Wageningen University –Department of Social Sciences

Development Economics Group
Msc thesis

Linking Small Farmers to Markets

Household Determinants of Smallholder Commercialization in Northern Ethiopia

Tesfalem Hadgembes Assay

International Development Studies
Specialization- Economics of Rural Development

August, 2008

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Linking Small Farmers to Markets

Household Determinants of Smallholder Commercialization in Northern Ethiopia

A thesis written for the partial fulfillment of master's degree in International Development Studies (specialization: Economics of Rural Development)

Tesfalem Hadgembes Assay

Supervisor: Dr. Nico Heerink
Wageningen University
Development Economics Group

Wageningen,
The Netherlands
August, 2008

Acknowledgment

I would like to say some words of appreciation to those who have contributed to the accomplishment of this thesis directly or indirectly. First, I would like to thank my supervisor, Dr. Nico Heerink for his valuable comments and cooperation from the very proposal through to the final draft. His thorough comments helped me put extra effort to improve the draft until it has the current shape. I would like to thank Guush Berhane, and Leonoor Akkermans, for their constructive comments, and Gian-Nicola Francesconi for suggesting the topic and supplying materials.

I would like to use this moment to forward my special thanks to NUFFIC for not only funding this thesis but also sponsoring my study in Wageningen which I would hardly think of doing on my own. Finally many thanks to Muluberhan Kifle, Gemechu waktola and Getu Tigre for their assistance in finding my way to Wageningen University.

Tesfalem,
August, 2008

Abstract

Development of favorable market institutions, trade and specialization are crucial for successful agricultural development. However, agriculture in Ethiopia is still subsistence-based agriculture to a large extent. Small farmers, who take the lion's share of the agricultural production, are not well integrated to the product and factor markets.

The aim of this thesis is first to characterize subsistence and semi-commercialized households; and then to identify the major household-level economic determinants of smallholder commercialization in the arid and semi-arid areas of Northern Ethiopia, Tigray region. As a major source of data, a structured questionnaire was administered to ninety farm households (55 semi-commercialized and 35 subsistence households) from *Golgol-Naele* village at the north-east of the Tigray region. The data is analyzed within the framework of non-separable farm household models with missing food and labor markets. Then Heckman selection equation is estimated for the decision to commercialize and the degree of commercialization using maximum likelihood and two-step estimation methods. A statistical comparison shows that subsistence and semi-commercialized households have significant heterogeneity in their access to irrigation, risk attitude, information access, productivity and livestock ownership. Further, the econometric result shows that irrigation access and fixed transaction costs(radio ownership) affect small farmers' decision to commercialize. The probability of market participation is likely to be higher if a household has access to irrigation and owns radio. Similarly, the major determinants of degree of commercialization are percentage of irrigable land and risk attitude of the household. A one percent increase in irrigable land has an effect of raising cash crop supply by 0.30 Birr per each kg of output produced while it raises area allocated to cash crop by 50 to 60 percent.

A vital policy implication of the results is that securing moisture availability should take priority in policies designed to promote commercialization in the dry lands of the Tigray region. While market initiatives like the recently implemented commodity exchange system can address transaction cost and risk problems primarily of the surplus regions of southern, central and western parts; water resource development should take priority in the moisture-scarce regions of northern and eastern Ethiopia.

Keywords: smallholder commercialization; market participation; farm household; Ethiopia; Tigray region

Contents

List of tables

List of figures

Chapter 1 Introduction

1.1 Background

According to neoclassical economic theory, development of favorable marketing institutions, trade and specialization are crucial for successful agricultural development. Yet, agriculture in a large part of the developing world is still subsistence-based to a large degree. A significant amount of land is devoted to subsistence cropping. For instance, around 440 million farmers are engaged in subsistence agriculture (von Braun, 1995).Efforts to integrating such farmers into the factors and product markets is part of the overall development endeavors.

There is sizeable body of literature on smallholder commercialization in developing countries. It addresses among others, the effects and determinants of agricultural commercialization on food security and nutrition (von Braun et al., 1994); cash crop and food crop synergies (e.g. Jayne, 1994; Govereh and Jayne, 1999); market participation decisions of small farmers (e.g. Goetz, 1992); and agricultural supply response of small farmers with missing markets (e.g. Key et al., 2000).

Specifically, in Ethiopia there have been studies underway at the International Food Policy Research Institute (IFPRI) on commercialization of small farmers. Their main focus has been on identifying efficient coordination mechanisms that can reduce transaction cost small farmers face and they have been based mainly on data either from national level surveys or household surveys from the surplus areas of the country. The development of such researches has contributed towards the establishment of a commodity exchange market system early this year (2008). However, such national level surveyed data could be too aggregative to represent heterogeneity of agriculture potential and other factors across the country not to mention the insufficiency of data from only the surplus areas. National agricultural commercialization policy designs should consider heterogeneity of farming and farmers in different regions of the country.

This thesis goes one step towards filling the gap by providing household level evidence from the Tigray region which typifies most of the dry lands in the northern and eastern parts of the country characterized by low agriculture potential and supply deficit. The approach of the thesis is to analyze the determinants of small farmers' market participation in cash crop supply at two levels: first their decision to participate, and then, their extent of participation.

1.2 Problem statement

Agriculture in Ethiopia accounts for over 80 percent of the national employment, 50 percent of the GDP, and more than 60 percent of the export (CSA, 2005).The sector has long been predominantly of traditional smallholder farming kind with high degree of subsistence orientation and thin markets.

Farmers are weakly connected to factor and product markets both domestically and internationally. Only 30 percent of the agricultural output is marketed (Bernard et al., 2006). Moreover, empirical studies show that cereal commercialization has been very low. For example, according to early estimates only 28 percent of total national agricultural output is estimated to have been commercialized. Recent estimates show that around 38 percent of *teff* (a widely consumed staple food in the country) production is commercialized (Bernard et al., 2006).

Recently, there are national policy interventions by the government to promote commercialization of small farmers into local as well as international markets (MOFED, 2005).Yet, Ethiopia is a country of diverse regions in terms of agricultural potential, demographic structure, and access to market & infrastructure (Chamberlin et al.,2006). Much of the southern, central and western parts of the country are endowed with relatively good agriculture potential, low population density (in the low lands) and relatively good access to markets and infrastructure (proximity to Addis Ababa). On the other hand most parts of northern and eastern Ethiopia have less agricultural potential due to moisture scarcity, population pressure, and relative remoteness from the centre.

National policies including commercialization policy should take into account such diversity when it comes to setting priorities.

Small farmers in the arid and semi-arid areas of northern Ethiopia such as the Tigray region, face not only high transaction costs but more notably also moisture scarcity. Commercialization interventions in such areas vary from those in other parts of the country, and need to be supported by evidence from micro-level research that examine the factors determining the integration of small farmers into markets.

1.3 Research objective

General objective

The general objective of the study is to characterize and identify the major household-level economic determinants of commercialization behavior of smallholder farmers in the Tigray region of Ethiopia.

General question

What are the major determinants of small farmers' decisions to commercialize and of their degree of commercialization in the Tigray region, Ethiopia? To what extent do the determinants of the commercialization decision and the degree of commercialization differ from one another?

Specifically the following **research questions** are identified:

i. Do household and production characteristics of subsistence and semi-commercialized farm households differ?

ii. What are the major household determinants of small farmers' **decisions to participate** in cash crop supply?

iii. What are the major household determinants of small farmers' **extent of participation** in cash crop supply?

iv. What policy implications can be inferred from the answers to these research questions?

1.4 Data and method of data analysis

The major source of the data used in the thesis is a structured questionnaire administered to 90 household farmers in the *Atsbi-Wemberta* district of the Tigray region, Ethiopia. The sample was selected in such a way that 60 percent of the total sample includes semi-commercialized [1] households whereas the other 40 percent comprise subsistence households.

The research questions are addressed within the framework of non-separable farm household models with missing food and labor markets. Heckman selection model, is used to estimate some hypothetically selected determinants of commercialization using maximum likelihood and two-step methods of estimation. Heckman model was chosen because of its capability to examine the determinants of decision to commercialize and extent of commercialization separately. The methodology is further elaborated in Sections 3.3.1 and 3.3.2.

1.5 Thesis organization

The rest of the thesis is organized as follows. In chapter two the literature related to agricultural commercialization will be briefly summarized while some concepts are also defined. In chapter three survey setting, data collection, sampling strategy, and the methodology used to address the research questions are discussed. In chapter four, the statistical and econometric analysis and their results are presented and discussed. Finally, in chapter five the thesis ends by discussing summary of the findings and their policy implications with some remarks on the limitation of the thesis.

[1] The distinction between semi-commercialized and subsistence households is defined in section 3.2

Chapter 2 Related literature and theoretical framework

2.1. Introduction

One strategy in the agenda of transforming traditional agriculture in to commercialized farming should be identifying and characterizing the determinants of agricultural commercialization both at micro and macro levels. The literature in the broad area of agricultural commercialization is rich and diverse. In this chapter an attempt is made to define agricultural commercialization and roughly asses a few of the major works on the determinants of commercialization.

2.2. Agricultural commercialization defined

Commercialization of subsistence agriculture may simply be perceived as a process of change from subsistence to commercial farming. There is no uniformity within and outside the economics domain on what the term subsistence refers to. Subsistence agriculture has been synonymous (usually with a connotation of backwardness and inefficiency) to terms such as smallholder, peasant, low income, resource poor, low input, and low technology farming. Hence, there are conceptual and measurement ambiguities with regard to the degree of subsistence and commercialization (Brunthrup and Heidhues, 2001). According to the authors the most commonly used definition of subsistence in agricultural economics is the percentage of production for own consumption, usually with the benchmark of 50 percent. On the other hand, agricultural commercialization is often associated with the size of cash crop availability relative to other non-cash produces. This has been a common concept used in empirical researches.

Commercialization may be broadly seen as a continuous process of transformation from subsistence agriculture in to a specialized commercial farming. Thus the end target of commercialization is more than the supply of surplus output to the market or the volume of cash crops produced.

Commercialization of subsistence agriculture may be broadly defined as the process of transformation of subsistence agriculture into a diversified (at sector level), specialized (at farm level), profit maximizing farm decision making where production (be it cash or staple) and input use decisions are market driven (Timmer, 1997; Pingali and Rosegrant, 1995; Von Braun, 1994).

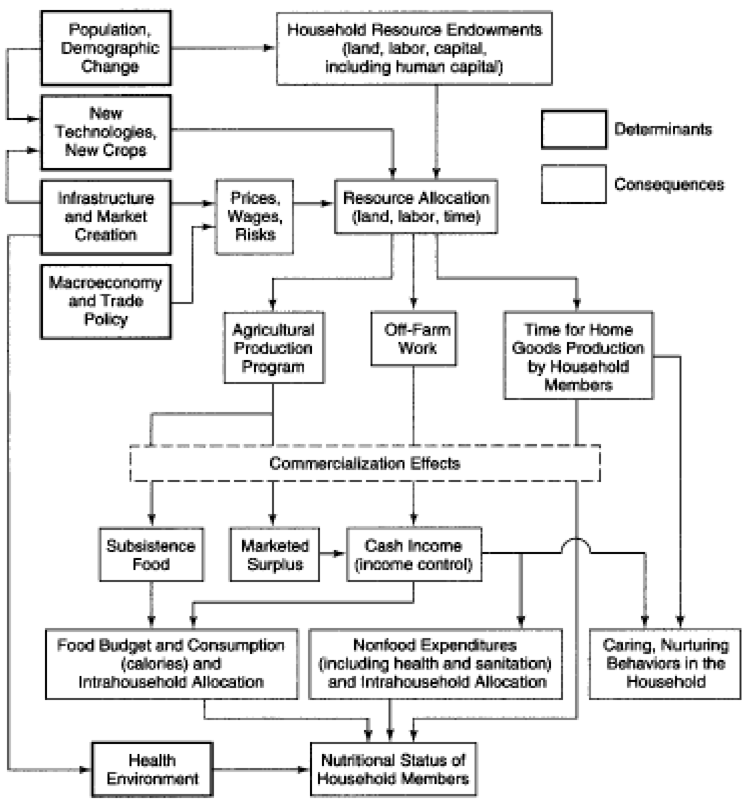
Von Braun et al. (1994) distinguish among three forms of commercialization: The first is commercialization of agriculture from the input and output side which can be measured as the ratio of value of output (input) to the total value of agricultural production. The second refers to commercialization of the rural economy which may be measured as the value of all goods and services used by the rural dwellers that come by way of market exchange be it in kind or in cash. The third distinction is about degree of integration of the rural economy into the cash economy that is when all transactions made are entirely in cash. The latter can be measured as the ratio of cash-transacted goods and services to total income.

