 5%. The bottom half of Table 2.10 contains the overall measures of covariate imbalance, such as Pseudo- R^2 values and mean/median bias as summary indicators of the distribution of the bias. Accordingly, the original mean bias of 27.9 in the unmatched sample

is reduced to < 3.5%. Similarly, the significant reduction of the Pseudo-R² original value in the unmatched sample (i.e., 0.21) also guarantees that the matching procedures were effective in reducing the existing bias and creating a comparable control group.

Table 2.9 Matching Quality Analysis: T-Test (Before/After) between Unmatched and Matched Treated and Control Groups

	Unmatched (mean and t-test)			Matched (t-test: treated and control groups)			
	Treated	Control	t-test	(1)	(2a)	(2b)	(3)
Age of the household head (=1 if >18 years)	0.99	0.99	-0.41	0.13	-0.41	-0.36	-0.39
Gender (=1 if male)	0.98	0.96	0.66	-0.21	-0.03	0.04	-0.00
Education (=1 if attended some formal school)	0.30	0.68	-6.10 ***	-0.07	-0.39	-0.22	-0.27
Household size (number of people in the house)	3.30	3.99	-2.47 **	0.20	0.36	0.23	0.37
Assets (number)	3.01	2.61	3.22 ***	0.23	0.17	-0.10	0.05
Farming experience (in years)	18.29	18.62	-0.36	-0.25	-0.20	-0.24	-0.10
Experience in AASs (in years)	15.24	13.63	1.69 *	-0.31	-0.13	-0.35	-0.16
Land holding (=1 if ≥ 1 ha)	0.32	0.36	-0.54	-0.07	0.04	-0.09	-0.03
Organization (=1 if a member of at least one organization)	0.99	0.86	3.38 ***	-0.00	0.34	-0.12	-0.06
Educated members (=1 if ≥ 1)	0.89	0.95	-1.80 *	-0.35	-0.20	-0.30	-0.32

***, ** and * denote significance at 1%, 5% and 10% level, respectively.

Note: (1) Nearest Neighbors matching; (2) Kernel matching; (2a) Normal (Gaussian) Kernel; (2b) Epanechnikov Kernel; (3) Radius matching.

Table 2.10 Matching Quality Analysis: Standardized Percentage Bias (Before/After Matching) and Mean/Median Bias as Summary Indicators of the Distribution of the Bias

	Unmatched		Matched		
	(1) to (3)	(1)	(2a)	(2b)	(3)
Age of the household age (=1 if >18 years)	-5.2	2.9	-6.2	-6.6	-7.0
Gender (=1 if male)	8.9	-3.2	-0.5	0.7	-0.1
Education (=1 if attended some formal school)	-80.6	-1.2	-6.7	-3.8	-4.7
Household size (number of people in the house)	-34.1	3.1	5.8	3.6	5.9
Assets (number of productive assets)	43.9	3.6	2.6	-1.5	0.8
Farming experience (number of years)	-5.1	-4.1	-3.4	-4.0	-1.7
Experience in AASs (number of years)	22.4	-5.0	-2.1	-5.8	-2.6
Land holding (=1 if ≥ 1 ha)	-7.2	-1.2	0.8	-1.5	-0.6
Organization (=1 if a member of an organization)	49.5	0.0	2.9	-0.9	-0.4
Educated members (=1 if ≥ 1)	-22.5	-6.0	-3.4	-5.3	-5.6
Mean	27.9 (24.5)	3.0 (1.9)	3.4 (2.2)	3.4 (2.1)	3.0 (2.6)
Median	22.4	3.2	3.2	3.7	2.2
Variance	602.1	3.4	4.7	4.5	7.0
Pseudo R2	0.21	0.003	0.003	0.003	0.003
LR chi2	68.4***	0.53	0.70	0.63	0.60

Standard deviation in parenthesis

Likelihood-ratio (LR) test of the joint insignificance of all the covariates: *** denote significance at 1% level.

Note: (1) Nearest Neighbors matching; (2) Kernel matching; (2a) Normal (Gaussian) Kernel; (2b) Epanechnikov Kernel; (3) Radius matching.

Regarding the common support condition, the *psgraph* (Figure 2.3) indicates the existence of a good overlap between the propensity scores of treatment and comparison respondents. Out of the 250 sampled respondents (i.e., 90 in treatment and 160 in comparison groups), 73 and 137 treatment and comparison respondents are on common support, respectively.

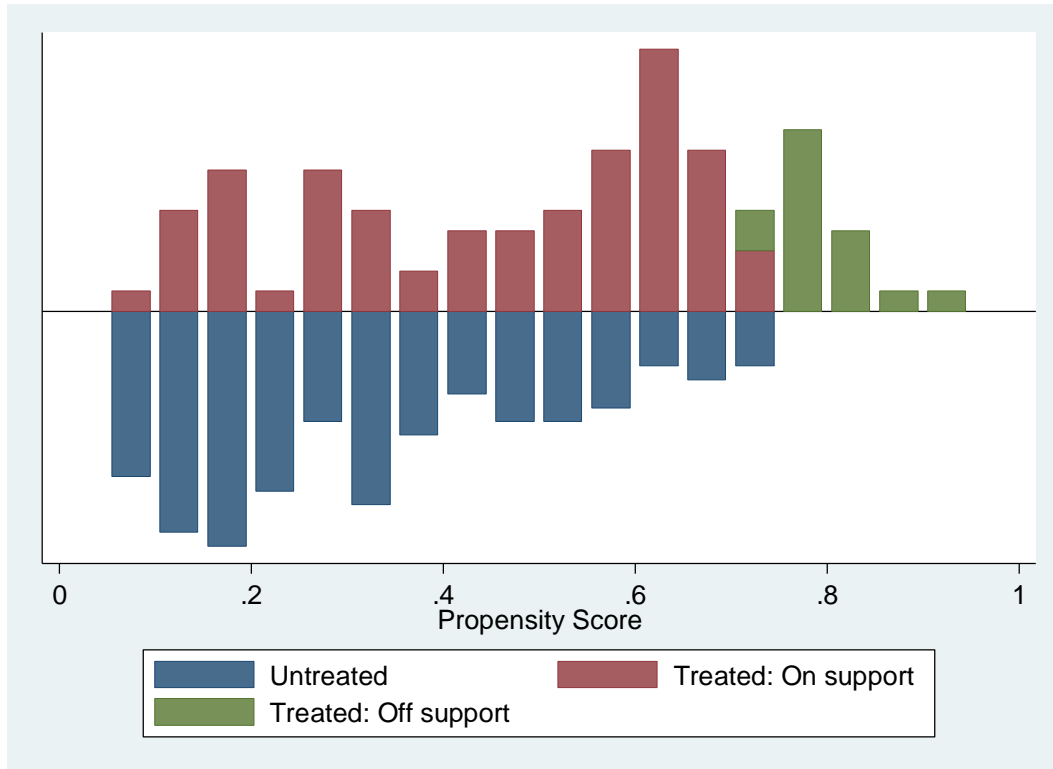


Figure 2.3 Histogram of Estimated Propensity Score between Participants (Treated) and Nonparticipants (Untreated) of FTC-Based Modular Training

2.4 Discussion

2.4.1 Determinants of Participation in FTC-Based Modular Training

The results of the logistic regression (Table 2.3) indicate that level of education, household size (in terms of number of people), asset ownership, experience in farming, and organizational membership are significant determinants of participation in modular training. In terms of education, the result shows that uneducated members of the community have better chance of participating in the training programs compared to the educated ones. This targeting might be of particular importance to rural dwellers as many of them do not have any formal education backgrounds. Household size is negatively related to participation probably because the more the number of members of the household, the better and diversified the opportunity to earn income outside of farming. For example, in the survey it is observed that some members of large households support their household by involving in off-/non-farm activities. This might have led to a reduced interest and participation in modular training.

Ownership of productive assets positively affects participation in modular training. This may be due to the fact that the implementation of knowledge and skills obtained from such training requires some assets. Hence, whereas those who have a relatively better stock of assets at their disposal are more likely to enroll in such training, those who are asset constrained do not. Finally, (formal/informal) organizational membership also positively influences participation in modular training. Since rural development requires the availability of complementary inputs (other than training), membership in and access to such organizations can play a great role in the adoption of productivity enhancing technologies and practices.

2.4.2 Impact of Modular Training on Farm Income of Rural Households

The FTC-based modular training programs can play positive roles in impacting rural people's livelihoods in Ethiopia. Although it has been shown recently that the general AASs in the country have positive contributions towards reducing headcount poverty and increasing consumption growth (Dercon et al., 2009), such analysis is based on the advice/visit component of AASs. It considers neither on-farm training – short-/long-term – nor (method/result) demonstration activities of the general AASs. Moreover, there is no specific attention given to the growing FTC-based provision of AASs in the country.

Of the few publications dealing with the FTC-based AASs in Ethiopia are IFPRI (2010) and Tefera et al. (2011). In their detailed assessment of AASs in the country, IFPRI (2010) investigate the pros and cons associated with public AASs, including the policy environment, FTC establishment and functioning, Agricultural Technical and Vocational Education and Training (ATVET) colleges, in order to provide feedbacks and recommendations on how to improve performance of the overall system. Likewise, though at smaller scale compared to IFPRI (2010), Tefera et al. (2011) made an assessment of IPMS-sponsored FTCs in IPMS Pilot Learning Woredas. The researchers document and provide evidence on the extent and efficiency of resource use (physical, human and other resources), training programs of the FTCs, and good practices and lessons for future reference and replication.

Evidently, both of these investigations do not provide any empirical evidence on the effectiveness of the FTCs in terms of improving rural households' income or other welfare indicators. The present study contributes in this regard by shading light on the causal effect of FTC-based modular training on farm incomes of rural people in Haramaya district of Oromia

region. In particular, this issue is investigated using various matching estimators in order to arrive at robust results of impact. Moreover, a range of matching quality tests is provided to ascertain the robustness and performance of the different matching techniques implemented. Before embarking on the discussion of impact related results, however, it seemed compelling to comment on a dominant crop in the farming system of the study area.

The estimation of the average treatment effect on the treated (ATT) shows the unique role of *chat* income on total farm income in the study areas. *Chat* (*Catha edulis*) is a perennial crop whose leaves and tender stems are chewed for their stimulating effects (Gebissa, 2010; Taffesse et al., 2011). The production of this crop is rapidly expanding in the country. For example, *chat* production is replacing cereal production in Habro district, western Hararghe zone. In this district, *chat* income represents 70% of household income (Feyisa & Aune, 2003). In the present study in east Hararghe zone, a similar pattern as in the west is observed. *Chat* is one of the dominant cash crops in the study area whose contribution amounts to about 55% of the gross crop income of the sampled respondents – 40% and 64% in treatment and comparison groups, respectively. This suggests that the respondents in the comparison areas obtain more income from this crop compared to those in the treatment areas. Moreover, during the field survey, it is observed that in general the comparison areas produced more high quality *chat* compared to the treatment areas.

Ironically, one might say, there is neither an extension package to promote the production and marketing of this crop nor is there a modular training organized to exploit the full potential of the crop. This is mainly because – in spite of its tremendous contributions to household income and social interactions – it is regarded as a stimulant crop and does not enjoy an official government support like other crops. Hence, a household income accounting that does not take into account

the specific circumstances related to *chat* production may not be as realistic as one might want it to be. Therefore, in the analysis, the ATT is estimated with and without this crop. However, the results pertaining to the policy aim of setting up FTCs in the study areas are discussed here.

Looking at farm income (excluding *chat*) in columns (1) to (3) of Table 2.8, one notices a clear convergence of the results obtained using the two Kernels, i.e., columns (2a) and (2b) and Radius matching. In general, all the matching algorithms implemented indicate a positive and statistically highly significant gain of farm income (excluding *chat*) by the participants of the modular training, which is between Birr 9,557.47 and Birr 10,387.53 per household, on average. Moreover, the ATT estimation based on the disaggregation of farm income into its components – crop and livestock incomes – also reveals quite similar patterns for crop income. The average crop income (excluding *chat*) is positive and statistically highly significant by all the matching methods, ranging from Birr 9,900.59 to Birr 10,685.63 per household. In terms of livestock income, however, the matching results are negative but not significant, suggesting a relatively poor performance of trained farmers in this regard. It has to be noted, nonetheless, that income from the livestock sector accounts for only about 6% of total farm income. This may partly be due to the definition of livestock income used in this study – income from the sale of livestock products and oxen rented out.

The results pertaining to farm income are broadly consistent with previous research conducted on the roles of farmer field schools (FFSs) on agricultural income in the country. Todo and Takahashi (2013) investigated the impact of FFSs established for the purpose of promoting participatory forest management and protection in Belete-Gera Regional Forest Priority Area in Oromia region. Their grand conclusion is that participation in the FFSs increases real income per worker by approximately US\$ 60-160, on average, in two years. The observed increase in

income, the authors justify, is related to the use of improved agricultural technologies and practices promoted by and taught at the FFSs. Although the FFSs approach is a small Japanese funded project, the results of the impact evaluation indicate their importance in improving agricultural income of rural people in the country.

The findings in the present study are also related to the output of research conducted on FFSs in East Africa (Kenya, Uganda, and Tanzania) and National Agricultural Advisory Services (NAADS) of Uganda. In the study in East Africa, Davis et al. (2012) evaluate the implications of FFSs for improved agricultural productivity and poverty reduction, and show that participation in FFSs has positive impact on crop production and income. Furthermore, the researchers indicate that the causal effect of FFSs varied among gender, education and land size categories. The research on the direct and indirect effects of the Ugandan NAADS program (Benin et al., 2011; Benin et al., 2012b) also documents positive effects on household agricultural income. In particular, Benin et al. (2011) show an increase of 37-95% (direct participants) and 27-55% (indirect participants) per capita gross agricultural revenues between 2004 and 2007. In general, the present study shows that FTC-based modular training has a positive and significant impact on household's farm income. The results are reasonably robust to changes in the matching methods. Moreover, the fact that all the impact estimates pass the various balancing tests indicates the plausibility of the results in the face of the assumptions underlying the PSM procedure.

2.5 Conclusion and Recommendation

In the present study, the causal effects of participating in an FTC-based modular training on household income are investigated. Through collecting data specifically for the purpose of impact evaluation and implementing rigorous evaluation methods, some previous criticisms on data quality and methodological rigor (e.g., Birkhaeuser et al., 1991; Waddington et al., 2010) are addressed. Moreover, evidence on the impact of a new approach to the organization and delivery of public AASs, i.e., the FTC-based AAS provision, is provided. The results in the present investigation could contribute towards the renewed discussion on the role of AASs in fostering agricultural development (Birner et al., 2009). In particular, by providing evidence on the impact of agricultural training, the study shows the need for due emphasis on this component of AASs in equipping farm households with the necessary knowledge and skills related to farming enterprise.

One of the findings relates to the decision to participate in the FTC-based training. The participation equation reveals some fundamental covariates influencing the participation of rural households in modular training. Accordingly, it is found that education, household size, asset ownership, experience in farming, and membership in organizations significantly affect the decision to take part in such training. Although the involvement of development agents in the selection of farmers into advisory services was pronounced in the country in the past (e.g., Belay, 2003), a takeover of this role by the *Kebele extension unit* in the FTC-based AASs is observed. Nonetheless, self-selection is limited in the present setting as well. This calls for an encouragement design to let the beneficiaries take the upper hand in the selection process as well as a reduced role of the development agents and/or *Kebele extension unit* in the process.

Regarding the crux of the investigation, i.e., the impact estimates, a good overall convergence of results is obtained using different matching methods, suggesting the efficiency of the matching procedures implemented. More specifically, all the matching techniques indicated a positive and significant gain of average farm income (excluding *chat*) by the participants of the FTC modular training. Furthermore, the analysis based on the breakdown of farm income into its components – crop and livestock – also shows that participation in modular training resulted in a positive and significant gain of crop income (excluding *chat*), on average. However, it appears that the results for livestock income are negative, though not significant.

In order to have confidence on the results of the matching, various matching quality assurance analyses were conducted, including post-matching t-tests of covariates, standardized percentage bias, and Likelihood-ratio test. First, it is observed that the t-tests after matching were non-significant for all the covariates involved in the matching process. Second, the standardized percentage bias for most of the covariates involved in the matching is found to be less than 5%. Third, it is found that the overall measure of standardized mean bias is less than 3.5%. Fourth, the post-matching Pseudo-R² values are very low compared to that for the unmatched sample. Finally, the non-significant Likelihood-ratio test result after matching also suggests the quality of matching implemented. In general, all the measures of matching quality assurance indicate that the matching procedures pass the test of balancing and that the matching methods are efficient.

Based on the results of this study, the following recommendations are suggested to improve the FTC-based provision of AASs in the study area in particular, and the country in general. To start with, there should be a reinforced awareness creation campaign regarding the importance of an FTC-based modular training, so that rural households may be committed in their participation in all aspects of the FTCs. In particular, young farmers need to be encouraged to partake in the

training. Since the default household heads are males, there should be a provision to target young and women farmers (although they are not household heads). As it is now, the modular training pays little attention to female farmers in the community. One operational suggestion in this regard requires a general shift of focus from household heads to members of the household. Furthermore, since the criterion of devotion can potentially exclude many young farmers, a minimum experience in farming and in general AASs should be sufficient to qualify young entrepreneurs to take part in the training.

It is also important to pay special attention to households who do not have enough family labor to work on the farm. Likewise, since asset ownership positively determines participation in training, a particular focus needs to be given to asset-constrained households. Similar to the case with assets, membership in organizations also plays crucial roles. Hence, not only should those with limited access to organizations/associations be specially considered, but also should be advised/oriented on how to become a member in such organizations (e.g., credit and saving, microfinance, farmers' associations, etc.).

In terms of the impact of training on farm income, although positive and significant impacts were found, there is a need to strengthen and support vegetable and root crop production. Moreover, since the results suggest the dominant role of *chat* production, the FTC-based AASs needs to encourage farmers to produce more cereals (such as sorghum and maize) by providing training and production inputs on them. Similarly, given that the study area is favorable for the production of small ruminants and cattle, there should be an increased effort towards the promotion of improved livestock production and management. In addition to these, other stakeholders of the agricultural innovation process (such as market intermediaries and

collaborators) should work towards broadening market incentives for farmers in the production of staple food and cash crops other than *chat*.

Finally, although the evaluation of the impact of modular training was based on rigorous statistical procedures, covering pre- and post- matching issues, there are still potentials for improvement and further investigation. The first aspect for improvement requires the inclusion of more FTCs from the different agro-ecological zones of the country and evaluate country-level impact estimates on selected outcome indicators, such as productivity, farm income, poverty, consumption expenditure, and asset accumulation. Such a mega-project can collect baseline data on important covariates, including outcome variables, and can be amenable to the use of various evaluation techniques. One potential technique can be combining propensity score matching with difference-in-difference, which is superior to using each technique single-handedly. Another issue for future consideration can be a move away from individual level evaluations to aggregate levels (such as village/peasant association, district, etc) or a combination of both of them. Whereas the former can enable a researcher to better capture spillover effects of a program or an intervention as well as any unobserved heterogeneity between treatment and comparison groups, the latter can help in checking the complementarity of individual and group-based approaches to impact evaluation.

3 Smallholder Farmers' Willingness to Pay for Agricultural Advisory Services: Investigating the Effect of Solemn Oath in Mitigating Hypothetical Bias in Best-Worst Discrete Choice Experiments in Haramaya District, Eastern Ethiopia

3.1 Introduction

In the majority of developing countries, the agricultural sector has been viewed as the most important platform where the battle for long term economic development can be won or lost. This is evident considering the contribution of the sector to food security, poverty reduction, and general economic development (Green et al., 2005; World Bank, 2007). Cognizant to this, there has been a concerted effort to support the sector and its collaborating institutions such as agricultural research, education, and advisory services. Nevertheless, in spite of evidence that investment in agriculture has positive impacts on agricultural growth and poverty reduction, since 1980s there has been a stagnation or decline in public spending on agriculture in most developing countries (Akroyd & Smith, 2007). Hence, some alternative ways of financing the sector are being sought and implemented. As a key supporting institution, agricultural advisory services (AASs) have been the subject of reform (Birner et al., 2009), which focuses on making them 'pluralistic' and effective in meeting the needs of smallholder farmers and other clients (Dinar, 1996; Rivera & Alex, 2005).

AASs encompass the entire set of organizations involved in supporting and facilitating the access of farmers to technical knowledge, skills, improved technologies, credit, markets, and other rural services (Birner et al., 2006; Swanson, 2008). As such, AASs contribute to local development and improved rural livelihoods in at least four ways: by transferring technical knowledge from research centers to farmers, and reporting farmers' feedback to researchers; by demonstrating the

performance of improved technologies (such as high yielding crop varieties, fertilizers, insecticides); by conducting seminars and tailor-made training; and by advising clients on their farm decision-making and problem-solving endeavors (Birkhaeuser et al., 1991). In terms of the historical evolution of AASs, there is a perceptible transition from a simple technology transfer paradigm to participatory approaches to market-oriented and cost-sharing models (Birner et al., 2006; Swanson & Rajalahti, 2010). However, one also notices a combination of two or more of the approaches being used depending on the needs and priorities of the clients as well as the institutional capacity of the implementing organizations.

One way of addressing the varied needs and priorities of farmers in agricultural development is through diversifying the institutional options of AAS provision (Rivera & Alex, 2004). The term ‘pluralistic’ has been coined to capture the emerging diversity of institutional options in providing and financing AASs (Birner et al., 2009). For a long time, AASs have been viewed as public goods and have been financed mainly by the public sector both in developed and developing countries (Birkhaeuser et al., 1991; Dinar, 1996). However, there have been extensive debates worldwide on the redefinition of the role of the public sector in the provision of these services (Ozor et al., 2011). Consequently, many developed countries have followed a paradigm shift towards decentralization, contracting/outsourcing, public-private partnerships, privatization, and increased involvement of non-governmental organizations (NGOs) in AAS provision (Rivera & Alex, 2004; Birner et al., 2009). This has resulted in an effective reduction of public expenditures on these services, and greater involvement of private and other stakeholders.

The paradigm shift has also entailed a fee-for-service arrangement (Dinar, 1996), where clients of AASs share, or pay in full, the cost of services and make service providers accountable for the

quality of services. In many developing countries, however, the fundamental premise of public AASs – that low-income farmers are unlikely to obtain technical information and technology unless it is provided by the public sector – largely dictates the provision of these services (Farrington, 1994). The Ethiopian national AAS is no exception to this. Since its formal inception in the early 1950s, it has been offered and financed primarily by the public sector. Although the private sector and NGOs are becoming increasingly important, they are often left out of AAS initiatives or coordination is weak (Byerlee et al., 2007; Ethiopian Economic Association, 2006), which may explain, in part, the poor performance of AASs for many decades. Moreover, there is hardly any empirical evidence available documenting the existence or lack of favorable environment for diversified provision of AASs in the country.

For example, previous studies on AASs in Ethiopia have focused on the farmers' attitudes towards AASs (Beyene, 2003); the use of production inputs such as seeds and fertilizer (Spielman et al., 2011); the involvement of farmers in AASs and constraints in the adoption of new technologies (Abesha et al., 2000; Belay, 2003; Belay & Abebaw, 2004; Gebremedhin et al., 2006); the institutional development, adopted approaches, and involved processes (Gebremedhin et al., 2006; Davis et al., 2009; Moguees et al., 2009; IFPRI, 2010); the decentralization of AASs and satisfaction with services provided (Cohen & Lemma, 2011); the evolving roles of public and private sectors in intensifying cereal production (Spielman et al., 2010); the political economy of AASs (Berhanu, 2012); the role of institutional services of credit, input supply, and advisory services (Gebremedhin et al., 2009); the work environment of agricultural advisors (Belay & Abebaw, 2004); the impact of AASs on various outcomes (Dercon et al., 2009); and, the role of learning from development agents (DAs) versus neighbors (Krishnan & Patnam, 2012).

However, considering the current economic crisis, budget limitations, the large number of people relying on agriculture for their livelihoods, the inefficiencies in public AAS provision, and the specialized needs of market-oriented producers, it is important, on the one hand, to evaluate the effectiveness of current AAS provisions in impacting the livelihoods of rural dwellers, and on the other, to consider alternative ways of providing and financing AASs. In an era where partnerships are being forged among public, private, NGOs, and other stakeholders in the provision of AASs, there is a heightened need to investigate the ability and willingness of clients to co-finance the provision of AASs. In this regard, there is very little empirical evidence (except those mentioned in the ‘literature review’ section) available to inform policy makers on how to effectively pluralize AAS provision in the country. Moreover, there is a dearth of empirical evidence to suggest which components of AASs to commercialize or privatize in the future. These investigations will contribute towards transforming the current advisory service to one that is demand-driven, efficient, and sustainable in the face of social, economic, and environmental challenges.

Recognizing the scarcity of empirical evidence on the valuation of AASs in the country and elsewhere, this choice experiment is proposed to estimate the willingness-to-pay (WTP) of smallholder farmers for improved FTC-based AASs. In doing so, the study seeks to provide explanations and some relevant recommendations on whether a fee-for-service arrangement can be a potential way out to diversify the provision of effective, coordinated, and demand-driven AASs in the country. The approach followed in the current study is as follows. First, the best-worst method is used to generate preference data (Louviere & Islam, 2008; Rose, 2014) combined with the solemn oath technique to deal with hypothetical bias (Jacquemet et al., 2013). To the best of my knowledge, there is not any previous study in the country, and elsewhere in

Africa, which employs such a technique. Second, following the recommended practices in choice experiments (Louviere et al., 2000; Hensher et al., 2005; Scarpa et al., 2011) and estimating both the multinomial logit (MNL) and random parameters logit (RPL) models, the WTP estimates are derived for advice, training, and demonstration, the three important services describing improved AAS provision through farmers' training centers (FTCs).

The remainder of the paper is structured as follows. In the next section, some literature review is provided pertaining to, among others, the best-worst approach to getting preference data and strategies to deal with hypothetical bias in choice experiments. Following this, the research methodology section presents the design and methods used in the study. Immediately following this is the section where the main findings of the study are presented and discussed, before concluding with some recommendations.

3.2 Literature Review

This section presents and discusses selected literature reviewed in the context of the present study. The starting point of the review is the discussion on the stated preference techniques with a focus on choice experiments. Following this, some applications of choice experiments in the context of developing countries are presented. Next, some insights are provided into the best-worst preference elicitation mechanism, which is implemented in this study. After this, the issue of status quo effect and the use of hypothetical baselines in choice experiments is discussed. Finally, the concept of hypothetical bias and strategies to deal with it are presented, focusing on the cheap talk script and solemn oath.

3.2.1 The Stated Preference Techniques

The stated preference (SP) techniques, such as the contingent valuation (CV) method and choice experiments (CEs), enable the estimation of the value of goods and services in the absence of a market for such goods and services (Bateman et al., 2002). Although CV was the most frequently used SP technique for eliciting preference, there is an increasing trend in the use of CEs in several disciplines, including marketing, health, transport, and environmental resource economics (Louviere et al., 2000; Bateman et al., 2002). The use of CEs is well established in developed countries, but its application in the context of developing countries is relatively a recent phenomenon.

A choice experiment enables researchers to estimate the value of a good or service by analyzing the attributes of the good or service in question. It requires respondents to state their choice over different combinations of attributes of alternatives and choose the one which they prefer (Mangham et al., 2009). Compared to other SP techniques that require respondents to rank or

rate alternatives, CEs present a reasonably straightforward task which more closely resembles a real-world decision (*ibid.*).

It is generally acknowledged that CEs are the most appropriate SP methods for the valuation of benefits generated from multiple characteristics and functions (i.e., attributes) of non-market goods and services. In relation to WTP for services, CEs present the possibility to estimate the marginal WTP for specific attributes and make comparisons among them (Bennett & Birol, 2010). In comparison with CV, which estimates WTP for a good or service as a whole, this more detailed information on WTP by attribute may be useful for those involved in policy decisions and setting resource allocation priorities (Mangham et al., 2009).

Nevertheless, there are some concerns over the use of CEs in preference elicitation. First, the assumption that a given good or service is the same as the sum of its attributes may not hold always. Second, the choice of study design may affect the welfare value estimates obtained through CEs. Third, in situations where several choice sets are used for a study, respondents may face cognitive burden and hence the quality of data obtained may be poor. Finally, it may be difficult to derive values of components of a policy or program implemented in a sequence (Bateman et al., 2002).

3.2.2 Application of Choice Experiments in Developing Countries

As indicated in the introductory section, the use of CEs in developing countries is limited. More importantly, the use of CEs in estimating the WTP for AASs is scarce. Among the recent studies employing CEs in developing countries (in a range of disciplines), the following can be mentioned. In Ethiopia, Carlsson et al. (2005), Agimass & Mekonnen (2011), and Tarfasa & Brouwer (2013) studied agricultural extension packages, fishermen's WTP for fisheries and

watershed management, and public benefits of urban water supply improvements, respectively. Other examples of application of CEs include: Ghana (Kruk et al., 2010), Zambia (Meenakshi et al., 2012), and Tanzania (Kruk et al., 2009). In addition to these, there are multicountry studies such as Blaauw et al. (2010) in Kenya, South Africa and Thailand, and Probst et al. (2012) in three West African cities.

The relatively limited application of CEs in developing countries may be partly due to the challenges associated with its application in such countries. Conducting CEs in developing countries can involve issues not encountered in developed countries (Mangham et al., 2009). Hence, the authors suggest cautious approaches towards identifying attributes, assigning attribute levels, designing choice sets, generating and pre-testing the questionnaire, and analyzing CEs data. Other researchers emphasize the importance of focus group discussions, the choice of reliable and realistic payment vehicle, the use of face-to-face interviews, the use of visual aids, the use of carefully selected and trained enumerators, the inclusion of follow-up questions in the survey instruments (to understand how or why respondents choose their preferred alternatives), and the advantage of keeping task complexity to the minimum for a successful implementation of CEs in developing countries (Bennett and Birol, 2010). These issues are addressed in greater detail in the methodology section.

