# Development and Diversification of Small-scale Agriculture in Afghanistan

Hayatullah Ahmadzai

Centre for Research in Economic Development and International Trade (CREDIT)

School of Economics

Thesis

Submitted to the University of Nottingham

for the Degree of Doctor of Philosophy

January 2019

## Abstract

Agriculture is central to the Afghanistan's economy. It continues to be a strategic sector in the economic development of Afghanistan in terms of its potential for contributing to household income, food security, and rural employment. The sector is dominated by resource-poor small-scale subsistence farmers emphasising food security through own production. The recent policy changes to transform the sector from a subsistence to a self-sufficient market-led system have presented opportunities but posed challenges and created uncertainties. Empirical evidence is required to assess the status of the sector, the effect of policies, and to guide future interventions. The self-contained essays in this thesis attempt to address these emerging concerns by analysing small-scale farmers production efficiency, diversification strategies, and market participation decisions.

The first essay investigates farm-level Technical Efficiency (TE) and empirically assesses how adopting Crop Diversification (CD) strategies by farmers affect production efficiency. A parametric Stochastic Frontier (SF) technique is employed to estimate production efficiency as well as identify potential sources of (in)efficiency. Our analysis suggests that substantial inefficiencies exist; there is room to expand farm revenues by more than a quarter by applying improved farm management strategies (such as crop diversification) without having to resort to greater use of production inputs or adopting expensive production technologies. Adopting a diversified portfolio of crop production by farm households significantly improves production efficiency and farm revenues, but the data confirm the low level of diversity in crop production. Production function estimates exhibits Constant Returns-to-Scale (CRS) meaning that doubling production inputs would lead to an equivalent increase in output.

The second essay addresses low diversity in crop production by identifying drivers of diversity in crop production, with emphasis on the allocation of labour between farm and non-farm activities and access to off-farm income. Our findings show that a third of farmers do not diversify, and the majority that do grow only two or three crops. Our empirical results confirm that a significantly lower degree of diversification is found for farm households with higher off-farm income consistent with the hypothesis that allocation of farm labour away to non-farm activities decrease diversity due to negative labour

effects, mainly because the opportunity cost of household labour is higher than the off-farm wages under imperfect markets implying non-separability between households' farm profits and off-farm earnings. Identification through instrumental variables confirms endogeneity in off-farm income revealing that unobserved factors such as risk-aversion behaviour of farmers drives household's decisions to diversify into both non-farm income and crop mix.

The third essay investigates factor market failures by testing separability in the household production and consumption decisions. Estimates of household labour demand rejects separability; labour demand decisions are strongly influenced by preferences and demographic compositions of household (endowments of labour) suggesting that there exist potential market failures in Afghanistan. Our analysis of input market participation reveals that the ownership of information and communication technologies and transport assets by the farm households, better access to roads, and proximity to permanent food markets, increase the likelihood of household's participation in factor markets. Transaction costs have critical implications for household's market participation and possibly causing market failures.

## Acknowledgement

The work presented in this thesis would not have been possible without the guidance of my supervisors, help from friends, and support from my family. I owe my deepest gratitude to my supervisors, Professor Oliver Morrissey and Professor Simon Appleton, for their exceptional guidance and support. I am very grateful for their patience, caring, and for providing me with a friendly atmosphere during my research journey at the University of Nottingham.

I am forever grateful to my wife and daughters for their unconditional love and support, and my parents and other family members back home for encouraging me to pursue my Ph.D. Their inspiration and emotional support have been beyond measure.

I would like to extend my appreciation to the Islamic Development Bank (IDB) for sponsoring my studies at University of Nottingham. I owe sincere and earnest thankfulness to the Central Statistical Organization (CSO) of Afghanistan, and in particular Mr. Gullajan Ahmadzai and Mr. Hayatullah Ahmadzai for their tremendous assistance in providing me with data to facilitate my research.

Special thanks go to all the staff of the School of Economics for their administrative support and creating an extraordinary learning and friendly environment.

Finally, my sincere thanks go to my office colleagues (from room B58) and friends at the School of Economics, University of Nottingham who have offered me advice and support and have made my stay and social life much easier and enjoyable.

*This work is dedicated to my parents, my wife, and my beloved daughters, Sheba and Leema.*

## TABLE OF CONTENTS

LIST OF TABLES

## LIST OF FIGURES

# 1 CHAPTER I: INTRODUCTION

The development of Afghanistan's agrarian-based economy depends on the speed with which agricultural growth is achieved, that ultimately depends on the speed of diversification out of staple food grain production. Growth in productivity and achieving self-sufficiency in basic food crops, and transformation of small-scale agriculture from subsistence-oriented monoculture production system to a self-sufficient market-oriented economy, are the main policies emphasised in tackling rural poverty and sustaining livelihoods and economic growth. Agrarian transition faces many constraints: poor infrastructure, lack of institutional support, inefficient marketing systems; and on-farm challenges such as low diversity in crop production, limited or no access to markets and market information, agriculture credit or extension services. Empirical evidence is necessary to understand the severity and effects of constraints and identify measures to support small-scale producers and guide future policy interventions.

The process of agriculture commercialisation starts with the household's production decisions that involve adding enterprises and high value cash crops (e.g. crops that are primarily geared towards production for the market) to the subsistence production portfolio. Commercialisation of agriculture means more than the selling of agricultural output to markets, it means that product choice and input utilisation decisions are based on the principles of profit maximisation (Pingali, 1997). Thus, diversity in crop production (a shift in production from mono-cropping to a mix cropping farming system) is the first step towards commercialization. Crop diversity is considered as a key farm management strategy in agricultural production due to the opportunities it offers for managing risk and adaptation to heterogeneous production conditions, as well as because of the increased income generation it allows through market participation. Joshi et al. (2006) add that diversifying farms towards vegetable production is more profitable and labour intensive, therefore it fits well in the small farm production systems because smallholders own more family labour. Thus, diversifying production not only augments farm income but also generates employment opportunities in rural areas.

While diversifying production and gains in productivity are essential for successful integration of small farms into markets, complementary services are required to overcome market related problems and increase the efficiency of marketing system to sustain and facilitate a diversified market-oriented production portfolio. Beyond production environment, the agrarian transition entails market orientation (i.e. production decisions based on market signals) and household choices regarding market participation in both input and output markets. Missing markets or imperfect market conditions, high transaction costs, and poor infrastructure are some of the most common barriers that makes it challenging for farmers, particularly smallholders, to be an active part of the agriculture economy and markets (Ouma et al., 2010; Barrett, 2008; de Janvry et al., 1991a). Satyasai and Viswanathan (1997) argues that the transition to commercialisation at the farm level can take place by shifting away from mono-cropping to production of a crop mix that involves cash crops and changing the share of marketed output or purchased inputs per unit of output. Commercialization of agricultural systems leads to greater market orientation of farm production and progressive substitution out of nontraded inputs in favour of purchased inputs; that is commercialization implies that both traded and non-traded inputs are valued in terms of their market value (Pingali and Rosegrant, 1995).

While agricultural commercialization is an imperative strategy for linking small-scale farmers to local markets, research suggests that commercializing smallholder resource-poor farms still presents the best window of opportunity for poverty reduction because these farms are typically operated by the rural poor (Ellis, 2008; Hazell et al., 2007). In this thesis we put greater emphasis on the farm household's crop diversification strategies and their market participation decisions, due to the fact that they are considered as key steps and milestones in transforming agriculture from subsistence towards a market-oriented agriculture economy.

## 1.1 Why is Crop Diversification Important?

In developing countries, income-source diversification is a key livelihood strategy for rural households (Ellis, 1998). This strategy is adopted largely by rural households to increase and sustain their incomes by spreading risk among farm and non-farm activities, and to secure employment to their own farms. Crop diversification as central element of the broader income diversification strategy has also received ample recognition in the literature among agricultural economists, particularly when farming is the main or only source of livelihood for farm households. Based on the World Bank (2014) statistics, agriculture employs about 40 percent of the total workforce in Afghanistan. For households whose primary and possibly only occupation is agriculture, crop diversification becomes an important strategy to deal with volatility and variability off farm income due to production and market risks, and to meet their dietary and consumption requirement. Hence, it is important to stress that the focus in this thesis is on crop diversification. Lack of data on different income sources, especially non-agriculture activities is a limitation that confines the analysis in this thesis to the analysis of diversity in the crop production, instead of the broader subject of income-source diversification.

Farm size is considerably small in Afghanistan and continue to decline due defragmentation, as a consequent productivity is also likely to decline particularly as farming is highly dominated by low-income staple food crops. Oushy (2010) argued that decrease in farm income among rice/wheat producers in Afghanistan due to the declining productivity has triggered a change toward farm diversification. Therefore, according to Oushy (2010), Afghan farmers need to diversify their farming system into mixed crop-livestock systems and shift production from staple crops to higher value commodities. Besides, employment opportunity in the non-farm sectors (services, manufacturing, and industry) especially in the context of rural areas in Afghanistan is inadequate, chances for households to diversify their income sources towards non-farm activities are quite low.

Majority of the Afghan rural population is either illiterate (national literacy rate is about 31%) and another significant proportion do not have sufficient education and skills to seek employment outside of the agriculture sector. Hence, agriculture diversification in general and crop diversification in particular is an important and perhaps the only window of opportunity of employment (especially in the short-run) and income as well as to break out of subsistence, poverty trap, and food insecurity. As evidenced by Woldenhanna and Oskam, (2001) and Barrett et al., 2001, lack of skills as an entry briar hinders farm household's participation in non-farm activities. Abdulai and CroleRees, (2001) established that households with educated heads are more likely to participate in the non-farm sector than those with illiterate heads.

There is a large body of literature that appraises crop diversification as an important strategy for income diversification. Rahman, (2009) recommended that crop diversification should be a desired strategy for agricultural growth in Bangladesh due to its significant efficiency gains. Makate et al., (2016) found that crop diversification improves productivity, income, food security, and nutrition at household level. Birthal et al., (2013) shows that crop diversification significantly reduced poverty in India and suggest that growers need to allocate at least 50% area to the production of high-value crops to escape poverty.

Other studies have demonstrated that household diversification decisions are context specific (Ellis, 2000). Vik and McElwee, (2011) demonstrate that social motivations are as important as economic motivations, that is, there are substantial differences in which motivations underpin different types of diversification. Weltin et al., (2017) showed the decision to diversify economic activities on or off the farm will largely depend on the agricultural business and household characteristics. Among the six different types of farm businesses they recognized, young farm households with organic production are most likely to diversify activities particularly on-farm, whereas farm types characterised by intensive livestock holding and part-time farm households are less likely to apply this strategy. Abdulai and CroleRees, (2001) showed that poorer households have fewer opportunities in non-cropping activities such as livestock rearing and non-farm work, and hence relative lack of capital makes it difficult for them to diversify away from subsistence agriculture. They also indicated that households in remote areas are less likely to participate in the non-cropping sector than their counterparts closer to local market. Ellis, (1998) described that diversification is an infinitely heterogeneous social and economic process, it is differentiated in its causes and effects by location, demography, vulnerability, income level, education and many other factors. Barrett et al., (2005) emphasized the central role that interhousehold heterogeneity in constraints and incentives plays in any sensible explanation of observed income diversification patterns in rural Africa. Based on their argument households that have neither access to non-farm activities nor sufficient productive non-labour assets (i.e., land and livestock) devote themselves entirely to on-farm agricultural production.

Yet, other studies consider the non-economic incentives of crop diversification. Pingali and Rosegrant (1995) indicated that Crop diversification can directly address serious environmental problems by providing a break in the monoculture system and improving crop system health. Lin, (2011) argued that farmers are facing growing stress from climate change, and that the greater implementation of diversified agricultural systems may be a productive way to build resilience into agricultural systems. Makate et al., (2016) established that crop diversification as a viable climate smart agriculture practice significantly enhances crop productivity and consequently resilience in rural smallholder farming systems. Besides, in regions characterised by dry climactic conditions including Afghanistan, diversity in crop production can be an effective copping strategy against draught risks.

In the context of Afghanistan, low productivity due to greater reliance on low-value staple crops is an emerging concern (Oushy, 2010). As a strategic and agricultural policy document, the Afghanistan National Development Framework (ANDF) of 2009 outlines four important pillars as the top priorities for the government: 1) increasing production and productivity of crops and livestock, through provision of better research and extension services and enhanced use of inputs; 2) economic regeneration through development of value chains; 3) natural resource management through natural resource surveillance, protection and conservation, and community management; and 4) institutional development through a process of reform and structural adjustment.

In light of the adaptation of a market-led approach as a strategy for economic transformation, there is an emerging concern about the viability and productivity of small-scale farming, particularly as Afghanistan's agriculture sector is going through transition. Thus, production and marketing decisions made by these smallholders, especially potential productivity and welfare gains that result from choosing a diversified market-oriented production and barriers that come in the way of substituting staple food grains with high-value market-oriented crops need to be empirically assessed.

Afghanistan's economy went under drastic policy changes since emerging out of conflict in 2001, posing uncertainties, risks, as well as opportunity. In the next section, we intend to briefly overview the status and progress of Afghan economy, particularly highlighting major changes that occurred in the agriculture sector.

## 1.2 An Overview of Afghanistan's Agricultural Economy

Afghanistan is a landlocked country within South and Central Asia. It shares the longest border (about 2,430 km) with Pakistan in the south and east. It is also bordered by Iran to the west, Tajikistan, Turkmenistan, and Uzbekistan to the north, and China to the north-east. Administratively, it is divided into 34 provinces which are subdivided into 398 districts. Afghanistan's population was estimated to be 34.65 million with an average annual growth rate of about 2.5% in 2016. More than 80 percent of the country's population, and nearly 90 percent of the poor, live in rural areas, and agriculture plays an important role in their livelihoods and income (World Bank, 2014).

After more than three decades of war and internal civil and political conflicts, Afghanistan continues to face severe economic hurdles and thereby remains one of the world's poorest countries. It's economy still remains heavily reliant on international development aid and grants (Figure 1.1). In 2016, Afghanistan received about $ 1.5 billions dollars (equivalent to $4.06 billions in nominal terms) of real net international aid which forms about 21% of the estimated total real GDP of about $7.6 billion dollars (equivalent to $20 billions dollars in nominal terms). Note that the GDP deflater data from the world indicator database of the World Bank (base year 2001) was used to calculate real values of the international aid, GDP, and GDP per capita.

Afghanistan's economy experienced an average annual growth of 9.6% over 2003 to 2012 (Figure 1.1), mainly driven by the significant amount of international aid but growth in the agriculture output, and revitalization of the industry and services sectors (particularly construction and transportation) was also significant. Real GDP per capita increased from $188 in 2002 to $282 in 2012 (equivalent to $188 to $669 in nominal terms). However, following the withdrawal of international security forces in 2014, the economic growth declined to 2.6% and 1.3% in 2014 and 2015 respectively; the lowest since 2004. Despite the moderate raise in economic growth in 2015/16, it is still considerably below the 9.6 percent average annual rate recorded in the period from 2003 to 2012. The per capita GDP also slightly declined in 2014 but started to raise in 2017 (World Bank, 2017).

Following the economic downturn or perhaps recession, the promising news for the country's economy is that this moderate improvement in growth reflects a recovery from the security transition that began in 2014 with a relatively reduced amount of international assistance being pledged into the country's economy.

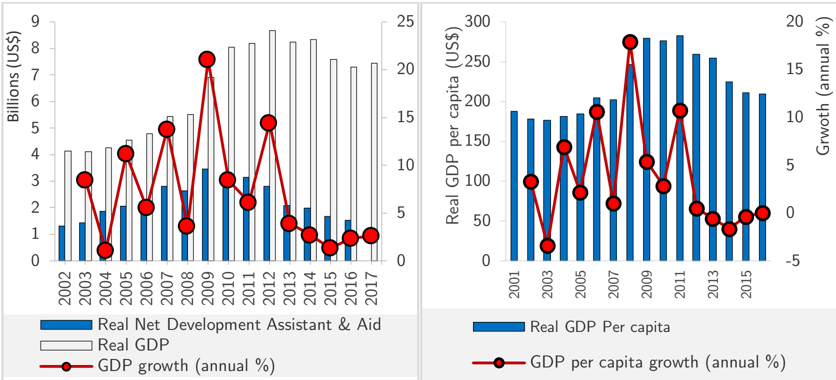



**Figure 1.1: a**) Real GDP, Net Official Development Assistance & Aid Received in Current US$ (LHS), & Annual GDP Growth (RHS), **b**) GDP Per Capita (LHS) & Growth (RHS)

*Source:* World Bank Microdata (World Development Indicators)

Assuming no further deterioration in the security environment, the annual economic growth rate is projected to reach 3.2 percent in 2018 with a growth rate of 2% in industry, 3.3% in services, and 1.5% in agriculture (World Bank, 2017). The expected growth in agriculture is substantially lower than 2016 (6%) primarily due to adverse agro-climatic conditions (e.g. low rainfall and higher temperatures over the recent wet season), leading to low agricultural production.

Among domestic sources, agriculture is central to the country's economy. It continues to be a strategic sector in the economic development of Afghanistan in terms of its potential for contributing to household income, food security, and rural employment. The sector is the second largest contributor to GDP growth after the services sector, that accounts for about a quarter (excluding the opium poppy economy) of the GDP and employs about 40% of the total national workforce (World Bank, 2014). However, with the revitalization of the industry and services sectors, the GDP share of agriculture has declined significantly from about 40% in 2002 to about 25% in 2016 (Figure 1.2). This notable structural change in GDP due to a faster growth in the industry and services sector exhibits a transformation from an agriculture-based economy towards a more industrialized and modernized economy. However, most of the jobs in the industry and services sectors typically require skilled or semi-skilled workers, therefore, despite the structural shifts, the main source of rural employment and growth in the short-term is probably still the agriculture sector.

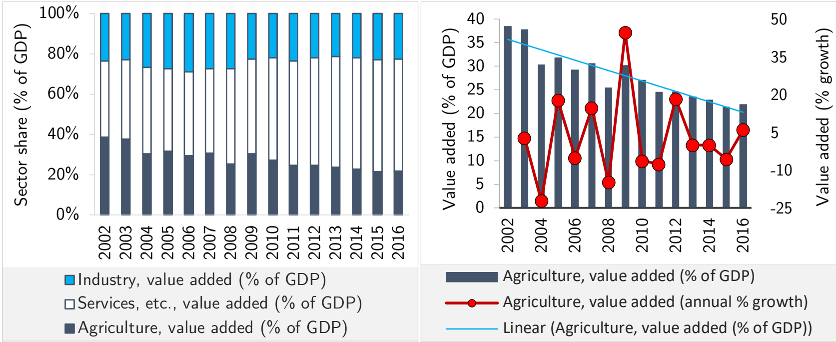



**Figure 1.2:** a) Sector Shares of Total Value (% of GDP), b) Agriculture Value Added - % of GDP (LHS) and Annual Growth in Value Added in Percent (RHS)
*Source:* World Bank Microdata (World Development Indicators)

Despite its substantial share in the country's economy, the performance of the agricultural sector remains poor and the country is not yet self-sufficient in the basic staple food crops (e.g. wheat) with almost 39% of the population living below the national poverty[1] line in 2013/14. The performance of agriculture may be attributed to the productivity gap due to factors such as lack of knowledge on the efficient utilization of available inputs, technical inefficiencies, lack of access to market information, incomplete or missing markets, and limited or no adoption of improved technologies, weak institutional support, particularly inadequate extension services in the remote areas are other obstacles to be addressed by the government.

Afghanistan has a large  trade deficit, with about $6.5 billions of imports and about $0.7 billions of exports. As a landlocked country, Afghanistan trades mainly with its neighbouring countries. Major export destinations include Pakistan, India, Iran, Turkey, UAE, China, and Russia, whereas main import destinations for the Afghan imports are Iran, Pakistan, China, Kazakhstan, Uzbekistan, Turkmenistan, UAE, and India.  In an attempt to reduce the reliance on trade with its immediate neighbouring countries and extend trade with other regional and international countries, Afghanistan recently signed

[1] Poverty rate is predicted to raise even more in 2016/17 due to sluggish economic growth and the deteriorating security situation in the country.

the Chahbahar[2] trade agreement with India and Iran to increase bilateral trade with India through Iran. In addition, Afghanistan became a member of the World Trade Organization (WTO) in July 2016 which will assist Afghanistan to increase its bilateral trade in international markets as well as cross-border regional trade with immediate neighbouring countries. Afghanistan's exports mainly depend on agricultural production as dried nuts and fruits, and other derivatives of agricultural production comprise majority of the exported commodities. Hence, commercialization and diversifying production at the farm level may be an important step in the long run to stabilize food supply and effectively respond to the changing market demand.

Of the total land (65 million ha), 58% (equivalent to 38 million ha) is agricultural land, whereas only about 12% (nearly 8 million ha) is arable (Figure 1.3). Only about 2 million hectares of the arable land is irrigated while the rest (6 million ha) is either rain-fed or left fellow (World Bank, 2014). This clearly indicate that precipitation and irrigation water can be a significant constraint that shrinks the annual cultivated land and affects production levels greatly, hence irrigation infrastructure development is a top priority in the long run to improve agriculture, particularly to bring additional land under cultivation[3].

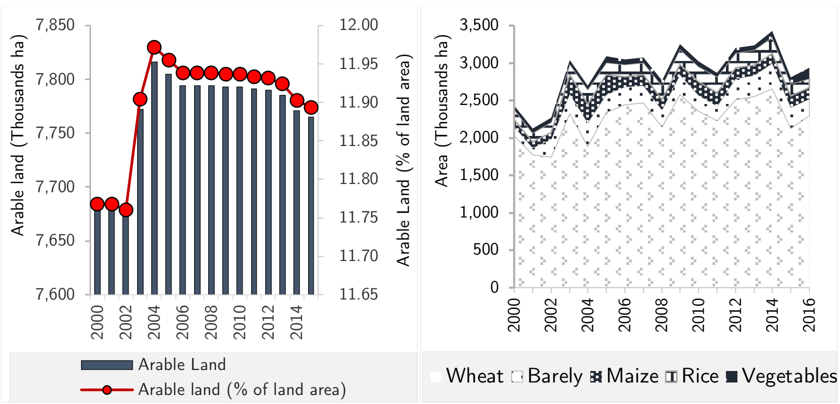



**Figure 1.3:** Arable Land (LHS) and Area under 5 Major Crops (RHS)
*Source:* Food and Agriculture Organization Database (FAOSTAT)

Agriculture is dominated by small-scale farm households with an average farm size of 7 Jeribs (equivalent to about 1.5 hectares) with 60% of farms smaller than 1 hectare and

[2] The Chahbahar trilateral transit agreement signed by India, Iran and Afghanistan in May 2016 that allows Indian goods to reach Afghanistan through Iran.

[3] As per the World Bank's estimates in 2014, Afghanistan could irrigate nearly 3 million ha land before the conflict, if the infrastructure is in place, up to 4.5 million ha land could eventually be brought under irrigation (World Bank, 2014).

nearly 90% smaller than 5 ha (or 25 Jeribs). Wheat dominates the production portfolio and occupies the major portion of the agriculture land (Figure 1.3). As the main staple food crop and a major source of calories, it plays a critical role in food security. With the world's highest annual per capita consumption (160 kg), wheat accounts for 57% of the caloric intake in Afghan diet[4]. Other major crops include maize, rice, barley, fodder crops, potato, and other high value crops such as vegetables and fruits. Although wheat is important for food security and self-sufficiency in terms of grains, market-based production requires farmers to diversify production at the farm level to enhance their cash income.

Greater heterogeneity in agro-climatic conditions allows greater diversity in food production. While Afghan farm households grow a very diverse portfolio of crops, in Figure 1.4, we plot production, yields, and gross value of production of the five major crops. Depending on the climatic condition, particularly the amount of annual precipitation, production varies greatly.

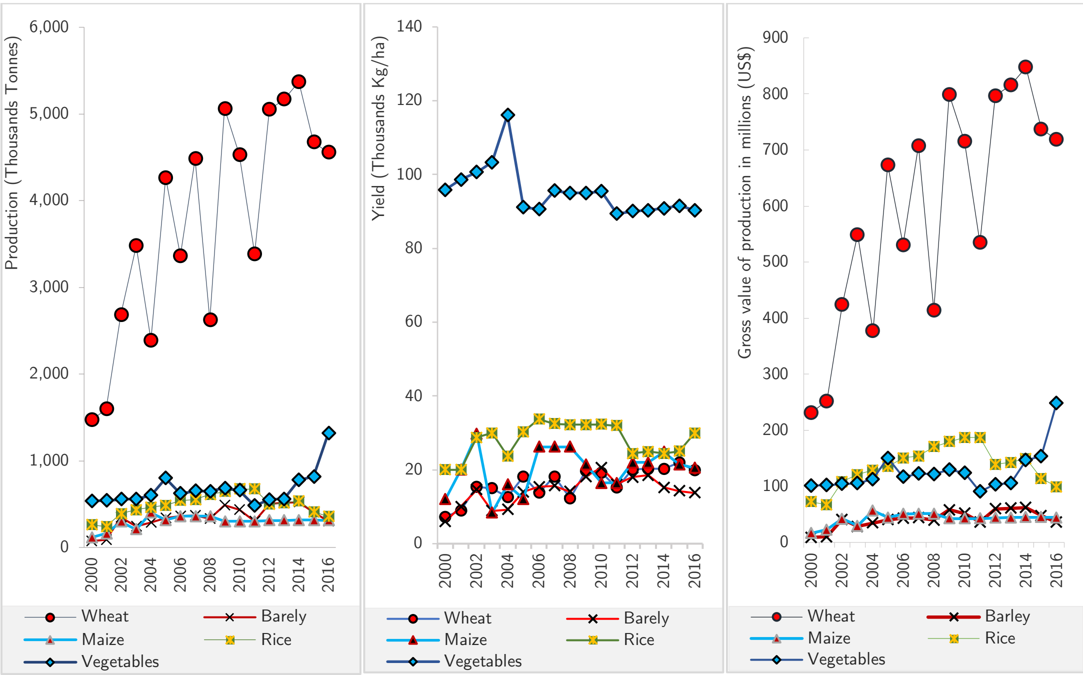



**Figure 1.4:** Five Major Crops: Production (MT), Yield (kg/ha), and Gross Value of Production (Based on 2004-2006 I$)
*Source:* Food and Agriculture Organization Database (FAOSTAT)

[4] At a 160 per capita consumption, Afghanistan with a 34.65 million people would need 5.5 million metric tons of wheat every year to be self-sufficient. However, as per the Figure 1.4, average annual wheat production from 2002 to 2016 is about 4 million MT ranging from 2.5 to 5.3 million MT. Afghanistan's domestic wheat production cannot suffice its own wheat demand, and therefore relays on imports.

Although yields are relatively stable, they increased slightly over the years (Figure 1.4), indicating greater potential and opportunity for economic development which can be achieved through increasing production efficiency and best management practices like crop diversification. Gross value of production, on the other hand, varies in line with production (Figure 1.4). While high value crops occupy relatively less land, it can be noted from Figure 1.4 that yields and gross value of production for vegetable crops are proportionally higher than most cereal crops. This is perhaps an indication of higher profits per unit for high value crops (such as fruits and vegetables) compared to basic staple food crops such as grains which are normally produced for home consumption. While staple food grains are mainly consumed at home, fruits and vegetables are the main cash crops traded in local markets and account for the major portion of exports.

### 1.3 The Data

The household level data used in the econometric analysis in this thesis come from the Afghanistan Living Condition Survey[5] (ALCS) conducted by the Central Organization (CSO) of Afghanistan. CSO has been collecting these data about the country for more than 10 years (previously known as the National Risk and Vulnerability Assessment). The data are mainly used by the government to report on the progress of governance and sector level development indicators. The data include both quantitative survey and in-depth qualitative information on several key indicators including farming and livestock production in Afghanistan

Geographically the survey covers all 34 provinces of the country. In total 35 strata were identified, 34 for the provinces of Afghanistan and one for the nomadic (Kuchi) population. The sampling frame used for the resident population in the recent surveys (2011/12, 2013/14, and 2016/17) was based on the pre-census household listing conducted by CSO in 2003-05, updated in 2009. Households were selected on the basis of a two-stage cluster design within each stratum. In the first stage Enumeration Areas (EAs) were selected as Primary Sampling Units (PSUs) with probability proportional to the EA size. Subsequently, in the second stage ten households were selected as the Ultimate Sampling

[5] ALCS is the only household or farm level data available for Afghanistan. CSO started collecting these data in 2003, and continued data collection in 2005, 2007/08, 2011/12, 2013/14, and more recently 2016/17.

Unit (USU). The design thus provided data collection for an average of 170 clusters (1,700 households) per month and 2,040 clusters (20,400 households) in the full year of data collection. For further discussion on sampling design and strategy, refer to the Afghanistan Living Condition Survey (ALCS) report available on the CSO website (Central Statistics Organization, 2017).

The data are representative at national and provincial level, covering roughly 20,000 households (about 157,262 persons) across the country in every round. The data are disaggregated for residential populations (urban, rural and nomad), and therefore unique in the sense that it also includes the nomadic (Kuchi) population of Afghanistan. Another distinguishing feature of the survey is the continuous data collection during a cycle of 12 months, which captures important seasonal variation in a range of indicators including agriculture. Using a structured household questionnaire, data were collected on a number of indictors including agriculture production, labour market, household assets, education, and other household socio-economic and socio-demographic characteristics.

The ALCS surveys (previously known as the National Risk and Vulnerability Assessment-NRVA) conducted before the 2011/12 round were not used in the analyses presented in this thesis because of the lack of comparability in methodologies, sampling procedure, questionnaires design, and more importantly the unavailability of data on key indicators/variables that are central to the econometric analysis in this thesis. As described in the National Risk and Vulnerability Assessment (NRVA) 2007/08 report, the radically different sampling design and a significant questionnaire revision in 2007 resulting in different measures prohibit any meaningful comparison of the NRVA 2003 and 2005 with the recent survey rounds (i.e. NRVA 2011/12, ALCS 2013/14, and ALCS 2016/17). In addition, the limitation of data collection to three months in the 2003 and 2005 rounds that prevented capturing information on seasonal variation in these two NRVA rounds, was another important factor that limits their comparability with recent surveys.

The National Risk and Vulnerability Assessment Report published by CSO in 2011/12, stated that although to an extent comparability between the 2007/08 wave and recent survey rounds (i.e. NRVA 2011/12, ALCS 2013/14, and ALCS 2016/17) was maintained for a number of key indicators, the sampling design differed between the recent rounds and 2007-08. The pre-census listing of households was updated in 2009 and the sampling frame for the recent rounds is based on this updated listing. In the 2007/08 round, smaller provinces and urban centres were over-sampled. More importantly, for the purpose of analysis conducted in this thesis, the 2007/08 questionnaires did not provide information on key variables such as production inputs. Except for the total land cultivated, the 2007/08 survey did not collect information on the quantities or expenditures on chemical fertilizers and pesticides, seeds, tractor hire, hired labour, expenditures on irrigation water, and other miscellaneous production costs. Information on quantalities of seed used in farming was provided only for two crops wheat and rice. While there was no information on the amount or expenditures of hired labour, the family own household labour coding/classification did not allow to distinguish between the labour that is actually applied to crop farming and labour used in broader agriculture industry (e.g. government extension services and agriculture industry workers).

Comparability between the recent rounds NRVA 2011/12, ALCS 2013/14, and ALCS 2016/17 was maintained as much as possible by a largely similar questionnaire design and content for reported indicators and data collection procedures. The sampling frame for all these three rounds was based on the CSO pre-census listing inartistically developed in 2003-05 but was updated in 2009. At the same time, methodologies were designed as to comply with international survey recommendations and best practices. Experience gained from the previous rounds of NRVA and application of international standards resulted in some methodological changes, but as much as possible the rigour of previous achievements was maintained in these rounds in order to ensure comparability over time.

At the time when the analysis in the first two essays of this thesis were conducted, there were only two rounds of the NRVA/ALCS datasets were available that actually maintained comparability, the 2011/12 and 2013/14 rounds. Among these two waves, the most recent 2013/14 round was chosen for the analysis of the first two essays. However, when 2016/17 dataset released by the CSO in mid-2018, the three datasets (2011/12, 2013/14, 2016/17) were combined to generate a pooled cross-section dataset from the three waves and subsequently the combined dataset was used in the analysis in the third essay.

## 1.4 Structure and Organization:

As economies grow and countries embark on commercial transformation, households' consumption patterns change. This can alter demand from food grains and push up demand for high-value crops in the long run which will subsequently create opportunities for small-scale growers. These emerging opportunities in a changing economic environment need to be assessed for the utility of the farm households. With a general narrative of transition from subsistence to a demand-driven market-led system in mind, this thesis explores the decision-making environment and choices made by resource-poor small-scale farmers. Specifically, the research in this thesis is organized in three self-contained chapters outlined as follow:

Chapter II concentrates on the analysis of the farm-level production efficiency with particular emphasis on the role of crop diversification strategies. The subject of measuring technical efficiency and subsequently the economic performance of the farming sector is important to both the households and policy makers. Their primary concern is to understand how far the output of the agriculture sector can be expected to increase by simply increasing the levels of efficiency, without absorbing further resources and adopting expensive production technologies. Empirical evidence of farmer specific efficiency analysis and identification of the potential factors (such as crop diversification) affecting it, can help address productivity gains basically by improving socio-economic characteristics and farm management practices. Improving efficiencies without increasing the level of inputs usage can lead to saving unnecessary production costs.

Analytically, farm-level Technical Efficiency (TE) is estimated by fitting a translog stochastic frontier model that allows to estimate farm-specific efficiency as well as identify potential sources of (in)efficiency with emphasis on the impact of crop diversification on TE. Crop diversification is measured by constructing the Herfindahl-Hershman and Transformed Herfindahl indices. The econometric analysis allows for potential endogeneity in crop diversification by applying a recent methodology that handle endogeneity issues in parametric stochastic frontier models.

Chapter III focusses on crop diversification and attempts to analyse microeconomic drivers of CD with a particular interest in the impact of non-farm income on the level of on-farm crop diversity. In Afghanistan, farming is characterised by relatively low yields and production inefficiencies due to misallocation of resources or inefficient farm management practices. Moreover, land holdings are mostly small and expected to decline over time (due to fragmentation into smaller farms). Under these circumstances, farmers are often forced to reconsider their decisions about allocation of the scarce farming resources among different crops. Given a demand driven market-led approach, food grain-based production system will no longer be adequate to effectively respond to the changing market demand and consumption patterns. Therefore, the challenging, and yet imperative task ahead for Afghan farm households and policy makers is to improve farm productivity, sustain farm incomes, and safeguard employment of farmers to their own land. This entails a shift from the existing food grain-based system by introducing high value horticultural crops such as vegetables and fruits.

The microeconomic determinants of crop diversification are estimated by applying a censored tobit model that allows to model non-participation as an optimal choice. Household's off-farm income as the main variable of interest in the analysis is potentially endogenous due to omitted unobserved household characteristics, thereby endogeneity bias is accounted for by using Instrumental Variable (IV) techniques. Since Afghanistan has heterogenous climate, it is vital to understand spatial variation of crop diversification and illustrate patterns in crop diversity attributed to geographical location. Hence, further descriptive analyses of the data are carried out to illustrate cropping activities and patterns across different agroecological regions and districts, and land holdings.

Chapter IV goes beyond production decisions and turns to the farm household's marketing decisions. It first aims to test for potential market failures by testing separation in household production and consumption decisions and then examines the implications of market access and transaction costs on the likelihood of households to participate in market, in an attempt to address potential market failures. Promoting market-orientation among farm households and integration into the market economy requires improving the ability of farm households to participate in markets, particularly smallholder resource-poor farmers. The essence of participation in output and input markets is based on the premise that crop yields, incomes and, hence, the livelihoods of smallholder farmers are likely to improve if they gain greater access to markets for inputs and outputs produced.

Transformation of subsistence agriculture to a market-led practice must be based upon the establishment of efficient and well-functioning markets and marketing systems that reduce transaction costs, mitigate risks, reduce search costs and extend information access to all players, particularly those living in rural areas of marginal productivity with poor public infrastructure.

The separability tests are theoretically grounded based on the household labour demand model. Under the assumptions that all current and future markets exist and that households take all prices as given, our analysis allows to model households' simultaneous production and consumption decisions into a recursive form in which production decisions can be made as independent of preferences of the farm household. Using three waves of repeated cross-sectional data from Afghanistan Living Condition Survey (ALCS), the analysis controls for time fixed effects by including a dummy variable for wave as well as location fixed effects by including district dummies. Separation is tested by the joint significance of household's size and composition in the household labour demand model. In an attempt to address potential market failures, we explore whether improving households' access to markets by reducing transaction costs would increase the probability of market participation. Transaction costs are proxied for by including household's ownership of Information and Communication Technologies (ICT), transport assets, time taken to reach the nearest permanent markets, and access to roads in the analysis. Identification through the Control Function (CF) approach and employing instrumental variable (IV) allow for endogeneity in ownership of ICT and transport equipment.

# 2 Chapter II: Crop Diversification and Technical Efficiency in Afghanistan

**Abstract:**

This chapter cantered on assessing the implications of Crop Diversification (CD) strategies on farm level production efficiency in Afghanistan. The empirical evidence from some 7,000 farm households suggest that adopting a diversified portfolio of crop production by the households significantly improves Technical Efficiency (TE) and farm revenues. These findings are particularly crucial as the evidence confirms the presence of a relatively low level of diversity in crop production. With the estimated mean level TE of 72%, our analysis reveals that substantial technical inefficiency exists and that there is an opportunity to expand farm revenues by 28% by applying better farm management strategies (e.g. crop diversification) and without having to resort to greater use of production inputs or the introduction of expensive production technologies. Identification through instrumental variables techniques confirms endogeneity in CD; causing a downward bias in the impact of CD on TE, hence the impact of CD is even greater once endogeneity bias is removed. That is, failing to account for endogeneity in the basic frontier model leads to a downward bias which is consistent with attenuation bias (measurement error in CD implies a bias towards zero, so one would predict IV coefficients greater in absolute size). ML estimation of Translog Stochastic Frontier (SF) model shows that land, labour, and other purchased inputs (fertilizer, seeds, pesticides usage) have positive impact on farm revenues. The results shown evidence of Constant Returns-to-Scale (CRS) meaning that doubling production inputs would lead to an equivalent increase in the output.

## 2.1 Introduction

Measuring economic performance of a farm requires an understanding of the production decisions that influence the levels of production efficiency. Technical Efficiency (TE) as a precondition for economic efficiency safeguards the economic viability and sustainability of a farm. Farm productivity can be improved by adopting technology such as introduction of new machinery, fertilizers and chemicals, and improved seed varieties. Alternatively, productivity can be enhanced by changing how production factors are combined to improve the efficiency by which inputs are being transformed into output such that higher outputs are produced from the same level of inputs and technology (Coelli, 1995). Production and farm management decisions by farmers also affect the level of technical efficiency and overall productivity of a farm, for instance decisions by farmers to shift away from specialization towards adopting a diversified production system.

Empirical research suggests that farmers in developing countries fail to exploit fully the production technology and production resources and often make inefficient decisions. This study attempts to measure farm-specific technical efficiency of crop farmers and identify potential factors determining technical efficiency using Stochastic Frontier Analysis (SFA) techniques. The main interest is to analyse the degree or extent of crop diversification and how it affects the levels of technical efficiency in Afghanistan.

As investment in the farm sector increases production and production efficiency, contributing to economic growth, farmers are likely to switch from subsistence agriculture based on self-sufficiency to profit and income-oriented decision making, henceforth farm output is accordingly more responsive to market trends. This transition from subsistence food production to a commercially oriented system typically involves crop diversification (Minot et al. 2006; Ibrahim et al. 2009; and Nguyen, 2014). Hence, the choice and the extent of crop diversification may depend on the degree of commercialization[6] of the farms (i.e. subsistence, semi-commercial, or fully commercial systems).

[6] Describes the extent of market participation. In substance farming system, production mainly takes place for the household consumption, in semi-commercial system part of the produce is sold to the market and part of it is consumed by the household, and in fully commercial system most of the production takes place for the market.

Farm level technical efficiency requires rational input allocation and improved farm management techniques to achieve the optimum output levels. This is vital for producers who intend to optimize their production decisions particularly under changing market conditions, high input costs, economic hardship and rapid technological progress. It is also relevant for policy makers interested in enhancing the farming sector's economic performance and competitiveness, promoting economic development and sustainable economic practices. Their primary concern is to understand how far the output for the agriculture sector can be expected to increase by simply increasing the levels of efficiency, without absorbing further resources.

Afghan economy was on the verge of collapse due to conflict and political instability and lack of a sound economic policy and inefficiencies of public institutions. However, since emerging out of conflict and establishing a modern economy in 2001, the international aid agencies began to pledge aid to support the economy, particularly agricultural economy, transition towards a fully market-led system began. Many challenges and uncertainties have resulted from policy changes made over the last fifteen years, all of which have influenced farming practices and production decision making in the country's faltering progress towards a market economy.

Against this backdrop and keeping the importance of recent policy changes in mind, it is important to investigate the levels of technical efficiency at the farm level and its determinants in Afghanistan. Identifying determinants of technical efficiency is a major task in efficiency analysis. It is also essential to examine how production decisions by farmers, particularly crop diversification strategies as a major factor, affect the level of technical efficiency.

The contribution of this study is twofold. Although the subject of technical efficiency is important, to the best of the author's knowledge there are no published studies that have investigated technical efficiency at the farm level using nationally representative data that consist of a large sample of households across all eight agro-ecological zones of Afghanistan. Therefore, the analysis in study represent an important contribution in the sense that it is aimed to greatly help rural farm households that depend on farming for their livelihood as well as policy makers in decision making related to production and productivity. Methodologically, in this study we allow for endogeneity in crop diversification using maximum likelihood stochastic frontier analysis. Hence, this study is intended not only to correct for potential endogeneity in CD and estimate its robust effects on the level of technical efficiency but also identify and evaluate the impact of other important factors, such as access to extension services, off-farm employment, agro-ecological zones, and other farm and household socio-economic characteristics.

The remainder of the chapter is structured as follows. Section 2.2 sets out the objective of the analysis presented in the chapter. Section 2.3 overviews relevant literature on production efficiency and crop diversity. Section 2.4 discuses estimation strategy. Section 2.5 describes data and variables used in the analysis. Section 2.6 presents empirical specification of the model. Section 2.7 contains hypothesis and specification tests. Section 2.8 presents econometric results. Section 2.9 contains robustness checks for our empirical results. Section 2.10 concludes the chapter with summarizing the findings of the study and recommendations.

## 2.2 Objective and Research Questions

In being a useful tool to diagnose farm economic problems, assessment of technical efficiency has drawn broad research interest. The assessment of farm level technical efficiency and the factors that affect it provides valuable information to improve farm management and economic performance. Avoiding sources of inefficiency and waste of resources is necessary for economic viability and sustainability in the long run. In this regard, analysing and measuring technical efficiency has important implications for economic performance, commercialization, technological innovation and the overall input use in the farming sector.

This study focuses on the farmers' decision-making processes that are required to incentivize farmers to cultivate a diverse portfolio of crops and reduce dependence on staple crops. The primary intention of this study is to estimate the level of technical efficiency and identify the potential factors determining it by answering the following empirical questions:

a) What are the levels of aggregate farm technical efficiency among Afghan crop farmers?

b) What is the status and extent of crop diversification among smallholder farmers and how does it affect the level of technical efficiency?
c) What are the implications of other important external factors such as access to extension services, agro-ecological zones, off-farm employment, education, and other farm and household characteristics?
d) Are there scale inefficiencies in the farming sector in Afghanistan?

While addressing these questions, it is important to note some of the limitations. The analysis in this study is based on the information generated from the household survey during a single year. Using cross-section data to analyse production decisions makes it difficult to draw concrete policy inferences on the level of technical efficiency that might be subject to change over time. However, a strength of the data is that it covers multiple seasons throughout the same year. A limitation of the data is that information is at the farm level and cannot be disaggregated by plot level or, in the case of inputs, by crop. Therefore, the analysis is limited to the estimation of an aggregate level production function.

The frontier techniques used in this study assume that all inputs required to produce output have been measured and included. However, as with other studies, it is possible to raise questions about whether all inputs have actually been accounted for, since farms that are apparently inefficient may just use less of certain unmeasured inputs. A more general problem is errors or inaccuracies in the measures of inputs, but we assume these are not systematic.

## 2.3 Literature Review

### 2.3.1 Concept and Measures of Technical Efficiency

Since the pioneering work by Farrell, (1957), a number of approaches to efficiency measurement have emerged. The two main approaches that have been extensively used in the efficiency literature are: 1) parametric Stochastic Frontier Analysis (SFA) initially proposed by Aigner et al. (1977) and Meeusen and Van den Broeck (1977); and 2) non-parametric Data Envelopment Analysis (DEA) initially developed by Charnes et al. (1978).

Choosing between the SFA and DEA approaches to measure efficiency has been controversial and depends upon the objective of the research, the type of industry and availability of data (Wadud and White, 2000). The nonparametric approach (DEA) does not rely on the definition of a functional form characterizing the underlying technology and therefore avoids misspecification problems. However, a drawback of this technique is that it is deterministic and ignores the stochastic error term which implies that deviations from the frontier are entirely attributed to inefficiency effects. As a result, technical efficiency ratings obtained from the nonparametric approach are generally lower than those obtained under the parametric SFA alternative (Coelli et al., 2005; Kumbhakar and Lovell, 2000; Wadud and White, 2000).

In contrast, the main advantage of the econometric or parametric SFA approach is that it incorporates a composed error structure with a two-sided symmetric term and a one-sided component which permits to distinguish between inefficiency and exogenous shocks. The one-sided component reflects inefficiency, while the two--sided error captures the random effects  and exogenous shocks outside the control of the production unit, including measurement errors and other statistical noise typical of empirical relationships (Aigner et al., 1977; Meeusen and Van den Broeck, 1977). In addition, SFA allows hypothesis testing and construction of confidence intervals (Wadud and White, 2000). The disadvantages of this approach are the need to assume a functional form for the frontier technology and for the distribution of technical inefficiency term of the composite error term.

This study adopts the stochastic frontier function approach since agricultural crop production exhibits random shocks and there is a need to separate the influence of stochastic factors (random shocks and measurement errors) from the effects of other inefficiency factors by assuming that deviation from the production frontier may not be entirely under the control of farmers.

Production efficiency is widely used in agricultural economics to assess the performance of farmers. Efficiency can be divided into two concepts, the technical efficiency (also called output-oriented efficiency), and allocative efficiency (also referred to as the input-oriented efficiency). Allocative efficiency can be viewed as the ability of a farm to use the inputs in optimal proportions given their respective prices and technology (i.e. obtaining optimal output or profits with the least cost of production). Technical efficiency, on the other hand, is the ability of a farming unit to produce a maximum level of output given the level of inputs (Farrell, 1957). In measuring output-oriented technical efficiency, the inputs are exogenously given, and the objective is to maximize output as the only choice variable.

To illustrate, assume the case where production involves two outputs ($q_1$ and $q_2$) and a single input ($x$) as depicted by Figure 2.1(a). Given the CRS property of the production function and assuming the input ($x$) quantity is fixed, the technology can be represented in two dimensions where the curve $ZZ'$ is the unit production possibility curve. $DD'$ is represents iso-revenue line. Point $A$, located below the possibility curve, corresponds to an inefficient producer because curve $ZZ'$ represents the upper bound of the production possibilities. Alternatively, all points along the production possibility curve represent farmers that are 100 percent technically efficient.

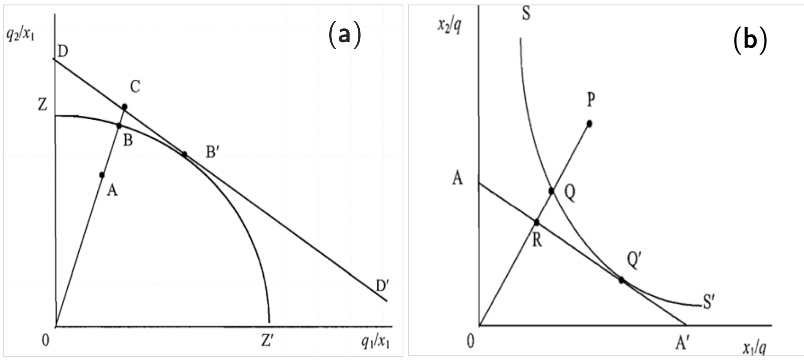



**Figure 2.1:** Technical & Allocative Efficiencies from an Output (a) and Input (b) Orientation

*Source:* Adopted from (Coelli et al., 2005)

In microeconomics of production, technical efficiency is defined as the maximum attainable level of output for given level of inputs, given the current range of alternative technology available to the farmer. In Figure 2.1(a), the distance AB represents technical inefficiency which is the amount by which output could be increased without requiring extra input. Thus, considering a farm producing at point $A$, the Farrell (1957) output-oriented technical efficiency can then be calculated as $TE = 0A/0B$. Allocative efficiency, on the other hand, can be calculated as $AE = 0B/0C$. Both measures have an output/revenue-increasing interpretation (similar to cost-reducing interpretation of allocative inefficiency in the input-oriented case). The overall efficiency could be defined as the product of these two measures $(0A/0B)$ x $(0B/0C)$ or $TE$ x $AE$.

Similarly Figure 2.1(b) illustrates input-oriented efficiency. Assume a farm uses quantities of two inputs ($X_1$and $X_2$) defined by point $P$ (where point $P$ represents an inefficient combination of quantities), to produce an output, the technical efficiency is represented by the distance $QP$ which is the amount by which all inputs could be proportionally reduced without reduction in output to achieve the technically efficient level of production (point $Q$ which is located on the isoquant curve represented by ($SS'$). If the input prices and input price ratio are represented by the slope of the isocost line ($AA'$), the allocative and technical efficiency measures can be calculated as: $AE = 0R/0Q$ and $TE = 0Q/0P$.

Given that this study is concerned with Afghanistan which is a developing country, the main concern maybe output shortfall rather than input over usage, therefore output oriented approach is preferred. Moreover, the lack of price data, particularly in case of inputs, implies that this study will not address allocative inefficiency.

### 2.3.2 Literature on Jointness and Non-jointness of Production Technology

When estimating production functions, one of the most debated issue among agricultural economists is the non-jointess or jointness of the production technology. Shumway et al., (1984) defined that non-jointness occurs if there exist individual production functions for each output or if there exist individual input requirement for each output, whereas jointness do not fit these criteria. Based on these definitions, non-jointness does not imply technical economies or diseconomies of scope whereas jointness brings cost complementarity (often referred to as economies of scope) among outputs, therefore producing them jointly is more inexpensive than producing them separately (Shumway et al., 1984; Leathers, 1991).

One of the distinguishing features of multi-production farms in agriculture is jointness in its production process. However, conventional empirical methods used to estimate production function and technical efficiency at the farm level have generally neglected jointness (or assumed non-joint production) in inputs or outputs. According to Shumway et al. (1984), Just et al. (1983), and Leathers (1991), there are two main sources of jointness in production: jointness is generated by technical interdependencies and allocatable (or quasi-fixed) fixed inputs.

Classic examples of the technical interdependencies include honey bees that pollinate crops or fruit trees, the pest-controlling effects of certain cropping patterns, or the impacts of crop rotations on nutrient balances and soil productivity. Shumway et al. (1984), Moschini

(1989), and Leathers, (1991) consider the presence of multiple outputs competing for an allocatable input that is fixed at the productive unit level (e.g. land) as an additional source of jointness in production. An increase or decrease in the production of one output changes the amount of land available for the supply of the others, thus creating a linkage among the outputs. Peterson, (2002) mentioned that a third type of jointness may arise due to non-allocatable inputs where multiple outputs are produced from the same non-allocatable input, for instance production of mutton and wool or the production of meat and manure which are jointly obtained from raising sheep and cattle. Peterson, (2002) further debated that the existence of non-allocatable inputs in agriculture can give rise to joint production, but not in fixed proportions, the proportions can vary depending on the choice of production technique. Further, the same amount of commodity output could be produced using less land and more labour.

Arnberg and Hansen, (2012) argue that "although the argument of non-jointness seems convincing for some inputs (e.g. fertilizer, pesticides, sowing seed, tractor fuel, etc.), true jointness seems probable for others (e.g., labour and capital). They added that in a short-run model, one might argue that it is reasonable to treat capital and possibly labour as fixed inputs". However, because typically there are important peak utilization capacity constraints around sowing and harvesting, jointness seems probable (i.e. if increasing production of crop 1 will require capital and labour in a peak period then production of other crops must be reduced). Leathers, (1991) demonstrated that while allocatable fixed inputs can cause, it is not true that they will necessarily lead to a joint production. They added that the issue of whether and to what extent allocatable fixed inputs cause joint production is an empirical question. Asunka and Shumway, (1996) stated that the presence of allocatable fixed inputs may cause truly joint technologies to appear non-joint in the short run as well as truly non-joint technologies to appear joint.

Depending on the source, jointness or non-jointness may have different implications on the empirical approach adopted to estimate a production function. Lence and Miller, (1998) discussed that there are three approaches have been used in the literature to circumvent the problem of estimating production functions in the absence of activity-specific input data: (i) the use of a single-equation joint production function summarizing the relationship among aggregate outputs and aggregate inputs, (ii) the use of duality between prices and quantities, and (iii) the use of a primal approach along with the plausible behavioural restrictions. However, all three approaches are plagued with problems and unresolved weaknesses.

The single-equation joint production function approach is very popular in the literature due to the fact that it has often been considered the most general representation of multiproduct-multifactor technologies (Christensen et al., 1973; Vincent et al., 1980; Koundouri et al., 2009). Although it implicitly recognizes that outputs are jointly produced, this approach is restrictive because it restricts each output to depend on all inputs and on all other outputs. It assumes that output is aggregate and the transformation function is separable in outputs and inputs. This assumption implies that the input mix can significantly change without affecting the slope of the production possibility curve (Christensen et al., 1973; Lence and Miller, 1998a). Just et al. (1983) proposes that one possible solution to this problem might be to estimate separate production functions instead of aggregate, with the amount of each input specifically used for producing, however an issue with this approach is that it requires a priori information on how the inputs are allocated across the outputs.

The duality approach to estimating multiproduct-multifactor models in the absence of activity-specific data is perhaps more widely used in practice. However, duality also has serious shortcomings for such purposes. First, the use of dual relationships implicitly assumes that the firm's objective is to maximise profits. Therefore, duality cannot be used when one is interested in testing competing behavioural hypotheses, for instance profit maximization versus expected utility maximisation (Lence and Miller, 1998a). Moreover, Mundlak (1996) debated that there may not be enough price variability to allow identification in cross-sectional analysis. Shumway et al. (1984) demonstrate that the dual method cannot yield allocation equations for fixed but allocatable inputs. Finally, duality is inefficient because it does not use all the information that is often available for econometric estimation (Mundlak, 1996; Lence and Miller, 1998a).

Shumway et al. (1984) criticizes the dual approach and advocate the use of a primal approach that is likely to address the estimation problems in dual methods, particularly when jointness is caused by allocatable fixed factors. This approach improves upon the dual approach by making complete use of available data. However, it shares with the duality approach the limitations of requiring behavioural assumptions, price data for those inputs with missing activity-specific allocations, and data on activity-specific allocations of other inputs (Mundlak, 1996; Lence and Miller, 1998a).

When analyzing multi-product multi-factor production function, a major issue is the lack of activity-specific input data (Lence and Miller, 1998b). While farm households may allocate inputs to specific crop activities, the available records are often limited to the aggregate level of inputs used in all the farm's activities. Just et al. (1983) described that most difficult problem in estimation of nonexperimental agricultural production functions is that input data typically are not available by crop. The difficulty of obtaining data on crop specific inputs partly arises from the joint nature of the two production processes as described above (Lovo, 2011). Using standard surveys, inputs are not usually recorded with sufficient detail because their allocation is affected by seasonality, and often only the total quantities available at household-level can be observed (Lovo, 2011). In the absence of such data, Lovo, (2011) resort to using a DEA approach to estimate household-level technical efficiency for different crops separately. Koundouri et al., (2009) adopts a single-equation joint production function (that summarises the relationship among aggregate outputs and aggregate inputs), due to the limitation of data.

As is often the case with agricultural data sets, the ALCS data used in this study do not provide input data at the crop or output level. In this study we construct an aggregate measure of revenue at the household level and since input and crop specific data are not available, we use the aggregate farm level input data to estimate a joint aggregate household level production function. Furthermore, in our case, it is mathematically complicated to allow for jointness due fixed inputs such as land in the stochastic frontier analysis framework.