2.3. Determinants of commercialization

A rather broader picture of the determinant factors of subsistence-versus commercialization of agriculture is outlined in Brunthrup and Heidhues (2001) where the determinants are set in three circles: farm internal factors (resource endowment and family specific characteristics), farm external - country internal (policies, institutions, markets etc.), and country external factors (ecology, history, culture, international environment etc.). However, such frameworks while vital for theoretical conceptualization of the determinants of commercialization are too broad for empirical analysis. A more practical and refined framework of classifying the determinants (and effects) of commercialization is used in Von Braun et al. (1994) as shown in Figure 1.

To begin with, the determinants in the upper part of Figure 1 may be distinguished into two levels -macro and micro level determinants. At macro level the major driving forces of agricultural commercialization include: population growth and urbanization, technological development, infrastructure and market development, macro economic and trade policies. Such factors in turn influence the household level resource endowment and allocation thereby affecting the commercialization behavior of households. For instance, introduction of irrigation technology to moisture deficit areas can shift household land and labor from subsistence production to production for markets. Similarly, availability of infrastructure can affect the price by reducing transaction cost hence greater opportunity for commercialization. Commercialization, as shown in the bottom of Figure 1, can then affect health, nutrition and food security level of the household.

Figure 1 Determinants of commercialization and nutritional status of households



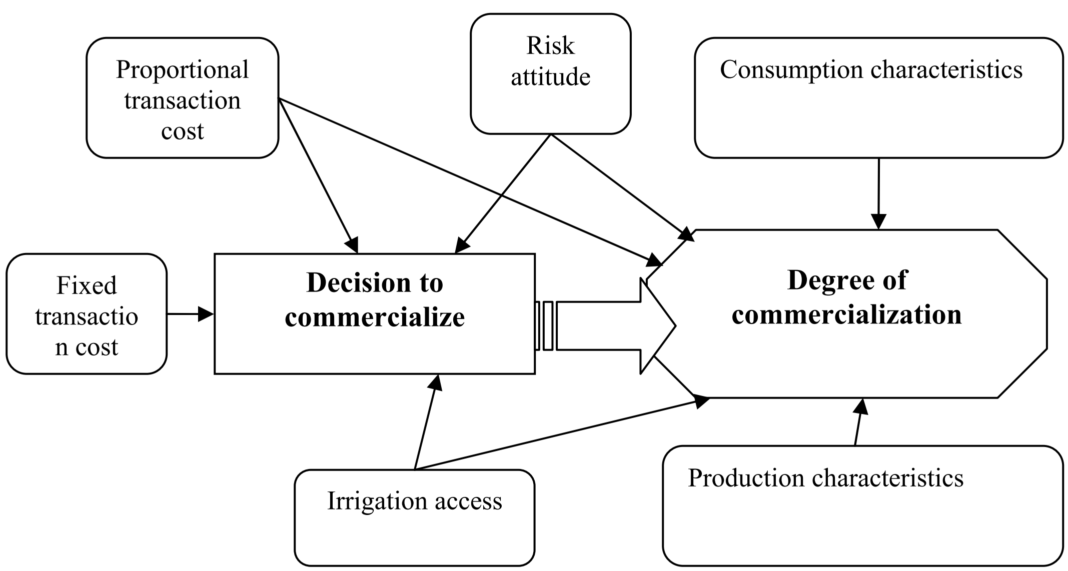
Source: Von Braun et al., 1994

2.4. Household level determinants

At the household level commercialization decisions are determined by various factors (varying at household and village levels) such as, agro-climatic conditions , risk behavior of household farmers, market and infrastructural access, resource availability at household and community level, institutions, socio-cultural conditions, and development of local goods and services, and rural credit and factor markets(Pender, 2006).Recent development of household models are considering also such decisive factors, like risk conditions, and income and consumption smoothing , transaction cost, and liquidity constraints to expand the understanding about the determinants of market participation decisions.

One important finding in the literature is that household determinants of commercialization may differ in decision to participate and in degree of participation (e.g. Goetz 1992). The theoretical framework adopted in this thesis (as sketched in Figure 2 and illustrated in section 3.3) separately addresses the household determinants of decision to commercialize and the determinants of degree of commercialization.

Figure 2 Conceptualizing household determinants of commercialization

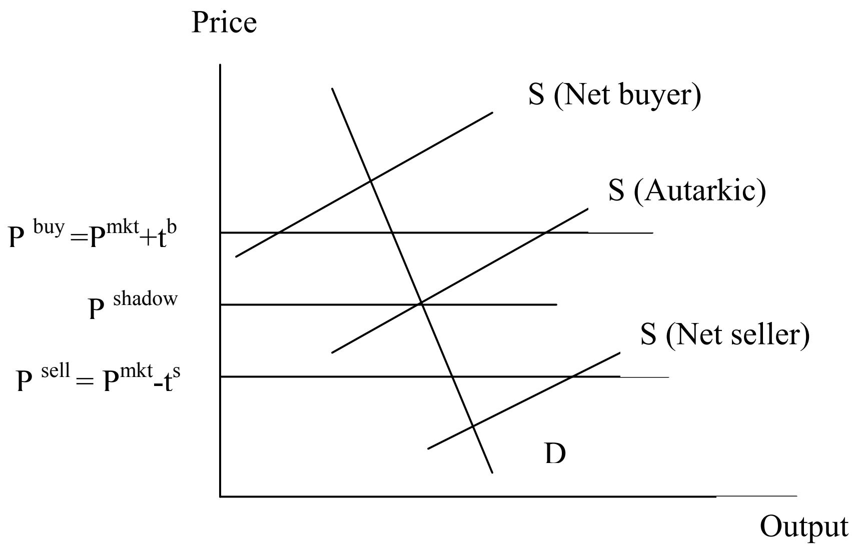


Participation decision

One specific line of development of the household level determinants in the literature hinges on the argument that transaction costs can make a remarkable difference in household farmers' decision to participate as seller, buyer or autarkic (Goetz 1992; Key et al. 2000; Vakis et al. 2003).Two types of transaction costs are identified -fixed transaction cost (FTC) and proportional transaction cost (PTC). FTC, which is independent of output traded, may include costs related to search, negotiation, bargaining, screening, enforcement and supervision. Fixed transaction cost is crucial in the decision of market participation.

The PTC refers to per unit cost of market accessibility which varies proportional to agricultural output traded. De Janvry and Sadoulet (2006) argue that market participation decision either as a buyer or seller does not exist in a certain price band that can be created by the presence of proportional transaction costs. Within such price interval farm households prefer to stay subsistence by adjusting their production and consumption decisions internally. This is illustrated in Figure 3 where three types of households are given having identical demand (D) in a given market. Given that p^{buy} is the buying price (market price plus buying transaction cost t^{b}); p^{sell} is the selling price (market price minus the selling transaction cost t^{s}); and S is output supplied .With the presence of such transaction costs, a price band is created between p^{buy} and p^{sell}. If the demand and supply of a household intersect within such price band, then it is not beneficial for the household to participate in the market. As a result the household remains autarkic by adjusting internally to a shadow price P^{shadow}.

Figure 3 The price band



In addition to transaction cost, risk preference of households can also play a role in determining market participation decision (Sadoulet and de Janvry, 1995; Fafchamps, 1992).The higher the risk aversion tendency of the household, the lower the probability of participating in cash crop production. In terms of Figure 3 the price band of such risk-averse household would be wider hence higher probability of staying autarkic.

Further, moisture availability may have significant effect on market participation behavior of farm households in the study area. A pilot survey made by the ILRI (International Livestock Research Institute) in *Atsbi-Wemberta* district identifies water resource development as a major factor in promoting market orientation of household farmers in the district (ILRI, 2004).The survey observes that with introduction of some water harvesting schemes like digging small ponds, the production of cash crops has increased. Moreover, farm households who have access to irrigation from earthen dams were observed to show some degree of shift from food crop production to cash crops such as vegetables (Hagos,2007).Most cash crops such as vegetables demand higher volumes of water than most food crops. Therefore, the higher the availability of water the higher the tendency to allocate land to cash crops. Hence, access to irrigation should be considered as a key determinant of commercialization decision.

Degree of participation

Among the participating households, still there can be variation in the extent of participation. Major determinants at this level can be the proportion of land accessible to irrigation, other characteristics in production, and characteristics in consumption, proportional transaction costs and risk attitude.

The degree of participation increase/decrease proportional to characteristics in consumption and production such as land size percentage of irrigable land, farm equipment, household size, labor index, etc (see section 3.3.4 below).The higher the proportional transaction cost the lower the proportion of cash crop supply. Similarly, the stronger the level of risk aversion, the smaller the degree of participation. If a household attaches a risk margin to its expected price due to its risk preference behavior (expected price = actual price + risk margin), the household s' extent of participation will vary proportional to such margin (Sadoulet and Dejanvry, 1995).

Chapter 3 Survey setting, data and methodology

3.1 Description of survey area

Tigray region (see map in Appendix) is located in the rear north of Ethiopia, East Africa. The region is set on a total estimated area of 50 thousand square kilometer bordering Sudan from the west and Eritrea from the north. About 65 percent of the land is under cultivation. The total population of the region by 2005 is estimated to be 4.3 million with annual growth rate of 3 percent. An estimated 81 percent of the population is rural inhabitant and over 90 percent of them are subsistence smallholder farmers growing predominantly cereals (CSA, 2005).

The region was selected for two reasons. One reason is that the region typifies the arid and semi-arid areas of northern and eastern Ethiopia. Most of the Tigray region falls under a semi-arid climatic zone. Most of the region has among the lowest agriculture potential in the country and is a deficit area. It is therefore interesting to compare the determinants of commercialization for such a region with those that are found in comparable studies in other parts of Ethiopia with more favorable agriculture potential. The other reason is local language access. The author wants to use the advantage of his knowledge of the local language which saves him enumeration and related costs.

There are 11 regions under the administration of federal government of Ethiopia among which Tigray region is one. Under the regional administration of Tigray there are four administrative zones (eastern, central, western and southern zones) which totally constitute 35 *woreda* or districts. Each district is divided into *tabia* or PAs (peasant associations) and each *tabia* is further classified into *Kushet* or villages *(CSA, 2005).* The survey was held at one of the PAs of *Atsbi-Wemberta* district in the eastern zone of the Tigray region.

Atsbi-Wemberta

Atsbi-Wemberta is a district located at about 860 km north of Addis Ababa, and further up to 65 km towards north east of Mekelle, the Regional capital, stretched over a total geographic area of 1223 km square. The estimated total population of the *woreda* in 2003/04 was 110578 or about 41,398 households. The altitude of the area ranges from 918 m.a.s.l. (meters above sea level) in the northern parts to 3069 m.a.s.l. in the southern parts of the *woreda*. The average annual rain fall in 2003/04 was recorded to be 642 mm/yr with uneven distribution over the agricultural year. The area receives short light showers during winter and relatively better rains (mostly short and intensive) during the summer season. There is a total area of 35305 ha potentially cultivable, out of which about 40 per cent is actually cultivated. The average household owns 0.5 ha of land mainly growing food cereals such as barley, wheat; pulses including fababean, field peas, lentil; and recently, with the introduction of water harvesting development, vegetables. Pulses and vegetables are the most marketed crops with relative better market prices (ILRI, 2004).

3.2. Data collection and sampling

A structured household questionnaire was developed as a major source of data for the statistical as well as econometric analysis. A total sample of 90 households is included in the survey. The sample is selected from one of the 16 PAs (Peasant Association) in the district called *Golgol-Naele,* using stratified random sampling technique, with semi-commercialized and subsistence farmers being the two strata. This PA is selected based on a suggestion from the Agricultural bureau of *Atsbi-Wemberta* district. Two PAs were suggested: *Hayelom* and *Golgol-Naele*- which are recognized to constitute subsistence and semi-commercialized farmers at the same time. Such coexistence of both types of farm households is vital for the objectives of the thesis. *Golgol-Naele* is chosen, for its relative accessibility by public transport.

Forty percent of the sample is from subsistence farm households while the other 60 percent are from semi-commercialized households. The criterion used to classify the households into subsistence and semi-commercialized group is whether the households have a marketed supply of cash crop or not. During the sampling period (December, 2007), s**ubsistence** farmers in this study are defined as households that did not supply any cash crop to the market while **semi-commercialized** households are households who have had a positive cash crop supply. All households in the PA are known and registered in the Bureau of Agriculture and rural Development of *Atsbi-Wemberta* district. The district agricultural bureau has agricultural development agents (DAs) in each PA including *Golgol-Naele.* The agricultural DAs at *Golgol-Naele*, who regularly observe the households' activities, have a list of the marketing behavior of the households in the PA. The sample is selected from this list by stratifying the households into subsistence and semi-commercialized groups and then randomly choosing from each stratum. Such sample selection allows not only to compare relevant characteristics of the subsistence and the semi-commercialized groups but also to identify the factors affecting the decision to participate and the degree of participation in commercialization.

3.3 Methodology

In this section the conceptual and empirical models used to address the research problem are presented.