3.2.3 Best-Worst Approach to Preference Elicitation

Unlike traditional choice experiment, where respondents are asked to state their best-preferred alternative among competing products or services, the best-worst approach to preference elicitation (Louviere and Islam, 2008; Lancsar et al., 2013; Rose, 2014) requires respondents to state their most- and least- preferred alternatives in a given choice task. Although this method is

gaining popularity recently, Scarpa et al. (2011) indicate that its application has yet to be explored in the context of nonmarket goods and services comprising various attributes. The basic idea behind the best-worst approach is the presentation of choice sets to respondents and then asking them to choose first their most-preferred alternative from all the options, and then choose their least-preferred alternative from the remaining options. Alternatively, the respondents can also be asked to choose their least-preferred option first and then the most-preferred one.

Some researchers and practitioners of SP emphasize the importance of best-worst approach compared to the traditional discrete choice setting (i.e., choosing the most favored alternative from a given set of options) or ranking/rating exercises for a number of reasons. First, best-worst approach can help to get more data on respondents' preferences than the 'pick one' traditional discrete choice (Marley & Louviere, 2005; Flynn et al., 2007; Vermeulen et al., 2010; Scarpa et al., 2011). Most importantly, in small samples, this approach can be an important way to get more data on each respondent's preferences (Cerroni et al., 2014). Second, the technique helps respondents to identify extreme alternatives and thereby enable them to respond more consistently throughout the choice task (Marley and Louviere, 2005; Vermeulen et al., 2010). Finally, in terms of cognitive effort, a choice set framed in best-worst approach is not only easy for respondents to respond, but also easy for researchers to implement (Marley and Louviere, 2005; Louviere and Islam, 2008).

On the basis of these considerations, the best-worst approach is implemented in the present choice experiment survey by asking each respondent to choose the most-preferred alternative first, and then the least-preferred one from the remaining options. Since estimations based on the opposite order, from worst-preferred to best-preferred, will not produce equivalent results

(Lancsar et al., 2013), the study required each respondent to follow the best-worst sequence consistently throughout the choice task.

3.2.4 The Status Quo Effect and the Use of Hypothetical Baselines in Choice Experiments

In addition to the problems of CEs mentioned earlier, there is what is known as the status quo (SQ) effect. Put simply, the SQ bias refers to the disproportionate choice of the SQ alternative (i.e., the current situation or any future scenario/baseline constructed by a researcher, which is usually associated with no payment for the goods/services) by respondents of CEs (Meyerhoff & Liebe, 2009). According to these researchers, there are many explanations for the occurrence of the SQ bias, including “mistrust in the providing organization, complexity of the choice task, or protest against the survey” (p. 516). In the context of the present study, the SQ effect may be of great concern. Since AASs have been provided free of charge, many respondents may resort to the SQ option for economic reasons. Moreover, the fact that not all respondents currently enjoy the same quality and quantity of AASs suggests a further difficulty to present a standard SQ option for all of them. In such cases, the use of hypothetical baseline may be a way out, although this may also be problematic.

Hypothetical baselines are descriptions of hypothetical (not actual) states which are intentionally introduced by the CEs researcher to move the respondent away from his/her current state. The researcher then poses a valuation scenario based on this new hypothetical baseline (Whittington & Adamowicz, 2010). There are at least four reasons why SP researchers may want to use hypothetical baselines. First, to understand what respondents would do if circumstances changed and people were confronted with a new policy or program. Second, to ask a sensible valuation question. Third, to conveniently implement the CEs (i.e., for the convenience of the researcher).

Finally, to remove himself from political or social controversies surrounding status quo conditions and thus take a fresh look at a policy problem from a different, perhaps more neutral perspective (*ibid.*).

Nevertheless, the use of hypothetical baselines also poses some risks as highlighted by Whittington & Adamowicz (2010). First, SQ bias is not captured. When a standard SQ is not easy to define and use in the CEs (as is the case with the present study), it is hard to evaluate the SQ bias. Second, hypothetical baselines may be more cognitively challenging. Since these baselines refer to a future point in time where there will be no services without payment, they may be cognitively more demanding compared to a standard SQ (which refers to an actual current situation). The challenge becomes especially severe when the respondents have no or a few years of education and experience with CEs. Third, the introduction of hypothetical baseline conditions may make the results of the valuation study less policy relevant. This may be due to the purely hypothetical nature of the trade-offs that the respondents make against the hypothetical baseline. Fourth, there may be ethical implications of hypothetical baselines. Considering the pros and cons of both the status quo and hypothetical baselines, the latter is used in this CEs.³⁴ In the next section, one of the daunting problems in CEs, namely hypothetical bias, is discussed together with some approaches to deal with it.

3.2.5 Hypothetical Bias and Strategies to Circumvent Hypothetical Bias

Although there are variations in terms of magnitude, depending on hypothetical or real environment, the existence of hypothetical bias – the gap between people’s commitments under real and hypothetical settings – is a widely acknowledged problem in stated preference studies

³⁴ From this point forward, therefore, the term ‘hypothetical baseline’ is used.

such as the CVM and CEs (List & Gallet, 2001; Murphy et al., 2005; Carlsson, 2010; Stachtiaris et al., 2011; Bosworth & Taylor, 2012; Jacquemet et al., 2013; Moser et al., 2014; de-Magistris & Pascucci, 2014). The primary reasons for this appear to be the lack of binding budget constraint³⁵ in hypothetical SP investigations (Jacquemet et al., 2013) and the hypothetical questions asked in reference to a possible future change in the quality/quantity of goods and services. Hence, it is a common observation that people tend to overstate, for example, WTP values in hypothetical experiments compared to real-life actions. Under such circumstances, there is a need to address the challenges posed by hypothetical bias in revealing truthful actions of people when confronted with hypothetical scenarios. In the present CEs to estimate the WTP for AASs, it seems highly important to address this issue given that a hypothetical baseline (as status quo) is used – against which respondents have to compare the other alternatives.

The literature on methods to deal with hypothetical bias discusses a range of *ex ante* and *ex post* techniques, including cheap talk (Cummings & Taylor, 1999), solemn oath (Carlsson et al., 2013; Jacquemet et al., 2013; de-Magistris & Pascucci, 2014), honesty priming (de-Magistris et al., 2013), cognitive dissonance (Alfnes et al., 2010), and calibration (Lusk & Schroeder, 2004; Whitehead & Cherry, 2007). From these techniques, the most widely used cheap talk script and the recently developed and yet increasingly useful solemn oath approach to circumvent hypothetical bias are discussed here. Since the cheap talk approach is widely explored in various contexts, in the present study the effectiveness of the solemn oath in revealing truthful responses from the respondents is tested. In doing so, whether the use of solemn oath translates to lower WTP estimates compared to not using it is tested. In implementing this, each respondent is also

³⁵ If the respondents consider that it is an experiment after all and that they will not pay any amount of money for the good or service (as a consequence of participating in the experiment), then, they can act reluctantly or sometimes irresponsibly. This may be reflected by the respondents indicating high premium prices that do not take into account their actual budget/available money.

explicitly reminded to consider his/her budget constraint throughout the CEs, since not considering this may result in higher hypothetical WTP estimates which may not reflect the true values.

i. The Cheap Talk Script

The seminal work of Cummings and Taylor (1999) introduced the concept of cheap talk approach as a mechanism to reduce and/or mitigate hypothetical bias in a referendum contingent valuation study and provided evidence of cheap talk script in reducing stated WTP and effectively eliminating hypothetical bias. Since then, it has gained wider acceptance and popularity among many CV researchers and is also being utilized in a number of choice experiments, although its application in the latter is limited compared to the former. The basic notion of cheap talk approach is the presentation of a script that describes hypothetical bias, why it occurs, and its consequences on the quality of data generated and preference estimated. Once a respondent is informed about such issues, the expectation is that he/she will behave as realistically/rationally as possible throughout the hypothetical valuation situation.

Two types of cheap talk versions can be found in the literature: 'hard' and 'soft' cheap talk. The distinguishing features of these two, as discussed by Carson & Groves (2011) are as follows. Whereas the hard cheap talk presents a respondent a statement indicating that some respondents do not tell the truth (i.e., they lie) in a survey, the soft cheap talk version does not contain such a statement. Instead, it informs the respondent the existence of deviation between survey responses and actual individual actions, which is mainly due to lack of careful budget consideration. Hence, the soft cheap talk requires each respondent to consider actual income constraint and payment obligations, and recognize this commitment during the survey. Although the cheap talk

script is not used in the present study, the issue of income constraint (from the soft cheap talk version) is borrowed and incorporated in the study design, in order to remind the respondents about this aspect.

Regarding the effectiveness of applying the cheap talk approach to various circumstances, the empirical literature indicates the existence of a mixed success. For example, Moser et al. (2014) document that hypothetical bias is significantly removed in only two out of five hypothetical apple attributes, emphasizing the existence of mixed effectiveness. On the other hand, whereas some scholars show that cheap talk does not reduce WTP effectively (Nayga et al., 2006; Blumenschein et al., 2008; Ami et al., 2011), others find that it actually increases WTP (Aadland & Caplan, 2006; Carlsson et al., 2011). Still, some demonstrate that cheap talk script affects the preferences of respondents who are less familiar with the good/service (List, 2001; Lusk, 2003; Aadland & Caplan, 2006; Ami et al., 2011; Tonsor & Shupp, 2011). In general, although not effective in all settings, the use of a cheap talk script can improve preference elicitation under certain circumstances. For example, long and detailed versions of the script are preferred to shorter ones (Cummings & Taylor, 1999; Aadland & Caplan, 2006; Jacquemet et al., 2011; Loomis, 2014).

ii. The Solemn Oath Script

A recent *ex ante* approach to deal with the problem of hypothetical bias in stated preference research is the use of an oath script. Mimicking a procedure quite often used in a courtroom, this approach requests respondents to swear on their honor to provide truthful answers before they embark on the experiment (Jacquemet et al., 2013; Carlsson et al., 2013; de-Magistris and Pascucci, 2014). The rationale behind using this procedure is substantiated by the social

psychology literature which provides an explanation based on the commitment theory³⁶, which postulates that making a strong promise is more likely to induce truth-telling.

The pioneering research in using an oath script to reduce hypothetical bias – and thereby improve preference elicitation – in stated preference studies can be traced to the work of Jacquemet et al. (2009), which was first published as *CES Working Paper* entitled “*Preference Elicitation under Oath*”.³⁷ In this seminal contribution, the authors employ a solemn oath as an approach to elicit truthful response in an incentive compatible second-price auction. As a mechanism to commit respondents to provide honest answers, the researchers ask them to sign on a paper containing an oath script before proceeding to the bidding experiment. The grand conclusion from their study is that taking an oath significantly influences the commitment required to bid sincerely (Jacquemet et al., 2013).

According to Jacquemet et al. (2009) and Carlsson et al. (2013), taking an oath may induce more truthful responses due to its ability to bind the commitment of the respondent to his/her responses/actions. More importantly, Jacquemet et al. (2013) emphasize the importance of having respondents sign on an oath script than verbally agree to provide honest answers. This is because, the authors indicate, the former appears to strengthen the commitment needed to provide candid answers more than the latter. However, since requesting respondents in a survey to take an oath to provide truthful answers can be a strange phenomenon in many

³⁶ Summarizing the crux of commitment theory, Jacquemet et al. (2013) identify three aspects which result in strong commitment to link intentions with actions: a) people should be free to take an oath (i.e., no coercion), b) the oath has to be expressed publicly, and c) the oath should have consequences to the person taking it.

³⁷ This work is later published as an article in the *Journal of Environmental Economics and Management* (Jacquemet et al., 2013).

circumstances,³⁸ especially in developing countries, it may make respondents uncomfortable or even skeptic. Consequently, requesting subjects to sign on an oath script may be neither plausible nor practicable. Nevertheless, the dearth of research in this area puts one in an inevitable position of having no definitive conclusion about not only the effect of an oath script in preference elicitation, but also the procedures to be followed in such exercises.

Following the footsteps of Jacquemet and his colleagues, a few researchers have explored the effect of an oath in eliciting truthful answers in stated preference research (examples include, Jacquemet et al. (2010), de Magistris & Pascucci (2012), and Carlsson et al. (2013)). In their CV survey regarding climate change (more precisely, reductions in greenhouse gas emissions) in Sweden and China, Carlsson et al. (2013) investigate the effect of an oath script, unlike Jacquemet et al. (2009), outside laboratory setting. Three noteworthy conclusions can be drawn from their investigation: a) respondents in both countries who took an oath are less inclined to state a zero WTP or an extremely high WTP value; b) the oath script has varying effects on respondents of different sociodemographic characteristics; and c) the oath seems to have reduced the variance of the WTP.

Another notable feature of this study is the way the researchers introduced the oath script in a neutral way using the word “promise” instead of “swear.” In the oath design, they asked respondents: “Do you feel that you can promise us to answer as truthfully as possible the questions that follow?” The alternatives were 1) “Yes, I promise to answer the questions in the survey as truthfully as possible,” or 2) “No, I cannot promise this” (Carlsson et al., 2013, p. 108). The researchers justify the use of this neutral approach to introducing the oath script because “...

³⁸ For example, in some cultures and religious beliefs, swearing and taking an oath (except for religious reasons) may be viewed as strange. It is customary in many places that people swear to tell truthful answers in a court. However, it is very uncommon to ask someone to take an oath to give truthful answers in a survey.

asking ... to swear to tell the truth is not customary in Swedish or Chinese court systems” (Carlsson et al., 2013, p. 108). Introducing an oath script in such manner seems to reduce the unease created on respondents by asking them to swear or sign on the script.

Jacquemet et al. (2010) expand the use of an oath script to investigate referendum voting for a public good – university-based wind power research. Similar to the approach used in their earlier work, they ask subjects in the oath treatment group to freely sign a form by which they commit themselves to tell the truth. Comparing votes in a real and hypothetical setting with and without an oath script, the researchers find that using an oath script eliminated hypothetical bias. This result is consistent with their previous finding wherein employing an oath script helped people bid more sincerely (Jacquemet et al., 2009). Nonetheless, the tendency to confine the use of this device of truthful response elicitation to lab-based studies leaves the room open for future studies exploring its effectiveness outside laboratory condition.

The studies reviewed above document the positive role played by an oath script in reducing and/or eliminating hypothetical bias in stated preference research. However, since the literature on the effect of solemn oath on preference elicitation is very shallow, making a definitive conclusion about its effect in various stated preference study contexts appears farfetched. For example, there are hardly any examples of an oath that works contrary to the studies presented above. This is mainly due to the absence of stated preference studies which employ solemn oath in a variety of environments. The only contrary finding, to the best of my knowledge, in this regard is the work of de-Magistris and Pascucci (2012), which was presented as a poster at the Agricultural and Applied Economics Association’s 2012 AAEA Annual Meeting, Seattle, Washington, August 12-14, 2012. In this poster, the authors demonstrate the effect of an oath script in hypothetical CEs. In particular, they assess, based on a sample of 106 respondents

randomly selected from a list of people responsible for purchasing food for their household, the preference of consumers for millennium bugs. Following Jacquemet et al. (2009, 2013) procedure for introducing the oath script and comparing three treatment groups (i.e., hypothetical CEs without any cognitive task group, short cheap talk group, and oath taking group), they found that solemn oath does not produce any effect on consumers' WTP.

In summary, one can draw several implications from the studies reviewed above. First, the context of the study (i.e., laboratory versus field) may be relevant for the particular results observed. Second, the type of experiment may also be important (for example, bidding, referenda, CV, or CEs). Third, the procedure in which the oath script is introduced to respondents might also matter. Evidently, the outcome of asking respondents to sign on an oath paper appears to be different from having them promise to tell the truth. Finally, sociodemographic factors also need to be taken seriously in such exercises.

Realizing these and other issues related to the implementation of an oath script to reduce and/or avoid hypothetical bias in stated preference research, some researchers underscore future investigations of the effectiveness of the oath script in a wide range of experimental studies, particularly discrete choice experiments in the field (de-Magistris and Pascucci, 2012; Jacquemet et al., 2013). The present CEs in AASs is one such study aimed at exploring the impact of an oath script introduced by combining the procedures set out by Jacquemet et al. (2009) and Carlsson et al. (2013). More specifically, the research employs a two-step procedure of introducing the oath script to the solemn oath taking group. First, it asks respondents the question: *Would you respond to all the questions in this experiment by telling us the truth, the whole truth, and nothing but the truth?* Second, all respondents who answer 'yes' to this question

will be requested to sign on the standard oath script (adapted from Jacquemet et al., 2013, p. 115). More on this is found in methodology section.

3.3 Research Methodology

This section presents a detailed account of the research methodology followed in this study. The section begins by presenting the sampling procedure followed to select study area and household heads. Following this, the survey design and methods are discussed, including the processes involved in the selection and refinement of attributes and their levels, scenario description, and experimental design to generate the choice sets. Finally, the empirical strategy employed in the estimation of the choice experiment data is discussed.

3.3.1 Sampling Procedures

For the purpose of this study, a multi-stage sampling procedure was employed to select study district, *Kebeles*/peasant associations (PAs)³⁹, and household heads. From the 18 districts in East Hararghe Zone of Oromia Region, Haramaya district was purposively selected based on representativeness to the major farming systems and agro-ecological zones, proximity to the collaborating institutions (i.e., Haramaya University and East Hararghe Zone Bureau of Agriculture and Rural Development), availability of established and functional FTCs, and personal experience in the area. In this district, there are 33 PAs, 31 of which have established farmers' training centers (FTCs). From these, 14 PAs have operational FTCs that provide some sort of agricultural advisory services (AASs). Relevant to the current study, however, only ten of these FTCs were actually providing full AASs including modular training (the remaining four provide demonstration services only). From the ten PAs with fully functional FTCs, two PAs – *Ifa Oromia* and *Biftu Geda* – were purposively selected based on representativeness to the rest of the PAs.

³⁹ *Kebele*/peasant association is the lowest administrative division in urban/rural areas of Ethiopia.

Regarding the selection of sample household heads, on the basis of the Haramaya District Bureau of Agriculture and Rural Development (BoARD) office, there are a total of 13, 916 households in the district, out of which 1,059 households reside in the two selected PAs. A list of households who participated in and completed modular training in 2008/09 was obtained from the district BoARD office. Accordingly, there were 315 households reportedly trained in the sampled PAs. From these, a total of 120 households were randomly sampled to be included in the study – 60 households in the oath treatment (OT) and 60 in the control treatment (CT). Later in the analysis, the 60 households in the OT are further differentiated into those who signed on the solemn oath form (OT_{signed}) and those who orally agreed to provide honest/truthful answers but did not sign on the form (OT_{not signed}).

3.3.2 Survey Design and Methods

Several stages are required to be completed in designing a survey instrument for CEs. The most common components of CEs survey design are: a) identifying relevant attributes, b) assigning levels to the attributes, c) describing scenarios, d) choosing experimental design and constructing choice sets, and e) gathering data and estimating preferences (Louviere et al., 2000; Hensher et al., 2005; Scarpa et al., 2011). In what follows, each of these components is discussed in relation to AASs.

i. Identifying Relevant Attributes

Identifying the attributes to be used in CEs is an iterative process which requires an in-depth knowledge about the good/service, users/clients, service providers, and other socio-economic and institutional dynamics of the study context in general. As a critical stage in the design process, a rigorous development of attributes guarantees immense validity to the study

(Mangham et al., 2009) by plainly communicating available alternatives to the decision makers (i.e., the respondents) and thereby enabling the researcher to draw sound conclusions which could be informative in policy-making (Coast et al., 2012). Nonetheless, some researchers in the field point out that the process of attribute development is viewed either as less important than the other components of the design – hence is not conducted meticulously – or poorly reported in published works (*ibid.*).

Mangham et al. (2009) suggest reviewing published and unpublished literature (including reports and policy documents) as a starting point or springboard in the search for relevant attributes. They hold the view that attributes identified in this initial step should be refined in a follow up focus group discussions (FGDs) and semi-structured interviews. It is important, the authors note, however, that the attributes identified through review of the literature have to be substantiated with adequate primary data and be adapted to the specific study context. This is one of the reasons why conducting qualitative/exploratory work is highly encouraged by prominent scholars in the field of SP research (Louviere et al., 2000). Advancing the issues of attribute selection, some researchers call for emphasis on other methods to be employed in the development of attributes for empirical investigation. Reviewing CEs mainly in the health sector, for example, Coast et al. (2012) present additional methods (other than literature review and FGDs/interviews already mentioned above), such as recommendations by professionals, expert reviews, and policy issues/questions.

Caution must be exercised in the choice of relevant attributes. Despite the absence of a ‘gold standard’ for attribute identification and development (Louviere et al., 2000), and in spite of the widely recognized difficulty to come up with an exhaustive list of attributes, it is important that the most important ones be included. Some practical suggestions in this regard include:

unambiguity, exhaustiveness, manipulability, cognitive difficulty, inter-attribute correlation, and policy relevance.

Attributes selected need to be unambiguous in that attribute definitions should be clearly articulated in order to capture and reflect the study setting (Mangham et al., 2009). As much as possible, the selected attributes should include all those that might be important for a respondent in arriving at a decision (Coast et al., 2012). Attributes should also be experimentally manipulable (*ibid.*) in that by varying the levels the attributes take, researchers should be able to estimate different results for various scenarios. Although there is no limit to the number of attributes to be included in an experiment, the greater the number of attributes, the greater the cognitive difficulty of completing the experiment (Mangham et al., 2009). With large number of attributes comes also what is known as inter-attribute correlation (Hensher et al., 2005) – the causal relationships and interconnections among attributes (Blamey et al., 2002). Hence, avoiding conceptual overlap among attributes facilitates the estimation of the main effect of a single attribute on the dependent variable (Mangham et al., 2009). Finally, policy-oriented research should include attributes which could produce policy-relevant information (Blamey et al., 2000).

On the basis of the above considerations and aim of establishing FTCs in the study area, four attributes were selected to be the most relevant descriptors of improved FTC-based AASs in the study context. These are: advice/visit, training, demonstration, and cost. Table 3.1 presents the attributes and descriptions of each of them.

ii. Assigning Attribute Levels

The identification of relevant attributes for CEs is followed by the assignment of attribute levels. Probably, the most important question to ask at this stage is how many levels to assign for each attribute (Hensher et al., 2005). There are a number of issues to be considered in having many as opposed to a few levels for each attribute. According to Mangham et al. (2009), for example, the levels chosen should be realistic, cognitively meaningful, and acceptable to the respondents. Assigning appropriate attribute levels also has implications for the types of estimates that can possibly be derived from a study. For example, a two level attribute only allows for the estimation of linear effects. If one is interested in estimating non-linear effects, then clearly one needs more than two levels for the attributes in question (Hensher et al., 2005).

In relation to easing the task of CEs for respondents in a developing country context, using fewer levels appears to be useful. This is evident looking into the 15 case studies (conducted in 12 countries) included in the book “*Choice Experiments in Developing Countries*” edited by Jeff Bennett and Ekin Birol. The book also contains enormous applications of different methods used to assign and refine attribute levels in CEs in developing countries. The most common ones are: reviewing the grey literature, consulting stakeholders, conducting FGDs, and conducting interviews (for example, Beharry-Borg & Scarpa (2010) in Trinidad and Tobago; Selassie & Kountouris (2010) in Ethiopia; Das et al. (2010) in India; and Villalobos & Huenchuleo (2010) in Chile). Details of the attribute levels used in the present study are given in Table 3.1.

In order to estimate the WTPs for the attributes describing improved AASs as realistically as possible, it is not enough to identify the attributes and assign levels to them. There is a need to choose and include a credible and realistic payment vehicle (Bennett & Birol, 2010). However,

the choice of the payment vehicle is context-specific. For example, payment vehicles which are commonly used in developed country contexts, such as utility bills, taxes, and voluntary donations, may be of little use in the context of developing countries. According to Bennett & Birol (2010), the reasons could be manifold. First, utilities such as electricity, gas or water may not be provided in a regular basis in developing countries. Hence, the regular collection of bills may not be adequately done. Second, using taxes as a means to collect money may result in protest answers due to lack of trust in the credibility of local authorities. Finally, voluntary donations as payment vehicles may suffer from strategic responses. Although difficult, some type of payment vehicle has to be found. In their CEs exercise to analyze farmers' choice between public goods and agricultural extension packages in Ethiopia, for example, Mekonnen et al. (2010) used an amount of money to be paid back at harvest time as their payment vehicle. Likewise, Scarpa et al. (2011) use an access fee (to be paid per person per visit) as a payment vehicle in their study dealing with the alpine grazing commons.

There are valuable suggestions put forward by researchers in the field concerning the choice of payment vehicles in CEs in the context of developing countries. The role of FGDs in the identification of appropriate payment vehicles is highly emphasized by Bennett & Birol (2010). Other practitioners suggest that immense care be taken in terms of the coverage, acceptability, and feasibility of the payment vehicle to be chosen. In the present CEs, the suggestions of these scholars are followed in order to select a realistic payment vehicle in FGDs. Since FGDs are also important in the other design components, a FGD was conducted in May 2013 with representatives of smallholder farmers in the study area. In what follows, the main issues discussed in the FGD are presented. Detailed aspects of the FGD are given in Appendix 1.

The main elements of the FGDs are: choice of attributes, assignment of attribute levels, selection of relevant payment vehicle, and scenario description. The discussion on attributes considered the main functions of FTC-based AASs (advice, training, and demonstration). Regarding attribute levels, the result of the discussion indicates that the farmers want a maximum of three advice/visits per year, three modular training (each with one month duration) per year, and three demonstrations per year. In relation to payment vehicle, the possibility to use either direct cash payments or registration fees was discussed. The unanimous decision of the discussants is to use direct cash payment to FTC administration on a yearly basis as a payment vehicle. In order to identify the range for the cost attribute, the discussants evaluated different cards depicting the highest, intermediate, and lowest level combinations of attributes (see Appendix 3 for the cards). Then, they indicated their individual bids for the different combinations, from which the upper- and lower-bound for this attribute were found (column (1), Table 3.1). Finally, the understandability, cognitive effort needed, and best-worst structure of the scenario description prepared for the OT and CT groups is discussed in the FGD. More on this is also found in the next section.

Table 3.1 Attributes and Attribute Levels for the CEs in FTC-based Provision of AASs

Attribute	Description	Attribute levels		
		(1)	(2)	(3)
Advice/visit	Number of advice/visit per year a farmer receives on crop, livestock, and NRM from Development Agents based at FTCs.	3	3	3
		2	2	2
		<i>0</i>	<i>0</i>	<i>0</i>
Training	Number of training per year – refers to participating in modular training (of 1 month duration) organized at FTCs to impart the necessary knowledge and skills in crop, livestock, and NRM.	3	3	3
		2	2	2
		<i>0</i>	<i>0</i>	1 <i>0</i>
Demonstration	It refers to the number of method or result demonstrations (per year) organized by FTCs to show the performance of improved technologies and practices in crop, livestock, and NRM.	3	3	3
		2	2	2
		<i>0</i>	<i>0</i>	<i>0</i>
Cost ^a	It refers to the amount of money (in ETB) – to be paid directly to the FTC administration on a yearly basis at the end of the farming season – associated with the various combinations of attributes comprising AASs.	150	250	300
		100	200	250
		50	150	200
		<i>0</i>	100	150
			<i>0</i>	100 <i>0</i>

^a On the day of the FGD (i.e., 31 May 2013), the official exchange rate was 1 Euro = 24.0897 ETB.

Note: Attribute levels in italics (i.e., the zero levels) represent the hypothetical baseline situation. (1) Levels discussed and agreed upon in FGD. (2) Levels used in Optimal Orthogonal Design to collect priors. (3) Levels used in the final experimental design (i.e., Bayesian where all the parameters are normally distributed).

iii. Describing Scenarios and Choice Cards

Scenario description in CEs entails the presentation of the CEs context under which trade-offs have to be made among the relevant attributes describing the good or service in question. It presents the baseline (or status quo) situation and the (hypothetical) improvements or management actions that will be taken in the foreseeable future. In a nutshell, scenario description introduces the good/service, the challenges faced in the long-term provision of the good/service, the importance of raising funds to solve some of the challenges faced, and the circumstances under which respondents have to make trade-offs between the attributes comprising the good/service.

The scenario description in the current study proceeds as follows. First, the FTC-based AASs are introduced, together with the context under which the CEs has to be done. Moreover, each of the attributes and their levels were described. Second, the scenario description is presented to the two treatment groups of the study, namely the oath treatment (OT) and the control treatment (CT). There are three common elements of the scenario descriptions for all the respondents in the treatment groups. First, they all are asked to imagine and act as if they have to pay the money right away (i.e., at the time of the interview). Second, they all are required to take into account the cash income constraint. Finally, they all are reassured that the data will be used only for the purpose of this study. The difference in the descriptions presented to the two groups relies on the use of a solemn oath script in the OT. Detailed aspects of the scenario description for the OT and CT groups are provided in Appendix 2.




After this, the profiles created by combining the levels of the four attributes were presented to the respondents as choice sets. Details on this are provided in the next section (under experimental design). Each choice set contained two hypothetical options – option A and option B – and an additional option describing the hypothetical baseline alternative. Table 3.2 presents an example choice set. After reminding the respondents that the research is interested only in their opinions and that there are no right or wrong answers⁴⁰, the respondents were then asked to make two choices for each choice set: first, the best-preferred alternative from the three alternatives, and then the worst-preferred alternative from the remaining two alternatives.

Finally, recognizing the fact that visual aids facilitate respondents' understanding of the attributes and their respective levels (Bennett & Birol, 2010) by enabling the enumerators to

⁴⁰ This appears to be a good practice in CEs related to developing countries as reported by many researchers. See the studies reported in Bennett & Birol (2010) for the details.

explain them plainly and thereby simplifying the choice task, pictorial representations of the attributes and their levels are prepared. The design of the visual aids also took into consideration the limited educational levels of the respondents and their lack of experience in CEs surveys. After testing the visuals in a FGD, they are incorporated in the final CEs survey. Examples of the cards used in the survey are provided in Appendix 3.