### 2.3.3 Diversity in Crop Production

The notion of diversification might have different inferences within the farming sector. Diversification might imply a shift away from monoculture to producing multiple crops on a single farm throughout the season/year. It could also be viewed as having many enterprises at the farm, for instance a larger mix of crops or a combination of livestock and crop units. This study is concerned with the first concept where diversification is defined as adding multiple crops (especially high value crops like vegetables, fruits, potato, etc.) to the planting practice at the farm.

Wheat is a major staple food crop in Afghanistan. National data by the Food and Agriculture Organization (FAO) suggest that wheat occupies more than 50% of the total land cultivated annually. However, the share of wheat in total revenue is proportionally lower compared to other crops, especially high value crops. This difference might indicate that adding high value horticulture crops to the production portfolio may be positively associated with the farmer's income. A study in Punjab of India confirms that incorporating horticultural crops in the production mix increases net expected returns while increasing the labour and working capital requirements (Chhatre et al. 2016). Van den Berg et al. (2007) concluded that diversification into high value vegetable crops would enable Chinese farms to sustain a reasonable income level. Guvele (2001) found that crop diversification reduces variability in income in Sudan.

Crop diversification could be viewed as a hedge against risks due to shocks such as extreme weather conditions, crop diseases and pests, and unexpected fall of market prices. The inherent characteristics of crop diversification that are widely accepted in the literature is that it reduces potential risk against uncertainty by reducing high dependency on monoculture, reduces losses due to diseases, weed and infestation, and increases soil fertility through crop rotation (Krupinsky et al., 2002).

Crop diversification is an environmentally sound and viable climate smart agriculture practice that is widely perceived to significantly enhance farm productivity and increases resilience in rural farming systems. According to Lin (2011) crop diversification improves soil fertility, controls for pests and diseases, and brings about yield stability, nutrition diversity, and health. It can also serve as a superior substitute for the use of chemicals to maintain soil fertility and control pests. Thus, crop diversification is considered as one of the most feasible, cost-effective, and ecologically sound practices that improves farm productivity and increases sustainability and resilience in farming systems.

Nevertheless, there is limited empirical evidence that explicitly studies the impact of crop diversification on technical efficiency, with mixed conclusions. For instance, Nguyen

(2014), Manjunatha et al. (2013), Ogundari (2013), Rahman (2009), and Coelli and Fleming (2004) concluded that crop diversification significantly improves technical efficiency of the farms in Vietnam, India, Bangladesh, Nigeria, and Papa New Guinea, respectively. On the other hand, Haji (2007) found no significance relationship between crop diversification and TE but has found that crop diversification significantly reduced allocative and economic efficiencies in Ethiopia. In addition, Llewelyn and Williams (1996) found that crop diversification significantly reduces technical efficiency in Indonesia. They argued that it is possible that the increased inefficiency with diversification may be transitory as farmers improve their ability to grow new crops as both the age and the Herfindahl index for diversification variables are statistically significant.

Given the mixed empirical evidence presented, it is important to evaluate the impact of crop diversification on technical efficiency especially in the case of Afghanistan where investment in the farm sector has substantially increased to transform farming from a subsistence to a diversified and commercialized system.

### 2.4 Methodology and Estimation Strategy

Since efficiency varies across producers, it is natural to seek determinants of efficiency variation. Early studies adopted a two-stage methodological approach, in which efficiency scores are derived from the estimation of a stochastic frontier function in the first stage, and estimated efficiencies are regressed against a vector of explanatory variables ($Z_i$) using OLS or Tobit regression in the second stage. However, the two-step approach has been criticized on the grounds that the household's knowledge of its level of technical efficiency or exogenous determinants of inefficiency ($Z_i$) might affect its input choices ($X_i$), hence efficiency might be dependent on the explanatory variables (Wang and Schmidt, 2002). Furthermore, even if $X_i$ and $Z_i$ are uncorrelated, ignoring the dependence between them and of the inefficiency with $Z_i$ will cause the first-step technical efficiency index to be underdispersed, so the results of the second-stage regression are likely to be downward biased (Kumbhakar and Wang, 2015).

Kumbhakar and Lovell (2000) and Battese and Coelli (1995) have advocated a single-stage simultaneous estimation approach through SFA in which explanatory variables are incorporated directly into the inefficiency error component. In this approach, either the mean or the variance of the inefficiency error component is hypothesized to be a function of the explanatory variables.

Following Aigner et al. (1977) and Meeusen and Van den Broeck (1977), the formulation of stochastic frontier model in terms of general production function could be specified as:

$$Y_i = f(X_i, \beta) + v_i - u_i = f(X_i, \beta) + \varepsilon_i \tag{2.1}$$

Where $Y_i$ is a scalar of output of the $i^{th}$ farmer, $X_i$ is the vector that collects direct production inputs (i.e. land, labour, fertilizer, etc), and β is a vector of parameters to be estimated. $\varepsilon_i$ is a composed error term where $v_i$ is a two-sided "noise" component assumed to be independently and identically distributed ( $iid$ ), symmetric, and distributed independently from $u_i$. It captures the effects of random shocks beyond the control of farmers (i.e. measurement errors as well as other noise). $u_i$ is a non-negative $(u_i \geq 0)$ technical inefficiency component of the error term that captures the factors that are under the control of the producer (i.e. determinants of inefficiency to be defined in the inefficiency model). $u_i$ is assumed to be independently and identically distributed as normal-half-normal distribution (Aigner et al., 1977). There are other possible specifications of the distributional assumptions on $u_i$ (i.e. truncated-normal distribution) suggested by Greene, (1980) and Lee (1983) which are still being used in empirical work. Jondrow et al. (1982), Battese and Coelli (1995), suggest that the half-normal model is the most common formulation. Other variants such as the truncated-normal model with heterogeneity in the mean allow for great flexibility in the modelling tools.

Since $u_i \geq 0$ , $\varepsilon_i = v_i - u_i$ is not symmetric, and $v_i$, $u_i$ are distributed independently of $X_i$, estimation of Equation (2.1) by Ordinary Least Square (OLS) provides consistent estimates of the parameters except for the constant $(\beta_0)$ since $E(\varepsilon_i) \neq 0$ (Kumbhakar and Wang, 2015). Further, OLS does not provide estimates for the farm-specific technical efficiency. In addition to obtaining estimates of the production technology parameters $(\beta' s)$ from $(X_i, \beta)$, the farmer-specific inefficiency $u_i$ and factors affecting it are the ultimate objectives of the efficiency estimation techniques. To estimate the farmer-specific efficiency, it is required that separate estimates of statistical noise $v_i$ and technical inefficiency $u_i$ are extracted from $\varepsilon_i$ for each producer.

In Equation (2.1) the inefficiency component $(u_i)$ of the error term is the log difference between the maximum (potential) and the actual output (i.e. $u_i = ln\, Y_i^* - ln\ Y_i$ ), therefore $u_i$ x 100% is the percentage by which actual output can be increased using the same inputs if production is fully efficient (Kumbhakar and Wang, 2015). In other words, it is the percentage of output that is lost due to technical inefficiency. The estimated value of $u_i$ is referred to as the output-oriented (technical) inefficiency, with a value close to 0 implying fully efficient. Rearranging Equation (2.1), we can derive the following equation for technical efficiency:

$$TE_i = \exp(-u_i) = \frac{Y_i}{Y_i^*} = \frac{Y_i}{f(x_i\ ;\ \beta)\ exp\{v_i\}} \qquad 2.2$$

Which defines the farm-specific technical efficiency as the ratio of observed output $(Y_i)$ to the frontier output $f(x_i\ ;\ \beta)\ exp\{v_i\}$ which is a maximum output feasible (under the current technology used) in an environment characterized by the stochastic elements specified by $(v_i)$. Because $u_i \geq 0$ , the ratio is bounded between 0 and 1, therefore a farm achieves maximum efficiency if, and only if, $TE_i = 1$. Otherwise $TE_i \leq 1$ is a shortfall of observed output from the maximum feasible output in an environment characterized by $v_i$ that is stochastic and varies across farmers (Kumbhakar and Lovell, 2000).

Using the conditional mean function, Jondrow et al. (1982) showed estimation of observation-specific technical efficiency $(u_i)$ conditional on the error term $(\varepsilon_i)$ as:

$$TE_i = \mathrm{E}(-u_i\,|\varepsilon_i) = \sigma^* \left[\frac{f^*(\varepsilon_i\,\lambda/\sigma)}{1-F^*(\varepsilon_i\,\lambda/\sigma)} - \frac{\varepsilon_i\lambda}{\sigma}\right] \qquad 2.3$$

Where $\sigma^{*2} = \sigma_u^2\sigma_v^2/\sigma^2$, $\lambda = \sigma_u/\sigma_v$, $f(.)$ is the standard normal density function, and $F^{\cdot}(.)$ is the distribution function, both functions being estimated at $\varepsilon\lambda/\sigma$. TE can be obtained by the method of Maximum Likelihood Estimation (MLE) which will simultaneously produce estimates of the variance parameters.

Using $\lambda$ parameterization, the log likelihood function for the (Aigner et al., 1977) model specified in Equation (2.1) assuming a half-normal distribution on $u_i$ is given as:

$$\ln(L) = -\left(\frac{N}{2}\right)(ln2\pi + ln\sigma^2) + \sum_{i=1}^{N}\left[\ln\phi\left[-\epsilon_i\,{}^{\lambda}/_{\sigma}\right] - \frac{1}{2}\,({}^{\varepsilon_i}/_{\sigma})^2\right] \qquad 2.4$$

Where $\sigma^2 = \sigma_u^2 + \sigma_v^2$ and $\lambda = \sigma_u/\sigma_v$ are the variance parameters which measures the fitness and correctness of the model. The variance parameter σ² indicates whether conventional production function would be a satisfactory representation of the data used or not. The ratio of standard errors $\lambda$, is an indicator of relative variability of the sources of variation (i.e. inefficiency and statistical noise). A value of $\lambda > 1$ implies that the discrepancy between the observed and maximum attainable level of output is dominated by variability emanating from technical inefficiency. A detailed mathematical derivation of Equations (2.3) and (2.4) are presented in Appendix 2.A at the end of this chapter.

Battese and Corra (1977) used the gamma parameterization in formulating the likelihood function, instead of $\lambda$. They argued that $\lambda$ could take any non-negative value, thus the gamma parameterization has an advantage in the numerical maximization process as it takes value between 0 and 1 and therefore it searches if the maximizing value are conveniently restricted to this (tight) parameter space. The log likelihood function for Equation (2.1) using gamma parameterization by (Battese and Corra, 1977) is given by:

$$ln(L) = -\left(\frac{N}{2}\right)\left(ln\left(\frac{\pi}{2}\right)\right) + ln\sigma^2\ ) + \sum_{i=1}^{N} ln\left[1 - \phi\left(\frac{\epsilon_i\sqrt{\gamma}}{\sigma^2}\sqrt{\frac{\gamma}{1-\gamma}}\right)\right] - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\epsilon_i^2 \qquad 2.5$$

Where $\sigma^2 = \sigma_u^2 + \sigma_v^2$ and $\gamma = \sigma_u^2/\sigma^2$ are the variance parameters. The gamma parameter could be used to test the presence of inefficiency in the model because it measures relative proportion of variability due to inefficiency ($u_i$) in total variability. In other words, it shows the percentage of the variation in output that is due to technical efficiency, ranging from 0 to 1, where a value close to 1 implies that a random component of the inefficiency significantly contributes to the production system (Battese and Corra, 1977; T. J. Coelli, 1995).

### 2.5 Description of Data and Variables

This study uses data from the Afghanistan Living Condition Survey (ALCS) conducted by the Central Statistics Organization (CSO) in 2013-14. Geographically the survey covered all 34 provinces of the country (Figure 2.2). In total 35 strata were identified, 34 for the provinces of Afghanistan and one for the nomadic (Kuchi) population. The data are representative at national and provincial level. It covered 20,786 households and 157,262 persons across the country. The data are unique in the sense that it also includes the nomadic (Kuchi) population of Afghanistan. Stratification by season was achieved by equally distributing data collection over 12 months within provinces which captures important seasonal variation in a range of indicators including agriculture. Using a structured questionnaire, data were collected on a number of indictors including agriculture production, labour market, household assets, education, and other household characteristics.

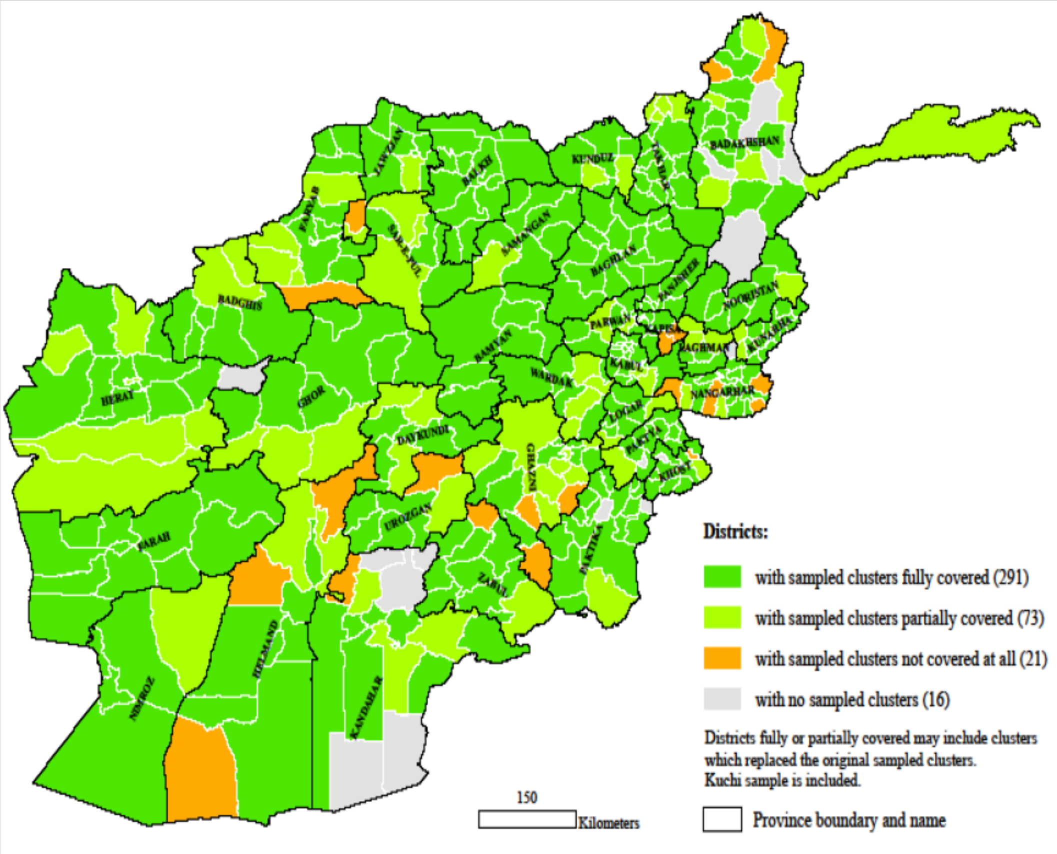



**Figure 2.2:** ALCS Coverage by Districts
*Source:* ALCS 2013/14 Survey Report

The reality of survey taking in Afghanistan imposed a number of deviations from the sampling design. In view of sustained levels of insecurity, clusters in inaccessible areas were replaced by clusters drawn from a reserve sampling frame that excluded insecure districts. In 182 out of 2,100 clusters (8.7 percent), originally sampled clusters could not be covered, in most cases due to security reasons. For a total of 182 clusters the coverage shifted in time or replacement clusters were selected. In addition, 19 clusters, representing 190 households, were not implemented and not replaced. Non-response within clusters was very limited. Only 845 (4.1 percent) of the households in the visited clusters were not available or refused or were unable to participate. In 841 of these non-response cases, households were replaced by reserve households listed in the cluster reserve list, leaving 4 (0.02 percent) households unaccounted for.

A limitation of the data, particularly for the purpose of this study is that the data could not be disaggregated by plot and crop. Therefore, the analysis is restricted to the estimation of an aggregate production function. Initial descriptive analysis of the data showed that as many as 9,642 households reported some involvement in agriculture. However, after accounting for missing values on key variables, the total number of usable observations was 7,052 households.

Referring to Equation (2.1), the dependent variable is aggregate physical output of crops weighted by prices of the respective crops. The dependent variable is the sum of revenue (measured in Afghan currency) of individual crops aggregated throughout the year for each farm household. Summary statistics of the dependent variable are presented in Table 2.1. The dependent variable was checked for potential outliers, and there seems to be no extreme values that influence the results. The price data used to weight physical output comes from the NRVA 2011-12 survey. Lack of price data on some crops and unavailability of price data at the same year in which the ALCS survey was conducted is a limitation. However, for the purpose of this study, the price data were only used to weight physical quantities of crops and to calculate annual aggregate revenues in Afghan currency (Afghani).

The input variables in the frontier and the variables in the inefficiency model are briefly described in the following sections.

### 2.5.1 Description of the Input Variables

Table 2.1 provides summary statistics for all the variables used in the analysis. The variable Land measured in Jeribs is the total land cultivated by the household in various seasons throughout the year. This includes both irrigated and rain-fed land owned or leased by the household that was actually cultivated throughout the year. The size of the agriculture holding is small in Afghanistan, and therefore availability of agriculture land is an important factor for production.

**Table 2.1:** Summary Statistics for Variables used in the Analysis

| Variable | Description | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| *Dependant Variable* | | | | | |
| Y | Aggregate Annual Revenue (10K AFN) | 5.83 | 9.04 | 0.011 | 12.80 |
| *Production Inputs* | | | | | |
| X1 | Land (Jeribs) | 7.035 | 9.12 | 0.10 | 90.00 |
| X2 | Labour (hours) | 63.79 | 62.37 | 1.00 | 417.85 |
| X3 | Seed Expenditure (AFN) | 2,354 | 3,617 | 0.00 | 45,000 |
| X4 | Fertilizer Expenditure (AFN) | 4,763 | 8,298 | 0.00 | 99,000 |
| X5 | Chemicals Expenditure (AFN) | 365.0 | 1,197 | 0.00 | 10,000 |
| X6 | Tractor Rental (AFN) | 2,504 | 4,296 | 0.00 | 60,000 |
| X7 | Other Expenditure (AFN) | 2,178 | 5,766 | 0.00 | 90,000 |
| *Sources or Factors of Efficiency/Inefficiency* | | | | | |
| Z1 | Diversification Index (0≤CDI<1) | 0.296 | 0.23 | 0.00 | 0.82 |
| Z2 | Opium Share by Province (%) | 0.033 | 0.10 | 0.00 | 0.48 |
| Z3 | Extension Services (1=access, 0=no) | 0.209 | 0.41 | 0.00 | 1.00 |
| Z4 | Land Quality (0=irrigated, 1=rainfed) | 0.231 | 0.42 | 0.00 | 1.00 |
| Z5 | Household Size (Persons) | 8.335 | 3.46 | 1.00 | 36.0 |
| Z6 | Household Head Age (Years) | 44.38 | 13.8 | 14.00 | 95.0 |
| Z7 | HH Head Sex (0=female, 1=male) | 0.996 | 0.06 | 0.00 | 1.00 |
| Z8 | HH Head Literacy (0=no, 1=literacy) | 0.321 | 0.47 | 0.00 | 1.00 |
| | HH Edu (no formal schooling) | 0.834 | 0.37 | 0.00 | 1.00 |
| | HH Edu (lower secondary) | 0.052 | 0.22 | 0.00 | 1.00 |
| Z9 | HH Edu (upper secondary) | 0.079 | 0.27 | 0.00 | 1.00 |
| | HH Edu (technical or teacher college) | 0.021 | 0.14 | 0.00 | 1.00 |
| | HH Education (university & postgrad) | 0.013 | 0.12 | 0.00 | 1.00 |
| Z10 | Off-farm Employment (1=yes, 0=no) | 0.125 | 0.33 | 0.00 | 1.00 |
| Z11 | Own Cattle (heads) | 1.601 | 2.04 | 0.00 | 31.0 |
| Z12 | Own Tractor (number) | 0.053 | 0.23 | 0.00 | 3.00 |
| Z13 | Own Oxen (number) | 0.232 | 0.62 | 0.00 | 9.00 |
| | Farm Size 1 (0.1-2 Jeribs) | 0.318 | 0.47 | 0.00 | 1.00 |
| | Farm Size 2 (>2-5 Jeribs) | 0.293 | 0.46 | 0.00 | 1.00 |
| Z14 | Farm Size 3 (>5-10 Jeribs) | 0.226 | 0.42 | 0.00 | 1.00 |
| | Farm Size 4 (>10-20 Jeribs) | 0.103 | 0.30 | 0.00 | 1.00 |
| | Farm Size 5 (>20 Jeribs & above) | 0.060 | 0.24 | 0.00 | 1.00 |
| | Agro-ecological Zone 1 (NEM) | 0.023 | 0.15 | 0.00 | 1.00 |
| | Agro-ecological Zone 2 (CM) | 0.137 | 0.34 | 0.00 | 1.00 |
| | Agro-ecological Zone 3 (HFL) | 0.043 | 0.20 | 0.00 | 1.00 |
| Z15 | Agro-ecological Zone 4 (SMF) | 0.202 | 0.40 | 0.00 | 1.00 |
| | Agro-ecological Zone 5 (HVSB) | 0.121 | 0.33 | 0.00 | 1.00 |
| | Agro-ecological Zone 6 (TP) | 0.064 | 0.24 | 0.00 | 1.00 |
| | Agro-ecological Zone 7 (NMF) | 0.169 | 0.37 | 0.00 | 1.00 |
| | Agro-ecological Zone 8 (EMF) | 0.241 | 0.43 | 0.00 | 1.00 |
| | N | | | | 7,052 |

*Source:* Author's calculations of the ALCS 2013-14 data

Farm labour is another important variable included in the analysis. Agriculture labour is coded based on the occupations and sub-categories in the survey including farm workers (those who are directly involved in production of crops or animal keeping), fishers, hunters, government extension workers, etc. Since this study deals with production, only the first type of labour is included, for three sources of labour supply involved in production: family labour, child labour and hired labour. Persons aged 14 and over are adult labourers and those below this threshold are child labour. However, the productivity of one unit of the child labour used in production may vary as compared to the productivity of adult labour, therefore households (about 133 households or 1.5% of the original sample) that reported child labour's involvement in production were not included in the analysis. Hired labour includes only those who were hired in by the farm.

Labour is treated as a variable input which is measured in hours. Household labour hours and hired labour hours were added. It is important to note that m

ajority of the households reporting hired labour did not report household labour and vice versa. Other inputs include expenditures on seed, chemical fertilizers, chemicals (i.e. pesticides and herbicides), tractor rental, and other expenditures measured (i.e. irrigation water) in Afghan currency (Afghani symbolized as AFN throughout this study).

It is important to note that some variables have shown wider variation across households leading to potential outliers. This study checked whether the inclusion or removal of these outliers has impacts on the results, it was found that results are slightly driven by outliners particularly in some input variables (including Land, labour, chemicals and other expenditures). Therefore 1% of the largest values of the labour and 0.5% in the other two variables (namely land and chemicals) were trimmed or winsorized. Percent of zero values in input variables ($X$) are reported in Table 2.B1 in Appendix 2.B.

### 2.5.2 The Determinants of Technical Efficiency

The objective of stochastic frontier models is not only to serve as a benchmark against which technical efficiency of producers is estimated, but also to explore how factors such as farm and household characteristics exert influence on the farmer's performance (Kumbhakar and Lovell, 2000). A number of potential sources of efficiency or inefficiency were identified and are briefly described in this section.

Crop diversification is the main variable of interest in this study. The concept of crop diversification implies production of multiple crops on the farm throughout the year by an individual household. The inverse Herfindahl-Hirschman Index (HHI) is used as a measure for crop diversification or specialization. The index captures the degree or extent of diversification for an individual farm household. In other words, the Transformed Herfindahl Index (THI) is calculated for each farm separately to measure the degree of diversification using the following equation:

$$CDI_i = 1 - HHI_i \ = 1 - \sum_{j=1}^{J} \left( \frac{Y_j}{\sum_{j=1}^{J} Y_j} \right)^2 \quad 0 \leq \ HHI_i \ \leq 1 \qquad 2.6$$

Where CDI denotes the index for crop diversification, $Y_j$ represents the revenue proportion occupied by the $j^{th}$ crop (for $j = 1, 2, \ldots, J)$ in total revenue earned by households annually. The HHI index ranges from (close to) zero, reflecting complete diversification (i.e. maximum number of crops), to one, reflecting complete specialization (i.e. just one crop). In order to help ease the interpretation of the results, a direct measure for crop diversification was constructed by subtracting the Herfindahl index from 1 (to create a Diversification index $CDI_i \ = 1 - HHI_i$) which ranges between 0 (specialization) and 1 (complete diversification). Any value above zero signifies diversification.

The average value of index for Crop Diversification (CDI) for the sample farms is 0.30 (equivalent to HHI=0.70) with a standard deviation of 0.233 (Table 2.1 above), implying presence of a relatively low level of crop diversification in the sample. The numbers equivalent or effective number which is the inverse of Herfindahl-Hirschman Index $(1/HHI_i)$ is useful in indicating the number of equal share crops consistent with the concentration. The effective number helps show the number or group of farmers with crop production that is equally-concentrated or diversified. The distribution of the index for crop diversification and effective or equivalent number are shown in Figure 2.3.

The Herfindahl-Hirschman index has been widely used as a measure of crop diversification (Coelli and Fleming, 2004; Lakner et al., 2015; Manjunatha et al., 2013; Ogundari, 2013; Rahman, 2009; and Weiss et al., 2002). Nguyen (2014) reported the average Herfindahl index of 0.75 for Vietnam which is slightly higher than the estimated average of 0.70 Afghanistan (corresponding to mean CDI of 0.30) whereas Rahman (2009) reported the average Herfindahl index of 0.60 for Bangladesh, Ogundari (2013) Herfindahl index of 0.46 in Nigeria, and Manjunatha et al. (2013) reported 0.55 in India.

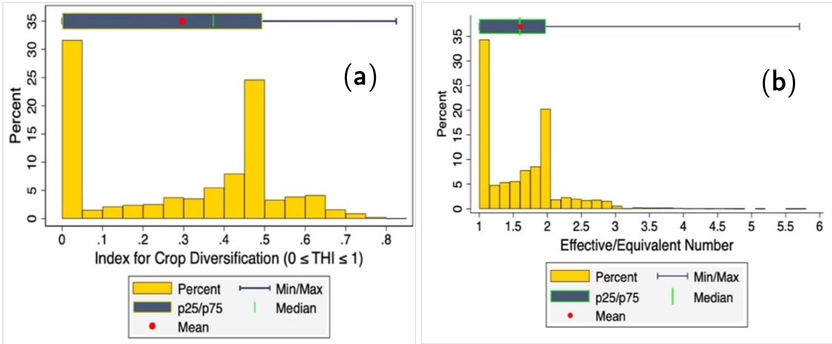



**Figure 2.3:** Distribution of THI Index (**a**), & Equivalent Number (**b**)
*Source:* Author's calculations of the ALCS 2013-14 data

Summary statistics and characteristics of relatively more diversified farms and less diversified farms are reported in Tables 2.B2 in Appendix 2.B. The farms in the sample ware divided in two sub-categories; those above the median level of CDI and below. The summary statistics show relatively higher total revenue for more diversified farms than those less diversified farms.

Afghanistan has a continental climate that is arid to semi-arid and is generally characterized by hot summers and cold winters. The wide range of altitude in Afghanistan leads to a great variation in climate within relatively small distances, which in turn affects the availability of water (rainfall), average annual temperature, and number of growing days. Temperature regimes are greatly modified by altitude – low sites are almost frost-free with very hot summers; the higher areas are arctic in winter (Thieme, 2006). The climatic types as listed by Khaurin (1996) which is also quoted by Thieme, (2006) are continental desert climate in the extreme north, Sub-tropical desert climate in the south, continental semi-arid Mediterranean climate in the north west, warm semi-arid mediterranean climate in the lower central and north west, continental semiarid to moist mediterranean with no winter frost in the north east central, dry steppe climate in the lower Kabul valley, alpine in high mountains, centre and north east.

Based on early work by Humlum (1959) later on revived by Dupree (1980), Afghanistan was divided into 11 geographical zones. However, recently a study by Maletta and Favre (2003) concluded that not all the 11 geographical zones have agricultural significance (i.e. some zones were classified as deserts). Based on ecological properties of land and climate, and some supplementary criteria about accessibility and prevailing agricultural activities, Maletta and Favre (2003) adopted the 8 agro-ecological zones (AEZ) scheme. These zones were constructed in the form of whole districts aggregations (Figure 2.4).

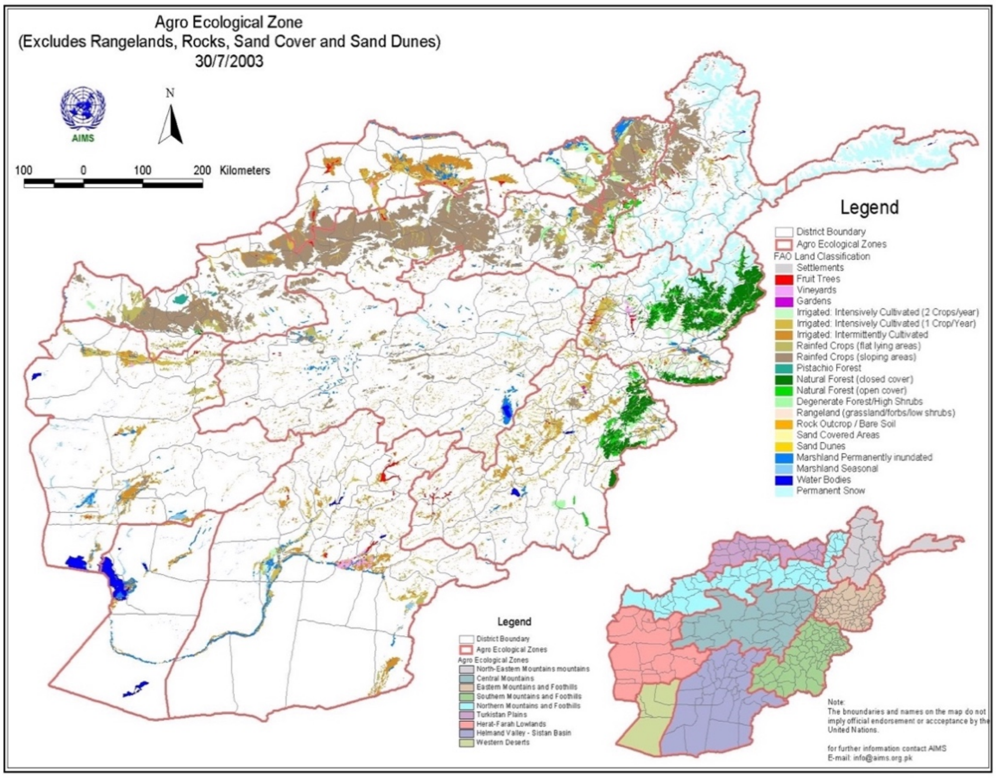



**Figure 2.4:** Agro-Ecological Zones of Afghanistan
*Source:* Adopted from Maletta and Favre (2003)

Annual participation, dry months and frost period (frost occurs when temperature drops below zero centigrade) across these 8 zones varies greatly. These variations, particularly the amount of annual rainfall may have potential effects on yield and the type of crops being grown. Table 2.2 summaries these climatic variations across the 8 zones.

**Table 2.2:** Agro-Ecological Zones of Afghanistan

| Agro-ecological Zone (AEZ) | Annual Precipitation (mm) | Dry months | Frost Months |
|---|---|---|---|
| North-Eastern Mountains (NEM) | 200-800 | 2-6 | 1-9 |
| Central Mountains (CM) | 200-800 | 2-6 | 1-9 |
| Heart-Farah Lowlands (HFL) | <100-300 | 6-12 | 0-3 |
| Eastern Mountains & Foothills (EMF) | 100-700 | 2-9 | 0-10 |
| Turkistan Plains (TP) | <100-400 | 5-8 | 0-2 |
| Helmand Valley-Sistan Basin (HVSB) | <100-300 | 6-12 | 0-3 |
| Southern Mountains & Foothills (SMF) | 100-700 | 2-9 | 0-10 |
| Northern Mountains & Foothills (NMF) | 200-800 | 2-9 | 0-8 |

*Source:* Adopted from Maletta and Favre, (2003) and Thieme (2006)

This study uses the eight agro-ecological zoning scheme to control for variation in crop production attributed to agro-climatic conditions. Afghanistan is generally categorized as a dry country where frequent droughts adversely affect farm production. Availability of irrigation water is important for crop production and varies greatly by agro-ecological zones (Table 2.2). These differences across agro-ecological zones are hypothesized to affect crop yields. In addition, the type and number of crops that can be grown are maybe restricted by climatic condition in certain zones that may in turn have implications for the extent of crop diversification.

Access to extension services is vital in assisting farmers in the production decision making process since it can be a reliable source of information, technical advice, trainings and improved farm management practices. Access to extension services is broadly believed in the literature to have a positive impact on the farm output and on the level of crop diversification. Table 2.B3 in the Appendix 2.B provide summary statistics and characteristics of farms with respect to access to extension services. The summary statistics show that farm revenues for farmers who have availed themselves of extension services are slightly higher than those who did not have contact with the extension services. In addition, farmers with access to extension services adopted a relatively diversified farming system than those who had no access.

About 21% of the sample farmers have access to extension services. Although relatively small number of farmers can avail of them, extension visits and training provided are important sources of information, farm management techniques, use and dissemination of innovation and technology. The survey directly provides data on whether farmers have had access to extension services or not. A binary variable was constructed which is equal to 1 if farmers have access and zero otherwise.

Farm size in Jeribs is the measure of the land variable. Impact of farm size on technical efficiency is investigated in the literature with mixed conclusions. Most of the empirical evidence suggests inverse relationship between the farm size and technical efficiency (i.e. smaller farm size is associated positively with the level of technical efficiency). Therefore, in context of Afghanistan where agriculture holding is relatively small, it is important to account for potential variability due to the farm size.

There are a number of recent studies that have identified and included off-farm employment in the inefficiency effect model. The impact of off-farm employment on technical efficiency is ambiguous. On one hand, off-farm employment shrinks the availability of labour for on-farm activities, especially if hiring agricultural labour incurs transaction costs, and therefore may negatively affect technical efficiency. On the other hand, off-farm employment enables households to increase their incomes, to overcome credit and insurance constraints and to increase their use of industrial inputs. Studies such as Essilfie et al. (2011) in Ghana, Haji (2007) in Ethiopia, Yang et al. (2016) and Zhang et al. (2016) in China, found that off-farm employment positively contributed to technical efficiency. On the other hand, studies conducted in North America and Europe concluded that technical efficiency is negatively related to off-farm employment due to reduction in labour supply to farm activities (Goodwin and Mishra, 2004; O'Neill et al., 2001).

Cattle ownership is used as a proxy for availability of animal manure at the farm. Animal manure is an important source of organic fertilizer, especially in the context of Afghanistan, and is generally believed to improve soil fertility. It is treated as a continuous variable being measured as the number of cattle heads owned by the farm at the time of the survey. Oxen and tractors are the two main sources of traction power used on the farm for ploughing and other farming activities. A dummy variable on whether a household owns a tractor, oxen or both was included in the model. It is generally believed that households who own a tractor or oxen or both might be cost effective, and therefore might have influence on the technical efficiency. On the other hand, oxen or tractor ownership may substitute for farm labour especially since some of the activities that are traditionally carried out by labour maybe completed by oxen or tractor.

As mentioned earlier, households own and cultivate either irrigated, rain-fed or a combination of both irrigated and rain-fed land to produce crops. Based on the descriptive statistics of the survey data, annual aggregate revenue for those household who cultivate irrigated land alone is much higher than those who operate a combination of both irrigated and rain-fed land. Therefore, it is a priori expected that households who own and operate rain-fed land may be less efficient compared to those who have access to irrigated land. To capture this variation attributed to the quality of land, a binary variable (equal to 0 for those who cultivated irrigated land alone, and 1 if the household cultivated rain-fed or a combination of both irrigated and rain-fed) was included in the analysis.

Another important source of (in)efficiency, especially in the context of Afghanistan, might be opium production by the farm household. Using the Afghanistan Ministry of Counter Narcotics annual data, an intensity variable is constructed to capture opium cultivation by province. The ALCS survey used in have also collected information on opium production from the households, however the reliability of the data might be a concern as production and trade of narcotics is illegal by the constitution, therefore households who actually produce opium might refrain from provision of data or provide misleading information. About 97.9% of households in the ALCS survey did not report growing opium. In general, there are certain zones and provinces where production of opium is relatively more common than other areas. Largely, opium production may have a direct connection with the security situation in the country (i.e. provinces that are opium free are relatively secure). Therefore, inclusion of this variable might also proxy for insecurity following that most of the opium infected areas are likely to be insecure. It may also capture unreported access to revenue as opium is a cash crop.

Household socio-economic characteristics such as household size, household head literacy and education (formal schooling), and household head sex, are generally included in the inefficiency effect model. Household's socio-economic characteristics are widely believed in the literature to affect efficiency. For instance, household size may affect labour supply. Household head education is used a proxy for farming experience and necessary skills of management. In the context of this study, in addition to the formal education by the household head, literacy rate is important and therefore was also included.

**Table 2.3:** Household Head Literacy and Education Levels

| Literacy/Education Level | Number of HH Heads | Percent |
|---|---|---|
| *Literacy Rate* | | |
| Can't Read & Write | 4,791 | 67.94 |
| Can Read & Write | 2,261 | 32.06 |
| *Formal Schooling* | | |
| No Formal Schooling | 5,326 | 83.42 |
| Lower Secondary | 364 | 5.16 |
| Upper Secondary | 560 | 7.94 |
| Technical or Teacher College (14 years) | 150 | 2.13 |
| University & Postgraduate | 95 | 1.35 |
| N | 7,052 | |

*Source*: Author's calculation from the ALCS 2013/14 data

The descriptive statistics as reported in Table 2.3 of the household head literacy rate and formal education attendance shows that literacy rate or level is important as 68% of the household's heads were reported to have no skill to read and write while 83.4% of them have not attended any type of formal schooling. This requires that these two aspects should be controlled for separately, especially since literacy rate is of more importance given the data.

### 2.6 Empirical Model Specification

Based on Equation (2.1), the translog stochastic frontier model initially developed by Aigner et al. (1977) and Meeusen and Van den Broeck (1977), can be specified as below:

$$lnY_i = \sum_{k=1}^{7} \beta_k \; lnX_{ik} + \frac{1}{2}\sum_{k=1}^{7}\sum_{j=1}^{7} \beta_{jk} \; lnX_{ik} \; lnX_{ij} \; + \; v_i - u_i \qquad 2.7$$

Where $Y_i$ represents aggregate revenue of the $i^{th}$ producer, $k$ represents the number of inputs used, $X_{ij}$ represents a set of seven input categories (mainly land, labour, seed, fertilizer, chemicals, tractor rental, and other expenditures) used by the $i^{th}$ farmer, and $\beta$ is a vector that collects unknown parameters to be estimated. In addition, $\varepsilon_i$ is the composed error term where $\varepsilon_i = v_i - u_i$ with $u_i \geq 0$. All input and output variables were transformed to their corresponding log values as denoted by $ln$ in 2.7. The random error $v_i$ accounts for the stochastic effects beyond the producer's control, measurement errors as well as other statistical noise, and $u_i$ captures production inefficiency due to factors that are in the control of the producer.

There are a number of distributional assumptions that could be made on the composed error term as mentioned earlier. In this study, two of the most commonly used distributional assumptions on the inefficiency component of the error term are made and results are cross checked and tested.

a) The Half-Normal Distribution: the normal-half normal case imposes the following restrictions on the error term:

   $v_i = iid \; N(0, {\sigma_v}^2)$
   $u_i = iid \; N^+(0, {\sigma_u}^2)$
   $v_i$ and $u_i$ are distributed independently of each other and of the regressors.

b) Truncated–Normal distribution: the truncated-normal case imposes the following restrictions on the error term:

$v_i = iid\ N(0, \sigma_v{}^2)$

$u_i = iid\ N^{+}(\mu, \sigma_u{}^2)$

$v_i$ and $u_i$ are distributed independently of each other and of the regressors, and $\mu$ is nonzero mean for $u_i$.

Given Equation (2.2) and the distributional assumption on the inefficiency component $(u_i)$ of the composed error term, the Battese and Coelli (1995) inefficiency model could be specified as:

$$u_i = \delta_0 + \sum_{i=1}^{14} \delta_i Z_i + w_i \qquad 2.8$$

Where $u_i$ is the inefficiency, $Z_i$ is the vector of exogenous variables (namely gender, age, literacy, education of the household head, household size, index for crop diversification, access to extension services, cattle ownership, oxen ownership, tractor ownership, off-farm employment, land quality, opium share by province, farm size, and agro-ecological zones) that are likely to affect efficiency, δ's are the parameters to be estimated, and $w_i$ is the error term of the efficiency model. As the dependent variable in Equation (2.8) is defined in terms of technical inefficiency, a farm-specific variable associated with the negative (positive) coefficient will have a positive (negative) impact on technical efficiency.

### 2.6.1 Maximum Likelihood Estimator (MLE)

The estimation of the model involves (i) estimating the parameters of the frontier function, and (ii) estimating inefficiency. There are various methods of estimation depending on the distributional assumptions for the error components. Early methods include Corrected Ordinary Least Square (COLS) and Corrected Mean Absolute Deviation (CMAD) which estimates technical efficiency without imposing any assumptions on the inefficiency component of the error term. However, these methods assume that the frontier function is deterministic, and the randomness of the model comes entirely from the variation in inefficiency. Therefore, deviations from the estimated frontier are entirely attributed to inefficiency, and there is no role for other randomness such as data errors (Kumbhakar and Wang, 2015).

On the other hand, the choice of distributional assumptions on the components of the error term is central to the ML estimation approach of the stochastic frontier model. After these distributional assumptions are imposed, the log-likelihood function of the model is derived, and numerical maximization procedures are used to obtain the ML estimates of the model parameters. Consequently, the maximum likelihood estimate of an unknown parameter is defined to be the value of the parameter that maximizes the probability (or likelihood) of randomly drawing a particular sample of observations. Aigner et al. (1977) focused on the implicit assumption that the likelihood of inefficient behaviour monotonically decreases for increasing levels of inefficiency. They parameterized the log-likelihood function for the half-normal model in terms of the variance parameters.

Maximizing a log likelihood function usually involves taking first derivatives with respect to the unknown parameters and setting them to zero. However, since these first order conditions are highly nonlinear and cannot be solved analytically for parameters, the likelihood function is maximized using an iterative optimization procedure.

### 2.7 Testing Hypothesis and Specification

Prior to undertaking the maximum likelihood estimation, it is important to check the validity of the stochastic frontier specification. Schmidt and Lin (1984) and Coelli (1995) proposed that in specifying the stochastic frontier model, a pre-test of the skewness of the OLS residual based on the Third Moment (M3T) should be carried out to test the null hypothesis of no skewness. The theory behind the test is that, for a production-type stochastic frontier model with the composed error $\varepsilon_i = v_i - u_i$ with $u_i \geq 0$ and $v_i$ distributed symmetrically around zero, the residuals from the corresponding OLS estimation should skew to the left (i.e. negative skewness). Thus, a negative skew of the third moment is an indication of the existence of efficiency effects. The Coelli (1995) test is given by:

$$M3T = m_3 / \sqrt{\frac{6m_2^3}{N}} \qquad 2.9$$

Where $m_2$ and $m_3$ are the second and the third sample moments of the OLS residuals, respectively. If the value of $M3T$ is statistically significant at the 1% level the frontier framework is supported. In our case, the computed value of the test statistic is -6.51.

Because it has a normal distribution, the corresponding critical value is 1.96, so the result confirms the rejection of the null hypothesis of no skewness in the OLS residuals. This result is further confirmed by significance of variance parameters ($\gamma$ and $\sigma^2$) in Table 2.5 where results of the stochastic frontier model are presented and the generalized log-likelihood ratio test for $\gamma$ presented in Table 2.4.

One of the drawbacks of the parametric SFA approach is having to specify functional form representing the production technology and imposing assumptions on the error components of the model. In addition, the stochastic frontier model imposes certain assumptions on the inefficiency term of the composed error term. It is important to ensure that the model specification correctly represents the data. It is therefore of interest to test the following hypothesis before presenting the results.

- Hypothesis 1: H0: $\beta_{jk} = 0$ the null hypothesis that identifies an appropriate functional form between the restrictive Cobb-Douglas and the translog production function. It specifies that the coefficients on square and interaction terms of input variables in Equation (2.7) are not statistically different from zero. The Cobb-Douglas production frontier is a special case of the translog frontier in which the coefficients of the second-order terms are zero, i.e., $\beta_{jk} = 0,\ j \leq k = 1, 2, \ldots, 7$.
- Hypothesis 2: $\mathrm{H0}$: $\gamma = 0$ in Equation (2.8) the null hypothesis that the inefficiencies are not stochastic and that the technical inefficiency effects are not present in the model at every level, so the joint effect of these variables on technical inefficiency is statistically insignificant. If this null hypothesis is not rejected, the Stochastic frontier model could be reduced to the OLS specification. In this case, if there is output difference among farmers given equal inputs, this difference is purely due to the difference in random shocks that are outside of the control of the farmer.
- Hypothesis 3: $\mathrm{H0}$: $\delta 0 = \delta 1 = \ \delta 2 \ldots \delta n = 0$ in Equation (2.8) the null hypothesis specifies that the influence of identified inefficiency factors (i.e. household socio-economic, farm-specific, and geographical factors) is zero.

- Hypothesis 4: H0: $u_i = iid\ N^+(0, \sigma_u{}^2)$ in Equation (2.7) the null specifying that half normal distribution better fits the model as opposed to the alternative case which assumes truncated normal distribution for the $u_i$.
- Hypothesis 5: $H0{:}\ \sum_1^7 \beta_i = 1$ in Equation (2.7) the null hypothesis specifying that there exists constant return to scale in the production function. A Wald test will be used to test whether the production function exhibits a constant, increasing, or decreasing returns to scale.

A Generalized log-likelihood ratio (LR) test can be used to test which specification better fits the data. The Generalized log-likelihood ratio test is given by:

$$LR = -2[\ln\{L(\mathrm{H0})\} / \ln\{L(\mathrm{H1})\}] = -2[\ln\{L(\mathrm{H0})\} - \ln\{L(\mathrm{H1})\}] \quad 2.10$$

Where L(H0) and L(H1) are the values of the likelihood functions under the null (H0) and alternative (H1) hypothesis respectively. The computed test statistics should be compared with critical values of the mixed chi-square distribution proposed by Kodde and Palm (1986). The LR and Wald tests are applied using the *lrtest* and *test* commands in Stata.

**Table 2.4:** Specification and Hypothesis Testing

| Null Hypotheses | Test Statistic | P-Value | Decision |
|---|---|---|---|
| *Functional Form (Translog vs Cobb-Douglas)* | | | |
| H0: $\beta jk = 0$ | LR= 593.43 | 0.000 | Reject H0 |
| *Specification of Frontier Model* | | | |
| H0: $\gamma=\delta 0=\delta 1=\ldots\delta n=0$ | LR= 171.49 | 0.000 | Reject H0 |
| H0: $\delta 0=\delta 1= \delta 2=\ldots\ \delta n=0$ | LR= 1,151.06 | 0.000 | Reject H0 |
| H0: $u_i = iid\ N^+(0, \sigma_u{}^2)$ | LR= 8.89 | 0.004 | Reject H0 |
| Testing for Constant Return to Scale | | | |
| H0: $\sum_1^7 \beta_i = 1$ | Wald($\chi$2)= 0.14 | 0.711 | Fail to reject H0 |

The results of the stochastic frontier model can be significantly affected by the choice of the functional form. The most widely used functional forms in estimating the production function are the Cobb-Douglas (restricted) and the translog (relatively more flexible). These two specifications of the stochastic frontier function were therefore selected and compared. The first hypothesis aims to test the choice of functional form using the generalized log-likelihood ratio test. The calculated LR statistic is 593.4 at 28 degrees of freedom which is greater than the $\chi^2$ critical value of 47.67 at 1% significance level, therefore the test rejects the Cobb-Douglas functional form in favour of translog production functional form. The test indicates that square and interaction terms in the translog model specified in Equation (2.7) are significantly different from zero, thus the translog model could not be reduced to the Cobb-Douglas specification.

The second null hypothesis can be tested using the generalized likelihood ratio test based on the value of log likelihood function under OLS and maximum likelihood estimation of stochastic frontier model. The computed LR test statistic is 171.5 at 1 degree of freedom which is greater than the $\chi^2$ crucial value of 5.41 at 1% significance level. Therefore, the null hypothesis that technical inefficiency effects are not present in the data is rejected at 1% significance level. Thus, the traditional average (OLS) production function is not an appropriate representation of the sample data. These findings confirm the results of M3T test presented earlier.

The third hypothesis is that the explanatory variables in the inefficiency model are simultaneously equal to zero. The LR test is used to calculate the test statistic using the log likelihood value of stochastic frontier model without explanatory variables of inefficiency effect model (H0) and the full frontier model with all explanatory variables of inefficiency effect model (H1). The computed LR test statistic is 1,151.06 at 27 degree of freedom which is greater than the $\chi^2$ crucial value of 46.35 at 1% significance level. Based on the calculated LR test statistic, the null hypothesis is rejected at 1% level of significance. Therefore, the explanatory variables associated with inefficiency effect model are jointly different from zero.

The fourth hypothesis was tested to validate the distributional assumption of the inefficiency term ($u_i$). Two models were constructed corresponding to the most common distributional assumptions of half normal and truncated normal for the one-sided error term as specified in section 3.5. The LR test is used to calculate the test statistic using the log likelihood value of stochastic frontier model assuming half normal destruction on the inefficiency term (H0) and the frontier model assuming truncated normal distribution on inefficiency term (H1). The calculated LR statistic is 8.9 at 1 degree of freedom which is greater than the $\chi^2$ critical value of 5.41 at 1% significance. Therefore, H0 is rejected implying that the truncated normal model is preferred to half-normal, however for comparison the results of both models are presented in Table 2.5 under the empirical results section.

The fifth hypothesis tests whether the production function exhibits constant return to scale. The computed Wald test statistic is 0.14 with a p-value of 0.711. Thus, the null hypothesis of constant return to scale cannot be rejected, implying that the specified production function exhibits constant return to scale.

## 2.8 Empirical Results and Discussion

### 2.8.1 Results of the Stochastic Frontier Model

The maximum likelihood estimates of parameters of the Stochastic Frontier Production Function (SFPF) and inefficiency model given by two-equation system (2.7) and (2.8) are simultaneously obtained using STATA and are reported in Tables 2.5 and 2.6. Both half normal (first column) and truncated normal (second column) specification of the inefficiency term $(u_i)$ were assumed and estimated. All seven inputs have the expected positive impact on the farm revenues.

The estimated value of $\sigma^2$ is positive and 3.83 which is statistically significant at 1% level. These values indicate that there exists sufficient evidence to suggest that technical inefficiencies are present in the data and that the differences between the observed (actual) and frontier (potential) output are due to inefficiency and not chance alone. Theoretically, this implies that the estimated model and distributional assumptions for the error terms are appropriate.

Gamma ($\gamma$) is the variance ratio, explaining the total variation in output from the frontier level of output attributed to technical efficiency. The estimated value of $\gamma$ (the ratio of the variance of output due to technical efficiency) is 0.902 for the preferred truncated normal model, indicating that about 90 percent of the difference between the observed and frontier output are primarily due to the inefficiency factors which can be managed and controlled by the farm households (Table 2.5).

The estimated coefficients for square terms (particularly labour, fertilizer, chemical and tractor rental squared) and several of the interaction terms are significantly different from zero indicating the rejection of the Cobb-Douglas model as an adequate representation of the data. It therefore justifies the non–linear functional form and that there exists important interaction among the variables.

### 2.8.2 Output Elasticities and Return to Scale

Since inputs and output variables were transformed to their corresponding log values, and were normalized by their respective sample means, therefore the estimated parameters are directly interpreted as partial elasticities at the sample mean. All slope coefficients or output elasticities of inputs had the expected signs and were found to be highly significant at 1 and 5% percent significance levels except for the variable of other farm expenditures. Coefficient estimates are quite similar for half- and truncated-normal speciation of the SFA (Table 2.5).

The results in Table 2.5 for the preferred truncated normal model show that land is the most important variable; the estimated coefficient is large and statistically significant at 1% with a positive sign which confirms the *priori* expectation. Expenditures on fertilizer and seed exhibits the second and third largest partial elasticities so is an important determinant of revenue. Other expenditures variable turned out to be insignificant at 5% level. Since farming is mostly subsistence and the farm size is small, other extra expenditures are quite uncommon and may not be a viable option especially for farmers that generate low cash income. All other purchased inputs are significant with the expected positive signs.

Returns to scale can be used to measure total resource productivity. The concept of returns to scale demonstrations how output responds to increase in all inputs together. The sum of the partial elasticities with respect to every input estimated by the maximum likelihood estimator of the translog stochastic production function is 0.99. This is roughly consistent with constant returns to scale which implies that an increase in all available inputs leads to an equal proportional increase in farm revenues.

Marginal effects of the explanatory variables at the mean could be obtained from the estimated production function by:

$$Marginal\ Effect\ of\ X_i = \frac{dy}{dx_i}\frac{\overline{x_\imath}}{\bar{y}} = b\frac{\overline{x_\imath}}{\bar{y}} \qquad 2.11$$

Where, $b$ =parameter estimate (partial elasticity associated with each independent variable), $\bar{x}$ = Mean of independent variable, $\bar{y}$ = Mean of dependent variable. The ME's measure a change in output at the mean as a result of one unit change in input .