3.3.1 Conceptual model

While investigating an economic problem at household level it is customary to assess the kit bags of microeconomic theory and when the problem is as complicated as farm household livelihoods in sub-Saharan Africa (SSA), a special microeconomic approach becomes necessary. The typical farm household in SSA is one on the one hand producing and on the other consuming, purchased and/or own produced good, and partially or not at all connected to product and factor markets. This very existence of market imperfections in the region and practically in a large share of the developing world makes separation of the consumption and production behaviors of farm households fuzzy. The development of farm household models in the agricultural economics literature is one step towards tackling such complexity of farm household economic behaviors.

Although the formal modeling of agricultural households has become popular in the past few decades with significant papers such as Barnum and Squire (1979), and Singh et al. (1986), and various works of Sadoulet, De Janvry, and Fafchamps, there have been some explorations into the peculiar nature of production and consumption decisions of peasant farmers by Chayanov in the early 20th century (Ellis 1993).

Farm household models in the literature are classified into two major categories based on their assumptions about presence of markets for products and factors. The first class, separable household models, assume existence of perfect product and factor markets hence clear separation of household production and consumption decisions (Singh et al. 1986). The second development is the non-separable models where some markets are missing hence,"…. *the household shadow price of at least one production-consumption good is not given exogenously by the market but instead is determined endogenously by the interaction between household demand and supply*" (Löfgren and Robinson 2006).There is also further extensions of the non-separable models incorporating risk and transaction costs into the analysis (De Janvry et al.,1991; Fafchamps 1992; Renkow et al.2004).

For the purpose of this thesis a non-separable farm household model with missing markets for food and labor is adopted below from De Janvry et al.1991.[2]

Basic assumptions and definitions

Consider a typical farm household allocating its land into production of two crops: food crop (d) and cash crop(r) over a single agricultural year time. The household applies two factors of production: labor (*l*) and other variable inputs (v) like fertilizer, and pesticide; and uses two fixed (quasi-fixed) factors land, and capital (farm equipments, etc). The household's consumption is limited to only three goods: food, other purchased goods and services (a), and leisure time. Markets for food and labor are missing hence they are

[2] I used a simplified elaboration in Elisabeth Sadoulet's lecture notes (http://are.berkeley.edu/courses/ARE251/fall2006/Hholds_handt_06.pdf) for this part. The notes are based on de Janvry et al. 1991.

treated as non-tradable goods valued at their shadow prices. The survey area is a food deficit area. All food produced is assumed to be consumed by the household. To reduce complexity of the model food supply (i.e. food purchase supplied from domestic surplus areas or food aid) is neglected. There is no formal labor market with exogenous wages. Farm households use family labor or reciprocity. Leisure consumed equals total time endowment (T) after time worked on farm is deducted, i.e. c_l=T-q_l. Finally, competitive markets exist for cash crops, other variable inputs (v), and purchased consumer goods (a), hence they may be treated as tradable goods with exogenously determined prices.

Definition of variables in the model,

q_d= food crop produced
q_r = cash crop produced with farm gate sale price p_r
q_l= labor used in farm production
q_v= other variable inputs with farm gate price p_v
z^q= fixed factors (land and capital) and producer characteristics
c_d = food crop consumed
c_a = other purchased consumer goods with farm gate price p_a
c_l = leisure consumed
z^h = exogenous household consumption characteristics
l_s =labor time worked off the farm
T =total time endowment
M =exogenous transfers
y =income

The household utility maximization problem is then set as follows with a well behaved utility function, U i.e. twice differentiable, increasing in its arguments, and strictly quasi-concave.

$$\underset{q_d, q_r, q_l, q_v, c_d, c_a, c_l}{Max} \quad U(c_a, c_d, c_l; z^h) \tag{1}$$

Subject to the following constraints:

Cash income constraint

$$p_v q_v + p_a c_a = p_r q_r + M \tag{1.1}$$

Production technology constraint

$$g(q_d, q_r, q_l, q_v; z^q) = 0 \tag{1.2}$$

Non-tradables equilibrium constraint

$$c_d = q_d$$
$$c_l = T - q_l \tag{1.3}$$

Non-negativity constraints,

$$q_d, q_r, ql_, q_v, z^q, c_d, c_a, c_l, z^h, l_s, T, M, y, , p_v, p_r, p_a \geq 0 \tag{1.4}$$

Forming the Lagrangian,

$$\underset{q_d, q_r, q_l, q_v, c_d, c_a, c_l}{Max} \quad L = \{U + \lambda(p_r q_r + M - p_v q_v - p_a c_a) + \varphi g$$
$$+ \mu_d (q_d - c_d)) + \mu_l (T - q_l - c_l)\} \tag{2}$$

Where, λ , φ , μ_d , μ_l are Lagrangian multipliers for the cash income constraint, technology constraint, equilibrium condition for food, and equilibrium condition for labor respectively.

Assuming interior solution except for cash crop production, the first order conditions (FOCs) for the Lagrangian equation are the following;

$$\frac{\partial L}{\partial q_r} = \lambda p_r + \phi g_r^{/} = 0 \quad \textit{Cash crop production} \tag{2.1}$$

$$\frac{\partial L}{\partial q_v} = \lambda p_v + \phi g_v^{/} = 0 \quad \textit{Variable inputs} \tag{2.2}$$

$$\frac{\partial L}{\partial q_d} = \phi g_d^{/} + \mu_d = 0 \quad \textit{Food production} \tag{2.3}$$

$$\frac{\partial L}{\partial q_l} = \phi g_l^{/} + \mu_l = 0 \quad \textit{Agricultural labor input} \tag{2.4}$$

$$\frac{\partial L}{\partial c_a} = U_a^{/} - \lambda p_a = 0 \quad \textit{Purchased consumer goods} \tag{2.5}$$

$$\frac{\partial L}{\partial c_d} = U_d^{/} - \mu_d = 0 \quad \textit{Food consumption} \tag{2.6}$$

$$\frac{\partial L}{\partial c_l} = U_l^{/} - \mu_l = 0 \quad \textit{Leisure consumption} \tag{2.7}$$

$$\frac{\partial L}{\partial \phi} = g = 0 \quad \textit{Techno}\log y \ \textit{constra}\,\text{int} \tag{2.8}$$

$$\frac{\partial L}{\partial \lambda} = p_r q_r + M - p_v q_v - p_a c_a = \quad \textit{Cash income constra}\,\text{int} \tag{2.9}$$

$$\frac{\partial \mathrm{L}}{\partial \mu_{\mathrm{d}}} = q_d - c_d = 0 => q_d = c_d \quad \textit{Food equilibrium} \tag{2.10}$$

$$\frac{\partial L}{\partial \mu_l} = T - q_l - c_l = 0 => T = q_l + c_l \quad \textit{Labor equilibrium} \tag{2.11}$$

The decision prices for both tradable and non-tradable goods can be defined as follows;

$P_d^* = \frac{\mu_d}{\lambda}$ -Shadow price for food

$P_l^* = \frac{\mu_l}{\lambda}$ -Shadow price for labor, and

$P_i^* = \overline{p_i}$ -Exogenously determined market prices for cash crop, variable inputs, and purchased consumer goods

Now that we have defined the (shadow) prices for food and labor, we can write the full income constraint by combining the FOCs of cash income constraint, food equilibrium, and labor equilibrium given above.

$$p_v q_v + p_a c_a + p_d^* c_d + p_l^* c_l = p_r q_r + p_d^* q_d + p_l^* (T - q_l) + M \qquad (3)$$

Therefore, the above FOCs for the Lagrangian can be rewritten as follows in terms of the decision prices p^*.That is optimizing the Lagrangian subject to the new full income constraint in equation (3), and the production technology constraint.

$$\frac{\partial L}{\partial q_i} = \lambda p_i^* + \phi g_i^{/} = 0 \; Where \; i = r, d \qquad Cash\ and food\ crop\ production$$

$$\frac{\partial L}{\partial q_j} = \lambda p_j^* + \phi g_j^{/} = 0 \; Where \; j = l, v \qquad Labor and\ \text{var}iable\ inputs use$$

$$\frac{\partial L}{\partial c_k} = U_k^{/} - \lambda p_k^* = 0 \; Where k = d, a, l \qquad Food, purchased\ goods and\ labor consumption$$

$$\frac{\partial L}{\partial \phi} = g = 0 \qquad Techno\log y\ constra\text{int}$$

$$\frac{\partial L}{\partial \lambda} = p_a^* c_a + p_d^* c_d + p_l^* c_l + p_v^* q_v + p_l^* q_l - p_r^* q_r - p_d^* q_d - p_l^* T - M = 0$$

$Full\ income\ constra\text{int}$

$$\frac{\partial L}{\partial \mu_d} = q_d - c_d = 0 \Rightarrow q_d = c_d \qquad Food\ equilibrim$$

$$\frac{\partial L}{\partial \mu_l} = T - q_l - c_l = 0 \Rightarrow T = q_l + c_l \quad Labor\ equilibrim$$

At this level the reduced cash crop supply function can be derived out of three conditions, that is the production decision, the consumption decision, and the "shadow market" equilibrium conditions for labor and food crop.

First, from the producer problem i.e. maximizing agricultural profit, the household's production (supply functions) and input demand (derived demand functions) can be derived

$$\underset{q_d, q_r, q_v, q_l}{Max} \quad \Pi = p_d^* q_d + p_r^* q_r - p_v^* q_v - p_l^* q_l \quad (4)$$

$$Subject\ to\ \ g(q_d, q_r, q_l, q_v; z^q) = 0 \quad (4.1)$$

The production decision can be written as follows

$$q_i = q_i(p_d^*, p_r^*, p_v^*, p_l^*; z^q) \ \ where\ i = d, r, v, l \Rightarrow output\ \ \sup ply\ and\ \ derived\ \ input\ \ demand \quad (4.2)$$

$$y^* = \Pi^* + p_l^* T + M \Rightarrow\ is\ the\ \ full\ income\ at\ optimum\ profit \quad (4.3)$$

Second, from the consumption decision i.e. utility maximization now with the decision prices p* (vector of decision prices including the defined shadow prices) and the full income y*, the demand functions can be derived.

$$\underset{c_d, c_a, c_l}{Max} \ U(c_d, c_a, c_l; z^h) \quad (5)$$

$$subject\ to\ \ y^* = \Pi^* + p_l^* T + M \quad (5.1)$$

Out of which the demand function is derived as,

$$c_k = c_k(p_d^*, p_a, p_l^*, y^*, z^h) \ \ , where\ \ k = d, a, l \quad (5.2)$$

Third, the equilibrium shadow price for food and labor can be derived by solving the "shadow market" equilibrium as follows.

$$c_d = q_d \Rightarrow c_d(p_d^*, p_a, p_l^*, y^*, z^h) = q_d(p_d^*, p_r^*, p_v^*, p_l^*; z^h) \text{ and} \quad (6.1)$$

$$c_l = T - q_l \Rightarrow c_l(p_d^*, p_a, p_l^*, y^*, z^h) = T - q_l(p_d^*, p_r^*, p_v^*, p_l^*; z^h) \quad (6.2)$$

From which the equilibrium shadow prices are solved as

$$p_j^* = p_j^*(p_r, p_v, p_a, z^q, z^h, T, M) \text{ where } j = d, l \quad (6.3)$$

Substituting p_j^* into (5.2) we get the reduced demand equations for all consumption goods; and by substituting p_j^* into (4.2) we get the reduced supply functions and derived demand equations for the variable inputs.

For the purpose of this thesis only the cash crop supply function is derived below.

$$q_r = q_r(p_r, p_v, p_a, z^q, z^h, T, M) \quad (7)$$

Equation (7) shows that cash crop supply depends not only on production characteristics but also on consumption characteristics plus exogenous factors. While in separable models only production affects consumption through the profit effect, in the non-separable household models the effect is bidirectional. In this equation q_r is the optimum cash crop supplied given prices and exogenous production and consumption shifters.

3.3.2 Empirical model[3]

The variables at the right hand side of equation (7) are the determinants of cash crop supply (commercialization).The first three variables; p_r, p_v, and p_a are prices. However, since the data is collected from one village, market prices can be assumed to be constant for all households. But transaction costs can vary among households. For example a household's membership in a cooperative can raise the price it receives relative to non-members (Bernard et al. 2006).Distance to the road side or market center also can bring

---

[3] For most of this section I used Verbeek, (2004: chapter 7) and Alison Burrell (AEP-50806 Lecture notes, 2007).

variation in the effective prices households get. Vegetables are the dominant cash crops produced by the households. The greater the distance to road the larger the (opportunity) cost of time involved in transportation and the higher the probability of spoilage hence the lower the price they will receive. Thus, fixed transaction cost and proportional transaction cost are included in place of the price variables. The next variable Z^q, represents quasi-fixed production factors and other producer characteristics. These include farm size, soil quality, share of irrigable land, farm equipment, livestock, education of the head, gender of the head, age of the head, and risk attitude of the head. The next variable Z^h, represents exogenous household consumption characteristics such as household size and labor index (number of working-age persons per household size) in the household. The next variable, T is total time endowment. It depends on the total household size and the share of working age persons in the household. The last variable M is exogenous income transfers.