Table 3.2 An Example of a Choice Card Used in the Collection of Prior Data

		Alternative 1	Alternative 2	Alternative 3
a. number of visits to FTCs/year (advice)		3	0	0
b. number of training (max. 1 month duration) per year		3	0	0
c. number of demonstrations/year		0	2	0
d. cost (in ETB)/year		150	200	0
Q1. Which of the three is your most preferred option? (mark ✓)				
Q2. Which of the remaining two is your least preferred option? (mark ✓)				

iv. Experimental Design and Designing Choice Sets

On the basis of the attribute levels obtained from the FGDs, and a slight change on the levels for the cost attribute (Table 3.1, column 2), an Optimal Orthogonal Design (Street & Burgess, 2007; Rose & Bliemer, 2009) was launched to generate 36 choice sets to collect prior estimates for the

final experimental design.⁴¹ The design was 98.13% efficient for estimating a main effects only indirect utility function and conditional logit model. The change in the levels for the cost attribute was necessary in order to capture the highest bids, avoid ‘silly cards’, and produce better experimental design. The 36 choice sets, blocked into two, were introduced to a sample of 30 household heads, excluding the 12 households who took part in the FGDs. This resulted in 18 choice sets per respondent per block, which translates to 15 complete questionnaires. The pilot survey took place as follows. After introducing the aim of the survey, some warm up questions regarding the three components of AASs were asked. Following this, the scenario description was introduced. After this, the choice task began in which each respondent had to indicate his/her most-preferred alternative first followed by the least-preferred one. Finally, some questions pertaining to socio-demographic characteristics were asked.

The final experimental design for the main survey was based on the coefficients estimated from the pilot survey. Moreover, considering the results of the pilot, one more level was added to the cost attribute and one more to the attribute ‘training’ (Table 3.1, column 3). A Bayesian D-efficient design (Sándor & Wedel, 2001) was used, where all the parameters are normally distributed (Table 3.3) to generate 12 choice sets per respondent. The benefit of this design is that it optimizes over a range of possible parameter values, under the assumption that the prior parameter values are only approximately known (Kessels et al., 2006; Rose & Bliemer, 2008; Bliemer et al., 2009). Bayesian designs are hence more robust to misspecification reducing the dependence of the design on the unknown parameter β .

⁴¹ The experimental designs (both for the collection of prior estimates as well as final survey) were generated using the NGENE software.

Table 3.3 Experimental Design

	Order in ranking-task sequence											
	1	2	3	4	5	6	7	8	9	10	11	12
Alternative 1												
<i>ADV</i>	2	2	3	3	3	0	0	2	2	0	0	3
<i>TRAIN</i>	2	2	3	1	0	0	3	2	1	1	3	0
<i>DEMO</i>	3	0	0	0	2	3	2	0	3	3	2	2
<i>COST</i>	150	200	100	250	300	150	200	250	150	300	100	100
Alternative 2												
<i>ADV</i>	2	2	0	0	0	3	3	2	2	3	3	0
<i>TRAIN</i>	1	1	0	2	3	3	0	1	2	2	0	3
<i>DEMO</i>	0	3	3	3	2	0	2	3	0	0	2	2
<i>COST</i>	200	150	250	150	100	200	150	100	250	100	300	300

Note: *COST* = cost (in Ethiopian birr)/year; *ADV* = number of FTC-based advice/visit per year; *TRAIN* = number of FTC-based modular training/year; *DEMO* = number of FTC-based demonstration/year. In addition to the two alternatives, there is Alternative 3, which is the hypothetical baseline situation that takes the value of zero for all the attributes.

v. Survey Administration

The choice sets were presented as part of the main survey which was administered from May to October 2013. The choice sets were introduced to a sample of 120 household heads – excluding those in the pilot survey and FGD participants. The identified households were divided into two treatment groups (i.e., OT and CT). The questionnaire, containing the choice cards, was administered by two trained enumerators supervised by trained supervisors from Haramaya University. Each enumerator assisted the respondents in the choice process and recorded the choices of the respondents on the questionnaire. Since most of the respondents were uneducated and with no previous experience in CEs, using other methods of gathering choice data (such as, computer-assisted self-administered system as in Moser et al. (2014)) was unfeasible. Moreover, since touch-screen laptops equipped with a touch-screen pen were not available (due to financial constraints), the survey was carried out in a paper-and-pencil style.

3.3.3 Empirical Strategy and Estimation

Similar to the traditional discrete choice experiment data, best-worst choice experiment data can be used to estimate an indirect utility function following the random utility theory (Train, 2009; Lancsar et al., 2013). Accordingly, the utility for an individual farmer n ($n = 1, \dots, 120$) facing a choice among j alternatives ($j = 3$ including the hypothetical baseline option) in choice set t ($t = 1, \dots, 12$) can be decomposed into a systematic (explainable) component, V_{ntj} , and a random/stochastic (error) component, ε_{ntj} , as

$$(1) \quad U_{ntj} = V_{ntj} + \varepsilon_{ntj}.$$

Assuming a linearly additive indirect utility function and incorporating the vector of the services comprising AASs (X_{ntj}), the systematic component of the utility function may be given as

$$(2) \quad V_{ntj} = \beta_j X'_{ntj}$$

where β_j are the vectors of coefficients to be estimated (through the maximum likelihood estimation (MLE) procedures).

According to Hensher et al. (2005) and Train (2009), the probability that alternative j will be chosen by farmer n in choice set/situation t can be expressed as the probability that the utility of alternative j exceeds the utility of all other alternatives q :

$$(3) \quad \begin{aligned} \text{Prob}_{ntj} &= \text{Prob} [U_{ntj} > U_{ntq}]; \forall q \neq j \\ &= \text{Prob} [V_{ntj} + \varepsilon_{ntj} > V_{ntq} + \varepsilon_{ntq}]; \forall q \neq j. \end{aligned}$$

Substituting $V_{nt} = \beta X'_{nt}$ and performing some algebraic manipulations gives

$$(4) \quad \text{Prob}_{ntj} = \text{Prob} [\beta_j X'_{ntj} - \beta_q X'_{ntq} > \varepsilon_{ntq} - \varepsilon_{ntj}]; \forall q \neq j.$$

This means, the probability that a farmer n chooses alternative j in choice set t is simply the probability that the difference in observed utility ($\beta_j X'_{ntj} - \beta_q X'_{ntq}$) is greater than the difference in unobserved utility ($\varepsilon_{ntq} - \varepsilon_{ntj}$).

When analyzing most discrete choice experiment data, the starting point is to assume that the error terms (ε_{nt}) are *independently and identically distributed (IID)* as extreme value type I across the respondents. This produces the McFadden's conditional logit (also known as multinomial logit (MNL)) model, with the associated *independence from irrelevant alternatives (IIA)* property. Consequently, the probability that a farmer n chooses alternative j from all available alternatives q in choice set t can be expressed by the logit formula (Train, 2009)

$$(5) \quad \text{Prob}_{ntj} = \frac{e^{\lambda V_{ntj}}}{\sum_{j=1}^q e^{\lambda V_{ntq}}}$$

where V_{nt} is defined in equation (2) above and λ is the scale factor – that is inversely proportional to the variance of the error term – given by $\lambda = \frac{\pi}{\sigma\sqrt{6}}$. When the ε_{nt} are IID, the variance of the random error, σ , is constant across individuals, and so is λ . Since λ cannot be estimated separately from the parameters of the explanatory variables of V_{nt} (Swait & Louviere, 1993), it is often normalized to one. For the MNL model, $\lambda = \pi^2/6$.

The best-worst choice experiment data in the present study can be modeled with the conditional logit by only using, for example, the farmers' best-preferred choice in each choice situation. Nevertheless, since one of the advantages of best-worst choice experiments is generating additional preference information (for example, second-best choice), one can use this extra

information to estimate discrete choice models. This can be done by noting that the best-preferred and worst-preferred choice questions produce an implied rank order over the alternatives. Then, this can be modeled using rank ordered logit (ROL) or otherwise known as “exploded logit” (Scarpa et al., 2011; Lancsar et al., 2013). ROL models the probability of a particular ranking of alternatives as the product of MNL models for the choice of best-preferred options (Chapman & Staelin, 1982). For example, the ranking of the three choices ($y_{1b} > y_r > y_{1w}$) from the generic alternatives (where y_{1b} is the first best choice, y_r is the residual, and y_{1w} is the first worst choice) is modeled as the product of the (MNL) probability of choosing y_{1b} as best from the set (y_{1b}, y_r, y_{1w}) multiplied by the probability of choosing y_r as best from the remaining alternatives (y_r, y_{1w})

$$(6) \quad \text{Prob}(y_{1b} > y_r > y_{1w}) = \frac{e^{V_{1b}}}{\sum_{j \in \{1b, r, 1w\}} e^{V_j}} * \frac{e^{V_r}}{\sum_{j \in \{r, 1w\}} e^{V_j}}.$$

ROL is called “exploded logit” because drawing on the IIA property of MNL models, it can be estimated by exploding the data from each choice set into statistically independent choice subsets (Lancsar et al., 2013). For example, if there are six alternatives in a given choice set, a respondent first chooses the best from the six alternatives, followed by the best from the remaining five alternatives, and so on until the best from the remaining two alternatives is chosen. Having gathered choice data in this manner, one can use the traditional conditional logit to estimate the ROL parameters from this expanded choice data.

Nevertheless, although the MNL model is the most commonly used discrete choice model, it has limitations with regard to allowing for random taste variations, unrestricted substitution patterns, and correlation in unobserved factors over time/repeated choice (see Train (2009) and Hensher *et al.* (2005) for a detailed description on these issues). A direct way to model preference

heterogeneity is to account for observed heterogeneity by including respondents' characteristics (Z_n) in the choice model. This can be done through interactions with X_{nt} or the alternative-specific constants (Hensher et al., 2005). Unobserved preference heterogeneity often is modeled using a mixed logit (ML) model – also known as random parameters logit (RPL) model (Revelt & Train, 1998; Train, 2009).

Hence, to analyze the best-worst choice experiment data in this study, the standard multinomial logit (MNL) – after ‘exploding’ the data – and the random parameters logit (RPL) models are estimated. Regarding the RPL model, following Revelt and Train (1998), an RPL model for a panel data is estimated. To illustrate this, let us consider a farmer n facing a sequence of t ($t = 12$) choice tasks. If the alternative that farmer n chooses in period t is denoted as $i(n, t)$, then the unconditional probability for the sequence of choices is given as (7)

$$\int_{\beta_n} \prod_{t=1}^{12} \frac{\exp(\beta'_n x_{nit})}{\sum_{j=1}^3 \exp(\beta'_n x_{njt})} f(\beta_n) d\beta_n.$$

While estimating the RPL model, first all the non-cost coefficients (i.e., advice, training, demonstration) are specified to be independently normally distributed. A fixed cost coefficient is a common assumption because it not only simplifies the calculation of the WTP, but also ensures that the WTP estimates are distributed the same as the corresponding non-cost parameters in the numerator (Asrat et al., 2010; Moser et al., 2014).⁴² However, there are some concerns in using a fixed cost coefficient. For example, Scarpa et al. (2008) suggest that an individual's response to costs may vary independently of observed sociodemographic characteristics.⁴³ One way out of

⁴² Since the WTP is the ratio between an attribute's coefficient and the cost coefficient. In our case, since each of the attributes in the numerator is distributed normally, this implies that the WTP is also distributed normally.

⁴³ For details on the issue of whether to use a fixed cost coefficient or not, as well as some solutions to the issue, see Scarpa et al. (2008) and Balcombe et al. (2009).

the problem of a fixed cost coefficient assumption, according to these authors, is to parameterize the model in “WTP space” rather than the usual “preference space” parameterization.

For estimating the random parameters, 150 Halton draws are used, rather than pseudo-random draws since this provides a more precise simulation for the RPL model (Hensher et al., 2005; Train, 2009). Finally, the WTPs for each attribute describing improved FTC-based AASs in the study area are calculated as the ratio between the coefficients of the random parameters and the coefficients of the cost (i.e., $WTP = -\beta_i/\beta_c$) as it is a common practice (Hensher et al., 2005).

3.4 Results and Discussion

The result and discussion section is structured as follows. First, the results pertaining to the descriptive statistics of sociodemographic characteristics of the respondents are presented. Following this, both the MNL and RPL model results are presented, but the RPL results regarding parameter estimates are discussed. Moreover, since there are observed heterogeneities around the parameter estimates, the results of interaction effects introduced into the RPL model are presented and discussed. Finally, since the very aim of the study is estimating the WTP for the services describing improved FTC-based AASs, the marginal WTP estimates and the effect of the solemn oath on reducing hypothetical WTP values are discussed.

3.4.1 Descriptive Statistics of Sociodemographic Characteristics of Respondents

The descriptive statistics of sociodemographic characteristics of the respondents – along with the tests for equality of distributions among the treatment groups – are presented in Table 3.4. The Kolmogorov-Smirnov test (K-S test) of equality of distributions reveals that there are significant differences between OT_{total} and CT as well as between OT_{signed} and CT for covariates, such as household size (measured by the number of people sharing the same house), access to daily and weekly markets (to buy production inputs and consumables as well as sell crop and livestock products), household assets (productive assets, household goods, and consumer durables), land size and land certification, value of livestock possessed, farm income per year (including crop and livestock incomes), and off-farm employment. Furthermore, the K-S test also indicates the existence of significant variations related to participation in FTCs (both in terms of years of participation and frequency of advice, training, and demonstration per year). As shown later on, some of these sociodemographic characteristics explain part of observed heterogeneity around

mean parameter estimates (for example, the effect of interacting years spent in AASs with attributes on preferences).

Table 3.4 Summary Statistics of Sociodemographic Characteristics of Respondents
(Standard Deviations in Parenthesis)

Characteristics	Treatment groups ^e			CT	K-S test (OT _{signed} vs CT)	K-S test (OT _{total} vs CT)
	OT					
	OT _{total}	OT _{signed}	OT _{not signed}			
Age (years)	39.0 (3.7)	39.5 (3.6)	38.5 (3.7)	41.4 (4.9)	0.361	0.056
Male household head ^a	98.3	99.6	100.0	100.0	1.000	1.000
Education (years)	1.8 (3.0)	1.7 (2.8)	2.0 (3.2)	1.0 (2.3)	0.764	0.665
Household size (number)	6.8 (1.1)	6.5 (1.2)	7.1 (1.0)	7.9 (1.1)	0.000	0.001
Married household head ^a	100.0	100.0	100.0	100.0	1.000	1.000
Access to basic facilities ^a						
Drinking water	85.0	79.3	90.3	100.0	0.318	0.494
Human health center	18.3	13.8	22.6	0.0	0.790	0.190
Daily market	1.7	3.5	0.0	73.3	0.000	0.000
Weakly market	8.3	10.3	6.5	100.0	0.000	0.000
Veterinary	100.0	100.0	100.0	95.0	1.000	1.000
Experience in farming (years)	23.6 (5.2)	23.7 (5.6)	23.5 (4.8)	22.4 (4.8)	0.361	0.190
Experience in AASs (years)	16.6 (3.6)	17.1 (3.8)	16.2 (2.4)	14.7 (4.4)	0.211	0.045
Experience with FTCs (years)	3.5 (0.6)	3.6 (0.6)	3.5 (0.6)	4.0 (1.0)	0.019	0.001
FTC participation (number/year)						
Advice ^f	25.5 (5.9)	25.2 (5.1)	25.8 (6.7)	7.2 (6.2)	0.000	0.000
Training (short-term & modular)	7.4 (2.4)	7.3 (1.8)	7.4 (2.8)	2.2 (1.1)	0.000	0.000
Demonstration	2.2 (0.7)	2.3 (0.5)	2.1 (0.9)	1.7 (0.8)	0.000	0.000
Evaluation of AASs (satisfaction) ^a						
<i>Advice</i>						
Very satisfied	3.3	3.5	3.2	50.0	0.000	0.000
Satisfied	80.0	79.3	80.7	50.0	0.059	0.008
Not satisfied	16.7	17.2	16.1	0.0	0.578	0.329
<i>Training</i>						
Very satisfied	76.7	62.1	90.3	47.4	0.732	0.010
Satisfied	23.3	37.9	9.7	52.6	0.727	0.010
Not satisfied	0.0	0.0	0.0	0.0	1.000	1.000
<i>Demonstration</i>						
Very satisfied	23.1	36.4	13.3	46.4	0.989	0.084
Satisfied	19.2	27.3	13.3	48.2	0.427	0.016
Not satisfied	57.7	36.4	73.3	5.4	0.075	0.000
Assets ^b	9,649.9 (8,711.0)	12,960.9 (10,451.0)	6,552.4 (5,174.9)	12,724.7 (22,514.8)	0.019	0.001
Land size (ha)	0.9 (0.4)	0.9 (0.3)	0.9 (0.4)	0.8 (0.3)	0.049	0.015
Land certified ^a	100.0	100.0	100.0	1.7	0.000	0.000
Value (Birr) of livestock possessed ^c	40,731.0 (13,115.4)	40,222.4 (12,512.5)	41,206.8 (13,845.1)	53,838.4 (17,961.9)	0.000	0.000
Farm income (Birr) ^d	18,185.2 (6,587.7)	19,652.9 (5,962.9)	16,812.2 (6,937.9)	28,013.5 (9,700.4)	0.000	0.000
Involvement in off-farm activities ^a	56.7	65.5	48.4	8.3	0.000	0.000
Have access to microfinance	11.7	6.9	16.1	8.3	1.000	1.000
Number of people in one's network	4.4 (1.6)	4.2 (1.2)	4.7 (1.9)	5.7 (1.4)	0.000	0.000
Experienced shock ^a	96.7	100.0	93.6	51.7	0.000	0.000
Sample size	60	29	31	60		

K-S test: Two-sample Kolmogorov-Smirnov test for equality of distribution functions. Exact *p-values* on the combined K-S reported.

^a Proportion/percentage of sample possessing the specific characteristic.

^b Value of assets (Birr) is estimated based on the prevailing market price for the good or materials to produce it. It comprises values of productive assets, household goods, and consumer durables bought or possessed during the last 12 months preceding the survey.

^c Computed based on the estimated value (Birr) of each livestock species had it been sold in the current market.

^d Calculated by summing up crop and livestock incomes (net of intermediate costs associated with the two). (i.e., a. Value of crop sold and total crop residues produced (in ETB) per household during the last 12 months minus cost of intermediate inputs associated with crop production, and b. income from sell of livestock products (Birr) per household net of intermediate costs associated with livestock production and management.)

^e OT = Oath Treatment; OT_{total} = total oath treatment (i.e., those who agreed to tell the truth or sign on the solemn oath form); OT_{signed} = those who agreed to tell the truth and signed on the solemn oath form; OT_{not signed} = those who agreed to tell the truth but did not sign on the solemn oath form; CT = Control/Baseline Treatment

^f This refers to the total number of advice received from the FTCs. It encompasses the number of advice farmers received by going to the FTCs as well as the number of advice that were provided to them (on their farm/home) by development agents coming from the FTCs.

3.4.2 Parameter Estimates in the Full Sample

Tables 3.5 & 3.6 report the parameter estimates for the treatment groups (OT and CT in the full sample) for the MNL and RPL models, respectively. All the models in this study are estimated in *NLOGIT 5* (Econometric Software, Inc. <http://www.limdep.com/>). The null hypothesis of preference equality across the two treatments has been tested with the likelihood ratio (LR) test, $-2(\log l_p - \sum \log l_t)$ distributed χ^2 with $K(T - 1)$ degrees of freedom, where $\log l_p$ is the log likelihood value for the pooled (p) model, $\log l_t$ are the log likelihood values for each of the treatments (t), K is the number of parameters/coefficients (5 in MNL and 9 in RPL), and T is the number of treatments, two in the case of this study (Louviere et al., 2000; Hensher et al., 2005). The hypothesis of preference equality among the two treatments is rejected for both the MNL ($\chi^2_5 = 204.77; p < 0.001$) and RPL model ($\chi^2_9 = 195.57; p < 0.001$).

The estimation of the RPL model is based on normally distributed parameters (advice, training, demonstration, cost, and hypothetical baseline) with 24 choices per respondent (best and second best from each choice card). Furthermore, 150 Halton draws are used with 100 maximum iterations for the maximum likelihood estimation (MLE) procedure. Moreover, the Scaled Multinomial Logit (SMNLOGIT) and Generalized Mixed Logit (GMXLOGIT) models are

estimated in order to test for the presence of scale effects (Swait & Louviere, 1993; Hoyos, 2010; Scarpa et al., 2011). The results indicate that there is no statistically significant difference in variance across the responses in the treatment groups (OT & CT). Moreover, no significant difference in scale is found between the respondents trained at FTCs in 2008 and 2009.⁴⁴ Similarly, using the two models, the existence of a ‘learning effect’ resulting from the order of presentation of the choice cards (Adamowicz et al., 1998; Hanley et al., 1998; Kjaer et al., 2006) is tested. The results confirm the lack of such effects. Since the coefficients from the SMNLOGIT and GMXLOGIT are broadly consistent – except for the additional scale parameter – with those in MNL and RPL, in the following section, the results from the latter models are presented and discussed.

In general, the parameter estimates are consistent in terms of sign in both models (MNL and RPL). Comparing the statistics reported in the lower part of Tables 3.5 & 3.6, the RPL model (log likelihood – 680.99 and – 835.42; AIC⁴⁵ 0.958 and 1.173; and BIC⁴⁶ 0.991 and 1.206) fits the data significantly better than the MNL (log likelihood – 690.85 and – 850.21; AIC 0.966 and 1.188; and BIC 0.985 and 1.206). Therefore, although the MNL model results are displayed ahead of the RPL outputs (for the sake of comparison and illustrating the base case estimation), the focus of the discussion is on the RPL results only.

⁴⁴ Note that the full sample comprises 60 respondents trained in 2008 and 60 respondents trained in 2009. Latter in the analysis, the respondents trained in 2009 were analyzed separately, in order to link the results of this choice experiment with the impact evaluation conducted on this sub-sample.

⁴⁵ Akaike’s Information Criterion (AIC) is given as $-2(\log l - K)$ and the normalized AIC is given as $-2(\log l - K)/n$, where $\log l$ is the log likelihood value, K is the number of parameters (model size), and n is the number of observations (sample size). In this study, the reported AIC values are the normalized ones. The decision rule is the lower the AIC, the better the model fit.

⁴⁶ Bayesian Information Criterion (BIC) is given as $-2(\log l) + \ln(n)K$ and the normalized BIC is obtained by dividing this by n , where $\log l$, n , and K are as defined for AIC. In this study, normalized BIC values are reported. The decision rule is the lower the BIC, the better the model fit.

Examining the two treatments (OT_{total} and CT) in Table 3.6, except for advice, the other attributes are of the expected sign and significant. The negative and significant coefficient for advice in OT_{total} can be related to the intensity of participation in FTC-based advice/visit. As can be seen in Table 3.4, these respondents received, on average, 25 rounds of advice/visit during the last 12 months preceding the survey. Therefore, they may be reluctant to participate in the new arrangement where they have to pay for getting advice from the FTCs (unlike those in the CT).⁴⁷ Furthermore, whereas the coefficient for the hypothetical baseline (i.e., *no payment, no AASs*) is positive but not significant at 5% in the OT_{total} , it is positive and highly significant in the CT, suggesting that many respondents in the CT have chosen it as their second best option (see Appendix 5), despite the relatively low level of dissatisfaction with the current service provisions.

⁴⁷ It may also be the case that despite the high number of advice, the quality (content), relevance, and timeliness of advice received might not have been as expected by the respondents.

**Table 3.5 MNL Model Estimation Results on the Treatment Groups in the Full Sample
(Trained in 2008 & 2009)**

Attribute parameters	OT			CT
	OT _{total}	OT _{signed}	OT _{not signed}	
Advice	- 0.591 (0.069) ***	- 0.595 (0.102) ***	- 0.590 (0.094) ***	0.335 (0.061) ***
Training	1.502 (0.091) ***	1.629 (0.136) ***	1.391 (0.123) ***	1.102 (0.063) ***
Demonstration	0.994 (0.075) ***	1.050 (0.112) ***	0.946 (0.102) ***	1.272 (0.078) ***
Hypothetical Baseline	0.789 (0.385) **	0.949 (0.563) *	0.640 (0.528)	2.692 (0.346) ***
Cost	- 0.00208 (0.00094) **	- 0.00239 (0.00138) *	- 0.00181 (0.00129)	- 0.00438 (0.00072) ***
LogL function	- 690.85	- 323.92	- 365.54	- 850.21
Parameters	5	5	5	5
Observations	1,440	696	744	1,440
AIC	0.966	0.945	0.996	1.188
BIC	0.985	0.978	1.027	1.206

Standard errors in parenthesis
 ***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Note: OT = Oath Treatment; CT = Control Treatment

**Table 3.6 RPL Model Estimation Results on the Treatment Groups in the Full Sample
(Trained in 2008 & 2009)**

Attribute parameters	OT			CT
	OT _{total}	OT _{signed}	OT _{not signed}	
Advice	- 0.656 (0.074) ***	- 0.731 (0.113) ***	- 0.610 (0.097) ***	0.329 (0.070) ***
Training	1.581 (0.100) ***	1.800 (0.156) ***	1.423 (0.133) ***	1.208 (0.080) ***
Demonstration	1.031 (0.079) ***	1.123 (0.120) ***	0.966 (0.104) ***	1.358 (0.084) ***
Hypothetical Baseline	0.684 (0.405) *	0.752 (0.617)	0.620 (0.538)	2.806 (0.363) ***
Cost	- 0.00234 (0.00098) **	- 0.00294 (0.00147) **	- 0.00189 (0.00132)	- 0.00482 (0.00076) ***
<i>Sd_ADV</i>	0.009 (0.057)	0.017 (0.098)	0.005 (0.070)	0.240 (0.048) ***
<i>Sd_TRAIN</i>	0.152 (0.089) *	0.016 (0.177)	0.240 (0.086) ***	0.271 (0.057) ***
<i>Sd_DEMO</i>	0.004 (0.050)	0.012 (0.076)	0.002 (0.069)	0.000031 (0.055)
<i>Sd_HB</i>	0.630 (0.120) ***	0.925 (0.188) ***	0.095 (0.550)	0.254 (0.179)
LogL function	- 680.99	- 310.43	- 363.57	- 835.42
Parameters	9	9	9	9
Observations	1,440	696	744	1,440
AIC	0.958	0.918	1.002	1.173
BIC	0.991	0.977	1.057	1.206
Pseudo-R ²	0.57	0.59	0.56	0.47

Standard errors in parenthesis
 ***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Note: OT = Oath Treatment; CT = Control Treatment. Sd_ADV = Standard deviation (Advice), Sd_TRAIN = Standard deviation (Training), Sd_DEMO = Standard deviation (Demonstration), Sd_HB = Standard deviation (Hypothetical Baseline).

Since the choice of the hypothetical baseline situation does not seem to arise as a consequence of the level of satisfaction with the current AASs, it requires a plausible justification. From Table 3.4, it is evident that the respondents in the CT are less better-off in terms of the extent/magnitude of advice, training, and demonstration they received compared to those in the OT_{total} . Hence, it may be the quantity of AASs received which influenced the respondents in the CT to incline more to that hypothetical baseline option. Furthermore, as Meyerhoff & Liebe (2009) noted, the choice of the status quo (hypothetical baseline in this study case) may also be related to protesting against the survey, complexity of the choice task or mistrust in the FTCs. It may also be the case that some respondents simply do not want to pay for improved AASs. In general, whereas training, demonstration, and cost are the main drivers that affect the probability of participation in paid AASs in OT_{total} , advice is an additional attribute which also influences the likelihood of participation in improved FTC-based AASs in the CT.

The within group comparison in the OT (OT_{signed} and $OT_{not\ signed}$) shows that the coefficient for the hypothetical baseline, though still positive, is no more significant. This may suggest that agreeing to tell the truth and/or signing on a solemn oath form may decrease the likelihood of respondents adhering to the hypothetical baseline compared to the CT. Moreover, in the OT, those who signed on the solemn oath form showed significant commitment in analyzing the cost attribute compared to those who did not sign.

The RPL model in Table 3.6 also reveals the existence of heterogeneity around the mean parameter estimates for advice (Sd_ADV), training (Sd_TRAIN), and hypothetical baseline

situation (Sd_{HB}) in some treatment groups/sub-groups (as indicated by the significant coefficients for the standard deviations of some of them). However, comparing OT_{total} and CT, there exists significant heterogeneity only in the CT for advice and training. The within sample comparison (OT_{signed} and $OT_{not\ signed}$) also shows that there is heterogeneity around the mean parameter estimates for training in the $OT_{not\ signed}$ group. Although the RPL model indicates the existence of heterogeneity, the sources of this heterogeneity are not clear. In order to deal with this, selected sociodemographic characteristics are interacted with the attributes that exhibited heterogeneity. The result is presented in Table 3.7 and discussed in the following section.

3.4.3 Exploring Heterogeneity around Mean Parameter Estimates

Although the RPL model shows the existence of heterogeneity, it does not indicate the source of this preference heterogeneity. In order to account for this in the full sample (Table 3.6), following the suggestions of Hensher et al. (2005) selected sociodemographic characteristics of the respondents were interacted with the attributes exhibiting significant heterogeneity around the mean parameter estimates. Applications of such a procedure in the context of developing countries can be found in Asrat et al. (2010), Kassie et al. (2010), Probst et al. (2012), and Owusu-Sekyere et al. (2014). Whereas Asrat et al. (2010) find contact with extension services, household resource endowment, and years of farming experience to be the major causes of heterogeneity in crop variety preferences, Kassie et al. (2010) document differences in age, occupation, and education to be the main factors explaining variations in preferences cattle buyers have for indigenous bull traits. Similarly, whereas Probst et al. (2012) find, among others, type of business (small business, restaurant), education, and age to explain heterogeneity in preferences for organic vegetables, Owusu-Sekyere et al. (2014) document household income,

gender, age, and education to be the causes of preference heterogeneity in beef food safety assurance labels.