**Table 2.5:** MLE Estimates for the Stochastic Frontier Model

| | Truncated-Normal | | Half-Normal | |
|---|---|---|---|---|
| Variable | Coefficient | SE | Coefficient | SE |
| *Dependent Variable (Total Aggregate Revenue in AFN)* | | | | |
| Constant | 0.147*** | 0.041 | 0.179*** | 0.041 |
| Ln Land ($X_1$) | 0.433*** | 0.025 | 0.431*** | 0.025 |
| Ln Labour ($X_2$) | 0.051** | 0.021 | 0.050** | 0.021 |
| Ln Seed Expenditures ($X_3$) | 0.131*** | 0.015 | 0.131*** | 0.015 |
| Ln Fertilizer Expenditures ($X_4$) | 0.199*** | 0.015 | 0.200*** | 0.015 |
| Ln Chemical Expenditures ($X_5$) | 0.038** | 0.015 | 0.039** | 0.015 |
| Ln Tractor Rental ($X_6$) | 0.117*** | 0.017 | 0.116*** | 0.017 |
| Ln other Expenditures ($X_7$) | 0.021* | 0.012 | 0.021* | 0.012 |
| 0.5 x Ln Land squared($X_1^2$) | 0.014 | 0.016 | 0.009 | 0.016 |
| 0.5 x Ln Labour squared ($X_2^2$) | 0.057*** | 0.014 | 0.057*** | 0.014 |
| 0.5 x Ln Seed Expenditures squared ($X_3^2$) | 0.035*** | 0.004 | 0.034*** | 0.004 |
| 0.5 x Ln Fertilizer Expenditures squared ($X_4^2$) | 0.037*** | 0.004 | 0.038*** | 0.004 |
| 0.5 x Ln Chemical Expenditures squared ($X_5^2$) | 0.018*** | 0.006 | 0.018*** | 0.006 |
| 0.5 Ln Tractor Rental squared ($X_6^2$) | 0.032*** | 0.005 | 0.032*** | 0.005 |
| 0.5 Ln other Expenditures squared ($X_7^2$) | 0.006* | 0.003 | 0.006* | 0.003 |
| Ln Land x Ln Labour | -0.021** | 0.010 | -0.021** | 0.010 |
| Ln Land x Ln Seed | -0.007** | 0.003 | -0.007** | 0.003 |
| Ln Land x Ln Fertilizer | 0.006** | 0.003 | 0.007** | 0.003 |
| Ln Land x Ln Chemicals | 0.001 | 0.004 | 0.001 | 0.004 |
| Ln Land x Ln Tractor Rental | -0.005 | 0.003 | -0.005 | 0.003 |
| Ln Land x Ln Other Expenses | 0.009*** | 0.003 | 0.009*** | 0.003 |
| Ln Labour x Ln Seed | 0.008*** | 0.003 | 0.008*** | 0.003 |
| Ln Labour x Ln Fertilizer | -0.010*** | 0.003 | -0.010*** | 0.003 |
| Ln Labour x Ln Chemicals | -0.001 | 0.004 | -0.002 | 0.004 |
| Ln Labour x Ln Tractor Rental | 0.003 | 0.003 | 0.003 | 0.003 |
| Ln Labour x Ln Other Expenses | -0.004* | 0.003 | -0.005* | 0.003 |
| Ln Seed x Ln Fertilizer | 0.001 | 0.001 | 0.001* | 0.001 |
| Ln Seed x Ln Chemicals | -0.003** | 0.001 | -0.002** | 0.001 |
| Ln Seed x Ln Tractor Rental | -0.001** | 0.001 | -0.001** | 0.001 |
| Ln Seed x Ln Other Expenses | -0.002*** | 0.001 | -0.002*** | 0.001 |
| Ln Fertilizer x Ln Chemicals | -0.000 | 0.001 | -0.000 | 0.001 |
| Ln Fertilizer x Ln Tractor Rental | -0.003*** | 0.001 | -0.003*** | 0.001 |
| Ln Fertilizer x Ln Other Expenses | 0.002*** | 0.001 | 0.002*** | 0.001 |
| Ln Chemicals x Ln Tractor Rental | -0.002* | 0.001 | -0.002* | 0.001 |
| Ln Chemicals x Ln Other Expenses | -0.001 | 0.001 | -0.001 | 0.001 |
| Ln Tractor Rental x Ln Other Expenses | -0.002*** | 0.001 | -0.002*** | 0.001 |
| (σ)2 | 0.371*** | 0.011 | 0.371*** | 0.012 |
| **γ** | 0.902*** | 0.038 | 0.889*** | 0.045 |
| Log-Likelihood | | -7,500.13 | | -7,504.22 |
| Chi2 (Prob.) | 4,675.44 | 0.000 | 4,675.44 | 0.000 |
| N | | 7,052 | | 7,052 |

*Significances is indicated by * p<0.10, ** p<0.05, *** p<0.010*

The computed marginal effects of the input variables are reported in Table 2.6. Land and labour have the largest impact as they are regarded as the most important factors for crop production. There is no data available on the rental rate of land, hence we cannot compare or benchmark the estimated marginal effects of land against the rental rate. However, the estimated ME for land appears to be reasonably high in relation to what a farm household would pay to rent one Jerib of land.

**Table 2.6:** Marginal Effects

| Variable | Elasticity | Marginal Effect |
|---|---|---|
| Land (X1) | 0.43 | 3,584.01 |
| Labour (X2) | 0.05 | 46.73 |
| Seed Expenditures (X3) | 0.13 | 3.25 |
| Fertilizer Expenditures (X4) | 0.20 | 2.44 |
| Chemical Expenditures (X5) | 0.04 | 6.04 |
| Tractor Rental (X6) | 0.12 | 2.73 |
| Other Expenditures (X7) | 0.02 | 0.55 |

*Source:* Author's calculation of the ALCS 2013/14 data

Labour turns out to have marginal effect of 46% (i.e. a unit change in labour will change the revenues by 46%). In comparison to the computed average hourly wage for agricultural labour which is about 38.5 per hour (or 256 Afghan daily), marginal effect of labour is fairly higher.

### 2.8.3 The Inefficiency Effect Model

Maximum likelihood estimator is used to estimate the $\delta$ coefficients of Equation (2.8) for technical inefficiency and the estimated results are presented in Table 2.7. A negative sign of the estimated parameters indicates a reduction in technical inefficiency or alternatively an increase in technical efficiency.

Estimated coefficients from the inefficiency effect model are all significant except for the sex, age, and education of household head, household size and off farm employment. The estimated coefficient for the index of crop diversification is negative and statistically significant at the 1 percent level under both half-normal and truncated-normal estimations of the $u_i$. This indicates that greater crop diversification index[7] (CDI) (which means lower HHI) is associated with higher level of technical efficiency at the farm level. The finding

[7] Using a binary variable for crop diversification, we re-estimated the SF model. The results are qualitatively consistent for CD; those that diversify are likely to be more efficient. These results are reported in Table 2.C1 in the Appendix for Chapter II.

that more diversified farms are more efficient is consistent with Nguyen (2014), Manjunatha et al. (2013), Ogundari (2013), Rahman (2009), and Coelli and Fleming (2004) for Vietnam, India, Bangladesh, Nigeria, and Papa New Guinea, respectively.

**Table 2.7:** Maximum Likelihood Estimation of the Inefficiency Model

| | Truncated-Normal | | Half-Normal | |
|---|---|---|---|---|
| Variable | Coefficient | SE | Coefficient | SE |
| Constant | 1.226*** | 0.430 | 1.089** | 0.459 |
| Head Sex (male) | -0.451 | 0.377 | -0.485 | 0.408 |
| Head Age (years) | 0.001 | 0.002 | 0.002 | 0.003 |
| Head Education (lower secondary) | 0.095 | 0.163 | 0.104 | 0.181 |
| Head Education (upper secondary) | 0.172 | 0.143 | 0.198 | 0.158 |
| Head Education (teacher college) | 0.102 | 0.243 | 0.113 | 0.272 |
| Head Education (university & postgrad) | -0.282 | 0.320 | -0.312 | 0.358 |
| Head Literacy (can read & write) | -0.026 | 0.093 | -0.028 | 0.103 |
| Household Size (persons) | 0.001 | 0.011 | 0.001 | 0.012 |
| Diversification Index (0<CDI<1) | -3.754*** | 0.347 | -4.334*** | 0.354 |
| Extension Services (1=yes) | -0.337*** | 0.098 | -0.388*** | 0.109 |
| Oxen and Yaks (number) | -0.169*** | 0.062 | -0.193*** | 0.069 |
| Tractor/Threshers (number) | -0.822*** | 0.255 | -0.930*** | 0.291 |
| Cattles (number) | -0.107*** | 0.022 | -0.120*** | 0.024 |
| Off-farm Employment (1=yes) | -0.027 | 0.100 | -0.041 | 0.111 |
| Opium share by province (%) | -0.914 | 0.724 | -1.024 | 0.849 |
| Farm Size (>2 to 5 Jeribs) | -0.380*** | 0.094 | -0.409*** | 0.102 |
| Farm Size (>5 to 10 Jeribs) | -0.340*** | 0.117 | -0.375*** | 0.126 |
| Farm Size (>10 to 20 Jeribs) | -0.054 | 0.157 | -0.072 | 0.173 |
| Farm Size (>20 & above Jeribs) | 0.353* | 0.203 | 0.371* | 0.225 |
| Land Quality (Low) | 0.236*** | 0.090 | 0.229** | 0.100 |
| Agro-ecological Zone 1 (CM) | -0.066 | 0.191 | -0.040 | 0.207 |
| Agro-ecological Zone 2 (HFL) | -0.054 | 0.225 | -0.026 | 0.246 |
| Agro-ecological Zone 3 (SMF) | -0.903*** | 0.209 | -0.977*** | 0.227 |
| Agro-ecological Zone 4 (HVSB) | -0.823*** | 0.234 | -0.878*** | 0.256 |
| Agro-ecological Zone 5 (TP) | -0.010 | 0.207 | 0.032 | 0.225 |
| Agro-ecological Zone 6 (NMF) | -0.231 | 0.183 | -0.216 | 0.199 |
| Agro-ecological Zone 7 (EMF) | -0.477** | 0.197 | -0.484** | 0.215 |
| N | | 7,052 | | 7,052 |

*Note: Table reports estimates of Equation (2.7). The omitted categories are: no formal schooling for education level, <2 Jeribs for farm size, agro-ecological zone 8 for AEZ, no access for extension services, none for literacy, none for off-farm employment, and irrigated and rainfed for of land quality; significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

Figure 2.5 further illustrates the effect of crop diversification; TE is increasing in crop diversification, that is the higher the intensity or extent of diversification, the higher the level of the technical efficiency achieved by the farm. Although diversifying crops may require additional management skills, it has advantages of greater utilization of inputs, and producing marketable crops while reducing reliance on production of a single staple crop mainly for home consumption.

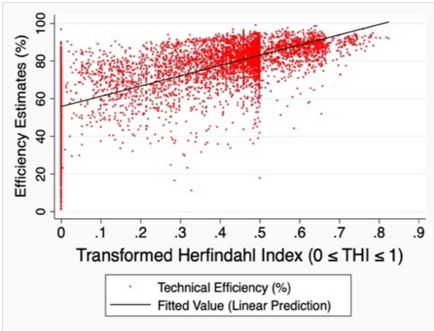



**Figure 2.5:**Distribution of TE by the Index of Crop Diversification (CD)
*Source:* Author's calculations of the ALCS 2013-14 data

In addition, Table 2.8 further summarizes the estimated ranges of technical efficiency and the distribution of farms according to their intensity of crop diversification. While about a third of the farms are almost perfectly specialized (mono-cropping), nearly half of the farms are experiencing the degree of diversification in the middle range (between 0.3 and 0.6). This leaves about 8% of the farms that are highly diversified (i.e. with CDI ranging from 0.6 to 1).

**Table 2.8:** Distribution of Farms According to CD and Estimated TE

| Index of Diversification | Mean TE (%) | No. of Farms | Percent of Farms |
|---|---|---|---|
| 0-<0.1 | 54.98 | 2,346 | 33.27% |
| 0.1-<0.2 | 67.42 | 316 | 4.48% |
| 0.2-<0.3 | 71.94 | 452 | 6.41% |
| 0.3-<0.4 | 76.97 | 637 | 9.03% |
| 0.4-<0.5 | 82.08 | 2,236 | 31.71% |
| 0.5-<0.6 | 84.60 | 562 | 7.97% |
| 0.6-<0.7 | 86.84 | 415 | 5.88% |
| 0.7-1 | 89.76 | 88 | 1.25% |
| N | | | 7,052 |

*Source*: Author's calculations of the ALCS 2013-14 data

The negative and significant effect of access to extension services on technical inefficiency implies that farmers who have had contact with extension services have higher technical efficiency, perhaps because they are helped to diversify. The descriptive analysis of diversification and extension services reveal that farmers who have access to extension services have implemented relatively more crop diversification than those who did not have access to extension services (Table 2.B3 in appendix 2.B). In a recent study, Makate et al. (2016) found that farmers with access to extension services had 38.4 % more chance of adopting a diversified cropping system than their counterparts (those without access to extension). Extension workers have technical knowledge on crop production and improved production management practices that can assist farmers to implement their crop diversification decisions. Elias et al. (2013) concluded that extension services increases farm productivity by 20% in Ethiopia. Mango et al. (2015) and Bozoğlu and Ceyhan (2007) found a positive impact of extension services on technical efficiency in Zimbabwe and Turkey respectively.

There seems to be an inverted U-shaped relationship between farm size and technical efficiency. Efficiency level rises initially with farm size (inefficiency is lower in farms with 2-10 Jeribs compared to <2 Jeribs) but appears to fall when farm size exceeds 20 Jeribs. Figure 2.6 shows the distribution of technical efficiency and the index of crop diversification by the farm size.

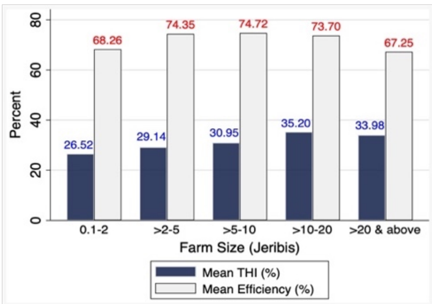



**Figure 2.6:** TE and Index of Crop Diversification by Farm Size
*Source:* Author's calculations of the ALCS 2013-14 data

It is evident that both crop diversification and technical efficiency initially follow the same pattern; as the farm size initially increases, the levels of crop diversification and technical efficiency also increase, but eventually when farm size is 20 or above efficiency fall and crop diversification levels out as the as farm size increases beyond 20 Jeribs. The computed average efficiency scores imply that medium sized farms are relatively more efficient. This may be due to the fact that medium level farms are more diversified. Findings on the relationship between farm size and efficiency vary in the literature. Oladeebo and Oyetunde (2013) and Bhatt and Bhat (2014) find an inverse relationship between farm size and technical efficiency. Manjunatha et al. (2013) and Mburu et al. (2014) concludes that increased farm size improves technical efficiencies. Helfand and Levine (2004) concluded that the relationship between farm size and efficiency is non-linear, with efficiency first falling and then rising with size and fall again when farm size is too large. Adhikari and

Bjorndal (2012) concluded that medium size farmers achieve a higher technical efficiency than large and small farm sizes, suggesting that productive efficiency can be increased with the encouragement of creating medium size holdings. Narala and Zala ( 2010) found that medium size farms are the most efficient in rice farming in Gujrat India, presumably due to medium farmers having agriculture as their main occupation and allocating their resources more effectively.

Another possible explanation of the observed inverted U-shaped relationship between farm size and technical efficacy may be due to the fact that the small farms may be incurring higher fixed costs. On the other hand, large farms are more likely to operate in diseconomies of scale and are more likely to suffer from resource misallocation and monitoring production activities.

While the mean technical efficiency across the entire country is estimated to be 72.64%, it varies across agro-ecological zones (Figure 2.7). Southern Mountains and Foothills (SMF) records the highest average level of 81%, followed by Helmand Valley and Sistan Basins (HVSB) of 79%, Eastern Mountain and Foothills (EMF) of 77%, Central Mountains (CM) of 66%, Heart-Farah Lowlands (HFL) of 65%, Northern Mountains and Foothills (NMF) 64%, Turkistan Plains (TP) 61%, and North Eastern Mountains (NEM) experienced the lowest level of 58%. However, there is only a statistically significant difference for 3 zones including SMF, HVSB and EMF having higher efficiency than the NEM zone (Table 2.5).

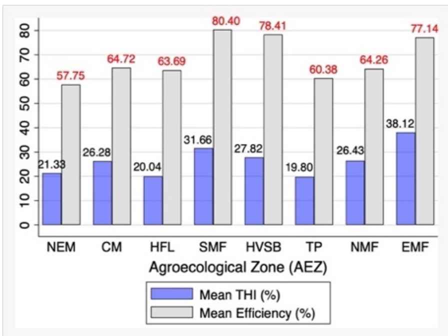



**Figure 2.7:** TE and Crop Diversification by Agro-ecological Zones
*Source:* Author's calculations of the ALCS 2013-14 data

The distribution of technical efficiency and degree of crop diversification across all 8 agro-ecological zone is shown in Figure 8 below which confirms that the most efficient agro-ecological zones (particularly SMF, HBVS, and EMF) are relatively more diversified on average as compared to those relatively less inefficient zones. We also mapped the intensity of crop diversification and estimated TE levels by districts (see Figure 2.B1 in Appendix 2.B).

Ownership of cattle, oxen and tractors by the households are positively correlated to the level of technical efficiency. Cattle and oxen ownership might imply availability of animal manure which is an important and cheap source of organic fertilizer (particularly in small-scale farming system) in soil that is widely believed to have a positive impact on soil fertility. To a certain degree, animal manure is considered as a good substitute for chemical fertilizers. Oxen and tractors ownership are also considered as important sources of the cheaper traction power available to farmers than those who hire tractor power and therefore farmers with greater number of oxen and tractors/thrashers maybe more efficient. In addition, tractor use may also indicate farm mechanization that ensures timely land preparation, planting and weeding.

Farmers who are operating a combination of rain-fed land and irrigated land were found to be more inefficient as compared to those who cultivated irrigated land alone. This suggests that rain-fed land is associated with lower crop yields. In addition, farmers who operate rain-fed land are less likely to diversify their production as was found in the descriptive statistics. This is because most crops, especially high value vegetables crops, require more irrigation water and therefore are not commonly produced on rain-fed land.

Opium intensity by province was found to be positively but insignificantly linked with technical efficiency. Insignificance may be due to a trade-off between effects of access to cash and insecurity. Production in provinces where farmers grow opium may be relatively more efficient compared to other regions, because farmers can purchase inputs (and sales of opium may inflate reported revenue), but opium affected provinces are likely to be more insecure. The descriptive statistics given in Table 2.B4 in Appendix 2.B confirms the higher farm revenues in provinces where more than 1% of opium production is reported as compared to those which are opium free or producing less than 1% of opium.

The insignificant efficiency factors include household head age, sex, literacy, and education levels, the size of household, and off-farm employment, indicating that these factors may not have significant impact on the farm technical efficiency.

### 2.8.4 Estimation of Technical Efficiency:

Based on Equations (2.7) and (2.8), farm-specific indices of technical efficiency were estimated assuming both half normal and truncated normal specification on the inefficiency component of the composed error term. It is evident from the results that the estimated technical efficiency estimates from the preferred truncated normal distribution range from 1.5% to 99.29%, with a sample mean of 71.9%. This reveals that there is substantial technical inefficiency in the Afghan farming sector. The main implication of this result is that farmers could increase their output by 28.1% on average without using additional resources, simply by improving technical efficiency. These estimates of technical efficiency are comparable with findings of other recent studies, for instance, Mwajombe and Mlozi (2015), Elias et al. (2013), Alam et al. (2012), Amaza et al. (2006), and Kudaligama et al. (2000) have estimated average efficiency levels of 72% in Tanzania, 78% in Bangladesh, 72% in Ethiopia, 65% in Nigeria, and 72% in India respectively.

**Table 2.9:** Range and Frequency of Technical Efficiency

| Efficiency | Truncated-Normal | | Half-Normal | |
|---|---|---|---|---|
| Range (%) | Number of farms | Percentage | Number of farms | Percentage |
| <25 | 153 | 2.17 | 163 | 2.31 |
| 25-<50 | 761 | 10.79 | 846 | 12.00 |
| 50-<60 | 646 | 9.16 | 695 | 9.86 |
| 60-<70 | 925 | 13.12 | 1,010 | 14.32 |
| 70-<80 | 1,560 | 22.12 | 1,786 | 25.33 |
| 80-<90 | 2,416 | 34.26 | 2,365 | 33.54 |
| 90-100 | 591 | 8.38 | 187 | 2.65 |
| Mean | 71.88 | | 69.87 | |
| SD | 17.47 | | 17.27 | |
| Minimum | 1.32 | | 1.39 | |
| Median | 77.20 | | 75.27 | |
| Maximum | 99.29 | | 97.53 | |
| N | 7,052 | | 7052 | |

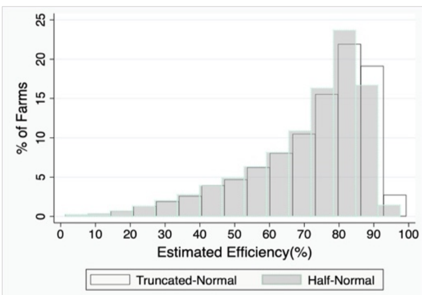



**Figure 2.8:** Distribution of Technical Efficiency

*Source:* Author's calculations of the ALCS 2013-14 data

The frequency distributions of the technical efficiency estimates are presented in Table 2.9. Moreover, distribution of the estimated technical efficiency and crop diversification with respect to the effective or equivalent number (i.e. the households with equal share of crops) is reported in Table 2.B5 in Appendix 2.B.

Distribution of the estimated efficiency indices estimated by the preferred truncated-normal model reveals that about 13% of the sample farmers realized less than 50% of the potential revenues, whereas about 43% of farms have achieved more than 80% technical efficiency. The remaining farmers were operating between the levels of 50% and 80% (Table 2.9 and Figure 2.8). The distribution of efficiency indices derived from the half-normal is quite similar to those of the truncated-normal case (Table 2.9).

### 2.8.5 Endogeneity in Crop Diversification

In the previous sections, crop diversification was assumed to be exogenous, however the decision to diversify may be an endogenous variable because the decision to adopt a diversified production system is likely to depend on unobservable variables. This means that failing to account for potential endogeneity may lead to endogeneity bias and consequently result in estimating an inconsistent effects of crop diversification on technical efficiency in the model presented in the previous section. The use of other input variable might also be endogenous, but since the main focus of this study is on analysing the impact of crop diversification on technical efficiency, hence this study will focus on the endogeneity issue in the crop diversification variable.

Due to its voluntary nature, the farmers self-select or choose whether to produce a single crop or a number of different crops. For instance, farmers who are relatively wealthier and have more technical knowledge on crop diversification as a viable strategy might be more likely to adopt crop diversification than their counterparts (such as those without access to extension), thus this unobserved selection bias may overstate the impact of crop diversification on technical efficiency. On the other hand, to the extent that CD is measured with error there may be attenuation bias so the basic SFA model may underestimate the impact of crop diversification on technical efficiency. In either case, there are unobserved factors in the error term $(u_i)$ that are correlated with the endogenous variable (CDI) that may result in biased estimates of the impact of crop diversification on technical efficiency in the basic SFA model.

To account for this potential selection bias due to endogeneity, the instrumental variables method is used. The IV for crop diversification used in this study is the mean value of the Crop Diversification Index (CD) for other farm households in the district which is calculated as follow:

$$Instrumental\ Varaible\ (IV) = \left(\frac{\sum X_d - X_i}{N_d - 1}\right)$$

Where $\sum X_d$ is the sum of the Diversification Index ($CD = 1 - HHI$)) in the district, $X_i$ is the CD value of the $i^{th}$ farm in the respective district, $N_d$ is the number of observations in the respective district (so $n-1$ for the district is the number of observations in the district excluding the ith farm itself). This means that the IV will differ slightly for each farm depending on the variance relative to that farm. On average, there are 40 farm households in each district. While constructing the IV, about 14 observations were dropped because there were too few observations in a few districts which made it impossible to calculate the average value of CD, as a result the sample was reduced from 7,052 to 7,038.

The extent or degree of crop diversification may be magnified through social interactions between farmers in the local neighbourhood. Farms that face similar demographic characteristics and preferences are likely to adopt similar production systems. For instance, a farm household located in a district where farmers have greater access to information and markets, and are therefore more likely to diversify, is more likely to adopt a diversified production system than a farm in a less diversified district. Observing that neighbours diversify would encourage a farmer to follow the example, so even relatively ‘low ability’ farmers are more likely to diversify. However, the fact that neighbours diversify should not in itself affect the efficiency of the farmer as factors associated with efficiency, such as farm size or ownership of livestock, are not affected by neighbours’ diversification.

Although there is a growing concern about the endogeneity issues in the stochastic frontier models, there is still limited work available in the literature to address it. Addressing the endogeneity issue is relatively more complicated in the stochastic frontier models due to the special nature of the error term. Standard Instrumental variable (IV) approaches cannot be used and the literature has yet to develop a strategy for addressing endogeneity with respect to the scaling factors in the one-sided error of a stochastic frontier model (Gronberg et al., 2015). Gronberg et al. (2015) attempt to solve the endogeneity problem in frontier models through pseudo-IV methodologies.

Guan et al., (2009) employs a two-step estimation methodology to handle the endogenous regressors in the frontier framework. Using GMM, in the first step, they estimate the consistent estimates of the frontier parameters, and then use the computed residuals from the first step as the dependent variable to get the maximum likelihood stochastic frontier estimates in the second step. Because the second step estimation is obtained by employing the standard stochastic frontier estimators, the efficiency estimates are inconsistent if the two-sided and one-sided error terms are correlated.

In a recent study, Amsler et al. (2016)  present a copula method, in which the more general correlation structures are allowable when modelling endogeneity. However, the copula approach is computationally intensive and complex which requires to choose a copula properly. Besides, this approach does not allow variables that affect inefficiency, which makes it less applicable when trying to understand the factors that TE.

The instrumental variable estimator used in this study follows the recent work of Karakaplan and Kutlu (2017) who developed a general ML based framework to handle the endogeneity problem in the stochastic frontier models. The endogenous stochastic frontier model is estimated using the *sfkk* command in STATA (Karakaplan, 2017). For further discussion and mathematical derivation see Karakaplan and Kutlu (2017) and Karakaplan, (2017). The stochastic frontier and inefficiency effect models were simultaneously estimated. For comparison, Table 2.10 provides estimates of both exogenous (column 1) and endogenous model (column 2) assuming a half-normal distribution for the inefficiency component ($u_i$) of the error term in Equation (2.7). The *sfkk* command only allows half-normal distribution for the $u_i$ term. The parameter estimates of the two models (exogenous and endogenous) are largely and qualitatively similar, however the estimated coefficient for the endogenous variable (crop diversification) slightly changes in size.

Table 2.10: Estimation of Endogenous Stochastic Frontier Model

| | Endogenous | | Exogenous | |
|---|---|---|---|---|
| variable | Coefficient | se | Coefficient | se |
| *Dependent Variable (Total Aggregate Revenue in AFN)* | | | | |
| Constant | 0.127*** | 0.045 | 0.159*** | 0.041 |
| Ln Land ($X_1$) | 0.413*** | 0.025 | 0.442*** | 0.025 |
| Ln Labour ($X_2$) | 0.052** | 0.021 | 0.053** | 0.021 |
| Ln Seed Expenditures ($X_3$) | 0.126*** | 0.015 | 0.131*** | 0.015 |
| Ln Fertilizer Expenditures ($X_4$) | 0.202*** | 0.015 | 0.201*** | 0.015 |
| Ln Chemical Expenditures ($X_5$) | 0.043*** | 0.015 | 0.039*** | 0.015 |
| Ln Tractor Rental ($X_6$) | 0.112*** | 0.017 | 0.116*** | 0.017 |
| Ln other Expenditures ($X_7$) | 0.035*** | 0.012 | 0.022* | 0.012 |
| 0.5 x Ln Land squared($X_1^2$) | -0.008 | 0.016 | 0.009 | 0.016 |
| 0.5 x Ln Labour squared ($X_2^2$) | 0.057*** | 0.014 | 0.065*** | 0.016 |
| 0.5 x Ln Seed Expenditures squared ($X_3^2$) | 0.033*** | 0.004 | 0.034*** | 0.004 |
| 0.5 x Ln Fertilizer Expenditures squared ($X_4^2$) | 0.038*** | 0.004 | 0.038*** | 0.004 |
| 0.5 x Ln Chemical Expenditures squared ($X_5^2$) | 0.019*** | 0.006 | 0.018*** | 0.006 |
| 0.5 Ln Tractor Rental squared ($X_6^2$) | 0.031*** | 0.005 | 0.032*** | 0.005 |
| 0.5 Ln other Expenditures squared ($X_7^2$) | 0.010*** | 0.003 | 0.006* | 0.003 |
| Ln Land x Ln Labour | -0.019* | 0.010 | -0.018* | 0.010 |
| Ln Land x Ln Seed | -0.006** | 0.003 | -0.007** | 0.003 |
| Ln Land x Ln Fertilizer | 0.009*** | 0.003 | 0.007** | 0.003 |
| Ln Land x Ln Chemicals | 0.001 | 0.004 | 0.001 | 0.004 |
| Ln Land x Ln Tractor Rental | -0.004 | 0.003 | -0.004 | 0.003 |
| Ln Land x Ln Other Expenses | 0.008*** | 0.003 | 0.009*** | 0.003 |
| Ln Labour x Ln Seed | 0.009*** | 0.003 | 0.009*** | 0.003 |
| Ln Labour x Ln Fertilizer | -0.010*** | 0.003 | -0.009*** | 0.003 |
| Ln Labour x Ln Chemicals | -0.003 | 0.004 | -0.003 | 0.004 |
| Ln Labour x Ln Tractor Rental | 0.004 | 0.003 | 0.002 | 0.003 |
| Ln Labour x Ln Other Expenses | -0.005* | 0.003 | -0.006** | 0.003 |
| Ln Seed x Ln Fertilizer | 0.001 | 0.001 | 0.001 | 0.001 |
| Ln Seed x Ln Chemicals | -0.002** | 0.001 | -0.002** | 0.001 |
| Ln Seed x Ln Tractor Rental | -0.001** | 0.001 | -0.001* | 0.001 |
| Ln Seed x Ln Other Expenses | -0.002*** | 0.001 | -0.002*** | 0.001 |
| Ln Fertilizer x Ln Chemicals | 0.000 | 0.001 | -0.000 | 0.001 |
| Ln Fertilizer x Ln Tractor Rental | -0.003*** | 0.001 | -0.003*** | 0.001 |
| Ln Fertilizer x Ln Other Expenses | 0.002*** | 0.001 | 0.002*** | 0.001 |
| Ln Chemicals x Ln Tractor Rental | -0.002* | 0.001 | -0.001 | 0.001 |
| Ln Chemicals x Ln Other Expenses | -0.001 | 0.001 | -0.001 | 0.001 |
| Ln Tractor Rental x Ln Other Expenses | -0.003*** | 0.001 | -0.002*** | 0.001 |
| ***The Inefficiency Effect Model*** | | | | |
| Constant | 1.052* | 0.492 | 1.025** | 0.465 |
| Head Sex (male) | -0.447 | 0.437 | -0.486 | 0.412 |
| Head Age (years) | 0.001 | 0.003 | 0.002 | 0.003 |
| Head Education (lower secondary) | 0.114 | 0.197 | 0.088 | 0.186 |
| Head Education (upper secondary) | 0.226 | 0.173 | 0.210 | 0.163 |

***Table 2:10 Continue***

| | | | | |
|---|---|---|---|---|
| Head Education (teacher college) | 0.168 | 0.301 | 0.160 | 0.279 |
| Head Education (uni & postgrad) | -0.204 | 0.382 | -0.369 | 0.370 |
| Head Literacy (can read & write) | -0.074 | 0.114 | -0.036 | 0.106 |
| Household Size (persons) | 0.004 | 0.013 | 0.003 | 0.013 |
| Diversification Index | -6.806*** | 1.027 | -4.481*** | 0.383 |
| Extension Services (1=yes) | -0.473*** | 0.123 | -0.421*** | 0.113 |
| Oxen & Yaks (number) | -0.201*** | 0.077 | -0.210*** | 0.070 |
| Tractor/Threshers (number) | -0.785*** | 0.303 | -1.051*** | 0.312 |
| Cattles (number) | -0.101*** | 0.027 | -0.117*** | 0.024 |
| Off-farm Employment (1=yes) | -0.009 | 0.122 | -0.044 | 0.117 |
| Opium share by province (%) | -0.060 | 0.870 | -0.876 | 0.860 |
| Farm Size (>2 to 5 Jeribs) | -0.306*** | 0.110 | -0.412*** | 0.104 |
| Farm Size (>5 to 10 Jeribs) | -0.295** | 0.133 | -0.324** | 0.128 |
| Farm Size (>10 to 20 Jeribs) | -0.032 | 0.183 | -0.017 | 0.175 |
| Farm Size (>20 & above Jeribs) | 0.276 | 0.263 | 0.404* | 0.227 |
| Land Quality (Low) | 0.038** | 0.122 | 0.220** | 0.102 |
| Agro-ecological Zone 1 (CM) | 0.029 | 0.220 | 0.004 | 0.210 |
| Agro-ecological Zone 2 (HFL) | -0.120 | 0.260 | -0.038 | 0.250 |
| Agro-ecological Zone 3 (SMF) | -0.851*** | 0.239 | -0.997*** | 0.231 |
| Agro-ecological Zone 4 (HVSB) | -0.953*** | 0.269 | -0.884*** | 0.260 |
| Agro-ecological Zone 5 (TP) | 0.050 | 0.242 | 0.037 | 0.229 |
| Agro-ecological Zone 6 (NMF) | -0.106 | 0.213 | -0.229 | 0.203 |
| Agro-ecological Zone 7 (EMF) | -0.315** | 0.228 | -0.483** | 0.219** |
| Log-Likelihood | | -5,795.52 | | -7,541.77 |
| Wald Chi2 | | 5,474.54 | | 4,801.98 |
| Prob. Chi2 | | 0.000 | | 0.000 |
| Mean Efficiency (%) | | 73.88% | | 69.87% |
| η Endogeneity test | | 25.13 | | |
| η Endogeneity test (p-value) | | 0.000 | | |
| N | | 7,038 | | 7,053 |

*Note: The omitted categories are: no formal schooling for education level, <2 Jeribs for farm size and agro-ecological zone 8 for AEZ, no access for extension services, none for literacy, none for off-farm employment, and irrigated & rainfed combined for of land quality; significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

Correcting for the potential endogeneity of the variable of crop diversification (CDI) decreases slightly (increases in absolute terms) its coefficient (from -4.95 to -6.81) in the inefficiency model (Table 2.9). Failing to account for the endogeneity issue underestimated the effect of crop diversification on technical efficiency in the standard exogenous stochastic frontier model presented in column 1 of Table 2.10. This is consistent with attenuation bias due to measurement error in CD so there was a downward bias in the estimation of the coefficient on CD in the basic SFA model. As a result, the average estimated level of technical efficiency by the endogenous model is also 4% (i.e. mean efficiency for exogenous model is 69.9% and 0.73.9% from endogenous model) higher than the estimated efficiency by the standard model assuming that crop diversification is exogenous, indicating that exogenous model underestimates the level of technical efficiency.

For further illustration we plotted the distribution of the estimated technical efficiency by the exogenous and endogenous models. These estimates largely overlap for less efficient farms; however, the two estimates are different for farms with higher efficiency levels (Figure 2.9). This is perhaps due to the fact that more efficient farms are highly diversified (as evidenced before), and when endogeneity is not corrected for the estimates by the exogenous model are biased downwards.

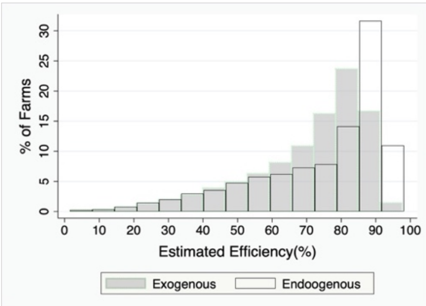



**Figure 2.9:** Estimated TE Indices by Endogenous and Exogenous Models
Source: Author's calculations of the ALCS 2013-14 data

Karakaplan and Kutlu (2017) offers an endogeneity test similar to Durbin-Wu-Hausman test as part of the sfkk estimation to test for endogeneity in stochastic frontier models. The eta ($\eta$) endogeneity test examines the joint significance of the components of the $\eta$ term. If the components are jointly significant, endogeneity is detected and must be corrected for, otherwise the model can be fit by traditional frontier models (for more details please see Karakaplan and Kutlu (2017). Under the null hypothesis, the tests assume that correction for the endogeneity is not necessary and that the exogenous estimation of crop diversification is valid. The estimated test statistic is $\chi^2 = 25.13$ with a p-value of 0.000 which rejects the null hypothesis of no endogeneity in CD at 1% statistical significance. This means that the Karakaplan and Kutlu (2017) test confirms endogeneity of CD and that issue of endogeneity has to be addressed in the exogenous model. Meanwhile, endogeneity is investigated by applying the Durbin and Wu-Hausman test. The calculated test statistic is 39.82 and rejects the null hypothesis of no endogeneity in crop diversification at 1% level.

The correlation of the instrument is also tested (results are reported in Table 2.C2 in Appendix 2.C). The instrument is strongly correlated with the endogenous variable (CD), conditional on the other covariates. This correlation is highly statistically significant (at 1%) indicating that the instrument is informative and strongly correlated with the endogenous variable. The endogenous variable was regressed on the instrument and all other covariates (i.e. all covariates included in the basic inefficiency model). The estimated coefficient for the instrumental variable is large (0.71) and statically significant at 1% level. A test of the joint significance of the instrument rejected the null hypothesis of weak instruments with an F-statistic of 1,075.57 (well above 10, the minimum value for an instrument to be strong) with $(Prob > F = 0.00)$. The instrument is sufficiently correlated with the diversification index but appears uncorrelated with the error term $(u_i)$. This means that the average value of CD for neighbouring farms in the district is likely to affect technical efficiency of the $i^{th}$ farm only through its impact on the crop diversification.

## 2.9 Robustness checks

Afghanistan is the largest producer of opium in the world which is primarily considered as a high-value cash crop. Based on the Afghanistan's Ministry of Counter Narcotics annual report, opium poppy cultivation in Afghanistan reached a sobering record high in 2013. The cultivation of poppy amounted to some 209,000 hectares, outstripping the earlier record in 2007 of 193,000 hectares, and representing a 36 percent increase over 2012. The prevalence of opium poppy cultivation may affect crop production patterns and diversification, especially in provinces where opium cultivation is highly concentrated. Moreover, these specific geographic areas where opium production is more common are likely to be more insecure as compared to those with no opiate production. As a result, this phenomenon could lead to systematic differences between households that are located in areas where opium is more common and areas that are opium free.

We included a control variable in our main empirical model to account for the variation due intensity of opium cultivation in some provinces. As a further robustness check, we now split our analytical sample into two sub-samples on the basis of the prevalence of opium poppy cultivation at the provincial level to investigate whether poppy cultivation has a significant bearing on our main empirical results presented earlier. We use information from the Afghanistan's Ministry of Counter Narcotics annual report published in 2013. The report contains information on area under opium poppy cultivation at the provincial levels. Main opium-cultivating provinces are Helmand, Kandahar, Farah, Nimiroz, Nangarhar, Urozgan, Badghis, Badakhshan and Dai Kundi with a total area of 48%, 14%, 12%, 8%, 8%, 5%, 2%, 1%, and 1% under the poppy cultivation respectively.

We estimate efficiency using SFA for each sub-sample separately. The results are presented in Tables 2.C3 and 2.C4 in the Appendix for Chapter II. The results yield quantitively different results across the two sub-samples, however, in qualitative terms the estimated coefficients do not appear to be substantially different between the two groups. Major disparities in estimates between two models are: total labour variable changes sign, other expenditures and household age variables becomes significant, and the coefficient for access to extension services become insignificant in the sub-sample of farm households in provinces with poppy cultivation (referred to as the opium sub-sample from this point forward). The rest of the coefficient estimates are largely and qualitatively similar.

The results from both sub-samples reveal that crop diversification is positively associated with the farm technical efficiency (negatively associated with inefficiency). However, the coefficient estimate for the diversification index (CEI) by the preferred endogenous model is considerably higher (in absolute terms) in the opium free sub-sample than the opium sub-sample. In other words, the direction of the bias is different between the two sub-samples, that is the endogeneity causes a downward bias in the opium free sub-sample consistent with the attenuation bias due to measurement error (i.e. the effect of crop diversification is even greater when endogeneity is accounted for), whereas for the opium sub-sample endogeneity causes an upward bias (e.g. the uninstrumented SFA model overestimates the effect of crop diversification to TE and once endogeneity biased is removed, the effect is smaller in absolute terms). Perhaps unobserved factors are linked to the fact that poppy cultivation is illegal by constitution and counter narcotic efforts from the government may restrict spreading poppy cultivation and instead encourage licit production, despite that these areas are less accessible by the government. It may will be because of the specialized entrepreneurship skills required for poppy production that may confine farmers to the production of other common crops, as is evidenced by the significance of the household age variable.

As for the total labour variable, the negative affect in the opium sub-sample could be due to the fact that opium producer households uses highly skilled hired labour which might be under reported or not reported at all in the ALCS household survey. Other expenditure variable which include other miscellaneous production costs is now significant in the opium sub-sample, perhaps these costs are crucial for poppy production farmers compared to other licit crops. It is not surprising to see that the household age variable becomes significant in opium sample, older farmers are likely to have more experience in farming and trading of opium and therefore they are expected be more involved in opium cultivation. Similarly, access to extension services appears to become insignificant in the opium sample as expected, because opium is an illegal crop and perhaps the extension agents are not able to travel to these areas due to insecurity.

In general, there appears to be no systematic and significant qualitative differences among the two sub-groups due to prevalence of opium poppy cultivation. Therefore, our main results (presented earlier) remain unaffected, especially as we include a control variable that captures the intensity of poppy cultivation at the provincial in our analysis which would be sufficient to handle any probable variation due to poppy cultivation.

### 2.10 Conclusion and Discussion

The focus of the analysis in this chapter is to estimate farm-level production efficiency and to investigate whether crop diversification strategies (a shift in production from mono-cropping to a mix-cropping system) by farm household improve technical efficiency of crop farming system in Afghanistan. We employ a recent estimation methodology developed by Karakaplan and Kutlu (2017) that allows to correct for the endogeneity bias in the stochastic frontier models, a major econometric issue in traditional basic stochastic frontier models.

The results of this study reveal that farming sector in Afghanistan experiences significant technical inefficiencies. Nearly 15% of farm households achieve less than 50% of the potential revenues, whereas about 23% percent realize 50-70% of the potential revenues.

Overall technical efficiency is estimated at 72% on average, indicating a substantial room for improving farm revenues via employing improved farm management practices and without using additional production resources and raising production costs. This finding is particularly crucial as the derived Constant Return to Scale (CRS) from the estimation of the preferred translog stochastic production function signifying that an increase in all inputs leads to equal proportional increase in revenues. Among inputs, land, household labour, fertilizer, seeds, tractor rental, and other expenditures were found to be positively and significantly contributing to production.

The empirical results from the preferred endogenous stochastic frontier model indicate that crop diversification, measured by the Transformed Herfindahl Index, is a key factor associated with higher levels of technical efficiency. Our findings agree with those of Ogundari (2013), Rahman (2009), and Coelli and Fleming (2004) for Bangladesh, Nigeria, and Papa New Guinea, respectively. This outcome is particularly crucial as the evidence suggests a relatively low level of diversity in crop production. Nearly 33% of the households do not diversify (i.e. monocropping) achieving 50% or less than 50% of the production efficiencies. The overall average value of the THI is estimated at 0.30 indicating low diversity compared to other countries with similar context such as Bangladesh, Nigeria, and India.

Aside from crop diversification, this study identified and examined the impact of a number of factors on technical efficiency. Farm households with better access to extension services appear to realize higher technical efficiencies. Ownership of farm assets (such as cattle, oxen, and tractor) by the farm households were found to have a significant positive impact on the production efficiency. Farm size appears to have an inverted U-shaped relationship with efficiency, revealing that efficiency first increase with the farm size but fall when farm size is larger than 20 Jeribs. Other factors were not consistently significant: off-farm employment, agro-ecological zones, intensity of opium production at the provincial level, and household characteristics.

Lastly, robustness checks were carried out to ensure the econometric specification of our empirical model best fit the data. Tests on the distributional assumptions for the composed error term validated that truncated normal distribution better fits the specification of the frontier model. The LR test confirmed the appropriateness of translog production function instead of a more restrictive Cobb-Douglas functional form. The robustness analysis to assess whether there are systematic differences in production systems between regions based on the prevalence of poppy cultivation reveals that there exist no qualitative differences among households located in provinces with opium and no opium cultivation.

A direct policy recommendation that can be generated from the findings of this study is that crop diversification should be given more credit and recognition by both farm households and policy makers, particularly in shifting production systems away from a mono- and staple-crop production to a mixed multiple and high value crop production system. As indicated by the significance of the extension services variable in the analysis carried out in this study, one way to improve crop diversification is to expand household's access to extension services. More generally, government investment in the development of rural infrastructure programs will not only increase production efficiencies at the farm level but will also complement crop diversification strategies by improving opportunities for technology diffusion, marketing, storage and resource supplies.

### 2.10.1 Further Research

This study can be extended to examine the trends of crop diversification and how crop diversification evolved over time. In addition, it is important to investigate the drivers of crop diversification. Crop Diversification might be restricted by farm size, agro-ecological zones and even the household habits of food consumption. For instance, wheat is the main staple food crop that accounts for about 60% of the caloric intake in the Afghan diet, thus replacing wheat might not be a choice for some households. Hence, it is worthwhile to further investigate how sensitive the efficiency estimates are to the farm size and agro-ecological zones.

Allocative efficiency as mentioned earlier is another important part of the total productivity of farms. Optimal use and allocation of inputs may potentially be an aspect that could improve overall productivity of farms. This could not be addressed given the absence of price data for inputs. Availability of data on input allocations and crop-specific inputs can also help address to allow for non-jointness (or jointness due to allocatable fixed inputs) of the production technology.

Production of high value cash crops with the basic objective of improving household cash income might require improved local and regional market opportunity. In fact, lack of access to markets maybe another restriction for diversifying farm production. Further, lack of a well-developed farm to market supply chain for the high value crops may make it difficult to move away from single crop production. Access to credit and other institutional aspects of farming might also affect both crop diversification and technical efficiencies.

If CD is a desired strategy for farmers in Afghanistan as was found in this sturdy, another line of research can focus on what drives or restricts crop diversification and investigate the crop choices and optimum combinations of annual crops to inform farmers on better crop mixes or enterprises to ensure productivity gains as well as food security.

# APPENDIX TO CHAPTER II

**Appendix 2.A:** DETAILED MATHEMATICAL DERIVATION OF SFA

The literature on stochastic frontier models begins with Aigner et al. (1977) normal-half normal model which assumes the following distribution for the components of the composite error term:

- $fv(v_{i)} = iid\ N(0, {\sigma_v}^2)$
- $u_i = iid\ N^+(0, {\sigma_u}^2)$
- $v_i$ and $u_i$ are distributed independently of each, and of the repressors.

The density function of half-normal destruction for the $u_i$ can be further illustrated by Figure (2.A1).

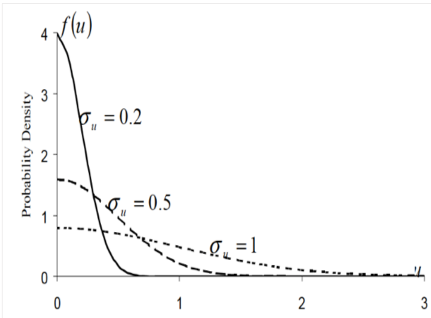



**Figure 2.A1:** Half-Normal Distribution

Assuming a half normal distribution for the inefficiency term of the composed error, the density function of $u_i$≥ 0 and $v_i$ are given by:

$$f(u) = \frac{2}{\sqrt{2\pi}\ \ \sigma_u} exp\left\{-\frac{u^2}{2\sigma_u^2}\right\}$$

$$f(v) = \frac{2}{2\pi\sigma_v} exp\left\{-\frac{v^2}{2\sigma_v^2}\right\}$$

Given the basic assumption of the stochastic frontier models that $u_i$ and $v_i$ are independent from each other, the join density function of ui and vi is the product of their individual density function, is given by:

$$f(u, v) = \frac{2}{2\pi\sigma_u\sigma_v} exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right\}$$

And because $\varepsilon_i = v_i - u$, the join density function of u and ε can be specified as:

$$f(u, \varepsilon) = \frac{2}{\sqrt{2\pi}\ \sigma_u\ \sigma_v}\ exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{(\varepsilon + u)^2}{2\sigma_v^2}\right\}$$

The marginal density function of $\varepsilon$ can be obtained by integrating u out of $f(u, \varepsilon)$ which yields:

$$f(\varepsilon) = \int_0^\infty \int (u, \varepsilon) du$$
$$= \frac{2}{\sqrt{2\pi}\ \sigma_u}\left[1 - \Phi\left(\frac{\varepsilon\lambda}{\sigma}\right)\right] exp\left\{-\frac{\varepsilon^2}{2\sigma^2}\right\}$$
$$= \frac{2}{\sigma}\phi\left(\frac{\varepsilon}{\sigma}\right)\Phi\left\{-\frac{\varepsilon\lambda}{\sigma}\right\}$$

Where $\sigma^2 = \sigma_u^2 + \sigma_v^2$ , $\lambda = {}^{\sigma_u}/_{\sigma_v}$ , and $\Phi(.)$ and $\phi(.)$ are the standard normal cumulative distribution density functions. The parameter of $\lambda$ represents degree of asymmetry of the distribution of the error term. The larger $\lambda$ is, the more pronounced the asymmetry will be. On the other hand, if $\lambda$ is equal to zero, then the symmetric error component dominates the one-side error component in the determination of $\varepsilon_i$. Therefore, the complete error term is explained by the random disturbance $v_i$, which follows a normal distribution. $\varepsilon_i$ therefore has a normal distribution.