A linear cash crop supply form of equation (7) is estimated to measure household decisions of commercialization. Because of some unobserved values of the dependent variable in the sample (i.e. the marketed quantity of cash crops by those households that do not participate in the cash crop market), a selection model is developed for empirical purpose. This method makes it possible to examine the households' decision to participate in cash crop production and their extent of participation separately. The justification for a selection model is that when the dependent variable is not observed for some of the sampled units and a simple OLS regression is made only on the observed values, the result will be biased estimates (Verbeek, 2004).That is, the effect of the factors that restrain the non-participant households from participation are omitted in the OLS regression.

The decision to participate and degree of participation in cash crop supply can be affected by different variables or by the same variables but with different signs. Therefore, two equations are identified.

The selection equation

The first equation is a binary probit equation that detects whether the household is participating or not and can be specified as follows,

$$S_i^* = x'_{1i}\beta_1 + \varepsilon_{1i} \qquad \varepsilon_{1i} \sim N(0,1) \tag{8}$$

$$S_i = 0 \quad if \qquad S_i^* \leq 0$$

$$S_i = 1 \quad if \qquad S_i^* > 0$$

Where, S_i is a binary variable representing participating or not participating in cash crop supply (i.e. whether $q_{r \leq}$ 0 or q_r >0); x_{1i} is a vector of hypothesized explanatory variables determining the decision of participation; β_1 is vector of coefficients; and ε_{1i} is an error term. The variance of the error term σ_1^2 is normalized to 1 because only S_i is observed; S_i^* is a latent variable.

The outcome equation

The second equation examines the determinants of the degree of participation and is specified as,

$$Q_{ri}^* = x'_{2i}\beta_2 + \varepsilon_{2i} \qquad \varepsilon_{2i} \sim N(0,\sigma_2^2) \tag{9}$$

$$Q_{ri} = Q_{ri}^* \qquad if \qquad S_i = 1$$

$$Q_{ri} = notobserved \qquad if \qquad S_i = 0$$

Where Q_{ri} measures the amount of cash crop supplied by the i[th] household, β_2 is a vector of coefficients, x_{2i} is a vector of explanatory variables determining the degree of participation and ε_{2i} is an error term. For the non-participating households Q_{ri}^* is not observed hence $S_i = 0$. For the participating households Q_{ri}^* is observed hence $S_i = 1$.

The error terms in equations (8) and (9) are assumed to have a bivariate normal joint distribution with zero expected value and variances of σ_1^2 and σ_2^2 and covariance of σ_{12} .That is

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim NID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right)$$

, Where, $\sigma_1^2 = 1$, due to normalization restriction in the selection equation.

3.3.3 Estimation

The selection model can be estimated in two ways: two-step estimation or maximum likelihood estimation methods (Verbeek, 2004).

Two step estimation method

The approach in the two-step technique is: first to obtain estimate of the Inverse Mill's Ratio (IMR) for all observations from the binary probit model, and then to regress the outcome equation by OLS with the IMR included as one of the explanatory variables.

Step one: Selection equation

First a standard probit model of the selection equation i.e. $\Pr(S_i = 1) = \Phi(x_{1i}^{'}\beta_1)$ is estimated using maximum likelihood method. The IMR is then estimated from the estimates and variables of the probit estimation, β_{1i} and x_{1i} (with IMR= $\dfrac{\varphi(x^{'}_{1i}\beta_1)}{\Phi(x^{'}_{1i}\beta_1)}$.

Step two: outcome equation

In step two OLS regression is run on the explanatory variables x_{2i} and the estimated IMR as an additional explanatory variable. What is estimated at this step is the amount of cash crop supply given that a household is participating. The conditional expected value of cash crop supply given that a household is participating, is computed as follows.

$$
\begin{aligned}
E\{Q_{ri} / S_i\} &= x'_{2i}\beta_2 + E\{\varepsilon_{2i} / S_i = 1\} \\
&= x'_{2i}\beta_2 + E\{\varepsilon_{2i} / \varepsilon_{1i} > -x'_{1i}\beta_1\} \\
&= x'_{2i}\beta_2 + \frac{\sigma_{12}}{\sigma^2_1} E\{\varepsilon_{2i} / \varepsilon_{1i} > -x'_{1i}\beta_1\} \\
&= x'_{2i}\beta_2 + \sigma_{12}\frac{\varphi(x'_{1i}\beta_1)}{\Phi(x'_{1i}\beta_1)}
\end{aligned}
$$

$$
E\{Q_{ri} / S_i\} = x'_{2i}\beta_2 + \sigma_{12}\frac{\varphi(x'_{1i}\beta_1)}{\Phi(x'_{1i}\beta_1)} \qquad (10)
$$

The second term of the right hand side of the last equality (10) .i.e., $\frac{\varphi(x'_{1i}\beta_1)}{\Phi(x'_{1i}\beta_1)}$, is what is known as the IMR(inverse Mill's ratio) or Heckman's lambda with a coefficient of $\sigma_{12}(=\rho_{12}\sigma_2)$.The latter is the covariance of the error terms of the two equations. It is evident now to see that what makes OLS estimates biased is the omission of the IMR from the regression. If σ_{12}=0, i.e. if there is no correlation between the disturbance terms of the two equations, there will not be sample selection bias. Therefore, the selection equation will be unimportant, hence OLS is unbiased. However if the two error terms are correlated, a hypothesis in this thesis' case, the selection equation becomes necessary.

The outcome equation for estimation purpose is specified as follows.

$$
Q_{ri} = x'_{2i}\beta_2 + \sigma_{12}\lambda_i + \eta_i \qquad (11)
$$

$$
where,\ \lambda_i = \frac{\varphi(x'_{1i}\beta_1)}{\Phi(x'_{1i}\beta_1)}
$$

η_i is an error term that is independent of x_{2i} and λ_i and is normally distributed.

Maximum likelihood estimation

The maximum likelihood (ML) estimates are obtained by maximizing the following log likelihood function with respect to β , σ_1^2 *and* σ_{12} .

$$LogL_3(\beta,\sigma_1^2,\sigma_{12}) = \sum_{i\in Io}\log p\{S_i = 0\} + \sum_{i\in I1}[\log f(Q_{ri}/S_i = 1) + \log P\{S_i = 1\}] \qquad (12)$$

Where the first part of the right hand side of equation (12) is the binary selection and the other part is the outcome equation.

Both methods are used to estimate the model. The two-step method is widely used in empirical works because it is computationally simple and yields consistent estimates although not efficient. The ML estimates are relatively efficient and consistent (Verbeek, 2004:231-232).Using the ML method also enables log likelihood ratio test for the correlation between the error terms of the selection and the outcome equations. The disadvantage of ML method is that the estimates tend to be biased for small sample size, presence of heteroskedasticity and/or erroneous distributional assumption.

3.3.4 Variables and hypothesis

Dependent variable

The dependent variable estimated is the commercialization index**.** One of the widely used measures of degree of commercialization in the literature is the commercialization index which may take different forms based on the purpose, assumption and nature of data availability. For this thesis the commercialization is measured as an index of sales and total produce. Therefore, the dependent variable(**Q_r)** ,estimated below is **sales-to-output ratio**, which is the Birr[4] value of agricultural output sold (cash crop) divided by the Birr value of total agricultural produce (all crops).A supplementary measure of degree of commercialization, **area allocated to cash crops**, is also estimated and compared.

[4] Birr is Ethiopian currency;1 Birr =0.06 Euro (July 2008)

Cash crops in this thesis are cabbage, onion, garlic, and fababean in descending order of their share of the marketed out put. Cabbage takes the highest share covering nearly 80 per cent of the total marketed output. Onion and garlic take 6 and 5 percent of the marketed output respectively. Pulses like Fababean, despite their long recognition as common cash crops in the study area, cover only 2 percent of the total output sold in the sample. Whereas cabbage, tomato, and garlic are entirely marketed, less than 25 percent of Fababean and onion produces are consumed at home. Crop choice and land allocation are further discussed in section 4.1.

Independent variables

Below are discussed the explanatory variables selected for the empirical model. The selection of determinants of the discrete choice i.e. whether to participate or not and the continuous decision i.e. the extent of participation are based on the transaction cost approach in the literature (discussed in chapter 2), and on previous studies in the study area.

Fixed (lump sum) transaction costs and proportional transaction costs affect the probability of market participation whereas the degree of participation is also affected by the proportional transaction cost. Following this line of argument, information access (radio ownership) is included to capture the impact of fixed transaction cost on the decision to participate. Radio ownership is expected to reduce fixed transaction cost. Two measures of proportional transaction cost, distance to all weather road and cooperative member ship are included to capture the effect of proportional transaction cost on the decision of participation.

Moreover, with the exception of some parts of the western zone large part of the study area is moisture-deficit rain fed agriculture. The low, erratic and unevenly distributed rain fall is one of the main barriers for agricultural production. Hence, it is hypothesized that water scarcity can be a barrier to participation in cash crop supply. Irrigation access is included in the selection equation to capture this (see section 2.4).

Finally, risk attitude of the households is also added in the selection equation (see section 2.4). Risk attitude is proxyed by two alternative measures: one is the household's own rating of risk behavior in farm decision making and the other is an indirect measure of risk attitude (see section 4.1.3).

Radio ownership, irrigation access, distance to all weather road, cooperative membership and risk attitude are included in the selection equation as major determinants of the decision to participate. The rest of the variables in the table are not included in the selection equation with the hypothesis that they will have zero effect on probability of participation. These variables are producer and consumer characteristics. They can affect degree of participation once households decide to participate. But, they cannot affect the probability of participation of households which in this thesis's assumption depends primarily on irrigation access, transaction cost and risk factors.

Once farm households decide to participate, their extent of participation is expected to be affected by production and consumption characteristics listed in Table 1 below together with their hypothesized effect. Some of the selector variables i.e. distance to all weather roads, cooperative membership and risk attitude are also included in the outcome equation assuming that they will also affect the extent of participation. To measure the effect of irrigation on the degree of participation, a percentage of irrigable land is considered to be the proper explanatory variable.

Table 1 Independent variables and hypothesis

Independent variable	Expected Effect: Decision to participate (S)	Expected Effect: Degree of participation (Q_r)	Explanations
Age of head		+	Experience adds to human capital
Head is male		?	Producer characteristic
Head is literate		+	Human capital
Household size		?	Time endowment(higher consumption needs more labor)
Labor index (number of working adults / household size		+	More laborers, less consumers create a larger surplus for market supply
Livestock owned (TLU)		+	Capital
Farm size		+	Capital
Good soil quality (%)		+	Natural Capital
Plot under irrigation (%)		+	Farmers with more irrigation access tend to be more commercial oriented
Have irrigation access	+		
Owns radio	+		Information access reduces fixed TC hence increase probability of participation
Distance to all weather road (in minutes)	-	-	Proportional transaction costs decrease commercialization
Have cooperative membership	+	+	Members of cooperatives receive more favorable prices and enjoy economies of scale in marketing
Head is risk taking	+	+	A risk averse household tends to grow more food crops for self consumption rather than produce for markets

Chapter 4 Results and discussion

4.1 Statistical descriptions

This section presents descriptive statistics of the survey data in a way of characterizing and comparing the subsistence and semi-commercialized households based on consumption characteristics, production characteristics, rural institutions, public goods, and transaction cost and risk attitude.

4.1.1 Consumption and production characteristics

Household characteristics

As can be seen in Table 2, the two groups on average do not show significant variation with regard to age, sex, household size, illiteracy level and marital status. A statistical test shows that at 5 percent level all the mean differences are not significantly different from zero.[5] Therefore, the two groups of households do not show significant variation in their household characteristics.

Table 2 Household characteristics

variable	Sub-sample	N	Mean	Std. Deviation	Std. Error Mean
Sex of head, Female=0, Male=1	semi-commercialized	55	0.82	0.39	0.05
	subsistence	34	0.85	0.36	0.06
Age of head in years	semi-commercialized	40	45.83	13.50	2.14
	subsistence	27	44.74	10.46	2.01
Size of household (number of house hold members older than 1 year)	semi-commercialized	55	5.49	2.16	0.29
	subsistence	34	6.00	2.39	0.41
Head is illiterate (Yes=0,No=1)	semi-commercialized	55	0.40	0.50	0.07
	subsistence	35	0.51	0.51	0.09
Married (Yes=0, No=1)	semi-commercialized	55	0.24	0.43	0.06
	subsistence	35	0.29	0.46	0.08

[5] see Appendix for all independent sample test outputs for this section

Land

The average area allocated to food crops i.e. wheat and barley, jointly by both groups is 0.65 ha per household whereas the average land allotted to all cash crops is 0.37 ha per household including primarily cabbage and fababean; secondarily onion, garlic, tomato and others in descending order of area coverage. Out of 31.8 ha total cultivated area by the semi-commercialized group, about 50 percent is allotted to cash crops. Total area cultivated by the subsistence group is 15.83 ha where only 19 percent of which is allocated to cash crops.