Although several interaction effects were explored in this study, the results for statistically significant interactions are presented in Table 3.7. The comparison between Tables 3.6 & 3.7 for OT_{total} and CT indicates that the RPL model with interactions (log likelihood – 660.63 and – 825.74; AIC 0.936 and 1.165; BIC 0.983 and 1.213; and Pseudo- R^2 0.48 and 0.58) fits the data better than the RPL without interactions (log likelihood – 680.99 and – 835.42; AIC 0.958 and 1.173; BIC 0.991 and 1.206; and Pseudo- R^2 0.47 and 0.57). In general, years of participation in AASs partly explains the heterogeneity in the preference for advice. Similarly, years spent in AASs, farm income, off-farm employment, and networks are some of the causes for preference heterogeneity in training. Regarding the hypothetical baseline, the explanatory variables for observed preference heterogeneity include income, exposure to shocks, and education. In the following, the results pertaining to attribute-wise interaction effects are discussed.

Table 3.7 RPL Model Estimation Results with Interactions (Full Sample – Trained in 2008 & 2009)

Attribute parameters	OT _{total}	CT
<i>Advice</i>	- 0.683 (0.076) ***	0.767 (0.14) ***
<i>X</i> years spent in AASs		- 0.026 (0.009) ***
<i>Training</i>	1.355 (0.283) ***	0.864 (0.2) ***
<i>X</i> years spent in AASs	0.052 (0.014) ***	
<i>X</i> farm income/year	- 0.000033 (0.0000088) ***	
<i>X</i> off-farm employment		- 0.298 (0.16) *
<i>X</i> network		0.062 (0.033) *
<i>Demonstration</i>	1.035 (0.079) ***	1.345 (0.083) ***
<i>Hypothetical Baseline</i>	2.357 (0.514) ***	2.789 (0.36) ***
<i>X</i> farm income/year	- 0.000086 (0.000018) ***	
<i>X</i> experienced shock		- 0.053 (0.028) *
<i>X</i> years of schooling	- 0.093 (0.036) ***	
<i>Cost</i>	- 0.00251 (0.001) **	- 0.00476 (0.00076) ***
<i>Sd_ADV</i>	0.007 (0.06)	0.177 (0.045) ***
<i>Sd_TRAIN</i>	0.13 (0.096)	0.21 (0.059) ***
<i>Sd_DEMO</i>	0.0002 (0.052)	0.0005 (0.055)
<i>Sd_HB</i>	0.44 (0.128) ***	0.073 (0.282)
LogL function	- 660.63	- 825.74
Parameters	13	13
Observations	1,440	1,440
AIC	0.936	1.165
BIC	0.983	1.213
Pseudo-R ²	0.58	0.48

Standard errors in parenthesis.

***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Note: OT = Oath Treatment; CT = Control Treatment. *Sd_ADV* = Standard deviation (Advice), *Sd_TRAIN* = Standard deviation (Training), *Sd_DEMO* = Standard deviation (Demonstration), *Sd_HB* = Standard deviation (Hypothetical Baseline).

Concerning the parameter advice in CT, the interaction (*advice*years spent in AASs*) reduced, but not completely removed, the original heterogeneity observed (i.e., comparing *Sd_ADV* for CT in Tables 3.6 & 3.7). The coefficient for this interaction is negative and highly significant, suggesting that the more the number of years of participation in general AASs, the less the probability to seek out for advice in the new FTC-based improved AAS provision. This may be due to previous accumulated knowledge and expertise as a result of several years of involvement in public AASs. It may also be because, given their large social networks and interconnectedness, farmers may obtain advice and information from model farmers, other fellow farmers, distant relatives or neighbours. Such communication is long recognized in agricultural extension and advisory services (e.g., in *Communication for Rural Innovations* by Leeuwis, 2008). Nonetheless, since new technologies and practices are continuously being generated, adapted and disseminated, farmers should be encouraged to maintain constant interactions, irrespective of their experience, with development agents operating in the area in particular, and the FTCs in general. This way, they can be up-to-date with current agricultural technologies and practices aimed at boosting production and productivity.

The significant heterogeneity around the coefficient of training in OT_{total} (though only at 10%) is completely removed by the inclusion of two interactions (*training*years spent in AAS* and *training*farm income/year*). In the CT, however, it is the interactions *training*off-farm employment* and *training*network* that resulted in a reduction of observed heterogeneity. In the OT_{total} , the interactions indicate that with increased number of years of involvement in AASs, the probability to participate in improved FTC training increases. This suggests that farmers seek out

technical skill acquisition and improvement training, in spite of their supposedly accumulated knowledge over time.⁴⁸ Nevertheless, as their farm income increases, the interaction shows the less likelihood of the farmers to take part in improved FTC training (as indicated by the negative coefficient for the interaction *training*farm income*). This may imply the existence of an inverse relationship between participation in FTC-training and farm income. It appears, from this, that income insecurity pushes farmers to the training. This reinforces the notion of recruiting low income households into the FTC training. Hence, one possible criterion in the selection of farmers can be household farm income, although the definition of low income needs to be carefully established.

From the other interactions, the negative coefficient for *training*off-farm employment* in the CT shows that the more the involvement in off-farm income generation schemes, the less the probability to participate in FTC training. One possible explanation for this can be the limited time available for farmers to involve in the modular training. Another explanation can be the fact that the farmers earn income from off-farm employment and do not bother too much about improving their farm income through FTC training. Off-/non-farm activities – such as skilled laborer, weaving/spinning, handicraft, traditional healer, trader, etc. – are important in rural livelihood diversification in most developing countries (Ellis, 1999). Concerning networks, the interaction shows that the higher the number of people in one's network, the higher the chance to take part in training (*training*network*). This can be due to the existence of ties among the network members (Granovetter, 1983) which facilitate information flow, or the presence of peer

⁴⁸ Unlike the case with advice, farmers seek out for FTC-based training even if they have many years of experience with AASs. This may be due to the fact that FTC-based training is a relatively new phenomenon and many farmers may have the desire to take part in it (especially the modular training).

group effects, in which the choices of some members of a group exert influences on the preferences of others.

Finally, the heterogeneity around the coefficient for the hypothetical baseline in OT_{total} is partly captured by the interactions *hypothetical baseline*farm income* and *hypothetical baseline*years of schooling*. Both of these interactions show negative sign for the respective coefficients. In particular, the higher the income realized from farm activities, the lower the probability to select the hypothetical baseline option. This means, the farmers with higher income will look for improved FTC services offered with payment. However, since the coefficient for the interaction *training*farm income* is negative (as discussed above), the farmers seem less likely to seek out for training, but more likely to participate in demonstration or advice as their income improves. Likewise, the interaction with years of schooling suggests that there is less chance of remaining in the hypothetical baseline for those farmers who have a relatively better education.

3.4.4 Parameter Estimates in the Sub-Sample

The MNL and RPL model results for the treatment groups (OT and CT) in the sub-sample (those trained in 2009) are given in Tables 3.8 & 3.9, respectively. Whereas the full sample in Tables 3.5 & 3.6 refers to respondents who participated in FTC training in 2008 and 2009, the sub-sample here refers to the restricted group of respondents who were trained in 2009. This sub-sample level analysis was undertaken in order to link the results of the choice experiment with that of impact evaluation conducted on this sample. Moreover, it also enables to explore preferences in a group of respondents who are more homogeneous in terms of, for example, experience with AASs. The summary statistics on the sociodemographic characteristics of the respondents in this sub-sample is given in Appendix 4. From this table, the K-S test indicates that

there exists, in general, a significant difference between the treatment groups (OT_{total} versus CT and OT_{signed} versus CT) for covariates indicating human capital (education, household size); physical capital (asset ownership, livestock possessed, certified land); financial capital (farm income, off-farm employment); social capital (networks); infrastructure and services (market access, participation in FTC-based services, level of satisfaction with FTC services); and exposure to long-term shocks.

Regarding the parameter estimates, the likelihood ratio test indicates that the hypothesis of preference equality across the two treatments in the sub-sample is rejected for both the MNL ($\chi^2_5 = 108.36$; $p < 0.001$) and RPL models ($\chi^2_9 = 109.45$; $p < 0.001$). In general, the RPL results are consistent with that of MNL in terms of sign and statistical significance. Comparing the statistics reported in the lower part of Tables 3.8 & 3.9, the RPL model (log likelihood – 339.00 and – 380.19; AIC 0.967 and 1.081; and BIC 1.024 and 1.138) appears to fit the data somehow better than the MNL (log likelihood – 343.80 and – 381.36; AIC 0.969 and 1.073; and BIC 1.001 and 1.105) when considering the criteria of log likelihood and maximum value for AIC. In terms of BIC and minimum value for AIC, the MNL is better. However, since the RPL model reveals the existence and/or lack of heterogeneity around the mean parameter estimates, the focus the discussion is on it.

Table 3.8 MNL Model Estimation Results on the Treatment Groups in the Sub-Sample Trained in 2009

Attribute parameters	OT			CT
	OT _{total}	OT _{signed}	OT _{not signed}	
Advice	-0.830 (0.103) ***	-0.866 (0.154) ***	-0.812 (0.140) ***	0.266 (0.091) ***
Training	1.179 (0.124) ***	1.275 (0.189) ***	1.116 (0.166) ***	1.303 (0.101) ***
Demonstration	0.761 (0.101) ***	0.763 (0.150) ***	0.769 (0.136) ***	1.426 (0.120) ***
Hypothetical Baseline	-0.947 (0.563) *	-1.345 (0.828)	-0.626 (0.772)	2.871 (0.519) ***
Cost	-0.00366 (0.00137) ***	-0.00491 (0.00201) **	-0.00262 (0.00188)	-0.00478 (0.00112) ***
LogL function	-343.80	-153.63	-188.08	-381.36
Parameters	5	5	5	5
Observations	720	336	384	720
AIC	0.969	0.944	1.006	1.073
BIC	1.001	1.001	1.057	1.105

Standard errors in parenthesis
 ***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Note: OT = Oath Treatment; CT = Control Treatment

Table 3.9 RPL Model Estimation Results on the Treatment Groups in the Sub-Sample Trained in 2009

Attribute parameters	OT			CT
	OT _{total}	OT _{signed}	OT _{not signed}	
Advice	-0.899 (0.111) ***	-0.897 (0.160) ***	-0.913 (0.155) ***	0.269 (0.092) ***
Training	1.245 (0.142) ***	1.307 (0.203) ***	1.181 (0.198) ***	1.346 (0.113) ***
Demonstration	0.791 (0.104) ***	0.777 (0.152) ***	0.810 (0.142) ***	1.457 (0.124) ***
Hypothetical Baseline	-1.112 (0.586) *	-1.433 (0.843) *	-0.813 (0.815)	2.942 (0.528) ***
Cost	-0.00393 (0.00141) ***	-0.00508 (0.00204) **	-0.00286 (0.00195)	-0.00484 (0.00114) ***
<i>Sd_ADV</i>	0.000046 (0.082)	0.014 (0.222)	0.164 (0.099)	0.009 (0.155)
<i>Sd_TRAIN</i>	0.327 (0.093) ***	0.230 (0.133) *	0.385 (0.139) ***	0.194 (0.083) **
<i>Sd_DEMO</i>	0.008 (0.071)	0.002 (0.098)	0.013 (0.105)	0.001 (0.058)
<i>Sd_HB</i>	0.319 (0.244)	0.081 (0.535)	0.457 (0.283)	0.002 (0.167)
LogL function	-339.00	-153.00	-184.06	-380.19
Parameters	9	9	9	9
Observations	720	336	384	720
AIC	0.967	0.964	1.006	1.081
BIC	1.024	1.067	1.098	1.138
Pseudo-R ²	0.57	0.59	0.56	0.52

Standard errors in parenthesis
 ***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Note: OT = Oath Treatment; CT = Control Treatment. Sd_ADV = Standard deviation (Advice), Sd_TRAIN = Standard deviation (Training), Sd_DEMO = Standard deviation (Demonstration), Sd_HB = Standard deviation (Hypothetical Baseline).

The RPL model results for the sub-sample (Table 3.9) broadly resemble the results discussed in the previous section (Table 3.6) in terms of sign and significance of the coefficients, except for the hypothetical baseline. However, in this sub-sample there is heterogeneity only in the parameter estimates for training. Moreover, the coefficient for the hypothetical baseline is, as expected, negative in OT (although it is positive in CT). This indicates that the respondents in the OT do not want the hypothetical baseline situation, although respondents in the CT still prefer to adhere to it (possibly because many of them have chosen it to be their second-best option as can be seen in Appendix 5). In general, training, demonstration, and cost are the main attributes – in order of magnitude of importance – that affect the probability of participation in paid AASs in this sub-sample. The negative and significant coefficient for advice in OT may be due to the reasons discussed in section 3.4.2.

The within group comparison in the OT (OT_{signed} and $OT_{not\ signed}$) also shows that in addition to being experienced with the services, agreeing to tell the truth and/or signing on a solemn oath form may have helped to significantly reduce (especially in OT_{signed}) selection of the hypothetical baseline option compared to the CT (although in $OT_{not\ signed}$ the coefficient for hypothetical bias is not significant but with the correct sign). Moreover, those who signed on the solemn oath form showed significant dedication in evaluating the cost attribute compared to those who did not sign.

3.4.5 Marginal Willingness to Pay (WTP) Estimates

The marginal WTPs – estimated based on the RPL values – for the different services describing improved FTC-based AASs are presented in Table 3.10. These WTPs are estimated for each household – on the basis of the individual betas – and then averaged. Concentrating at the WTP estimates in the full sample (top half of the table), whereas the highest premium is associated with training, followed by demonstration in the OT, the pattern is reversed in the CT (i.e., respondents in the CT want more demonstration services than training). Moreover, there is a small premium price associated with advice in the CT, suggesting that all the three components of improved FTC-based AASs are important in this group (though their importance varies in terms of magnitude).⁴⁹ Although receiving at least one advice/visit from development agents was shown to significantly reduce poverty and increase consumption growth in the country (Dercon et al., 2009)⁵⁰, in this study no strong evidence of preference for advice provided through the FTCs is found (except for farmers in the CT group who received significantly less number of advice, on average). In the present investigation, it appears that the respondents in the OT group want to be compensated or subsidized to get advice from the FTCs. This may be associated with the low quality (in terms of timeliness, relevance, and content) of current advice as reflected by the low level of satisfaction with the service.

⁴⁹ A similar pattern also emerges at the sub-sample level (bottom half of Table 3.10).

⁵⁰ In fact, the study by Dercon and his colleagues is based on the public AASs, where the farmers get advice/visit free of charge.

Table 3.10 Marginal Willingness to Pay Estimates (in Birr) for Advice, Training, and Demonstration in the Full Sample and Sub-Sample

	(1) OT _{total}	(2) t ^a	(3) OT _{signed}	(4) t ^b	(5) OT _{not signed}	(6) t ^c	(7) CT
<i>Full sample: trained in 2008 & 2009</i>							
Advice	-280.61	79.92 (0.000)	-249.12	50.41 (0.000)	-322.74	64.20 (0.000)	68.99
Training	677.61	72.86 (0.000)	612.86	51.67 (0.000)	755.65	42.99 (0.000)	250.44
Demonstration	440.81	1.6e+04 (0.000)	382.18	3.7e+03 (0.000)	511.24	3.4e+04 (0.000)	281.55
<i>Sub-sample: trained in 2009 only</i>							
Advice	-228.97	2.4e+04 (0.000)	-176.40	7.9e+03 (0.000)	-318.45	4.8e+03 (0.000)	55.52
Training	316.94	3.50 (0.001)	257.63	2.62 (0.012)	417.07	7.40 (0.000)	277.82
Demonstration	201.54	7.0e+03 (0.000)	152.94	7.5e+04 (0.000)	282.99	249.60 (0.000)	301.18

Note: OT = Oath Treatment; CT = Control Treatment.

^a Based on the assumption that the absolute value of t-statistic (mean differences between CT and OT_{total}, i.e., mean(CT) – mean(OT_{total}) is not different from 0). *p*-values in parenthesis.

^b Based on the assumption that the absolute value of t-statistic (mean differences between CT and OT_{signed}, i.e., mean(CT) – mean(OT_{signed}) is not different from 0). *p*-values in parenthesis.

^c Based on the assumption that the absolute value of t-statistic (mean differences between CT and OT_{not signed}, i.e., mean(CT) – mean(OT_{not signed}) is not different from 0). *p*-values in parenthesis.

Although not shown in the table, the difference of means between OT_{not signed} and OT_{signed} (for the full sample) is not different from zero for advice (| t | = 1.1e+03, *p* = 0.000); is not different from zero for training (| t | = 10.28, *p* = 0.000); and, is not different from zero for demonstration (| t | = 3.3e+03, *p* = 0.000). The same is true for the sub-sample (trained in 2009): the difference of means is not different from zero for advice (| t | = 1.2e+03, *p* = 0.000); is not different from zero for training (| t | = 5.81, *p* = 0.000); and, is not different from zero for demonstration (| t | = 1.2e+03, *p* = 0.000). However, since the coefficient for the cost attribute is not significant for the OT_{not signed} treatment in both the full sample and sub-sample (Tables 3.6 & 3.9, respectively), the WTP estimates for this group and the comparisons with OT_{signed} and CT have to be approached with caution.

For example, it is actually observed in the data that only about 3% of the respondents in the OT_{total} indicated that they are very satisfied with the current provision of advice. Compared to the corresponding figures for training (about 77%) and demonstration (about 23%), this is clearly a low level of satisfaction (Table 3.4). Rationally, one would expect the respondents to involve highly in this improved advice provision through the FTCs. However, it seems that the observed level of satisfaction with current provision of advice, among other factors, may have influenced the respondents not to put too much weight on this attribute, in spite of the fact that the choice experiment promised an improved advice provision on crop, livestock, and natural resource management through the FTCs.

The between-treatment comparison of mean WTP in the full sample indicates the existence of variations in preference for the services describing improved FTC-based AASs. Looking at the WTPs for training and demonstration between OT_{total} and CT, it is apparent that the OT_{total} mean values are larger than the corresponding CT values. The t-statistic shows that the difference is statistically significantly different from zero (column (2) of Table 3.10). The pattern of larger WTPs in OT remains largely the same as one moves along the comparisons OT_{signed} versus CT and $OT_{not\ signed}$ versus CT as indicated in columns (4) and (6) of the same table. However, the comparison between mean WTPs in the OT_{signed} versus $OT_{not\ signed}$ indicates that the mean WTP for those respondents who signed on the solemn oath form is lower than those who did not. Although this may suggest that signing on a solemn oath form has a better effect in reducing mean WTPs than not signing, since the coefficient for the cost attribute is not significant for the $OT_{not\ signed}$ group (Table 3.6), this remark has to be taken with caution. A similar caution also applies to the above comparison between $OT_{not\ signed}$ and CT. In general, however, in the full sample it is evident that signing on an oath form and/or agreeing to tell the truth in the

experiment had no effect in reducing mean WTPs when compared to the respondents who are not exposed to such hypothetical bias reduction treatment (i.e., OT_{total} versus CT).

At the sub-sample level (bottom half of Table 3.10) and considering only the magnitude of the estimates, however, when the groups are more homogenous in terms of experience in AASs, there is evidence of reduced marginal WTPs as a result of exposure to the solemn oath script. First, the mean WTPs for training and demonstration in OT_{signed} are lower than the corresponding values for the CT (column (4) of Table 3.10). Second, the mean WTP for demonstration in the $OT_{not\ signed}$ is smaller than the CT. Third, comparing OT_{signed} and $OT_{not\ signed}$, the mean WTPs for training and demonstration are lower in the former compared to the latter. This reinforces the views of Jacquemet et al. (2013) that signing on an oath form strengthens the commitment needed to reveal candid answers than orally agreeing to do so. Finally, the mean WTP for demonstration is lower in the OT_{total} compared to CT. These findings are in general agreement to previous studies documenting the positive role of solemn oath in preference elicitation in stated preference research (Jacquemet et al., 2010; Carlsson et al., 2013; Jacquemet et al., 2013; Stevens et al., 2013; de-Magistris & Pascucci, 2014). In particular, quite similar to the results of the choice experiment conducted by de-Magistris & Pascucci (2014), in this study lower WTP values are found for training and demonstration among those who signed on the solemn oath form. Hence, the oath is more effective when it is signed.

Unlike the studies reported above, the WTPs for a sub-group of respondents who did not sign on the oath form (but orally agreed to tell truthful answers) are also estimated.⁵¹ The results are mixed – whereas agreeing to tell the truth reduced mean WTP for demonstration, compared to the CT, it increased mean WTP for training. At aggregate level also (comparing OT_{total} & CT),

⁵¹ Unfortunately, cost is not significant in $OT_{not\ signed}$ group.

the oath appears to have reduced the mean WTP for demonstration but increased that for training. Evidence on mixed effectiveness of the solemn oath procedure can also be found in the climate-change related study of Carlsson et al. (2013). In that study, the authors show that although the oath decreased the shares of extremely high WTPs and zero WTPs, it significantly decreased the conditional WTP in China, but not in Sweden.

Although the primary objective of the present study is deriving the WTP estimates for the components describing improved FTC-based AASs, a further investigation is also made to examine the changes in the WTPs on the basis of changes in socio-demographic characteristics (SDCs) of the respondents in the treatment groups. Doing such an analysis has at least two implications. On the one hand, it enables to have a deeper understanding of the existence and/or lack of variation in WTPs across the respondents. And, on the other, it helps to uncover the reasons/determinants of such variations in WTPs. In what follows, the univariate summary statistics on the relationship between selected SDCs and changes in the WTP for advice, training, and demonstration – calculated on *STATA* 11 – are discussed for the ‘full sample’ (trained at FTCs in 2008 and 2009) and ‘sub-sample’ (trained at FTCs in 2009 only). The remaining discussion on this matter (i.e., for those trained at FTCs in 2009 and signed on the solemn oath form) is provided in Appendix 6.

i. Changes in the WTP in Relation to Socio-demographic Characteristics (‘Full Sample’)

The changes in the WTP (for advice, training, and demonstration) as a result of changes in some SDCs and their categories (i.e., human capital, physical capital, economic/financial capital, social capital, access to services, and others) are given in Table 3.11 for those households trained at FTCs in 2008 and 2009 (i.e., ‘the full sample’ according to Table 3.10). Generally, the factors

responsible for the lowest level of negative WTP (in the case of advice) and the lowest WTP (in the case of training and demonstration) are related mainly to human capital, physical capital, and social capital.

From Table 3.11, it is evident that human capital (household size < 5 members), physical capital (assets: 15,000 – 20,000 birr; land holding: 0.25 – 0.5 ha), social capital (network size > 10 people), and intensity of FTC-based advice (5 – 15 per year) describe the households who indicated the lowest level of negative WTP for the FTC-based advice in the OT group. In addition to these, it is found that uneducated households with a rich experience in farming and in on-farm AASs, who possess livestock whose value is > 60,000 birr, but earn an annual income of < 10,000 birr also showed a lower level of negative WTP for getting advice from the FTCs.

In relation to the WTP for training, the same kinds of variables (and their categories) that resulted in the lowest level of negative WTP for advice are also associated with the smallest amount of WTP in the OT group. Moreover, households with no formal education background but with many years of experience in farming and in AASs, and those with livestock possession valued at > 60,000 birr also indicated a lower WTP for training (compared to the other households in the respective categories). Likewise, households who participated in an FTC-based training (short-term and modular) 4 – 8 times a year but earn small amount of farm income (i.e., < 10,000 birr/year) also showed a lower WTP for training.

Looking at the changes in the WTP for demonstration, one also notices an exactly same kind of variable categories (as for training) being responsible for the lowest amount of WTP in the OT group. It can be concluded, hence, that the SDCs (and their categories) determining the lowest level of WTP for training and demonstration are largely the same. Nevertheless, as indicated in

the earlier sections of the discussion on WTP, the comparison of WTP for training and demonstration between OT group and CT group shows that the solemn oath script did not result in lower WTPs at the 'full sample' level.

Table 3.11 Changes in the WTP (birr) Based on Some SDCs (Full Sample – Trained in 2008 and 2009)

	Advice		Training		Demonstration	
	OT _{total}	CT	OT _{total}	CT	OT _{total}	CT ^b
Age (years) ^a : 15 – 64	-287.2	69.0	686.6	250.4	448.9	281.6
Male household head	-287.8	69.0	687.9	250.4	450.0	281.6
Education ^a : No formal education	-275.4	69.4	661.1	252.8	428.3	281.6
Primary school (1 – 4)	-304.3	62.2	709.3	243.1	478.9	
Primary and lower secondary (5 – 8)	-316.1	71.5	749.8	237.4	499.5	
Secondary school (9 – 12)	-322.8		782.2		511.2	
Household size (number) ^a : < 5	-248.8		613.0		381.8	
5 – 10	-287.8	69.0	687.9	250.4	450.0	
Experience in farming (years): < 15		66.3		278.6		281.6
15 – 25	-294.6	67.8	708.5	252.1	461.9	
25 – 30	-279.3	79.3	657.7	234.6	435.0	
> 30	-267.7	47.1	659.4	271.8	414.5	
Experience with on-farm AASs (years) ^a : 5 – 15	-300.1	74.3	720.4	250.6	471.5	281.6
15 – 25	-277.3	62.1	660.8	250.3	431.5	
Value of household assets (birr): < 5,000	-302.9	67.3	726.3	245.6	476.5	281.6
5,000 – 10,000	-282.3	84.1	659.1	254.5	440.3	
10,000 – 15,000	-285.9	76.5	713.7	240.0	446.6	
15,000 – 20,000	-248.6	52.3	613.0	268.9	382.1	
> 20,000	-262.5	70.9	644.5	259.8	405.5	
Land holding size (ha): < 0.25		68.4		225.7		281.6
0.25 – 0.5	-249.4		613.3		382.4	
0.5 – 0.75	-285.9	67.5	670.4	251.6	446.7	
> 0.75	-289.3	70.3	701.7	252.7	452.6	
Certified land	-287.2	32.9	686.6	237.1	448.9	281.6
Value of livestock possessed (birr) ^a : 10,000 – 30,000	-285.9	81.8	690.2	240.3	446.7	281.6
30,000 – 60,000	-288.3	65.8	686.6	254.8	450.8	
> 60,000	-273.9	72.8	674.7	243.4	425.2	
Farm income (birr) ^a : < 10,000	-273.7	21.2	657.5	280.9	425.3	281.6
10,000 – 20,000	-290.0	69.9	697.8	264.7	453.9	
20,000 – 30,000	-287.5	73.4	680.0	247.0	449.5	
30,000 – 40,000	-322.8	62.7	798.9	243.0	511.2	
Involved in off-farm activities	-279.5	75.6	671.6	219.2	435.3	281.6
Have access to microfinance credit	-322.8	56.7	778.7	250.9	511.2	281.6
Have access to daily market	-248.9	63.3	613.0	255.2	381.9	281.6
Number of people in one's network: < 5	-284.1	91.6	676.2	230.5	443.5	281.6
5 – 10	-295.6	62.7	712.5	255.3	463.7	
> 10	-248.7	63.0	612.3	284.5	382.3	
Number of FTC-based advice per year: < 5	-322.7	70.4				
5 – 15	-248.9	70.0				
15 – 30	-284.2	63.4				
> 30	-298.2					
Number of FTC-based training per year: < 4			786.9	248.9		
4 – 8			650.8	257.5		
> 8			695.5			
Number of FTC-based demonstration per year: < 2					455.9	281.6
2 – 4					442.8	
> 4					511.2	
Satisfied with current advice provision	-288.9	68.5				
Satisfied with current provision of training			686.6	249.9		
Satisfied with current provision of demonstration					399.8	281.6
Experienced long-term shock	-288.5	65.1	689.2	243.0	451.2	281.6

Note:

OT_{total} = Oath Treatment; CT = Control Treatment. Highlighted in bold are the smallest levels of negative WTPs (in the case of *advice*) and the lowest WTPs for *training* and *demonstration* in the OT (i.e., for those who signed on the solemn oath form and/or orally agreed to tell the truth in the CE).

^a The following categories (not shown in the table) have no observations up on which to compute the WTPs: age (< 15 and > 64 years); education (some college/university); household size (> 10 members); experience with on-farm AASs (< 5 years); household assets (< 10,000 birr); value of household assets (< 10,000 birr); and, farm income (> 40,000 birr).

^b Except for *education* (9 – 12), *household size* (< 5), and *land holding size* (0.25 – 0.5), all categories of relevant variables have the same WTP for demonstration in this treatment group (due to the fact that the WTP for demonstration does not vary much among the individual households in this group).

ii. Changes in the WTP in Relation to Socio-demographic Characteristics ('Sub-Sample')

Table 3.12 presents the changes in WTP according to some SDCs for a sub-group of the respondents trained at FTCs in 2009. From the table, it is evident that the lowest level of negative WTP (for advice) and the lowest WTPs (for training and demonstration) are associated with human capital and physical capital. In relation to the indicators of human capital, households with less than five members and who are highly experienced in farming (i.e., farming experience > 30 years) showed the lowest level of negative WTP for FTC-based advice. Regarding physical capital, the investigation shows that the lowest negative WTPs for advice are indicated by the households who manage small plots of land (i.e., 0.25 to 0.5 ha) and whose asset possession is valued at 10,000 to 15,000 birr and > 20,000 birr.