The marginal density function $f(\varepsilon)$ is asymmetrically distributed with mean and variance of:

$$E(\varepsilon) = -E(u) = -\sigma_u\sqrt{\frac{2}{\pi}}$$
$$V(\varepsilon) = \frac{\pi - 2}{\pi}\sigma_u^2 + \sigma_v^2$$

The log-likelihood function for the normal - half normal stochastic frontier model is:

$$\ln(L) = -\left(\frac{N}{2}\right)(ln2\pi + ln\sigma^2) + \sum_{i=1}^{N}\left[\ln\phi\left[-\epsilon_i\, {}^{\lambda}/_{\sigma}\right] - \frac{1}{2}\ ({}^{\varepsilon_i}/_{\sigma})^2\right]$$

Meanwhile, Jondrow et al. (1982) also computed the expected value of $u_i$ conditional on the composed error term for the case in which the asymmetric error term follows an exponential distribution. They provided the following result:

$$f(u, \varepsilon) = \frac{f(u, \varepsilon)}{f(\varepsilon)}$$

$$= \frac{1}{\sqrt{2\pi}\ \ \sigma_*}\ exp\left\{-\frac{(u-\mu_*)^2}{2\sigma_*^2}\right\}\Big/ exp\left\{-\frac{(u-\mu_*)^2}{2\sigma_*^2}\right\}$$

Where $\mu_* = -\varepsilon\sigma_u^2/\sigma^2$ and $\sigma_*^2 = \sigma_u^2\sigma_v^2/\sigma^2$. Since $f(u|\varepsilon)$ is distributes as $N^+(\mu_*, \sigma_*^2)$ the mean of this distribution can serve as; point estimated of ui which is given by:

$$E(u_i|\varepsilon_i) = \mu_{*_i} + \sigma_* \left[\frac{\phi\left(\mu_{*_i}/\sigma_*\right)}{1-\Phi\left(-\mu_{*_i}/\sigma_*\right)}\right]$$

$$= \sigma_* \left[\frac{\phi(\varepsilon_i\lambda/\sigma)}{1-\Phi\big((\varepsilon_i\lambda/\sigma)\big)} - \frac{\varepsilon_i\lambda}{\sigma}\right]$$

Therefore, the estimates of $u_i$ can be obtained from the following specification:

$$TE_i = exp\{\hat{u}\} = exp\{-E(u_i|\varepsilon_i)\}$$

## **Appendix 2.B:** Detailed Descriptive Statistics

**Table 2.B1:** Percent of Zero Values in Input Variables

| Variable | Non-zero values | Zeros | %of Zero Values |
|---|---|---|---|
| Total Revenue (Y) | 7,052 | None | 0% |
| Land (Jeribs) | 7,052 | None | 0% |
| Labour (hours) | 7,052 | None | 0% |
| Seed Expenditures (AFN) | 4,582 | 2,470 | 35% |
| Fertilizer Expenditures (AFN) | 4,935 | 2,117 | 30% |
| Chemicals Expenditures (AFN) | 1,594 | 5,458 | 77% |
| Tractor Rental Expenditures (AFN) | 4,078 | 2,974 | 42% |
| Other Cost Expenditures (AFN) | 3,048 | 4,004 | 57% |
| N | | | 7,052 |

**Table 2.B2:** Characteristics of Households with CDI below & above Median (0.37)

| | Specialized | Diversified | Two-Tailed T Test | |
|---|---|---|---|---|
| | Mean | Mean | Difference | SE |
| Aggregate Annual Revenue (AFN) | 46,884 | 69,621 | -22,737*** | -10.6 |
| Land (Jeribs) | 6.77 | 7.30 | -0.530** | -2.44 |
| Labour (hours) | 60.16 | 67.41 | -7.253*** | -4.89 |
| Seed Expenditure (AFN) | 2,389 | 2,320 | 69.33 | -0.8 |
| Fertilizer Expenditure (AFN) | 3,582 | 5,945 | -2,363*** | -12.1 |
| Chemicals Expenditure (AFN) | 228.0 | 502.1 | -274*** | -9.68 |
| Tractor Rental (AFN) | 2,460 | 2,549 | -88.7 | -0.87 |
| Other Expenditure (AFN) | 797 | 749 | 47.98 | -0.91 |
| Herfindahl Index (HHI) | 0.91 | 0.50 | 0.414*** | -163 |
| Household Size (persons) | 7.91 | 8.76 | -0.842*** | -10.3 |
| Head Age (years) | 44.22 | 44.55 | -0.325 | -0.99 |
| Head Sex (1=male, 0=female) | 1.00 | 1.00 | -0.0014 | -1.00 |
| Extension Services (1=access, 0=No) | 0.19 | 0.23 | -0.043*** | -4.49 |
| Head Literacy (1=yes, 0=otherwise) | 0.31 | 0.34 | -0.0315*** | -2.83 |
| Off-farm Employment (1=yes, 0=No) | 0.12 | 0.13 | -0.0199** | -2.53 |
| Cattle (number) | 1.33 | 1.87 | -0.533*** | -11.1 |
| Tractors (number) | 0.05 | 0.06 | -0.0102 | -1.85 |
| Oxen (number) | 0.23 | 0.24 | -0.00851 | -0.58 |
| N | | | | 7,052 |

*Significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

**Table 2.B3:** Characteristics of Farms by Access to Extension Services

| | Access | | No Access | |
|---|---|---|---|---|
| Variable | Mean | SD | Mean | SD |
| Aggregate Annual Revenue (AFN) | 59,238 | 94,208 | 57,992 | 89,415 |
| Land (Jeribs) | 6.25 | 9.15 | 7.24 | 9.10 |
| Labour (hours) | 61.54 | 65.10 | 64.38 | 61.63 |
| Seed Expenditure (AFN) | 2,289 | 4,157 | 2,372 | 3,460 |
| Fertilizer Expenditure (AFN) | 4,969 | 7,057 | 4,709 | 8,596 |
| Chemicals Expenditure (AFN) | 284.1 | 876.7 | 386.4 | 1,267 |
| Tractor Rental (AFN) | 2,443 | 4,293 | 2,520 | 4,297 |
| Other Expenditure (AFN) | 716.1 | 2,151 | 787.8 | 2,237 |
| Herfindahl Index (HHI) | 0.69 | 0.23 | 0.71 | 0.23 |
| Crop Diversification Index (CDI) | 0.31 | 0.23 | 0.29 | 0.23 |
| Household Size (persons) | 9.06 | 3.87 | 8.14 | 3.31 |
| Head Age (years) | 45.28 | 14.20 | 44.15 | 13.63 |
| Head Sex (1=male, 0=female) | 1.00 | 0.03 | 1.00 | 0.07 |
| Head Literacy (1=yes, 0=otherwise) | 0.44 | 0.50 | 0.29 | 0.45 |
| Off-farm Employment (1=yes, 0=No) | 0.20 | 0.40 | 0.11 | 0.31 |
| Cattle (number) | 1.91 | 1.85 | 1.52 | 2.08 |
| Tractors (number) | 0.06 | 0.25 | 0.05 | 0.23 |
| Oxen (number) | 0.15 | 0.55 | 0.25 | 0.63 |
| N | 1,473 | | 5,579 | |

**Table 2.B4:** Characteristics of farm Households in Provinces with less and more than 1% of Opium Production

| | Less Than 1% | | 1% & more than 1% | |
|---|---|---|---|---|
| Variable | Mean | SD | Mean | SD |
| Aggregate Annual Revenue (AFN) | 51,902 | 82,492 | 82,181 | 112,000 |
| Land (Jeribs) | 6.88 | 9.41 | 7.61 | 7.90 |
| Labour (hours) | 62.44 | 62.44 | 68.88 | 61.89 |
| Seed Expenditure (AFN) | 2,236 | 3,660 | 2,800 | 3,415 |
| Fertilizer Expenditure (AFN) | 3,497 | 5,609 | 9,536 | 13,457 |
| Chemicals Expenditure (AFN) | 146.5 | 531.2 | 1,188 | 2,216 |
| Tractor Rental (AFN) | 2,130 | 4,036 | 3,916 | 4,909 |
| Other Expenditure (AFN) | 758 | 2,115 | 829 | 2,577 |
| Herfindahl Index (HHI) | 0.70 | 0.23 | 0.71 | 0.24 |
| Crop Diversification Index (CDI) | 0.30 | 0.23 | 0.29 | 0.24 |
| Household Size (persons) | 8.20 | 3.37 | 8.85 | 3.73 |
| Head Age (years) | 44.9 | 13.80 | 42.4 | 13.4 |
| Head Sex (1=male, 0=female) | 1.00 | 0.07 | 1.00 | 0.03 |
| Extension Services (1=access, 0=No) | 0.22 | 0.42 | 0.17 | 0.37 |
| Head Literacy (1=yes, 0=otherwise) | 0.35 | 0.48 | 0.21 | 0.41 |
| Off-farm Employment (1=yes, 0=No) | 0.13 | 0.33 | 0.11 | 0.32 |
| Cattle (number) | 1.69 | 2.13 | 1.27 | 1.60 |
| N | 5,573 | | 1,479 | |

Table 2.B5: Distribution of TE and CD by Equivalent Number (1/HH)

| Equivalent Number | Mean TE (%) | Mean CD | No. of Farms |
|---|---|---|---|
| 1 | 54.12 | - | 2,198 |
| 1-<2 | 78.33 | 0.59 | 3,843 |
| 2-<3 | 85.51 | 0.59 | 689 |
| 3-<4 | 88.76 | 0.71 | 93 |
| 4-<5 | 91.22 | 0.77 | 19 |
| 5-<6 | 92.28 | 0.82 | 3 |
| N | | | 7,052 |

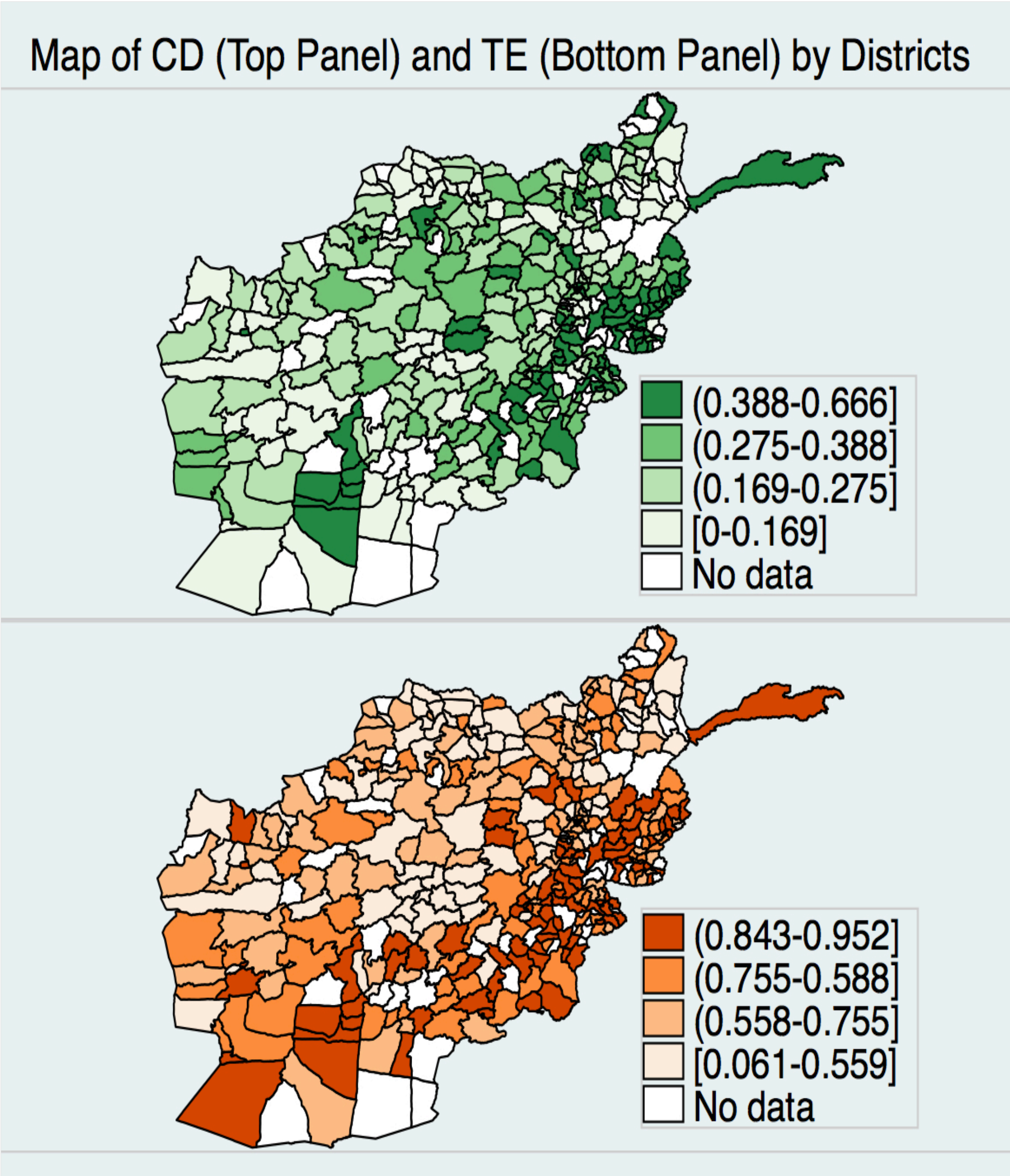



**Figure 2.B1:** Geographical Distribution of CD and TE by Districts
*Source:* Author's calculations of the ALCS 2013-14 data

## Appendix 2.C: Robustness Analysis

**Table 2.C1: SF Model with a Binary Variable for Crop Diversification**

| Variable | Coefficient | SE | | Coefficient | SE |
|---|---|---|---|---|---|
| *Dependent Variable (Total Aggregate Revenue in AFN)* | | | *The Inefficiency Model* | | |
| Constant | 0.138*** | 0.040 | Constant | 1.196*** | 0.418 |
| Ln Land ($X_1$) | 0.458*** | 0.026 | Head Sex (male) | -0.433 | 0.366 |
| Ln Labour ($X_2$) | 0.029 | 0.021 | Head Age (years) | -0.00003 | 0.002 |
| Ln Seed Expenditures ($X_3$) | 0.133*** | 0.015 | Head Edu (lower sec) | 0.113 | 0.159 |
| Ln Fertilizer Expenditures ($X_4$) | 0.194*** | 0.015 | Head Edu (upper sec) | 0.202 | 0.138 |
| Ln Chemical Expenditures ($X_5$) | 0.046*** | 0.014 | Head Edu (college) | 0.170 | 0.243 |
| Ln Tractor Rental ($X_6$) | 0.100*** | 0.017 | Head Edu (uni & grad) | -0.238 | 0.307 |
| Ln other Expenditures ($X_7$) | 0.027** | 0.012 | Head Literacy (1=yes) | -0.006 | 0.090 |
| 0.5 x Ln Land $(X_1)^2$ | 0.024 | 0.016 | Household Size (persons) | 0.004 | 0.011 |
| 0.5 x Ln Labour $(X_2)^2$ | 0.069*** | 0.014 | Diversity (binary, 1=yes) | -1.652*** | 0.119 |
| 0.5 x Ln Seed Expense $(X_3)^2$ | 0.036*** | 0.004 | Ext Services (1=yes) | -0.286*** | 0.095 |
| 0.5 x Ln Fertilizer Expense $(X_4)^2$ | 0.037*** | 0.004 | Oxen & Yaks (n) | -0.137** | 0.057 |
| 0.5 x Ln Chemical Expense $(X_5)^2$ | 0.020*** | 0.006 | Tractor/Threshers (n) | -0.797*** | 0.249 |
| 0.5 Ln Tractor Rental $(X_6)^2$ | 0.028*** | 0.005 | Cattles (number) | -0.111*** | 0.022 |
| 0.5 Ln other Expenditures $(X_7)^2$ | 0.007** | 0.003 | Off-farm Emp (1=yes) | -0.049 | 0.101 |
| Ln Land x Ln Labour | -0.018* | 0.010 | Farm Size (>2-5J) | -0.338*** | 0.093 |
| Ln Land x Ln Seed | -0.007** | 0.003 | Farm Size (>5-10) | -0.181 | 0.115 |
| Ln Land x Ln Fertilizer | 0.006** | 0.003 | Farm Size (>10-20J) | 0.080 | 0.153 |
| Ln Land x Ln Chemicals | 0.002 | 0.004 | Farm Size (>20J) | 0.560*** | 0.194 |
| Ln Land x Ln Tractor Rental | -0.006** | 0.003 | Land Quality (Low) | 0.267*** | 0.088 |
| Ln Land x Ln Other Expenses | 0.008*** | 0.003 | AEZ 1 (CM) | 0.127 | 0.182 |
| Ln Labour x Ln Seed | 0.008*** | 0.003 | AEZ 2 (HFL) | 0.007 | 0.213 |
| Ln Labour x Ln Fertilizer | -0.011*** | 0.003 | AEZ 3 (SMF) | -0.683*** | 0.200 |
| Ln Labour x Ln Chemicals | -0.005 | 0.004 | AEZ 4 (HVSB) | -0.867*** | 0.213 |
| Ln Labour x Ln Tractor Rental | 0.003 | 0.003 | AEZ (TP) | 0.208 | 0.199 |
| Ln Labour x Ln Other Expenses | -0.006** | 0.003 | AEZ 6 (NMF) | -0.140 | 0.174 |
| Ln Seed x Ln Fertilizer | 0.001 | 0.001 | AEZ 7 (EMF) | -0.496*** | 0.189 |
| Ln Seed x Ln Chemicals | -0.003*** | 0.001 | | | |
| Ln Seed x Ln Tractor Rental | -0.002** | 0.001 | | | |
| Ln Seed x Ln Other Expenses | -0.002*** | 0.001 | | | |
| Ln Fertilizer x Ln Chemicals | -0.001 | 0.001 | | | |
| Ln Fertilizer x Ln Tractor Rental | -0.003*** | 0.001 | | | |
| Ln Fertilizer x Ln Other Expenses | 0.002*** | 0.001 | | | |
| Ln Chemicals x Ln Tractor Rental | -0.001 | 0.001 | | | |
| Ln Chemicals x Ln Other Expenses | -0.001 | 0.001 | | | |
| Ln Tractor Rental x Ln Other Expenses | -0.002*** | 0.001 | | | |
| $(\sigma)^2$ | 0.267*** | 0.000 | | | |
| **$\gamma$** | 0.899*** | 0.000 | | | |
| Log-Likelihood | -7,495.62 | | | | |
| Mean Efficiency | 71.94% | | | | |
| N | 7,059 | | | | |

*Note: significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

**Table 2.C2:** Testing the Correlation of the Instrumental Variable (IV)

| Variable | Coefficient | SE |
|---|---|---|
| *Dependent Variable - Crop Diversification Index (CDI=1-HHI)* | | |
| Instrument (IV) | 0.710*** | 0.022 |
| Ln Land ($X_1$) | 0.015 | 0.012 |
| Ln Labour ($X_2$) | -0.005 | 0.006 |
| Ln Seed Expenditures ($X_3$) | -0.005 | 0.004 |
| Ln Fertilizer Expenditures ($X_4$) | 0.007* | 0.004 |
| Ln Chemical Expenditures ($X_5$) | -0.008 | 0.005 |
| Ln Tractor Rental ($X_6$) | -0.001 | 0.005 |
| Ln other Expenditures ($X_7$) | -0.023*** | 0.003 |
| 0.5 x Ln Land squared($X_1^2$) | -0.014** | 0.006 |
| 0.5 x Ln Labour squared ($X_2^2$) | 0.001 | 0.004 |
| 0.5 x Ln Seed Expenditures squared ($X_3^2$) | -0.001 | 0.001 |
| 0.5 x Ln Fertilizer Expenditures squared ($X_4^2$) | 0.001 | 0.001 |
| 0.5 x Ln Chemical Expenditures squared ($X_5^2$) | -0.004* | 0.002 |
| 0.5 Ln Tractor Rental squared ($X_6^2$) | -0.000 | 0.001 |
| 0.5 Ln other Expenditures squared ($X_7^2$) | -0.006*** | 0.001 |
| Ln Land x Ln Labour | -0.008*** | 0.003 |
| Ln Land x Ln Seed | -0.002** | 0.001 |
| Ln Land x Ln Fertilizer | -0.002*** | 0.001 |
| Ln Land x Ln Chemicals | -0.001 | 0.001 |
| Ln Land x Ln Tractor Rental | 0.000 | 0.001 |
| Ln Land x Ln Other Expenses | 0.002** | 0.001 |
| Ln Labour x Ln Seed | -0.000 | 0.001 |
| Ln Labour x Ln Fertilizer | -0.001 | 0.001 |
| Ln Labour x Ln Chemicals | 0.001 | 0.001 |
| Ln Labour x Ln Tractor Rental | -0.002** | 0.001 |
| Ln Labour x Ln Other Expenses | -0.001 | 0.001 |
| Ln Seed x Ln Fertilizer | 0.000* | 0.000 |
| Ln Seed x Ln Chemicals | -0.001** | 0.000 |
| Ln Seed x Ln Tractor Rental | -0.000 | 0.000 |
| Ln Seed x Ln Other Expenses | -0.000 | 0.000 |
| Ln Fertilizer x Ln Chemicals | -0.000 | 0.000 |
| Ln Fertilizer x Ln Tractor Rental | 0.000** | 0.000 |
| Ln Fertilizer x Ln Other Expenses | 0.001*** | 0.000 |
| Ln Chemicals x Ln Tractor Rental | -0.000 | 0.000 |
| Ln Chemicals x Ln Other Expenses | -0.001** | 0.000 |
| Ln Tractor Rental x Ln Other Expenses | 0.000 | 0.000 |
| Head Sex (male) | 0.038 | 0.039 |
| Head Age (years) | -0.000 | 0.000 |
| Head Education (lower secondary) | -0.013 | 0.012 |
| Head Education (upper secondary) | 0.004 | 0.010 |
| Head Education (teacher college) | -0.012 | 0.017 |
| Head Education (Uni & postgrad) | 0.010 | 0.021 |
| Head Literacy (can read & write) | 0.004 | 0.007 |
| Household Size (persons) | 0.001 | 0.001 |
| Extension Services (1=yes) | -0.014** | 0.006 |

| | | *Table 2.C2 Continued* |
|---|---|---|
| Oxen and Yaks (number) | 0.021*** | 0.004 |
| Tractor/Threshers (number) | 0.007 | 0.011 |
| Number of Cattles (number) | 0.001 | 0.001 |
| Off-farm Employment (1=yes) | 0.031*** | 0.007 |
| Opium share by province (%) | 0.108*** | 0.042 |
| Farm Size (>2 to 5 Jeribs) | 0.010 | 0.009 |
| Farm Size (>5 to 10 Jeribs) | 0.023 | 0.014 |
| Farm Size (>10 to 20 Jeribs) | 0.058*** | 0.022 |
| Farm Size (>20 & above Jeribs) | 0.085** | 0.034 |
| Land Quality (Low) | -0.080*** | 0.008 |
| Agro-ecological Zone 1 (CM) | 0.036** | 0.018 |
| Agro-ecological Zone 2 (HFL) | -0.035* | 0.021 |
| Agro-ecological Zone 3 (SMF) | 0.021 | 0.018 |
| Agro-ecological Zone 4 (HVSB) | -0.041** | 0.020 |
| Agro-ecological Zone 5 (TP) | -0.074*** | 0.019 |
| Agro-ecological Zone 6 (NMF) | 0.019 | 0.017 |
| Agro-ecological Zone 7 (EMF) | 0.083*** | 0.018 |
| Constant | 0.088* | 0.046 |
| Log-Likelihood | | 1,682.54 |
| R2 | | 0.331 |
| Test of Endogeneity-Durbin (score) chi2(1)a | | 39.82 |
| Test of Endogeneity-Wu-Hausman F(1,7009)a | | 39.68 |
| Test of Weak IV- F statistic | | 1,075.57 |
| N | | 7,038 |

*Note: The omitted categories are: no formal schooling for education level, <2 Jeribs for farm size, no access for extension services, cannot read & write for literacy, none for off-farm employment, irrigated & rainfed combined for land quality, and agro-ecological zone 8 for AEZ; significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010. H0: CDI is exogenous, rejected. The P-value is (p=0.000)b H0: Instrument is weak, rejected. The P-value is (p=0.000).*

**Table 2.C3:** Estimates of the Stochastic Frontier Model for sub-sample of HHs in Provinces with Greater Prevalence of Opium Production

| | Exogenous | | Endogenous | |
|---|---|---|---|---|
| Variable | Coefficient | SE | Coefficient | SE |
| *Dependent Variable (Total Aggregate Revenue in AFN)* | | | | |
| Constant | -0.112* | 0.065 | -0.134* | 0.065 |
| Ln Land ($X_1$) | 0.453*** | 0.043 | 0.470*** | 0.045 |
| Ln Labour ($X_2$) | -0.057** | 0.027 | -0.077*** | 0.028 |
| Ln Seed Expenditures ($X_3$) | 0.090*** | 0.026 | 0.104*** | 0.027 |
| Ln Fertilizer Expenditures ($X_4$) | 0.165*** | 0.025 | 0.172*** | 0.026 |
| Ln Chemical Expenditures ($X_5$) | 0.120*** | 0.020 | 0.135*** | 0.021 |
| Ln Tractor Rental ($X_6$) | 0.144*** | 0.030 | 0.150*** | 0.031 |
| Ln other Expenditures ($X_7$) | 0.053*** | 0.017 | 0.036** | 0.018 |
| 0.5 x Ln Land squared($X_1^2$) | 0.144*** | 0.036 | 0.150*** | 0.036 |
| 0.5 x Ln Labour squared ($X_2^2$) | 0.003 | 0.033 | -0.002 | 0.033 |
| 0.5 x Ln Seed Expenditures squared ($X_3^2$) | 0.030*** | 0.007 | 0.033*** | 0.007 |
| 0.5 x Ln Fertilizer Expenditures squared ($X_4^2$) | 0.040*** | 0.007 | 0.043*** | 0.007 |
| 0.5 x Ln Chemical Expenditures squared ($X_5^2$) | 0.046*** | 0.009 | 0.051*** | 0.009 |
| 0.5 Ln Tractor Rental squared ($X_6^2$) | 0.035*** | 0.008 | 0.035*** | 0.009 |
| 0.5 Ln other Expenditures squared ($X_7^2$) | 0.017*** | 0.005 | 0.012** | 0.005 |
| Ln Land x Ln Labour | 0.003 | 0.023 | 0.008 | 0.023 |
| Ln Land x Ln Seed | -0.015** | 0.006 | -0.016** | 0.006 |
| Ln Land x Ln Fertilizer | -0.009 | 0.006 | -0.010 | 0.006 |
| Ln Land x Ln Chemicals | -0.011* | 0.006 | -0.012* | 0.007 |
| Ln Land x Ln Tractor Rental | -0.007 | 0.006 | -0.007 | 0.006 |
| Ln Land x Ln Other Expenses | -0.000 | 0.005 | 0.001 | 0.005 |
| Ln Labour x Ln Seed | -0.014** | 0.006 | -0.016** | 0.007 |
| Ln Labour x Ln Fertilizer | -0.018*** | 0.006 | -0.020*** | 0.007 |
| Ln Labour x Ln Chemicals | 0.003 | 0.006 | 0.002 | 0.006 |
| Ln Labour x Ln Tractor Rental | 0.001 | 0.006 | 0.002 | 0.006 |
| Ln Labour x Ln Other Expenses | -0.008* | 0.005 | -0.010** | 0.005 |
| Ln Seed x Ln Fertilizer | -0.002 | 0.002 | -0.001 | 0.002 |
| Ln Seed x Ln Chemicals | 0.005*** | 0.002 | 0.005*** | 0.002 |
| Ln Seed x Ln Tractor Rental | -0.001 | 0.002 | -0.001 | 0.002 |
| Ln Seed x Ln Other Expenses | -0.001 | 0.001 | -0.001 | 0.001 |
| Ln Fertilizer x Ln Chemicals | -0.002 | 0.002 | -0.002 | 0.002 |
| Ln Fertilizer x Ln Tractor Rental | 0.001 | 0.001 | 0.000 | 0.001 |
| Ln Fertilizer x Ln Other Expenses | 0.002 | 0.001 | 0.003* | 0.001 |
| Ln Chemicals x Ln Tractor Rental | -0.008*** | 0.002 | -0.008*** | 0.002 |
| Ln Chemicals x Ln Other Expenses | -0.002 | 0.001 | -0.002* | 0.001 |
| Ln Tractor Rental x Ln Other Expenses | 0.001 | 0.001 | 0.002 | 0.001 |
| ***The Inefficiency Model*** | | | | |
| Constant | 0.353 | 0.954 | 0.267 | 0.993 |
| Head Sex (male) | -0.071 | 0.932 | -0.419 | 0.973 |
| Head Age (years) | -0.012** | 0.005 | -0.012** | 0.005 |
| Head Education (lower secondary) | -0.459 | 0.414 | -0.655 | 0.466 |
| Head Education (upper secondary) | 0.080 | 0.343 | 0.044 | 0.351 |
| Head Education (teacher college) | -0.002 | 0.905 | -0.071 | 0.988 |

| | | | | |
|---|---|---|---|---|
| | | | | ***Table 2.C3 Continue*** |
| Head Education (university & postgrad) | -0.767 | 0.951 | -0.696 | 1.091 |
| Head Literacy (can read & write) | 0.187 | 0.177 | 0.238 | 0.179 |
| Household Size (persons) | -0.002 | 0.022 | -0.003 | 0.023 |
| Diversification Index (0≤ THI≤ 1) | -4.610*** | 0.499 | -2.343*** | 0.519 |
| Extension Services (1=yes, 0 otherwise) | -0.118 | 0.237 | -0.053 | 0.240 |
| Number of Oxen and Yaks (number) | -0.307** | 0.148 | -0.380** | 0.153 |
| Number of Tractor/Threshers (number) | -0.685 | 0.462 | -0.851 | 0.590 |
| Number of Cattles (number) | 0.004 | 0.046 | -0.011 | 0.053 |
| Off-farm Employment (1=yes, 0 otherwise) | 0.404* | 0.212 | 0.213 | 0.222 |
| Farm Size (>2 to 5 Jeribs) | -0.526*** | 0.182 | -0.698*** | 0.206 |
| Farm Size (>5 to 10 Jeribs) | -0.100 | 0.210 | -0.247 | 0.223 |
| Farm Size (>10 to 20 Jeribs) | 0.523* | 0.307 | 0.302 | 0.309 |
| Farm Size (>20 & above Jeribs) | 1.434*** | 0.393 | 1.150*** | 0.377 |
| Land Quality (Low) | 0.029 | 0.288 | 0.245 | 0.326 |
| Agro-ecological Zone 1 (CM) | 1.189*** | 0.396 | 1.456*** | 0.418 |
| Agro-ecological Zone 2 (HFL) | 0.007 | 0.354 | 0.204 | 0.391 |
| Agro-ecological Zone 4 (HVSB) | -0.244 | 0.325 | -0.157 | 0.360 |
| Agro-ecological Zone 6 (NMF) | 0.181 | 0.279 | 0.110 | 0.302 |
| Agro-ecological Zone 7 (EMF) | -0.794 | 0.553 | -2.021* | 1.035 |
| $(\sigma)^2$ | 0.237*** | 0.000 | | |
| $\boldsymbol{\gamma}$ | 0.857*** | 0.000 | | |
| Log-Likelihood | | -1,667.22 | | -1,092.10 |
| eta Endogeneity Test | - | - | 34.48*** | 0.00 |
| eta1_THI | - | - | 0.720*** | 0.123 |
| Mean Efficiency | | 70.03% | | 0.710 |
| N | | 1,890 | | 1,885 |

*Note: The omitted categories are: No formal schooling for education level, <2 Jeribs for farm size, no access for extension services, can't read & write for literacy, none for off-farm employment, irrigated & rainfed combined for land quality, and AEZ 8 for AEZ; significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

**Table 2.C4:** Estimates of the Stochastic Frontier Model for Subsample of HHs in Provinces with no Opium Production

| | Exogenous | | Endogenous | |
|---|---|---|---|---|
| Variable | Coefficient | SE | Coefficient | SE |
| *Dependent Variable (household's aggregate annual revenue in AFN)* | | | | |
| Constant | 0.323*** | 0.051 | 0.244*** | 0.053 |
| Ln Land ($X_1$) | 0.395*** | 0.033 | 0.393*** | 0.033 |
| Ln Labour ($X_2$) | 0.059** | 0.029 | 0.051* | 0.030 |
| Ln Seed Expenditures ($X_3$) | 0.155*** | 0.019 | 0.151*** | 0.019 |
| Ln Fertilizer Expenditures ($X_4$) | 0.250*** | 0.018 | 0.252*** | 0.019 |
| Ln Chemical Expenditures ($X_5$) | 0.097*** | 0.026 | 0.107*** | 0.026 |
| Ln Tractor Rental ($X_6$) | 0.053** | 0.021 | 0.046** | 0.021 |
| Ln other Expenditures ($X_7$) | 0.025 | 0.017 | 0.039** | 0.017 |
| 0.5 x Ln Land $(X_1)^2$ | -0.025 | 0.017 | -0.038** | 0.017 |
| 0.5 x Ln Labour $(X_2)^2$ | 0.069*** | 0.015 | 0.067*** | 0.015 |
| 0.5 x Ln Seed Expenditures $(X_3)^2$ | 0.033*** | 0.005 | 0.032*** | 0.005 |
| 0.5 x Ln Fertilizer Expenditures $(X_4)^2$ | 0.047*** | 0.004 | 0.048*** | 0.004 |
| 0.5 x Ln Chemical Expenditures $(X_5)^2$ | 0.034*** | 0.010 | 0.037*** | 0.010 |
| 0.5 Ln Tractor Rental $(X_6)^2$ | 0.018*** | 0.006 | 0.015*** | 0.006 |
| 0.5 Ln other Expenditures $(X_7)^2$ | 0.004 | 0.004 | 0.007* | 0.004 |
| Ln Land x Ln Labour | -0.021* | 0.011 | -0.018* | 0.011 |
| Ln Land x Ln Seed | -0.005 | 0.003 | -0.004 | 0.003 |
| Ln Land x Ln Fertilizer | 0.009*** | 0.003 | 0.012*** | 0.003 |
| Ln Land x Ln Chemicals | -0.003 | 0.005 | -0.002 | 0.005 |
| Ln Land x Ln Tractor Rental | -0.001 | 0.003 | 0.000 | 0.003 |
| Ln Land x Ln Other Expenses | 0.011*** | 0.003 | 0.010*** | 0.003 |
| Ln Labour x Ln Seed | 0.013*** | 0.003 | 0.013*** | 0.003 |
| Ln Labour x Ln Fertilizer | -0.012*** | 0.003 | -0.013*** | 0.003 |
| Ln Labour x Ln Chemicals | -0.004 | 0.005 | -0.004 | 0.005 |
| Ln Labour x Ln Tractor Rental | 0.002 | 0.003 | 0.003 | 0.003 |
| Ln Labour x Ln Other Expenses | -0.004 | 0.003 | -0.004 | 0.003 |
| Ln Seed x Ln Fertilizer | 0.002** | 0.001 | 0.002** | 0.001 |
| Ln Seed x Ln Chemicals | -0.003*** | 0.001 | -0.003** | 0.001 |
| Ln Seed x Ln Tractor Rental | -0.000 | 0.001 | -0.000 | 0.001 |
| Ln Seed x Ln Other Expenses | 0.000 | 0.001 | 0.001 | 0.001 |
| Ln Fertilizer x Ln Chemicals | 0.002 | 0.001 | 0.002 | 0.001 |
| Ln Fertilizer x Ln Tractor Rental | -0.003*** | 0.001 | -0.004*** | 0.001 |
| Ln Fertilizer x Ln Other Expenses | 0.002* | 0.001 | 0.001 | 0.001 |
| Ln Chemicals x Ln Tractor Rental | -0.001 | 0.001 | -0.001 | 0.001 |
| Ln Chemicals x Ln Other Expenses | 0.000 | 0.001 | 0.001 | 0.001 |
| Ln Tractor Rental x Ln Other Expenses | -0.004*** | 0.001 | -0.004*** | 0.001 |
| ***The Inefficiency Model*** | | | | |
| Constant | 0.968 | 0.797 | 1.106 | 0.888 |
| Head Sex (male) | -0.502 | 0.461 | -0.469 | 0.514 |
| Head Age (years) | 0.005 | 0.003 | 0.006 | 0.003 |
| Head Education (lower secondary) | 0.214 | 0.196 | 0.241 | 0.223 |
| Head Education (upper secondary) | 0.250 | 0.173 | 0.289 | 0.197 |
| Head Education (teacher college) | 0.237 | 0.281 | 0.299 | 0.324 |

| | | | | |
|---|---|---|---|---|
| | | | | ***Table 2.C4 Continue*** |
| Head Education (university & postgrad) | -0.266 | 0.378 | -0.059 | 0.420 |
| Head Literacy (can read & write) | -0.014 | 0.120 | -0.092 | 0.138 |
| Household Size (persons) | -0.006 | 0.015 | -0.004 | 0.017 |
| Diversification Index (THI) | -4.048*** | 0.409 | -7.384*** | 1.108 |
| Extension Services (1=yes, 0 otherwise) | -0.315*** | 0.118 | -0.479*** | 0.138 |
| Number of Oxen and Yaks (number) | -0.229*** | 0.081 | -0.255*** | 0.093 |
| Number of Tractor/Threshers (number) | -1.035*** | 0.317 | -0.859** | 0.344 |
| Number of Cattles (number) | -0.136*** | 0.027 | -0.113*** | 0.030 |
| Off-farm Employment (1=yes, 0 otherwise) | -0.022 | 0.128 | -0.030 | 0.144 |
| Farm Size (>2 to 5 Jeribs) | -0.373*** | 0.120 | -0.238* | 0.133 |
| Farm Size (>5 to 10 Jeribs) | -0.522*** | 0.157 | -0.401** | 0.169 |
| Farm Size (>10 to 20 Jeribs) | -0.366* | 0.202 | -0.225 | 0.219 |
| Farm Size (>20 & above Jeribs) | -0.095 | 0.252 | -0.026 | 0.284 |
| Land Quality (Low/irrigated and rainfed combined) | 0.331*** | 0.111 | 0.049 | 0.139 |
| Agro-ecological Zone 1 (CM) | -0.071 | 0.637 | 0.049 | 0.139 |
| Agro-ecological Zone 2 (HFL) | 0.458 | 0.671 | -0.138 | 0.709 |
| Agro-ecological Zone 3 (SMF) | -0.797 | 0.643 | 0.224 | 0.752 |
| Agro-ecological Zone 5 (TP) | 0.263 | 0.639 | -0.861 | 0.713 |
| Agro-ecological Zone 6 (NMF) | 0.049 | 0.634 | 0.073 | 0.712 |
| Agro-ecological Zone 7 (EMF) | -0.364 | 0.637 | -0.035 | 0.705 |
| $(\sigma)^2$ | 0.383*** | 0.000 | | |
| $\boldsymbol{\gamma}$ | 0.873*** | 0.000 | | |
| Log-Likelihood | | -5,601.98 | | -4,338.39 |
| eta Endogeneity Test | - | - | 39.15*** | 0.000 |
| eta1_cdi | - | - | -0.514*** | 0.082 |
| Mean Efficiency | | 69.21% | | 75.02% |
| N | | 5,169 | | 5,160 |

*Note: The omitted categories are: No formal schooling for education level, <2 Jeribs for farm size, no access for extension services, can't read & write for literacy, none for off-farm employment, irrigated & rainfed combined for land quality, and AEZ 8 for AEZ; significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

# 3 Chapter III: Non-farm Income and Crop Diversification in Afghanistan

**Abstract:**

Using data from 8,613 farm households collected by the Afghanistan Living Condition Survey (ALCS) in 2013/14, the analysis in this chapter estimates micro-economic drivers of diversity in crop production with particular emphasis on the implications of household's access to non-farm income on the level of Crop Diversification (CD). CD is measured by the Composite Entropy Index (CEI) as this incorporates crop (revenue) shares weighted by the total number of crops grown. The level of CD is relatively low, with a mean CEI of 0.29 (where zero is no diversification): a third of farmers do not diversify, and the majority that do, grow only two or three crops. Our econometric analysis reveals that while landholding size, access to irrigation, quality of land, household size, and ownership of tractor, oxen, and cattle by the farm households significantly increase the level of CD, a significantly lower degree of CD is found for farm households in communities with low access to all-season derivable road and households with higher non-farm income. This is consistent with the hypothesis that allocation of farm labour away to non-farm activities decrease diversity due to negative labour effects, mainly because the opportunity cost of household labour is higher than the off-farm wages under imperfect markets implying non-separability between households' farm profits and off-farm earnings. Identification through instrumental variables confirm endogeneity in off-farm income revealing that unobserved that risk-aversion behaviour of farmers drive household's decision to diversify into both non-farm income and crop mix.

## 3.1 Introduction

A crucial component in the farming business is to understand the decision-making environment and behaviour of farm households, particularly subsistence smallholders that are often exposed to various types of risk and uncertainties. Understanding these decisions such as allocation of limited resources among diverse crops requires empirical evidence. Traditionally, crop diversification is regarded as a management strategy, particularly in the context of subsistence farming, where farmers choose the appropriate crop mix to sustain their consumption (diet), livelihoods, and income. Previous studies have demonstrated the economic value of crop diversification as an alternative strategy that farmers can utilize to survive and even increase incomes. Given the importance of crop diversity, this paper aims to explore farm households' decisions with regard to the magnitude of crop diversification at the micro level in Afghanistan. It analyses the status, patterns, and extent of crop diversification, and the empirical relationship between crop diversification and household socio-economic, demographic, farm, and farmer characteristics. The key focus is to examine the impact of the household off-farm income on the level of crop diversity.

Heterogeneity in farmer crop portfolio in a given location and under certain socio-economic circumstances is an important empirical discussion. Even in the presence of high-return alternatives both on- and off-farm, a large number of farm households still engage in producing low value food commodities (mainly staple food grains), and crop portfolio choices vary among similar households (Stoeffler, 2016). Farmer's knowledge, technical know-how, and production management practices have significant implications on their income and costs. Without incurring additional costs, there is a high potential for many farm households to improve their productivity and income just by adding high value crops to their production agenda.

Afghanistan's agriculture is highly dominated by the production of staple food grains. Wheat occupies the major portion of the agriculture land, followed by other grains such as maize, barley and rice. There is evidence that the grain-based production systems may not continue to contribute as significantly in countries with a policy focus on raising incomes and production of high value market crops, generating employment opportunities, and alleviating poverty (Joshi et al., 2007). The decrease in farm income among rice/wheat producers in Afghanistan due to the declining productivity has triggered a change toward farm diversification. Afghan farmers need to diversify their farming system into mixed crop-livestock systems and shift production from staple crops to higher value commodities (Oushy, 2010). Thus, both farmers and agricultural policy makers require solid empirical understanding of the production decision-making environment, farmers socio-economic characteristics, and behaviour in order to respond to the changing market demand and consumption patterns. Adding high value horticultural crops such as vegetables and fruits to the existing production system by subsistence farmers is a well-documented production strategy that enhances productivity while improving and sustaining farm incomes and consumption requirements (Joshi et al., 2007; Kumar and Gupta, 2015; Kurosaki, 2003; Weinberger and Lumpkin, 2007).

Recent economic growth was accompanied by significant changes in agriculture production and consumption patterns, whilst other economic sectors such as services and manufacturing industries have been revitalized. With revitalization of other sectors, and improving levels of education, farm households may be enabled to diversify into off-farm activities. This line of reasoning in turn signifies the importance of understanding the potential impact of household off-farm earnings on the extent of crop diversification.

Markets for particular commodities are imperfect and often fail to facilitate efficient trade of farm produce, forcing farmers to adjust their production decisions to compensate for losses due to such market risks. These decisions often involve the choice and degree of crop or enterprise diversification. Hence, to make informed decisions, both farmers and policy makers require empirical evidence that analyses the production environment, particularly whether adopting crop diversification under certain socio-economic conditions is an economically viable strategy.

Development theory suggests that when markets for farm produce are missing, farmers tend to produce mostly staple food crops (primarily grains) mainly for home consumption to be self-sufficient. However, as access to markets become available farm households aim to diversify their production into high value marketable crops to earn cash income. There is a historic evidence in the agricultural economics literature that agricultural diversification in most of the South Asian countries has been demand-driven rather than an outcome of the government policy (Joshi et al., 2007). This implies that farmer's decision on the extent of crop diversification are driven by market conditions and transaction costs.

The remainder of the chapter is structured as follows. Section 3.2 sets out the objective of the analysis presented in the chapter. Section 3.3 overviews relevant literature on diversity and its measures and off-farm income. Section 3.4 discuses methodology and theoretical framework. Section 3.5 presents estimation strategy and empirical specification. Section 3.6 contains decisions on the determinants of crop diversity. Section 3.7 describes data and variables used in the analysis. Section 3.8 presents empirical results. Section 3.9 contains robustness checks for our main results. Section 3.10 summarizes the findings of the study and recommendations.

## 3.2 Objectives and Research Questions

Current studies on production efficiency find a significant and positive relationship between crop diversification and farm level technical efficiency in Afghanistan (Ahmadzai, 2017). Tavva et al., (2017) finds that higher share of land under the wheat crop (in total land) reduce the efficiency of farms in 7 districts of Afghanistan. Broader research also confirms that crop diversification significantly improves farm level technical efficiency in other countries with similar economic context (Coelli and Fleming, 2004; Manjunatha et al., 2013; Ogundari, 2013; Rahman, 2009). However, Afghanistan's agriculture sector is still dominated by production of staple food crops (mainly wheat) and the country's production system remains highly undiversified. It, therefore, calls for significant transformation in agriculture system to diversify towards high value crops such as vegetables and fruits. This transformation to a diversified system that consists of high-value crops will assist farmers to improve production efficiencies, improve and sustain farm income, meet changing dietary requirements, mitigate price and production risks associated with mono-cropping, and improve soil quality through crop rotation.

Using nationally representative household level data, in this study we attempt to analyse the status, patterns, and extent of diversity in crop production, and to investigate empirical relationship between crop diversification and household socio-economic, demographic, farm, and farmer characteristics in Afghanistan. More precisely, the study aims to address the following specific research questions and objectives:

- What inspires farmers' decision to adopt a diversified crop portfolio? Which factors influence the extent of diversification and crop choices of smallholder Afghan farmers?
- How do geographical and socio-economic characteristics of farmers who adopt crop diversification and their counterparts who do not diversify differ?
- Examine heterogeneity in crop diversification based on differences in household off-farm income. How does heterogeneity in off-farm income influence its estimated impact on crop diversity?

There are currently no studies that explicitly focus on empirical relationships between crop diversification and household socio-economic, farm and regional characteristics in Afghanistan and its spill over effects. Understanding this empirical relationship can improve decision-making process at the farm level and generate useful insights and implications for the Afghan policy-makers.

### 3.3 Overview of Related Literature

Although crop diversification is an important part of production decision-making for a farming unit, surprisingly, it has received little empirical attention. Much of the literature on crop diversity adopts an exploratory approach to investigate cropping patterns, trends and factors that affect the decision and extent of crop diversification. There are a small number of empirical studies with econometric analysis of the determinants of producing one or multiple crops (Birthal et al., 2013; Stoeffler, 2016). In this section, we will split the previous findings of literature into two sub-sections focusing on measurement of crop diversity and empirical estimation techniques used in the crop diversity studies. We will then provide a summary of the literature assessing the empirical relationship between household non-farm income and crop intensity.

#### 3.3.1 Concept and Measures of Crop Diversification

The nature of agriculture production is highly volatile and is subject to random shocks often affecting farming output. Crop diversification as a potential mitigation strategy can be devised to hedge against production and market risks. It may be regarded as the reallocation of some of the farm's productive resources, such as land, labour, and other production inputs into different portfolios of activities (i.e. adding new crops to the existing cropping system, a combination of crop and livestock production, value-added post-harvest activities, etc.).

here are two common and complementary approaches to crop diversification in agriculture, namely horizontal and vertical diversification (Behera et al., 2007). Horizontal diversification, which is the primary approach to crop diversification in production agriculture, takes place through crop intensification by adding new crops (usually high-value crops) to existing production line or cropping systems. Vertical diversification, under which, farmers and processors add value to agriculture produce through value-added activities such as processing, branding, packaging, and other post-harvest activities to enhance the marketability of farm product. In the context of this study, crop diversification is defined as a shift in production portfolio away from mono-cropping to adopting a multiple cropping system. In developing economies, this shift in production usually occurs as farmers move away from producing staple towards high-value food commodities such as fruits and vegetables.

Depending on the objective and research question, there are several methods that are widely used in the literature to measure the extent of crop diversification. The most common method for measuring the degree of diversification is the calculation of a vector of income/revenue shares related to different income sources. While this approach puts diversification and income changes directly into the relationship, a relevant part of information related to different aspects of diversification is neglected such as the number of crops grown. Other studies rely on a multidimensional perspective by employing a set of different statistical indices, which allow for a multidimensional analysis of diversification behaviour (Asfaw et al., 2016; Barrett and Reardon, 2000). Table 3.1 provides information on the calculation of these diversity indices, their interpretation, and usage.

The diversity methods that measure crop or species richness (such as count and Margalef Index see Table 3.1) are usually used in the ecological research to capture spatial biodiversity of crops and the richness of genetic resources. A limitation of measuring crop diversity at the parcel level in terms of the number of crops produced is that it masks across-parcel heterogeneity in the distribution of parcel land over the components of crop portfolio. Limiting crop diversity analysis to a subset of main crops may equivalently conceal production diversity that could represent an important contribution to household income and food security (Covarrubias, 2015). Count measures provide a general level of overall diversity on a farm, but do not account for whether the farm is growing high value cash crops or staple crops, and what percentage of resources are allocated to which crops (Turner, 2014). Bezabih and Sarr, (2012) used the count measure to study linkages between risk preferences and environmental uncertainty in Ethiopia. Bartolini et al., (2014) used a count variable to measure diversity of on-farm activities to study diversification towards multifunctional activities in Tuscany Italy.

**Table 3.1:** Measures of Crop Diversification

| Method | **Formula** | **Interpretation** | Concept |
|---|---|---|---|
| Crop count | $D_i = N$ | $D_i \geq 0$ | Richness[i] |
| Margalef Index (MI) | $D_i = \frac{(N-1)}{ln(A_i)}$ | $D_i \geq 0$ | Richness |
| Herfindahl-Hirschman Index (HHI) | $D_i = \sum_{n=1}^{N}(P_n)^2$ | $0 \leq D_i \leq 1$ | Relative abundance[ii] |
| Simpson/Transformed Herfindahl Index (THI) | $D_i = 1 - \sum_{n=1}^{N}(P_n)^2$ | $0 \leq D_i \leq 1$ | Relative abundance |
| Berger-Parker (BP) | $D_i = 1/max(P_n)$ | $D_i \geq 0$ | Inverse dominance[iii] or proportional abundance |
| Shannon-Weaver/the Entropy Index (EI) | $D_i = -\sum_{n=1}^{N} P_n \; log\;(P_n)$ | $D_i \geq 0$ | Evenness, proportional abundance |
| Modified Entropy Index (MEI) | $D_i = -\sum_{n=1}^{N}[P_n \; log_N P_n]$ | $0 \leq D_i \leq 1$ | Evenness[iv], proportional abundance |
| Composite Entropy Index (CEI) | $D_i = MEI * \left[1 - \frac{1}{N}\right]$ | $0 \leq D_i \leq 1$ | Evenness, proportional abundance |

Notes: $D_i$ is the value of the diversity index for $i^{th}$household, $N$=Number of crops grown by the $i^{th}$ household, $A_i$ =Total gross revenue of all crops for the $i^{th}$ household, $P_n$=revenue share of $n^{th}$ crop for the $i^{th}$ household. The concepts are defined as: i) Richness is a simple count of species or crops which does not take into account their abundance or relative distribution; ii) Relative abundance refers to how common or rare a species is relative to other species in a defined location or community; iii) Dominance is the degree to which a crop is more numerous than its competitors in an ecological community, or makes up more of the biomass; and iv) Evenness refers to how close the number of each species in an environment is; a measure of the extent to which household revenue is distributed evenly or disproportionally over the number of crops produced.

Given the objective of this study, the Composite Entropy Index (CEI) was selected as a primary measure for crop diversification. In addition to revenue shares of individual crops, CEI gives due weighting to the total number of crops grown by the farm household. This is important as the revenue share captures the relative importance of crops based on their economic value which may largely vary depending on the type of crops (i.e. the value of the index will be higher for households that grow larger number of high value crops). Thus, the CEI index is sensitive to the changes in the number of crops and their respective revenues. Moreover, CEI is easier to interpret as it provides a standard scale bounded by 0 and 1.

While the CEI index possesses all the desirable properties of Entropy and Modified Entropy Indices as explained in Table 3.1, it is adjusted by the number of crops. The detailed formula of CEI is given by:

$$D_i = -\left[\sum_{n=1}^{N} P_n \; log_N P_n\right] \; \left[1-\frac{1}{N}\right] = -\sum_{n=1}^{N} \frac{\ln P_n}{\ln N} \; \left[P_n - \frac{P_n}{N}\right] \qquad 3.1$$

Where $D_i$ represents composite entropy index, $P_n$ is the share of revenue from the nth crop (for $n = 1, 2, \ldots, N$) grown by the ith farmer, and $N$ is the number of total crops grown by the ith farm household in a given year. The computed value of the index increases with level of diversification which ranges from 0 implying no diversification (i.e. mono-cropping) to 1 implying the highest level of CD.

There are a number of studies that used CEI as a measure for crop diversification: Jadhav and Deshmukh, (2014), Mandal and Bezbaruah, (2013), and Acharya et al., (2011), for Marathwada region of Maharashtra, Assam Plains, and Karnataka state of India respectively. Mesfin et al., (2011), Weiss and Briglauer, (2000), De and Chattopadhyay, (2010), Malik and Singh, (2002), McNamara and Weiss, (2005), Mishra et al., (2004), Stoeffler, (2016), and Cutforth et al., (2001) used the entropy index as a measure for crop/farm diversity in Eastern Ethiopia, Austria, West Bengal and Haryana of India, Federal State of upper Austria, USA, Burkina Faso, and Saunders county in USA respectively. Other studies have used Herfindahl and Transformed Herfindahl (or the Simson Index) indices to measure crop or income diversification. These studies include

Ayieko, (2015), Babatunde and Qaim, (2009), Ibrahim et al., (2009), Rahman, (2009), and Barrett et al., (2005).

For the purpose of sensitivity of results to using different measures of crop diversification, alternative measures such as Transformed Herfindahl Index (THI) (measuring the relative abundance) was also used to test the model for robust estimates of the determinants of crop diversification. Like CEI, THI is also bounded by 0 to 1, with 0 representing the lowest level of diversification. If there is just one crop, then $P_n$ would be 1 and the computed THI will be 0. As the number of crops increases, the share of $P_n$ decreases and so does the sum of the squared share, so that THI approaches 1. Assume there are $N$ sources of revenue, then THI falls between 0 and $1 - 1/N$. Thus, the closer the computed THI is to 0, the higher the specialization, and the further it is from zero, implies the more the diversification.

### 3.3.2 Estimation Techniques and Factors Affecting Crop Diversity

Most studies on factors that influence the adoption and extent of farm diversification decisions of farm households in developing countries identify farm household socio-economic, demographic, regional, farm, farm characteristics as important (Bowman and Zilberman, 2013; Mishra et al., 2004; Ellis, 2000, 1998). The empirical relationship of these factors and diversity in crop production is analysed using various econometric techniques (Table 3.2).

The type of econometric technique depends on the type of dependent variable (i.e. different measures for crop diversification presented in Table 3.1). In Table 3.2 we summarize previous studies, the estimation techniques, and methods used to measure crop diversification. Because most of the indices used to measure crop diversification can be censored at one or both sides, Tobit analysis is one of the most common methods used in crop diversification analysis. While majority of these studies use Tobit analysis, some other studies argue that the decision to diversify crop portfolio involves selectivity bias and therefore use Heckman two-stage model. These studies include Kimhi and Chiwele, (2000), Rehima et al., (2013), Kanyua et al.,( 2013), Kumara et al., (2016) Seng, (2014), Omiti et al., (2009).

Most of the studies on developing countries listed above find a significant relationship between crop diversity and standard determinants such as farm characteristics (i.e. farm size, land quality, landscape, and access to irrigation), farmer characteristics (such as age, sex, and education of the farm operator), access to infrastructure and services (such as access to roads, market, transport equipment, and extension services), and geographical characteristics capturing differences in cultural and physical conditions.

**Table 3.2:** Previous Studies on Crop Diversification

| Study | Country | Sample | Measure of CD | Estimator |
|---|---|---|---|---|
| Mandal and Bezbaruah, (2013) | Assam Plains | 342 | CEI | Two-limit Tobit Analysis |
| Mesfin et al., (2011) | Ethiopia | 167 | Modified Entropy Index | Two-limit Tobit Analysis |
| Abdalla et al., (2013) | Sudan | 200 | Shannon Entropy Index | Tobit Analysis |
| Kumar et al., (2012) | Eastern India | 2,885 | THI | Heteroskedastic Tobit Analysis |
| Cavatassi et al., (2012) | Hararghe Ethiopia | 699 | Count, Shannon Index, & Berger-Parker | Poisson and Instrumental Variable Tobit |
| Stoeffler, (2016) | Burkina Faso | 229 | Count, Entropy, & Berry indices | Probit and MNL |
| Aneani et al., (2011) | Ghana | 300 | Count | Multinomial Logit Analysis |
| Hitayezu et al., (2016) | Kwazulu-Natal | 152 | HHI | logit Transformation |
| Bartolini et al., (2014) | Tuscany | 72,686 | Number of activity | Zero-inflated Negative Binomial |
| Van Dusen and Taylor, (2005) | SNP Mexico | 281 | Count of crops | Poisson |
| Acharya et al., (2011) | Karnataka India | - | CEI | OLS |
| Benin et al., (2004) | Ethiopian Highlands | 739 | Margalef Index | Two-stage, Probit &CLAD |

The descriptive analysis of the data for Afghanistan shows that 33% of the farmers do not diversify, whereas majority of the farmers who actually diversify grow only two or three crops. This indicates that modelling crop diversification may not require two-step modelling techniques (e.g. Heckman or double hurdle) to assess household's decisions on whether to diversify and the extent to which they diversify in two separate steps. Two-step estimators are specifically designed to allow each step to be affected by different factors. Given that our dependent variable (CEI) is censored and the fact that diversification appears to be the outcome of a single decision process, Tobit analysis is likely to better fit the data.

### 3.3.3 Off-farm Income and Magnitude of Crop Diversification

Heterogeneity in the motivation and constraints faced by rural households plays a key role in households' diversification behaviour. As per the development economic literature, these motives can be driven by "pull" and "push" factors. With a fall in agriculture income and when farm income alone cannot provide sufficient livelihood, farm household may be "pushed" to diversify into non-agriculture activities to stabilize their incomes given the variability of farm income (Minot et al., 2006; Mishra and Goodwin, 1997). McNamara and Weiss, (2005) stated that if farm income falls below the household reservation wage, household members will allocate time to off-farm labour. Meanwhile, households are maybe pushed by higher returns to labour and or capital particularly in the less risky nature of investment in the off-farm sector (Ellis, 1998; Kilic et al., 2009).

On the other hand, household may be "pulled" into farming business and on-farm diversification when prevailing market conditions for agriculture commodities present opportunities that offer them a comparative advantage (Ayieko, 2015). Pull factors generate opportunities for diversification of income sources related to commercial agriculture, improved infrastructure, and better market access.

Given the literature on the constraints and motivation of on- and off-farm diversification, limited attention has been devoted to assessing whether a causal relationship between off-farm income and crop diversification exist. Weiss and Briglauer, (2000) indicated that the existence of additional off-farm income reduces the degree of diversification because part-time farms (i.e. farmers who engage in both farm production and non-farm activities) have less labour time to allocate to the production of a broad agricultural product mix. More importantly, off-farm income is considered by farm households as a strategy to diversify employment risks and thus reduces the necessity to diversify on the farm.

Similarly, Mishra et al., (2004) reported an inverse relationship between off-farm income and the level of crop diversification for US farm households. They argued that time allocation of farmer and family labour between farm and off-farm alternatives influences on-farm enterprise diversification. If the household members are working full time on the farm, this may be an indicator that the comparative advantage for their labour is on the farm. Hence, they would be more likely to diversify on-farm enterprises to increase profit. Mishra and Goodwin, (1997) pointed out that If farmers are risk averse, greater farm income variability should increase off-farm labour supply to sustain incomes. Hitayezu et al., (2016) also reported a negative relationship between off-farm income and intensity of crop diversification in the Midlands region of Kwazulu-Natal of South Africa. They argued that access to off-farm work increases the opportunity cost of on-farm diversification efforts.

On the contrary, Cavatassi et al., (2012) found that participation in non-farm activities is positively associated with the number of crops grown by households. They argue that household motivation in off-farm activities are driven by the liquidity constraints which enhance diversity by allowing households to purchase inputs. Similarly, Girish and Mehta (2003) investigated the empirical relationship between the magnitude of diversification and socio-economic factors and showed that non-farm income significantly increased level of crop diversification in Himachal Pradesh of India. Based on their explanations, because non-farm income significantly contributes to the overall income and well-being of households, it increases crop diversity through this income effect.

### 3.4 Methodology and Theoretical Framework

This section presents a theoretical framework to inform the empirical analysis. The most fundamental theoretical question that researchers in the field of agricultural economics continue to ask is, which farmers diversify and why? Motives for crop diversification by the farm households may vary depending on the objectives pursued by them. Farmers may adopt a more diversified cropping system to stabilize their income or minimize production risks caused by adverse farming conditions and shocks. As for the evidence in the broader literature on rural livelihood diversification, a number of studies have pinpointed the socio-economic rationale of farm households for pursuing a diversified crop portfolio.

In assessing diversification strategies by the rural households, Reardon et al., (2001) and Barrett et al. (2005) argued that heterogeneous constraints and incentives play a fundamental role in determining livelihood diversification patterns. Wealthy and poor farmers behave differently considering the diversification decision depending on the endowment of initial assets. Rich farmers with engagement in capital intensive activities may see diversification as a method for increasing return on agricultural capital and therefore they aim to maximise their profits, whereas poor farmers with engagement in higher labour-intensive activities may have a different incentive, that is to mitigate production and market risks. In Burkina Faso, rich farm households have more diversified crop portfolio and mostly engage in producing high-yield and high-value crops, whereas poor farmers mostly produce basic food grains (Stoeffler, 2016).

Van Dusen and Taylor (2005) argue that diversification is driven by the output and input factor market conditions and decreasing return to scale. For a farm household, missing or incomplete markets (usually as a result of high transaction costs) implies optimal allocation of the scarce production resources between multiple crops. This rationale can be explained by Figures 3.1a and 3.1b below. Figure 3.1a illustrates imperfect market for crop $j$ with a Production Possibility Frontier (PPF) that characterises the technologically efficient production mixes available to a household that aims to allocate scarce resources between crops $j$ and $h$. Under perfect market conditions where there is a market for both commodities and risk is absent, farmer's decision is guided by the (exogenously given) market price (shown by the $M^*$ line with a slope of$-P_j/P_h$), and optimality with perfect markets implies a corner solution at ($Q_h^*$, 0), with production of one crop ($h$). However, when there is missing market for crop $j$ and risk is present (and in absence of insurance market), the household decision on allocating resources among two crops is determined by a subjectively valued shadow price ($P_j'$) which is shaped by the household's marginal utility and availability of production resources. This defines a new downward price line ($M'$) which leads to a new optimal crop diversification solution ($Qj'$, $Qh'$) as the household shifts from using the exogenous prices ($P_j$, $P_h$) and producing only crop $j$ to producing at the constrained level $Qh'$ corresponding to the point of tangency between the price line $M'$ and the PPF shown by $Q'$ (Hitayezu et al., 2016; Van Dusen and Taylor, 2005). In the graph in Figure 3.1a, the household is assumed to produce two crops, but the results can easily be extended to producing multiple crops.

Production risk may also drive the household decision towards diversification. For instance, previously it was assumed that the absence of risk and insurance market would lead the households to produce only crop $h$ under perfect market condition in Figure 3.1a. However, assuming crop $h$ is characterised by high-yield risk, then household decision on producing crop $h$ will not be determined by the exogenous market price (the $M^*$), but rather based on the household subjective level of risk-preferences. In this case the household's tendency to produce only one crop ($h$) may decline, and the relevant price line would resemble $M'$ in Figure 3.1a which implies the household would be induced to produce crop $j$ to spread perceived level of risk (Van Dusen and Taylor, 2005).