As Table 3 and 4 demonstrate, with regard to input applications (fertilizer, pesticide, and improved seed) and soil quality, the semi-commercialized group takes advantage over the subsistence group. However, the mean differences are not statistically significant from zero at 5 percent level. The significantly visible difference is on irrigation access. As shown in Table 13, the mean difference of percentage of irrigable land between semi-commercialized and subsistence groups (43.8) is significantly different from zero at 1 percent level of significance (t = 8.33). Moreover, mean difference in land productivity i.e. output produced per hectare of cultivated land (2966- 1065=1901 kg/ha) shows significant variation between the sub-samples at 1 percent level favoring the semi-commercialized group.

Table 3 Farm characteristics (percentages)

Sub-sample	Mean cultivated land in ha	Good soil quality in ha	Ferti-lizer applied in ha	Improved seed applied in ha	Irrigable land in ha	Owned land In ha	Rented in land in ha	Share-cropped in land in ha
Semi-commercialized	0.58	(47 %)	(72 %)	(75 %)	(41%)	(66%)	(14%)	(14%)
Subsistence	0.48	(40 %)	(62 %)	(68 %)	(11%)	(83%)	(7%)	(7%)

About 30 percent of the land cultivated by the semi-commercialized group is either rented in or sharecropped in, implying the involvement of this group in some form of land markets. More than 80 percent of the land cultivated by the subsistent group is owned land. Thus, the semi-commercialized group has some kind of incentive to rent in additional land for cultivation. However; it is difficult to interpret such incentive in association with commercialization as the effect of farm size on commercialization is not significant (see section 4.2).

As shown in Table 4 the average cultivated area by the semi-commercialized and subsistence groups are 0.58 ha and 0.48 ha respectively. The mean difference is not significant at 10 percent level. With regard to land input use, both groups have a negligible difference. All such differences are not statistically significant at 10 percent level.

Table 4 Farm characteristics (means)

Area in hectare	Sub-sample	N	Mean	Std. Deviation	Std. Error Mean
Cultivated area	Semi-Commercialized	55	0.58	0.41	0.06
	Subsistence	33	0.48	0.25	0.04
Area with good soil quality	Semi-Commercialized	38	0.40	0.28	0.05
	Subsistence	16	0.38	0.21	0.05
Area fertilizer/pesticide applied	Semi-Commercialized	49	0.47	0.44	0.06
	Subsistence	25	0.38	0.21	0.04
Area improved seed applied	Semi-Commercialized	52	0.47	0.43	0.06
	Subsistence	25	0.44	0.22	0.04
Area partially/fully irrigable	Semi-Commercialized	47	0.27	0.25	0.04
	Subsistence	8	0.23	0.10	0.03

Labor

As illustrated in Table 5, while the subsistence group possesses a slightly greater size of labor force it appears to have less participation in off-farm work. However, the mean difference in labor and off-farm employment (off-farm work in the past six months and off-farm work in the past 12 months) between the two groups is not significantly different from zero at 5 percent level.

Table 5 Labor and off-farm participation

Sub-sample	Mean Number of Working Adults	Mean Days of Off-farm work in six months	Mean Months of Off-farm work in a year
Semi-Commercialized	2.7 (1.0)	30.6 (43)	3.2 (4.3)
Subsistence	3.1 (1.7)	20.1 (31)	2.8 (6)

Note: Standard deviations between brackets

Crop choice

The main crops harvested by the sampled households include wheat, barley, cabbage, and Fababean. Barley and wheat remain the dominant food crops jointly covering more than **three-fourth** of the area cultivated by the subsistence group and nearly **half** of the area cultivated by the semi-commercialized group. As shown in Table 6 the two relatively significant marketed crops are cabbage and Fababean with combined area coverage of 32 percent of the area cultivated by the semi-commercialized sub-sample. Average cash crop produced per household by the semi-commercialized sub-sample is about 1009 kg out of which more than 90 percent is marketed.

Table 6 Crop supply (averages per household)

	Mean output produced in kg per household		Mean output sold in kg per household		Mean cultivated areas in ha per household	
Crop type	Subsistence	Semi-Commercialized	Subsistence	Semi-Commercialized	Subsistence	Semi-Commercialized
barley	208.57	97.27	0	8.18	0.172	0.090
wheat	156.57	195.09	0	19.09	0.178	0.190
All food crops	365.14	292.36	*0*	27.27	0.351	0.280
cabbage	34.29	842.58	0	842.58	0.007	0.137
fababean	27.71	55.36	0	28.00	0.045	0.045
onion	-	84.55	-	66.36	-	0.019
garlic	-	58.76	-	58.76	-	0.038
tomato	-	13.64	-	13.64	-	0.010
All Cash crops	62.00	1054.89	-	1009.35	0.052	0.250
All crops	427.14	1347.25	-	1218.44	0.402	0.530
Total	**989.43**		**74.49**		**0.48**	

Total area cultivated by the subsistence group is 15.83 ha out of which only about 10 percent is allotted to cash crops, fababean taking the largest share. The subsistence group produced an average output of about 62 kg of cash crops per household.

Table 7 Crop supply (percentage sold and output per ha)

Crop type	Percentage of output sold		output produced per hectare in kg/ha	
	Subsistence	Semi-Commercialized	Subsistence	Semi-Commercialized
barley	0	8	1211	1079
wheat	0	10	878	1028
All food crops	***0***	9	1040	1047
cabbage	0	100	4800	6163
Faba bean	-	51	622	1218
onion	-	78	-	4346
garlic	-	100	-	1554
tomato	-	100	-	1339
All Cash crops	-	93	950	3871
All crops	**-**	75	1024	2482
Total	64.0		**2003**	

Farm assets

The average aggregate livestock ownership for the subsistence group is 3.4 in TLU[6], the weighted aggregate farm equipment possession being 2.1 units[7] by the same group. For the semi-commercialized group the aggregate livestock and aggregate farm equipment ownership is 4.5 points in TLU and 2.3 units respectively. The mean difference of the livestock ownership between the semi-commercialized group and subsistence group (1.12 TLU) is significantly different from zero at 5 percent level (t=2.023). Table 8 displays the average count of some of the livestock and farm equipments owned by the households.

[6] The TLU (Tropical Livestock Unit) is calculated average of weights assigned to owned livestock as follows: ox/cow/=1,heifer/bull=0.9,donkey=0.7, calf=0.3, sheep/goat=0.15,and poultry=0.005.

[7] Farm equipment is also measured by a weighted average of farm equipments owned

Table 8 Average assets ownership per household

Sub-sample	Plough tool/Hammer/hoe	Cow/Ox/bull/ heifer	Donkey	Sheep	Stove/ jewelry	Rooms
semi-commercialized	3.5	3	1	8	1	2
Subsistence	3.2	2.2	0.7	6	1.3	1.9

4.1.2 Rural institutions, public goods and transaction cost

A statistical test (see Appendix) shows that at 5 percent level, the two groups have no significant difference in access to rural institutions such as credit institutions, and cooperatives. Both have membership in cooperatives and borrow money at similar rates. Similarly, as Table 10 shows, the two groups have on average equal access to road infrastructure, and extension services such as marketing consultation. All mean differences are not significantly different from zero at 5 percent level.

Table 9 Rural institutions

Sub-sample	Members in a Cooperative	Mean Months of membership	Mean frequency of loans in three years
semi-commercialized	56% N=55	16.2 (6.7)	0.8 (0.55)
subsistence	63% N=35	19.5 (17.2)	0.76 (0.5)

Note: standard deviations in bracket

Table 10 Public goods

Sub-sample	Average Minutes to market center	Average Minutes to All weather road	Average times of Extension visits in a year	Received Marketing consultation
Semi-commercialized	31 (17)	33.3 (18.6)	10.4 (10.3)	36% N=55
subsistence	32.3 (21.5)	32.7 (22.5)	8.9 (18)	18% N=35

Note: standard deviations in bracket

Radio ownership significantly varies (at 1 percent level) across the two groups; the semi-commercialized group is observed to own on average more radios than the subsistence group.

Table 11 Radio ownership

Radio Ownership	Sub-sample	N	Mean	Std. Deviation	Std. Error Mean
owns radio/tape recorder yes=0 No=1	Semi-Commercialized	55	0.24	0.43	0.06
	Subsistence	31	0.58	0.50	0.09

Therefore, the two measures of proportional transaction cost, cooperative membership and distance to road do not significantly vary while fixed transaction cost (measured by radio ownership) significantly varies across the two groups of households. The limited variation in distance to road/market center could probably be because the data was collected from one peasant association.

4.1.3 Risk attitude

Risk attitude of the household was measured in two ways. One is by asking the households to rate their own behavior in overall farm decision making as 0=risk-takers,1=risk-neutral or 2= risk-averse .The other is by providing households with four circumstances where they have to choose what combination of cash and food crop they would decide to grow. The question asked was the following. *Suppose you could grow two types of crops on your plot, a cash crop and a food crop. Suppose by the next harvest time there would be a 50% chance that the price of cash crop will be four times higher than that of the food crop, after covering transportation and other costs, and a 50% chance that it will be half the price of the food crop. Which crop would you decide to grow?* 0= Cash crop only, 1= More cash crop/less food crop,2= More food crop/Less cash crop , or 3= Food crop only. ‘Cash crop only’ shows risk-taking and ‘food crop only’ shows risk-aversity behaviors.

As shown in Table 12 the subsistence group scored 1.7 points that is closer to the risk-aversity scale of 2 in the first measure and a little more than half way towards risk-aversity in the second measure. The difference between the two groups in their risk preference behavior is statistically significant at 5 percent level. On average the semi-commercialized group tends to be less risk averse than the subsistence group.

Table 12 Risk attitude

Sub-sample	Number of Neutral or risk averse household heads	
	Self -stated risk by the head (scaled b/n 0=risk taker, 2=Risk averse)	Indirectly revealed risk attitude of the head (Scaled b/n 0=risk taker,3=risk averse)
Semi-Commercialized	1.4 (0.7)	1.2 (0.8)
subsistence	1.7 (0.5)	1.7 (0.8)

Note: standard deviations in bracket

The variables that show significant variation between the two groups of households are summarized in Table 13.From the statistical tests made to compare the two groups of sampled households, a significant variation is found in irrigation access, risk preference, radio ownership, livestock ownership, and land productivity. Irrigation access generates the highest magnitude of difference between the two groups with an effect size of 0.7(see Table 13).

Table 13 Summary of significant variables and their effect size

variable	Sub-sample	N	Mean	Std. Deviation	Std. Error Mean	Mean difference	Std. Error difference	t-value	Effect size*
Percentage of total land cultivated that is irrigable	Commercialized	51	50.63	31.20	4.37	44	5.2	8.3	0.68
	Subsistence	33	6.8	16.9	2.93				
Self stated risk	Semi-Commercialized	54	1.44	0.66	0.09	-0.27	0.12	-2.2	0.23
	Subsistence	31	1.71	0.46	0.08				
Revealed risk	Semi-Commercialized	54	1.24	0.82	0.11	-0.40	0.19	-2.2	0.23
	Subsistence	31	1.65	0.84	0.15				
Productivity in kg/ha	Semi-Commercialized	55	2966	3333	449	1900.8	462	4.1	0.47
	Subsistence	32	1065	606	107				
Own radio Yes=0, No=1	Semi-Commercialized	55	0.24	0.43	0.06	-0.34	0.11	-3.2	0.40
	Subsistence	31	0.58	0.50	0.09				
Livestock in TLU	Semi-Commercialized	55	4.51	3.36	0.45	1.12	0.56	2.0	0.21
	Subsistence	33	3.39	1.85	0.32				

* Effect size (r = $\sqrt{\frac{t^2}{t^2 + df}}$) is a standardized measure of the size of the observed effect. Practically a bench mark is given as follows: r=0.1=small effect, r=0.3=medium effect and r=0.5 =large effect (Field, 2005).

4.1 Econometric results

The variables used in the econometric estimation are presented in Table 14. The average values and standard deviations of the variables are calculated for the total sample.