With respect to the changes in WTP for training and demonstration, the analysis reveals that the lowest WTPs for training are related to physical capital (i.e., 'value of livestock possessed > 60,000 birr'). The variables (and variable categories) responsible for the lowest level of demonstration in the OT group are exactly the same as those responsible for the lowest level of negative WTP for advice. Comparing OT group and CT group, one also notices a lower WTP for training and demonstration (for some variable categories) in the OT group, reinforcing the earlier

claim that, for example, the oath script helped to obtain a lower WTP estimate for demonstration in OT_{total} .

Table 3.12 Changes in the WTP (birr) Based on Some SDCs (Sub-Sample – Trained in 2009)

	Advice		Training		Demonstration	
	OT _{total}	CT ^b	OT _{total}	CT	OT _{total}	CT ^c
Age (years) ^a : 15 – 64	-252.2	55.5	342.7	277.8	222.3	301.2
Male household head	-254.8	55.5	344.9	277.8	224.7	301.2
Education ^a : No formal education	-244.8	55.5	334.6	277.6	215.6	301.2
Primary school (1 – 4)	-318.3		451.5	286.0	282.4	
Primary and lower secondary (5 – 8)	-318.9		343.0	274.6	283.1	
Household size (number) ^a : < 5	-176.4		278.0		152.9	
5 – 10	-254.8	55.5	344.9	277.8	224.7	301.2
Experience in farming (years): < 15		55.5		304.3		301.2
15 – 25	-277.9		389.4	277.0	245.8	
25 – 30	-241.9		310.2	277.0	213.0	
> 30	-176.4		265.5	275.7	152.9	
Experience with on-farm AASs (years) ^a : 5 – 15	-283.1	55.5	394.6	275.0	250.5	301.2
15 – 25	-240.9		323.8	281.6	212.1	
Value of household assets (birr): < 5,000	-318.5	55.5	371.2	272.3	283.2	301.2
5,000 – 10,000	-271.1		381.1	281.1	239.5	
10,000 – 15,000	-176.2		282.3	283.0	152.9	
15,000 – 20,000	-318.3		494.6	276.3	283.0	
> 20,000	-176.4		259.0	278.1	152.9	
Land holding size (ha): < 0.25		55.5		259.6		301.2
0.25 – 0.5	-176.3		282.8		152.9	
0.5 – 0.75	-257.6		365.6	279.6	227.2	
> 0.75	-252.2		325.3	280.1	222.4	
Certified land	-252.2		342.7		222.3	
Value of livestock possessed (birr) ^a : 10,000 – 30,000	-233.1	55.5	356.7	266.4	204.9	301.2
30,000 – 60,000	-256.8		350.1	277.7	226.4	
> 60,000	-246.8		221.6	280.7	218.5	
Farm income (birr) ^a : < 10,000	-247.4		389.5		218.0	
10,000 – 20,000	-274.7	55.5	347.2	280.9	243.0	301.2
20,000 – 30,000	-228.1		311.7	273.0	200.3	
30,000 – 40,000				285.7		
Involved in off-farm activities	-219.0	55.6	311.2	277.3	192.0	301.2
Have access to microfinance credit		55.5		281.5		301.2
Have access to daily market	-176.3	55.5	282.4	278.3	152.9	301.2
Number of people in one's network: < 5	-266.8	55.5	356.0	276.7	235.7	301.2
5 – 10	-196.7		279.3	276.0	171.5	
> 10	-318.5		492.5	323.4	283.1	
Number of FTC-based advice per year: < 5		55.5				
5 – 15	-247.4					
15 – 30	-247.5					
> 30	-270.9					
Number of FTC-based training per year: < 4				278.2		
4 – 8			378.9	276.1		
> 8			321.7			
Number of FTC-based demonstration per year: < 2					283.0	301.2
2 – 4					215.6	
> 4						
Satisfied with current advice provision	-247.4	55.5				
Satisfied with current provision of training			342.7	277.5		
Satisfied with current provision of demonstration					187.2	301.2
Experienced long-term shock	-247.5	55.5	343.4	281.8	217.9	301.2

Note:

OT_{total} = Oath Treatment; CT = Control Treatment. Highlighted in bold are the smallest levels of negative WTPs (in the case of *advice*) and the smallest WTPs for *training* and *demonstration* in the OT (i.e., for those who signed on the solemn oath form and/or orally agreed to tell the truth in the CE).

^a The following categories (not shown in the table) have no observations up on which to compute the WTPs: age (< 15 and > 64 years); education (9 – 12 and some college/university); household size (> 10 members); experience with on-farm AASs (< 5 years); value of livestock (< 10,000 birr); and, farm income (> 40,000 birr).

^b Except for *household size* (< 5), *land holding size* (0.25 – 0.5), *certified land*, *farm income* (< 10,000), and *number of FTC-based advice* (> 30), all categories of relevant variables have the same WTP for advice in this treatment group (due to the fact that the WTP for advice does not vary much among the individual households in this group).

^c Except for *household size* (< 5), *land holding size* (0.25 – 0.5), *certified land*, and *farm income* (< 10,000), all categories of relevant variables have the same WTP for demonstration in this treatment group (due to the fact that the WTP for demonstration does not vary much among the individual households in this group).

3.5 Conclusion and Recommendation

In the present choice experiment, farmers' WTP for attributes comprising improved FTC-based AAS provision in Ethiopia are investigated. By using the best-worst method to eliciting farmers' preferences and estimating both the MNL and RPL models, the farmers' WTP for advice, training, and demonstration are analyzed. Most importantly, the study shows that the use of the solemn oath script (Jacquemet et al., 2013) helps to obtain lower WTP values, and thereby reduce hypothetical bias in hypothetical choice experiments such as this one. However, to what extent the hypothetical bias was reduced is not crystal clear since one needs to compare hypothetical and real choice experiments to arrive at this. Although conducting a real choice experiment in AASs can be hugely unfeasible, one can inform the farmers that a follow up study involving a real payment will be conducted, in order to somehow mimic the 'real' behavior of the farmers. In doing so, it would be possible to compare the results of the 'truly' hypothetical CEs with this one and examine the extent of hypothetical bias reduction.

The novelty of the approach followed in this study is the use of the best-worst method to generate preference data (Louviere & Islam, 2008; Scarpa et al., 2011) combined with the solemn oath approach to reduce hypothetical bias (e.g., Jacquemet et al., 2013) in the context of non-market good/service valuation in the field in a developing country. Such a combined approach has not been employed in any previous study. Moreover, the use of the solemn oath procedure is slightly different from previous studies in that a two-step process was followed. First, the respondents were asked whether they agree to tell the truth in the survey. Following this, those respondents who agreed to tell the truth were asked to sign on the solemn oath form. This procedure enabled to analyze the effect of signing on an oath form separately from the effect of merely agreeing to tell the truth. Although the results show that signing on an oath form

is better than simply agreeing to tell the truth (as also indicated by Jacquemet et al., 2013) – in terms of reducing WTPs – agreeing to tell the truth is also better than nothing at least in some way.

In addition to this, the present study shows that interacting selected sociodemographic characteristics with the attributes can explain part of the observed heterogeneity. In this regard, of the several interaction effects explored, it is found that human capital (education), financial capital (farm income, off-farm employment), social capital (networks), years of participation in AASs, and long-term shocks explain observed heterogeneity around advice, training, and hypothetical baseline. No evidence is found for the effect of physical capital (such as, land holding size, value of household assets, value of livestock possessed), other human capital indicators (household size, age), and access to institutions and services (market, microfinance credit).

Based on the findings and general observations in this study, the following recommendations are suggested. To start with the policy-relevant implications of the study, since the coefficients and marginal WTP estimates for advice are negative and not significant in the OT group, the provision of this service through the FTCs emerges to be strongly disliked by the participants. Therefore, the provision of this service should be strengthened and reorganized, considering its importance in poverty reduction and consumption growth. This may require stakeholders' full participation in deciding the contents, time, mode of delivery, feedback loop and other related aspects. Moreover, since an inverse relationship between farm income and participation in FTC-based AASs (i.e., training) is observed, the current FTC-based AASs should focus more on low-income farmers. In this regard, there should be an improved targeting of FTC-training towards the poorest segments of the population.

In addition to these, there should be a differential targeting scheme in order to encourage young and less-experienced households; households with low number of members; and, asset-constrained households. Such an encouragement design should also consider addressing members of the household instead of the usual household head based approach. This can improve, for example, the involvement of women and younger members of the household in the FTC-based provision of AASs. A specific attention should also be given to improve the farmers' access to microfinance credit. Moreover, since the positive coefficients and marginal WTP estimates for training and demonstration indicate, indirectly, the impact of these services, an organized and needs-based FTC services should be designed and expanded to the wider rural communities.

Advancing the recommendations, the following limitations of the study and other methodological issues should be considered in future CEs. The first limitation relates to the use of hypothetical baselines as status quo. Although the use of hypothetical baselines is justifiable in situations where a standard status quo option is difficult to define, such baselines can add to the complexity of the choice task and, therefore, be cognitively more challenging (Whittington & Adamowicz, 2010). The challenge becomes even more daunting when hypothetical baseline is coupled with low levels of education of respondents, a few (or no) previous experience with choice experiments, and hypothetical nature of the study. Perhaps such complications may have contributed to the observed choice of the hypothetical baseline by some respondents as their second-best alternative in this study. Nonetheless, future research in developing countries can explore the pros and cons associated with the use of hypothetical baselines in hypothetical CEs.

The second issue revolves around the use of the solemn oath. Given that the use of the solemn oath script in CEs is at its infancy, and more so is in the context of developing countries, it is

highly recommended to encourage future research to employ this approach in real and hypothetical CEs and analyze the extent of hypothetical bias reduction. Furthermore, combining two or more methods of circumventing hypothetical bias can also provide enriched information on the effectiveness of the various methods in bias reduction or elimination. However, since taking an oath in a survey is unusual in most cultures, a subtle way of introducing the oath script (as in Carlsson et al., 2013) should be pursued. In addition, an approach that compares signing on an oath form with a mere agreement to reveal truthful answers should also be explored in large samples.

Likewise, since this study suggests the suitability of the best-worst approach to applications in developing countries, future research should explore this preference elicitation procedure in a wide range of environmental valuation exercises in various settings. In particular, since the lowly educated most rural households in developing countries can easily identify extreme values in CEs (Marley & Louviere, 2005; Vermeulen et al., 2010), the best-worst approach can be beneficial, provided that it is carefully designed and implemented.

Finally, since making trade-offs among the levels describing each attribute in a best-worst format in developing countries can be challenging for rural dwellers, it is suggested that choice experiment surveys be conducted independently. They should not be mixed with or be part of other larger surveys, as is the case with the present study. For example, integrating a choice experiment survey in national or local survey may result in poor quality choice data, since choice experiment survey by itself is cognitively very demanding and time consuming, at least in the context of developing countries. Future choice experiment research in AASs can also be conducted by exclusively considering both current users and non-users of FTC-based AASs. In

doing so, one can examine the potential for scaling up FTC-based AASs to farmers who are not currently beneficiaries of such services.

References

- Aadland, D., & Caplan, A. J. (2006). Cheap talk reconsidered: New evidence from CVM. *Journal of Economic Behavior & Organization*, 60(4), 562–578.
- Abesha, D., Waktola, A., & Aune, J. B. (2000). Agricultural Extension in the Drylands of Ethiopia. Drylands Coordination Group. Report no. 9, September 2000. Norway.
- Abro, Z. A., Alemu, B. A., & Hanjra, M. A. (2014). Policies for Agricultural Productivity Growth and Poverty Reduction in Rural Ethiopia. *World Development*, 59, 461–474.
- Adamowicz, W., Boxall, P., Williams, M., & Louviere, J. (1998). Stated preference approaches for measuring passive use values: choice experiments and contingent valuation. *American Journal of Agricultural Economics*, 80(1), 64–75.
- Agimass, F., & Mekonnen, A. (2011). Low-income fishermen's willingness-to-pay for fisheries and watershed management: An application of choice experiment to Lake Tana, Ethiopia. *Ecological Economics*, 71 (2011), 162–170.
- Agresti, A. (1996). *An introduction to categorical data analysis* (Vol. 135). New York: Wiley.
- Agresti, A. (2007). *An introduction to categorical data analysis*. Hoboken, NJ: Wiley.
- Agresti, A. (2014). *Categorical data analysis*. John Wiley & Sons.
- Akroyd, S., & Smith, L. (2007). The decline in public spending to agriculture - does it matter? Oxford Policy Management (OPM) briefing notes 2007 - 02. <http://siteresources.worldbank.org/EXTRESPUBEXPANAAGR/Resources/opm1.pdf> [accessed on 12 Feb 2013]
- Alene, A. D., & Hassan, R. M. (2003, August). Measuring the impact of Ethiopia's new extension program on the productive efficiency of farmers. In *Proceedings of the 25th International Conference of Agricultural Economists (IAAE)* (Vol. 16, p. 22).
- Alfnes, F., Yue, C., & Jensen, H. H. (2010). Cognitive dissonance as a means of reducing hypothetical bias. *European Review of Agricultural Economics*, 37(2), 147–163.
- Amare, M., Asfaw, S., & Shiferaw, B. (2012). Welfare impacts of maize–pigeonpea intensification in Tanzania. *Agricultural Economics*, 43(1), 27–43.
- Ami, D., Aprahamian, F., Chanel, O., & Luchini, S. (2011). A test of cheap talk in different hypothetical contexts: The case of air pollution. *Environmental and Resource Economics*, 50(1), 111–130.
- Anaglo, J. N., Boateng, S. D., & Boateng, C. A. (2014). Gender and Access to Agricultural Resources by Smallholder Farmers in the Upper West Region of Ghana. *Journal of Education and Practice*, 5(5), 13–19.
- Anderson, J. R. (2008). Agricultural advisory services. <https://openknowledge.worldbank.org/handle/10986/9041> [accessed on 17 Oct 2012]
- Arpino, B., & Aassve, A. (2013). Estimating the causal effect of fertility on economic wellbeing: data requirements, identifying assumptions and estimation methods. *Empirical Economics*, 44(1), 355–385.
- Asenso-Okyere, K., Davis, K., & Aredo, D. (2008). *Advancing agriculture in developing countries through knowledge and innovation: Synopsis of an international conference*. Intl Food Policy Res Inst.
- A Smith, J., & E Todd, P. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125(1), 305–353.

- Asrat, S., Yesuf, M., Carlsson, F., & Wale, E. (2010). Farmers' preferences for crop variety traits: Lessons for on-farm conservation and technology adoption. *Ecological Economics*, 69(12), 2394–2401.
- Ayele, G., Bekele, M., & Zekeria, S. (2006). Productivity and Efficiency of Agricultural Extension Package in Ethiopia. The Ethiopian Development Research Institute (EDRI). Research Report 5. Addis Ababa, Ethiopia June, 2006.
- Balcombe, K., Chalak, A., & Fraser, I. (2009). Model selection for the mixed logit with Bayesian estimation. *Journal of Environmental Economics and Management*, 57(2), 226–237.
- Ballantyne, A. O. (1987). Extension work in the small farm sector. *Agricultural Administration and Extension*, 24(3), 141–147. doi:10.1016/0269-7475(87)90099-7
- Baloch, A. M., & Thapa, B. G. (2014). Agricultural extension in Balochistan, Pakistan: Date palm farmers' access and satisfaction. *Journal of Mountain Science*, 11(4), 1035–1048.
- Bateman, I. J., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., ... Özdemiroglu, E. (2002). Economic valuation with stated preference techniques: a manual. *Economic Valuation with Stated Preference Techniques: A Manual*.
- Becerril, J., & Abdulai, A. (2010). The impact of improved maize varieties on poverty in Mexico: a propensity score-matching approach. *World Development*, 38(7), 1024–1035.
- Beharry-Borg, N., & Scarpa, R. (2010). Valuing quality changes in Caribbean coastal waters for heterogeneous beach visitors. *Ecological Economics*, 69(5), 1124–1139.
- Belay. (2003). Agricultural extension in Ethiopia: the case of participatory demonstration and training extension system. *Journal of Social Development in Africa*, 18(1). doi:10.4314/jsda.v18i1.23819
- Belay, K. (2002). Constraints to agricultural extension work in Ethiopia: The insiders' view. *South African Journal of Agricultural Extension*, 31(1), 63–79.
- Belay, K., & Abebaw, D. (2004). Challenges Facing Agricultural Extension Agents: A Case Study from South-western Ethiopia. *African Development Review*, 16(1), 139–168. doi:10.1111/j.1467-8268.2004.00087.x
- Belay, K., & Manig, W. (2004). Access to rural land in Eastern Ethiopia: Mismatch between policy and reality. *Journal of Agriculture and Rural Development in the Tropics and Subtropics (JARTS)*, 105(2), 123–138.
- Benin, S., Nkonya, E., Okecho, G., Randriamamonjy, J., Kato, E., Lubade, G., & Kyotalimye, M. (2011). Returns to spending on agricultural extension: the case of the National Agricultural Advisory Services (NAADS) program of Uganda†. *Agricultural Economics*, 42(2), 249–267.
- Benin, S., Nkonya, E., Okecho, G., Randriamamonjy, J., Kato, E., Lubade, G., & Kyotalimye, M. (2012). Impact of the National Agricultural Advisory Services (Naads) program of Uganda: Considering Different Levels of Likely Contamination with the Treatment. *American Journal of Agricultural Economics*, 94(2), 386–392. doi:10.1093/ajae/aar094
- Bennett, J., & Birol, E. (Eds.). (2010). *Choice experiments in developing countries: Implementation, challenges and policy implications*. Cheltenham: Edward Elgar.
- Bennett, J., & Birol, E. (2010). Concluding remarks and recommendations for implementing choice experiments in developing countries. *Choice Experiments in Developing Countries: Implementation, Challenges and Policy Implications*, 297–306.
- Berhane, G., & Gardebreek, C. (2011). Does microfinance reduce rural poverty? Evidence based on household panel data from northern Ethiopia. *American Journal of Agricultural Economics*, 93(1), 43–55.

- Berhanu, K. (2012). The Political Economy of Agricultural Extension in Ethiopia: Economic Growth and Political Control. FAC Political Economy of Agricultural Policy in Africa (PEAPA). Working Paper 042. <http://www.futre-agricultures.org/> [accessed on 22 March 2013]
- Berhanu, K., & Poulton, C. (2014). The Political Economy of Agricultural Extension Policy in Ethiopia: Economic Growth and Political Control. *Development Policy Review*, 32(s2), s199–s216.
- Beyene, F. (2003). Estimating Attitude of Farmers Toward Maize Extension Package Program: Package Policy Assessment in Hararghe Highlands (Eastern Ethiopia). *AIAEE*.
- Bezu, S., & Barrett, C. (2012). Employment dynamics in the rural nonfarm sector in Ethiopia: Do the poor have time on their side? *Journal of Development Studies*, 48(9), 1223–1240.
- Bezu, S., & Holden, S. (2014). Are Rural Youth in Ethiopia Abandoning Agriculture? *World Development*, 64, 259–272.
- Birkhaeuser, D., Evenson, R. E., & Feder, G. (1991). The economic impact of agricultural extension: A review. *Economic Development and Cultural Change*, 39(3), 607–650.
- Birner, R., Davis, K., Pender, J., Nkonya, E., Anandajayasekeram, P., Ekboir, J. M., ... & Kisamba-Mugerwa, W. (2006). *From "best practice" to "best fit": a framework for designing and analyzing pluralistic agricultural advisory services*(No. 4). International Food Policy Research Institute (IFPRI).
- Birner, R., Davis, K., Pender, J., Nkonya, E., Anandajayasekeram, P., Ekboir, J., ... Cohen, M. (2009). From Best Practice to Best Fit: A Framework for Designing and Analyzing Pluralistic Agricultural Advisory Services Worldwide. *The Journal of Agricultural Education and Extension*, 15(4), 341–355. doi:10.1080/13892240903309595
- Blaauw, D., Erasmus, E., Pagaiya, N., Tangcharoensathein, V., Mullei, K., Mudhune, S., ... Lagarde, M. (2010). Policy interventions that attract nurses to rural areas: a multicountry discrete choice experiment. *Bulletin of the World Health Organization*, 88(5), 350–356.
- Blamey, R. K., Bennett, J. W., Louviere, J. J., Morrison, M. D., & Rolfe, J. (2000). A test of policy labels in environmental choice modelling studies. *Ecological Economics*, 32(2), 269–286.
- Blamey, R. K., Bennett, J. W., Louviere, J. J., Morrison, M. D., & Rolfe, J. C. (2002). Attribute causality in environmental choice modelling. *Environmental and Resource Economics*, 23(2), 167–186.
- Bliemer, M. C., Rose, J. M., & Hensher, D. A. (2009). Efficient stated choice experiments for estimating nested logit models. *Transportation Research Part B: Methodological*, 43(1), 19–35.
- Blumenschein, K., Blomquist, G. C., Johannesson, M., Horn, N., & Freeman, P. (2008). Eliciting Willingness to Pay Without Bias: Evidence from a Field Experiment*. *The Economic Journal*, 118(525), 114–137.
- Bosworth, R., & Taylor, L. O. (2012). Hypothetical Bias in Choice Experiments: Is Cheap Talk Effective at Eliminating Bias on the Intensive and Extensive Margins of Choice? *The BE Journal of Economic Analysis & Policy*, 12(1).
- Brehanu, A., & Fufa, B. (2008). Repayment rate of loans from semi-formal financial institutions among small-scale farmers in Ethiopia: Two-limit Tobit analysis. *The Journal of Socio-Economics*, 37(6), 2221–2230.

- Bryson, A., Dorsett, R., & Purdon, S. (2002). The use of propensity score matching in the evaluation of active labour market policies. Working Paper Number 4. Department for Work and Pensions.
- Byerlee, D. R., Spielman, D. J., Alemu, D., & Gautam, M. (2007). Policies to promote cereal intensification in Ethiopia: A review of evidence and experience. <http://ageconsearch.umn.edu/handle/42406> [accessed on 19 Dec 2012]
- Byers, A. L., Allore, H., Gill, T. M., & Peduzzi, P. N. (2003). Application of negative binomial modeling for discrete outcomes: A case study in aging research. *Journal of Clinical Epidemiology*, *56*(6), 559–564.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31–72.
- Cameron, A. C., & Trivedi, P. K. (1986). Econometric models based on count data. Comparisons and applications of some estimators and tests. *Journal of Applied Econometrics*, *1*(1), 29–53.
- Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics using stata* (Vol. 5). College Station, TX: Stata Press.
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (No. 53). Cambridge university press.
- Carlsson, F. (2010). Design of stated preference surveys: Is there more to learn from behavioral economics? *Environmental and Resource Economics*, *46*(2), 167–177.
- Carlsson, F., Kataria, M., Krupnick, A., Lampi, E., Löfgren, A., Qin, P., & Sterner, T. (2013). The truth, the whole truth, and nothing but the truth—A multiple country test of an oath script. *Journal of Economic Behavior & Organization*, *89*, 105–121.
- Carlsson, F., Köhlin, G., Mekonnen, A., & Yesuf, M. (2005). *Are Agricultural Extension Packages What Ethiopian Farmers Want? A Stated Preference Analysis* (Working Papers in Economics No. 172). Göteborg University, Department of Economics.
- Carlsson, F., Martinsson, P., & Akay, A. (2011). The effect of power outages and cheap talk on willingness to pay to reduce outages. *Energy Economics*, *33*(5), 790–798.
- Carson, R. T., & Groves, T. (2011). 15 Incentive and information properties of preference questions: commentary and extensions. *The International Handbook on Non-Market Environmental Valuation*, 300.
- Cavatassi, R., Lipper, L., & Narloch, U. (2011). Modern variety adoption and risk management in drought prone areas: insights from the sorghum farmers of eastern Ethiopia. *Agricultural Economics*, *42*(3), 279–292.
- Cavatassi, R., Salazar, L., González-Flores, M., & Winters, P. (2011). How do Agricultural Programmes Alter Crop Production? Evidence from Ecuador. *Journal of Agricultural Economics*, *62*(2), 403–428. doi:10.1111/j.1477-9552.2010.00279.x
- Cerroni, S., Notaro, S., Raffaelli, R., & Shaw, W. D. (2014). Subjective Risks, Scientific Information, and Food Choices: A Test of Scenario Adjustment in Hypothetical Choice Experiments. *Scientific Information, and Food Choices: A Test of Scenario Adjustment in Hypothetical Choice Experiments* (March 7, 2014).
- Chapman, R. G., & Staelin, R. (1982). Exploiting rank ordered choice set data within the stochastic utility model. *Journal of Marketing Research*, *19*(3), 288–301.
- Coast, J., Al-Janabi, H., Sutton, E. J., Horrocks, S. A., Vosper, A. J., Swancutt, D. R., & Flynn, T. N. (2012). Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Economics*, *21*(6), 730–741.

- Cohen, M. J., & Lemma, M. (2011). Agricultural Extension Services and Gender Equality. An Institutional Analysis of Four Districts in Ethiopia. Development Strategy and Governance Division, International Food Policy Research Institute. Ethiopia Strategy Support Program II (ESSP II), ESSP II Working Paper 28.
<http://www.ifpri.org/sites/default/files/publications/ifpridp01094.pdf> [accessed on 24 Feb 2013]
- Cummings, R. G., & Taylor, L. O. (1999). Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. *American Economic Review*, 89(3), 649–665.
- Cunguara, B., & Darnhofer, I. (2011). Assessing the impact of improved agricultural technologies on household income in rural Mozambique. *Food Policy*, 36(3), 378–390.
- Cuong, N. V. (2013). Which covariates should be controlled in propensity score matching? Evidence from a simulation study. *Statistica Neerlandica*, 67(2), 169–180.
- Das, S., Birol, E., & Bhattacharya, R. N. (2010). 10. Informing efficient solid waste management to improve local environmental quality and public health in West Bengal, India. *Choice Experiments in Developing Countries: Implementation, Challenges and Policy Implications*, 171.
- Davis. (2008). Extension in sub-Saharan Africa: Overview and assessment of past and current models and future prospects. *Journal of International Agricultural and Extension Education*, 15(3), 15–28.
- Davis, K., & Nkonya, E. (2008). Developing a methodology for assessing the impact of farmer field schools in East Africa. In *Proceedings of the 24th Annual Meeting of the Association for International Agricultural and Extension Education (AIAEE)* (pp. 93–99).
- Davis, K., Nkonya, E., Kato, E., Mekonnen, D. A., Odendo, M., Miiro, R., & Nkuba, J. (2012). Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa. *World Development*, 40(2), 402–413. doi:10.1016/j.worlddev.2011.05.019
- Davis, K., Swanson, B., & Amudavi, D. (2009). Review and Recommendations for Strengthening the Agricultural Extension System in Ethiopia. Final Report for the Bill and Melinda Gates Foundation and the Government of Ethiopia. Addis Ababa: IFPRI.
- Dehejia, R. H., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94(448), 1053–1062.
- Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics*, 84(1), 151–161.
- Deininger, K., Ali, D. A., Holden, S., & Zevenbergen, J. (2008). Rural land certification in Ethiopia: Process, initial impact, and implications for other African countries. *World Development*, 36(10), 1786–1812.
- de-Magistris, T., Gracia, A., & Nayga, R. M. (2013). On the Use of Honesty Priming Tasks to Mitigate Hypothetical Bias in Choice Experiments. *American Journal of Agricultural Economics*, 95(5), 1136–1154.
- de Magistris, T., & Pascucci, S. (2012). Consumers’ preferences for the millennium bugs. Does “solemn oath” mitigate the hypothetical bias in choice experiment?. In *2012 Annual Meeting, August 12-14, 2012, Seattle, Washington* (No. 124834). Agricultural and Applied Economics Association.
- de-Magistris, T., & Pascucci, S. (2014). The effect of the solemn oath script in hypothetical choice experiment survey: A pilot study. *Economics Letters*, 123(2), 252–255.

- Dercon, S., De Weerd, J., Bold, T., & Pankhurst, A. (2006). Group-based funeral insurance in Ethiopia and Tanzania. *World Development*, 34(4), 685–703.
- Dercon, S., Gilligan, D. O., Hoddinott, J., & Woldehanna, T. (2009). The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages. *American Journal of Agricultural Economics*, 91(4), 1007–1021. doi:10.1111/j.1467-8276.2009.01325.x
- Deschamps-Laporte, J.-P. (2013). The impact of extension services on farming households in Western Kenya. <http://www.oru.se/PageFiles/63931/WP%205%202013%20Revised.pdf> [accessed on 2 April 2014]
- Devendra, C. (2007). Small farm systems to feed hungry Asia. *Outlook on Agriculture*, 36(1), 7–20. doi:10.5367/000000007780223641
- Dillon, A. (2011). Do Differences in the Scale of Irrigation Projects Generate Different Impacts on Poverty and Production? *Journal of Agricultural Economics*, 62(2), 474–492. doi:10.1111/j.1477-9552.2010.00276.x
- Dinar, A. (1996). Extension commercialization: how much to charge for extension services. *American Journal of Agricultural Economics*, 78(1), 1–12.
- D'souza, G., & Ikerd, J. (1996). Small farms and sustainable development: is small more sustainable? *Journal of Agricultural and Applied Economics*, 28(1), 73–83.
- Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of Development Economics*, 4, 3895–3962.
- Efa, N., Gorman, M., & Phelan, J. (2005). Implications of an extension package approach for farmers' indigenous knowledge: The Maize Extension Package in south-western Ethiopia. *Journal of International Agricultural and Extension Education*, 12(3), 67–78.
- El Benni, N., & Finger, R. (2013). The effect of agricultural policy reforms on income inequality in Swiss agriculture—An analysis for valley, hill and mountain regions. *Journal of Policy Modeling*, 35(4), 638–651.
- Ellis, F. (1999). *Rural livelihood diversity in developing countries: evidence and policy implications*. London: Overseas Development Institute.
- Emana, B., & Hadera, G. (2007). Constraints and opportunities of horticulture production and marketing in eastern Ethiopia. *Dry Land Coordination Group (DCG) Report*, 46.
- Ethiopian Economic Association. (2006). *Evaluation of the Ethiopian Agricultural Extension with Particular Emphasis on the Participatory Demonstration and Training Extension System (PADETES)*. Ethiopian Economic Association / Ethiopian Economic Policy Research Institute.
- FAO. (2014). *FAO: Responsible agricultural investments. Principles for Responsible Investment in Agriculture and Food Systems*. Prepared by the Committee on World Food Security. Forty-first Session “Making a Difference in Food Security and Nutrition”. Rome, Italy, 13-18 October 2014. Retrieved November 15, 2014, <http://www.fao.org/cfs/cfs-home/resaginv/en/> [accessed on 7 Sep 2014]
- Farrington, J. (1994). Public sector agricultural extension: Is there life after structural adjustment. *Natural Resource Perspectives*, No. 2, November 1994. ODI Incorporated. <http://www.odi.org.uk/resources/docs/2987.pdf> [accessed on 14 Jan 2012]
- Feyisa, T. H., & Aune, J. B. (2003). Khat expansion in the Ethiopian highlands: Effects on the farming system in Habro district. *Mountain Research and Development*, 23(2), 185–189.
- Flynn, T. N., Louviere, J. J., Peters, T. J., & Coast, J. (2007). Best–worst scaling: what it can do for health care research and how to do it. *Journal of Health Economics*, 26(1), 171–189.