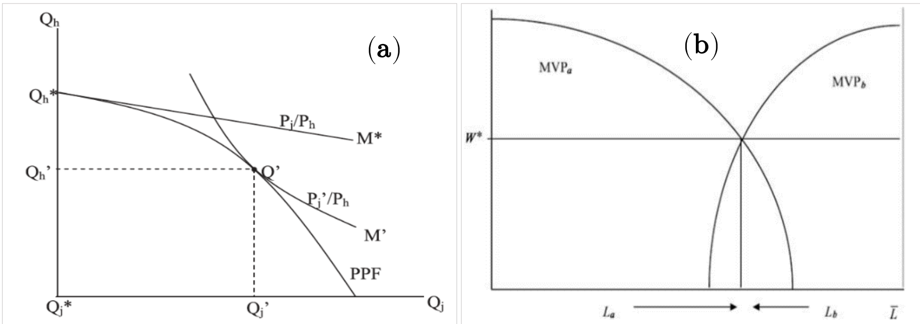



**Figure 3.1:** (**a**) PPF for two Crops (*j* & *h*) Under Perfect (*) and Missing (') Markets. (**b**) MVP of Crops a & b vs. a Fixed Factor of Production

*Source:* Adapted from (Hitayezu et al., 2016; Van Dusen and Taylor, 2005)

A missing market brings the production of a good directly into the household's utility function (via the subsistence constraint); therefore, factors affecting the utility function also affect crop allocations (Van Dusen and Taylor, 2005). Missing or incomplete markets for inputs also play a role in household's diversification decision. Incomplete markets for assets such as land, labour, credit or insurance are major causes of diversification behaviour (Barrett et al., 2001; Van Dusen and Taylor, 2005). For instance, if hired labour is unavailable and cannot be substituted by family labour, the household's decision is driven by the fixed family labour hours available to them (Van Dusen and Taylor, 2005).

McNamara and Weiss (2005) argue that with respect to on-farm diversification, economies of scale and scope of the agricultural enterprise mix are important; if the cost function exhibits economies of scope the households would produce goods jointly instead of separately. McNamara and Weiss (2005) and Van Dusen and Taylor (2005) suggested that the household's rationale for diversifying crop activities may be due to Decreasing

Returns to Scale (DRS) by a given set of production technology. This can be illustrated by Figure 3.1b which portrays a production factor (family labour $L$, assuming it is fixed at $\overline{L}$) being allocated between two crop activities ($L_a$ and $L_b$). A decreasing marginal value product (MVP), such as fixed endowment of land, soil quality, distance from market, or other input that results in a decreasing MVP of labour, shows DRS with respect to the family labour (depicted on the horizontal axis). For instance, if hired labour is not available or cannot be substituted for family labour , the household is left with the endowment of family labour available for crop production. In this case, if the household could only allocate labour to producing two crops, it would do so until the marginal value product of labour is equated between two activities at an endogenous 'shadow' family wage, $w^*$.

### 3.4.1 The Agriculture Household Model

Farm household decisions on crop choices and extent of diversification can also be understood in the context of the farm household model initially developed by Singh et al. (1986) which assumes farm households are both consumers and producers of agricultural goods subject to constraints. A number of studies in the recent literature focusing on on-farm diversity adopted this approach to explore the decision of farm households with regard to the intensity of farm or crop diversification (Hitayezu et al., 2016; Cavatassi et al., 2012; Benin et al., 2004; Van Dusen and Taylor, 2005; Van Dusen, 2000).

In general, there are two motives and objectives that households pursue to practice crop/farm diversification that can be conceived by the potential gains in the expected utility and minimization of the coefficient of variation or risk (McNamara and Weiss, 2005). However, an empirical comparison of these frameworks (Herath, 1980) indicates that the expected utility framework is more representative for the actual behaviour. The analytical model used for this study draws upon the household model applied to study on-farm crop diversification. In case of on-farm and crop diversification, this approach is also adopted by Van Dusen (2000), Van Dusen and Taylor (2005), Cavatassi et al., (2012), and Hitayezu et al., (2016). For the purpose of this study, the household model is based on the original model of Van Dusen and Taylor (2005)

Proceeding to the household model, consider an agricultural household that maximises utility over a set of consumption goods produced on the farm ($C_f$), a set of purchased non-farm commodities ($X_{nf}$), and leisure ($l$). The expected utility gained from various combinations and levels of consumption goods directly depends on the vector of preferences of the household, denoted by $\Phi^{hh}$, shaped by household socio-economic, cultural, and other exogenous factors. This maximization problem can be written as:

$$\underset{C_f, C_{nf}, L, X, A}{max} \qquad U(C_f, C_{nf}, l|\ \Phi^{hh}) \qquad 3.2$$

Subject to the following constraints facing the household:

$$p_f(Q_f - C_f) - C(Q_f|\Phi^f) + Y_{nf} \ = p_{nf}C_{nf} + w(L_f + L_{nf}) \qquad 3.3$$

$$Q_f = f(\alpha, L, X_f|A, \Phi^f) \qquad 3.4$$

$$T = (L_f + L_{nf}) + l \qquad 3.5$$

$$Y_{nf} = y(L_{nf}|\Phi^{nf}) \qquad 3.6$$

The utility is constrained by the general budget constraint (Equation 3.3) such that the maximum expenditures of time $w(L_f + L_{nf})$ and money $p_{nf}C_{nf}$ cannot exceed the total income of a farm household in a given decision-making period (in the case of this study a season or year). Total household income is composed of farm income $p_f(Q_f - C_f)$ net of production costs $C(Q_f|\Phi^f)$, and off-farm income denoted by $Y_{nf}$ that includes remittances, stocks carried over, and other transfers which are exogenous to the season's crop choices.

The amount of agriculture produce consumed by the household $(C^f)$ or sold $(Q_f - C_f)$ are chosen from the crop(s) output $Q_f$ (for crop $j = 1, 2, 3, \dots J$ that household choose) which is constrained by the given production technology embedded here in the cost function $C(Q_f|\Phi^f)$ where $\Phi^f$ is a vector collecting exogenous farm characteristics. Household decisions about the number of crops and the quantity is constrained by the fixed technology constraint (Equation 3.4) such that the quantity of goods produced on the farm $Q_f$ is a function of purchased inputs ($X_f$), Labour ($L^f$), a given area of land ($A$) which is allocated to different crops (here denoted by $\alpha$ or the set of share of land allocated between $J$ crops such that $\sum_{j=1}^{J} \alpha = 1$), and exogenous characteristics of the farm $\Phi^f$. According to Benin et al., (2004), each set of area shares implies a level or combination of crop outputs, then the objective function in Equation.3.1 can be re-expressed as:

$$\max_{h} \quad V(C_f, C_{nf}, l \mid \Phi^{hh}) \tag{3.7}$$

Where $h = \ ((\alpha_{1,}\alpha_{2,}, \alpha_{3,} \ldots \alpha_n) \geq 0, C_f, C_{nf}, X, and\ L))$. The allocation of labour is constrained by the household total labour time (Equation 3.4) which is denoted by $(T)$ available for off-and on-farm activities (denoted by $L^{\mathrm{f}}$ and $L^{nf}$) and leisure $(l)$.

Assuming that households maximize utility, and markets for farm goods function perfectly, then production decisions by farm households can be made separately from the consumption decisions. Thus, the level of crop diversification is driven by net returns which are determined by market wage, input and output prices $(w$, $p_x$, and $p_f)$, and farm physical characteristics $(\Phi^f)$. However production and consumption decisions cannot be separated under imperfect market conditions, then the household optimal choice $h^* = (\alpha^*, L^*\ C_f^*, C_{nf}^*, X^*)$ can be expressed as a reduced form function of land holding size, exogenous income, and household, farm, and market characteristics (Benin et al., 2004) and it follows that:

$$h^* = h^*[\alpha^*(A, Y_{nf}, \Phi^{hh}, \Phi^f, \Phi^m\ )] \tag{3.8}$$

Assuming that households do not explicitly value crop diversification (i.e. it is not reflected explicitly in the utility function itself) and that it is the outcome of choices made in a constrained optimization problem rather than an explicit choice (Benin et al., 2004; Van Dusen and Taylor, 2005), then crop diversification (D), can be expressed as a derived demand function given by:

$$D = D[\alpha^*(A, L, Y_{nf}, \Phi^{hh}, \Phi^f, \Phi^{nf}, \Phi^m)] \tag{3.9}$$

Where $D$ represents CD measured by the composite entropy index of crop diversity at the household level. Equation (3.9) indicate that crop diversification is a function of the initial endowments of labour $(L)$, land $(A)$, exogenous non-farm income $(Y_{nf})$, farm household characteristics $(\Phi^{hh})$, farm characteristics $(\Phi^f)$, and market conditions $(\Phi^m)$. The unit of analysis is the farm household that decides the level of diversification given a number of objectives and constraints.

## 3.5 Estimation Strategy

### 3.5.1 Identification

The expected causal relationship between off-farm income and extent of crop diversity can be either positive or negative. In the context of subsistence small scale farming system, farming often fails to provide sufficient livelihood for the households. While farming may still remain their primary source of income, households often seek alternative means of income by participating in off-farm activities. This results in the reallocation of production resources among on- and off-farm activities. Based on this argument, the off-farm income may lead to a lower level of crop diversification due to negative labour effects. This is consistent with the narrative that allocation of farm labour away to off-farm activities decrease diversity due to negative labour effects, particularly when the opportunity cost of household labour is higher than the off-farm wages under imperfect markets implying non-separability between households' farm profits and off-farm earnings as argued by Chavas et al., (2005).

On the contrary, off-farm income may have a positive impact on the level of crop diversity due to income effects. Because increased off-farm income will increase household's capability to purchase sufficient production inputs necessary for different crops and may ease cash constraints. Thus, it will motivate the intensity of crop diversification.

The sample data suggest that as CD increases, the share of off-farm in total income falls whereas the share of farm income increases to about 50% for three or more crops (Figure 3.2) and see mean differences of household characteristics by access to off-farm income in Table 3.A1 in Appendix 3.A. However, the descriptive information cannot imply any negative or positive causal correlation between off-farm income and the number of crops grown by a household. The causal impact of off-farm income on the intensity of crop diversity is hypothesized to be mixed (i.e. negative or positive). Thus, these ambiguous implications of off-farm income signify the importance of assessing the empirical relationship between off-farm income and level of crop diversity.

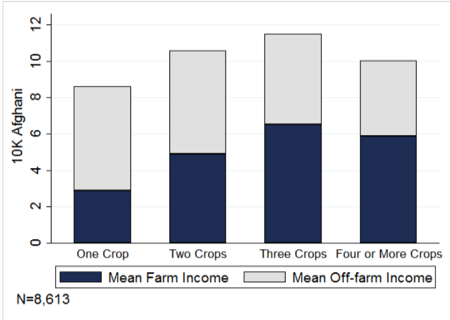



**Figure 3.2:** On- and Off-farm Income Against Number of Crops
*Source:* Author's calculations of the ALCS 2013-14 data

Meanwhile, there might be a third category of unobserved factors affecting both on-farm diversity (i.e. crop diversification) as well as diversification towards off-farm activities. Subsistence farmers are typically assumed to be risk-averse, and such behaviour may encourage farmers to diversify into both crop diversification and off-farm activities. Given that earning additional off-farm income might also be used as a diversification strategy by households to spread risk outside the farming sector, one would expect the parameter estimate of off-farm income to be biased upwards if endogeneity is not allowed for. Another example of these unobserved factors could be the entrepreneurial ability and relative efficiency that can influence both farmers' decisions about the extent of crop diversification and diversification towards off-farm income. Relative efficiency generates a downward bias in the coefficient off off-farm income if endogeneity not allowed for. Another source of endogeneity may be the presence of measurement error attributed to the recall of the extent of non-farm income earned by the household (Zereyesus et al., 2017). In the presence of measurement errors, one would expect the coefficient of off-farm income to be biased towards zero. Hence, we allow the off-farm income variable to be endogenous and use instrumental variables to identify its true effect on the intensity of crop diversity.

The cross-section household level data used in this study do not control for unobserved household fixed effects, so instrumental variable (IV) techniques are employed to control for the potential endogeneity bias in off-farm income. The estimation of the endogenous model requires the use of an IV to be included in the reduced-form equation but excluded from the structural model of crop diversification. It is therefore required that the IVs should be significantly correlated with the endogenous variable (off-farm income) but should not directly affect the level of crop diversification.

Two instruments are used to control for the endogeneity bias in off-farm income. Firstly, the share of aggregate off-farm income in the total income for all households in a given district. According to Diiro and Sam (2015), this instrument captures the status of local non-farm labour market; higher share of non-farm income signifies high prevalence of non-farm employment opportunity at district level which in turn translates into greater potential for households to diversify into off-farm activities.

Kilic et al. (2009) use share of non-farm employment within a district as an instrument for off-farm income, noting that, because the instrument is constructed at the district level as opposed to the village level, when regional fixed effects are controlled for it is unlikely for the instrument to have a direct effect on the farming decisions of households. Diiro and Sam (2015) argue that, controlling directly for family labour, the only pathway for the instrument to influence household decisions is through the household non-farm income activities. Smale et al. (2016) studied the relationship between off-farm work and farm output and used share of total non-farm earnings (business and salary) in total income by location as an instrument for off-farm income. In assessing the impact of off-farm income on farmer's liquidity constraints, Gebregziabher et al. (2012) used unemployment rate at the district level to control for potential endogeneity in off-farm income. Similarly, in examining the relationship between participation in non-agricultural labour activities and farming production decisions, Stampini and Davis (2009) used a dummy variable for the existence of off-farm employment opportunities in the commune.

Controlling for the household's family labour and regional fixed effects by including household size and agro-ecological dummies in the analysis, we expect the existence of non-farm employment opportunity will affect household decisions to diversity only through the off-farm income channel. It is important to note that data in the sample comes from 349 districts and 34 provinces and on average there are about 50 farm households being surveyed in each district.

Secondly, we use district level lagged values of off-farm income from year 2011/12 to instrument for off-farm income. Lagged off-farm income comes from the National Risk & Vulnerability Assessment (NRVA) survey conducted by CSO in 2011/12. Off-farm income from the past is expected to positively affect farmer's current non-farm activities. Diiro and Sam (2015) uses off-farm income from previous years as an instrument to control for endogeneity in off-farm income. According to them, income from previous years represents an important form of financial endowment that assists farm households to invest in productive farm assets. One might argue the generation of income is a dynamic process and that transitory values of past income will influence current farming decisions. However, we use district level aggregate lagged income (not household level) as instrument to capture the overall non-farm employment status. There is also evidence that farmers, particularly small holders do not leave cash money on the table to transfer them from one season to another (Duflo et al., 2008).

### 3.5.2 Econometric Specification

Because not every farm household diversify or choose to diversify, a censoring issue underlies the empirical model. Although theoretically the dependent variable (CEI) is censored on both sides because it is bounded by 0 and 1, practically there are no computed values for CEI that are 1. Since the dependent variable is censored at 0 for 33% of the sample (i.e. non-diversifiers), the Tobit model was employed to deal with the censorship at zero of the dependent variable (CEI). Conventional regression methods (i.e. OLS) fail to account for the qualitative difference between zero observations and continuous observations. Zero values of the CEI/THI indices may occur for various reasons. Even though farmer's may be potential diversifiers, they may not be able to diversify due to constraints such as soil type, climate or farm size. Households may choose to remain non-diversifiers if production of certain crops offer a comparative advantage in market or production of a particular staple food crop required for food security. In these cases, zero observations represent a corner solution which is an optimal choice by the farmers not to diversify. Therefore, the zero observations are important to be accounted for. Tobit model originally developed by Tobin, (1958) with left-censored values of the dependent variable (CEI) is specified as:

$$y_i^* = x_i' \, \beta + \varepsilon_i \qquad 3.10$$

Where $\beta$ is a row vector that collects unknown parameters to be estimated, $x_i$ is a column vector of the explanatory variables that effect the extent of crop diversification,

$e_i$~$N(0, \sigma 2)$, and $y_i^*$ is a latent variable that is not directly observed, but takes on the following values:

$$y_i = \begin{cases} 0 & y_i^* < 0 \\ y_1^* & 0 \le y_i^* \le 1 \end{cases} \qquad 3.11$$

Here $y_i$ represent the observed values of the dependent variable (CEI). However, as household Off-Farm Income (OFY) is likely to be endogenous as explained above, the following endogenous Tobit model is estimated using instrumental variables:

$$CD \qquad y_{1i}^* = \beta y_{2i} + \gamma x_i' + u_i \qquad 3.12$$

$$\text{OFY} \qquad y_{2i} = x_{1i}\Pi_1 + x_{2i}\Pi_2 + v_i \qquad 3.13$$

Where $\beta$ and γ are row vectors of unknown parameters to be estimated (structural parameters), $x_{1i}$ is a 1 x k1 column vector of exogenous variables that affect the level of crop diversification, $y_{1i}^*$ is a latent variable that is not directly observed, but takes the values shown in (3.10) depending on $y_i$ (the observed values of the CEI). The equation for OFY denoted by $y_{2i}$ is written in the reduced form where $x_{2i}$ is a 1 x k2 vector of additional instrumental variables, and $\Pi_1$ and $\Pi_2$ are matrices of reduced-form parameters, $u_i$ and $v_i$ are error terms that are assumed to be jointly identically and independently distributed $(u_i, v_i) \sim N(0, \sigma^2)$. The endogeneity bias due to unobserved factors as explained earlier will generate an association between the error term $(u)$ in Equation (3.12) and $(v)$ in Equation (3.13) (the reduced form) that will mask the true effect of off-farm income on crop diversity.

In this representation $y_i^*$ captures the unobserved difference between the latent utility gained from crop diversification and the utility gained from choosing a single crop (non-diversified system). The latent utility is assumed to be determined by a linear function of observed household demographic, socio-economic, regional and the farm characteristics plus an observable error term $(u_i)$. The endogenous Tobit model can be estimated using the two-step estimator proposed by Newey (1987) or Maximum Likelihood Estimator (MLE) techniques.

The estimated tobit coefficients are the marginal effects of a change in $x_i$ with respect to $y^*$, the unobservable latent variable, and show the effect of a change in a given independent variable $(x)$ on the expected value of the latent variable, holding all other independent variables constant (Greene, 2012). However, such an interpretation may not have a quantitative meaning or may not be of interest since $y^*$ is unobserved (e.g. we only observe $y^*$ if it is above a threshold, in our case zero). We are interested in the effect of $x$ on the observable $y$ (or change in the censored outcome). Depending on the purpose of the study, there are three values of interest after fitting a tobit model: 1) Marginal effects of $x$ on the index or latent variable $E[y_i^*|x_i']$; 2) the expected value of y, conditional on y being positive, $E[y_i|y_i > 0, x_i]$ ; and 3) the unconditional expected value of y, $E[y_i|y_i > 0, x_i]$ (Greene, 2012).

The Expected value of the latent variable ($y^*$) is simply the estimated coefficient of the tobit model:

$$\frac{\partial E[y_i^*|x_i]}{\partial x_i} = x_i'\ \beta \tag{3.14}$$

Expected value of the truncated subpopulation or those who actually diversify (i.e. where y or CEI is greater than zero) is given by:

$$\frac{\partial E[y_i|y_i > 0, x_i]}{\partial x_i} = \beta_k\ \left\{1 - \lambda\left[\frac{x_i'\ \beta}{\sigma} + \lambda(\alpha)\right]\right\} \tag{3.15}$$

Where $\lambda(\alpha) = \frac{\phi\ \left(\frac{x_i\ \beta}{\sigma}\right)}{\Phi\left(\frac{x_i\ \beta}{\sigma}\right)}$ that is also referred to as the inverse mills ratio, $\phi(.)$ is Normal Probability Density Functions (PDF), and $\Phi(.)$ is the normal Cumulative Distribution Function (CDF). Unconditional expected value for observations (that may be censored or uncensored) $y$ on $x$ is given by[8]:

$$\frac{\partial E[y_i|x_i]}{\partial x_i} = \ \Phi\left(\frac{x_i'\ \beta}{\sigma}\right)\beta_k \tag{3.16}$$

Because we intend to estimate the determinants of the extent of crop diversification for both single croppers (i.e. with the CEI value of zero) and diversifiers (i.e. with the CEI

[8] As per McDonald and Moffitt (1980), Equation (3.15) can be further decomposed and rewritten as:

$$\frac{\partial E[y_i|x_i]}{\partial x_i} = \beta\{\Phi_i[1 - \lambda_i(\alpha_i + \lambda_i)] + \phi_i(\alpha_i + \lambda_i)\}$$

Where $\alpha_i = x_i'\ \beta\ /\sigma$, $\Phi_i = \Phi_i(\alpha_i)$, and $\lambda_i = \phi_i/\Phi_i$. Taking the two parts separately, this result decomposes the slope vector into:

$$\frac{\partial E[y_i|x_i]}{\partial x_i} = Prob[y_i > 0]\frac{\partial E[y_i|y_i > 0, x_i]}{\partial x_i} + E[y_i|y_i > 0, x_i]\ \frac{\partial Prob[y_i > 0,]}{\partial x_i}$$

Thus, a change in $x_i$ has two effects: It affects the conditional mean of $y_i^*$ in the positive part of the distribution, and it affects the probability that the observation will fall in that part of the distribution.

value greater than zero), our interest is therefore in the estimation of the unconditional expected value given by Equation (3.16) (i.e. partial effects of explanatory variable with respect to the observed $y$ being censored or uncensored). For discrete choice variables, $E(y)$ is evaluated at alternative discrete choice values of $x_k$. Marginal effects are for the entire sample. The effects on the uncensored observations will be greater.

In addition, we estimated Equations 3.12 and 3.13 using a probit model treating crop diversification as a binary variable (i.e. 0 for non-diversifiers and 1 diversifiers) instead of CEI to examine the likelihood of diversifying by the households. Since the survey only reports the total number of crops grown by the household, it is unclear whether household grow different crops on the same unit of land in a certain season/year (e.g. crop intensification or inter-cropping) or they actually grow different crops on different units of land (i.e. crop diversification). Thus, CEI index might be a better measure that aims to capture the relative proportion of revenue associated with individual crops grown by the household and also account the number of crops, whereas the binary measure of CD only captures the fact whether households diversify of not (it does not account for the extent of diversification for those who actually diversify). Nevertheless, the anecdotal evidence[9] suggests that inter-cropping is not common in Afghanistan, particularly since grain crops (such as wheat, maize, barely, and rice) occupy the absolute majority of the land which are highly unlikely to be inter-cropped with other crops. Hence, we also employ a probit model to estimate the likelihood of growing single crop as opposed to multiple crops.

### 3.6 Standard Determinants of Crop Diversity

The intensity of crop diversification may be driven or constrained by a number of different factors. These can be grouped into household demographic and socioeconomic characteristics, access to rural infrastructure and services, and regional differences. The direction and degree of influence of these factors depend on household choices, access to and allocation of production resources, and motives for crop diversification. This section briefly summarizes these factors and discusses their potential expected causal relationship

[9] Inter-cropping may rarely happen if households own fruit orchards. Some fodder crops, especially clover and alfalfa are maybe intercropped with gardens.

with the dependent variable, the crop diversity measured by the CEI index, following the literature review above.

As per Equation (3.9), in theoretical model under subsection 3.5.1, the vector $\Phi^{hh}$ comprises a set of household characteristics. In the literature on crop diversification, household head age, gender, level of education, and household size are broadly evidenced to have influence on the intensity of crop diversification. Age of household may proxy for farmers experience and capabilities to do physical labour work. Older farmers are likely to have gained more experienced with farm management techniques and production. According to Mesfin et al., (2011) older farmers may be less risk-averse and therefore age has a negative influence on the level of crop diversification. Ibrahim et al., (2009) suggested that farmers try new crops as they age and gain more experience overtime. Ownership and access to farm assets and farm land can vary by the gender of the household head. In some cases, participation of females in crop diversification may be restricted by access to particular resources, therefore it is ex-ante hypothesized that male farmers have better access to resources to diversify.

Household head education is included to test whether more educated farmers have a higher propensity to diversity because of their technical skills and knowledge. The level of education of the head may have an ambiguous influence on a household's decision to diversify. More education is likely associated with employment outside farming, with a negative influence on crop diversity by withdrawing labour from farming. Alternatively, higher education would be associated with better management skills and productivity, allowing them to engage in the production of a variety of crops.

Household size, the number of adults living in the household, represents the pool of family labour available for farming activities (Van Dusen and Taylor, 2005) and affects farm labour supply. Larger households may be more flexible in allocation of labour time to various activities. Mesfin et al., (2011) stated that larger household size allows the household flexibility to pool resources and share risk by taking advantage of household returns to scale and labour supply when needed in peak seasons. Against this backdrop, one would expect a positive relationship between household size and the level of CD.

In addition to household demographic factors above, household socioeconomic variables such as household income, ownership of land and farm assets, and livestock wealth are important determinants of crop diversification. Livestock ownership by the farm households, as a proxy of wealth, may have ambiguous effects on the intensity of crop diversification (Benin et al., 2004; Van Dusen, 2000). However ownership of oxen is likely to increase the level of crop diversity by ensuring the availability of power for ploughing when needed (Benin et al., 2004). In addition, cattle ownership as a proxy for availability of animal manure, is an important source of organic fertilizer that may positively affect crop intensity. Other factors such as input and output prices are also expected to affect farmer's decision on the intensity of crop diversification (Singh et al., 1986; Van Dusen and Taylor, 2005). However, Van Dusen and Taylor (2005) argued that there are insufficient price variations in cross-section data, therefore prices are unlikely to affect crop diversity decisions in the short-run .

As per Equation (3.9) in theoretical framework, farm characteristics ($\Phi^f$) including land holding size, landscape characteristics, quality of soil and land, and availability and access to sufficient irrigation water are likely to affect the decision and magnitude of crop diversity. Pope and Prescott (1980) argue that the relationship between diversification and farm size is an indicator of trade-offs between risk reduction and economics of size, that is, if there are large-scale economies in an enterprise, then one might expect larger farmers to be more specialized. On the other hand, farmers (particularly farmers with small land holdings) may attempt to diversify to reduce production risks. Ayieko (2015) stated that land under cultivation by a farm household can result either in diversification or specialization, depending on the phase of the agricultural transformation process.

The variable land measured in hectares is the total land cultivated by the household in various seasons throughout the year. This includes both irrigated and rain-fed land owned or leased by the household that was actually cultivated throughout the year. Farms are generally small in Afghanistan. While average farm size is 1.6 hectares (equivalent to 7.9 Jeribs), majority (62%) of the farmers cultivated 1 or less than 1 hectare of land (Table 3.3), demonstrating that availability of farm land is an important and limiting factor for production that affects land allocation decisions. Distribution of diversity indices when farm households in the sample are grouped by the size of landholding are displayed in Table 3.3.

**Table 3.3:** Distribution of CEI and THI by Farm Size (ha)

| Farm Size (ha) | CEI | THI | Number of farms | % of Farms |
|---|---|---|---|---|
| Up to 1 | 0.27 | 0.26 | 5,327 | 62% |
| 1 to 2 | 0.32 | 0.30 | 1,931 | 22% |
| 2 to 3 | 0.35 | 0.34 | 462 | 5% |
| More than 3 | 0.34 | 0.33 | 893 | 10% |
| Overall | 0.29 | 0.28 | | |
| N | | | 8,613 | |

*Source*: Author's calculations of the ALCS 2013-14 data

From Table 3.3, farm size and crop diversification follow an inverse u-shaped relationship. Both indices CEI and THI of diversity initially increase with the farm size but starting to fall when farm size is beyond 3 ha. Based on distribution of CEI across farm size in Table 3.3, land allocation among crops is hypothesized to have positive effect on crop diversity.

The topographic features of farm land such as slope and landscape characteristics of the farm land are also controlled for in the regression analysis. According to Van Dusen and Taylor (2005), the altitude and slope (steepness or flatness) of the farm land proxies for agro-climatic niches within farms. In assessing crop diversity, Cavatassi et al., (2012) included number of plots with different slopes in their analysis to control for variability of production conditions. In this study, we included a landscape dummy variable that equals to zero if the terrain is Valleys & Hills and 1 if it is open plain.

Soil and land quality are conjectured to affect production decisions and crop diversity. Initial analysis of the data reveals that farmers own and operate two types of land, irrigated and unirrigated (which is mainly rain-fed). If land quality is heterogeneous and yields depend on land quality, the likelihood of diverse crops is low, as yields in rain-fed agriculture are substantially lower. To control for variations in land quality, a dummy variable equal to 0 if farmers have and operate a combination of both irrigated and rain-fed land and equals to 1 if farmers cultivate irrigated land alone, is included. It is hypothesized that farmers with rain-fed land are less likely to diversify.

Access to infrastructure and services are other important determinants of crop diversity. Distance to local market and nearest all-season roads, as proxies for transaction costs and market development, are important determinates of crop diversification (Benin et al., 2004). Turner (2014) indicated that farms lacking access to transport infrastructure do not allocate land to marketable or cash crops. It is hypothesized that the further a farm is located from the market and drivable roads, the longer the travel time to market and the higher are the transportation costs, the lower the level of crop diversification. Transaction costs are typically grouped into variable (e.g. transportation costs) and fixed (e.g. access to market information) transaction costs (Key et al., 2000; Seng, 2014). Following Heltberg and Tarp (2002), Benin et al. (2004), and Seng (2014), we use households access to television, mobile phones, and radio and transportation equipment as a proxies for fixed transaction costs, and distance to markets which varies among households as a proxy for variable transaction costs. Household's possession of transport equipment and their access to radio, TV, and mobile phones are conjectured to reduce transaction costs (i.e. search and information costs) and induce crop diversity.

Access to extension services is vital in assisting farmers in the production decision making process since it can be a reliable source of information, technical advice, trainings and improved farm management practices. Although relatively few farmers avail themselves of extension services (about 18% of the farmers in the sample have access to extension services), it is generally perceived as an important factor to control for. To an extent, extension services may depend on the country's agricultural policy, that is in some countries extension services may encourage farmers to produce certain staple crops to achieve self-sufficiency and ensure food security, whereas in other countries polices may target production for market. It is therefore difficult to priori predict the impact of extension services on the magnitude of crop divarication.

Afghanistan's climate is generally characterised by hot and dry summers and unequal distribution of rainfall throughout the year. Majority of the rainfall is accumulated over the spring season. While the main source of irrigation water for the irrigated land is the running water in rivers, canals, kariz[10], the irrigation water in rivers significantly decreases during the summer seasons, often leading to a water shortage. In addition, water

[10] As per the ALCS data, about 70% of irrigation water comes from rivers, Kariz, canals in spring of 2013. During the water shortages, farmers often use alternative means such as deep well pump to irrigate their crops which is costlier due to fuel costs.

requirement of crops increases in the summer seasons due to hot and dry weather. As per the descriptive statistics, about 45% of the farmers indicated that they did not have access to sufficient irrigation water. To account for variations in access to irrigation water throughout the year, a dummy variable equal to 0 for farms with insufficient irrigation water and 1 for availability of sufficient irrigation water was included. It is hypothesized that lack of sufficient irrigation water may restrict farmers to grow "certain[11] " crops.

Heterogeneity with respect to regional conditions may also largely affect level of crop diversity. Based on early work by Humlum (1959) revived by Dupree (1973), Afghanistan was divided into 11 geographical zones. However, recently a study by Maletta and Favre (2003) concluded that not all the 11 zones have agricultural significance (i.e. some zones were classified as deserts). Based on ecological properties of land and climate, and some supplementary criteria about accessibility and prevailing agricultural activities, Maletta and Favre (2003) adopted the 8 agro-ecological zones scheme (Figure 3.3a). These zones were constructed in the form of whole districts aggregations. Thus, in this study, we adapt the eight agro-ecological zoning scheme.

Since the eight agro-ecological zones are formed based on the aggregation of whole districts, we mapped the level of crop diversity by districts to illustrate the district-wise and zone-wise crop diversity (Figure 3.3a). Average district level CEI was first computed and classified into 4 categories. Given four levels of the CEI, the map shows the most diversified districts with green colour (CEI=0.36-0.67), and the least diversified districts with light green colour (CEI= 0-0.17).The grey areas on the map represent areas with no data (Figure 3.3b). These areas are either areas with no agricultural significance (i.e. deserts and mountains) or could not be covered by the survey. In addition, these areas may represent the households that were surveyed but did not report any involvement in agriculture activities (i.e. they are non-agricultural households and did not report crop production) as discussed earlier.

[11] Irrigation needs of crops vary from crop to crop. Generally, vegetables and fruits require more irrigation than cereals like wheat and barley. Descriptive analysis show that wheat, barley and melons are the most common crops grown in rain-fed land that require comparatively less irrigation.

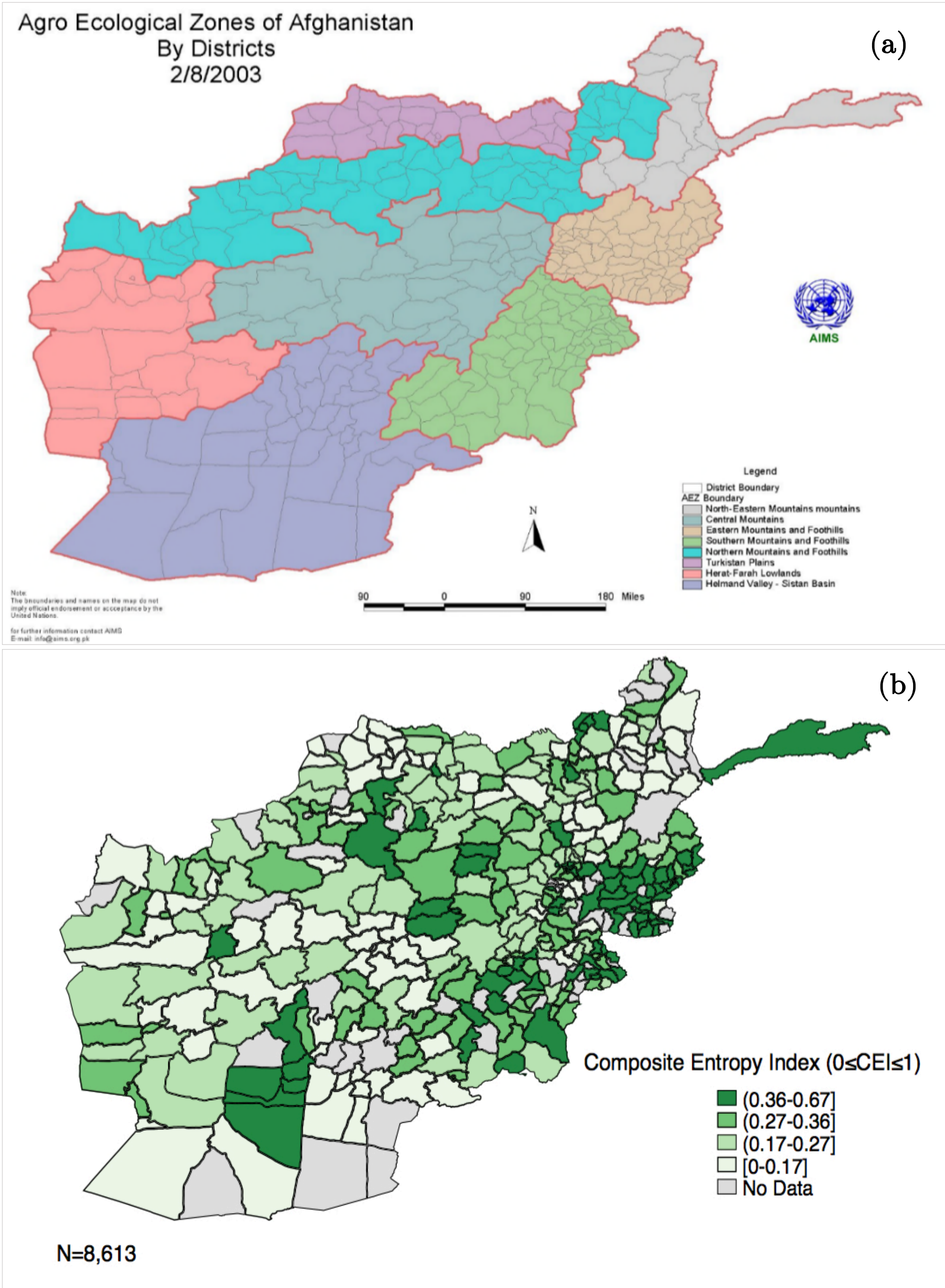



**Figure 3.3:** Agro-Ecological Zones; b) Map of CD at District Level

*Source:* 3.3a) Adapted from (Maletta and Favre, 2003); 3.3b) Author's calculations of the ALCS 2013-14 data

Availability of irrigation water and rain or snowfall throughout the year, crop yields, farm size, market infrastructure and conditions, and even cultural aspects of farmers may vary greatly by agro-ecological regions that may result in different levels of the extent of crop diversification. Heterogeneity in agroecology and regional differences captures these variations in physical and cultural environments. Among other unobserved climatic and cultural factors, the nature and extent of diversification is expected to differ across regions due to wide heterogeneity in farm size. Using different indices, Table 3.4 summarizes the degree of crop diversification by farm size across different agro-ecological zones. From

Table 3.4, we note that crop diversification consistently increases with the farm size but starting to decrease as farm size increase beyond 3 ha.

**Table 3.4:** CD by Farm Size (ha) Across AEZ's using Different Indices

| | Farm Size (ha) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | up to 1 | | 1-2 | | 2-3 | | more than 3 | | overall | |
| AEZ | CEI | THI | CEI | THI | CEI | THI | CEI | THI | CEI | THI |
| NEM | 0.15 | 0.15 | 0.28 | 0.27 | 0.31 | 0.31 | 0.28 | 0.28 | 0.20 | 0.20 |
| CM | 0.23 | 0.23 | 0.32 | 0.30 | 0.43 | 0.41 | 0.42 | 0.40 | 0.26 | 0.25 |
| HFL | 0.15 | 0.14 | 0.28 | 0.27 | 0.27 | 0.26 | 0.36 | 0.36 | 0.21 | 0.20 |
| SMF | 0.31 | 0.30 | 0.36 | 0.35 | 0.38 | 0.37 | 0.39 | 0.37 | 0.33 | 0.32 |
| HVSB | 0.20 | 0.20 | 0.28 | 0.28 | 0.38 | 0.38 | 0.41 | 0.40 | 0.27 | 0.27 |
| TP | 0.10 | 0.09 | 0.20 | 0.19 | 0.19 | 0.18 | 0.26 | 0.25 | 0.20 | 0.19 |
| NMF | 0.29 | 0.18 | 0.28 | 0.27 | 0.33 | 0.33 | 0.35 | 0.34 | 0.26 | 0.25 |
| EMF | 0.37 | 0.37 | 0.48 | 0.48 | 0.50 | 0.51 | 0.54 | 0.54 | 0.39 | 0.39 |
| Overall | 0.27 | 0.27 | 0.32 | 0.31 | 0.35 | 0.35 | 0.35 | 0.34 | 0.29 | 0.28 |
| N | | | | | | | | | | 8,613 |

*Source:* Author's calculations of the ALCS 2013-14 data

Moreover, intensity of crop diversification largely varies across agro-ecological zones. Eastern Mountains and Foothills appears to be the most diversified region. This can be explained by the availability of favourable agro-ecological conditions in this region to grow different crops/varieties and the existence of relatively better market conditions. On the other and Turkistan Plains appears to be the least diversified. Turkistan plains are maybe very specialized because traditionally wheat production is very common in the area.

### 3.7 Data, Summary Statistics, and Description of Variables

The data used to undertake the analysis in this chapter comes from the Afghanistan Living Condition Survey (ALCS) conducted by the Central Statistics Organization (CSO) in 2013/14 which were also used in the Chapter II. The data include both quantitative survey and in-depth qualitative information on several key indicators including farming and livestock production in Afghanistan.

Initial descriptive analysis of the data showed that as many as 9,642 households reported some involvement in agriculture. However, after accounting for missing values on key variables defined in this chapter, the total number of usable observations reduced to 8,853 households. Furthermore, the sample of agricultural households was further investigated to assess if the household who only grow a single crop on a very small amount of land (i.e. gardens) are systematically different from those who operate a relatively larger amount of land and grow major crops such as wheat, rice, cotton etc (see Table 3.A2 and 3.A3 and Figure 3.A1 Appendix 3.A). Based on the t-test (see Table 3.A4 in Appendix 3.A), the mean difference was found to be significant between these two categories, indicating that farmers who only maintain production in their garden may not be regular full-time farmers but grow some vegetables while undertaking off-farm activities as their main occupation. These farmers were therefore excluded from the sample, reducing the sample employed from 8,853 to 8,613 households.

Initially the total land variable was measured in Jeribs but to avoid small parameter estimates of the land variable, it was rescaled to hectares (1 hectare is equivalent to 5 Jeribs). Similarly, off-farm income which was originally measured in Afghani (AFN), was rescaled and measured in 10,000 AFN.

The descriptive statistics on the type of crop shows that there are a total of 22 different crops grown throughout the year (a typical agriculture year involves 1, 2, or 3 planting seasons). However, food grains such as wheat, maize, barley, and rice are the major crops. On average, wheat accounts for about 49.5% of the total value of revenue (physical output weighted by their respective prices), followed by maize (12%), rice (11.42%), potato (5.5%), and onion (5.17%). High value crops such as fruits and vegetables occupy a smaller share of the total revenues. Table 3.A1 and Figure 3.A1 in Appendix 3.A presents the frequency and total revenue of different crops grown.

Table 3.5 provides summary statistics for the dependent and all independent variables used in the analysis. Two different measures of crop diversification (Table 3.1) are used, CEI and THI indices as dependent variables, constructed based on the revenue share of individual crops that a household grow in different seasons throughout the year. Physical output of crops was weighted by their respective prices to calculate revenues (measured in Afghan currency symbolized by AFN) of individual crops.

The price data used to calculate revenues comes from the NRVA 2011-12 survey that collects data on prices of different agriculture commodities at the district-level. Lack of price data on some crops and unavailability of price data at the same year in which the ALCS survey was conducted is a limitation. However, for the purpose of this study, the price data were only used to convert physical quantities to revenues for individual crops, that are used in the CEI and THI calculations.

**Table 3.5:** Summary Statistics for Variables used in the Analysis

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| *Dependent Variable* | | | | |
| Composite Entropy Index (0≤CEI≤1) | 0.295 | 0.233 | 0.000 | 0.830 |
| Transformed Herfindahl Index (0≤THI≤1) | 0.283 | 0.232 | 0.000 | 0.830 |
| *Explanatory Variables* | | | | |
| Off-farm Income (in 10,000 AFN) | 5.519 | 11.05 | 0.000 | 480.0 |
| Total Land (Ha) | 1.564 | 4.227 | 0.020 | 211.2 |
| Transport Equipment (1=access, 0=otherwise) | 0.450 | 0.498 | 0.000 | 1.000 |
| Communication Equipment (1=access, 0=otherwise) | 0.798 | 0.402 | 0.000 | 1.000 |
| Cattle Ownership (N) | 1.477 | 1.943 | 0.000 | 31.00 |
| Oxen & Yaks (N) | 0.248 | 0.635 | 0.000 | 9.000 |
| Tractor & Thresher (N) | 0.052 | 0.231 | 0.000 | 4.000 |
| Land Quality (1=all irrigated, 0=irrigated & rainfed) | 0.437 | 0.496 | 0.000 | 1.000 |
| Landscape (1=open plain, 0=hills & valleys) | 0.753 | 0.431 | 0.000 | 1.000 |
| Sufficient Irrigation Water (1=access, 0=otherwise) | 0.448 | 0.497 | 0.000 | 1.000 |
| Household Size (persons) | 8.124 | 3.474 | 1.000 | 36.00 |
| Head Edu: No Formal Schooling (1=yes,0 otherwise) | 0.769 | 0.422 | 0.000 | 1.000 |
| Head Edu: Primary & Lower sec (1=yes, 0=otherwise) | 0.116 | 0.320 | 0.000 | 1.000 |
| Head Edu: Upper Secondary (1=yes, 0=otherwise) | 0.079 | 0.270 | 0.000 | 1.000 |
| Head Edu: Teacher College (1=yes, 0=otherwise) | 0.023 | 0.150 | 0.000 | 1.000 |
| Head Edu: Uni & Postgrad (1=yes, 0=otherwise) | 0.013 | 0.115 | 0.000 | 1.000 |
| Household Head Sex (0=F, 1=M) | 0.995 | 0.067 | 0.000 | 1.000 |
| Household Head Age (Years) | 44.11 | 13.90 | 13.000 | 98.00 |
| Extension Services (1=access, 0=otherwise) | 0.184 | 0.387 | 0.000 | 1.000 |
| Distance to Nearest Road (km) | 2.513 | 8.876 | 0.000 | 100.0 |
| Time to Market (1=Not reachable, 0=otherwise) | 0.044 | 0.204 | 0.000 | 1.000 |
| Time to Market (1=Less than 1h, 0 otherwise) | 0.548 | 0.498 | 0.000 | 1.000 |
| Distance to Market (1=More than 1h, 0 otherwise) | 0.408 | 0.492 | 0.000 | 1.000 |
| Agro-Ecological Zone 1: (1=NEM, 0=0therwise) | 0.023 | 0.151 | 0.000 | 1.000 |
| Agro-Ecological Zone 2 (1=CM, 0=otherwise) | 0.166 | 0.372 | 0.000 | 1.000 |
| Agro-Ecological Zone 3: (1=HFL, 0=otherwise) | 0.040 | 0.197 | 0.000 | 1.000 |
| Agro-Ecological Zone 4: (1=SMF, 0=otherwise) | 0.198 | 0.399 | 0.000 | 1.000 |
| Agro-Ecological Zone 5: (1=HVSB, 0=otherwise) | 0.105 | 0.306 | 0.000 | 1.000 |
| Agro-ecological Zone 6: (1=TP, 0=otherwise) | 0.068 | 0.252 | 0.000 | 1.000 |
| Agro-ecological Zone 7: (1=NMF, 0=otherwise) | 0.183 | 0.387 | 0.000 | 1.000 |
| Agro-ecological Zone 8: (1=EMF, 0=otherwise) | 0.216 | 0.412 | 0.000 | 1.000 |
| *Instruments* | | | | |
| IV1- Share of OFY in Total Income within District | 0.519 | 0.294 | 0.000 | 1.000 |
| IV2-Lag District Level OFY in 2011/12 (10K AFN) | 507.6 | 568.1 | 11.975 | 9,090 |
| N | 8,613 | | | |

*Source:* Author's calculations of the ALCS 2013-14 data

A considerable proportion (roughly 62%) of the sample households are engaged in off-farm activities, with a sample mean of 55K AFN of off-farm earnings per household. For households who actually have access to non-farm activities, the off-farm income is highly variable and ranges from a minimum of 10K to a max of 480K AFN with a standard deviation of 130K AFN. Some farm households clearly have significant opportunities to transfer and spread risks to off-farm activities. Besides, as discussed earlier, farmers with the highest off-farm income are probably part-time farmers who only maintain a garden to produce some vegetables for the household consumption. Their main occupation is outside the farm sector and perhaps do not allocate considerable amount of labour hours to farming.

Given the focus on diversification, the sample is divided into two sub-groups; diversifiers and non-diversifiers. T-test and chi2 tests were conducted to evaluate the mean difference between diversifiers and non-diversifiers. Summary statistics and mean difference of non-diversifiers and diversifiers are presented in Table 3.A5 in Appendix 3.A, the characteristics of the two-sub groups are significantly different from each other. Total annual revenue and farm income for diversifiers are significantly greater than those of non-diversifiers. Similarly, ownership of farm assets (cattle, oxen, and tractors) and use of purchased seed, fertilizer and expenditures are higher significantly higher for diversifiers. On the contrary, off-farm income and distance to the nearest road are insignificant but higher for the non-diversifiers. This is perhaps because non-diversifiers allocate a greater portion of their labour to off-farm activities which may will be their main activity.

### 3.7.1 Status and Patterns of Crop Diversification

The concept of crop diversification in this study implies production of multiple crops on the farm throughout the year by an individual household. The Composite Entropy and Transformed Herfindahl indices are used to measure the level of crop diversification or specialization. Descriptive statistics of number of crops grown by households show that 33% (equivalent to 2,830 out of 8,613) households grow one crop, 48% of the farmers grow two crops, 16.5% grow three crops, and about 3.5% grow four or more, with a sample average of 1.92 crops (Figure 3.4a). For more summary statistics see Appendix 3.A (distribution of households based on the number of crops in Table 3.A6, cropping activities across agro-ecological zones in Table 3.A7, and detailed summary statistics of variables by the number of crops are presented in Table 3.A8).

Similar statistics were generated using CEI and THI indices to measure crop diversification. Average CEI and THI for the overall sample were calculated to be 29.5% and 28.3% with standard deviation of 0.23 respectively (Table 3.5 of the summary statistics), whereas the CEI and THI among diversifiers is 0.44% and 42%. This suggests a low level of crop diversification relative to other comparable countries (see Table 3.7 in the following page).

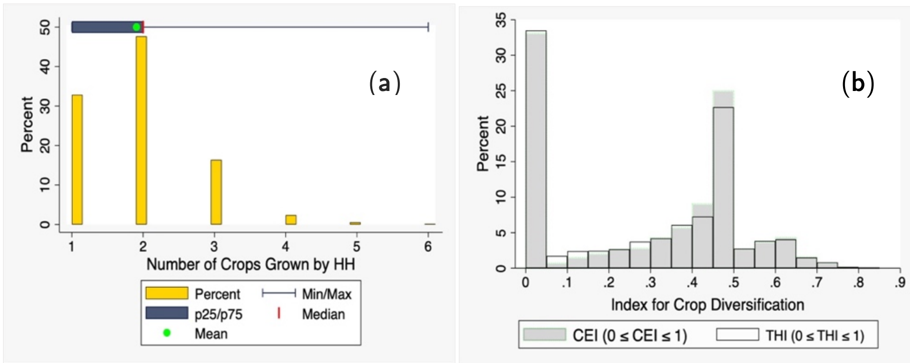



**Figure 3.4:** Distribution of Farms with Respect to **a**) Number of Crops, **b**) CEI & THI

*Source:* Author's calculations of the ALCS 2013-14 data

The distribution of CEI and THI are presented in Table 3.6 and Figure 3.4b. The computed value of the CEI for 33% of the households is zero indicating that they did not diversify (i.e. growing only one crop), whereas for 52% of farms the value of CEI is between 0.1 and 0.50, and for the remaining 15% CEI falls between 0.50 and 0.82. The distribution of the THI is quite similar to the CEI (Table 3.6). Since THI is mainly based on the revenue shares of individual crops that households grow and CEI accounts for variation in the number of crops in addition to the share of revenue of crops, one would expect slightly lower estimates of the THI.

**Table 3.6:** Distribution of CEI and THI

| | CEI | | THI | |
|---|---|---|---|---|
| Range | No. of Farms | % of Farms | No. of Farms | % of Farms |
| 0 | 2,911 | 33.80 | 3,034 | 35.23 |
| 0.1-<0.2 | 304 | 3.53 | 419 | 4.86 |
| 0.2-<0.3 | 470 | 5.46 | 552 | 6.41 |
| 0.3-<0.4 | 858 | 9.96 | 888 | 10.31 |
| 0.4-<0.5 | 2,827 | 32.82 | 2,496 | 28.98 |
| 0.5-<0.6 | 655 | 7.60 | 653 | 7.58 |
| 0.6-<0.7 | 512 | 5.94 | 481 | 5.58 |
| 0.7-<0.82 | 76 | 0.88 | 90 | 1.04 |
| N | 8,613 | | 8,613 | |

*Source:* Author's calculations of the ALCS 2013-14 data

For compression with other countries, Table 3.7 illustrates the relatively low level of diversification of Afghanistan's farming sector: the average value of the index for crop diversification of 0.30 in Afghanistan is comparatively lower than most other regions except Cambodia (Table 3.7).

**Table 3.7:** Comparison of CD Across Different Countries

| Study | Country of Study | CD Measure | Avg. CD value |
|---|---|---|---|
| Acharya et al., (2011) | Karnataka, India | CEI | 0.66 |
| Mandal and Bezbaruah, (2013) Benin et al. (2004) | Assam Plains | CEI | 0.39 |
| Hitayezu et al., (2016) | Kwazulu-Natal, South Africa | THI | 0.44 |
| Kumar et al., (2012) | Eastern India | THI | 0.44 |
| (Seng, 2014) | Cambodia | THI | 0.12 |
| Mofya-Mukuka and Hichaambwa, (2016) | Zambia | THI | 0.37 |

*Note:* The reported HHI for South Africa is subtracted from 1 to produce THI for better comparison.

### 3.8 Empirical Results and Discussion

Using CEI and a binary variable (for diversifiers vs non-diversifiers) as the dependent variables based on Equations (3.12) and (3.13), the estimated results from the Instrumental Variable tobit and probit models are reported in column 3 of Tables 3.8 and 3.9, with the basic uninstrumented tobit and probit models in column 2 (to assess the effect of allowing for endogeneity). Estimated results by tobit and probit models do not yield dramatically different results (i.e. the direction and magnitude of marginal effects of all explanatory variables are largely and qualitatively similar). As the estimated raw coefficients from the tobit and probit models do not have an economic interpretation, unconditional marginal effects (also referred to as the unconditional expected value) are reported to show the effects of the independent variables on the overall level of diversification (i.e. for both non-diversifiers and diversifiers). In case of Probit, we present average marginal effects of the explanatory variables. For continuous variables, the marginal effects measure the change in probability of the observed y given a one unit change of the independent variables, holding all other variables at their mean. For discrete choice variables a change from 0 to 1, leaving all other variables constant at their mean.

The reduced-form model (Equation 3.13) for off-farm income is estimated using OLS and presented as the first stage estimates in Table 3.8. Conditional on other covariates, the results of the first stage demonstrate strong correlation between the two instruments and the endogenous off-farm income. Both instruments, share of the aggregate off-farm income in total income within a district and the district level lagged off-farm income from 2011/12, are positive and significantly correlated with the endogenous off-farm income at 1% level as expected. The strong correlation of instruments with the endogenous variable imply that instruments are relevant. Potential endogeneity in off-farm income is investigated by applying the Wald test of exogeneity. The calculated test statistic is 142.49 and 96.22 (based on of IV-Tobit and IV-Probit models respectively) and rejects the null hypothesis of no endogeneity in off-farm income at 1% significance level. This indicate that household nonfarm income is endogenous conditional on IV validity.

Test of validity of instruments was conducted using the Amemiya-Lee-Newey overidentification test estimator. The null hypothesis of over identification test is that the instruments are jointly valid, and that the excluded instruments are correctly excluded from the estimated equation. Rejections of the null hypothesis will mean that the instruments are not valid and vice-versa. As per the test statistic in Tables 3.8 and 3.9, the result of Amemiya-Lee-Newey is insignificant in both cases of Tobit and Probit estimations, thus establishing the validity of the instruments.

Estimated parameters from both the uninstrumented and the instrumented tobit models illustrate negative and significant impact of household non-farm income on the extent or likelihood of crop diversification. This implies that holding other variables at their mean, an increase of 10,000 Afghani in off-farm income (equivalent to almost 20% of mean off-farm income) decreases CD by 0.002 (a small effect corresponding to a reduction in CEI at the mean from 0.295 to 0.293). This is consistent with the hypothesis that allocation of farm labour away to off-farm activities decrease diversity due to negative labour effects, particularly when the opportunity cost of household labour is higher than the off-farm wages under imperfect markets implying non-separability between households' farm profits and off-farm earnings The impact of non-farm income on the level of crop diversification is even greater (an increase of 10,000 in off-farm income reduces CEI by 0.015) when endogeneity is controlled for (the effect corresponds to a reduction in CEI at the mean from 0.295 to 0.28). In other words, increasing off-farm income by the mean value (or going from zero to 55K) would reduce mean CEI from 0.295 to 0.21.

Similarly, the results from both uninstrumented and endogenous probit models reveal a negative and significant effect of off-farm income on diversification; an increase of 10K Afghani in off-farm income would decrease the likelihood of diversification (growing multiple crops or probability of $y=1$) by 0.3 and 7 percentage points in case of uninstrumented and endogens models respectively. In other words, increasing off-farm income by the mean value (or going from zero to 55K) would reduce the likelihood of $y = 1$ by 1.65 and 38.5 percent points based on the estimates of uninstrumented and instrumental variable probit estimates respectively.