Table 14 Descriptive statistics of the variables used in the econometric analyses

Variable	Description	Obs.	Mean	St.dev	Min.	Max.
CIBIRR	Sales-to output ratio in Birr/kg	88	0.45	0.40	0	1
CIAREA	Area allocated to commercial crops(area allotted to commercial crops/total area)	88	0.46	0.27	.02	1
AGEH	Age of the household head in years	90	45.40	10.59	23	75
SEXH	Sex of the head is female =0,or male=1	89	0.83	0.38	0	1
HHSIZE	Number of household members older than one year	89	5.69	2.25	2	11
LABORINDEX	Number of fully working adults divided by HHSIZE	90	0.52	0.17	0.2	1
ILLITRATEH	Head is illiterate yes=0, No=1	87	0.43	0.50	0	1
FARMSIZE	Total Cultivable land(owned+ share cropped+ rented)	88	0.56	0.38	0	2.3
GOODSOIL	Percentage of cultivated land with good soil quality	87	0.98	0.36	0.3	4
LIVESTOCK	Tropical livestock units owned in TLU*	88	4.09	2.92	0	14
IRRIGLAND	Percentage of cultivated land that is irrigable	90	0.34	0.33	0	1
LNROAD	Minutes to an all weather road**	86	3.29	0.71	1.6	4.5
COOPMEM	Have cooperative membership Yes=0, No=1	88	0.41	0.49	0	1
RAVERSE	Head is risk taking Yes=1,No=0	90	0.38	0.49	0	1
IRIGDUMMY	Have irrigation access Yes=o No=1	90	0.34	0.48	0	1
RADIOWN	Household owns radio yes=0 ,or No=1	86	0.36	0.48	0	1

Notes:*The TLU is calculated average of weights assigned to owned livestock as follows: ox/cow/=1 Heifer/Bull=0.9, donkey=0.7, calf=0.3, sheep/goat=0.15 and poultry=0.005.
** **LNROAD** is natural logarithm of distance to all weather roads (in minutes)

4.2.1 Determinants of commercialization

The determinants of the two variables of interest to the analysis, i.e. decision to participate and extent of participation are analyzed in this section using Heckman selection model. Two-step and maximum likelihood estimation methods are used to estimate the factors determining **sales-to-output ratio** and **area allocated to cash crops** The results are presented in Table 15 and 16. The tables can be read as four cases of output i.e. two estimates for each of the two measures of commercialization. In Table 15 there are two-step estimates for sales-to-output ratio and two-step estimates for area allocated to cash crops. In Table 16 there are ML estimates for sales- to-output ratio and ML estimates for area allocated to cash corps.

The Wald ($\chi 2$) is significant at 1 percent level in all four cases. Both estimation methods give more or less consistent estimates of the independent variables in both sales-to-output ratio and area allocated to cash crop equations. A negligible difference is that RADIOWN is not significant in the ML estimates of area allocated to ash crops while in the same equation two other variables becomes significant: LIVESTOCK in the outcome equation and RAVERSE in the selection equation. Moreover, COOPMEM becomes significant in the ML estimates of sales-to-output ratio in the selection equation. RAVERSE is not significant in the selection equation for both estimates but it is significant in the outcome equations.

A test for the equality of the coefficients of the variables common to both selection equation and outcome equation (i.e. COOPMEM, LNWEZEROAD and RAVERSE) shows that the common variables have significantly different effects in the two equations at 10 percent level or below for both sales-to-output ratio and area allocated to commercial crops.

Table 15 Heckman results (Two step method)

Variables	Sales-to-output ratio Coefficient	Sales-to-output ratio Marginal effect	Sales-to-output ratio Elasticity	Area allocated to cash crops Coefficient	Area allocated to cash crops Marginal effect	Area allocated to cash crops Elasticity
Determinants of degree of participation						
AGEH	-0.002	-0.002	-0.119	-0.003	-0.003	-0.445
	(0.004)	(0.004)	(0.259)	(0.007)	(0.007)	(0.994)
SEXH	0.013	0.013	0.018	0.150	0.150	0.434
	(0.091)	(0.091)	(0.121)	(0.164)	(0.164)	(0.491)
HHSIZE	-0.027	-0.027	-0.249	-0.017	-0.017	-0.336
	(0.019)	(0.019)	(0.175)	(0.035)	(0.035)	(0.729)
LABORINDEX	-0.159	-0.159	-0.125	0.126	0.126	0.217
	(0.220)	(0.220)	(0.176)	(0.409)	(0.409)	(0.693)
ILLITRATEH	-0.001	-0.001	-0.001	0.030	0.030	0.043
	(0.072)	(0.072)	(0.048)	(0.136)	(0.136)	(0.197)
FARMSIZE	0.007	0.007	0.006	-0.059	-0.059	-0.114
	(0.112)	(0.112)	(0.099)	(0.220)	(0.220)	(0.422)
SOILGOOD	0.106	0.106	0.131	0.047	0.047	0.126
	(0.110)	(0.110)	(0.135)	(0.211)	(0.211)	(0.566)
LIVESTOCK	-0.003	-0.003	-0.020	0.029	0.029	0.408
	(0.013)	(0.013)	(0.084)	(0.025)	(0.025)	(0.367)
IRRIGLAND	**0.320***	**0.320***	**0.168***	**0.606***	**0.606***	**0.695****
	(0.135)	**(0.135)**	**(0.079)**	**(0.247)**	**(0.247)**	**(0.444)**
LNWEZEROAD	-0.034	-0.034	-0.176	-0.008	-0.008	-0.085
	(0.044)	(0.044)	(0.225)	(0.088)	(0.088)	(0.983)
COOPMEM	0.078	0.078	0.049	0.128	0.128	0.178
	(0.066)	(0.066)	(0.042)	(0.131)	(0.131)	(0.193)
RAVERSE	**0.158***	**0.158***	**0.091***	**0.295***	**0.295***	**0.372****
	(0.068)	**(0.068)**	**(0.043)**	**(0.136)**	**(0.136)**	**(0.242)**
constant	**0.783***			-0.144		
	(0.265)			(0.492)		
Determinants of Decision of participation						
IRIGDUMMY	**-3.209***			**-3.209***		
	(0.611)			**(0.611)**		
RADIOWN	**-0.821***			**-0.821***		
	(0.441)			**(0.441)**		
LNWEZEROAD	0.506			0.506		
	(0.371)			(0.371)		
COOPMEM	-0.793			-0.793		
	(0.539)			(0.539)		
RAVERSE	0.635			0.635		
	(0.554)			(0.554)		
constant	0.260			0.260		
	(1.196)			(1.196)		
Wald $\chi2$ (15) =40.5 **Prob > $\chi2$ = 0.0004** ρ **= 0.463**				**Wald $\chi2$ (15)=31.66** **Prob > $\chi2$ =0.0072** ρ **=1.000**		

Note; *significant at ≤ 5%; ** significant at 10%

Further, the correlation between the selection equation and outcome equations (ρ) is significant and positive for both sales-to-output ratio and area allocated to cash crop (see the LR tests in table 16).Therefore, the selection bias hypothesized earlier in chapter 3 is indeed statistically significant. The positive correlation can be interpreted as follows. The unobserved elements in the error term of the selection equation that are likely to increase the probability of participation in cash crop supply are also likely to increase the degree of participation in the out come equation and vice versa.

Decision to participate

Household decision to participate in cash crop supply is analyzed by the binary probit model, i.e. the bottom of Table 15 and 16. A binary dependent variable was created to capture the decision of participation or no-participation. Then, five selection variables namely irrigation access(IRRIGDUMMY), risk averse head **(**RAVERSE), cooperative membership(COOPMEM),distance to all weather road(LNWEZEROAD) and, radio ownership (RADIOWN) are included to measure the effect of decision to participate. The rest of the variables in the table are not included in the binary probit estimation with the hypothesis that they will have zero effect on the selection equation.

Table 16 Heckman results (ML method)

Variables	Sales-to-output ratio			Area allocated to cash crops		
	Coefficient	Marginal effect	Elasticity	Coefficient	Marginal effect	Elasticity
Determinants of degree of participation						
AGEH	-0.002	-0.002	-0.133	-0.004	**-0.004**	**-0.512**
	(0.004)	(0.004)	(0.264)	(0.003)	**(0.003)**	**(0.359)**
SEXH	0.003	0.003	0.004	0.055	0.055	0.122
	(0.088)	(0.088)	(0.119)	(0.099)	(0.099)	(0.219)
HHSIZE	-0.025	-0.025	-0.233	-0.012	-0.012	-0.189
	(0.018)	(0.018)	(0.173)	(0.021)	(0.021)	(0.318)
LABORINDEX	-0.152	-0.152	-0.123	0.199	0.199	0.261
	(0.211)	(0.211)	(0.172)	(0.179)	(0.179)	(0.235)
ILLITRATEH	-0.004	-0.004	-0.003	-0.017	-0.017	-0.018
	(0.072)	(0.072)	(0.0480	(0.089)	(0.089)	(0.098)
FARMSIZE	0.037	0.037	0.033	-0.113	-0.113	-0.167
	(0.115)	(0.115)	(0.104)	(0.124)	(0.124)	(0.184)
SOILGOOD	0.106	0.106	0.133	0.201	0.201	0.512
	(0.111)	(0.111)	(0.139)	(0.218)	(0.218)	(0.536)
LIVESTOCK	-0.006	-0.006	-0.038	**0.031***	**0.031***	**0.332***
	(0.013)	(0.013)	(0.083)	**(0.014)**	**(0.014)**	**(0.159)**
IRRIGLAND	**0.321***	**0.321***	**0.172***	**0.545***	**0.545***	**0.476***
	(0.131)	**(0.131)**	**(0.077)**	**(0.122)**	**(0.122)**	**(0.140)**
LNWEZEROAD	-0.028	-0.028	-0.149	-0.006	-0.006	-0.052
	(0.045)	(0.045)	(0.233)	(0.052)	(0.052)	(0.444)
COOPMEM	0.064	0.064	0.042	0.086	0.086	0.091
	(0.066)	(0.066)	(0.043)	(0.080)	(0.080)	(0.087)
RAVERSE	**0.146***	**0.146***	**0.086***	**0.182***	**0.182***	**0.175***
	(0.065)	**(0.065)**	**(0.040)**	**(0.076)**	**(0.076)**	**(0.080)**
constant	**0.754***			-0.012		
	(0.253)			(0.228)		
Determinants of decision of participation						
IRIGDUMMY	**-3.461***			**-3.544***		
	(0.635)			**(0.582)**		
RADIOWN	**-0.735****			-0.602		
	(0.511)			0.577		
LNWEZEROAD	0.473			0.537		
	(0.342)			(0.423)		
COOPMEM	**-0.981****			-0.358		
	(0.535)			0.412		
RAVERSE	-0.596			**1.228***		
	(0.641)			**(0.695)**		
constant	1.24(1.4)			-0.65(1.1)		
Log likelihood = -5.4 **Wald $\chi2$ (12) =30.65 Prob > $\chi2$ =0.00** **ρ =0.747 (0.266) LR (ρ =0)** **$\chi2$ (1) =2.10 , Prob > $\chi2$ = 0.15**				**Log likelihood = -3.08** **Wald $\chi2$ (12) =7.76e+06** **Prob > $\chi2$ = 0.00** **ρ =1.000 (0.000) LR(ρ =0)** **$\chi2$ (1) =20.13 , Prob > $\chi2$ = 0.00**		

Note: *significant at ≤ 5% ; **significant at 10%m ; Standard deviations in brackets

The result shows that the probability of participation is significantly affected by irrigation access at 5 percent level followed by radio ownership at 10 percent level .The probability of a household's participation decision is highly dependent on whether the household has more of its land holding accessible to irrigation or not. Other things being equal, the prime barrier to farmers' connection with the market is thus strongly associated with moisture scarcity.

The second variable determining the decision to participate is fixed transaction cost as measured by radio ownership. If a household owns a radio, a proxy for information access, it is significantly likely to be participating in cash crop supply.

Proportional transaction cost, as measured by distance to all weather roads and cooperative membership are not observed to affect the decision to participate in cash crop supply. Cooperative membership appears to have significant effect on sales-to-output ratio (bottom of Table 16 column 2), but since it is not significant in all three other cases it could be neglected. Further, risk attitude does not significantly affect commercialization decisions of the households.

Determinants of degree of participation

The extent of participation is measured by the outcome equation i.e. the top of Table 16 and 16. The results show that the share of land that is irrigated significantly affects the degree of participation. The higher the percentage of irrigable land the household cultivates the higher the proportional increase in cash crop supply and area allocated to commercial crop. In fact, a one percentage point increase in cultivation of irrigable land raises sales-to-output ratio by 0.30 Birr per kg. A one percentage point's rise in irrigable land raises area allocated to cash crop per total area cultivated by 50 to 60 percent.

The other significant determinant of cash crop supply is risk attitude of the household. Risk-taking households have a tendency to allocate proportionally more land to cash crops and also to supply more cash crops to the market than risk-averse households. The full effect of a household's move from a state of risk-aversity to non-risk-aversity is to increase sales-to output ratio by about 0.15 Birr per kg of output produced and to raise area allocated to commercial crop per total cultivated land by 0.2 to 0.3 points.