- Gebissa, E. (2010). Khat in the Horn of Africa: Historical perspectives and current trends. *Journal of Ethnopharmacology*, 132(3), 607–614.
- Gebremedhin, B., Hoekstra, D., & Tegegne, A. (2006). Commercialization of Ethiopian agriculture: Extension service from input supplier to knowledge broker and facilitator. <http://cgspace.cgiar.org/handle/10568/564> [accessed on 23 Nov 2011]
- Gebremedhin, B., Jaleta, M., & Hoekstra, D. (2009). Smallholders, institutional services, and commercial transformation in Ethiopia. *Agricultural Economics*, 40(s1), 773–787.
- Godtland, E. M., Sadoulet, E., De Janvry, A., Murgai, R., & Ortiz, O. (2004). The impact of farmer field schools on knowledge and productivity: a study of potato farmers in the Peruvian Andes. *Economic Development and Cultural Change*, 53(1), 63–92.
- Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological Theory*, 1(1), 201–233.
- Greene, W. (2008). Functional forms for the negative binomial model for count data. *Economics Letters*, 99(3), 585–590.
- Green, R. E., Cornell, S. J., Scharlemann, J. P. W., & Balmford, A. (2005). Farming and the Fate of Wild Nature. *Science*, 307(5709), 550–555. doi:10.1126/science.1106049
- Haji, J. (2007). Production efficiency of smallholders' vegetable-dominated mixed farming system in eastern Ethiopia: A non-parametric approach. *Journal of African Economies*, 16(1), 1–27.
- Hanley, N., Wright, R. E., & Adamowicz, V. (1998). Using choice experiments to value the environment. *Environmental and Resource Economics*, 11(3), 413–428.
- Harttgen, K., & Vollmer, S. (2013). Using an asset index to simulate household income. *Economics Letters*, 121(2), 257–262.
- Hazell, P., Poulton, C., Wiggins, S., & Dorward, A. (2010). The future of small farms: Trajectories and policy priorities. *World Development*, 38(10), 1349–1361.
- Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, 65(2), 261–294.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4), 605–654.
- Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of Labor Economics*, 3, 1865–2097.
- Heckman, J. J., & Robb, R. (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics*, 30(1), 239–267.
- Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). *Applied Choice Analysis: A Primer*. Cambridge University Press.
- Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.
- Holden, S. T., Deininger, K., & Ghebru, H. (2009). Impacts of low-cost land certification on investment and productivity. *American Journal of Agricultural Economics*, 91(2), 359–373.
- Holden, S. T., Deininger, K., & Ghebru, H. (2011). Tenure insecurity, gender, low-cost land certification and land rental market participation in Ethiopia. *The Journal of Development Studies*, 47(1), 31–47.
- Hoyos, D. (2010). The state of the art of environmental valuation with discrete choice experiments. *Ecological Economics*, 69(8), 1595–1603.

- IFPRI. (2010). In-depth assessment of the public agricultural extension system of Ethiopia and recommendations for improvement (No. 1041). International Food Policy Research Institute (IFPRI).
- Imbens, G. M., & Wooldridge, J. M. (2008). *Recent developments in the econometrics of program evaluation*. National Bureau of Economic Research. <http://www.nber.org/papers/w14251> [accessed on 26 Feb 2013]
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, 86(1), 4–29.
- Jacquemet, N., James, A. G., Luchini, S., & Shogren, J. F. (2011). Social psychology and environmental economics: a new look at ex ante corrections of biased preference evaluation. *Environmental and Resource Economics*, 48(3), 413–433.
- Jacquemet, N., James, A., Luchini, S., & Shogren, J. F. (2010). Referendum under Oath. Available at SSRN 1805162. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1805162 [accessed on 9 March 2013]
- Jacquemet, N., Joule, R. V., Luchini, S., & Shogren, J. (2009). *Eliciting preferences under oath*. CES Working paper 09 (43).
- Jacquemet, N., Joule, R.-V., Luchini, S., & Shogren, J. F. (2013). Preference elicitation under oath. *Journal of Environmental Economics and Management*, 65(1), 110–132.
- Jensen, L. P., Picozzi, K., de Almeida, O. da C. M., da Costa, M. de J., Spyckerelle, L., & Erskine, W. (2014). Social relationships impact adoption of agricultural technologies: the case of food crop varieties in Timor-Leste. *Food Security*, 6(3), 397–409.
- Kalinda, T. H., Shute, J. C., & Filson, G. C. (1998). Access to agricultural extension, credit and markets among small-scale farmers in southern Zambia. *Development Southern Africa*, 15(4), 589–608.
- Kassie, G. T., Abdulai, A., & Wollny, C. (2010). Implicit Prices of Indigenous Bull Traits in Crop-Livestock Mixed Production Systems of Ethiopia*. *African Development Review*, 22(4), 482–494.
- Kessels, R., Goos, P., & Vandebroek, M. (2006). A comparison of criteria to design efficient choice experiments. *Journal of Marketing Research*, 43(3), 409–419.
- Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2010). *Handbook on impact evaluation: quantitative methods and practices*. World Bank Publications.
- Kjaer, T., Bech, M., Gyrd-Hansen, D., & Hart-Hansen, K. (2006). Ordering effect and price sensitivity in discrete choice experiments: need we worry? *Health Economics*, 15(11), 1217–1228.
- Krishnan, P., & Patnam, M. (2012). Neighbours and Extension Agents in Ethiopia: Who matters more for technology diffusion? <http://www.theigc.org/publications/working-paper/neighbours-and-extension-agents-ethiopia> [accessed on 10 Dec 2013]
- Krishnan, P., & Patnam, M. (2014). Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption? *American Journal of Agricultural Economics*, 96(1), 308–327.
- Kruk, M. E., Johnson, J. C., Gyakobo, M., Agyei-Baffour, P., Asabir, K., Kotha, S. R., ... Dzodzomenyo, M. (2010). Rural practice preferences among medical students in Ghana: a discrete choice experiment. *Bulletin of the World Health Organization*, 88(5), 333–341.
- Kruk, M. E., Paczkowski, M., Mbaruku, G., de Pinho, H., & Galea, S. (2009). Women's preferences for place of delivery in rural Tanzania: a population-based discrete choice experiment. *Am J Public Health*, 99(9), 1666–72.

- Lancsar, E., Louviere, J., Donaldson, C., Currie, G., & Burgess, L. (2013). Best worst discrete choice experiments in health: Methods and an application. *Social Science & Medicine*, 76, 74–82.
- Larsen, A. F., & Lilleør, H. B. (2014). Beyond the Field: The Impact of Farmer Field Schools on Food Security and Poverty Alleviation. *World Development*, 64, 843–859.
- Lechner, M. (1999). Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany After Unification. *Journal of Business & Economic Statistics*, 17(1), 74–90.
- Leeuwis, C. (2008). *Communication for rural innovation: Rethinking agricultural extension*. John Wiley & Sons.
- Lee, W.-S. (2013). Propensity score matching and variations on the balancing test. *Empirical Economics*, 44(1), 47–80.
- LeMay-Boucher, P. (2009). Beninese and Ethiopian Informal Insurance Groups: A Comparative Analysis. *Development Policy Review*, 27(3), 333–347.
- Lemma, M. (2007). *The Agricultural Knowledge System in Tigray, Ethiopia*. Margraf Publishers, Weikersheim.
- Lemma, T., Sehai, E., & Hoekstra, D. (2011). Status and capacity of Farmer Training Centers (FTCs) in the Improving Productivity and Market Success (IPMS) Pilot Learning Woredas (PLWs). *International Livestock Research Institute (ILRI), Addis Ababa, Ethiopia*.
- Leones, J. P., & Rozelle, S. (1991). Rural Household Data Collection in Developing Countries: Designing Instruments and Methods for Collecting Off-Farm Income Data. No. 128160, *Working Papers in Agricultural Economics*. Cornell University.
- Leuven, E., & Sianesi, B. (2003). Psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *Statistical Software Components S*, 432001.
- List, J. A. (2001). Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sports cards. *American Economic Review*, 91(5), 1498–1507.
- List, J. A., & Gallet, C. A. (2001). What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental and Resource Economics*, 20(3), 241–254.
- Loomis, J. B. (2014). Strategies for Overcoming Hypothetical Bias in Stated Preference Surveys. *Journal of Agricultural and Resource Economics*, 39(1), 34–46.
- Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: analysis and applications*. Cambridge University Press.
- Louviere, J. J., & Islam, T. (2008). A comparison of importance weights and willingness-to-pay measures derived from choice-based conjoint, constant sum scales and best–worst scaling. *Journal of Business Research*, 61(9), 903–911.
- Lusk, J. L. (2003). Effects of cheap talk on consumer willingness-to-pay for golden rice. *American Journal of Agricultural Economics*, 85(4), 840–856.
- Lusk, J. L., & Schroeder, T. C. (2004). Are choice experiments incentive compatible? A test with quality differentiated beef steaks. *American Journal of Agricultural Economics*, 86(2), 467–482.
- Maertens, A., & Barrett, C. B. (2013). Measuring Social Networks' Effects on Agricultural Technology Adoption. *American Journal of Agricultural Economics*, 95(2), 353–359.

- Mangham, L. J., Hanson, K., & McPake, B. (2009). How to do (or not to do) ... Designing a discrete choice experiment for application in a low-income country. *Health Policy and Planning*, 24(2), 151–158. doi:10.1093/heapol/czn047
- Marley, A. A., & Louviere, J. J. (2005). Some probabilistic models of best, worst, and best–worst choices. *Journal of Mathematical Psychology*, 49(6), 464–480.
- Marshall, R. W. (2012). The Impact of the Extension Service on Minority-Owned Small Farm Operations. *Journal of Extension*, 50(1), 1–5.
- McIntosh, C., Sarris, A., & Papadopoulos, F. (2013). Productivity, credit, risk, and the demand for weather index insurance in smallholder agriculture in Ethiopia. *Agricultural Economics*, 44(4-5), 399–417.
- Meenakshi, J. V., Banerji, A., Manyong, V., Tomlins, K., Mittal, N., & Hamukwala, P. (2012). Using a discrete choice experiment to elicit the demand for a nutritious food: Willingness-to-pay for orange maize in rural Zambia. *Journal of Health Economics*, 31(1), 62–71. doi:10.1016/j.jhealeco.2012.01.002
- Mekonne, A., Yesuf, M., & Carlsson, F. (2010). "Farmers' Choice between Public Goods and Agricultural Extension Packages in Ethiopia: A Stated Preference Analysis". In *Choice Experiments in Developing Countries: Implementation, Challenges and Policy Implications*. Cheltenham, UK: Edward Elgar Publishing.
- Meyerhoff, J., & Liebe, U. (2009). Status quo effect in choice experiments: Empirical evidence on attitudes and choice task complexity. *Land Economics*, 85(3), 515–528.
- Mogues, T., Cohen, M. J., Birner, R., Lemma, M., Randriamamonjy, J., Tadesse, F., & Paulos, Z. (2009a). Agricultural extension in Ethiopia through a gender and governance lens. International Food Policy Research Institute ESSP2 Discussion Paper 007. <http://www.ifpri.org/publication/agricultural-extension-ethiopia-through-gender-and-governance-lens> [accessed on 5 Dec 2012]
- Moser, R., Raffaelli, R., & Notaro, S. (2014). Testing hypothetical bias with a real choice experiment using respondents' own money. *European Review of Agricultural Economics*, 41(1), 25–46. doi: 10.1093/erae/jbt016
- Mulatu, E., & Belete, K. (2001). Participatory varietal selection in lowland sorghum in eastern Ethiopia: impact on adoption and genetic diversity. *Experimental Agriculture*, 37(02), 211–229.
- Mulatu, E., & Kassa, H. (2001). Evolution of smallholder mixed farming systems in the Harar Highlands of Ethiopia: The shift towards trees and shrubs. *Journal of Sustainable Agriculture*, 18(4), 81–112.
- Murphy, J. J., Allen, P. G., Stevens, T. H., & Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics*, 30(3), 313–325.
- Mutuc, M., Rejesus, R. M., & Yorobe Jr, J. M. (2013). Which farmers benefit the most from Bt corn adoption? Estimating heterogeneity effects in the Philippines. *Agricultural Economics*, 44(2), 231–239.
- Nayga, R. M., Woodward, R., & Aiew, W. (2006). Willingness to pay for reduced risk of foodborne illness: a nonhypothetical field experiment. *Canadian Journal of Agricultural Economics/Revue Canadienne D'agroeconomie*, 54(4), 461–475.
- Noltze, M., Schwarze, S., & Qaim, M. (2013). Impacts of natural resource management technologies on agricultural yield and household income: The system of rice intensification in Timor Leste. *Ecological Economics*, 85, 59–68.

- Nyariki, D. (2009). Household Data Collection for Socio-Economic Research in Agriculture: Approaches and Challenges in Developing Countries. *J. Soc Sci*, 19(2), 91–99.
- Olajide, B. R. (2011). Assessment of farmers' access to agricultural information on selected food crops in Iddo District of Oyo State, Nigeria. *Journal of Agricultural & Food Information*, 12(3-4), 354–363.
- Owusu-Sekyere, E., Owusu, V., & Jordaan, H. (2014). Consumer preferences and willingness to pay for beef food safety assurance labels in the Kumasi Metropolis and Sunyani Municipality of Ghana. *Food Control*, 46, 152–159.
- Ozor, N., Garforth, C. J., & Madukwe, M. C. (2011). Farmers' Willingness to Pay for Agricultural Extension Service: Evidence from Nigeria. *Journal of International Development*, 25(3), 382–393. doi:10.1002/jid.1849
- Pankhurst, A., & Mariam, D. H. (2000). The iddir in Ethiopia: historical development, social function, and potential role in HIV/AIDS prevention and control. *Northeast African Studies*, 7(2), 35–57.
- Probst, L., Houedjofonon, E., Ayerakwa, H. M., & Haas, R. (2012). Will they buy it? The potential for marketing organic vegetables in the food vending sector to strengthen vegetable safety: A choice experiment study in three West African cities. *Food Policy*, 37(3), 296–308. doi:10.1016/j.foodpol.2012.02.014
- Rauniyar, G. P., & Goode, F. M. (1992). Technology adoption on small farms. *World Development*, 20(2), 275–282.
- Ravallion, M., & Chen, S. (2005). Hidden impact? Household saving in response to a poor-area development project. *Journal of Public Economics*, 89(11), 2183–2204.
- Reardon, T. (1997). Using evidence of household income diversification to inform study of the rural nonfarm labor market in Africa. *World Development*, 25(5), 735–747.
- Rehman, F., Muhammad, S., Ashraf, I., Ch, K. M., & Ruby, T. (2013). Effect of Farmers' Socioeconomic Characteristics on Access to Agricultural Information: Empirical Evidence from Pakistan. *Young (Up to 35)*, 52, 21–67.
- Revelt, D., & Train, K. (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics and Statistics*, 80(4), 647–657.
- Rivera, W. M., & Alex, G. (2004). Extension system reform and the challenges ahead. *The Journal of Agricultural Education and Extension*, 10(1), 23–36.
- Rivera, W. M., & Alex, G. (2005). Extension reform: The challenges ahead. In *Proceedings of the Association for International Agricultural and Extension Education Annual Meeting, San Antonio, Tex., USA, May* (pp. 25-31).
- Rocha, C., Burlandy, L., & Maluf, R. (2012). Small farms and sustainable rural development for food security: The Brazilian experience. *Development Southern Africa*, 29(4), 519–529.
- Rose, J. M. (2014). Interpreting discrete choice models based on best-worst data: a matter of framing. In *Transportation Research Board 93rd Annual Meeting* (No. 14-5072).
- Rose, J. M., & Bliemer, M. C. (2008). *Stated preference experimental design strategies*. Elsevier.
- Rose, J. M., & Bliemer, M. C. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, 29(5), 587–617.
- Rosenbaum, P. R., & 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., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33–38.
- Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2), 135–146.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology; Journal of Educational Psychology*, 66(5), 688–701.
- Salvati, L., & Carlucci, M. (2013). The impact of mediterranean land degradation on agricultural income: A short-term scenario. *Land Use Policy*, 32, 302–308.
- Sandor, Z., & Wedel, M. (2001). Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, 38(4), 430–444.
- Scarpa, R., Notaro, S., Louviere, J., & Raffaelli, R. (2011). Exploring scale effects of best/worst rank ordered choice data to estimate benefits of tourism in alpine grazing commons. *American Journal of Agricultural Economics*, 93(3), 813–828.
- Scarpa, R., Thiene, M., & Train, K. (2008). Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps. *American Journal of Agricultural Economics*, 90(4), 994–1010.
- Selassie, G. G., & Kountouris, Y. (2010). 4. Fishing permit price and wetland conservation: a choice experiment on the value of improved environmental quality of Lake Awassa, Ethiopia. In *Choice Experiments in Developing Countries: Implementation, Challenges and Policy Implications*, 50.
- Severini, S., & Tantari, A. (2013). The impact of agricultural policy on farm income concentration: the case of regional implementation of the CAP direct payments in Italy. *Agricultural Economics*, 44(3), 275–286.
- Shiferaw, B., Kassie, M., Jaleta, M., & Yirga, C. (2013). Adoption of improved wheat varieties and impacts on household food security in Ethiopia. *Food Policy*, 44, 272–284.
- Sianesi, B. (2004). An evaluation of the Swedish system of active labor market programs in the 1990s. *Review of Economics and Statistics*, 86(1), 133–155.
- Smith, J. (2000). *A critical survey of empirical methods for evaluating active labor market policies*. Research Report, Department of Economics, University of Western Ontario.
- Spielman, Dawit Kelemwork, & Dawit Alemu. (2011). Seed, fertilizer, and agricultural extension in Ethiopia. Ethiopian Strategy Support Program II (ESSP II). ESSP II Working Paper 020. March 2011. International Food Policy Research Institute.
- Spielman, D. J., Byerlee, D., Alemu, D., & Kelemework, D. (2010). Policies to promote cereal intensification in Ethiopia: The search for appropriate public and private roles. *Food Policy*, 35(3), 185–194. doi:10.1016/j.foodpol.2009.12.002
- Stachtiaris, S., Drichoutis, A., & Klonaris, S. (2011). The “more is less” phenomenon in Contingent and Inferred valuation. MPRA Paper No. 29456. March 2011.
- Stevens, T. H., Tabatabaei, M., & Lass, D. (2013). Oaths and hypothetical bias. *Journal of Environmental Management*, 127, 135–141.
- Street, D. J., & Burgess, L. (2007). *The construction of optimal stated choice experiments: Theory and methods* (Vol. 647). John Wiley & Sons.
- Swait, J., & Louviere, J. (1993). The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, 30(3), 305–314.

- Swanson, B. E. (2008). *Global review of good agricultural extension and advisory service practices*. Food and Agriculture Organization of the United Nations. Rome.
- Swanson, B. E., & Rajalahti, R. (2010). *Strengthening Agricultural Extension and Advisory Systems: Procedures for Assessing, Transforming, and Evaluating Extension Systems*. The International Bank for Reconstruction and Development/The World Bank.
- Taffesse, A. S., Dorosh, P., & Asrat, S. (2011). Crop Production in Ethiopia: Regional Patterns and Trends. Ethiopian Strategy Support Program II (ESSP II). ESSP II Working Paper No. 0016. March 2011. International Food Policy Research Institute.
- Tarfasa, S., & Brouwer, R. (2013). Estimation of the public benefits of urban water supply improvements in Ethiopia: a choice experiment. *Applied Economics*, 45(9), 1099–1108. doi:10.1080/00036846.2011.613793
- Tefera, T. L., Kirsten, J. F., & Perret, S. (2003). Market Incentives, Farmers' Response and a Policy Dilemma: A Case Study of Chat Production in the Eastern Ethiopian Highlands. *Agrekon*, 42(3), 213–227.
- Tefera, T. L., Sehai, E., & Hoekstra, D. (2011). Status and Capacity of Farmer Training Centers (FTCs) in the Improving Productivity and Market Success (IPMS) Pilot Learning Woredas (PLWs). *International Livestock Research Institute (ILRI), Addis Ababa, Ethiopia*.
- Teshome, G., Kassa, B., Emanu, B., & Haji, J. (2013). Determinants of Rural Household Savings in Ethiopia: The Case of East Hararghe Zone, Oromia Regional State. *Journal of Economics and Sustainable Development*, 4(3), 66–75.
- Thangata, P., & Mequaninte, T. (2011). The Impact of Agricultural Extension Services on Farm Household Efficiency in Ethiopia. <http://addis2011.ifpri.info/files/2011/10/The-Impact-of-Agricultural-Extension-Services-on-Farm-Household-Efficiency-in-Ethiopia.pdf> [accessed on 7 June 2012]
- Todo, Y., & Takahashi, R. (2013). Impact Of Farmer Field Schools On Agricultural Income And Skills: Evidence From An Aid-Funded Project In Rural Ethiopia. *Journal of International Development*, 25(3), 362-381.
- Tonsor, G. T., & Shupp, R. S. (2011). Cheap Talk Scripts and Online Choice Experiments: "Looking Beyond the Mean." *American Journal of Agricultural Economics*, 93(4), 1015–1031.
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.
- Tripp, R., Wijeratne, M., & Piyadasa, V. H. (2005). What should we expect from farmer field schools? A Sri Lanka case study. *World Development*, 33(10), 1705–1720.
- Vermeulen, B., Goos, P., & Vandebroek, M. (2010). Obtaining more information from conjoint experiments by best–worst choices. *Computational Statistics & Data Analysis*, 54(6), 1426–1433.
- Villalobos, P., & Huenchuleo, C. (2010). Ecosystem service valuation of Ruil (*Nothofagus alessandrii*) forests in central Chile: an application of the choice experiment method. In *Choice Experiments in Developing Countries—Implementation, Challenges and Policy Implications*. Cheltenham: Edward Elgar, 151–167.
- Waddington, H., Snilstveit, B., White, H., & Anderson, J. (2010). The Impact of Agricultural Extension Services. 3ie Synthetic Reviews - SR009 Protocol January 2010. <http://www.3ieimpact.org/media/filer/2012/05/07/009%20Protocol.pdf> [accessed on 12 March 2011]

- Whitehead, J. C., & Cherry, T. L. (2007). Willingness to pay for a green energy program: a comparison of ex-ante and ex-post hypothetical bias mitigation approaches. *Resource and Energy Economics*, 29(4), 247–261.
- Whittington, D., & Adamowicz, W. (2010). *The Use of Hypothetical Baselines in Stated Preference Surveys*. Economy and Environment Program for Southeast Asia (EEPSEA).
- Wiggins, S., Kirsten, J., & Llambí, L. (2010). The future of small farms. *World Development*, 38(10), 1341–1348.
- Winters, P., Maffioli, A., & Salazar, L. (2011). Introduction to the Special Feature: Evaluating the Impact of Agricultural Projects in Developing Countries. *Journal of Agricultural Economics*, 62(2), 393–402. doi:10.1111/j.1477-9552.2011.00296.x
- World Bank. (2007). *World Development Report 2008. Agriculture for Development*. Retrieved from http://siteresources.worldbank.org/INTWDR2008/Resources/WDR_00_book.pdf [accessed on 13 Sep 2012]

Appendices

Appendix 1 FGDs to Test Attributes, Attribute Levels, Payment Vehicle, and Scenario Description

A recommended practice in many discrete choice experiments is conducting focus group discussions (FGDs) especially during the early stages of the design process. Such FGDs usually aim at presenting and discussing proposed attributes and their levels, payment vehicle, and scenario description. With this in mind, at the beginning of our survey, we conducted an enumerator-moderated FGD with 12 representative farmers at the Haramaya District Bureau of Agriculture and Rural Development (BoARD).

We started the FGD by explaining the aim and context of the CEs to the participants. After this, we introduced to the discussants the proposed four attributes describing FTC-based AASs, i.e., advice, training, demonstration, and cost. With the help of visual cards⁵² depicting pictorial representations of the attributes, the discussants evaluated whether the attributes are comprehensive, realistic, and understandable. Furthermore, they were encouraged to imagine and put forward other attributes which they think are left out and should be included in the CEs. Nonetheless, other than indicating that the presented attributes are unambiguous, the discussants did not suggest other attributes. This may be probably because we already presented the attributes which correspond to the main functions of the FTCs.

Following the discussion on the attributes, we asked the discussants to suggest levels for each of the attributes. Since we wanted to know the views of the discussants on this matter, we merely

⁵² These are black-and-white pictures used for the purpose of the FGD only. The understandability of the cards was also discussed in the FGD. The respondents suggested that it would be nice to have colored pictures. During the final survey, therefore, we used colored and laminated cards.

asked them what levels they think are appropriate, rather than showing them our original propositions concerning attribute levels. Consequently, we were able to find out the levels and reasons for them. For the attribute ‘advice’, the discussants wanted to get advice from FTCs for a maximum of three times a year instead of the four visits we had in our proposal. This means, they preferred to have one visit for crop production, one for livestock production and management, and one for NRM. Hence, the final levels for this attribute become 0, 2, and 3.

In relation to the attribute ‘training’, based on the information we obtained from the Zonal and District BoARD, the normal duration of one modular training is three months. This training encompasses crop production, livestock production and management, and NRM. In our CEs, however, instead of the standard three months long modular training which covers crop, livestock, and NRM simultaneously, we presented the discussants three separate training each with one month duration. The result of the discussions shows that this alternative proposal is preferred to the standard one. Hence, the levels for the attribute ‘training’ become 0, 2, and 3 but the duration of each of the training becomes one month instead of the one week duration proposed.

Similar to the attribute ‘advice’, the maximum level suggested for the attribute ‘demonstration’ varied slightly from the one we proposed. Most of the discussants preferred to have one demonstration per crop, livestock, and NRM, making the maximum number of demonstrations three. Hence, the final levels for this attribute become 0, 2, and 3. We also obtained the views of experts (such as Development Agents, Head of the District Extension Department) on the plausibility of the attributes and their levels proposed by the FGD discussants. They confirmed that the suggestions are of practical importance in terms of implementation, given the size and capacity of each FTC.

The final attribute discussed was ‘cost’. To facilitate this discussion, we presented different cards depicting the highest, intermediate, and lowest level attribute combinations to each of the discussants and asked them to provide monetary values individually. The rationale behind asking the discussants individually is to minimize the possibility of distortion of individual preferences due to the existence of some dominant people in the group (i.e., group influences on individual outcomes) and to get individual bids for the different combinations. After getting individual evaluations, the discussants were regrouped and their respective monetary indications were discussed in the FGD, in order to arrive at the upper-bound for the cost attribute. Using **Card 4** (i.e., the combination depicting the highest level of advice, training, and demonstration), the upper-bound for the cost attributed agreed up on by the discussants is about 150 birr/year. Likewise, on the basis of **Card 5**, the discussants agreed to attach roughly half of the money they suggested as an upper-bound for an intermediate level combination of these attributes.⁵³ Therefore, in a four relatively equi-distant levels, the attribute ‘cost’ takes 0, 50, 100, and 150 (all in birr) as its levels.

Following the discussion on the cost attribute, we discussed about the type of payment vehicle to be used. Initially, two payment mechanisms, i.e., direct cash payments to FTC administration on a yearly basis at the end of the farming season and registration fees at the beginning of the season, were suggested to the group. Although the discussants were also brainstormed about other potentially usable payment vehicles, the unanimous decision of the discussants points towards the use of direct cash payments to FTCs as payment mechanisms.

The final part of the FGD focused on the scenario description to be used in the experiment. First, we discussed the understandability of the respective scenario descriptions prepared for the two

⁵³ See Appendix 3 for the cards.

treatment groups, i.e., OT and CT. We emphasized on the occurrence of hypothetical bias in hypothetical choice experiments, the problems it creates on analysis and results, and the ways to avoid and/or mitigate the bias. In this regard, we discussed the importance of an ex-ante hypothetical bias reduction/elimination mechanism, namely the solemn oath script. Most importantly, since the respondents can be skeptical about taking an oath in a survey, we discussed in detail the reasons for taking an oath and the ways of administering the oath script (e.g., oath taken in public versus individually/in private). Since we highlighted that the response of each respondent will be kept in secret and will be used only for the purposes of the study, most of the discussants were convinced about not only taking an oath, but also signing on the oath form. We suspect that the use of moderators selected from within the community has also had a positive effect on the discussants commitment and agreement.

The other issue concerning scenario description is the level of effort, i.e., cognitive burden, required to understand hypothetical trade-offs. In particular, we discussed why we intended to use hypothetical baselines against which trade-offs are to be made. Moreover, we explained the fact that the respondents have to assume that they will have to pay the stipulated amount of money right during the survey. Moreover, the concept ‘cash income constraint’ was raised in the discussion to test whether or not it is comprehensible by the discussants. Provided that an enumerator or a data collector reminds them about these issues, the discussants mentioned that these can also be overcome in the real data collection process.