This illustrates that failing to account for the endogeneity of the household nonfarm income underestimates its negative impact on the intensity of crop diversification. This finding suggests that instrumenting for off-farm income controls for the bias due to unobserved factors, such as risk-aversion behaviour of farmers, that positively influence both nonfarm earnings and magnitude of crop diversification (Corr (off-farm income, $\varepsilon$) $> 0$ in Equation 3.10). The estimated effect of off-farm income by the uninstrumented models is biased upwards (more negative) compared to endogenous estimations; this may be due to measurement error, or unobserved risk aversion. When we allow for endogeneity in IV-tobit and IV-probit, effect of off-farm income is more negative (upward bias removed). More productive farms appear to have adopted greater CD and less off-farm income, and farms with off-farm income diversify less.

Our findings of negative impact of off-farm income are consistent with the conclusions of earlier studies that assessed the impact of non-farm income on crop diversity. Abdalla et al. (2013) found a significant reduction in the degree of crop diversification is mainly associated with the engagement in off-farm activities during the agricultural season in Sudan. Weiss and Briglauer (2000) found that the existence of additional off-farm income reduces the degree of diversification and argued that engaging in non-farm activities is used as a strategy by Upper-Austrian farmers to diversify employment risks and thus reduces the necessity to diversify on the farm. Mishra et al., (2004) also found a negative relationship between off-farm income and crop diversity and pointed out that off-farm-

**Table 3.8:** Unconditional Marginal Effects of the Tobit and IVTobit

| *Dependent variable: CEI* | Tobit | | Instrumental Variable Tobit | | | |
|---|---|---|---|---|---|---|
| | | | 1st stage | | 2nd stage | |
| Variable | ME | SE | Coefficient | SE | ME | SE |
| Off-farm Income (10K AFN) | -0.002*** | 0.000 | - | - | -0.015*** | 0.001 |
| Total Land (Ha) | 0.004*** | 0.001 | -0.062** | 0.027 | 0.003*** | 0.001 |
| Transport Equip. (1=access) | 0.020*** | 0.006 | 0.914*** | 0.251 | 0.031*** | 0.007 |
| Communication Equip (1=yes) | 0.015** | 0.007 | 0.373 | 0.298 | 0.026*** | 0.008 |
| Cattle Ownership (N) | 0.006*** | 0.001 | -0.137** | 0.059 | 0.004** | 0.002 |
| Oxen & Yaks (N) | 0.037*** | 0.005 | -0.696*** | 0.191 | 0.024*** | 0.005 |
| Tractor & Thresher (N) | 0.035*** | 0.012 | 0.407 | 0.496 | 0.051*** | 0.013 |
| Land Quality (1=good) | 0.050*** | 0.008 | -0.279 | 0.340 | 0.050*** | 0.009 |
| Landscape (1=open plain) | 0.055*** | 0.007 | 0.777*** | 0.270 | 0.064*** | 0.007 |
| Irrigation Water (1=access) | 0.024*** | 0.006 | 0.469** | 0.234 | 0.029*** | 0.006 |
| Household Size (persons) | 0.006*** | 0.001 | 0.588*** | 0.035 | 0.014*** | 0.001 |
| Head Edu (1=primary & sec) | 0.014* | 0.009 | 1.396*** | 0.350 | 0.041*** | 0.010 |
| Head Edu (2=upper sec) | 0.028*** | 0.010 | 3.584*** | 0.419 | 0.088*** | 0.013 |
| Head Edu (1=teacher college) | 0.0001 | 0.018 | 2.107*** | 0.736 | 0.042** | 0.021 |
| Head Edu (1=uni & grad) | 0.014 | 0.023 | 6.786*** | 0.952 | 0.128*** | 0.031 |
| Head Sex (1=male) | 0.101*** | 0.034 | 0.026 | 1.618 | 0.080** | 0.040 |
| Head Age (years) | -0.000 | 0.000 | 0.002 | 0.008 | 0.0002 | 0.000 |
| Extension Services (1=access) | -0.012* | 0.007 | -1.107*** | 0.294 | -0.017** | 0.008 |
| Distance to Road (km) | -0.003*** | 0.001 | -0.044 | 0.039 | -0.005*** | 0.001 |
| Time to Market (1=< 1h) | 0.007 | 0.014 | 0.372 | 0.559 | 0.030** | 0.015 |
| Time to Market (2=>1h) | 0.025* | 0.013 | -0.317 | 0.549 | 0.016 | 0.015 |
| Agro-Ecological Zone 1 (CM) | 0.081*** | 0.016 | -0.877 | 0.768 | 0.041** | 0.000 |
| Agro-Ecological Zone 2 (HFL) | -0.001 | 0.019 | 1.051 | 0.927 | -0.023 | 0.023 |
| Agro-Ecological Zone 3 (SMF) | 0.141*** | 0.017 | 1.220 | 0.785 | 0.130*** | 0.020 |
| Agro-Ecological Zone 4 (HVSB) | 0.039** | 0.018 | -0.799 | 0.878 | -0.049** | 0.022 |
| Agro-Ecological Zone 5 (TP) | -0.020 | 0.017 | -0.517 | 0.867 | -0.059*** | 0.021 |
| Agro-Ecological Zone 6 (NMF) | 0.103*** | 0.017 | -0.700 | 0.773 | 0.065*** | 0.020 |
| Agro-Ecological Zone 7 (EMF) | 0.184*** | 0.017 | -1.103 | 0.777 | 0.162*** | 0.020 |
| IV1- Share of Off-farm Income in Total Income within District | - | - | 10.658*** | 0.478 | - | - |
| IV2-Lag District Level OFY | - | - | 0.001*** | 0.000 | - | - |
| Constant | - | - | -5.952*** | 1.876 | - | - |
| Log-Likelihood | -3,981.40 | | - | | -35,949.00 | |
| Pseudo R-Square | 0.121 | | | | - | - |
| Wald Test of exogeneity (chi2, p-value) | - | - | | | 142.25*** | 0.000 |
| Amemiya-Lee-Newey statistic (chi2, p-value) | - | - | - | - | 0.500 | 0.479 |
| Left censored observations(N) | 2,830 | | - | - | 2,830 | |
| Uncensored observations (N) | 5,782 | | - | - | 5,782 | |
| N | 8,613 | | | | 8,613 | |

*Notes: The omitted categories are: no access transport and communication equipment, rain-fed & irrigated land for land quality, hills and valleys for landscape, no access to sufficient water for irrigation, no formal schooling for education, no access for extension services, female for HH head sex, not reachable for distance to market, and AEZ 8 for AEZ. significance indicated by * p<0.10, ** p<0.05, *** p<0.010.*

**Table 3.9:** Average Marginal Effects of the Probit and Instrumental Variable Probit

| *Dependent variable: (binary=0 no diversity, 1=diversification)* | Probit | | **Instrumental Variable Probit** First Stage | | IV-Probit | |
|---|---|---|---|---|---|---|
| | ME | se | Coefficient | SE | ME | se |
| Off-farm Income (10K AFN) | -0.003*** | 0.000 | - | - | -0.070*** | 0.009 |
| Total Land (Ha) | 0.028*** | 0.003 | -0.062** | 0.027 | 0.052*** | 0.015 |
| Transport Equip (1=access) | 0.022* | 0.012 | 0.926*** | 0.251 | 0.102*** | 0.031 |
| Communication Equip (1=yes) | 0.008 | 0.014 | 0.394 | 0.298 | 0.071** | 0.034 |
| Cattle Ownership (N) | 0.010*** | 0.003 | -0.144** | 0.060 | 0.012 | 0.008 |
| Oxen & Yaks (N) | 0.067*** | 0.010 | -0.685*** | 0.191 | 0.099*** | 0.025 |
| Tractor & Thresher (N) | 0.108*** | 0.029 | 0.522 | 0.503 | 0.311*** | 0.082 |
| Land Quality (1=good) | 0.097*** | 0.017 | -0.288 | 0.341 | 0.226*** | 0.042 |
| Landscape (1=open plain) | 0.128*** | 0.012 | 0.806*** | 0.268 | 0.344*** | 0.035 |
| Irrigation Water (1=access) | 0.039*** | 0.011 | 0.469** | 0.235 | 0.106*** | 0.029 |
| Household Size (persons) | 0.011*** | 0.002 | 0.589*** | 0.035 | 0.060*** | 0.010 |
| Head Edu. (1=primary & sec) | 0.032* | 0.016 | 1.378*** | 0.351 | 0.194*** | 0.047 |
| Head Edu. (2=upper sec) | 0.074*** | 0.019 | 3.572*** | 0.420 | 0.435*** | 0.058 |
| Head Edu. (1=teacher college) | 0.013 | 0.036 | 2.103*** | 0.737 | 0.222*** | 0.083 |
| HH Head Edu. (1=uni & grad) | 0.021 | 0.046 | 6.757*** | 0.954 | 0.523*** | 0.145 |
| HH Head Sex (1=male) | 0.165* | 0.086 | -0.049 | 1.620 | 0.232 | 0.176 |
| HH Head Age (years) | -0.0001 | 0.000 | 0.002 | 0.008 | 0.001 | 0.001 |
| Extension Services (1=access) | -0.005 | 0.014 | -1.123*** | 0.295 | -0.043 | 0.035 |
| Distance to Road ( km) | 0.002** | 0.001 | -0.022* | 0.013 | 0.001 | 0.002 |
| Time to Market (1=<1 hr) | -0.009 | 0.026 | 0.262 | 0.564 | 0.069 | 0.057 |
| Time to Market (2=>1 hr) | 0.044* | 0.026 | -0.448 | 0.553 | 0.051 | 0.055 |
| Agro-Ecological Zone 1 (NEM) | 0.232*** | 0.039 | -0.922 | 0.769 | 0.267*** | 0.087 |
| Agro-Ecological Zone 2 (CM) | 0.032 | 0.047 | 1.011 | 0.929 | -0.069 | 0.147 |
| Agro-Ecological Zone 3 (HFL) | 0.360*** | 0.039 | 1.273 | 0.787 | 0.732*** | 0.094 |
| Agro-Ecological Zone 4 (SMF) | 0.068 | 0.043 | -0.842 | 0.881 | -0.328*** | 0.108 |
| Agro-Ecological Zone 5 (TP) | 0.016 | 0.045 | -0.562 | 0.869 | -0.162 | 0.101 |
| Agro-Ecological Zone 6 (NMF) | 0.232*** | 0.039 | -0.804 | 0.775 | 0.284*** | 0.088 |
| Agro-Ecological Zone 7 (EMF) | 0.332*** | 0.039 | -1.096 | 0.779 | 0.604*** | 0.096 |
| IV1- Share of Off-farm Income in Total Income within District | - | - | 10.612*** | 0.485 | - | - |
| IV2-Lag District Level OFY | - | - | 0.001*** | 0.000 | - | - |
| Constant | - | - | -5.833*** | 1.881 | - | - |
| Log-Likelihood | -4,936.36 | | -32,065.16 | | -36,900.66 | |
| chi2 (p-value) | 1,034.15*** | 0.000 | - | - | 621.97*** | 0.000 |
| Pseudo R-square | 0.095 | - | - | - | | |
| Wald Test of exogeneity (chi2, p-value) | - | - | - | - | 96.22*** | 0.000 |
| Amemiya-Lee-Newey statistic (chi2, p-value) | - | - | - | - | 0.548 | 0.459 |
| N | 8,613 | | 8,613 | | 8,613 | |

*Note: The omitted categories are: no access for transport equipment and communication equipment, rain-fed and irrigated combined for land quality, hills and valleys for landscape, no access to sufficient irrigation water for irrigation, no formal schooling for education ,no access for extension services, female for HH head sex, not reachable for distance to market, and AEZ 8 for AEZ. significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

- income diversifies a farm operator's income portfolio and reduces the need for on-farm enterprise diversification. Our findings are in contrast of Cavatassi et al., (2012) that found positive causal relationship between off-farm income and level of diversity in Hararghe Ethiopia and argued that the anticipated relationship between participation in non-farm activity and diversity depends largely on the motivation of the households. If participation in off-farm activities is primarily done with the intent of relaxing liquidity - constraints, it may enhance diversity by allowing households to purchase inputs. However, if off-farm income is regarded as an alternative to agricultural production and thus takes away labour from crop production it may lead to lower diversity.

Land holding size (i.e. total land cultivated by farm household) significantly increases crop diversity. Holding all variables at their mean, increase in land by one hectare increases CEI by 0.003 at the mean (or increase 2.2 percent points in the likelihood of $y = 1$ in the case of iv-probit) based on the IV-Tobit estimation (alternatively an increase in land by the mean value or 1.6 ha, would increase mean CD from 0.295 to 0.30. Based on the IV-Probit estimates, an increase of 1ha in land will increase the likelihood of going from no diversification to diversification by 5.2 percent points (corresponds to an increase of 8.32% if land is increased by the mean value of 1.6ha). This small effect for a relatively large increase in land is perhaps due the fact that farm households with the largest land size are may be commercial farmers that tend to specialize. The overall positive effect of land size on crop diversity indicates that households with a relatively larger land size have the flexibility to allocate land among a variety of crops and therefore diversify.

These findings are consistent with those of Sichoongwe et al., (2014) for Zambia, Hitayezu et al., (2016) for South Africa, Kasem and Thapa (2011) for Thailand, and McNamara and Weiss (2005) for Austria. In assessing the impact of farm size on the level of crop diversity. Pope and Prescott (1980) found a positive and quadratic relationship between farm size and diversity for California crop farmers and offered the argument that there is a trade-off between scale economies and risk reduction. That is, if there are large-scale economies in an enterprise, then one might expect larger farms to be more specialized.

These claims are further supported by the descriptive analysis of the data (Figure 3.5). With larger farm size, farm income increases whereas off-farm income falls, indicating that farmers with larger farm sizes may allocate more labour to farming and therefore stick with farming, whereas farmers with smaller size of land are part-time farmers that may engage in off-farm activities as their main source of livelihood. On the other hand, as farm size increases, the number of crops grown initially increases, but starts to decline when land size is beyond three hectares, supporting the hypothesis that households with the largest farm size may specialize.

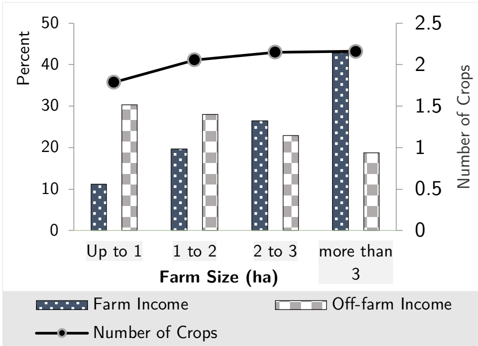



**Figure 3.5:** Farm-, Off-farm Income, & Number of Crops by Farm Size (ha)
*Source:* Author's calculations of the ALCS 2013-14 data

The estimated ME for the distance to road variable and time taken to reach market differ between the Tobit and Probit models; estimated ME from the tobit are significant whereas MEs from Probit are insignificant. This is probably due the fact that distance to road and time to market aren't considerably relevant in probit model that analysis the household's decision to go from no diversification to diversification but are significant in Tobit using CEI that also captures the extent of diversity which increases with the number of crops.

Farm households living in communities with poor access to all-season roads exhibit lower crop diversity, whereas households with better access to roads and permanent food markets maintain higher level of crop diversity. Improved access to roads and closer proximity to markets implies lower transaction costs due to better market infrastructure, transport and storage facilities. In addition, improved access to roads and local markets provide farming communities with better access to market information on prices for output and inputs. High-value horticultural crops such as vegetables and fruits are perishable and require sustained supply chain in order for the households to sell them in local markets. Thus, farmers located close to markets and with high road density are more likely to diversify into producing cash crops. Rao et al., (2008) finds a significant and positive impact of road density on diversification towards high value horticultural crops in India.

Turner (2014) indicated that Mozambican farmers lacking access to transport infrastructure do not allocate land to marketable cash crops. Mesfin et al., (2011) finds that households that had access to information on market prices, supply, and demand are more diversified. Our findings confront with those of Sichoongwe et al., (2014) that found a positive and significant impact of distance to market on crop diversity in Zambia and indicated that farmers located further from markets may diversify for food security as their access to market is limited.

Other proxies for transaction costs such as the ownership of transport equipment by the households and their access to communication equipment such as television, mobile phone, and radio were also found to have a significant and positive influence on the extent of crop diversity (in both tobit and probit models). This further supports the argument that lower transaction costs enhance crop diversification. Better access to market information on input and output prices as proxy for fixed transaction costs (i.e. search costs) assists farmers in production decision making. Ownership of transport equipment introduces efficiency to the cost function through availability of low-cost means of transport. Seng (2014) found that ownership of transport equipment significantly increases crop diversity in Cambodia and claimed that ownership of transport equipment reduces variable transaction costs (e.g. transport costs), providing incentive for the farmers to diversify crop portfolio, particularly increasing the production of cash crops for the purpose of selling in the market.

Households with greater number of livestock (cattle and oxen) maintain higher level of crop diversity. Cavatassi et al., (2012) pointed out that owners of oxen tend to plant greater number of crops which is perhaps due the mechanical power provided by the oxen that makes the cultivation easier. Benin et al. (2004) and Abay et al., (2009) found that oxen ownership contributes positively to crop diversity through ensuring draught power for ploughing when needed by the households. Ownership of larger cattle herd increases the amount of manure produced at the farm that enhances soil fertility through adding organic materials to the soil and thereby positively influences the crop intensification process. Farm households that own tractors maintain higher degree of diversity. Tractor ownership by the farm households contributes to utilizing lands more efficiently and increases production efficiency through availability of cheaper and timely traction power at the time of cultivation. In addition, household may use tractors to transport their produce to the market. Our results agree with previous studies (Abay et al., 2009; Abdalla et al., 2013; Dube, 2016).

Agricultural extension services appear to have a significant negative impact on the extent of crop diversification. This is perhaps due to the policy emphasis on achieving self-sufficiency in producing staple grain food crops. While grain, particularly wheat, is the major source of nutrition, Afghanistan still imports a substantial quantity of wheat flour so there is an aim to produce more grains domestically. This is consistent with the findings of Mesfin et al., (2011) for farm households in eastern Ethiopia where number of extension visits were found to decrease crop diversity. They argued that the negative impact of extension services is associated with the extension system favouring specialization at macro level and overlooks the role of crop diversification in risk minimization. Similarly, Abay et al., (2009) found negative association between extension services and crop diversity in Northern Ethiopia and concluded that the agricultural policy incentivise production of legume and cereal crops in Tigray.

There appears to be a significant and positive relationship between land quality and crop diversification. Farmers operating on irrigated land alone are significantly more diversified than their counterparts who operate a combination of irrigated and rain-fed land. In addition, households with stable access to sufficient irrigation water throughout the year appear to be more diversified. Afghanistan in general is a dry country and the main source of irrigation is running water in canal. During the hot months of summer, irrigation water often decreases that in turn has an adverse impact on farming. As a result, farmers are restricted to grow limited number of crops, particularly since many vegetables require greater amount of irrigation. Mesfin et al., (2011) confirms that irrigation intensity has a positive effect on crop diversity by enabling farmers to grow vegetables along other grains.

Farmers operating in the plains or on flat land diversify more compared to farmers with land in valleys and hills. Altitude and slope of land effects physical conditions of farming which translates into the household decisions on the number and type of crops they choose to grow. Cavatassi et al., (2012) indicated that variability in slope of the farm land leads to greater variability in diversity. Our results are in contrast of those of Van Dusen and

Taylor (2005) who found that Mexican farms located in areas with steep slope are more diversified.

Except for the age of the household head, household characteristics are positively and significantly associated with crop diversification, in line with findings in the literature: household head education; household size, as a proxy for the labour supply; and households headed by a male. There is no statistically significant association between household head age and crop diversity.

We control for eight agro-ecological regions: Eastern Mountains and Foothills (EMF), Southern Mountain and Foothills (SMF), and Central Mountains (CM) were the most favourable for crop diversification compared to the reference zone (Norther Eastern Mountains). Farm households in Turkistan Plain (TP) and Helmand Valley and Sistan Basin (HVSB) zones are the least diversified. Among other heterogenous unobserved effects such as climatic, physical conditions, and cultural conditions, the level to off-farm employment/income, access to farm land, market development infrastructure and market conditions, and road density are expected to greatly vary from region to region.

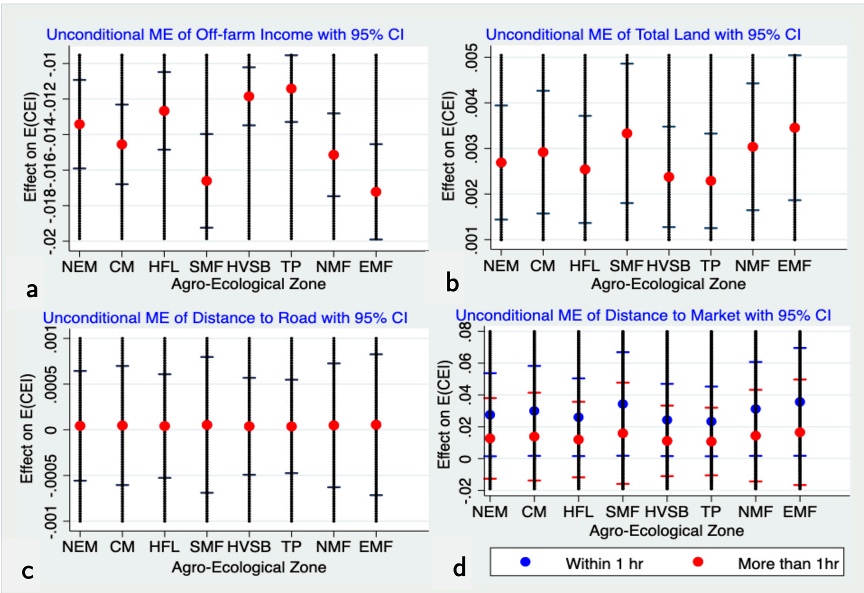



**Figure 3.6:** Unconditional ME's from the Tobit Model of **a**) Off-farm income, **b**) Cropped Area, **c**) Time Taken to Reach Food Market, and **d**) Distance to Road on the Expected Value of CEI Across Agro-ecological Zones

*Source:* Author's calculations of the ALCS 2013-14 data

In Figure 3.6, we plot the unconditional marginal effects of off-farm income, total cropped land, and household's distance to market and roads against the expected value of CEI across agro-ecological zones to show varying effects across agro-ecological regions. In EMF and SMF regions, land size and proximity to permanent food markets have the largest positive impact on crop diversity (Figure 6b and 6c), whereas the negative impact of off-farm income and distance to road is the least for EMF and SMF zones (Figure 3.6a and 3.6d). For the significantly least diversified zones such as TP and HVSB these effects follow the opposite direction.

### 3.9 Robustness and Specification Tests

The model was checked for potential multicollinearity problem using Variance Inflation Factor (VIF) and Tolerance values. Based on the computed values for VIF and Tolerances, the calculated values for all variables except for age-squared are less than the cut-off point of 10 for VIF and greater than the cut-off of 0.1 for tolerances respectively, indicating no serious multicollinearity problem in data. As the coefficient was insignificant, age squared was dropped from the model.

For the purpose of sensitivity of results to using different measures of crop diversification, we re-estimated the IV Tobit model using an alternative measure of crop diversity, the Transformed Herfindahl Index (THI) of relative abundance as a dependent variable. The results are reported in Table 3.B1 in Appendix 3.B and are largely similar to those using CEI as the dependent variable.

The descriptive analysis (particularly Figure 3.5) may suggest that the relationship between off-farm income and crop diversification could be more complex (i.e. not linear). For this reason, we test the linearity/non-linearity of the off-farm income by including a square term of the off-farm income on our model. The results from the preferred iv-tobit model show that the square term is non-significant, hence the effect appears linear and our main results are unaffected.

Tobit model assumes homoscedastic and normality of the error term. Despite that all standard errors reported with the main results presented earlier are robust, we complement our analysis by reporting the results of Powell's (1984) Censored Least Absolute Deviation (CLAD) and 2SLS models. The CLAD estimator relies significantly less on distributional assumption and if heteroscedasticity is present in the data the CLAD estimator will produce consistent results. The point estimates obtained using CLAD and 2SLS estimators in Table 3.9 are qualitatively similar to our main results reported in Table 3.9.

**Table 3.10:** Marginal Effects from CLAD and 2SLS Estimators

| *Dependent variable: CEI* | CLAD | | 2SLS | |
|---|---|---|---|---|
| | ME | SE | ME | SE |
| Off-farm Income (10K AFN) | -0.002*** | 0.000 | -0.013*** | 0.001 |
| Total Land (Ha) | 0.014*** | 0.001 | 0.003*** | 0.001 |
| Transport Equip.(1=access) | 0.026*** | 0.006 | 0.027*** | 0.006 |
| Communication Equip. (1=access) | 0.034*** | 0.007 | 0.024*** | 0.007 |
| Cattle Ownership (N) | 0.004*** | 0.001 | 0.004** | 0.001 |
| Oxen & Yaks (N) | 0.047*** | 0.004 | 0.023*** | 0.005 |
| Tractor & Thresher (N) | -0.024** | 0.012 | 0.044*** | 0.012 |
| Land Quality (1=good) | 0.081*** | 0.008 | 0.049*** | 0.008 |
| Landscape (1=open plain) | 0.041*** | 0.006 | 0.057*** | 0.007 |
| Sufficient Irrigation Water (1=access) | 0.035*** | 0.005 | 0.028*** | 0.006 |
| Household Size (persons) | 0.007*** | 0.001 | 0.013*** | 0.001 |
| Head Edu (1=primary & lower sec) | 0.015* | 0.008 | 0.034*** | 0.009 |
| Head Edu (2=upper secondary) | 0.018* | 0.010 | 0.073*** | 0.011 |
| Head Edu (1=teacher college) | 0.018 | 0.017 | 0.033* | 0.019 |
| Head Edu (1=uni & graduate) | 0.002 | 0.021 | 0.103*** | 0.025 |
| HH Head Sex (1=male) | 0.077* | 0.041 | 0.069* | 0.040 |
| HH Head Age (years) | -0.000 | 0.000 | 0.0002 | 0.000 |
| Extension Services (1=access) | -0.003 | 0.007 | -0.019*** | 0.007 |
| Distance to Nearest Road (10 km) | -0.005*** | 0.001 | -0.004*** | 0.001 |
| Distance to Market (< 1 hr) | -0.017 | 0.012 | 0.026* | 0.014 |
| Distance to Market (> 1 hr) | 0.007 | 0.012 | 0.011 | 0.014 |
| Agro-Ecological Zone 1 (CM) | 0.203*** | 0.022 | 0.021 | 0.019 |
| Agro-Ecological Zone 2 (HFL) | -0.064** | 0.026 | -0.043* | 0.023 |
| Agro-Ecological Zone 3 (SMF) | 0.276*** | 0.022 | 0.092*** | 0.019 |
| Agro-Ecological Zone 4 (HVSB | 0.238*** | 0.023 | -0.066*** | 0.022 |
| Agro-Ecological Zone 5 (TP) | -0.028 | 0.024 | -0.092*** | 0.022 |
| Agro-Ecological Zone 6 (NMF) | 0.214*** | 0.022 | 0.046** | 0.019 |
| Agro-Ecological Zone 7 (EMF) | 0.361*** | 0.022 | 0.140*** | 0.019 |
| R-Square | | | 0.559 | |
| Pseudo R-Square | 0.099 | | | |
| Durbin-Wu-Hausman exogeneity test (chi2, p-val) | - | - | 162.18*** | 0.000 |
| Sargan-Hansen overidentification (chi2, p-va) | | | 0.042 | 0.837 |
| Cragg-Donald Wald F Statistic Weak identification (ch2, p-value) | | | 258.06*** | 0.000 |
| N | 8,445 | | 8,613 | |

*Notes: The omitted categories are: no access for transport equipment and communication equipment, rain-fed & irrigated for land quality, hills and valleys for landscape, no access to sufficient irrigation water for irrigation, no formal schooling for education, no access for extension services, female for head sex, not reachable for distance to market, & AEZ 8 for AEZ, Significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

Further to the Amemiya-Lee-Newey overidentification test presented earlier, a set of minimum distance version weak-instrument-robust tests were also carried out to examine the validity of the excluded instruments. These tests include Anderson-Rubin (AR), Conditional Likelihood Ratio (CLR), the Lagrange Multiplier (LM), overidentification (J), and a combination of the LM and J over identification (K-J) tests. These tests were carried out using the *rivtest* command in Stata 15. The confidence intervals for the off-farm income coefficient produced from the weak-instrument tests in Table 3.10 are not wider than the non-robust Wald confidence intervals, indicating that instruments are strong and that point estimates are robust to possible weak instrument bias. For further discussion on the weak-instrument-robust tests in limited dependent variable models see Finlay and Magnusson (2009). The test also rejects that the estimated coefficient for the endogenous off-farm income is zero.

**Table 3.11:** Weak-Instrument-Robust Tests

| Test | H0 | Test Statistic | P-value | 95% Confidence Set |
|---|---|---|---|---|
| CLR | beta=b0 | 234.92 | 0.000 | [-0.024151, -0.018043] |
| AR | beta=b0 | 235.35 | 0.000 | [-0.024948, -0.017379] |
| LM | beta=b0 | 234.72 | 0.000 | [-0.024151, -0.018043] |
| J | E(Zu)=0 | 0.630 | 0.4289 | |
| LM-J | H0 rejected at 5% level | | | [-0.024417, -0.017911] |
| Wald | beta=b0 | 163.33 | 0.000 | [-0.024716, -0.018143] |

*Note: beta is coefficient on the endogenous regressor, $E(Z\ u) = 0$ indicate zero covariance is the exogeneity of the instruments where $Z$ are the instruments and $u$ is the disturbance in the structural equation*

*Source:* Author's calculations of the ALCS 2013-14 data

To investigate whether crop diversity is a multidimensional decision, and to analyse whether there is disparity in the effect of explanatory variables that influence household's choices of the extent of diversification we also estimated a Multinomial Logit Model (MNL) using a dichotomous or discrete choice variable classifying households in terms of the number of crops they grow. For this purpose, the household were classified into four discrete categories based on the number of crops they grow; non-diversifiers and diversifiers with 2, 3, and 4 or more crops. The MNL model carried out in this study passes the check of the Independence of Irrelevance Alternative (IIA) assumption for different categories of the discrete choice dependent variable. The MNL model permits the analysis of multivariate decision across more than two groups allowing the determination of choice probabilities for different categories of number of crops.

The MNL model used non-diversifiers (i.e. household who grow only one crop) as the base or reference category, therefore the estimated coefficient for each category of the diversification measures the change relative to non-diversifiers. Using the Maximum Likelihood estimation, the estimates of the MNL model are presented in Table 3.B3 in Appendix 3.B. The results of the MNL model indicated that decision by household to choose a particular level of crop diversification activity is not a multidimensional or multivariate decision, and there is no significant disparity in the effect of most of the explanatory variables on the extent of crop diversity, suggesting that the analysis could be reduced to a single decision process that can be analysed using a Tobit model. As a result, our results from the IV Tobit model are maintained.

### 3.9.1 Robustness Test on the Basis of Prevalence of Poppy Cultivation

As debated in Chapter II, cultivation of poppy crop is an important aspect of farming in the context of Afghanistan that may generate systematic differences in household characteristics and their management strategies (e.g. crop diversification) across the regions, especially since opium poppy cultivation is relatively more common in some zones or provinces than others.

Using information from Afghanistan's Ministry of Counter Narcotics (MCN) on major poppy producing areas, we divide our analytical sample into two categories based on the intensity of poppy cultivation at the provincial level: 1) Households in main opium producing provinces (mentioned in Chapter II) were assigned in one category, and 2) Households in other provinces that were opium free according to the MCN report published in 2013 were assigned to another category. Subsequently, we run our analysis for each category separately aiming to investigate the concern as to what extent crop diversification and other household's socioeconomic characteristics can differ between opium infected and opium free areas/provinces. The results estimated by both Tobit and Probit models for the two distinct categories are presented in Table 3.B3 and 3.B4 respectively.

Our results from the two sub-samples suggest no dramatic qualitative differences in the estimates among the two sub-groups, though there are some quantitative differences in the estimated parameters among the opium-infected and opium-free areas. The major disparities in the estimated parameters among two groups are: ownership of transport equipment and tractors or threshers by the household, and access to extension services are insignificant in the opium infected sub-sample, whereas they are significantly associated with CD in the opium free sub-sample.

Ownership of transport equipment and tractors/threshers by the households has no significant impact on CD for the opium sub-sample but significantly and positively associated with CD for the opium free sub-sample. This is possibly because household that engage in poppy cultivation don't use their own transport equipment to transport opium produce due to security and safety reasons. In addition, opium farming may require more skilled labour to conduct farming/harvesting activities manually than activities done by tractors or threshers. Access to extension services has no significant effect for opium sub-sample but a significant affect opium free areas. As debated in Chapter III, farm households in provinces where poppy is a common practice are not accessible by the extension agents due security concerns.

Moreover, crop diversification is significantly lower in major poppy producing agro-ecological zones (AEZ) such as Helmand Valley and Sistan Basin (HVSB), Heart-Farah Low Lands (HFL), and Central Mountains (CM), perhaps farm households specialize in opium production in these areas due to the extra income that poppy offers.

Off-farm income is consistently significant and negatively associated with the level of crop diversification in all models. The results from both sub-samples show that farm households that engage in off-farm income maintain significantly lower diversity in crop production. In general, there are no significant qualitative differences in the estimated parameters across the two sub-samples, therefore our main results presented earlier are uneffaced.

### 3.9.2 Robustness Test Based on Proximity to or Remoteness from Urban Centres

Proximity to or remoteness from the urban centres is another critical aspect in the context of this study that may alter crop diversification strategies by farmers. While the narrative central to the analysis presented in this chapter adopts the theory that marketization increases crop diversification, a concern may arise that market orientation may actually motivate farmers to engage in production of specialized crops as marketization may offer competitive advantages for certain agriculture commodities. Conversely, substance farmers in remote areas may engage in crop diversification, so to be able to meet their dietary requirement from own production since their access to markets is limited. At the meantime, consistent with the narrative of marketization- diversification presented in this chapter, if closeness to urbanization or marketization truly increases crop diversification, then it could be the case that farm households closed to urban centres are probably diversifying way more as compared to those located in remote areas with less access to markets, as a result there might be significant differences across households. In either "marketization-specialization" or "marketization-diversification" case, it is important to carry out a robustness check to ensure our main results are not driven by this spatial aspect of farming.

We therefore run a further robustness check and split our analytical sample into two sub-groups: farm households situated within 1 our 2 hours from or to the main urban centres[12] are assigned to one group[13] and farms located in remote areas (e.g. households not located within 1 or 2 hours from the main urban centres) are assigned to another group. Subsequently, we run our econometric models for each group separately to evaluate if there are systematic differences among the two groups. The results for these groups estimated by tobit and probit models are presented in Tables 3.B5 and 3.B6 respectively in the Appendix for Chapter III.

Overall, our analysis reveal no substantial qualitative dissimilarities for the estimated coefficients across the two sub-samples. The magnitude or size of estimated coefficients may vary between the two sub-sample but in general the results are similar to those of our main results presented earlier. Some of the major disparities in estimates between the two groups are: access to or ownership of communication equipment, cattle, and oxen

[12] Urban centres are all the provincial centre markets in 34 provinces. Initially, we intended to split the simple based on proximity to the five large commercial cities (Kabul, Kandahar, Herat, Jalalabad, and Mazar-i-Sharif, however this will substantially reduce the sub-sample (to about 75 households) not allowing enough observations or farm households that are actually engaged in farming.

[13] We also intended to include households that are situated within 1-2 hours to long-distance highways in this sub-group, we therefor explored the road shapefiles which allow us to place the location of the road and the province or districts that the roads go through, however the exact location of the farm household is unknown in the ALCS survey to allow the estimated distance between farm household and the major road(s). This effect is somewhat captured by dividing the sample based on the proximity to all major provincial markets, because most of highways are built to connect these provincial urban centres. This is another to consider splitting the sample by proximity to all 34 urban centres, instead of splitting it by proximity to the 5 large commercial cities.

by the household and land quality, and distance to roads (note: however, tractor ownership is significant in both models) are insignificant in the sub-sample that contains households located within 1-2 hours from or to the urban centres (referred to as the urban sub-sample from this point forward), but are significant in the model analysing the remote sub-sample (referred to the remote sub-sample from this point forward). The rest of the estimates are largely similar between the two groups.

Firstly, off-farm income, a variable of interest in this study, is constantly significant and associated negatively with the extent of crop diversification in all models. Even though the coefficient estimate is slightly higher in size (in absolute terms) for the remote sub-sample, the sign and significance remain the same in both groups. It is understandable for the remote sub-group that allocating more labour to off-farm activities may have larger impact, because for this group agriculture is mostly be their main activity for livelihood, whereas for the urban sub-sample off-farm income might be considered rather as a diversification strategy.

It is not shocking to see that ownership of communication equipment, cattle, and oxen by the households in the urban sub-sample are insignificant. Majority of them may live very close to urban centres or have good access to market information through other alternative channels, as a result ownership of communication equipment such as TV, radio, and mobile phones may not be the primary source of information for them. As for the cattle ownership and oxen ownership, it is possible that households close to urban areas may not keep significant livestock on the farm due lack of facilities and may not use them as their primary source of fertilization (manure) and traction power. Note that tractor ownership, however, is highly significant in both sub-samples which may indicate that households near urban areas are mainly using tractors for farming activities because they have less labour available to farming.

Having completed this robustness check, we are confident to say that our main results still remain unaffected and valid, as is evidenced by the similarity of the estimates among the two groups; proximity to or remoteness do not cause systematic differences among households in terms of their crop diversification strategies. Our main results are conditioning on distance to roads, time taken to reach to permeant food markets, and ownership of communication and transport equipment by the households which are sufficient to capture any variations associated with urbanization or marketization.

### 3.10 Conclusion and Discussion

Using a nationally representative survey from 8,613 households, we investigated the status, patterns, and determinants of the extent of diversity in crop production in Afghanistan with a particular interest in the impact of off-farm income on CD. Our results show that about a third of farmers do not diversify, and the majority that do, grow only two or three crops. The computed value of the diversity index measured by composite entropy index establishes the presence of a relatively level of crop diversity in Afghanistan which greatly varies across the 8 agro-ecological regions.

The results of the preferred IV-Tobit model revealed lower level of diversity in crop production for households with higher off-farm income. This is consistent with the hypothesis that allocation of farm labour away to non-farm activities decrease diversity due to negative labour effects, mainly because the opportunity cost of household labour is higher than the off-farm wages under imperfect markets implying non-separability between households' farm profits and off-farm earnings. Identification through instrumental variable techniques, reveal even greater impact of non-farm income on crop diversity suggesting that unobserved factors such as risk-aversion behaviour of farmers may drive household's decision towards diversification of both non-farm activities and crop diversification. Our finding of the negative impact of off-farm income on CD are consistent with Weiss and Briglauer, (2000), Mishra et al., (2004), and Hitayezu et al., (2016).

Other factors that significantly determine the intensity of crop diversity include household characteristics (sex and level of education of the household head, and household size), farm characteristics (land size, land quality, access to sufficient irrigation water, and landscape), transaction costs (proxied for by distance to market, nearest road, ownership of transport and communication equipment), ownership of livestock units and tractors, receipt of extension services, and regional factors. Among these factors, land, household ownership of transport and communication equipment, proximity to markets, ownership of cattle, oxen, and tractors, household size and household head education appear to have a positive significant impact on crop diversity. On the contrary, distance of farm households from roads and access to extension services were found to be negatively associated with the level of diversity in crop production. Surprisingly, low diversity is found for household with access to extension services, our intuition from this is that this might be due the emphasis of agricultural policies on the production of staple crops that are vitally important for food security.

Lastly, our robustness analysis of crop diversification based on the prevalence of poppy cultivation and proximity to or remoteness from urban centre reveal no significant and systematic differences between households across different locations. Despite some quantitative differences in estimates among the sub-samples (i.e. households in provinces with opium production vs household in provinces with no opium prevalence, and household located within 1-2 hours to urban centres vs remote households that are not located within 1-2 hours from urban centres). This in turn validate the accuracy of our main results.

This research is intended to contribute to the understanding of smallholder decision-making in relation to crop portfolio diversification and factors affecting it. It particularly has important implications for household's decisions about allocation of resources such as land and labour among on- and off-farm activities, especially since engagement of farm households in the non-farm activities reduces crop diversity. In general, smallholder resource-poor farmers are cautiously risk-averse and try to spread risk over a diverse profile of both on-and off-fam activities, particularly if farming business experiences high volatility. Policies associated with increasing opportunities for off-farm income do not contribute to crop diversification, therefore if crop diversification is the objective, policies must focus on farmers. Farmers that receive advice from extension agents appear to diversify less, thus it is may be viable to revisit the extension services programs if future policies aim to encourage crop diversification as a potential strategy for risk mitigation and income sustainability.

Crop diversification as an effective farm management strategy, can help small-scale farmers to mitigate potential risks associated with mono-cropping and reallocate productive resources away from low-value food grains towards high value horticultural crops to help improve and sustain household income. Policies that incentivise farmers' access to regional and international markets through better forward and backward linkages can ease the diversification process. Investment in rural infrastructure development such as roads, storage and transportation facilities, and other means to reduce transaction costs is an equally important aspect to stabilize supply chain and thereby ensure crop diversity.

### 3.10.1 Future research

Farmers access to credit is an important area that can have implications on decision making at the fam level, particularly the level of crop diversity. Lack of access to affordable financial micro-credit can constraint crop diversification process as it my increase the need for cash to purchase extra inputs such as seeds, agro-chemicals, labour, and other equipment for the cultivation and harvest. This research could be further extended by investigating empirical relationship between farmer's access to cheap credit or loans and crop intensity. In addition, including more precise indicators for market development and integration in the analysis carried out in this study could further assist to derive constructive policy implications for crop diversity and the transformation processor of agriculture towards commercialization.

Another potential area for the future research is to analyse the empirical implications of land fragmentation on farm households' decision-making process and crop intensification. As farm land size is considerably small in Afghanistan and is expected to further shrink over time due to rapid increase in population and urbanization, the implications of farm size can alter over time and the overall well-being of farming remains an important aspect that needs to be empirically addressed. This is especially of great interest as crop diversity significantly increases with farm size.

Another line of research could explore possible implications of the market conditions on farmers diversification decisions, especially farmers choices to grow market-oriented crops. Missing or imperfect and poorly functioning markets due to high transaction costs can reduce farmers chances to participate in local markets to sell their surplus produce and buy necessary inputs from the markets.

## APPENDIX TO CHAPTER III

**Appendix 3.A:** Detailed Descriptive Analysis

**Table 3.A1:** Characteristics of Farm Household by Access to Off-farm Income

| | No Off-farm Income | | Off-farm Income | | T-Test Mean Difference | |
|---|---|---|---|---|---|---|
| Characteristic | Mean | SD | Mean | SD | Difference | t-val |
| Total Land (Ha) | 2.030 | 6.16 | 1.290 | 2.42 | 0.74*** | -6.45 |
| Farm income (10K) | 8.230 | 9.13 | 2.380 | 4.42 | 5.85*** | -33.87 |
| THI (0≤THI≤1) | 0.300 | 0.23 | 0.270 | 0.24 | 0.02*** | -4.7 |
| CEI (0≤CEI≤1) | 0.312 | 0.22 | 0.280 | 0.24 | 0.03*** | -5.36 |
| Cattle ownership (N) | 1.550 | 2.32 | 1.430 | 1.68 | 0.11* | -2.42 |
| Oxen ownership (N) | 0.310 | 0.72 | 0.210 | 0.57 | 0.09*** | -6.11 |
| Tractor ownership (N) | 0.040 | 0.2 | 0.060 | 0.25 | -0.02*** | (-4.04) |
| Distance to road (10km) | 2.640 | 3.27 | 1.970 | 2.94 | 0.67*** | -9.57 |
| Opium share by province (%) | 0.060 | 0.13 | 0.010 | 0.06 | 0.04*** | -17.26 |
| Irrigation water (1=access) | 0.410 | 0.49 | 0.470 | 0.5 | -0.06*** | (-5.20) |
| Communication Equip. (1=access) | 0.740 | 0.44 | 0.830 | 0.37 | -0.10*** | (-10.24) |
| Transport Equip (1=access) | 0.470 | 0.5 | 0.440 | 0.5 | 0.03** | -2.9 |
| Extension Services (1=access) | 0.120 | 0.33 | 0.220 | 0.41 | -0.10*** | (-12.5) |
| Landscape (1=open plain) | 0.410 | 0.49 | 0.450 | 0.5 | -0.04*** | (-4.06) |
| N | 3,184 | | 5,429 | | 8,613 | |

*Source:* Author's calculations of the ALCS 2013-14

**Table 3.A2:** Crop Revenue Share and Growing Frequency

| Ranked by growing frequency | | Ranked by the share of revenue | |
|---|---|---|---|
| Crop | Growing Frequency | Crop | Revenue share (%) |
| Wheat | 7,961 | Wheat | 49.50 |
| Maize | 2,783 | Rice | 11.79 |
| Fodder | 1,564 | Maize | 11.42 |
| Potatoes | 1023 | Potatoes | 5.49 |
| Rice | 549 | Onions | 5.17 |
| Barley | 548 | Cotton | 3.01 |
| Beans | 419 | Melons | 2.70 |
| Onion | 377 | Fodder Crops | 2.46 |
| Other Vegetables | 224 | Beans | 1.76 |
| Tomatoes | 202 | Tomatoes | 1.58 |
| Millet | 179 | Other Vegetables | 1.49 |
| Sugar beet/cane | 128 | Barley | 1.46 |
| Melons | 121 | Okra | 0.64 |
| Cotton | 113 | Millet | 0.50 |
| Okra | 105 | Eggplant | 0.33 |
| Eggplant | 41 | Other Fruits | 0.23 |
| Courgette | 40 | Tree Fruits | 0.12 |
| Tree fruits | 13 | Sugar beet/cane | 0.12 |
| Cumin | 9 | Nuts | 0.10 |
| Flax | 8 | Cumin | 0.07 |
| Nuts | 7 | Flax | 0.04 |
| Other fruits | 7 | Courgette | 0.02 |
| N | 8,613 | | 8,613 |

*Source:* Author's calculations of the ALCS 2013-14 data

**Table 3.A3:** Single Croppers (Farmers who Grow one and Only one Crop)

| Crop | Growing frequency | Percent | % of total sample (N=8,864) |
|---|---|---|---|
| Wheat | 2,539 | 82.73 | 28.70 |
| Maize | 129 | 4.20 | 1.46 |
| Fodder crops | 125 | 4.07 | 1.41 |
| Potatoes | 93 | 3.02 | 1.05 |
| Barley | 40 | 1.30 | 0.45 |
| Other vegetables | 23 | 0.75 | 0.26 |
| Melons | 22 | 0.72 | 0.25 |
| Onions | 21 | 0.68 | 0.24 |
| Rice | 19 | 0.62 | 0.21 |
| Tomatoes | 14 | 0.46 | 0.16 |
| Millet | 13 | 0.42 | 0.15 |
| Beans | 12 | 0.39 | 0.14 |
| Cotton | 10 | 0.33 | 0.11 |
| Okra | 3 | 0.10 | 0.03 |
| Sugar beet/cane | 3 | 0.10 | 0.03 |
| Cumin | 2 | 0.07 | 0.02 |
| Nuts | 1 | 0.03 | 0.01 |
| Courgette | 1 | 0.03 | 0.01 |
| Flax | - | - | - |
| Eggplant | - | - | - |
| Tree fruits | - | - | - |
| Other fruits | - | - | - |
| Total | 3,070 | | |

*Note: Note total sample size is 8,853. Single croppers (farm household who grow one and only crop) except for farmers who grow basic staple crops such as wheat, maize, rice, and barley are excluded. This means that 240 observations are dropped, reducing the sample from 8,853 to 8,521 households. These farmers are assumed to be part-time farmers who are mainly involved in off-farm activities but growing garden crops. Source: Author's calculations of the ALCS 2013-14*

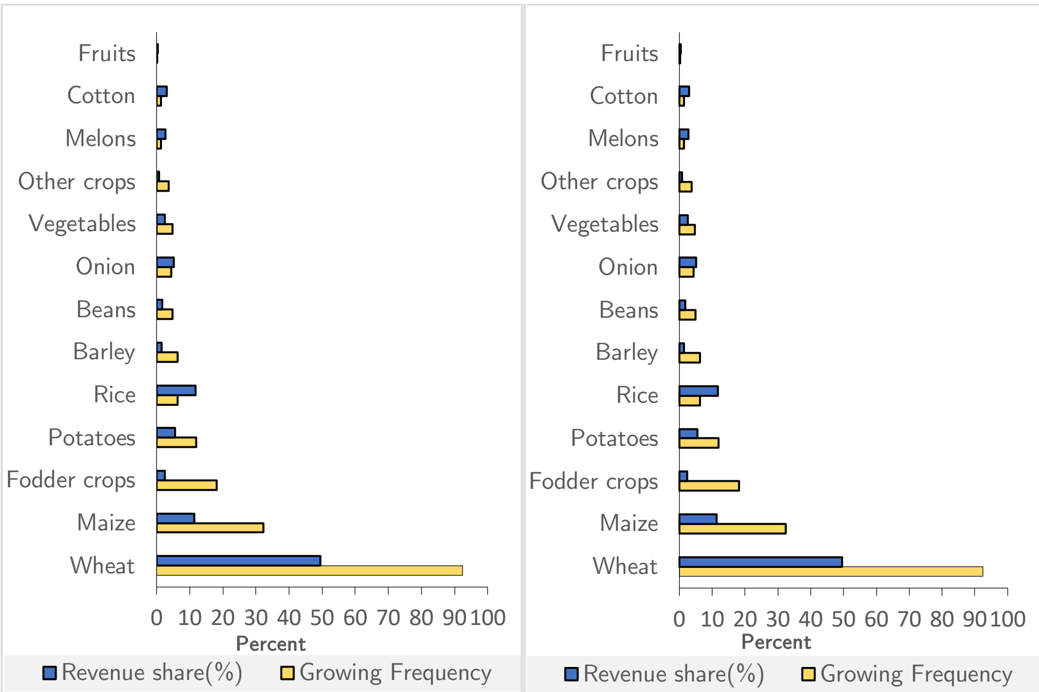



**Figure 3.A1:** a) Crop Revenue Share and Growing Frequency b) single croppers

*Source:* Author's calculations of the ALCS 2013-14

**Table 3.A4:** Characteristics of HH's that Grow one and Only one Crop (except for basic staple crops such as wheat, maize, rice, barley, potato, cotton, and onion) vs Households who Grow a Mix of Crops

| | Mix of Crops | | One and Only One crop | | T-test | |
|---|---|---|---|---|---|---|
| Variable | Mean | SD | Mean | SD | Difference | t |
| Total Land (Jeribs) | 1.560 | 4.23 | 0.66 | 1.11 | 0.90*** | -10.66 |
| Distance to Road (km) | 2.220 | 3.09 | 1.94 | 2.84 | 0.28 | -1.48 |
| Off-farm Income (10K AFN) | 5.520 | 11.1 | 8.08 | 9.88 | -2.56*** | (-3.95) |
| Farm Income (10K AFN) | 4.540 | 7.15 | 2.63 | 5.06 | 1.91*** | -5.70 |
| CDI (0≤THI≤1) | 0.280 | 0.23 | 0.00 | 0.00 | 0.28*** | -112.97 |
| CEI (0≤THI≤1) | 0.290 | 0.23 | 0.00 | 0.00 | 0.29*** | -117.60 |
| Cattle Ownership (Number) | 1.480 | 1.94 | 0.96 | 1.23 | 0.52*** | -6.33 |
| Oxen Ownership (Number) | 0.250 | 0.64 | 0.08 | 0.36 | 0.17*** | -6.88 |
| Tractor Ownership (Number) | 0.050 | 0.23 | 0.02 | 0.13 | 0.04*** | -4.13 |
| Extension Services (1=access) | 0.180 | 0.39 | 0.20 | 0.40 | -0.01 | (-0.47) |
| Transport Equip. (1=yes) | 0.450 | 0.50 | 0.38 | 0.49 | 0.07* | -2.10 |
| Communication Equip. (1=yes) | 0.800 | 0.40 | 0.87 | 0.34 | -0.07** | (-3.30) |
| Irrigation water (1=access) | 0.450 | 0.50 | 0.54 | 0.50 | -0.09** | (-2.74) |
| Landscape (Open Plan=1) | 0.440 | 0.50 | 0.52 | 0.50 | -0.08* | (-2.43) |
| Fertilizer Expenditures (AFN) | 4,630 | 8900 | 2,402 | 3,525 | 2,228*** | -10.20 |
| N | 8,613 | | 240 | | 8,853 | |

*Source:* Author's calculations of the ALCS 2013-14

**Table 3.A5:** T-test & Pearson $\chi 2$ Tests of Mean Difference Between Non-diversifiers & Diversifiers

| | Non-Diversifiers | | Diversifiers | | | T-Test |
|---|---|---|---|---|---|---|
| Variable | Mean | SD | Mean | SD | Difference | T-Value |
| Annual Revenue (AFN) | 28,639 | 54,631 | 79,795 | 137,597 | -51,155*** | (-24.5) |
| Total Land (Ha) | 1.14 | 2.22 | 1.77 | 4.91 | -0.63*** | (-8.3) |
| Off-farm Inc.(10K AFN) | 5.72 | 8.58 | 5.42 | 12.07 | 0.30 | -1.33 |
| Farm Inc.(10K AFN) | 2.9 | 4.63 | 5.34 | 7.98 | -2.44*** | (-17.8) |
| Own Cattle (N) | 1.18 | 1.69 | 1.62 | 2.04 | -0.44*** | (-10.6) |
| Oxen and Yak (N) | 0.2 | 0.53 | 0.27 | 0.68 | -0.07*** | (-5.1) |
| Tractors (N) | 0.02 | 0.14 | 0.07 | 0.26 | -0.05*** | (-10.5) |
| Distance Road (10 km) | 2.59 | 3.42 | 2.04 | 2.89 | 0.55*** | -7.42 |

Pearson chi2 test for categorical variables

| | | All | Specialized | Diversified | $\chi 2$ | P-val |
|---|---|---|---|---|---|---|
| Transport Equip. | No Access | 54.95 | 54.77 | 55.04 | 0.056 | 0.813 |
| | Access | 45.05 | 45.23 | 44.96 | | |
| Communication Equip. | No Access | 20.23 | 25.48 | 17.66 | 72.05 | 0.000 |
| | Access | 79.77 | 74.52 | 82.34 | | |
| Irrigation Water | No Access | 55.21 | 62.01 | 51.88 | 78.97 | 0.000 |
| | Access | 44.79 | 37.99 | 48.12 | | |
| Land Quality | Irrigated only | 24.66 | 31.24 | 21.44 | 98.12 | 0.000 |
| | Combined | 75.34 | 68.76 | 78.56 | | |
| Household Head Sex | Female | 0.45 | 0.78 | 0.29 | 9.851 | 0.002 |
| | Male | 99.55 | 99.22 | 99.71 | | |
| Landscape | Hills & Valleys | 56.29 | 64.45 | 52.29 | 114.2 | 0.000 |
| | Open Plain | 43.71 | 35.55 | 47.71 | | |
| Distance to Market | Not reachable | 4.35 | 5.76 | 3.67 | 26.41 | 0.000 |
| | Less than 1hr | 54.81 | 52.12 | 56.13 | | |
| | More than 1hr | 40.83 | 42.12 | 40.2 | | |
| Head Education | No School | 76.86 | 81.1 | 74.79 | 46.77 | 0.000 |
| | Primary | 11.59 | 10.18 | 12.28 | | |
| | Secondary | 7.92 | 5.8 | 8.96 | | |
| | T. College | 2.29 | 1.84 | 2.51 | | |
| | Uni & Postgrad | 1.35 | 1.1 | 1.47 | | |
| Extension Services | No Access | 81.64 | 84.13 | 80.43 | 17.45 | 0.000 |
| | Access | 18.36 | 15.87 | 19.57 | | |
| Agro-ecological zone | NEM | 2.33 | 4.28 | 1.38 | 419.6 | 0.000 |
| | CM | 16.6 | 18.98 | 15.44 | | |
| | HFL | 4.03 | 6.15 | 2.99 | | |
| | SMF | 19.84 | 11.52 | 23.91 | | |
| | HVSB | 10.47 | 12.79 | 9.34 | | |
| | TP | 6.79 | 8.87 | 5.78 | | |
| | NMF | 18.33 | 21.73 | 16.67 | | |
| | EMF | 21.6 | 15.69 | 24.49 | | |
| N | | 8,613 | 2,830 | 5,783 | | |