As can be seen in Table 15 and 16, there is no evidence for significant effects of proportional transaction cost, consumption and production characteristics, except for livestock ownership in the equation for area allocated to cash crop (top of Table 16 column 5).

Summarizing, the most important result of the econometric analysis is that irrigation access and fixed transaction cost affect small farmers' commercialization decision, whereas percentage of irrigable land and household risk preference are significantly important factors in the households' degree of commercialization. There is no statistical evidence for effects of proportional transaction cost and risk attitude on the commercialization decision. Proportional transaction costs have no significant effect on degree of participation either. Further, other producer and consumer characteristics do not bring significant variation in degree of commercialization. In conclusion, from both statistical and econometric analyses there is a consistent result that irrigation access, fixed transaction cost and risk attitude of households are the major determinants of commercialization.

Chapter 5 Summary of findings and policy implications

5.1 Summary of findings

This thesis addresses one of the key agricultural development issues in Ethiopia, agricultural commercialization, by providing household level evidence from the Tigray region. The objective is first to characterize the difference between sampled subsistence and semi-commercialized farm households; and then to identify the household determinants of small farmers' decision to participate in commercialization (cash crop supply) and their degree of participation, using Heckman selection model. The analysis is made on 90 households which are randomly selected from two groups of households i.e. subsistence and semi-commercialized households. The subsistence farmers are households which have not supplied any cash crop to the market during the sampling period (December 2007) whereas the semi-commercialized groups are those who have had a positive supply.

From the statistical comparisons, the two groups of households show significant heterogeneity in their irrigation access, radio ownership, risk attitude, land productivity, and livestock ownership. The semi-commercialized households have relatively larger proportion of irrigable land and they own more radios and livestock than the subsistence households. Moreover, the former cultivate more productive land than subsistence households. The semi-commercialized group also tends to be more risk-taking than subsistence households. However, the subsistence and semi-commercialized households do not show significant heterogeneity in most of their consumption and production characteristics such as age, sex, household size, farm size, input use, soil quality, off-farm work participation, farm assets ownership, and access to rural institutions.

For the econometric analysis, two commercialization indices are constructed, i.e. sales-to–output ratio and area allocated to cash crop production, to measure degree of commercialization. The factors affecting the two variables are estimated using two-step and the maximum likelihood estimation methods.

With regard to the decision to commercialize, the result shows that irrigation access, and fixed transaction costs are significant determinants. Households who have access to irrigation are more likely to be engaged in the production of cash crops for markets than households relying on rain fed agriculture.

Fixed transaction costs also play a significant role. Households with better information access (facing less fixed transaction cost) are more likely to participate in cash crop production. There is no statistical evidence that proportional transaction cost and risk attitude affect the commercialization decision.

With regard to the determinants of degree of commercialization, percentage of irrigated land and risk attitude of the households are significant determinants. Irrigation access affects not only the decision of commercialization but also equally determines degree of commercialization. This puts more emphasis on the role of water resource availability in the issue of commercialization in the semi arid areas such as the Tigray region.

Risk attitude also affects the degree of commercialization significantly. Risk-taking households are observed to have a higher degree of commercialization than risk averse households.

5.2 Policy implications

The most important conclusion that can be drawn from the results of the thesis is that there is a pressing need for more water resource development efforts in the Tigray region. Practically, shortage of rain is not only a barrier to agricultural commercialization but also the root of all the curses of drought and famine in the arid and semiarid areas of the country. The results show more water availability guarantees more cash crop supply. Rain-fed agriculture is least expected to be commercialized. Large-scale irrigation development schemes (e.g., utilizing big rivers) could be costly in the short-run. Moreover, the experience in Ethiopia and other African countries is that such large schemes have not been efficient (Rahmeto, 1999).Therefore, other cost-effective small scale water resource development schemes should be emphasized. This includes reevaluation of the existing water harvesting experience practiced so far. For instance about 44 earthen dams were built in different zones around the Tigray region over the last decade or so and on average 1418 hectares of land is irrigated per year (Hagos, 2007).

The results also imply the need for interventions to minimize transaction cost and risk avoidance. Irrigation access is a necessary but not sufficient condition for effective commercialization. About half of the cultivated land by the semi-commercialized households in the sample, who have relatively more access to irrigation, is still allocated to food crops. Risk and transaction cost are also found to be crucial factors in determining commercialization. Market-oriented production requires information about markets. However, small farmers often face information asymmetry in the factor and product markets which forces them into production for self-sufficiency. Further, risks of price variability keep many farmers autarkic (Sadoulet and de Janvry, 1995). In surplus years most farmers supply their produces causing a glut in the market which yields them depressed prices. When crops fail prices rise sharply as supply shrinks. Such risk of price variability limits many risk averse farmers into food crop production for subsistence. Initiatives like the recently implemented commodity exchange system can have a vital role in addressing these problems.

Broadening the implications to national level, market initiatives like the commodity exchange system can address transaction cost problems primarily for the surplus areas in the southern, central and western regions.However; water resource development should be given priority in the moisture-scarce northern and eastern parts of the country for better commercialization effects.

5.3 Limitations

Finally, there are two drawbacks of the thesis that may be addressed in further works. First, the idea about the sample in the proposal was to include three zones from Tigray region so as to have more village level heterogeneity. Unfortunately, the available budget limited the sample to only one village. Consequently it was not possible to examine the effects of village level variables such as product and factor prices specified in the cash crop supply equation. Moreover the effect of exogenous income is missing. The non-response rate for remittance (a measure of exogenous income) is more than 75 percent. Therefore it is dropped from the model. Second, the thesis assumes that there is a clear line between subsistence and semi-commercialized households and that the subsistence group produces food crops for consumption at home while the commercialized group produces mainly cash crops for markets. Recent review of the literature in the area by Gebre-Madhin et al. (2007), suggests that the subsistence /seller (or food crop/cash crop) dichotomy is a rather strong assumption in that it excludes buying and rebuying (arbitrage) from the analysis.

References

Barnum H.N. and L. Squire (1979).An econometric application of the theory of the farm household. ***Journal of Development Economics***, vol.6, pp.79-102.

Bernard. T, E., Gebre-Medhin and A.S. Tafesse (2006).***Heterogeneous impacts of cooperatives on smallholder commercialization behavior: evidence from Ethiopia***. Draft unpublished paper.

Bruentrup M, and F. Heidhues (2002). ***Subsistence agriculture in development: its role in process of structural change***. Discussion Paper No1.Institute of Agricultural Economics and Social Sciences in the Tropics and Subtropics. Stuttgart: University of Hohenheim.

Chamberline,J, J. Pender and B. Yu (2006).***Development domains for Ethiopia: capturing the geographical context of smallholder development options***. Paper submitted for ESSP(Ethiopian Strategy support Program) policy conference “Bridging, balancing and scaling up: advancing the rural growth agenda in Ethiopia” 6-8 June, Addis Ababa, Ethiopia

CSA (2005). ***National Statistics***. Addis Ababa: Central Statistical Agency of Ethiopia.

De Janvry, A. and E. Sadoulet (2006).Progress in the modeling of rural households' behavior under market failures. In A.de and R. Kanbur (Eds.) ***Poverty, Inequality and Development: Essays in Honor of Erick Thorbecke***.Kluwer publishing, chap. 8.

De Janvry, A M. Fafchamps, and E. Sadoulet (1991).Peasant household behavior with missing markets: Some paradoxes explained. ***The Economic Journal***, Vol.101, No.409, pp.1400-1417.

Ellis Frank (1993). ***Peasant Economics: Farm Households in Agrarian Development***. Cambridge: Cambridge University Press.

Hagos, Eyasu Yazew (2007). ***Development and Management of Irrigated Lands in Tigray, Ethiopia.*** Wageningen:Wageningen University PhD thesis.

Fafchamps (1992).Cash Crop Production, Food Price Volatility, And Rural Market Integration In The Third World. ***American Journal of Agricultural Economics,*** Vol. 74, No. 1. (Feb., 1992), pp.90-99**.**

Field Andy (2005).***Discovering statistics using SPSS,*** 2nd ed., London: Sage publications.

Gabre-Madhin Eleni, Dawit Alemu, and Samson Dejene (2007).***From farmer to Market: smallholder commercialization of food crops in Ethiopia***. Addis Ababa: International food Policy Research Institute-Ethiopian Strategy Support Program (IFPRI_ESSP) working paper (unpublished).

Goetz, S.J. (1992). A Selectivity Model Of Household Food Marketing Behavior In Sub-Saharan Africa. ***American Journal of Agricultural Economics*** 74:444–52

Govereh, J and T.S. Jayne (1999). Effects of Cash Crop Production On Food Crop Productivity In Zimbabwe: Synergies Or Trade-Offs? East Lansing, Michigan: ***Michigan State University international development working paper*** No 74.

ILRI (2004).***Atsbi Wemberta Pilot Learning Site Diagnosis and Program Design***. International Livestock Research Institute, Unpublished.

Jayne T. S. (1994). Do High Food Marketing Costs Constrain Cash Crop Production? Evidence from Zimbabwe. ***Economic Development and Cultural Change***, Vol. 42, No. 2. (Jan.,1994), pp. 387-402.

Key, N., E. Sadoulet, and A. de Janvry (2000). Transaction Costs And Agricultural Household Supply Response. ***American Journal of Agricultural Economics:*** 82:245–59.

MOFED (2005).***Ethiopia: Building on progress: A plan for accelerated and sustained development to end poverty (PASDEP 2005/06-2009/10).***Addis Ababa: Ministry of Finance and economic development.

Löfgren H. and S. Robinson (2006). To Trade Or Not To Trade: Non-Separable Farm Household Models In Partial And General Equilibrium. *Washington: International Food Policy Research Institute, TMD Discussion Paper No. 37.*

Pender J. (2006).***Smallholder commercialization of staple food production***. Unpublished draft paper.

Pingali, P.L and M.W. Rose grant (1995). Agricultural commercialization and Diversification: processes and policies, ***Food policy****.20 (3):171-186.*

Rahmato D. (1999). ***Water resource development in Ethiopia: issues of sustainability and participation***. Addis Ababa: FDS (Forum for Social Studies) discussion papers.

Renkow , M, G. Hallstroma, D. Karanja (2004).Rural infrastructure, transactions costs and market participation in Kenya .***Journal of development economics:*** 73 (2004) 349– 367. Amsterdam: Elsevier Science.

Sadoulet E.and A. de Janvry (1995). ***Quantitative Development Policy Analysis***. Baltimore: The John Hopkins University press.

Singh, I., Squire, L., Strauss, J. (Eds.), (1986). ***Agricultural Household Models: Extensions, Applications, and Policy***. Baltimore: Johns Hopkins University Press.

Timmer, C.P. (1997).Farmers and Markets: the political economy of new paradigms. ***American journal of agricultural economics***, 79(May 1997):621-627.

Vakis, R., E. Sadoulet, and A. de Janvry (2003).***Measuring Transactions Costs from Observed Behavior: Market Choices in Peru.*** Department of Agricultural &Resource economics, UC Berkeley working papers (University of California, Berkeley).

Verbeek, Marno (2004). ***A Guide to Modern Econometrics***, 2nd ed., Chichester, England: John Wiley and Sons.

Von Braun.J. (1994).Introduction and overview. In J. Von Braun and E. Kennedy (Eds), ***Agricultural commercialization, economic development, and nutrition*** Baltimore: Johns Hopkins University Press.

Von Braun. J, H. Bouis and E. Kennedy(1994) Conceptual framework. In J. Van Braun and E. Kennedy (Eds), ***Agricultural commercialization, economic development, and nutrition*** .Baltimore: Johns Hopkins University Press.