The last aspect of scenario description discussed was the best-worst approach to preference elicitation. Since we intended to obtain the most-preferred and least-preferred choices of a respondent per choice set, we asked the FGD participants if this is understandable and practicable. Most of them indicated that identification of extreme values within a card can be

straightforward, given the constraints they face in the choice situation (e.g., cash income and payment on the spot).

Appendix 2 Scenario Description for CEs in AASs in Haramaya District, Ethiopia

The general scenario description presented to the two treatment groups reads as follows.⁵⁴

Farmers' training center (FTC) based provisions of AASs have been organized and financed by the government to provide you with advice, training, and demonstration on crop, livestock, and natural resource management. To ensure the continued provision of these services – in the face of dwindling budgets to undertake public provision – farmers' full participation is required in all activities of the FTCs, including participatory planning and financing. Since the ultimate aim of establishing FTCs is to hand over the management and financing of services to clients of the FTCs, there is a need to investigate farmers' preferences for the different combinations/types of the characteristics comprising AASs.

Now, we are going to illustrate to you the three main characteristics of AASs provided by the FTCs. The first characteristic is the number of visits per year to FTCs to get advice on crop, livestock or NRM. The number of visit ranges from 0 to 3 as you can see in this card (show **Card 1**). The second characteristic is the number of training per year. It refers to participating in training (max. 1 month duration) organized at FTCs to impart knowledge and skills in crop, livestock, and NRM. The number of training ranges from 0 to 3 (show **Card 2** for the options included). The third characteristic is the number of demonstrations per year. It refers to method and result demonstrations organized by FTCs to show the performance of improved technologies

⁵⁴ Since most of the respondents are farmers with low educational background, we have included some instructions to the interviewers on how to proceed with the CEs. Moreover, we translated the scenario description (as well as the whole questionnaire) to *Amharic*.

and practices in crop, livestock, and NRM. It ranges from 0 to 3 (show **Card 3** for the options included).

These different types of services can be provided in several combinations that have different costs. The amount of money (in ETB) has to be paid directly to the FTC administration on a yearly basis at the end of the farming season. An example of a combination depicting the highest level for the three types of services is shown in **Card 4** (show the card to the respondent). There are many different intermediate level combinations with a range of costs (show **Card 5** as an example of such intermediate level combination of characteristics to the respondent). Finally, without financial support by the clients, the AASs will not be provided and the situation will be the one presented in **Card 6** (show the card to the respondent).

You can now express your preferences by choosing different combinations of services in **12** cards similar to this one (show the example choice card depicted in **Card 7**). In each choice card, you have to evaluate three different combinations and choose your most preferred and least preferred options.

Additional Information for the Oath Treatment – Solemn Oath Script

Before you proceed with the experiment, we would like you to imagine and act as if you will have to pay the cost now/today. Moreover, please make your choice considering the cash income constraint you have.

Since telling the truth is an integral part of your religious faith, we would like to ask you the following question:

Would you respond to all the questions in this experiment by telling us the truth, the whole truth, and nothing but the truth? (yes=1, no=0)

If 'yes', show the paper containing the formal oath script (which is adapted from Jacquemet et al., 2013, p. 115). Reassure the respondent that his/her name will not be written down on the paper but the respondent ID will be used instead. Then, request the respondent to sign on the paper to confirm his/her response to the above question.

Haramaya University

Department of Rural Development and Agricultural Extension

SOLEMN OATH

Topic: Choice Experiment to Estimate the Willingness to Pay for AASs in East Hararghe
Zone of Oromia Region, Ethiopia

I, the undersigned (*respondent's ID*), swear upon my honour that, during the whole experiment, I will tell the truth and always provide honest answers.

.....
Place/PA/Village

.....
Signature

...../...../.....
Date

Continue the experiment with respondents who have answered ‘yes’ to the previous question and also signed on the ‘oath’ paper (OT_{signed}) as well as those who responded ‘yes’ but declined to sign on the paper ($OT_{\text{not signed}}$).

We highly appreciate your cooperation in providing genuine responses. There is not a right or wrong answer! We are just interested in your answer! We promise that your responses will be kept strictly confidential and will only be used for the purpose of this study.

Now is your turn. The choice task begins.

Additional Information for the Control Treatment



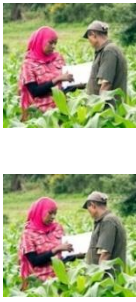

Before you proceed with the experiment, we would like you to imagine and act as if you will have to pay the cost now/today. Moreover, please make your choice considering the cash income constraint you have.

We highly appreciate your cooperation in providing genuine responses. There is not a right or wrong answer! We are just interested in your answer! We promise that your responses will be kept strictly confidential and will only be used for the purpose of this study.


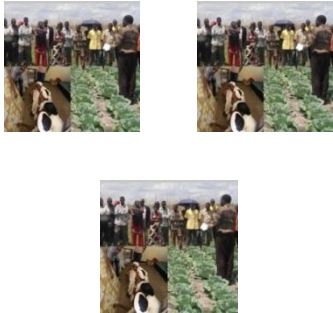


Now is your turn. The choice task begins.

Appendix 3 Illustrative Show Cards Used in the Choice Experiment Survey





Card 1. Advice/Visit

	<p>You will have access to FTCs [...] to receive technical information and advice on crop, livestock, and NRM. The duration of each of your visits will be arranged in an ideal manner so that you will explain the problems you face in your farming enterprise and get potentially useful solutions to solve them.</p>		<p>You will not have access to FTCs to receive technical information and advice on crop, livestock, and NRM.</p>
			
	<p>3 <i>visits/year</i></p>	<p>2 <i>visits/year</i></p>	<p>No <i>visit</i></p>




Card 2. Training

	<p>You will participate in [...] organized at FTCs to impart you the skills in dealing with crop, livestock, and NRM. The duration of each of the training will be arranged in an ideal manner so that you will involve in theoretical as well as practical aspects of farming enterprise.</p>	<p>You will not participate in any training organized at FTCs to impart you the skills in dealing with crop, livestock, and NRM.</p>	
			
	<p>3 <i>training/year</i></p>	<p>2 <i>training/year</i></p>	<p>No <i>training</i></p>




Card 3. Demonstration

	<p>You will participate in [...] demonstrations per year organized by FTCs to show you the performance of improved practices in crop, livestock, and NRM. The duration of each of the demonstrations will be arranged in an ideal manner to help you in your farming enterprise.</p>	<p>You will not participate in any demonstration organized by FTCs to show the performance of improved practices in crop, livestock, and NRM.</p>	
			
	<p>3 <i>demonstrations/year</i></p>	<p>2 <i>demonstrations/year</i></p>	<p>No <i>demonstration</i></p>




Card 4. The Highest Level of Advice/Visit, Training, and Demonstration in FTC-based AASs

		
<p>3 <i>visits/year</i></p>	<p>3 <i>training/year</i></p>	<p>3 <i>demonstrations/year</i></p>










Card 5. An Example of Intermediate Level of Advice/Visit, Training, and Demonstration in FTC-based AASs

		
<p>3 <i>visits/year</i></p>	<p>2 <i>training/year</i></p>	<p>2 <i>demonstrations/year</i></p>

Card 6. The Lowest Level of Participation in Advice/Visit, Training, and Demonstration in FTC-based AASs

		
<p><i>No</i> visit</p>	<p><i>No</i> training</p>	<p><i>No</i> demonstration</p>

Card 7. A Sample Choice Card

Attribute	Alternative 1	Alternative 2	Alternative 3
Number of visits to FTCs/year (advice)			
Number of training/year (max. 1 month duration)			
Number of demonstrations/year			
Cost (in ETB)/year	100	250	0
Choose your most preferred option, please! (mark ✓)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Choose your least preferred option, please! (mark ✓)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Appendix 4 Summary Statistics of Sociodemographic Characteristics of Respondents Trained at FTCs in 2009 (Standard Deviations in Parenthesis)

Characteristics	Treatment groups ^c			CT	K-S test (OT _{signed} vs CT)	K-S test (OT _{total} vs CT)
	OT					
	OT _{total}	OT _{signed}	OT _{not signed}			
Age (years)	37.7 (4.2)	37.8 (4.1)	37.6 (4.4)	39.3 (2.7)	0.643	0.536
Male household head ^a	100.0	100.0	100.0	100.0	1.000	1.000
Education (years)	3.2 (3.5)	3.4 (3.3)	3.0 (3.8)	1.1 (2.4)	0.068	0.035
Household size (number)	7.0 (1.1)	6.8 (1.1)	7.1 (1.1)	7.7 (1.2)	0.135	0.035
Married household head ^a	100.0	100.0	100.0	100.0	1.000	1.000
Access to basic facilities ^a						
Drinking water	96.7	100.0	93.8	100.0	1.000	1.000
Human health center	13.3	7.1	18.8	0.0	1.000	0.958
Daily market	0.0	0.0	0.0	96.7	0.000	0.000
Weakly market	3.3	0.0	6.3	100.0	0.000	0.000
Veterinary	100.0	100.0	100.0	100.0	1.000	1.000
Experience in farming (years)	21.9 (4.9)	21.1 (4.3)	22.6 (5.4)	21.2 (4.1)	0.791	0.291
Experience in AASs (years)	15.7 (4.0)	15.9 (3.8)	15.6 (4.4)	15.2 (3.2)	0.476	0.393
Experience with FTCs (years)	3.5 (0.6)	3.5 (0.5)	3.5 (0.6)	4.0 (0.2)	0.020	0.003
FTC participation (number/year)						
Advice	26.5 (6.4)	26.6 (4.4)	26.4 (7.9)	9.5 (6.8)	0.000	0.000
Training	7.6 (2.5)	7.4 (2.0)	7.8 (2.9)	2.2 (1.2)	0.000	0.000
Demonstration	2.2 (0.9)	2.2 (0.6)	2.3 (1.1)	1.5 (0.6)	0.000	0.000
Evaluation of AASs (satisfaction) ^a						
<i>Advice</i>						
Very satisfied	3.3	0.0	6.3	62.1	0.000	0.000
Satisfied	83.3	78.6	87.5	37.9	0.062	0.003
Not satisfied	13.3	21.4	6.3	0.0	0.689	0.899
<i>Training</i>						
Very satisfied	96.7	100.0	93.8	24.1	0.000	0.000
Satisfied	3.3	0.0	6.2	75.9	0.000	0.000
Not satisfied						
<i>Demonstration</i>						
Very satisfied	9.1	0.0	13.3	27.6	0.674	0.708
Satisfied	4.6	14.3	0.0	65.5	0.070	0.000
Not satisfied	86.3	85.7	86.7	6.9	0.001	0.000
Assets ^b	6,568.4 (6,368.9)	6,914.8 (6,378.6)	6,265.3 (6,553.7)	10,805.0 (7,267.0)	0.005	0.001
Land size (ha)	0.9 (0.4)	0.9 (0.3)	1.0 (0.4)	0.8 (0.2)	0.126	0.129
Land certified ^a	100.0	100.0	100.0	3.3	0.000	0.000
Value (Birr) of livestock possessed ^c	40,112.0 (12,546.2)	39,932.1 (10,622.6)	40,269.4 (14,369.0)	53,120.2 (19,257.0)	0.020	0.003
Farm income (Birr) ^d	19,324.7 (5,842.8)	20,657.7 (4,511.0)	18,158.3 (6,727.2)	29,443.2 (11,191.7)	0.001	0.000
Involvement in off-farm activities ^a	46.7	42.9	50.0	3.3	0.073	0.003
Have access to microfinance	10.0	7.1	12.5	6.7	1.000	1.000
Number of people in one's network	4.5 (1.0)	4.1 (0.8)	4.8 (1.1)	5.6 (1.1)	0.000	0.002
Experienced shock ^a	100.0	100.0	100.0	36.7	0.000	0.000
Sample size	30	14	16	30		

K-S test: Two-sample Kolmogorov-Smirnov test for equality of distribution functions. Exact *p-values* on the combined K-S reported.

^a Proportion/percentage of sample possessing the specific characteristic.

^b Value of assets (Birr) is estimated based on the prevailing market price for the good or materials to produce it. It comprises values of productive assets, household goods, and consumer durables bought or possessed during the last 12 months preceding the survey.

^c Computed based on the estimated value (Birr) of each livestock species had it been sold in the current market.

^d Calculated by summing up crop and livestock incomes (net of intermediate costs associated with the two). (i.e., a. Value of crop sold and total crop residues produced (in ETB) per household during the last 12 months minus cost of intermediate inputs associated with crop production, and b. income from sell of livestock products (Birr) per household net of intermediate costs associated with livestock production and management.)

^e OT = Oath Treatment; OT_{total} = total oath treatment (i.e., those who agreed to tell the truth or signed on the solemn oath form); OT_{signed} = those who agreed to tell the truth and signed on the solemn oath form; OT_{not signed} = those who agreed to tell the truth but did not sign on the solemn oath form; CT = Control/Baseline Treatment (i.e., those who are not exposed to the solemn oath treatment).

Appendix 5 Choice of Hypothetical Baseline as Best, Second-best, and Worst

	OT _{total}	OT _{signed}	OT _{not signed}	CT
<i>Full sample (trained in 2008 & 2009)</i>				
Best	1	0	1	3
Second-best	812	398	414	976
Worst	626	298	328	458
Total	1,439	696	743	1,437
<i>Sub-sample (trained in 2009)</i>				
Best	0	0	0	0
Second-best	366	158	208	486
Worst	354	178	176	234
Total	720	336	384	720

Appendix 6 Changes in the WTP in Relation to Socio-demographic Characteristics (‘Sub-Sample’, Signed on the Solemn Oath Form)

The table below presents the changes in WTPs on the basis of some SDCs for those households trained at FTCs in 2009 and signed on the solemn oath form. From the table, it appears that the WTP to receive advice from FTCs does not vary with SDCs primarily because of the absence of any significant variation in WTP for advice across the respondents in the OT group. The same reasoning also applies to the constant WTP for demonstration in the OT group. Regarding training, however, there are changes in the WTP based on some SDCs. Most importantly, the lowest level of WTP for training is associated with financial capital (i.e., farm income category 10,000 – 20,000 birr/year). Other variables, such as human capital (household size, experience in farming and in AASs), physical capital (assets, land holding size, and value of livestock), and social capital (networks) are also important in describing the lower WTP for training in this treatment group. It is also evident from the table that signing on the solemn oath form helped obtain a lower WTP value for training and demonstration (similar to the observation in Table 3.10 and the discussion that followed it).

WTP Changes with SDCs for those Trained at FTCs in 2009 and those who Signed an Oath Form

	Advice		Training		Demonstration	
	OT _{tot} ^b	CT ^c	OT _{tot}	CT	OT _{tot} ^d	CT ^e
Age (years) ^a : 15 – 64	-176.4	55.5	257.6	277.8	152.9	301.2
Male household head	-176.4	55.5	256.1	277.8	152.9	301.2
Education ^a : No formal education	-176.4	55.5	257.6	277.6	152.9	301.2
Primary school (1 – 4)				286.0		
Primary and lower secondary (5 – 8)				274.6		
Household size (number) ^a : < 5	-176.4		278.0		152.9	
5 – 10		55.5	256.1	277.8		301.2
Experience in farming (years): < 15		55.5		304.3		301.2
15 – 25	-176.4		236.8	277.0	152.9	
25 – 30			266.2	277.0		
> 30			265.5	275.7		
Experience with on-farm AASs (years) ^a : 5 – 15	-176.4	55.5	250.9	275.0	152.9	301.2
15 – 25			258.7	281.6		
Value of household assets (birr): < 5,000		55.5		272.3		301.2
5,000 – 10,000	-176.4		248.4	281.1	152.9	
10,000 – 15,000			282.3	283.0		
15,000 – 20,000				276.3		
> 20,000			259.0	278.1		
Land holding size (ha): < 0.25		55.5		259.6		301.2
0.25 – 0.5	-176.3		282.8		152.9	
0.5 – 0.75			261.9	279.6		
> 0.75			250.4	280.1		
Certified land	-176.4		257.6		152.9	
Value of livestock possessed (birr) ^a : 10,000 – 30,000	-176.3	55.5	264.1	266.4	152.9	301.2
30,000 – 60,000			254.2	277.7		
> 60,000			272.2	280.7		
Farm income (birr) ^a : < 10,000	-176.3		281.2		152.9	
10,000 – 20,000		55.5	228.8	280.9		301.2
20,000 – 30,000			264.0	273.0		
30,000 – 40,000				285.9		
Involved in off-farm activities	-176.4	55.6	257.6	277.3	152.9	301.2
Have access to microfinance credit		55.5		281.5		301.2
Have access to daily market	-176.3	55.5	282.4	278.3	152.9	301.2
Number of people in one's network: < 5	-176.4	55.5	256.1	276.7	152.9	301.2
5 – 10			259.6	276.0		
> 10				323.4		
Number of FTC-based advice per year: < 5		55.5				
5 – 15	-176.3					
15 – 30						
> 30						
Number of FTC-based training per year: < 4				278.2		
4 – 8			256.1	276.1		
> 8			258.2			
Number of FTC-based demonstration per year: < 2						301.2
2 – 4					152.9	
> 4						
Satisfied with current advice provision	-176.4	55.5				
Satisfied with current provision of training			257.6	277.5		
Satisfied with current provision of demonstration					152.9	301.2
Experienced long-term shock	-176.4	55.5	257.6	281.8	152.9	301.2

Note:

OT_{tot} = Oath Treatment; CT = Control Treatment.

^a The following categories (not shown in the table) have no observations up on which to compute the WTPs: age (< 15 and > 64 years); education (9 – 12 and some college/university); household size (> 10 members); experience with on-farm AASs (< 5 years); value of livestock (< 10,000 birr); and, farm income (> 40,000 birr).

^b Except for *education* (1 – 8), *experience in farming* (< 15), *value of household assets* (15,000 – 20,000), *land holding size* (< 0.25), *farm income* (30,000 – 40,000), *access to microfinance*, *number of people in one's network* (> 10), and *number of FTC-based advice* (< 5), all categories of relevant variables have the same WTP for advice in this treatment group (due to the fact that the WTP for advice does not vary much among the individual households in this group).

^c Except for *household size* (< 5), *land holding size* (0.25 – 0.5), *certified land*, and *farm income* (< 10,000), all categories of relevant variables have the same WTP for advice in this treatment group (due to the fact that the WTP for advice does not vary much among the individual households in this group).

^d Except for *education* (1 – 8), *experience in farming* (< 15), *value of household assets* (< 5,000 and 15,000 – 20,000), *land holding size* (< 0.25), *farm income* (30,000 – 40,000), *access to microfinance*, *number of people in one's network* (> 10), and *number of FTC-based demonstration* (< 2 and > 4), all categories of relevant variables have the same WTP for demonstration in this treatment group (due to the fact that the WTP for demonstration does not vary much among the individual households in this group).

^e Except for *household size* (< 5), *land holding size* (0.25 – 0.5), *certified land*, and *farm income* (< 10,000), all categories of relevant variables have the same WTP for advice in this treatment group (due to the fact that the WTP for advice does not vary much among the individual households in this group).

Appendix 7 Focus Group Discussions (FGDs) in the CE in AASs

At the beginning of the field work, a FGD comprising 12 representative farmers was conducted at the Haramaya District Bureau of Agriculture and Rural Development. This enumerator-moderated FGD proceeded as follows.

1. First, the aim and context of the CE was explained.
2. Second, the four attributes were introduced and the discussants were asked whether or not they are realistic and understandable.
 - Moreover, the discussants were encouraged to imagine and put forward other attributes which they think are relevant to be considered.
3. After obtaining the complete list of attributes, the discussants were asked to suggest clear and unambiguous levels for each of the attributes.
 - The suggestions regarding the attributes and their levels were compared with our original proposition, and necessary adjustments were made accordingly, so as to make the CE realistic given the particular context under consideration.
 - During the discussion on attribute levels, a special consideration was taken about the cost attribute.
 - However, since the type of payment vehicle to be used is relevant in relation to the cost attribute, the FGD took a detour here and participants were directed to discuss about possible payment vehicles.
4. To kick-off the discussion on payment vehicles, two payment mechanisms, i.e., direct cash payments to FTC administration on a yearly basis at the end of the farming season and registration fees at the beginning of the season, were suggested to the group. The discussants were also brainstormed about other potentially usable payment vehicles.

5. After identifying a realistic and reasonable payment vehicle, the FGD was redirected to the discussion on the assignment of levels to the cost attribute.
 - Just before embarking on this, nevertheless, the group discussion was paused for a while and, using the visual aids depicting the highest and intermediate level attribute combinations (i.e., Cards 4 and 6), each of the discussants were asked individually⁵⁵ about the amount of money they would like to associate to these formulations.
 - After this, the individuals were regrouped and their respective monetary indications were discussed in the FGD, in order to arrive at both the upper- and intermediate-bounds for the cost attribute.
 - Furthermore, the clarity of the pictorial representations of the attributes and their levels (e.g., the size of the pictures, colored vs black-and-white) were tested in the FGD.
6. Having identified and refined the relevant attributes and their levels in the FGDs, the next step was to conduct key informant interviews with DAs, members of the FTC administration, as well as representatives of regional and district bureaus of agriculture and rural development, in order to collect further insights and opinions about the attributes, attribute levels, and payment vehicle. This exercise helped to ensure that policy-relevant attributes and levels are included in the experiment.
7. The final part of the FGD was on the scenario description used in the experiment. There were three issues put forth here.

⁵⁵ The rationale behind this is to minimize the possibility of distortion of individual preferences due to the existence of some dominant people in the group (i.e., group influences on individual outcomes).

- First, the understandability of the respective scenario descriptions for the control and oath groups was discussed. For example, the reasons for taking an oath and the way of administering it (i.e., public versus private) were mentioned and doubts were clarified.
- Second, the level of effort, i.e., cognitive burden, required to understand hypothetical trade-offs was discussed, in order to ensure task simplicity.
 - A particular emphasis was made on the fact that respondents have to assume that they will have to pay the stipulated amount of money right during the survey.
 - Moreover, the concept ‘cash income constraint’ was raised to test whether or not it is comprehensible by the discussants.
- Finally, whether the respondents understand the best-worst approach to preference elicitation was discussed.

FGDs in the CE in AASs

Place:

Date: / /

<i>Attributes</i>	<i>Realistic?</i> <i>(yes=1, no=0)</i>	<i>Understandable?</i> <i>(yes=1, no=0)</i>	<i>Comments</i>
Advice (Card 1)			
Training (max. 1 month duration) (Card 2)			
Demonstration (Card 3)			
Cost (in ETB) – direct cash payments to FTCs at the end of the farming season			

<i>Attribute combinations</i>	<i>Realistic?</i> <i>(yes=1, no=0)</i>	<i>Understandable?</i> <i>(yes=1, no=0)</i>	<i>Comments</i>
The highest level of advice/visit, training, and demonstration (Card 4)			
An example of intermediate level of advice/visit, training, and demonstration (Card 5)			
The lowest level of participation in advice/visit, training, and demonstration (Card 6)			

<i>Attributes</i>	<i>Cost (in ETB)</i>		<i>Comments</i>
	<i>Minimum</i>	<i>Maximum</i>	
Advice (Card 1) 1 visit/year 2 visits/year 4 visits/year			
Training (max. 1 month duration) (Card 2) 1 training/year 2 training/year 3 training/year			
Demonstration (Card 3) 1 demonstration/year 2 demonstrations/year 4 demonstrations/year			
The highest level of advice/visit, training, and demonstration (Card 4)			
An example of intermediate level of advice/visit, training, and demonstration (Card 5)			
The lowest level of participation in advice/visit, training, and demonstration (Card 6)			

Appendix 8 Household Survey Questionnaire - Treatment (Oath) Group

*Agricultural Advisory Services in Eastern Ethiopia: Access, Impact, and Willingness to Pay*⁵⁶
Muluken G. Wordofa

To the enumerator! Please use the following paragraph to start your conversation with each of the respondents.

“Greetings! My name is ___ (*name of the enumerator*) ___. I am a member of the data collection team for the research which aims to investigate the access to farmer training center (FTC)-based provision of agricultural advisory services (AASs); the impact of FTC-based AASs on household agricultural income; and, the preferences for future improvements of the characteristics comprising AASs at FTCs (i.e., advice, training, and demonstration). The research is conducted jointly by the University of Trento (Italy), Haramaya University (Ethiopia), East Hararghe Zone Bureau of Agriculture and Rural Development (BoARD), and Haramaya District BoARD, and focuses on household heads in East Hararghe Zone of Oromia Region, Ethiopia. Since the data you provide to us is of immense value, we highly appreciate your honest responses to each of the questions included in this questionnaire. We assure you that any information you provide to us, including your personal details, will be kept strictly confidential and will be used only for the purpose of this study. On behalf of the research team, I would like to thank you in advance for your time and active participation in this survey.”

Part 1. Survey identification

1	Region: Oromia	7	Household ID:	
2	Zone: East Hararghe	8	Name of interviewer:	ID:
3	District: Haramaya	9	Date of interview: ____/____/____ (dd/mm/yy)	
4	PA/Kebele:	10	Time interview started:	
5	Village:	11	Time interview ended:	
6	Name of the HH head:			

Part 2. Household characteristics

1. HH member: relationship to the head	2. Age (in years)	3. Sex (<i>male=0, female=1</i>)	4. Education (number of years of schooling)	5. Main occupation* <i>Code (a)</i>	6. Gross monthly income (in ETB)
1. Head					
2. Wife					
3. Son					
4. Daughter					
5. Other (specify					

⁵⁶ This questionnaire is constructed based on the research project documentations obtained (with permission) from the International Food Policy Research Institute (IFPRI). These documents and datasets are: Ethiopian Rural Household Survey (ERHS); Food and Water Security under Global Change – Developing Adaptive Capacity with a Focus on Rural Africa; Policies for Sustainable Land Management in the Ethiopian Highlands dataset 1998-2000; and, Assessing the Potential of Farmer Field Schools (FFS) to Fight Poverty and Foster Innovation in East Africa.

* **Code (a):** farmer=1, domestic worker=2, manual worker=3, tailor=4, weaver=5, craftsworker/potter=6, blacksmith=7, driver/mechanic=8, skilled factory worker=9, teacher=10, health worker=11, public administrator=12, trader=13, student=14, looking for work=15, herding=16, other=17 (specify.....)

7. Additional questions about the HH head

- a. Marital status (single=1, married=2, divorced=3, widowed=4, other=5 (specify.....))
- b. Ethnic group (Oromo=1, Amhara=2, Tigrayan=3, Gurage=4, other=5 (specify))
- c. Languages spoken (Oromiffa=1, Amharic=2, Tigrigna=3, Guragigna=4, other=5 (specify))
- d. Religion (Islam=1, Orthodox =2, Protestant =3, Nonreligious=4, other=5 (specify ...))
- e. Family size (number of people living in the HH including the HH head)
- f. How long have you been involved in farming? (in number of years)

Part 3. Access to basic facilities

3.1 We would like to know the type of facility/service you have access to, the distance from your home, and whether you pay for it.

a. Type of facility/service	b. Do you have access to it? (yes=1, no=0)	c. Since when? [year in E.C.]	d. Distance from home [in km]	e. If you pay for it, how much per month (in ETB)?
piped water				
electricity				
telephone (land line)				
kindergarten				
primary school				
secondary school				
health center (clinic, pharmacy)				
post office				
mosque/church				
bank				
FTC				
NGO office				
daily market				
periodic (at least 1/week) market				
veterinary medicine				
microfinance institution				
others (specify)				

3.2 We would like to know the most important services/investments you would want the government/NGOs to do for you/your community.

a. What services, investments, or developments would you want the government, community, private sector, or NGOs to do for your family/village/PA to be better off? Identify and rank the three most important ones.	
1. Services	2. Rank [1=most important, 5=least important]
1.	
2.	
3.	

3.3 Now, we are moving to one of the most important parts of the survey. We would like to ask you about your participation in FTC-based provision of AASs (i.e., advice, training, and demonstration).

- 1. Overall, how many years have you participated in public AASs (i.e., agricultural extension)?
- 2. How many years have you participated in FTC-based AASs? Give us some details in the following tables.

a. AASs	b. In the last 12 months:			
	1. Have you received ..[...].. from FTCs? (yes=1, no=0)	2. If yes, how many times?	3. Have you received on farm ..[...].. from DAs? (yes=1, no=0)	4. If yes, how many times?
1. advice				
2. training				
3. demonstration				

AASs	c. Evaluation of current services provided			d. In the future, if the current services were to be improved both in quantity and quality, rank your order of preference among the three. [Rank: 3= <i>most preferred</i> , 1= <i>least preferred</i>]	e. In the future, if the improved services were to be paid for, would you be willing to pay for: (yes=1, no=0)
	1. Were the services provided to you on time? (yes=1, no=0)	2. In your opinion, how relevant were the services?*	3. In general, how satisfied are you with the current service provision?***		
1. advice					
2. training					
3. demonstration					

*Code (a): very relevant=1, relevant=2, not relevant=3

**Code (b): very satisfied=1, satisfied=2, not satisfied=3

3. For participants of FTC-based training:

a. What motivated you to join the training?

.....

b. What were your expectations before joining the training?