Source: Author's calculations of the ALCS 2013-14

**Table 3.A6:** Groups of Farm Household by the Number of Crops

| Number of Crops | Number of Farms | Percent |
|---|---|---|
| 1 | 2,830 | 32.86 |
| 2 | 4,110 | 47.65 |
| 3 | 1,410 | 16.37 |
| 4 or more | 269 | 3.12 |
| Mean | 1.91 | |
| Min | 1 | |
| Median | 2 | |
| Max | 6 | |
| N | 8,613 | |

*Source*: Author's calculations of the ALCS 2013-14

**Table 3.A7:** Cropping activity by HH across Agro-Ecological Zones (AEZ)

| AEZ | Grains | Fodder crops | Potato | Vegetables | Beans | Onions | Industrial crops | Melons | Fruits | Nuts |
|---|---|---|---|---|---|---|---|---|---|---|
| NEM | 275 | 6 | 17 | 0 | 8 | 14 | 0 | 2 | 0 | 0 |
| CM | 1,690 | 358 | 569 | 16 | 38 | 9 | 5 | 2 | 2 | 2 |
| HFL | 453 | 26 | 0 | 29 | 3 | 19 | 3 | 21 | 1 | 2 |
| SMF | 2,223 | 697 | 254 | 79 | 134 | 118 | 7 | 29 | 7 | 1 |
| HVSB | 1,367 | 5 | 5 | 14 | 1 | 2 | 49 | 20 | 2 | 0 |
| TP | 741 | 137 | 1 | 34 | 1 | 4 | 31 | 16 | 0 | 0 |
| NMF | 2,325 | 168 | 55 | 123 | 8 | 85 | 133 | 28 | 3 | 0 |
| EMF | 2,946 | 168 | 122 | 317 | 226 | 126 | 20 | 3 | 6 | 2 |
| Overall | 12,020 | 1,565 | 1,023 | 612 | 419 | 377 | 248 | 121 | 21 | 7 |

*Source:* Author's calculations of the ALCS 2013-14

**Table 3.A8:** Summary Statistics of Variables by Number of Crops

| Variable | All Farmers | | 1 crop | | 2 Crops | | 3 Crops | | 4 or more crops | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| CEI (0≤CEI≤1) | 0.298 | 0.232 | - | - | 0.401 | 0.113 | 0.515 | 0.121 | 0.621 | 0.112 |
| THI (0≤CEI≤1) | 0.286 | 0.232 | - | - | 0.379 | 0.131 | 0.506 | 0.133 | 0.631 | 0.117 |
| Total Land (Ha) | 1.576 | 4.248 | 1.164 | 2.256 | 1.704 | 5.318 | 1.814 | 2.716 | 2.561 | 6.803 |
| Household Head Sex (0=F, 1=M) | 0.995 | 0.068 | 0.992 | 0.089 | 0.997 | 0.058 | 0.998 | 0.046 | 1.000 | - |
| Household Head Age (Years) | 44.16 | 13.912 | 43.74 | 14.12 | 44.102 | 13.9 | 44.92 | 13.62 | 45.33 | 13.56 |
| Household Head Age Square (Years) | 2,144 | 1,333 | 2,112 | 1,362 | 2,138 | 1,322 | 2,203 | 1,318 | 2,238 | 1,267 |
| No Formal Schooling (1=Yes, 0 Otherwise) | 0.771 | 0.420 | 0.818 | 0.386 | 0.750 | 0.433 | 0.760 | 0.427 | 0.654 | 0.476 |
| Primary & Lower Secondary (1=Yes, 0=Otherwise) | 0.115 | 0.319 | 0.100 | 0.300 | 0.118 | 0.322 | 0.134 | 0.341 | 0.138 | 0.345 |
| Upper Secondary (1=yes, 0=Otherwise) | 0.078 | 0.268 | 0.054 | 0.226 | 0.091 | 0.287 | 0.077 | 0.266 | 0.138 | 0.345 |
| Teacher College (1=yes, 0=Otherwise) | 0.022 | 0.148 | 0.017 | 0.129 | 0.026 | 0.160 | 0.016 | 0.124 | 0.056 | 0.230 |
| University & Postgrad (1=yes, 0=Otherwise) | 0.013 | 0.115 | 0.011 | 0.104 | 0.015 | 0.121 | 0.014 | 0.118 | 0.015 | 0.121 |
| Household Size (Persons) | 8.135 | 3.483 | 7.568 | 3.305 | 8.362 | 3.599 | 8.478 | 3.283 | 8.632 | 3.755 |
| Quality of Land (1=High, 0=Low) | 0.749 | 0.433 | 0.675 | 0.469 | 0.799 | 0.401 | 0.760 | 0.427 | 0.699 | 0.460 |
| Cattle Ownership (Heads) | 1.484 | 1.950 | 1.194 | 1.711 | 1.518 | 1.838 | 1.860 | 2.557 | 1.955 | 1.757 |
| Oxen Ownership (Number) | 0.250 | 0.638 | 0.206 | 0.533 | 0.214 | 0.571 | 0.418 | 0.895 | 0.372 | 0.803 |
| Tractor Ownership (Number) | 0.053 | 0.232 | 0.022 | 0.145 | 0.068 | 0.260 | 0.070 | 0.279 | 0.041 | 0.198 |
| Access to Info Equipment (0=No, 1=Yes) | 0.797 | 0.403 | 0.740 | 0.439 | 0.817 | 0.387 | 0.838 | 0.368 | 0.848 | 0.360 |
| Own Transport Equipment (0=No, 1=Yes) | 0.452 | 0.498 | 0.456 | 0.498 | 0.427 | 0.495 | 0.523 | 0.500 | 0.420 | 0.494 |
| Access to Irrigation (0=No, 1=Yes) | 0.446 | 0.497 | 0.373 | 0.484 | 0.481 | 0.500 | 0.452 | 0.498 | 0.628 | 0.484 |
| Landscape 1: Valleys & Hills (1=Yes, 0=Otherwise) | 0.562 | 0.496 | 0.642 | 0.479 | 0.497 | 0.500 | 0.617 | 0.486 | 0.439 | 0.497 |
| Landscape 2: Open Plain (1=Yes, 0=Otherwise) | 0.438 | 0.496 | 0.358 | 0.479 | 0.503 | 0.500 | 0.383 | 0.486 | 0.561 | 0.497 |
| Total Off-Farm Income (10,000 AFN) | 5.511 | 11.088 | 5.714 | 8.654 | 5.650 | 11.382 | 4.974 | 14.610 | 4.127 | 5.726 |
| Extension Services (1=Access, 0=Otherwise) | 0.182 | 0.386 | 0.155 | 0.362 | 0.212 | 0.409 | 0.158 | 0.365 | 0.145 | 0.353 |
| Distance to Market (1=Not Reachable, 0 Otherwise) | 0.043 | 0.202 | 0.056 | 0.230 | 0.038 | 0.191 | 0.035 | 0.183 | 0.026 | 0.159 |
| Distance to Market (1=Less than 1h, 0 Otherwise) | 0.548 | 0.498 | 0.519 | 0.500 | 0.566 | 0.496 | 0.523 | 0.500 | 0.695 | 0.461 |
| Distance to Market (1=More than 1h, 0 Otherwise) | 0.410 | 0.492 | 0.425 | 0.494 | 0.396 | 0.489 | 0.443 | 0.497 | 0.279 | 0.449 |
| Distance to Nearest Road (10 km) | 2.220 | 3.089 | 2.608 | 3.438 | 2.045 | 2.895 | 2.167 | 2.990 | 1.243 | 2.117 |
| N | 8,613 | | 2,830 | | 4,104 | | 1,410 | | 269 | |

*Source:* Author's calculations of the ALCS 2013-14

**Appendix 3.B:** Robustness Analysis

**Table 3.B1:** ME of Tobit and IV-Tobit Model using THI as a Dependent Variable

| Dependent variable: THI | Tobit | | IV-Tobit | |
|---|---|---|---|---|
| | ME | SE | ME | SE |
| Off-farm Income (10,000 AFN) | -0.002*** | 0.000 | -0.015*** | 0.001 |
| Total Land (Ha) | 0.004*** | 0.001 | 0.003*** | 0.001 |
| Transport Equipment (1=access) | 0.020*** | 0.006 | 0.030*** | 0.007 |
| Communication Equipment (1=access) | 0.015** | 0.007 | 0.025*** | 0.008 |
| Cattle Ownership (N) | 0.006*** | 0.001 | 0.003** | 0.002 |
| Oxen & Yaks (N) | 0.036*** | 0.004 | 0.024*** | 0.005 |
| Tractor & Thresher (N) | 0.031*** | 0.011 | 0.046*** | 0.013 |
| Land Quality (1=good) | 0.049*** | 0.008 | 0.049*** | 0.009 |
| Landscape (1=open plain) | 0.052*** | 0.006 | 0.061*** | 0.007 |
| Sufficient Irrigation Water (1=access) | 0.024*** | 0.006 | 0.029*** | 0.006 |
| Household Size (persons) | 0.006*** | 0.001 | 0.013*** | 0.001 |
| HH Head Edu (1=primary & lower sec) | 0.014* | 0.008 | 0.039*** | 0.010 |
| HH Head Edu (2=upper secondary) | 0.027*** | 0.010 | 0.085*** | 0.013 |
| HH Head Edu (1=teacher college) | 0.0001 | 0.017 | 0.040* | 0.020 |
| HH Head Edu (1=university & graduate) | 0.015 | 0.023 | 0.123*** | 0.030 |
| HH Head Sex (1=male) | 0.097*** | 0.033 | 0.077** | 0.039 |
| HH Head Age (years) | 0.000 | 0.000 | 0.000 | 0.000 |
| Extension Services (1=access) | -0.012* | 0.007 | -0.016** | 0.008 |
| Distance to Nearest Road (10 km) | -0.004*** | 0.001 | -0.005*** | 0.001 |
| Distance to Market (1=less than 1 hr) | 0.007 | 0.013 | 0.029** | 0.015 |
| Distance to Market (2=more than 1 hr) | 0.023* | 0.013 | 0.015 | 0.014 |
| Agro-Ecological Zone 1 (NEM) | 0.074*** | 0.016 | 0.036* | 0.019 |
| Agro-Ecological Zone 2 (CM) | -0.004 | 0.019 | -0.025 | 0.022 |
| Agro-Ecological Zone 3 (HFL) | 0.130*** | 0.017 | 0.120*** | 0.020 |
| Agro-Ecological Zone 4 (SMF) | 0.038** | 0.017 | -0.046** | 0.021 |
| Agro-Ecological Zone 5 (TP) | -0.028* | 0.017 | -0.065*** | 0.021 |
| Agro-Ecological Zone 6 (NMF) | 0.097*** | 0.016 | 0.061*** | 0.020 |
| Agro-Ecological Zone 7 (EMF) | 0.182*** | 0.017 | 0.161*** | 0.020 |
| Log-Likelihood | -3,900.48 | | -35,874.70 | |
| Wald Test of exogeneity (chi2, p-value) | - | - | 168.73*** | 0.000 |
| Amemiya-Lee-Newey statistic (chi2, p-value) | - | - | 0.230 | 0.629 |
| Left censored observations(N) | 2,830 | | 2,830 | |
| Uncensored observations (N) | 5,782 | | 5,782 | |
| N | 8,613 | | 8,613 | |

*Note: significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010. Marginal Effects for factor levels is the discrete change from the base level. The omitted categories are: no access for transport equipment and communication equipment, irrigated & rainfed combined for land quality, hills and valleys for landscape, no access to sufficient irrigation water for irrigation, no formal schooling for education, no access for extension services, female for HH head sex, not reachable for distance to market, and agro-ecological zone 8 for AEZ.*

**Table 3.B2:** Maximum Likelihood Estimation of the MNL Choice Model

| | Two Crops | | Three Crops | | Four or more crops | |
|---|---|---|---|---|---|---|
| Variable | b | se | b | se | b | se |
| Off-farm Income (in 10K AFN) | -0.016*** | 0.003 | -0.042*** | 0.005 | -0.094*** | 0.014 |
| Total Land (Ha) | 0.198*** | 0.019 | 0.215*** | 0.020 | 0.222*** | 0.021 |
| Transport Equip (1=access) | 0.083 | 0.060 | 0.220*** | 0.081 | 0.143 | 0.166 |
| Communication Equip (1=own | -0.054 | 0.068 | 0.282*** | 0.097 | 0.281 | 0.195 |
| Cattle Ownership (N) | 0.026 | 0.016 | 0.095*** | 0.019 | 0.071** | 0.032 |
| Oxen & Yaks (N) | 0.216*** | 0.051 | 0.487*** | 0.058 | 0.358*** | 0.104 |
| Tractor//Threshers (N) | 0.566*** | 0.153 | 0.666*** | 0.181 | 0.940*** | 0.333 |
| Land quality(1=irrigated & rain fed combined) | 0.491*** | 0.081 | 0.491*** | 0.106 | -0.308 | 0.204 |
| Landscape (1=open plain) | 0.595*** | 0.066 | 0.494*** | 0.087 | 0.862*** | 0.163 |
| Sufficient Irrigation (1=access) | 0.165*** | 0.056 | 0.095 | 0.076 | 0.864*** | 0.152 |
| Household Size (persons) | 0.048*** | 0.009 | 0.070*** | 0.012 | 0.091*** | 0.021 |
| Head Edu (1=primary) | 0.120 | 0.086 | 0.250** | 0.110 | 0.313 | 0.205 |
| Head Edu (2=secondary) | 0.399*** | 0.107 | 0.234* | 0.142 | 0.810*** | 0.221 |
| Head Edu (3=teacher college) | 0.189 | 0.181 | -0.359 | 0.268 | 0.753** | 0.333 |
| Head Edu (4=uni & postgrad) | 0.116 | 0.237 | 0.266 | 0.308 | 0.042 | 0.570 |
| Head Age (years) | -0.008 | 0.011 | -0.002 | 0.015 | 0.026 | 0.029 |
| Head Age Squared | 0.000 | 0.000 | 0.000 | 0.000 | -0.000 | 0.000 |
| Extension Services (1=access) | 0.098 | 0.070 | -0.338*** | 0.098 | -0.668*** | 0.195 |
| Distance to Nearest Rd (10 km) | -0.009 | 0.009 | -0.022* | 0.012 | -0.115*** | 0.033 |
| Distance to Market (<1 hr) | -0.139 | 0.125 | 0.009 | 0.186 | 0.361 | 0.413 |
| Time to Market (> 1hr) | 0.129 | 0.123 | 0.330* | 0.183 | 0.235 | 0.412 |
| Agro-Ecological Region 1 (CM) | 0.987*** | 0.189 | 0.946*** | 0.238 | 0.110 | 0.474 |
| Agro-Ecological Zone 2 (HFL) | 0.292 | 0.219 | -0.670** | 0.319 | -0.755 | 0.619 |
| Agro-Ecological Zone 3 (SMF) | 1.705*** | 0.195 | 1.735*** | 0.246 | -0.366 | 0.526 |
| Agro-Ecological Zone 4 (HVSB) | 0.581*** | 0.203 | -1.925*** | 0.327 | -16.339 | 451.209 |
| Agro-Ecological Zone 5 (TP) | 0.254 | 0.210 | -1.307*** | 0.308 | -2.141*** | 0.644 |
| Agro-Ecological Zone 6 (NMF) | 1.060*** | 0.191 | 0.552** | 0.245 | 0.535 | 0.461 |
| Agro-Ecological Region 7 (EMF) | 1.503*** | 0.193 | 1.338*** | 0.246 | 1.582*** | 0.460 |
| Constant | -1.834*** | 0.315 | -3.309*** | 0.432 | -5.069*** | 0.866 |
| Base Outcome | One Crop | | | | | |
| Log-Likelihood | -8,728.57 | | | | | |
| chi2 | 1,915.298 | | | | | |
| p | 0.000 | | | | | |
| Pseudo R2 | 0.099 | | | | | |
| N | 8,624 | | | | | |

*Note: significance levels by * p<0.10, ** p<0.05, *** p<0.010 Marginal Effects for factor levels is the discrete change from the base level. The omitted categories are: no access to transport equipment and communication equipment, irrigated & rainfed combined for land quality, hills and valleys for landscape, no access to sufficient irrigation water, no formal education, none for extension services, not reachable for distance to market, and agro-ecological zone 8 for AEZ,*

**Table 3.B3:** Unconditional ME's from Tobit & IV-Tobit Models for Sub-samples Based on the Prevalence of Opium Production at the Province Level

| *Dependent variable: CEI ($0 \leq CE \leq 1$)* | **Sub-sample: provinces with significant opium** | | | | | | **Sub-sample: provinces with no opium production** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Tobit | | Instrumental Variable Tobit | | | | Tobit | | Instrumental Variable Tobit | | | |
| | | | 1st stage | | 2nd stage | | | | 1st stage | | 2nd stage | |
| Variable | ME | se | Coefficient | se | ME | se | ME | se | Coefficient | se | ME | se |
| Off-farm Income (10K AFN) | -0.006*** | 0.001 | - | - | -0.031*** | 0.004 | -0.001** | 0.001 | - | - | -0.008*** | 0.001 |
| Total Land (Ha) | 0.018*** | 0.003 | 0.049 | 0.130 | 0.018*** | 0.005 | 0.004*** | 0.001 | -0.063*** | 0.019 | 0.003** | 0.001 |
| Transport Equip. (1=own) | 0.015 | 0.014 | -0.103 | 0.284 | 0.012 | 0.015 | 0.006 | 0.007 | 1.025*** | 0.304 | 0.014* | 0.007 |
| Communication Equip. (1=own) | -0.022 | 0.014 | 0.677*** | 0.222 | -0.006 | 0.014 | 0.032*** | 0.009 | 0.012 | 0.328 | 0.038*** | 0.009 |
| Cattle Ownership (N) | 0.004 | 0.004 | 0.003 | 0.091 | 0.005 | 0.004 | 0.007*** | 0.001 | -0.143** | 0.064 | 0.006*** | 0.001 |
| Oxen & Yaks (N) | 0.031** | 0.013 | -1.030*** | 0.265 | 0.011 | 0.014 | 0.040*** | 0.005 | -0.706*** | 0.170 | 0.032*** | 0.005 |
| Tractor & Thresher (N) | -0.015 | 0.025 | 1.733 | 1.285 | 0.027 | 0.042 | 0.027*** | 0.010 | 0.021 | 0.766 | 0.035*** | 0.011 |
| Land Quality (1=good) | 0.095*** | 0.028 | 0.106 | 1.163 | 0.097*** | 0.038 | 0.071*** | 0.009 | 0.141 | 0.300 | 0.072*** | 0.009 |
| Landscape (1=open plain) | 0.091*** | 0.016 | -0.379 | 0.366 | 0.083*** | 0.018 | 0.039*** | 0.007 | 0.796** | 0.312 | 0.045*** | 0.007 |
| Irrigation Water (1=access) | -0.010 | 0.012 | 0.495* | 0.281 | -0.001 | 0.014 | 0.035*** | 0.006 | 0.463 | 0.304 | 0.038*** | 0.007 |
| Household Size (persons) | 0.005*** | 0.002 | 0.400*** | 0.080 | 0.014*** | 0.003 | 0.007*** | 0.001 | 0.656*** | 0.204 | 0.012*** | 0.002 |
| Head Edu (1=primary& low sec) | 0.040** | 0.020 | 0.319 | 0.519 | 0.063** | 0.025 | -0.001 | 0.010 | 1.590*** | 0.461 | 0.015 | 0.010 |
| Head Edu (2=upper sec) | 0.074** | 0.029 | 1.809** | 0.717 | 0.151*** | 0.041 | 0.018* | 0.011 | 3.836*** | 1.153 | 0.051*** | 0.015 |
| Head Edu (1=teacher college) | 0.065 | 0.049 | 2.034 | 1.587 | 0.140** | 0.069 | -0.004 | 0.018 | 2.122*** | 0.514 | 0.017 | 0.018 |
| Head Edu (1=uni & grad) | -0.055 | 0.043 | 4.965*** | 1.672 | 0.089 | 0.068 | 0.028 | 0.025 | 6.943*** | 1.423 | 0.086*** | 0.030 |
| HH Head Sex (1=male) | 0.038 | 0.117 | 0.698 | 1.680 | 0.009 | 0.126 | 0.087** | 0.042 | -0.114 | 1.245 | 0.077* | 0.044 |
| HH Head Age (years) | -0.001** | 0.000 | -0.0001 | 0.009 | -0.0001 | 0.000 | 0.0001 | 0.000 | 0.006 | 0.014 | 0.001** | 0.000 |
| Extension Services (1=access) | 0.012 | 0.016 | 0.718 | 0.449 | 0.028 | 0.021 | -0.025*** | 0.008 | -1.745*** | 0.421 | -0.030*** | 0.008 |
| Distance to Road (10 km) | 0.001 | 0.001 | -0.019 | 0.015 | 0.001 | 0.001 | 0.000 | 0.000 | -0.022** | 0.010 | -0.000 | 0.000 |
| Time to Market (1= $\leq$ 1 hr) | 0.015 | 0.027 | 1.329*** | 0.360 | 0.067*** | 0.025 | 0.022 | 0.017 | 0.170 | 0.410 | 0.033** | 0.017 |
| Distance to Market (2=> 1hr) | -0.003 | 0.026 | 0.513 | 0.330 | 0.021 | 0.023 | 0.039** | 0.017 | -0.735* | 0.395 | 0.030* | 0.016 |
| Agro-Ecological Zone 1 (CM) | -0.121*** | 0.030 | -4.974*** | 1.051 | -0.255*** | 0.056 | 0.109*** | 0.033 | 0.961 | 0.656 | 0.100*** | 0.032 |
| Agro-Ecological Zone 2 (HFL) | -0.051 | 0.034 | -3.201*** | 1.173 | -0.231*** | 0.062 | 0.010 | 0.040 | 6.066 | 3.880 | 0.057 | 0.051 |

***Table 3.B.3 Continue***

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Agro-Ecological Zone 3 (SMF) | - | - | - | - | - | - | 0.117*** | 0.033 | 2.355*** | 0.671 | 0.123*** | 0.032 |
| Agro-Ecological Zone 4 (HVSB) | -0.009 | 0.034 | -3.048** | 1.185 | -0.218*** | 0.064 | - | - | - | - | - | - |
| Agro-Ecological Zone 5 (TP) | - | - | - | - | - | - | -0.036 | 0.034 | 0.670 | 0.689 | -0.043 | 0.032 |
| Agro-Ecological Zone 6 (NMF) | 0.153*** | 0.037 | -1.823** | 0.831 | 0.041 | 0.052 | 0.084** | 0.034 | 0.357 | 0.667 | 0.076** | 0.032 |
| Agro-Ecological Zone 7 (EMF) | 0.255*** | 0.039 | -3.431*** | 1.175 | 0.135** | 0.066 | 0.146*** | 0.033 | -0.141 | 0.767 | 0.149*** | 0.032 |
| IV1- Share of Off-farm Income in Total Income within District | - | - | 8.490*** | 0.709 | - | - | - | - | 11.516*** | 0.613 | - | - |
| IV2-Lag District Level Off-farm Income | - | - | 0.001** | 0.000 | - | - | - | - | 0.001** | 0.000 | - | - |
| Constant | - | - | -3.180 | 1.960 | - | - | - | - | | | - | - |
| Log-Likelihood | -1,220.58 | | - | | -8,263.77 | | -2,484.10 | | - | | -26,803.60 | |
| Pseudo R-Square | 0.183 | | - | - | - | - | 0.137 | | - | - | - | - |
| Wald Test exogeneity (chi2, p) | - | - | - | - | 74.74 *** | 0.000 | - | - | - | - | 35.38 *** | 0.000 |
| Left censored observations(N) | 1,021 | | - | - | 1,021 | | 1,809 | | - | - | 1,809 | |
| Uncensored observations (N) | 1,227 | | - | - | 1,227 | | 4,554 | | - | - | 4,554 | |
| N | 2,249 | | 2,249 | | 2,249 | | 6,364 | | 6,364 | | 6,364 | |

*Note: The omitted categories are: no access for transport equipment & communication equipment, rain-fed land & Irrigated combined for land quality, hills and valleys for landscape, no access to sufficient irrigation water for irrigation, no formal schooling for education, no access for extension services, female for head sex, not reachable for time to market, and AEZ 8. significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

**Table 3.B4:** Average ME's from Probit & IV-Probit Models for Sub-samples Based on the Prevalence of Opium Production at the Province Level

| Dependent variable: CEI *(0≤CEI≤1)* | Sub-sample: provinces with significant opium | | | | | | Sub-sample: provinces with no opium production | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Probit | | Instrumental Variable Probit | | | | Probit | | Instrumental Variable Probit | | | |
| | | | 1st stage | | 2nd stage | | | | 1st stage | | 2nd stage | |
| Variable | ME | se | Coefficient | se | ME | se | ME | se | Coefficient | se | ME | se |
| Off-farm Income (10K AFN) | -0.013*** | 0.003 | - | - | -0.129*** | 0.011 | -0.003** | 0.001 | - | - | -0.052*** | 0.007 |
| Total Land (Ha) | 0.079*** | 0.014 | 0.049 | 0.130 | 0.156*** | 0.034 | 0.021*** | 0.006 | -0.063*** | 0.019 | 0.050*** | 0.016 |
| Transport Equip (1=own) | 0.040 | 0.028 | -0.100 | 0.284 | 0.069 | 0.063 | -0.009 | 0.013 | 1.026*** | 0.304 | 0.025 | 0.038 |
| Communication Equip (1=own) | -0.046* | 0.028 | 0.680*** | 0.223 | -0.028 | 0.060 | 0.030* | 0.016 | 0.011 | 0.328 | 0.121*** | 0.046 |
| Cattle Ownership (N) | 0.003 | 0.009 | 0.003 | 0.091 | 0.013 | 0.017 | 0.012*** | 0.003 | -0.143** | 0.064 | 0.024** | 0.010 |
| Oxen & Yaks (N) | 0.027 | 0.027 | -1.031*** | 0.265 | -0.027 | 0.059 | 0.076*** | 0.011 | -0.707*** | 0.170 | 0.159*** | 0.034 |
| Tractor & Thresher (N) | -0.047 | 0.068 | 1.733 | 1.286 | 0.058 | 0.190 | 0.113*** | 0.033 | 0.023 | 0.765 | 0.357*** | 0.095 |
| Land Quality (1=good) | 0.165*** | 0.058 | 0.103 | 1.163 | 0.332* | 0.174 | 0.132*** | 0.019 | 0.140 | 0.300 | 0.349*** | 0.050 |
| Landscape (1=open plain) | 0.179*** | 0.030 | -0.373 | 0.366 | 0.341*** | 0.077 | 0.093*** | 0.014 | 0.793** | 0.310 | 0.291*** | 0.042 |
| Irrigation Water (1=access) | -0.029 | 0.026 | 0.496* | 0.281 | -0.020 | 0.060 | 0.054*** | 0.012 | 0.462 | 0.303 | 0.163*** | 0.035 |
| Household Size (persons) | 0.009** | 0.004 | 0.400*** | 0.080 | 0.057*** | 0.011 | 0.013*** | 0.002 | 0.656*** | 0.204 | 0.064*** | 0.011 |
| Head Edu (1=primary & low sec) | 0.069* | 0.042 | 0.314 | 0.519 | 0.238** | 0.105 | 0.005 | 0.018 | 1.590*** | 0.461 | 0.119** | 0.052 |
| Head Edu (2=upper sec) | 0.154*** | 0.055 | 1.807** | 0.717 | 0.609*** | 0.153 | 0.051** | 0.020 | 3.836*** | 1.153 | 0.342*** | 0.069 |
| Head Edu (1=teacher college) | 0.201* | 0.106 | 2.033 | 1.587 | 0.673*** | 0.257 | -0.000 | 0.036 | 2.121*** | 0.514 | 0.142 | 0.096 |
| HH Head Edu (1=uni & grad | -0.063 | 0.129 | 4.960*** | 1.671 | 0.436 | 0.267 | 0.034 | 0.045 | 6.942*** | 1.424 | 0.443*** | 0.152 |
| HH Head Sex (1=male) | 0.007 | 0.236 | 0.691 | 1.680 | -0.115 | 0.481 | 0.146* | 0.083 | -0.112 | 1.244 | 0.292 | 0.203 |
| HH Head Age (years) | -0.002* | 0.001 | -0.000 | 0.009 | -0.002 | 0.002 | 0.001 | 0.000 | 0.006 | 0.014 | 0.003** | 0.001 |
| Extension Services (1=access) | 0.053 | 0.036 | 0.712 | 0.449 | 0.165* | 0.088 | -0.031** | 0.015 | -1.742*** | 0.421 | -0.119*** | 0.041 |
| Distance to Road (10 km) | 0.002 | 0.002 | -0.018 | 0.015 | 0.003 | 0.004 | 0.001 | 0.001 | -0.022** | 0.010 | 0.000 | 0.002 |
| Time to Market (1=<1hr) | 0.020 | 0.052 | 1.333*** | 0.361 | 0.257** | 0.108 | 0.012 | 0.031 | 0.169 | 0.410 | 0.092 | 0.077 |
| Time to Market (1=>1hr) | -0.000 | 0.050 | 0.518 | 0.330 | 0.099 | 0.102 | 0.064** | 0.030 | -0.734* | 0.395 | 0.110 | 0.077 |
| Agro-Ecological Zone 1 (CM) | -0.189*** | 0.063 | -4.994*** | 1.050 | -0.769*** | 0.184 | 0.249*** | 0.081 | 0.957 | 0.656 | 0.548*** | 0.175 |
| Agro-Ecological Zone 2 (HFL) | -0.065 | 0.068 | -3.170*** | 1.172 | -0.831*** | 0.198 | 0.035 | 0.095 | 6.064 | 3.882 | 0.412 | 0.277 |

*Table 3.B4 Continue*

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Agro-Ecological Zone 3 (SMF) | - | - | - | - | - | - | 0.279*** | 0.081 | 2.349*** | 0.671 | 0.740*** | 0.178 |
| Agro-Ecological Zone 4 (HVSB) | 0.003 | 0.066 | -3.000** | 1.184 | -0.796*** | 0.197 | - | - | - | - | - | - |
| Agro-Ecological Zone 5 (TP) | - | - | - | - | - | - | -0.042 | 0.086 | 0.664 | 0.690 | -0.140 | 0.184 |
| Agro-Ecological Zone 6 (NMF) | 0.287*** | 0.053 | -1.805** | 0.830 | 0.221 | 0.163 | 0.155* | 0.083 | 0.352 | 0.667 | 0.308* | 0.178 |
| Agro-Ecological Zone 7 (EMF) | 0.363*** | 0.059 | -3.428*** | 1.175 | 0.407* | 0.212 | 0.230*** | 0.081 | -0.149 | 0.766 | 0.577*** | 0.177 |
| IV1- Share of Off-farm Income in Total Income within District | - | - | 8.579*** | 0.694 | - | - | - | - | 11.509*** | 0.614 | - | - |
| IV2-Lag District Level Off-farm Income | - | - | 0.001* | 0.000 | - | - | - | - | 0.001** | 0.000 | - | - |
| Constant | - | - | -3.217 | 1.958 | - | - | - | - | -7.811*** | 1.595 | - | - |
| Log-Likelihood | -1,306.60 | | - | | -8,353.65 | | -3,443.39 | | - | | -27,752.89 | |
| Wald (Chi2, p-value) | 327.75 (0.000) | | - | | 479.62 (0.000) | | 566.69 (0.000) | | - | | 625.91 (0.000) | |
| Pseudo R-Square | 0.1567 | | - | | - | | 0.094 | | - | | - | |
| Wald Test of exogeneity (chi2, p) | - | | - | | 56.58*** (0.000) | | - | | - | | 27.38*** (0.000) | |
| N | 2,249 | | 2,249 | | 2,249 | | 6,364 | | 6,364 | | 6,364 | |

*Note: The omitted categories are: no access for transport equipment, no access for communication equipment, rain-fed land alone for land quality, hills and valleys for landscape, no access to sufficient irrigation water for irrigation, no formal schooling for education ,no access for extension services, female for head sex, not reachable for distance to market, and agro-ecological zone 8. significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010.*

**Table 3.B5:** Unconditional ME's from Tobit & IV-Tobit Models for Sub-samples Based on Proximity to Commercial Urban Centres

| *Dependent variable: CEI (0≤CEI≤1)* | **Sub-sample: HH's in communities within 1-2 hours to urban centres (provincial capitals)** | | | | | | **Sub-sample: HH's in communities in remote areas (*NOT* in communities within 1-2 hrs to urban centres )** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Tobit | | Instrumental Variable Tobit | | | | Tobit | | Instrumental Variable Tobit | | | |
| | | | 1st stage | | 2nd stage | | | | 1st stage | | 2nd stage | |
| Variable | ME | se | Coefficient | se | ME | se | ME | se | Coefficient | se | ME | se |
| Off-farm Income (10,000 AFN) | -0.0002 | 0.000 | - | - | -0.013*** | 0.004 | -0.003*** | 0.000 | - | - | -0.017*** | 0.001 |
| Total Land (Ha) | 0.006*** | 0.002 | -0.003 | 0.078 | 0.005** | 0.002 | 0.004** | 0.002 | -0.059*** | 0.018 | 0.003** | 0.001 |
| Transport Equip. (1=own) | 0.029 | 0.020 | 0.983 | 1.371 | 0.051** | 0.025 | 0.022*** | 0.007 | 0.898*** | 0.227 | 0.032*** | 0.007 |
| Communication Equip. (1=own) | -0.034 | 0.024 | -1.282 | 1.540 | -0.030 | 0.026 | 0.024*** | 0.008 | 0.585*** | 0.159 | 0.037*** | 0.008 |
| Cattle Ownership (N) | 0.004 | 0.006 | -0.457 | 0.325 | -0.004 | 0.004 | 0.006*** | 0.001 | -0.107*** | 0.041 | 0.004*** | 0.001 |
| Oxen & Yaks (N) | 0.019 | 0.015 | -1.616* | 0.908 | 0.002 | 0.016 | 0.039*** | 0.005 | -0.597*** | 0.124 | 0.028*** | 0.005 |
| Tractor & Thresher (N) | -0.015 | 0.031 | -1.115 | 4.062 | -0.011 | 0.055 | 0.040*** | 0.010 | 0.765 | 0.478 | 0.062*** | 0.012 |
| Land Quality (1=irrigated) | 0.005 | 0.028 | 0.262 | 1.120 | 0.030 | 0.027 | 0.051*** | 0.010 | -0.313 | 0.306 | 0.053*** | 0.010 |
| Landscape (1=open plain) | 0.181*** | 0.021 | -0.375 | 0.964 | 0.159*** | 0.021 | 0.047*** | 0.007 | 0.805*** | 0.241 | 0.057*** | 0.008 |
| Irrigation Water (1=access) | 0.047*** | 0.017 | 0.779 | 1.225 | 0.042** | 0.021 | 0.021*** | 0.006 | 0.554*** | 0.180 | 0.029*** | 0.007 |
| Household Size (persons) | 0.007*** | 0.002 | 1.549* | 0.829 | 0.023** | 0.010 | 0.006*** | 0.001 | 0.400*** | 0.044 | 0.012*** | 0.001 |
| Head Edu (1=primary & low sec) | 0.011 | 0.025 | 2.961* | 1.625 | 0.040 | 0.030 | 0.021** | 0.009 | 1.284*** | 0.376 | 0.045*** | 0.011 |
| Head Edu (2=upper sec) | 0.029 | 0.023 | 12.360** | 5.863 | 0.197*** | 0.064 | 0.033*** | 0.011 | 2.068*** | 0.364 | 0.069*** | 0.013 |
| Head Edu (1=teacher college) | 0.032 | 0.032 | 2.656 | 1.648 | 0.060* | 0.035 | -0.0001 | 0.018 | 2.341*** | 0.509 | 0.043** | 0.020 |
| Head Edu (1=uni & grad) | -0.062 | 0.050 | 10.621** | 4.382 | 0.093 | 0.087 | 0.042* | 0.025 | 6.259*** | 1.302 | 0.143*** | 0.037 |
| HH Head Sex (1=male) | 0.223*** | 0.060 | -1.339 | 2.867 | 0.183** | 0.090 | 0.085** | 0.039 | 0.031 | 1.017 | 0.068 | 0.043 |
| HH Head Age (years) | -0.000 | 0.001 | -0.034 | 0.066 | -0.0001 | 0.001 | -0.0001 | 0.000 | 0.008 | 0.006 | 0.000 | 0.000 |
| Extension Services (1=access) | -0.039* | 0.020 | -3.787*** | 1.209 | -0.077*** | 0.021 | -0.002 | 0.007 | -0.607** | 0.261 | -0.0001 | 0.008 |
| Distance to Road (10 km) | -0.005*** | 0.001 | -0.070*** | 0.024 | -0.006*** | 0.001 | 0.001** | 0.000 | -0.015* | 0.008 | 0.0002 | 0.000 |
| Time to Market 1 (≤ 1 hr) | - | - | - | - | - | - | 0.019 | 0.015 | 0.402 | 0.303 | 0.038*** | 0.015 |
| Time to Market 2 (> 1 hr) | - | - | - | - | - | - | 0.027* | 0.015 | -0.207 | 0.276 | 0.020 | 0.014 |
| Agro-Ecological Zone 1 (CM) | -0.075 | 0.122 | 6.536 | 5.037 | -0.106 | 0.117 | 0.086*** | 0.020 | -1.411** | 0.602 | 0.048** | 0.022 |

***Table 3.B.5 Continue***

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Agro-Ecological Zone 2 (HFL) | -0.135 | 0.128 | 14.462 | 13.692 | -0.004 | 0.196 | -0.002 | 0.023 | -0.452 | 0.722 | -0.037 | 0.026 |
| Agro-Ecological Zone 3 (SMF) | -0.070 | 0.123 | 8.496** | 4.103 | 0.035 | 0.117 | 0.151*** | 0.020 | 0.617 | 0.661 | 0.137*** | 0.022 |
| Agro-Ecological Zone 4 (HVSB) | -0.101 | 0.124 | 6.900 | 5.704 | -0.164 | 0.121 | 0.040* | 0.022 | -1.111 | 0.690 | -0.040 | 0.024 |
| Agro-Ecological Zone 5 (TP) | -0.317*** | 0.122 | 8.160* | 4.602 | -0.292** | 0.117 | -0.006 | 0.021 | -1.060 | 0.653 | -0.045* | 0.023 |
| Agro-Ecological Zone 6 (NMF) | 0.004 | 0.123 | 6.880 | 4.931 | -0.037 | 0.118 | 0.095*** | 0.021 | -1.204** | 0.611 | 0.062*** | 0.023 |
| Agro-Ecological Zone 7 (EMF) | 0.018 | 0.123 | 1.609 | 2.868 | 0.015 | 0.114 | 0.194*** | 0.021 | -0.846 | 0.659 | 0.182*** | 0.023 |
| IV1- Share of Off-farm Income in Total Income within District | - | - | 18.996*** | 5.035 | - | - | - | - | 9.540*** | 0.342 | - | - |
| IV2-Lag District Level Off-farm Income | - | - | 0.002** | 0.001 | - | - | - | - | 0.001*** | 0.000 | - | - |
| Constant | - | - | -0.370 | 0.231 | - | - | - | - | -0.244*** | 0.076 | - | - |
| Log-Likelihood | -352.29 | | - | | -4,773.82 | | -3,541.74 | | - | | -29,236.25 | |
| Pseudo R-Square | 0.303 | | - | | - | | 0.1188 | | - | | - | |
| Wald Test of exogeneity (chi2, p) | - | | - | | 42.87***(0.000) | | - | | - | | 108.03***(0.000) | |
| Left censored observations(N) | 312 | | - | - | 312 | | 2,520 | | - | - | 2,520 | |
| Uncensored observations (N) | 684 | | - | - | 684 | | 5,097 | | - | - | 5,097 | |
| N | 996 | | 996 | | 996 | | 7,617 | | | | 7,617 | |

*Note: The omitted categories are: no access for transport equipment, no access for communication equipment, rain-fed land alone for land quality, hills and valleys for landscape, no access to sufficient irrigation water for irrigation, no formal schooling for education ,no access for extension services, female for head sex, not reachable for distance to market, and agro-ecological zone 8. significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

**Table 3.B6:** Average ME's from Probit & IV-Probit Models for Sub-samples Based on Proximity to Commercial Urban Centres

| *Dependent variable: CEI (0≤CEI≤1)* | **Sub-sample: HH's in communities within 1-2 hours to urban centres (provincial capitals)** | | | | | | **Sub-sample: HH's in communities in remote areas (*NOT* in communities within 1-2 hrs to urban centres )** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Tobit | | Instrumental Variable Tobit | | | | Tobit | | Instrumental Variable Tobit | | | |
| | | | 1st stage | | 2nd stage | | | | 1st stage | | 2nd stage | |
| Variable | ME | se | Coefficient | se | ME | se | ME | se | Coefficient | se | ME | se |
| Off-farm Income (10K AFN) | -0.001 | 0.001 | - | - | -0.040*** | 0.012 | -0.006*** | 0.001 | - | - | -0.083*** | 0.006 |
| Total Land (Ha) | 0.031*** | 0.012 | -0.003 | 0.078 | 0.052*** | 0.019 | 0.028*** | 0.007 | -0.059*** | 0.018 | 0.059*** | 0.018 |
| Transport Equip. (1=own) | 0.037 | 0.037 | 0.997 | 1.371 | 0.148* | 0.079 | 0.027** | 0.013 | 0.898*** | 0.227 | 0.116*** | 0.035 |
| Communication Equip. (1=own) | -0.063 | 0.042 | -1.293 | 1.541 | -0.113 | 0.088 | 0.020 | 0.015 | 0.586*** | 0.159 | 0.115*** | 0.037 |
| Cattle Ownership (N) | 0.002 | 0.012 | -0.465 | 0.326 | -0.016 | 0.015 | 0.011*** | 0.003 | -0.107*** | 0.041 | 0.018** | 0.009 |
| Oxen & Yaks (N) | 0.033 | 0.037 | -1.617* | 0.909 | 0.012 | 0.063 | 0.067*** | 0.011 | -0.597*** | 0.124 | 0.119*** | 0.029 |
| Tractor & Thresher (N) | 0.000 | 0.078 | -1.097 | 4.061 | 0.003 | 0.187 | 0.137*** | 0.032 | 0.764 | 0.478 | 0.429*** | 0.084 |
| Land Quality (1=good) | 0.062 | 0.054 | 0.225 | 1.117 | 0.202** | 0.090 | 0.093*** | 0.019 | -0.313 | 0.306 | 0.240*** | 0.048 |
| Landscape (1=open plain) | 0.349*** | 0.046 | -0.330 | 0.965 | 0.527*** | 0.099 | 0.103*** | 0.013 | 0.805*** | 0.241 | 0.316*** | 0.037 |
| Sufficient Water (1=access) | 0.110*** | 0.033 | 0.780 | 1.224 | 0.169** | 0.071 | 0.031*** | 0.012 | 0.554*** | 0.180 | 0.107*** | 0.033 |
| Household Size (persons) | 0.016*** | 0.005 | 1.547* | 0.828 | 0.077*** | 0.029 | 0.010*** | 0.002 | 0.400*** | 0.044 | 0.053*** | 0.006 |
| Head Edu (1=primary&low sec) | 0.011 | 0.049 | 2.994* | 1.627 | 0.132 | 0.104 | 0.043** | 0.017 | 1.284*** | 0.376 | 0.229*** | 0.051 |
| Head Edu (2=upper se) | 0.071 | 0.049 | 12.362** | 5.862 | 0.605*** | 0.141 | 0.077*** | 0.021 | 2.069*** | 0.364 | 0.381*** | 0.064 |
| Head Edu (1=teacher college) | 0.098 | 0.092 | 2.662 | 1.648 | 0.283 | 0.193 | 0.009 | 0.040 | 2.341*** | 0.509 | 0.242** | 0.099 |
| Head Edu (1=uni & grad) | -0.178 | 0.114 | 10.611** | 4.370 | 0.192 | 0.244 | 0.076 | 0.048 | 6.259*** | 1.302 | 0.666*** | 0.170 |
| Head Sex (1=male) | 0.512*** | 0.175 | -1.285 | 2.856 | 0.615* | 0.362 | 0.132 | 0.082 | 0.030 | 1.018 | 0.214 | 0.207 |
| Head Age (years) | -0.001 | 0.001 | -0.034 | 0.066 | -0.001 | 0.003 | -0.000 | 0.000 | 0.008 | 0.006 | 0.002* | 0.001 |
| Extension Services (1=access) | -0.057 | 0.043 | -3.816*** | 1.209 | -0.223*** | 0.074 | 0.015 | 0.015 | -0.607** | 0.261 | 0.035 | 0.042 |
| Distance to Road (10 km) | -0.007*** | 0.002 | -0.070*** | 0.024 | -0.012*** | 0.003 | 0.002** | 0.001 | -0.015* | 0.008 | 0.002 | 0.002 |
| Time to Market 1 (≤ 1 hr) | - | - | - | - | - | - | 0.003 | 0.027 | 0.402 | 0.303 | 0.099 | 0.065 |
| Time to Market 2 (> 1 hr) | - | - | - | - | - | - | 0.046* | 0.026 | -0.207 | 0.276 | 0.080 | 0.063 |
| Agro-Ecological Zone 1 (CM) | -0.009 | 0.131 | 6.678 | 5.029 | -0.154 | 0.278 | 0.240*** | 0.040 | -1.411** | 0.602 | 0.318*** | 0.102 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | ***Table 3.B6 Continue*** |
| Agro-Ecological Zone 2 (HFL) | -0.224 | 0.160 | 14.590 | 13.678 | 0.007 | 0.531 | 0.033 | 0.050 | -0.452 | 0.722 | -0.158 | 0.121 |
| Agro-Ecological Zone 3 (SMF) | -0.049 | 0.136 | 8.549** | 4.103 | 0.180 | 0.264 | 0.381*** | 0.041 | 0.618 | 0.661 | 0.841*** | 0.108 |
| Agro-Ecological Zone 4 (HVSB) | -0.188 | 0.139 | 7.197 | 5.667 | -0.555* | 0.304 | 0.068 | 0.047 | -1.111 | 0.690 | -0.336*** | 0.117 |
| Agro-Ecological Zone 5 (TP) | -0.718*** | 0.136 | 8.287* | 4.602 | -1.035*** | 0.370 | 0.059 | 0.049 | -1.059 | 0.653 | -0.100 | 0.115 |
| Agro-Ecological Zone 6 (NMF) | 0.082 | 0.130 | 7.112 | 4.912 | 0.032 | 0.274 | 0.219*** | 0.042 | -1.204** | 0.611 | 0.295*** | 0.103 |
| Agro-Ecological Zone 7 (EMF) | 0.036 | 0.133 | 1.675 | 2.869 | 0.066 | 0.260 | 0.341*** | 0.041 | -0.845 | 0.659 | 0.712*** | 0.104 |
| IV1- Share of Off-farm Income in Total Income within District | - | - | 19.517*** | 4.954 | - | - | - | - | 9.541*** | 0.343 | - | - |
| IV2-Lag District Level Off-farm Income | - | - | 0.002** | 0.001 | - | - | - | - | 0.001*** | 0.000 | - | - |
| Constant | - | - | -21.329*** | 7.873 | - | - | - | - | -4.156*** | 1.239 | - | - |
| Log-Likelihood | -493.58 | | - | | -4,921.87 | | -4,362.63 | | - | | -30,050.59 | |
| Wald (Chi2, p) | 227.12 (0.000) | | - | | 250.13 (0.000) | | 728.38 (0.000) | | - | | 1029.72 (0.000) | |
| Pseudo R-Square | 0.210 | | - | | - | | 0.0962 | | | | - | |
| Wald Test exogeneity (chi2, p) | - | | - | | 65.46*** (0.000) | | - | | - | | 72.21*** (0.000) | |
| N | 996 | | | | 996 | | 7,617 | | 7,617 | | 7,617 | |

*Note: The omitted categories are: no access for transport equipment and for communication equipment, rain-fed & irrigated combined for land quality, hills and valleys for landscape, no access to irrigation or irrigation water, no formal schooling for education, no access for extension services, female for head sex, not reachable for distance to market, and AEZ 8 for AEZ. significance levels indicated by * p<0.10, ** p<0.05, *** p<0.010*

# 4 Chapter IV:
# Tests for Separability and Factor Market Participation in Afghanistan


**Abstract:**

We test for market failures by testing separability in household's production and consumption decisions and analyse household factor market participation in Afghanistan. Under the assumptions that all current and future markets exist and that farmers treat all prices as given, our analysis allows to model households' simultaneous production and consumption decisions into a recursive form in which production decisions can be made as independent of preferences of the farm household. Estimates of the household labour demand model rejects separability; labour demand decisions are strongly influenced by preferences and demographic compositions of household (i.e. endowments of labour) suggesting that there exist potential market failures in Afghanistan. Exploring input market participation, econometric results reveal that ownership of Information and Communication Technologies (ICT) and transport assets by households have a strong positive influence on the use of inputs. In addition, households living in communities with better access and within a closer radius of markets are more likely to participate in factor markets and spend more on purchased inputs. Identification through control function approach confirms endogeneity of ICT and transport equipment in the log-normal hurdle models analysing participation in chemical and fertilizer and tractor rental markets, however we reject endogeneity of ICT and transport equipment in the case of hired labour. Correcting for endogeneity bias revealed a negative association between the error terms in the reduced form and structural models.

## 4.1 Introduction

Improving farm productivity, crop yields, and market-oriented production to improve rural incomes entails improved access to input and output markets. However, small-scale subsistence and semi-subsistence farmers often face a number of barriers and constraints that make it difficult for them to become part of the commercial agriculture economy. One of the limiting constraints faced by farmers, especially subsistence farmers, is lack of market access due to higher transaction costs (Ouma et al., 2010). These costs associated with market transactions often result in lower input utilization by farm households, and in many instances they can be the primary reason generating market distortions that lead to market failures (de Janvry et al., 1991b). To explore input markets and household's marketing decisions in Afghanistan, we test whether household production and consumption decisions are consistent with the hypothesis of separability and use the results to investigate the presence of potential market failures or missing markets. Moreover, in an attempt to address potential market failures, we extend our analysis to empirically assesses the critical implications of market access and transaction costs on farmers' input utilization decisions.

In the context of Afghanistan, barriers such as poor infrastructure development, poor access to all-season roads and district and provincial markets, limited or no access to farm assets such as transport equipment, and lack of market information make it difficult or even impossible for small-scale farmers to sell their surplus produce and transport production inputs from the respective markets. As a result, farmers are often forced to under-utilize production inputs that in turn significantly decrease crop yields and production efficiency. Oushy, (2010) suggested that Afghan farmers require knowledge of market trends and opportunities as well as the necessary skills to manage their farms in an increasingly competitive environment. Therefore, it is essential to assess farm household behaviour and decisions regarding the extent of input use, especially in the context of high transaction costs and potential missing markets or market failures.

Market failures or missing markets affect household behaviour and decisions that subsequently affect welfare outcomes. Analysing household behaviour under imperfect market conditions helps observe and understand different strategies that households devise to mitigate the welfare costs that market failures impose (Vakis et al., 2004). Household decisions under perfect markets imply separability between production and consumption decisions. This means that households can solve recursively first the production problem and then, based on the profit (income) from the production stage, make consumption choices. On the contrary, under imperfect markets, production and consumption decisions are non-separable; this implies that household production decisions are affected by consumption preferences or jointly determined (Benjamin, 1992; Bowlus and Sicular, 2003; Dillon and Barrett, 2017; LaFave and Thomas, 2016). In order to better understand household labour demand, production and consumption decisions, and their investment choices, as well as formulate and evaluate relevant policies, it is essential to model the opportunities and constraints they face (LaFave and Thomas, 2016). Thus, in this study we attempt to provide evidence based on household's behaviour in relation to the market conditions and identify factors underlying separability of the household's production and consumption decisions.

Market participation for both inputs and outputs is a prerequisite and a key step towards commercialization of rural farms. In order to break out of the subsistence poverty trap and improve rural farm incomes, agricultural development policies must aim to identify and address barriers to market participation and potential missing markets (Barrett, 2008). Market imperfections and high transaction costs are generally thought of as the main limiting factors that hinder the exchange of goods in the local markets. de Janvry et al. (1991b) debated that rural markets are often imperfect and transaction costs can be so high that farmers are unable to participate in markets. Key et al. (2000) argued that the existence of high transaction costs including costs related to search and information, transportation, bargaining, monitoring, and contract enforcement implies that some households will opt for self-sufficiency instead of market participation.

In many developing economies, lower crop yields due to the under-utilization of the production inputs and imperfect markets for both inputs and outputs are generally responsible for slow productivity growth and income generation. As in most low-income countries, input application in Afghanistan lags substantially behind the world average. For illustration, take fertilizer usage as an example, the average consumption of commercial fertilizer is negligible and far below the world average and the average of south

Asia (Figure 4.1). Consequently, the question arises as to what factors limit the application of fertilizer and other inputs and are input markets failing? could improving household access to markets by reducing transaction costs improve market participation? Analysing the main drivers and constraints of market participation helps to design effective policies and interventions to expand agriculture input use and output marketing opportunities.

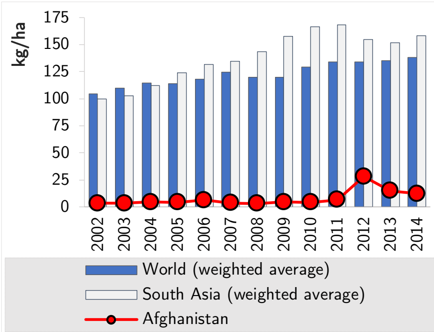



**Figure 4.1:** Fertilizer Consumption (Kg/ha of Arable Land) in Afghanistan, South Asian Region, and World (Weighted Average)

*Source:* World Bank Microdata, Development Indicators

Poor infrastructure development and weaker institutions together cause the costs of transaction to substantially rise, that in turn greatly alter farm household production and marketing related decisions. In most remote areas, smallholder farm households struggle to overcome the cost of entering the market due to the absence of sufficient means available to them (Barrett, 2008). Furthermore, in many instances, these resource-poor farm households do not possess the level of asset endowments required to guard them against adverse agro-ecological conditions and other production, market related, and political risks and shocks (Donovan and Poole, 2014). Besides, lack of access to reliable price information as well as information on potential exchange partners and players is yet another constraints making it hard for them to enter markets (Ouma et al., 2010). Analysing farm household behaviour in the context of imperfect market conditions requires an empirical understanding of both production decision-making at the farm level and market conditions, especially in low-income countries where production is carried out by smallholder farm households that make production and consumption decisions together.

Agricultural policy in Afghanistan encourages market-led development to ensure resource-poor peasant farmers are effectively a part of the broader agricultural economy so as to improve their incomes and livelihoods. Despite the recent economic growth in the country, a number of concerns and questions are raised about poorly functioning factor markets.

A major concern is that potentially incomplete markets and high transaction costs may hamper the overall commercialization process. Thus, empirical evidence to generate information about these factors affecting smallholder farmers' marketing decisions is required to better understand the decision-making environment.

The remainder of the chapter is structured as follows. Section 4.2 overviews relevant literature on input market participation and potential market failures. Section 4.3 discuses methodology and theoretical framework. Section 4.4 presents estimation strategy and econometric specification. Section 4.5 describes data and variables used in the analysis. Section 4.6 presents econometric results. Section 4.7 contains robustness checks for our econometric results. Section 4.8 summarizes the findings of the study and recommendations.

## 4.2 Overview of Literature

The majority of the rural poor in developing countries directly or indirectly depend on small-scale farming for their livelihoods and improving access to local markets remains a challenge for policy makers. Often living in remote areas with poor infrastructure, they face high transaction costs that significantly reduce their incentives for market participation (Barrett, 2008; Fischer and Qaim, 2012). Many barriers such as lack of sound institutional and physical infrastructure necessary to ensure low-cost access to competitive and well-functioning markets on the one hand, and diseconomies of scale on the other, impede smallholder market participation significantly (Lapar et al., 2003).