Appendix A-SPSS outputs for independent sample tests between commercialized and subsistence groups

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
		Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
sex of head	Equal variances assumed	.732	.395	-.421	87	.675	-.035	.083	-.199	.129
	Equal variances not assumed			-.429	74.314	.669	-.035	.081	-.196	.127
age of head in years	Equal variances assumed	3.857	.054	.352	65	.726	1.084	3.082	-5.071	7.240
	Equal variances not assumed			.370	63.680	.713	1.084	2.934	-4.777	6.946
size of household(older than 1 year)	Equal variances assumed	.245	.622	-1.038	87	.302	-.509	.490	-1.484	.466
	Equal variances not assumed			-1.013	64.712	.315	-.509	.502	-1.512	.494
Illiterate	Equal variances assumed	1.350	.248	-1.058	88	.293	-.11429	.10797	-.32886	.10029
	Equal variances not assumed			-1.052	71.178	.296	-.11429	.10859	-.33079	.10222
Married	Equal variances assumed	1.029	.313	-.518	88	.606	-.04935	.09523	-.23860	.13990
	Equal variances not assumed			-.511	68.948	.611	-.04935	.09667	-.24220	.14350

Group Statistics

	Subsample	N	Mean	Std. Deviation	Std. Error Mean
Total cultivated area ha	Commercialized	55	.578455	.4128580	.0556698
	Subsistence	33	.479773	.2458359	.0427945
Total area with good soil ha	Commercialized	38	.395066	.2793238	.0453123
	Subsistence	16	.384063	.2057038	.0514260
Total area fertilizer/pecticd appliedon ha	Commercialized	49	.470663	.4411440	.0630206
	Subsistence	25	.382500	.2091027	.0418205
Total area imp seed applied on ha	Commercialized	52	.473558	.4312253	.0598002
	Subsistence	25	.437500	.2194928	.0438986
Total area partially/fully irrigable ha	Commercialized	47	.269734	.2472442	.0360643
	Subsistence	8	.225000	.0972846	.0343953
number of fully working adults(1)+working child(.5)	Commercialized	41	2.7317	1.11858	.17469
	Subsistence	29	3.0517	1.65478	.30729
number of days in six month	Commercialized	55	30.58	43.489	5.864
	Subsistence	33	20.06	30.548	5.318
number of months last year	Commercialized	55	9.627	48.9797	6.6044
	Subsistence	33	2.788	6.0558	1.0542

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Upper	Lower
Total cultivated area ha	Equal variances assumed	4.709	.033	1.245	86	.216	.0986818	.0792435	-.0588491	.2562127
	Equal variances not assumed			1.405	86.000	.164	.0986818	.0702175	-.0409059	.2382696
Total area with good soil ha	Equal variances assumed	.382	.539	.142	52	.888	.0110033	.0775547	-.1446217	.1666282
	Equal variances not assumed			.161	38.038	.873	.0110033	.0685407	-.1277457	.1497523
Total area fertilizer/pecticd appliedon ha	Equal variances assumed	3.817	.055	.944	72	.348	.0881633	.0933686	-.0979637	.2742902

	Equal variances not assumed			1.166	71.754	.248	.0881633	.0756343	-.0626198	.2389464
Total area imp seed applied on ha	Equal variances assumed	3.619	.061	.393	75	.695	.0360577	.0916670	-.1465524	.2186678
	Equal variances not assumed			.486	74.687	.628	.0360577	.0741832	-.1117329	.1838483
number of fully working adults(1)+working child(.5)	Equal variances assumed	5.487	.022	-.966	68	.337	-.32002	.33123	-.98097	.34094
	Equal variances not assumed			-.905	45.683	.370	-.32002	.35347	-1.03165	.39162
number of days in six month	Equal variances assumed	1.206	.275	1.220	86	.226	10.521	8.626	-6.627	27.670
	Equal variances not assumed			1.329	83.753	.187	10.521	7.916	-5.222	26.264
number of months last year	Equal variances assumed	1.464	.230	.797	86	.428	6.8394	8.5847	-10.2264	23.9052
	Equal variances not assumed			1.023	56.724	.311	6.8394	6.6880	-6.5546	20.2333

Group Statistics

	Subsample	N	Mean	Std. Deviation	Std. Error Mean
times extension visits in 3 yrs	Commercialized	54	10.35	10.342	1.407
	Subsistence	33	8.94	18.133	3.157
Marketing consultation	Commercialized	52	.63	.486	.067
	Subsistence	28	.82	.390	.074
distance to all weather road	Commercialized	53	33.3019	18.55199	2.54831
	Subsistence	33	32.7273	22.46841	3.91125
distance to seasonal road	Commercialized	14	10.5000	6.41812	1.71532
	Subsistence	9	21.3333	15.84298	5.28099
distance to market center	Commercialized	53	31.0377	17.16320	2.35755
	Subsistence	33	32.2727	21.47276	3.73793
self stated risk	Commercialized	54	1.44	.664	.090
	Subsistence	31	1.71	.461	.083
Revealed risk	Commercialized	54	1.24	.823	.112
	Subsistence	31	1.65	.839	.151
own radio	Commercialized	55	.2364	.42876	.05781
	Subsistence	31	.5806	.50161	.09009
cooperative/network membership	Commercialized	55	.44	.501	.067
	Subsistence	33	.36	.489	.085

months of membership in 3yrs	Commercialized	31	28.68	11.700	2.101
	Subsistence	21	29.76	11.823	2.580
borrowed money in 3 yrs	Commercialized	55	.29	.497	.067
	Subsistence	33	.27	.452	.079
source of credit	Commercialized	41	.44	.838	.131
	Subsistence	24	.42	.830	.169
times borrowed in 3 yrs	Commercialized	41	1.10	.300	.047
	Subsistence	24	1.04	.204	.042
number of separate rooms	Commercialized	54	2.06	1.140	.155
	Subsistence	33	1.91	.765	.133
oxen	Commercialized	55	.82	.841	.113
	Subsistence	33	.67	.595	.104
donkey	Commercialized	55	1.02	.952	.128
	Subsistence	33	.73	.801	.139

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Upper	Lower
times extension visits in 3 yrs	Equal variances assumed	.698	.406	.463	85	.644	1.412	3.049	-4.651	7.476
	Equal variances not assumed			.409	44.914	.685	1.412	3.456	-5.549	8.374
Marketing consultation	Equal variances assumed	16.029	.000	-1.751	78	.084	-.187	.107	-.399	.026
	Equal variances not assumed			-1.870	66.460	.066	-.187	.100	-.386	.013
distance to all weather road	Equal variances assumed	.849	.360	.129	84	.898	.57461	4.46462	-8.30377	9.45299
	Equal variances not assumed			.123	58.452	.902	.57461	4.66816	-8.76819	9.91742
distance to market center	Equal variances assumed	1.821	.181	-.294	84	.769	-1.23499	4.19565	-9.57850	7.10852

	Equal variances not assumed			-.279	56.974	.781	-1.23499	4.41929	-10.08456	7.61457
self stated risk	Equal variances assumed	10.656	.002	-1.967	83	.053	-.265	.135	-.533	.003
	Equal variances not assumed			-2.164	79.827	.033	-.265	.123	-.509	-.021
Revealed risk	Equal variances assumed	1.267	.264	-2.166	83	.033	-.404	.187	-.776	-.033
	Equal variances not assumed			-2.155	61.655	.035	-.404	.188	-.780	-.029
own radio	Equal variances assumed	8.943	.004	-3.361	84	.001	-.34428	.10244	-.54799	-.14057
	Equal variances not assumed			-3.216	54.648	.002	-.34428	.10705	-.55884	-.12972
cooperative/network membership	Equal variances assumed	1.907	.171	.666	86	.507	.073	.109	-.144	.290
	Equal variances not assumed			.670	68.825	.505	.073	.109	-.144	.289
months of memebership in 3yrs	Equal variances assumed	.298	.588	-.327	50	.745	-1.084	3.321	-7.754	5.585
	Equal variances not assumed			-.326	42.784	.746	-1.084	3.328	-7.796	5.627
borrowed money in 3 yrs	Equal variances assumed	.256	.614	.172	86	.864	.018	.106	-.192	.229
	Equal variances not assumed			.176	72.603	.861	.018	.103	-.188	.224
source of credit	Equal variances assumed	.044	.835	.104	63	.917	.022	.215	-.407	.451
	Equal variances not assumed			.104	48.693	.917	.022	.214	-.408	.453
times borrowed in 3 yrs	Equal variances assumed	2.809	.099	.808	63	.422	.056	.069	-.082	.194

	Equal variances not assumed			.891	61.473	.377	.056	.063	-.070	.181
number of separate rooms	Equal variances assumed	5.600	.020	.653	85	.515	.146	.224	-.299	.592
	Equal variances not assumed			.717	84.176	.476	.146	.204	-.260	.553
oxen	Equal variances assumed	5.112	.026	.907	86	.367	.152	.167	-.181	.484
	Equal variances not assumed			.987	83.543	.327	.152	.154	-.154	.457
donkey	Equal variances assumed	.036	.850	1.469	86	.145	.291	.198	-.103	.684
	Equal variances not assumed			1.534	76.619	.129	.291	.190	-.087	.668

Group Statistics

	Subsample	N	Mean	Std. Deviation	Std. Error Mean
percentage of total land cultivated that is irrigable	Commercialized	51	50.627451	31.1999749	4.3688703
	Subsistence	33	6.818182	16.8549966	2.9340783

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Upper	Lower
percentage of total land cultivated that is irrigable	Equal variances assumed	13.024	.001	7.388	82	.000	43.8092692	5.9294566	32.0136912	55.6048472
	Equal variances not assumed			8.325	79.883	.000	43.8092692	5.2626841	33.3359587	54.2825796

Group Statistics

	Subsample	N	Mean	Std. Deviation	Std. Error Mean
Total Livestock in TLU	Commercialized	55	4.5112	3.35548	.45245
	Subsistence	33	3.3882	1.84756	.32162
wighted farm equipments	Commercialized	54	2.3056	1.10495	.15036
	Subsistence	31	2.1290	.77425	.13906
number of separate rooms	Commercialized	54	2.06	1.140	.155
	Subsistence	33	1.91	.765	.133
oxen	Commercialized	55	.82	.841	.113
	Subsistence	33	.67	.595	.104
donkey	Commercialized	55	1.02	.952	.128
	Subsistence	33	.73	.801	.139

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Upper	Lower
Total Livestock in TLU	Equal variances assumed	6.633	.012	1.766	86	.081	1.12300	.63589	-.14111	2.38711
	Equal variances not assumed			2.023	85.515	.046	1.12300	.55511	.01938	2.22662
wighted farm equipments	Equal variances assumed	2.101	.151	.785	83	.435	.17652	.22492	-.27083	.62388
	Equal variances not assumed			.862	79.583	.391	.17652	.20481	-.23109	.58414
number of separate rooms	Equal variances assumed	5.600	.020	.653	85	.515	.146	.224	-.299	.592
	Equal variances not assumed			.717	84.176	.476	.146	.204	-.260	.553
oxen	Equal variances assumed	5.112	.026	.907	86	.367	.152	.167	-.181	.484

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Upper	Lower
	Equal variances not assumed			.987	83.543	.327	.152	.154	-.154	.457
donkey	Equal variances assumed	.036	.850	1.469	86	.145	.291	.198	-.103	.684
	Equal variances not assumed			1.534	76.619	.129	.291	.190	-.087	.668

Group Statistics

	Subsample	N	Mean	Std. Deviation	Std. Error Mean
productivity	Commercialized	55	2966.3301	3332.50230	449.35452
	Subsistence	32	1065.5093	606.19970	107.16198

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Upper	Lower
productivity	Equal variances assumed	10.087	.002	3.189	85	.002	1900.82082	596.13878	715.53723	3086.10442
	Equal variances not assumed			4.115	59.979	.000	1900.82082	461.95581	976.76495	2824.87670

Group Statistics

	Subsample	N	Mean	Std. Deviation	Std. Error Mean
Total Livestock in TLU	Commercialized	55	4.5112	3.35548	.45245
	Subsistence	33	3.3882	1.84756	.32162
distance to all weather	Commercialized	53	33.3019	18.55199	2.54831

road	Subsistence	33	32.7273	22.46841	3.91125
self stated risk	Commercialized	54	1.44	.664	.090
	Subsistence	31	1.71	.461	.083
Revealed risk	Commercialized	54	1.24	.823	.112
	Subsistence	31	1.65	.839	.151
own radio	Commercialized	55	.2364	.42876	.05781
	Subsistence	31	.5806	.50161	.09009
percentage of total land cultivated that is irrigable	Commercialized	51	50.627451	31.1999749	4.3688703
	Subsistence	33	6.818182	16.8549966	2.9340783
productivity	Commercialized	55	2966.3301	3332.50230	449.35452
	Subsistence	32	1065.5093	606.19970	107.16198

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality			
		F	Sig.	t	df	Sig. (2-tailed)	Mear Differen
Total Livestock in TLU	Equal variances assumed	6.633	.012	1.766	86	.081	1.12
	Equal variances not assumed			2.023	85.515	.046	1.12
distance to all weather road	Equal variances assumed	.849	.360	.129	84	.898	.57
	Equal variances not assumed			.123	58.452	.902	.57
self stated risk	Equal variances assumed	10.656	.002	-1.967	83	.053	-.
	Equal variances not assumed			-2.164	79.827	.033	-.
Revealed risk	Equal variances assumed	1.267	.264	-2.166	83	.033	-.
	Equal variances not assumed			-2.155	61.655	.035	-.
own radio	Equal variances assumed	8.943	.004	-3.361	84	.001	-.34
	Equal variances not assumed			-3.216	54.648	.002	-.34
percentage of total land cultivated that is irrigable	Equal variances assumed	13.024	.001	7.388	82	.000	43.8092
	Equal variances not assumed			8.325	79.883	.000	43.8092
productivity	Equal variances assumed	10.087	.002	3.189	85	.002	1900.82
	Equal variances not assumed			4.115	59.979	.000	1900.82

Appendix B-Map of survey area