.....

c. Have you shared the contents of the training with somebody not in the training group? (yes=1, no=0)
 If 'yes', name at least one person

.....

d. In general, how have you been selected into FTC training? (Who decides which HHs in the community would participate in FTC training?) (the development agent=1, PA administration=2, woreda administration=3, FTC administration=4, the community (we all decided together)=5, do not know=6, other=7 (specify))

3.4 Farmer expectations about improved farm outcomes due to participation in FTC-based AASs

a. How much do you expect your farm output and general economic situation will improve in the future because of participation in AASs? [rate using the five-level Likert scale provided]					
Type of AAS	Very much [5]	Somewhat [4]	Undecided [3]	Not really [2]	Not at all [1]
1. advice/visit					
2. training					
3. demonstration					

b. In general, do you think FTC-based AASs will satisfy your needs for improved farm management, decision-making and overall farm performance? [rate using the five-level Likert scale provided]					
Type of AAS	Very much [5]	Somewhat [4]	Undecided [3]	Not really [2]	Not at all [1]
1. advice/visit					
2. training					
3. demonstration					

c. How do you expect life to be one year from now? (improved=1, about the same=2, deteriorated=3)					
d. How do you expect life to be five years from now?(improved=1, about the same=2, deteriorated=3)					
e. Do you have any person(s) whom you aspire to be like? (yes=1, no=0)			f. If 'yes', how many are they?		
g. If 'yes', who are they? (those similar to me in their status=1, from me in their status=3, others=4 (specify))			those above my status=2, those very far away		
h. Do you have any person(s) whose characteristics and behaviors influence yours (i.e., role models) either positively or negatively? (yes=1, no=0)			i. If 'yes', how many are they?		
j. If 'yes', who are they? (those similar to me in their status=1, from me in their status=3, others=4 (specify))			those above my status=2, those very far away		

Part 4. Choice experiment in AASs

4.1 Scenario description

Farmer training center (FTC) based provisions of AASs have been organized and financed by the government to provide you with advice, training, and demonstration on crop, livestock, and natural resource management. To ensure the continued provision of these services – in the face of dwindling budgets to undertake public provision – farmers’ full participation is required in all activities of the FTCs, including participatory planning and financing. Since the ultimate aim of establishing FTCs is to hand over the management and financing of services to clients of the FTCs, there is a need to investigate farmers’ preferences for the different combinations/types of the characteristics comprising AASs.

Now, we are going to illustrate to you the three main characteristics of AASs provided by the FTCs. The first characteristic is the number of visits per year to FTCs to get advice on crop, livestock or NRM. The number of visit ranges from 0 to 3 as you can see in this card (show **Card 1**).

The second characteristic is the number of training per year. It refers to participating in training (max. 1 month duration) organized at FTCs to impart knowledge and skills in crop, livestock, and NRM. The number of training ranges from 0 to 3 (show **Card 2** for the options included).

The third characteristic is the number of demonstrations per year. It refers to method and result demonstrations organized by FTCs to show the performance of improved technologies and practices in crop, livestock, and NRM. It ranges from 0 to 3 (show **Card 3** for the options included).

These different types of services can be provided in several combinations that have different costs. The amount of money (in ETB) has to be paid directly to the FTC administration on a yearly basis at the end of the farming season.

An example of a combination depicting the highest level for the three types of services is shown in **Card 4** (show the card to the respondent).

There are many different intermediate level combinations with a range of costs (show **Card 5** as an example of such intermediate level combination of characteristics to the respondent).

Finally, without financial support by the clients, the AASs will not be provided and the situation will be the one presented in **Card 6** (show the card to the respondent).

You can now express your preferences by choosing different combinations of services in **12** cards similar to this one (show the example choice card depicted in **Card 7**). In each choice card, you have to evaluate three different combinations and choose your most preferred and least preferred options.

4.2 Solemn oath

Before you proceed with the experiment, we would like you to imagine and act as if you will have to pay the cost now/today. Moreover, please make your choice considering the cash income constraint you have.

Since telling the truth is an integral part of your religious faith, we would like to ask you the following question:

*Would you respond to all the questions in this experiment by telling us the truth, the whole truth, and nothing but the truth?
..... (yes=1, no=0)*

If ‘yes’, show the paper containing the formal oath script (which is adapted from Jacquemet et al., 2013:115). Reassure the respondent that his/her name will not be written down on the paper but the respondent ID will be used instead. Then, request the respondent to *sign on the paper to confirm his/her response to the above question.*

Continue the experiment with respondents who have answered ‘yes’ to the previous question and also signed on the ‘oath’ paper as well as those who responded ‘yes’ but declined to sign on the paper.

We highly appreciate your cooperation in providing genuine responses. There is not a right or wrong answer! We are just interested in your answer! We promise that your responses will be kept strictly confidential and will only be used for the purpose of this study.

Haramaya University
Department of Rural Development and Agricultural Extension

SOLEMN OATH

Topic: Choice Experiment to Estimate the Willingness to Pay for AASs in East Hararghe Zone of Oromia Region, Ethiopia

I, the undersigned (*respondent's ID*), swear upon my honour that, during the whole experiment, I will tell the truth and always provide honest answers.

.....
Place/PA/Village

.....
Signature

...../...../.....
Date

Now is your turn. The choice task begins.

Card 1

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	3	0	0
b. number of training (max. 1 month duration) per year	0	3	0
c. number of demonstrations/year	2	2	0
d. cost (in ETB)/year	100	300	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 2

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	0	3	0
b. number of training (max. 1 month duration) per year	3	0	0
c. number of demonstrations/year	2	2	0
d. cost (in ETB)/year	100	300	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 3

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	0	3	0
b. number of training (max. 1 month duration) per year	1	2	0
c. number of demonstrations/year	3	0	0
d. cost (in ETB)/year	300	100	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 4

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	2	2	0
b. number of training (max. 1 month duration) per year	1	2	0
c. number of demonstrations/year	3	0	0
d. cost (in ETB)/year	150	250	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 5

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	2	2	0
b. number of training (max. 1 month duration) per year	2	1	0
c. number of demonstrations/year	0	3	0
d. cost (in ETB)/year	250	100	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 6

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	0	3	0
b. number of training (max. 1 month duration) per year	3	0	0
c. number of demonstrations/year	2	2	0
d. cost (in ETB)/year	200	150	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 7

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	0	3	0
b. number of training (max. 1 month duration) per year	0	3	0
c. number of demonstrations/year	3	0	0
d. cost (in ETB)/year	150	200	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 8

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	3	0	0
b. number of training (max. 1 month duration) per year	0	3	0
c. number of demonstrations/year	2	2	0
d. cost (in ETB)/year	300	100	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 9

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	3	0	0
b. number of training (max. 1 month duration) per year	1	2	0
c. number of demonstrations/year	0	3	0
d. cost (in ETB)/year	250	150	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 10

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	3	0	0
b. number of training (max. 1 month duration) per year	3	0	0
c. number of demonstrations/year	0	3	0
d. cost (in ETB)/year	100	250	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 11

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	2	2	0
b. number of training (max. 1 month duration) per year	2	1	0
c. number of demonstrations/year	0	3	0
d. cost (in ETB)/year	200	150	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Card 12

	alt1	alt2	alt3
a. number of visits to FTCs/year (advice)	2	2	0
b. number of training (max. 1 month duration) per year	2	1	0
c. number of demonstrations/year	3	0	0
d. cost (in ETB)/year	150	200	0
Q1. Which of the three is your most preferred option? (mark ✓)			
Q2. Which of the remaining two is your least preferred option? (mark ✓)			

Part 5. Household asset inventory

5.1 We would like to know if you own any tools (e.g., farm implements), large household goods or valuables. We would also like to know how much it is worth, and when and where you obtained it. (*Probe for items not listed*)

a. Item	b. How many?	c. Original price (ETB)	d. Year purchased (E.C.)	e. Where was this obtained?**** <i>Code (d)</i>	f. If someone were to buy one, similar item (or the materials to make it), how much would the person have to pay? (ETB)
<i>Productive assets*</i> <i>Code (a)</i>					
<i>HH goods**</i> <i>Code (b)</i>					
<i>Consumer durables***</i> <i>Code (c)</i>					

* **Code (a):** plough (*maresha*)=1, sickle (*machid*)=2, pick axe (*doma*)=3, axe (*metrebia*)=4, hoe (*doma/mekotkocha*)=5, spade (*akafa*)=6, traditional beehive=7, modern beehive=8, water pump (hand/foot)=9, water pump (diesel)=10, grain mill (stone)=11, grain mill (diesel)=12, chopper (*konchera, gejera, mencha*)=13, other=14 (specify

****Code (b):** charcoal/wood stove=1, kerosene stove=2, leather/wood bed=3, modern chair=4, modern table=5, metal bed=6, animal cart=7, other=8 (specify

*****Code (c):** mobile telephone=1, radio=2, television=3, jewellery (gold, silver)=4, bicycle=5, wristwatch=6, refrigerator=7, car=8, others=9 (specify

******Code (d):** village market=1, another village=2, local town=3, regional center=4, Addis Ababa=5, other=6 (specify

5.2 During the past 12 months, did you spend anything on building a new house or improving/repairing your house or other building? (**yes=1, no=0**) If yes, give us details in the following table.

a. Part of dwelling	b. How much did it cost in total (ETB)?
New house	
Repair:	
Wall	
Roof	
Floor	
Door	
Window	
Fence	
Other:	

Code (a)

Crop	Code	Crop	Code	Crop	Code
Teff	1	Black/Mixed Teff	21	Barley (Gebis)	41
Wheat (Durrah, Sinde)	2	Maize (Bekolo/Bahirmashla)	22	Sorghum / Mashila	42
Zengada	3	Oats	23	Horse Beans (Bakela)	43
Linseed (Telba)	4	Groundnuts (Lew)	24	Sesame (Selit)	44
Black Pepper (Kundoberbere)	5	Lentils (Mesir)	25	Vegetables	45
Coffee	6	Chat	26	Enset	46
Bananas	7	Grass	27	Gesho	47
Haricot Beans (Boloke)	8	Eucalyptus	28	Field Peas	48
Shiefera/Haleko	9	Fenugreek (Abish)	29	Dagussa	49
Beet Root (Key Sir)	10	Sunflower	30	Carrot	50
Potatoes	11	Ginger (Jinjibel)	31	Sugarcane	51
Selata (Lettuce)	12	Tobacco	32	Tikl Gommen	52
Pineapple (Ananas)	13	Pumpkin (Duba)	33	Avocado	53
Onions (Shinkurt)	14	Spinach (Quosta)	34	Garlic (Nech Shinkurt)	54
Yam	15	Fasolia	35	Fruit	55
Mango	16	Kocho	36	Hamicho	56
Chick Peas (Shimbra)	17	Cow Peas (Ater)	37	Orange	57
Godere	18	Adengware	38	Sweet Potatoes (Sikuar Dinich)	58
Tomato	19	Guaya (Vetch)	39	Nueg	59
Cabbage (Gomen)	20	Paddy, Rice	40	Sinar/Gerima	60

2. Do you store any cereals, pulses, or any other food crops or cash crops at present? (yes=1, no=0) If yes, give us details in the table below. (Use **Code (a)** above)

a. Stored crop		b. Are you currently storing any of these crops for:	
1. Crops <i>Code (a)</i> above	2. Quantity in kg	1. future sale? (yes=1, no=0)	2. future consumption? (yes=1, no=0)

3. In general, where did you sell your crop output?(this village=1, another village=2, middle men/traders=3, local town market=4, regional center=4, cooperative=5, Addis Ababa=6, other=7 (specify))
4. Did you use fertilizer on your land? (yes=1, no=0)
5. Did you use manure or compost on your land? (yes=1, no=0)
6. What are the major sources of marketing information? (neighbours in community=1, seed merchant=2, family/friends=3, local radio/TV=4, newspaper=5, extension workers=6, mobile phone=7, traders=8, own observations=9, other=10 (specify))
7. Was your farm output larger than last year? (yes=1, no=0)
8. Was your crop output affected because:
- someone in the family was too ill at critical periods of the farm operation? (yes=1, no=0)
 - outside labour was not available at the right time? (yes=1, no=0)
 - oxen were not available at the right time? (yes=1, no=0)
 - you could not get fertilizer, seeds, etc. at the right time? (yes=1, no=0)
9. About crop residues harvested in the last 12 months
- How much did you earn from selling crop residue (in ETB)?

- b. Have you given any part of the crop residue production to others as Kiray or rent? (yes=1, no=0)
 - 1. If yes, give the estimated value in ETB.....
- c. If the entire crop residue were sold, how much money would it fetch (in ETB)?

6.3 Agricultural inputs – family and hired labour, and labour sharing groups (last 12 months)

- 1. Excluding labour sharing, that is, apart from working with a work group, were any members of your household involved in your own land? (yes=1, no=0) If yes,
 - a. how many are they in total?
 - b. how many of them are adults (>=18 years old)?
 - c. for how many days?
- 2. Did you hire in any labour from outside the HH to work on your farm? (yes=1, no=0) If yes,
 - a. number of people hired
 - b. total number of days worked
 - c. total payment (in cash) per person per day
- 3. In the last 12 months, did you call for a work party [e.g., debbo, wenfel, jiggi, ...] for your household?..... (yes=1, no=0)
 - a. If no, why did you not call for a work party? (not customary in this area=1, no need for work party this year=2, too busy to work for others=3, enough male labour=4, enough income to hire labour=5, land too small=6, other=7 (specify))
 - b. If yes, give details.
 - 1. How many work parties did you call in the last 12 months?
 - 2. Why did you call for a work party? (it is customary to do so=1, quick completion of task=2, only way to get large amount of labour=3, cannot afford paid labour=4, no paid labour available=5, group is best way of completing task=6, other=7 (specify))
 - 3. Type of task work party was called for (ploughing/digging/land preparation=1, weeding/watering/pruning=2, harvesting=3, threshing/storage/processing=4, construction=5, other=6 (specify))
 - 4. Will you have to reciprocate this in terms of your labour or any other way? (yes=1, no=0)
 - 5. How many people participated in your work parties? a. males b. females
 - 6. How many days did your work parties last?

6.4 Agricultural inputs – expenditures

- 1. In the following table, give us the expenses related to inputs for crop agriculture. [Probe for items not in the list]

a Type of expenditure during the last 12 months	b Total amount paid (in ETB)
fertilizer	
pesticides (incl. fungicides and herbicides)	
seeds and young plants (seedlings)	
labour for crop production (not intermediate input)	
transport related to crop production and sale	
rent for oxen, tractor, harvester, or combine services	
others (specify	

6.5 Livestock ownership

- 1. Do you have any livestock? (yes=1, no=0) If yes, give us some details in the following table.

a Type of livestock	b Number owned and present at your farm/home	c Number not owned but cared for	d Number owned but away	e If you would sell one of the..[...].today, how much would you receive from the sale (in ETB)?
calves				
oxen				
heifer				
cows				
sheep				
goats				
horses				
donkeys				
mules				
camels				
cross breed cow				
cross breed oxen				
cross breed heifer				
cross breed calves				
exotic or fresian cows				
exotic or fresian heifer				
chicken				
others				

2. Did any of your livestock suffer from any of the following factors in the last 12 months?

(1=moderately, 2=severely, 3=not at all)

- a. Lack of drinking water
- b. Lack of grazing land
- c. Animal diseases
- d. Other (specify)

6.6 Livestock income and expenditure

1. We would like to know your gross income from the sale of **livestock products** in the last 12 months.

a Type of livestock product	b Amount sold	c Unit	d Total revenue obtained from the sale of ..[...].
meat (excluding live animals)			
hides/skins			
butter/cheese			
milk/cream/yoghurt			
dungcakes			
eggs			
others			

2. In general, where did you sell your LS outputs? (this village=1, another village=2, middle men/traders=3, local town market=4, regional center=4, cooperative=5, Addis Ababa=6, other=7 (specify

3. Was your LS output better than last year? (yes=1, no=0)

4. Did you get any income from hiring out oxen during the last 12 months? (yes=1, no=0)

a. If yes, how much (in ETB)?

5. In the last 12 months, have you had any expenditure related to livestock? (yes=1, no=0) If yes, give us some details in the following table.

a Type of expenditure	b Cash value (in ETB) [If in kind, give estimated cash value]
labour related to LS production	
feed and water	
veterinary services/medicine	
other expenses	

Part 7. Off-farm income and business activities

1. In the last 12 months, did you work off the household’s land either on someone else’s land or in some other employment, against payment in cash, including as part of food for work, or as part of a labour sharing arrangement (e.g., debbo, wonfel, etc.)? (yes=1, no=0) If yes:

a. where? (this village=1, another village=2, local town market=3, regional center=4, Addis Ababa=5, other=6 (specify))

b. give us the specifics of the work, days worked, and total amount of money earned.

Specify the kind of work	c Days worked	d Total amount earned (in ETB)
farm worker (for pay)		
professional (teacher, government worker, administration, health worker, clerical)		
skilled labourer (i.e., builder, carpenter, etc)		
trader		
soldier		
driver/mechanic		
unskilled non-farm worker		
domestic servant (yebet serategna)		
food-for-work		
weaving/spinning		
milling		
handicraft (incl. pottery)		
traditional healer		
religious teacher/preacher		
transport (by pack animal)		
collecting & selling firewood or dungcakes		
other		

2. Have you received any other income (such as remittances, gifts, food aid/other aid, payment for health or education, any other transfers) in the last 12 months?
(*yes=1, no=0*) If yes, give details in the following table.

a Type of receipt	b Who sent you the transfer?*	c Amount (in ETB)	d Have you ever given a gift to this person? (<i>yes=1, no=0</i>)	e Do you work together in labour sharing arrangements? (<i>yes=1, no=0</i>)	f Are you partners in a sharecropping or land renting arrangement? (<i>yes=1, no=0</i>)	g Are you partners in oxen-sharing arrangement? (<i>yes=1, no=0</i>)	h Are you members of the same <i>iqqub</i> ? (<i>yes=1, no=0</i>)	i Are you both members of the same <i>iddir</i> ? (<i>yes=1, no=0</i>)
remittance								
gift								
inheritance								
donation/aid								
dowry								
transfer for school costs								
compensation								
other								

**Code (a): non-resident HH member=1, relative=2, friend/neighbour=3, other=4 (specify)*

3. Have you given anything to another household (such as remittances, payment for someone else's health or education, any other transfers) in the last 12 months? (Do not include regular or statutory contributions to *iddir*, but focus on gifts directly to other households) (*yes=1, no=0*) If yes, give details.

a Type of receipt	b To whom did you send the transfer?*	c Amount (in ETB)	d Have you ever received a gift from this person? (<i>yes=1, no=0</i>)	e Do you work together in labour sharing arrangements? (<i>yes=1, no=0</i>)	f Are you partners in a sharecropping or land renting arrangement? (<i>yes=1, no=0</i>)	g Are you partners in oxen-sharing arrangement? (<i>yes=1, no=0</i>)	h Are you members of the same <i>iqqub</i> ? (<i>yes=1, no=0</i>)	i Are you both members of the same <i>iddir</i> ? (<i>yes=1, no=0</i>)
remittance								
gift								
inheritance								
donation/aid								
dowry								
transfer for school costs								
compensation								
other								

Part 8. Household consumables

8.1 Non-food items: please give us details of the non-food item purchases you made for use by the household.

Non-food item (last 12 months) (non-frequently consumed items)	Expenditure (in ETB)	Non-food item (last month) (frequently consumed items)	Expenditure (in ETB)
Clothes		Transport	
Shoes		Laundry soap/omo, hand soap	
Kitchen utensils/equipments		Chat	
Furniture (table, chair, bed, etc)		Cigarettes, tobacco, suret, gaya	
Other assets (bicycle, car, refrigerator, etc)		Cosmetics (Hair oil, perfume, etc)	
Modern medical treatment and medicines		Donations to the church/mosque	
Traditional medicine and healers		Involuntary (forced contributions)	
School fees		Contributions to IDDIR	
Other educational expenses (exercise books, pens, pencils, uniforms, maintenance, club fees, etc)		Other voluntary contributions (including Erteban/ligsina/irdata)	
Gold, dowry for spouse (ceremonial expenses), and other luxury items		Firewood, matches, batteries, charcoal, kerosene, candles, incense, sendel	
Tax and rent (crop, land, oxen, etc)		Other expenses	
Compensation and/or penalties			
Labour cost/salary			
Other expenses			

8.2 Food items: we would like to ask you about all the food items that were bought for consumption in the last month.

Food type (last month)	Expenditure (in ETB)	Food type (last month)	Expenditure (in ETB)
Cereals (teff, sorghum, barley, maize, etc)		Alcoholic beverages (Araqi/Kathikala, beer, tella, tej, etc)	
Vegetables (tomato, cabbage, etc)		Coffee/ashara, tea	
Root crops (carrot, ginger, potatoes, etc)		Spices (Turmeric (Ird), Black Pepper (Kundoberbere), Berbere, etc)	
Fruits (avocado, banana, orange, etc)		Salt, sugar	
Meat (beef, goat, sheep, chicken, etc)		Prepared food (bread, injera, shiro, etc)	
Eggs		Others	
Dairy products (milk, butter, cheese, yoghurt, etc)			
Cooking/edible oil			
Honey			
Soft Drinks (Karibo/Keredo, coke, etc)			

8.3 Have you consumed food from your own harvest or your own stock during the last month? (yes=1, no=0)

8.4 Generally, where do you purchase: (**Code (a):** this village=1, another village=2, local town market=3, regional center=4, Addis Ababa=5, other=6 (specify ...))

- e. food grains?
- f. fruits or vegetables?
- g. meat and/or dairy products?
- h. sugar, salt and/or cooking oil?
- i. processed foods such as sodas, other beverages or packaged foods such as biscuits?

8.5 During the last month, have you given away any food to other people, such as neighbours or family members not belonging to your HH? (yes=1, no=0)

8.6 Consumption habits

- a. During the past 12 months, did your household suffer from shortage of food to eat? (yes=1, no=0) If 'no', go to question 'f'. If 'yes', proceed to question 'b'.
- b. How many months in the last 12 months did you have problems satisfying the food needs of your household?
- c. In the months with shortage of food, did you eat foods that you ordinarily would not eat, i.e., "less preferred foods"? (yes=1, no=0)
- d. Compared to your usual diet, did you cut back quantities served per meal to:
 - 1. adult males? (yes=1, no=0)
 - 2. adult females? (yes=1, no=0)
 - 3. boys? (yes=1, no=0)
 - 4. girls? (yes=1, no=0)
- e. During the worst month, how many times a day did:
 - 1. adults in your HH eat?
 - 2. children in your HH eat?
- f. During a good month, how many times a day did:
 - 1. adults in your HH eat?
 - 2. children in your HH eat?

Part 9. Shocks

9.1 Health status

1. In the last 12 months, have you suffered any illness, disease or injury? (yes=1, no=0) If yes, give us details in the following table.

a What kind of disease?	b How many days in total?	c Treatment			d Cost of drugs (in ETB)
		1 Have you received? (yes=1, no=0)	2 From where?*	3 Cost (in ETB) per treatment/medication	
malaria					
TB					
diarrhea					
flu					
HIV/AIDS					
schistosomiasis					
diabetes					
others					

*Code (a): public clinic=1, private clinic=2, public hospital=3, private hospital=4, traditional healer=5, other=6 (specify

9.2 Long-term shocks and coping mechanisms

1. We would like to learn about shocks in the last *five years*. Has this HH been affected by a serious shock – an event that led to a serious reduction in your asset holdings, caused your HH income to fall substantially or resulted in a significant reduction in consumption? (yes=1, no=0) If yes, we would like to know more about it.

a. Specifics of shock	b. In what year did the shocks occur?	c. Did the shocks result in:							d. How widespread was this shock? <i>Code (g)</i>
		loss of productive assets? (yes=1, no=0)	loss of HH income? (yes=1, no=0)	reduction in HH consumption? (yes=1, no=0)	asset and income loss? (yes=1, no=0)	asset loss and reduced consumption? (yes=1, no=0)	income loss and reduced consumption? (yes=1, no=0)	asset, income loss and reduced consumption? (yes=1, no=0)	
<i>Natural shocks Code (a)</i>									
<i>Production shocks Code (b)</i>									
<i>Market shocks Code (c)</i>									
<i>Political and social shocks Code (d)</i>									
<i>Criminal shocks Code (e)</i>									
<i>Idiosyncratic shocks Code (f)</i>									

Code (a): 1: Drought, 2: Flood, 3: Hailstorm, 4: Fire outbreak, 5: Landslide

Code (b): 6: Pests of crops before harvest, 7: Crop loss during storage, 8: Animal disease

Code (c): 9: Large increase in input prices, 10: large decline in output prices, 11: Inability to sell agricultural products, 12: Inability to sell nonagricultural products

Code (d): 13: land distribution by government, 14: Forced migration, 15: Discrimination for political reasons, 16: Forced contributions,

17: Arbitrary taxation, 18. Discrimination for social reasons

Code (e): 19: Destruction or theft of tools or inputs for production, 20: Theft of crops, 21: Theft of livestock

Code (f): 22: Loss of job by family member, 23: Death of family member, 24: Illness of family member, 25: Separation of family member[s],

26: Dispute with extended family, 27: Dispute with others in village, 28: Other (specify)

Code (g): only my HH=1, some HH in village=2, all HH in village=3, many HH in district=4

2. What action did you take against these shocks? (*did nothing=1, sold livestock=2, borrowed from relatives=3, borrowed from bank=4, received food aid=5, participated in food for work=6, H head migrated to other rural area=7, HH plus others migrated to rural area=8, migrated to urban area=9, sought off-farm employment=10, ate less=11, other=12 (specify)*)

3. Have you received any aid in the past 12 months? (*yes=1, no=0*) If yes, give some details in the following table.

a Type of aid	b Month	c Quantity/amount
cash for work		
food for work		
food (emergency relief)		
income generating scheme/training		
other		
d. Did you or any member of this HH participate in the <i>Productive Safety Nets Program</i> in the last 12 months? (<i>yes=1, no=0</i>)		

Part 10. Participation in organizations, village life and decision-making

1. When there is a village assembly/meeting of any kind, how often do you participate? (*always=1, sometimes=2, rarely=3, never=4*)
 - a. Do you usually speak up/raise issues in those meetings? (*yes=1, no=0*)
2. When there is a village, *woreda* or national election of any kind, do you vote? (*yes=1, no=0*)
3. Are you a member of at least one association or group? (*yes=1, no=0*)
4. Do you belong to an *iqqub*? (*yes=1, no=0*)
5. Are you a member of the village council? (*yes=1, no=0*)
6. Do you have a leadership position in the kebele/PA administration? (*yes=1, no=0*)
7. How many times in the past month have you attended a Church/Mosque or other religious service?
8. Do you belong to at least one *iddir*? (*yes=1, no=0*)
 - a. If not, why not? (*no need=1, cannot afford to join=2, not allowed to join=3, other=4 (specify)*)
 - b. If yes, how many *iddir* are you a member of?
 - c. How much do you contribute to an *iddir* each month (in ETB)?
 - d. How much money does the *iddir* give you in time of death of a family member of a relative? (in ETB)
 - e. Besides payments at the time of a death, does the *iddir* provide cash payouts/gifts for events/circumstances (e.g., in case of fire, loss of oxen/LS, destruction of house, wedding, illness, harvest loss, etc)? (*yes=1, no=0*)
 - f. Does the *iddir* give any loans? (*yes=1, no=0*)
 1. If yes, for what purposes? (*loan for any purpose=1, loan in case of funeral=2, loan in case of fire=3, loan in case of loss of oxen/LS=4, loan in case of destruction of house=5, loan in case of wedding=6, loan in case of illness=7, loan in case of harvest loss=8, other=9 (specify)*)
 - g. During the last 12 months, have you received any payouts or loans from the *iddir*? (*yes=1, no=0*)
 - h. For what other reasons beyond funeral support have you joined this *iddir*? (*to obtain loans=1, to provide protection outside funeral insurance=2, to participate in social activities with others=3, to make contacts/friends=4, other=5 (specify)*)
9. Do you belong to any savings/credit organization? (e.g., microfinance) (*yes=1, no=0*)
10. In the past 12 months, have you taken out a loan in cash? (*yes=1, no=0*) If yes, give us details.
 - a. Source of loan (*money lender/arata=1, relative=2, friend/neighbor=3, from iqqub=4, from iddir=5, from the cooperative=6, from other local organizations=7, from a bank=8, from government/ministry/kebele=9, from microcredit program/NGO=10, other=11 (specify)*)
 - a. Why did you want to obtain a loan? (*to buy farm or other tools/implements=1, to buy inputs (e.g., seeds, fertilizer, pesticides)=2, to buy LS=3, to pay for hired labour=4, to pay rent/taxes=5, to start an off-farm business (like weaving)=6, to buy food/goods for the HH=7, to pay for travel expenses=8, to pay for building materials=9, to pay for health expenses=10, to pay for educational expenses=11, for wedding=12, for funeral=13, repay other debts=14, other=15 (specify)*)
 - b. Total amount of loan (in ETB)
11. In the past 12 months, have you given out any loan in cash? (*yes=1, no=0*)

- a. If yes, total amount (in ETB)
12. Do you have a bank account? (yes=1, no=0)
13. If yes, how much money do you have in this account (in ETB)?

Part 11. Networks

1. We want to know about **THREE** most important people you can rely on in time of need, either within this village or elsewhere.

	a Age	b Number of oxen he/she owns	c Does he/she own more or less land than you? Code (a)	d Are any of his/her plots of land next to/near yours? (yes=1, no=0)	e Size of his/her HH compared to yours? Code (b)	f Is he/she a neighbour? (yes=1, no=0)	g Is he/she a relative? (yes=1, no=0)	h Is he/she a member of the same iddir as you? (yes=1, no=0)
1.								
2.								
3.								

Code (a): more=1, about the same=2, less=3
 Code (b): larger=1, about the same=2, smaller=3

	i Are you both members of the same labour sharing group? (yes=1, no=0)	j Are you partners in a sharecropping or land renting arrangement? (yes=1, no=0)	k Are you partners in oxen-sharing or similar arrangement? (yes=1, no=0)	l Are you members of the same <i>iqqub</i> ? (yes=1, no=0)	m Has he/she already helped you in the past? (yes=1, no=0)	n Can he/she turn to you in case of need? (yes=1, no=0)	o Have you helped him/her in the past? (yes=1, no=0)
1.							
2.							
3.							

2. Beyond these THREE people, how many people can you rely on in time of need?

On behalf of the research team, I would like to thank you very much for your time and active participation in this survey!