Despite the disadvantages they face, there is evidence that smallholders successfully participate in local markets. Barrett (2008) suggests that interventions aimed at facilitating smallholder organization, at reducing the costs of intermarket commerce, and at improving poorer households access to improved technologies and productive assets are central to stimulating smallholder market participation. Poulton et al., (2010) argue that small family farms may have an advantage because of their greater local knowledge of the locally demand commodities. Narrod et al., (2009) provide a number of examples of small-scale family farms that successfully participate in local markets through collective action and institutional support.

The most significant barriers to smallholder market participation are argued to be transaction costs including search, information, transportation, bargaining, monitoring, and contract enforcement costs (de Janvry et al., 1991b; Goetz, 1992; Holloway et al., 2000). A number of empirical studies assess the influence of transaction costs on household decisions to participate in the output market (e.g., Key et al., (2000), Makhura, (2001), Ouma et al., (2010), and Jagwe, (2011), Mather et al., (2013). A common finding is that transaction costs proxied by distance from the market, access to or ownership of farm assets such as transport equipment, and farm households' access to information and communication technologies have a significant impact on the decision to market their produce.

A few studies explicitly focus on the role of transaction costs in input market participation, e.g. Winter-Nelson and Temu (2005), Alene et al., (2008), Liverpool-Tasie, (2014), and Ricker-Gilbert et al., (2011). Winter-Nelson and Temu (2005) and Alene et al., (2008) argued that market participation is a two-stage decision-making process, where in the first stage households decide to participate in the input market, and in the second stage they decide on the intensity of the inputs used. Fixed transaction costs affect the decision to participate but not the intensity of participation, and non-participation is unobserved due to incidental truncation (hence these studies employ a sample selection model which assumes non-participation is the outcome of prohibitive fixed transactions costs). Variable transaction costs significantly determine both the household decisions to participate in market and the degree of input use.

In assessing household decisions to participate in fertilizer markets, Ricker-Gilbert et al., (2011) and Liverpool-Tasie, (2014) adopted the same conceptual framework underlining that input utilization is the outcome of a two-stage decision (i.e. participation in market and extent of use) where fixed transaction costs affect only the first stage, not the second stage. However, they argued that zero values of the input use (i.e. non-participation) is an optimal choice and therefore used a double hurdle model that is designed to allow the possibility that different factors might affect each stage. Both studies found that distance to markets and roads, access to farm assets, communication and transport equipment, and other proxies for transaction costs significantly affect market participation decisions and quantity of the inputs used.

Even though past studies have focused on the impact of transaction costs on households factor marketing decisions (Alene et al., 2008; Jagwe, 2011; Mottaleb et al., 2014; Ouma et al., 2010), they have not addressed the possible endogeneity problem in observable transaction costs. In most of these studies transaction costs were proxied for by distance to markets, ownership of transport assets (e.g. bike, motorbike, vehicles) and access to information and communication technologies (i.e. mobile phones, radios and TV, and internet services). However, household unobserved factors could possibility be simultaneously associated with the access to transport and ICT equipment and their marketing decisions. Thus, one major contribution of the current study is to allow for endogeneity in transaction costs and estimate their unbiased casual effects on household marketing decisions.

Imperfect market conditions and potential market failures or missing markets are other severe conditions that prohibits smallholders from market participation, that could be as a result of high transaction costs or non-competitive market prices, or legal barriers (de Janvry et al., 1991b; Dillon and Barrett, 2017). Imperfect markets, market failures or missing markets affect household behaviour (i.e. different condition leads to different outcomes such as separability and non-separability of production and consumption decisions) and consequently affect their welfare outcomes (Le, 2010; Vakis et al., 2004). When markets are incomplete or not competitive, consumption and production decisions are non-separable: production depends on the price of consumer goods and household preferences. On the contrary, under complete market conditions households are price takers, production decisions are made to maximize profits without reference to the consumption preferences, while consumption choices take into account the income from production (Benjamin, 1992; Dillon and Barrett, 2017; LaFave and Thomas, 2016).

Correct modelling of household production and consumption decisions requires a thorough understanding of behaviours (whether separable or non-separable). The relevant literature offers a number of different tests that aim to assess the separation hypothesis (Le, 2010; Vakis et al., 2004). Jacoby, (1993), Abdulai and Regmi, (2000), and Grimard, (2000) used a structural form approach that involves two steps; in the first step production function is estimated and shadow wage or marginal product for labour is derived and compared with the market price. Other studies including the seminal work of Benjamin, (1992), Bowlus and Sicular, (2003), LaFave and Thomas, (2016), and more recently Dillon and Barrett, (2017) used a reduced form approach which aims to test whether variables that affect consumption decisions also affect the labour allocation and production decisions. We summarize these studies and their findings in Table 4.1.

**Table 4.1:** Studies that Tested the Hypothesis of Separation

| Study | Country of study | Type of test | Findings |
|---|---|---|---|
| Benjamin, (1992) | Java, Indonesia | Reduced form approach | Fail to reject separation |
| Jacoby, (1993) | Peruvian Sierra | Structural form approach | Reject separation |
| Grimard, (2000) | Côte d'Ivoire | Structural form approach | Reject separation |
| Abdulai and Regmi, (2000) | Nepal | Structural form approach | Reject separation |
| Bowlus and Sicular, (2003) | Zouping County, China | Reduced form approach | Fail to reject separation |
| LaFave and Thomas, (2016) | Central Java, Indonesia | Reduced form approach | Reject separation |
| Dillon and Barrett, (2017) | Sub-Saharan Africa | Reduced form approach | Reject separation |

The reduced form approach involves whether household endowment of labour significantly affects their labour demand. Some of these studies have raised concerns about potential econometric issues due to household level unobserved heterogeneity in the size of the household when estimating the household labour demand function to test the hypothesis of separation (i.e. changes in household demographic composition may be related to the demand). These unobserved changes in the household composition mainly arise from births of new members of the households but could also be as a result of death and aging of household members as well as migration into and out of the household (LaFave and Thomas, 2016). Some of these studies employed econometric techniques such as fixed effect models and instrumental variable approach to correct for this bias. Using a longitudinal data, LaFave and Thomas, (2016) and Bowlus and Sicular, (2003) used fixed effect techniques along with instrumental variables, whereas Grimard, (2000) used instrumental variable techniques to control for possible endogeneity in the household size. Using a cross-sectional sample, in a recent study Dillon and Barrett, (2017) defined the household size as the prime aged members of the household (members aged above 15 years) and excluded children from the analysis in an attempt to reduce the bias associated with potential endogeneity in the household composition.

Unfortunately our data lack the presence of good contemporaneous instruments, so we follow the recent study by Dillon and Barrett, (2017) and exclude children from our analysis as that should largely mitigate the bias due to unobservables, particularly the unobserved changes in the household composition associated with new births or the children socio-demographic group.

Other studies have raised a suspicion regarding potential endogeneity in the cultivated land area, as decisions regarding land and labour use may both be determined by other common factors omitted from the regression (Bowlus and Sicular, 2003). One way to tackle this problem is to include control variables related to land quality and household human capital, as well as controls to account for regional fixed effects. Following Bowlus and Sicular, (2003), we include covariates that control for land quality, landscape characteristics, age, literacy and education of the household head along with district fixed effects to try to avoid potential endogeneity problem in the land variable.

### 4.3 Concept and Theoretical Framework

To ease modelling and the interpretation of results, it is important to understand and clearly define the concepts of transaction costs, market participation and market failures. A number of studies have defined and contextualized transaction costs. Holloway et al., (2000) distinguish transaction costs between tangible (i.e. transportation costs, communication costs, legal costs) and intangible (uncertainty, moral hazard, etc.) costs. Pingali et al., (2005) contextualize transaction costs from the point they occur (e.g. Information costs arise ex ante, negotiation costs occur at exchange, while monitoring costs occur ex post of a transaction). Key et al., (2000) broadly categorizes transaction costs into two sub-categories: 1) Fixed Transaction Costs (FTC's), and 2) Proportional or Variable Transaction Costs (VTC's). FTC's are invariant to the quantity of an input purchased such as screening and search or information costs, while VTC's vary with the volume of inputs traded such as the cost of transportation. Because FTC's are one-off costs incurred, thus they may increase entry barriers but are unlikely to affect the quantity of the input used by households once the entry costs are paid for. VTC's on the contrary, increase with the amount of input used by farm households resulting in the raise of the input prices for buyers and lowers the price effectively received by sellers, creating a "price threshold" within which some households find it unprofitable to either sell or buy.

The heterogenous and volatile nature of transaction costs has challenged researchers attempting to measure and assess their impact on household's marketing decisions. When transaction costs are adequately high to prevent exchanges from occurring, then costs associated with transactions are unobserved (Alene et al., 2008). Information on transaction cost are also hard to collect in a survey particularly if farmers have no access to transportation and information equipment as there would be no paid out costs to observe (Alene et al., 2008; Key et al., 2000). In addition, when farmers transport their produce to the market or inputs from the market using their own transportation means, it would be difficult to measure the actual transport costs (Alene et al., 2008). Thus majority of the literature that studied transaction costs resorted to the observable factors that proxy for transaction costs such as ownership and access to transport and information equipment, distance to roads and markets, etc. (Winter-Nelson and Temu, 2005; Alene et al., 2008; Ouma et al., 2010).

Given the two distinct categories of transaction costs (i.e. FTC's and VTC's), we follow Winter-Nelson and Temu, (2005), Alene et al., (2008), and Ouma et al., (2010) and divide transaction costs into two categories. We use access to or ownership of transport equipment by households (bike, motorbike, or vehicles) and access to information and communication equipment (radio, TV, mobile phones, and internet services) as a proxy measures for FTCs, with farm or household's distance to all-season drivable roads and time taken to reach nearest permanent market as proxies for VTCs. Input markets may be subject to different transaction costs than the output markets which may impose different constraints on the households input utilization and intensity decisions. For instance, farm households may have to travel a longer distance to purchase inputs because input markets are usually located in the province centres, whereas outputs could be marketed at the village or district centres. This longer distance in turn imposes higher travel costs.

Most studies conceptualize that market participation is the outcome of a two-step decision process, namely participation and intensity of the volume of inputs applied by the farm households (Alene et al., 2008; Liverpool-Tasie, 2014; Ricker-Gilbert et al., 2011; Winter-Nelson and Temu, 2005). The rate of market participation is the percentage of farmers that actually purchase inputs from markets, whereas intensity of input use is the level of a particular input applied by farm households. Thus, in this study participation is defined as the percentage of farm households who actually reported a positive value of purchased inputs, while extent of participation is defined is the quantity of inputs applied conditional on the first stage.

de Janvry et al., (1991) and Dillon and Barrett, (2017) distinguishes between three different cases of market failures. If the exchange of goods is legally prohibited or rendered infeasible by some non-market force, then markets are truly missing. In the second situation, markets are functional, however exchange of goods takes place at non-competitive prices (i.e. prices that do not equate marginal profit and marginal costs), then markets are functional but are failing. The third condition of market failure may occur when markets exist and operate at the competitive and market-clearing prices but welfare outcomes for households are sufficiently low or sub-optimal so require interventions to improve wellbeing. Market failures that mismatch supply and demand can be induced by different factors such as legal restrictions, weak enforcement of contracts, transaction costs, and poor access to infrastructure. The design of interventions to tackle the market failure issue also depends on the type of the situation confronted as explained above. For instance, policy instruments to target completely missing markets may involve removal of legal restrictions or imposing property rights, whereas the later situation may require interventions aiming at increasing investment in public infrastructure to reduce transaction costs (roads, access to telecommunication, etc.), termination of collusion and formation of oligopolistic situation, education and provision of extension services, and possibly government subsidies.

In the context of subsistence or semi-subsistence agriculture systems, production decisions are made in a complex environment where production is carried out by households that both demand and supply labour. Under complete and competitive markets, these households exchange (hire in and hire out) their desired amount of labour freely to maximize profits. In this case, households are profit-maximizers and the amount of labour employed to carry out production would in theory be independent from their consumption decisions and household's endowment and preferences of labour therefore should not affect the production allocation of labour. This independence implies that household decisions are recursive such that households first aim to make optimum production choices, and consumption decisions are made in the second stage based on the profits and income from the first stage (Benjamin, 1992; Bowlus and Sicular, 2003; De Janvry and Sadoulet, 2006; Dillon and Barrett, 2017; LaFave and Thomas, 2016; Le, 2010). Alternatively, if the separation hypothesis fails (i.e. production and consumption decisions are non-separable), then it is an indication that markets are dysfunctional or are failing.

Following De Janvry and Sadoulet, (2006), we illustrate the concept of non-separability and the role of transaction costs in Figure (4.2). Consider the following hypothetical situation where we assess the impact of transaction costs on the market for a particular input; take as a second market failure; inexistence of a land market. Let the demand for the input of labour be denoted by $D(p, L_q)$ for households $i = 1,2,3$ with different farm sizes. To make comparison across households feasible, we assume that all households face the same supply denoted by $S(p, L_c)$ which is determined by the level of household labour endowments $(L)$ as a supplier. Let $p^m$ denote market price, $p^v$denote effective price of sale defined as market price net of transaction costs $(t_p^v)$, and $p^a$ denote the effective purchase price defined as market price $(p^m)$ plus the transaction costs $(t_p^a)$ for the input incurred in buying.

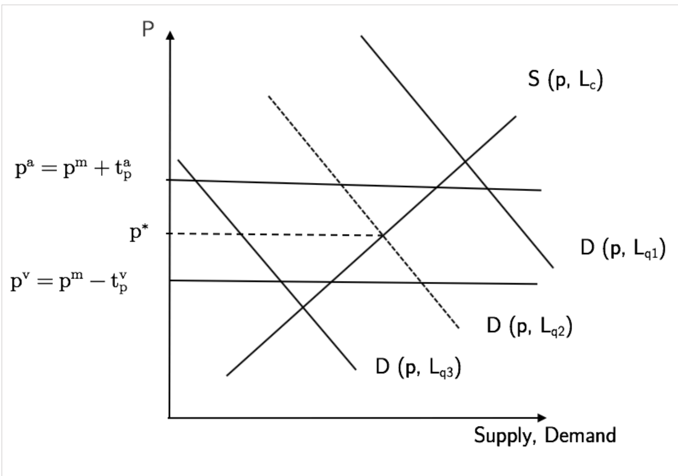



**Figure 4.2:** Variable Transaction Cost and Market Participation

*Source:* Adopted from De Janvry and Sadoulet, (2006) Household decision to purchase inputs from the market depend on the relative positions of households' supply and demand functions which are shaped by the level of household endowments of productive resources $(L_c)$ and demand characteristics $(L_q)$. Because variable transaction costs add to the market price and as a result there exists a non-zero price interval forcing households represented by the dashed line of $L_{q2}$ not to purchase inputs from the market. For these households their internal equilibrium defines a new shadow price $p^*(L_q, L_c)$ specific to each of them, as a result their behaviour is of a non-separable type where it is optimum for them to adjust production and consumption decisions and remain self-sufficient. Hence, both heterogeneity in household's endowments and differences in transaction costs $t_p^v$ and $t_p^a$ correspond to heterogeneity in the input market participation.

### 4.3.1 The Hypothesis of Separation in Agriculture Household Model

We define our utility function based on the standard time allocation model so that it simplifies and reflects the theory underlying the hypothesis of separation more clearly. The outline of this model is theoretically grounded in the generic household model (Singh et al., 1986) as articulated by Benjamin, (1992), and later applied by Bowlus and Sicular (2003), Le (2010), LaFave and Thomas (2016), and Dillon and Barrett (2017), to test for the separation hypothesis.

Consider a farm household that aims to maximize its utility represented by a strictly increasing and concave utility function (4.1). The utility is derived from the preferences over consumption $(C)$ and leisure $(L^l)$ which is conditional on household preference shifters $(Z)$ such as household endowments. The household is endowed with a fixed amount of labour $(\overline{L})$ which is supplied to the farm work $(L^f)$ to produce output that can be consumed by the household or sold to the market at the market price $(p)$, and off-farm work $(L^m)$ to receive market wages $(w)$. Households can also hire labour from the market, here denoted by $(L^h)$ at a market wage $(w)$ and purchase non-labour inputs $(X)$ such as seeds, fertilizer, etc. at the market price of $p_x$. Household's total land is denoted by $A$, which consists of household own land $(\bar{A})$ and land rented in $(A^r)$.

$$MAX_{C,A,,L^m,L^h,X}\ U(pF(A,L,X) + wL^m, L^l|Z) \quad 4.1$$

Subject to:

$$pC \leq \; pF(A, L, X) + wL^m - p_a A^r - \; wL^h - p_x X \tag{4.2}$$

$$0 \leq L^m \leq L^M \tag{4.3}$$

$$L \equiv L^f + L^h \tag{4.4}$$

$$\overline{L} \geq L^f + L^m + L^l \tag{4.5}$$

$$A = \; \bar{A} + A^r \tag{4.6}$$

$$L^l, L^f, L^h, L^m, C, A, X \geq 0 \tag{4.7}$$

Market imperfections are introduced into the model as the upper and lower constraints on the market labour: $0 \leq L^m \leq L^M$ where $L^M$ is the maximum number of hours a farm household can work in the labour market. The farm household faces imperfections if either the lower constraint or the upper constraint is binding ($L^m = 0$ or $L^m = M$). Then the hypothesis of separation holds. However, the farmer faces no imperfection if neither the lower nor the upper constraint is binding ($0 < L^m < L^M$), thus the farm household's behaviour is consistent with the separation (Le, 2010).

Based on the Langrangian function, the FoC for labour ($L$) can be calculated as in Equation (4.8) and FOC for $L^m$ can be derived as in (4.9 and 4.10) depending on the market conditions and separation:

$$w^* = \; pF_L(A, L, X) \tag{4.8}$$

$$w^* = w \qquad \text{if} \quad 0 < L^m < L^M \tag{4.9}$$

$$w^* \neq w \qquad \text{if} \quad L^m = 0 \;\; \text{or} \;\; L^m = L^M 0 \tag{4.10}$$

Where $w^*$ is:

$$w^* = \frac{U_l \; pF(A, L, X) + wL^m - p_a A^r - \; wL^h - p_x X; Z)}{U_c \; pF(A, L, X) + wL^m - p_a A^r - \; wL^h - p_x X; Z)} \tag{4.11}$$

Where $F_L$ in (4.8) is the derivative of output with respect to labour and $w^*$ is called the shadow wage or the opportunity cost of time. If the separation hypothesis holds, then the constraints are not binding as in Equation (4.9) and therefore $w^* = w$. Pluging $w$ for $w^*$ in Equation (4.8) we get $w = pF_L(A, L, X)$ where $Z$ does not appear implying that the choice of labour does not depend on the preference shifters. In other words, under complete and competitive market conditions, labour allocations in production are not affected by the household endowments of labour. In this case the farm household hires in labour or supplies labour to the market, and exchanges other inputs at exogenous, market-clearing prices, so that it allocates labour to maximize farm profits first, and then makes consumption choices conditional on the profit from production. On the contrary, in the case of non-separation where $w^* \neq w$ when the labour market constraint is binding $(\mathrm{i.e.}\ \ L^m = 0\ or\ L^m = L^m)$, then $L$ can be derived by substituting Equations (4.8) into (4.11) such that:

$$w^* \neq w \ \Rightarrow pF_L(A, L, X) = \frac{U_l\ pF(A, L, X) + wL^m - p_a A^r - \ wL^h - p_x X; Z)}{U_c\ pF(A, L, X) + wL^m - p_a A^r - \ wL^h - p_x X; Z)} \qquad 4.12$$

Where preference shifters $(Z)$ do appear in the Equation (4.12), meaning that labour allocation in the first stage of production is affected by the household endowments, thus the production and consumption decisions are not separate from each other. For further discussion on the theoretical model see Le, (2010).

As discussed earlier, there are two sets of tests available in the literature to analyse the above relationship. The first set of these test implemented by Benjamin, (1992), Bowlus and Sicular, (2003), Le, (2010), and more recently by LaFave and Thomas, (2016) and Dillon and Barrett, (2017) involve a reduced form approach that tests whether variables that affect consumption decisions (household preference shifters denoted by $Z$) also affect the labour allocation decisions in production stage, while the second set of these tests implemented by Jacoby, (1993), Abdulai and Regmi, (2000), and Grimard, (2000) involve a structural form approach testing the relationship between $w$ and $w^*$. In the later approach, since $w^*$ cannot be observed, it should be estimated using a production function (Le, 2010). The marginal product of labour from production function is equivalent to $w^*$ which then can be compared to the market price $w$ to test the hypothesis of separation (Le, 2010; Vakis et al., 2004).

Because the second approach involves estimation of a production function, questions arise with regard to choosing the correct functional form and due to the fact that endogeneity of variables in the production function may contaminate the results (Le, 2010). We choose the reduced form approach to test the hypothesis of separation without the need to estimate the production function. One point to bear in mind is that using the reduced form approach, rejection of the separation hypothesis may not be directly interpreted as a test for a market failure in a specific input market, as failure in any market will induce non-separable behaviour (Dillon and Barrett, 2017; Vakis et al., 2004). Similarly, rejecting the separation hypothesis can indicate failure of multiple markets simultaneously as relative prices of inputs or outputs (not absolute prices) may generate distortions resulting in market failure (Dillon and Barrett, 2017). Moreover, failing to reject the null hypothesis which implies consistency with the recursive or separation behaviour, may not mean that complete markets actually exist, it may rather be the result of household decision to allocate resources in a manner that make up for missing markets (LaFave and Thomas, 2016).

### 4.3.2 Market Participation and Agriculture Household Model

For simplicity and to avoid complications, we redefine our theoretical model and utility function to better accommodate the impact of transaction costs on household's marketing decisions. Following Key et al., (2000), Winter-Nelson and Temu, (2005), and Alene et al., (2008), and Ricker-Gilbert et al., (2011) input use by the farm households can be modelled as a two-step decision process: 1) household decision whether to purchase inputs from the market, and 2) household decision on the extent or level of expenditures on inputs. These household decisions can be analysed using a generic static household model in which utility is a function of net revenue:

$$MAX(U) = U(p_q Q - w_i X_i) \tag{4.13}$$

Subject to:

$$F(Q, X; Z) = 0 \tag{4.14}$$

Where Equation (4.13) is the objective utility function and Equation (4.14) represents production technology constraint in which $p_q$ and $Q$ represent output price and volume, $w_i$ and $X_i$ represent unit price and quantity of the $i^{th}$ input used, and the vector of $Z$ collects household characteristics. Production technology is represented by a well-behaved production function such that $\partial Q/\partial X_i > 0$ and $\partial^2 Q/\partial^2 X_i < 0$.

The utility function in (4.13) can be expanded to accommodate transaction costs explicitly. Let $VTC_o$ and $VTC_i$ denote variable transaction costs for unit of output and input respectively, so that the adjusted output price becomes $p_q' = (p_q - VTC_o)$ which is a downward adjustment in the output price, and the adjusted input price becomes $w_i' = (w_i - VTC_i)$ which reflects an upward adjustment in the input price (i.e. an increase in price due to VTCs).

Households market their surplus produce which is assumed to be equal to total output produced less total output consumed ($Q_i^s = Q_i - Q_i^0$) and purchase required inputs from the market which is assumed to be equal to total input applied to production less own input ($X_i^b = X_i - X_i^0$). This illustrates that for purchased modern inputs (i.e. from the market) such as certified seed, chemical fertilizer, and pesticides, the household relies entirely on the market (i.e. $X_i^b = X_i^0$), whereas in the case of the labour input, total input may equal to the sum of the hired and own labour, hence $X_i^b = X_i - X_i^0$ (Goetz, 1992). Let $FTC_o$ and $m_q$ be fixed transaction costs incurred at selling and the quantity of the output sold to the market, $FTC_i$ and $m_i$ be the fixed transaction costs for inputs incurred at buying and the volume of input purchased from markets respectively, then the objective function in (4.13) can be redefined to accommodate transaction costs such that:

$$\begin{gathered} MAX(U) = U(p_q Q^c + (p_q - VTC_o)Q^s - w_i X_i^0 - (w_i + VTC_i)X_i^b \\ -FTC_0(m_q) - FTC_i(m_i) \end{gathered} \qquad 4.15$$

Where

$$m_i = \begin{cases} 1 & X_i^b > 0 \\ 0 & otherwise \end{cases}$$

Taking the first order condition of the objective function will yield a reduced form of the input demand conditional on the market participation which implies that for households that actually purchase inputs from the market, the quantity is unaffected by the FTC. This means that once entry costs are paid, then fixed transaction costs do not affect the rate or quantity or expenditures on the inputs being purchased by the households.

$$m_i = f(p_q, VTC_o, w_i, VTC_i, FTC_o, FTC_i; Z) \qquad 4.16$$

$$X_i^b = f(p_q, VTC_o, w_i, VTC_i; Z) \qquad if \quad X_i^b > 0 \qquad 4.17$$

Equations (4.16) and (4.17) represent input market participation and input demand by the household, where participation is a function of prices, fixed and variable transaction costs, household characteristics, whereas input demand is a function of prices, variable transaction costs, and household characteristics but not fixed transaction costs.

### 4.4 Estimation Strategy and Econometric Specification

The estimation strategy and econometric techniques in this section build on the conceptual model presented in the previous section. To test the hypothesis of separation, we estimate the labour demand Equation in (4.18) using ordinary least square (OLS). As discussed in the previous sections, under the complete and competitive market conditions, the separation hypothesis specifically states that labour demand is invariant to the household endowments of labour (i.e. household size and composition are jointly statistically indistinguishable from zero). Rejecting the null hypothesis (separation) in favour of the alternative (non-separation) implies that markets for multiple inputs such as credit, insurance or land are failing as multiple market failures are required to generate distortions in factor markets because relative prices - not absolute price - are what matters in determining the efficient allocation of resources (Dillon and Barrett, 2017), whereas failing to reject the hypothesis of separation implies the presence of complete and competitive markets. To empirically test the hypothesis of separation, we estimate the following econometric model:

$$lnL_i = \alpha_i + \beta lnA_i + \delta_0 lnN_i + \sum_{s=1}^{S} \delta_s \frac{N_i^s}{N_i} + \sum_{k=1}^{K} \theta_k X_i + \sum_{j=1}^{J} \gamma_j D_j + W_t + \varepsilon_i \qquad 4.18$$

$$H_0\colon \delta_0 = \delta_s = 0$$

$$H_A\colon \delta_0 \neq \delta_s \neq 0$$

Where $L_i$ represents the total labour employed (household own labour and hired labour) by the $i^{th}$ household measured in person-days, $A$ is the total amount of land cultivated by the farm households, $N_i$ is the size of the household for $i^{th}$ household, $N_i^s$ are the household composition or structure variable such as age-sex demographic groups, $X$ collects additional control variables such as land quality, $D_j$ represents dummies to control for regional variation, and $W_t$ represent the year dummies for the repeated cross section. Since we do not have data on wages in our survey to include as a variable in (4.18), we follow Bowlus and Sicular (2003) and Dillon and Barrett (2017) and rely on the district and year fixed effects in (4.18) to mitigate difficulties arising from complex wage structures. The null hypothesis of separation ($\delta_0 = \delta_s = 0$) states that household structure variables (e.g. age-sex demographic groups) and the estimated coefficient of the household size are jointly indistinguishable from zero. Rejection of the null hypothesis implies non-separability of the household's production and consumption decisions.

The estimation strategy follows the empirical approach seminally outlined by Benjamin (1992) and recently applied by Dillon and Barrett (2017). We define four sex-age demographic groups that are included in (4.18) along with the household size. Following Dillon and Barrett (2017), prime age comprises of household members aged between 14-64 years and elderly sex-age group comprises of household members aged above 64. Household members (males and females) aged below 14 are excluded from the regression to avoid mixing children and adults and to try and reduce concerns about the productivity differences and more importantly to mitigate concerns of potential endogeneity problem in the household size. However, children's contribution to the agriculture labour demand (total labour days – the dependent variable) is accounted for, assuming that each child day is equivalent to half of an adult work day.

As a large percentage of the households in our sample do not purchase inputs from the market, it is plausible to argue that the Heckman sample selection models may better represent the data. However, this model assumes that the zero values for the input use (i.e. for household who did not actually participate in the market) are as a result of incidental truncation where the zero values are unobserved. In the context of Afghanistan, it is safe to argue that use of inputs is very common among farmers and that they are aware of their economic benefits. The zero observations may therefore be an optimal outcome as farmers may not purchase inputs due to market conditions or unfavourable agronomic and climatic conditions. In this context, a corner solution model seems to be more appropriate than the sample selection. As pointed out earlier under the theoretical framework section, the use of inputs can be an outcome of a sequential two-step decision process namely participation in the input market and intensity of expenditures. Therefore, we choose a Lognormal Double Hurdle (LDH) model proposed by Cragg (1971) which is more flexible compared to the standard Tobit model as it is designed to allow that there might be different factors that affect the first stage decision of participation and the second stage decision that determine the probability of participation. The LDH model can also allow us to consider that the same factor can potentially affect participation and expenditures in different ways (unlike the sample selection model that requires a strict exclusion restriction).

We hypothesize that fixed transaction costs are likely to affect the first stage, but not the second stage decision related to the intensity of the expenditures. Once the entry costs are paid, the household decisions on the amount of expenditures on the inputs is unaffected by them. Following Winter-Nelson and Temu (2005), Alene et al., (2008), and Liverpool-Tasie (2014), we use ownership or use of ICT and transport equipment by households as proxies for fixed transaction costs, with time taken to reach nearest permanent market and distance to nearest all-season driveable road as proxies for variable transaction costs. Given the dependent variable (i.e. the decision to participate in input market and extent of expenditures), access to ICT equipment can mitigate the one-off information or search cost, whereas ownership of transport equipment may mitigate transportation costs. However, given, the dependent variable (expenditures on inputs), distance to roads and markets proxy for proportional costs; the longer the distance to markets and roads, the higher are the costs incurred to transport input or outputs.

One potential problem in our analysis is the estimation bias due to endogeneity in the use of ICT and transport equipment ownership by the farm households. To remove the possible endogeneity bias and capture the true casual effect, we allow these variables to be endogenous and use Instrumental Variable (IV) technique to overcome the endogeneity problem. Chowdhury (2006) stated that the use of the telephone is possibly correlated with the household unobservable characteristics that may also be correlated with their market participation decisions. Thus, the estimated coefficient for the use of ICT equipment could possibly suffer from the endogeneity problem due the omitted variable bias. It can be hard to priori anticipate the direction of the bias as these unobserved characteristics may simultaneously increase or decrease the use of both ICT and transport equipment and market participation. However, it is plausible to assume that households that patriciate in markets are more likely to own or use ICT and transport equipment too, thus one would expect the coefficient estimate of the ICT and transport equipment variables to be biased upward. Similarly, ownership of the transport equipment by household is likely to be endogenous. Household unobserved characteristics may be correlated with both the ownership of transport equipment and market participation in a similar passion leading to an upward estimation bias.

We choose two instrumental variables to correct for the potential endogeneity bias in the use of ICT equipment by farm households, namely: 1) whether the farm household has access to electricity and 2) mean of off-farm income of other farmers at the community level. Access to electricity involves electrification from household own, private, and public electricity sources[14] of power. One could argue that electrification may signal regional investment and that households located closer to local markets may have better access to electricity and input markets, and therefore access to electricity may be correlated to the household input use decisions. However, in reality and given the data on access to electricity, the major sources of power are solar, community generator (hydro) and use of battery which are mostly common throughout the country regardless of whether household are located close to local market places. This means that the primary source of power is not from the public (government) source which is more likely to be more accessible by households that live near the local market centres. Thus, household access to electricity could not be directly correlated with their decisions to participate in the input markets, however it is directly linked to using ICT equipment. We also control for potential regional variations by including district fixed effects in our structural model. Therefore, access to electricity may affect household marketing decisions only through using ICT technologies that play a vital role in reducing search costs. Our second IV is the mean of off-farm income of other farmers in the community which is constructed as:

$$Mean\ off-farm\ income\ of\ other\ farms = \frac{\sum OFY_s \ - OFY_i}{N_s - 1} \qquad 4.19$$

Where $\sum OFY_s$ is the sum of off-farm income at Shura/community level, $sum\ OFY_i$ is the off-farm income of the $i^{th}$ farmer in the community, and $N_s$ is the number of farm households/observations in the respective community. The Mean of off-farm income of other farms in the household is intended to capture the status of local non-farm employment; higher non-farm employment in the community signifies high prevalence of

[14] Household power sources of electricity are: electric grid (6%), government generator (0.16%), private generator engine (1.2%), private generator hydro (2%), community generator engine (1%), community generator hydro (12%), solar (52.5%), wind (0.5%), and battery (13%).

non-farm employment opportunity at the local level which may in turn translates into greater potential for households to use ICT equipment. While we control for the household own off-farm income, household and farm assets, and other district level fixed effects directly in our structural model, our assumption is that the instrument will affect market participation only through the use of the ICT equipment channel.

Similarly, we instrument the ownership of or access to transport equipment using two instruments. Firstly, we use the number of road and bridge construction/rehabilitation projects being implemented in the community within 12 months. Controlling directly for distance between farm and local markets and road density within the community and other district level fixed effects in our analysis, the only remaining pathway for this instrument to influence household decisions to participate in input market is through the farm's accessibility to local input markets via improved roads development. Secondly, we use mean of the off-farm income of other farms within the community defined earlier as an instrument to remove bias due to unobservable that may affect both household decisions to own/use transport equipment and participate in input markets. In assessing households machinery investment decisions, Ji et al., (2012) used similar instruments (mean off-farm employment time and wage for other household in the district) to account for possible endogeneity in household's decisions to invest in farm machinery.

We use a Control Function (CF) approach to correct for possible endogeneity problem in the use of ICT and transport equipment. The CF approach requires the use of Instrumental Variables (IV) that should be included in the reduced form estimation but not included in the structural model of the factor market participation and demand equations and that they should satisfy the orthogonality condition. The CF technique entails that the endogenous variable is regressed on the instruments in the reduced form estimation and subsequently generalized residuals from the reduced form estimation are estimated and used as an independent variable in the structural model in addition to the actual endogenous variables themselves (Petrin and Train, 2010; Wooldridge, 2015).

Given the LDH model and the binary nature of endogenous variables in the context of corner solution, we choose control function because it is more efficient for binary outcome endogenous variables which other instrumental variable techniques (such as 2SLS, GMM, ivprobit) do not estimate efficiently. In addition, the CF approach is efficient even for weak instruments (Tadesse and Bahiigwa, 2015; Wooldridge, 2007). The CF approach is more efficient due to the prevalence of zeros in our structural equation, giving it the properties of a non-linear corner solution. Similar estimation strategy was applied by Winter-Nelson and Temu (2005), Ricker-Gilbert et al., (2011), Liverpool-Tasie (2014), Tadesse and Bahiigwa (2015), and Ragasa and Mazunda (2018) who analysed household's marketing decisions in the input or output markets.

Using the CF approach, in the first stage we estimate the following reduced form equation using a Probit model where we regress the binary endogenous variables (use or ownership of the ICT and transport equipment) on a number of controls and IVs:

$$\Pr(T_i = 1|M) = \alpha + \gamma Z_i + \varphi M_i + v_i \qquad 4.20$$

Where $T_i$ represents the endogenous variables of transaction costs for the $i^{th}$ household proxied for by the ownership of transport and ICT equipment, $M_i$ represents the vector of explanatory variables that affect $T_i$, and $Z_i$ represents instrumental variables that are not included in $X_i$ (or explanatory variables) in the structural model,$v_i$ represents the error term that follows a normal Probit distribution $N(0,1)$. Following Wooldridge (2015), the generalized residuals after the reduced form Equation (4.20) estimated by Probit model can be obtained as:

$$\widehat{gr_i} = T_i\ \lambda(Z_i\ \gamma) - (1 - T_i)\ \lambda\ (-Z_i\gamma) \qquad i = 1,2,3, \ldots N \qquad 4.21$$

Where $\widehat{gr_i}$ is the generalized residual obtain from Equation (4.20), and $\lambda = \phi(.)/\Phi(.)$ is the inverse mills ratio.

In the second step of the CF approach, we use the generalized residuals $\widehat{gr_i}$ obtained from the reduced form Equations (4.20) as additional regressors in the structural models estimated by the LDH models (i.e. the residuals are used as explanatory variable in the first hurdle-Probit regression of the LDH model). Following Wooldridge (2002), the general form of the LDH model can be written as:

$$\text{Hudle 1:}\ \ Pr(y_i = 0|x) = 1 - \Phi(x\gamma) \qquad 4.22$$

$$\text{Hurdle 2:}\ \ Log(y)\ |(x, y > 0) \sim Normal(x\beta + \sigma^2) \qquad 4.23$$

Where the decision to participate is governed by the Probit model in hurdle1 and the extent of expenditure is estimated by the truncated model. The LDH model in the second hurdle assumes that $log(y)$ follows a normal distribution for $y > 0$. Given the general LDH model in (4.22) and (4.23), our empirical model[15] takes the following form:

$$Pr(y_{1i}^* = 1|x_i) = \alpha_i + \ \delta\widehat{gr}_i + \theta T_i + \beta X_{1i} + W_t + D_j + u_{1i} \quad \text{Participation} \qquad 4.24$$

$$y_{2i}^* = exp(\alpha_i + \theta T_i + \ \beta X_{2i} + W_t + D_j + u_{2i}) \qquad \text{Extent of expenditures} \qquad 4.25$$

$$\begin{array}{ll} y_i = 1 & Only \ \ if \quad y_{1i}^* > 0 \ \ and \ \ y_{2i}^* > 0 \\ y_i = 0 & Otherwise \end{array}$$

Where $y_{1i}^*$ is a latent variable denoting the household's decision to participate in the input market (e.g. participation=1, and 0 otherwise) and $y_{2i}^*$ is a latent variable presenting the expenditures on the $i^{th}$ input purchased by the farm household, $y_i$ represents the actual observed dependent variable, which is the expenditure on $i^{th}$ input purchased by the household, $\widehat{gr}_i$ is the residual obtained from the reduced form Equation (4.20), $X_i$ is a vector of controls including household demographic and socioeconomic characteristics, $D_j$ represents district dummies that capture regional variation for $j^{th}$ district, $W_t$ represent the year dummies in our pooled cross-sectional sample, $\varepsilon_i$ is the error term. If the coefficient on $(\widehat{gr})$ is significantly different from zero in the structural model (4.24) then transaction costs (represented by ownership of transport equipment and use of ICT equipment) are endogenous in a farmer's decision to purchase inputs from the market.

The participation and extent of expenditure Equations in (4.24) and (4.25) are assumed to be independent (Hsu and Liu, 2008; Wooldridge, 2010) of each other, and are estimated using a Maximum Likelihood (ML) estimation procedure. The log-likelihood of function of the LDH model can be written as:

$$\ln(L) = [1 - \ \Phi(\beta X_{1i})] + \ln[\Phi(\beta X_{1i})] + \{(\phi \ [\ln(y_{2i}) - \ \beta X_{2i}/\delta]) - \ln(\delta) - \ln(y_{2i})\} \qquad 4.26$$

---

[15] We did not include the subscript of $(t)$ in our equations because each farm household is repeated in the survey only once, however since we use a repeated cross-sectional data that are collected in different years, we added a dummy variable representing individual survey year (indicated by $W_t$ in Equations (4.24) and (4.25).

Where $\phi(.)$ and $\Phi(.)$ are the normal probability density function (pdf) and cumulative distribution function (cdf) respectively. As in most non-linear models, coefficient estimates from the LDH model are directly hard to interpret, we estimate Average Partial Effects (APE) of the explanatory variables on the participation probability and expected expenditures level (in the second stage) given the positive decisions on participation. The APE of participation from the hurdle 1 Probit model in Equation (4.24) is:

$$Pr(y_{1i} = 1|X_{1i}) = \Phi\left(\frac{-\beta X_{1i}}{\delta}\right) \qquad 4.27$$

The average partial effects in the input purchase decision in hurdle 1 in Equation (4.24) represent the probability of input market participation for changes in corresponding explanatory variables. The expected value of expenditures in hurdle 2 (estimated using truncated regression) conditional on a positive purchase decision is given by:

$$P(y_{2i}|X,\ y_{2i} > 0) = \exp(\ \alpha + \theta T_i +\ \beta X_{2i} + W_t + D_j + \delta^2/2) \qquad 4.28$$

For the extent of expenditures model in hurdle 2, the conditional APE's show the conditional expectations for strictly positive expenditures on inputs with respect to the change in independent variables (for dummy variables, change implies switching from zero to one) evaluated at the ML estimates. Because the dependent variable in the second hurdle is in logarithmic form, the conditional APE's can be interpreted as elasticities for log-transformed continuous variables when $y_{2i}$ is strictly positive, whereas for discrete choice variable the APE measure percentage changes in the dependent variable when the variable shifts from zero to one, ceteris paribus. APE and Standard errors of the estimated marginal effects are computed using the margins command and delta method. The maximum likelihood estimation of the log normal hurdle model was obtained in Stata® 15 using *probit* and *truncreg* commands for participation and extent of expenditures, respectively.

### 4.5 The Data, Summary Statistics, and Description of Variables

This study uses repeated cross-section data from the Afghanistan Living Condition Survey (ALCS) conducted by the Central Statistics Organization (CSO) in 2011/12, 2013/14, and 2016/17. The surveys include both quantitative data and in-depth qualitative information on several key indicators including farming and livestock production in

Afghanistan. Each survey covered all 34 provinces of the country. In total 35 strata were identified each year, 34 for the provinces and one for the nomadic (Kuchi) population.

The surveys use largely similar structured questionnaires, hence data on similar indicators and variables were collected every year on sectors including agriculture, livestock, labour, household assets, income, and expenditures. One limitation is that it is not possible to disaggregate data at the crop level, restricting our analysis to aggregate farm level estimation. A strength of the surveys, however, is that they are representative at the national and provincial level, and continuous data collection over a cycle of 12 months allow to capture potential variations across the seasons. The surveys also include district and community level questionnaires that aim to collect data on development priorities, projects being implemented within the community, access to education, as well as district level market prices.

Each survey covered about 20,786 households and roughly 157,262 persons across the country. In total, the pooled sample from three years covers about 61,452 households. About 50% of the households reported some level of engagement in farming (i.e. with positive agriculture production and cultivated land area), reducing the analytical sample to roughly 30,000 households. Moreover, after accounting for missing values on key variables especially labour, our total usable sample became 21,189.

Before presenting the summary statistics on key variables used in the analysis, we introduce and define each variable and the measure in Table 4.2. For the first part of the study where we test the hypothesis of separation, the dependent variable is total (own and hired) labour measured in person-days. The dependent variables in the analysis of market participation in the second part of this study include a set of standard variables that are theoretically expected to influence household's decisions of participation in the factor markets and intensity of expenditures on each input. Household level summary statistics on the key variables used in the analysis are presented in Table 4.3. Columns 1,2, and 3 present means and standard deviations of key variable for each wave 2011/12, 2013/14, and 2016/17 respectively, while column 4 reports descriptive statistics for the pooled sample.

**Table 4.2:** Description of Variables used in the Analysis

| **Variable** | **Description** | **Measure** |
|---|---|---|
| *Dependent variables* | | |
| Total labour days | Total labour (own & hired) employed by farm | Labour days |
| Fertilizers & chemicals use | Whether farm HH uses fertilizers & chemicals | (Binary, 1=use) |
| Rent tractor | Whether the farm HH hire tractor | (Binary, 1=rent) |
| Hire labour | Whether the farm HH hire labour | (Binary (1=hire) |
| Fertilizer & chemical expense | HH spending on fertilizers and pesticide | Afghani |
| Tractor expenditures | HH expenditures on hiring tractor | Afghani |
| Labour expenditures | HH expenditures on labour hire | Afghani |
| *Explanatory variables* | | |
| ICT equipment | Whether HH owns ICT equipment such as TV, mobile, radio and internet | 1=own,0 otherwise |
| Transport equipment | Whether HH owns transport equipment such as car, bike, and motorbike | 1=own, 0 otherwise |
| Time taken to reach market | Time taken to reach the nearest permanent market by car | Hours |
| Distance to road | Nearest all-season drivable road to the community | Km |
| Total land | Total land cultivated annually | Jeribs |
| Off-farm income | HH income from non-farm activities | 10k Afghani |
| HH size | Number of members of the household | Count/persons |
| HH head literacy | Whether HH head can read and write | 1=can read & write |
| HH head education | HH head's highest formal education | Years |
| HH head age | Age of the household head | Years |
| HH head age square | Square of the age of the household head | Years |
| Land type | Whether the cultivated land is all irrigated or combined irrigated & rain-fed | 1= all irrigated, 0=combined |
| Landscape | Terrain or slope of the cultivated land (i.e. hills, valleys, and open plain) | 2=open plan, 1= valleys, 0=hills |
| Number of livestock | Number of livestock owned by HH (cows, sheep, goat, donkey, and horses) | Count |
| Number of oxen | Number of oxen owned by the farm HH | Count |
| Number of tractors | Number of tractors owned by the HH | Count |
| Electricity cost | HH spending on electricity | Afghani |
| *Instrumental variables* | | |
| Road/bridge project | Whether road/bridge project completed in the community in the last 12 months | (binary, 1=yes) |
| Mean off-farm income of other farms in the community | Mean of the off-farm income of other farmers at the Shura/community level | Afghani 10K |
| Access to electricity | Whether HH has access to electricity | (binary, 1=yes) |

**Table 4.3:** Summary Statistics of Variables used in the Analysis

| | 2011/12 | | 2013/14[16] | | 2016/17 | | Pooled | |
|---|---|---|---|---|---|---|---|---|
| VARIABLES | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Total labour (days) | 14.60 | 13.90 | 13.15 | 14.69 | 13.92 | 15.78 | 13.95 | 14.69 |
| Own labour (days) | 11.83 | 8.950 | 11.30 | 11.09 | 10.68 | 7.35 | 11.35 | 9.323 |
| Hire labour (days) | 3.777 | 11.04 | 2.704 | 9.953 | 4.01 | 14.49 | 3.490 | 11.76 |
| Fertilizer & chemicals (1=use) | 0.702 | 0.457 | 0.731 | 0.443 | 0.723 | 0.448 | 0.717 | 0.450 |
| Tractor hire(1=yes) | 0.560 | 0.496 | 0.579 | 0.494 | 0.542 | 0.498 | 0.561 | 0.496 |
| Labour hire (1=yes) | 0.323 | 0.468 | 0.194 | 0.396 | 0.217 | 0.412 | 0.253 | 0.435 |
| Fertilizer & chemical expense (AFN) | 3,246 | 5,029 | 5,324 | 9,880 | 5,043 | 8,661 | 4,399 | 7,933 |
| Tractor rent (AFN) | 1,974 | 3,455 | 2,688 | 5,354 | 2,515 | 4,880 | 2,350 | 4,542 |
| Labour exp. (AFN) | 944 | 2,759 | 672.2 | 2,474 | 1,027 | 3,714 | 877.9 | 2,966 |
| ICT equip (1=own) | 0.826 | 0.379 | 0.813 | 0.390 | 0.845 | 0.362 | 0.827 | 0.378 |
| Transport (1=own) | 0.470 | 0.499 | 0.456 | 0.498 | 0.544 | 0.498 | 0.485 | 0.500 |
| Distance to rd (km) | 3.420 | 6.613 | 2.527 | 9.052 | 1.284 | 5.559 | 2.561 | 7.310 |
| Time to mkt (>4 h) | 0.266 | 0.442 | 0.095 | 0.294 | 0.087 | 0.281 | 0.163 | 0.369 |
| Time to mkt (1-4h) | 0.201 | 0.401 | 0.354 | 0.478 | 0.000 | 0.000 | 0.197 | 0.398 |
| time to mkt (<1h) | 0.533 | 0.499 | 0.550 | 0.497 | 0.913 | 0.281 | 0.640 | 0.480 |
| Total land (Jeribs) | 7.628 | 24.53 | 8.303 | 23.28 | 8.098 | 12.32 | 7.972 | 21.50 |
| Off-farm income (10K AFN) | 3.527 | 6.744 | 5.375 | 11.71 | 4.822 | 8.169 | 4.472 | 9.028 |
| Livestock owned (N) | 15.70 | 30.73 | 11.84 | 23.43 | 14.40 | 25.67 | 14.10 | 27.25 |
| Oxen (own=1) | 0.212 | 0.409 | 0.158 | 0.365 | 0.191 | 0.393 | 0.189 | 0.391 |
| Tractor/thresher (N) | 0.024 | 0.155 | 0.054 | 0.233 | 0.033 | 0.181 | 0.036 | 0.191 |
| Land (1=all irrigated) | 0.711 | 0.453 | 0.767 | 0.423 | 0.686 | 0.464 | 0.723 | 0.448 |
| HH size (count) | 8.188 | 3.485 | 8.348 | 3.447 | 8.482 | 3.424 | 8.318 | 3.458 |
| Prime male share | 0.474 | 0.144 | 0.476 | 0.149 | 0.473 | 0.147 | 0.474 | 0.147 |
| Prime fem share | 0.476 | 0.134 | 0.475 | 0.135 | 0.464 | 0.137 | 0.473 | 0.135 |
| Elderly fem share | 0.017 | 0.063 | 0.018 | 0.063 | 0.025 | 0.072 | 0.020 | 0.066 |
| Elderly male share | 0.032 | 0.084 | 0.032 | 0.082 | 0.038 | 0.088 | 0.034 | 0.084 |
| Head education (yrs) | 1.942 | 4.030 | 2.202 | 4.285 | 2.224 | 4.218 | 2.101 | 4.166 |
| Head literacy(1=yes) | 0.280 | 0.449 | 0.318 | 0.466 | 0.309 | 0.462 | 0.300 | 0.458 |
| Head age (yrs) | 42.76 | 13.54 | 44.43 | 13.73 | 44.72 | 13.38 | 43.82 | 13.59 |
| HH head age square | 2,012 | 1,281 | 2,162 | 1,316 | 2,179 | 1,282 | 2,105 | 1,295 |
| Electricity(1=access) | 0.543 | 0.498 | 0.824 | 0.381 | 0.949 | 0.220 | 0.742 | 0.437 |
| Road project (1=yes) | 0.265 | 0.441 | 0.258 | 0.437 | 0.175 | 0.380 | 0.238 | 0.426 |
| Off-farm inc. other farms (10K AFN) | 4.692 | 4.394 | 6.440 | 5.675 | 5.908 | 5.456 | 5.583 | 5.184 |
| Electricity cost (AFN) | 70.44 | 233.9 | 87.97 | 291.7 | 61.77 | 251.5 | 73.81 | 258.7 |
| Observations | | 8,663 | | 6,876 | | 5,650 | | 21,189 |

*Source:* Author's calculation of the ALCS Data

[16] It should be noted that Afghanistan's economy suffered from an economic downturn in 2013/14 as the majority of international assistance withdrew from the country. This may have affected the estimated averages for certain variables presented in Table 4.3.

The descriptive statistics on the input use by the farm households show that roughly 2/3 of the farmers in the pooled simple purchase fertilizers and pesticides from the market, whereas about 56% hire tractor for ploughing or other farming activities. However, a relatively lower percentage (about 25% of the total sample) hire labour from the market. Perhaps household demand for the hired labour is seasonal and therefore they hire labour only during specific seasons such as planting or harvest seasons. Although the geographical coverage of the sample slightly changes (due to changes in security situation) from year to year, the average use of these inputs does not seem to fluctuate much over the years (Table 4.3). The average participation of the household in the input markets is relatively similar to other countries. Liverpool-Tasie, (2014) reported that about 70 % of households used chemical fertilizers in the Kano state of Nigeria. Dillon and Barrett, (2017) reported that on average 30% of Ethiopian households hire labour from the market, 40% in Malawi, 48.8% in Niger, 30% in Tanzania, and 45% households hire labour from the market in Uganda.

For households that have actually participated in the market, the estimated average expenditures on chemical and fertilizer in the pooled sample is 4,399 Afghani, tractor rental 2,350 Afghani, and average expenditure on labour hire is 878 Afghani (Table 4.4 and Figure 4.3). Overall, the averages are higher for the recent survey year, which may indicate a relatively higher application of inputs in the farm and expenditures (or it could simply be due to inflation). Note that in our econometric estimation, we control for time fixed effects by including a dummy for the survey year which will allow to control for fluctuations in the inflation rate from year to year.

The use of ICT equipment is quite common across the country, almost all the districts in the sample appear to be under the coverage of telecommunication facilities; in total 83% of farmers reported the use at least one of ICT equipment such as TV, radio, mobile phones, and internet services (roughly 50% of the sample farmers used mobile phones). Average use of ICT varies slightly from year to year (by about 1-3%), possibly because the sample size (the subsample of the agricultural households) and the geographical coverage changed over the years. The percentage owning transport equipment is relatively low at about 48%.

**Table 4.4:** HH Market Participation & Expenditures by the Input Type

| Input | Year | Percent of households: Non-users | Percent of households: Users | Expenditures (AFN) |
|---|---|---|---|---|
| Fertilizers & Chemicals | 2011/12 | 29.77 | 70.23 | 3,246.06 |
| | 2013/14 | 26.87 | 73.13 | 5,324.49 |
| | 2016/17 | 27.71 | 72.29 | 5,043.08 |
| | pooled | 28.28 | 71.72 | 4,399.36 |
| Tractor rental | 2011/12 | 43.97 | 56.03 | 1,974.22 |
| | 2013/14 | 42.1 | 57.9 | 2,687.50 |
| | 2016/17 | 45.79 | 54.21 | 2,514.89 |
| | pooled | 43.85 | 56.15 | 2,349.74 |
| Hired labour | 2011/12 | 67.65 | 32.35 | 944.21 |
| | 2013/14 | 80.58 | 19.42 | 672.20 |
| | 2016/17 | 78.29 | 21.71 | 1,026.57 |
| | pooled | 74.68 | 25.32 | 877.95 |

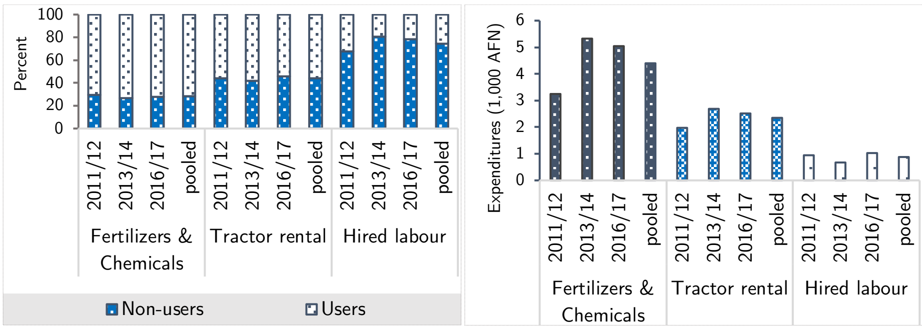
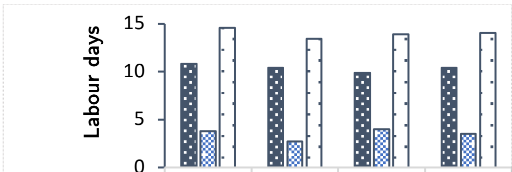



**Figure 4.3:** % of Users & Non-users, & Expenditures (AFN) by Input

*Source:* Author's calculation of the ALCS Data

The majority of households are located close to markets: about 64% stated that market was easily accessible (within community or less than one hour drive), 20% reported markets are within 1-4 hours of drive, but for the remaining 16% markets are not easily accessible (i.e. more than 4 but up to 12 hours of drive). While average distance to the nearest all-season road is about 2.56 km for the overall sample, average of distance to roads is consistently lower (1.28km) in 2016/17 compared to the average of 2.53 km in 2013/14 and 3.40 km 2011/12. This is an indication that road density and accessibility have increased over time because there are intensive road development projects underway across the country, as about 23% of the household in the pooled sample reported that there is a road/bridge construction or rehabilitation project being implemented within the community within the last 12 months.

Turning to the farm labour demand, we plotted the major types of labour applied by the farm households across different survey years to show the composition of total labour days employed over the years. As in Figure 4.4, household own labour comprises the major portion (75%) of the total labour days in our pooled sample. It can also be noted from the data that own or family labour and hired labour are somehow substitutes, as higher number of hired labour was reported when lower number of own labour is employed and vice versa.




| | 2011/12 | 2013/14 | 2016/17 | pooled |
|---|---|---|---|---|
| Own Labour | 10.82 | 10.44 | 9.91 | 10.46 |
| Hired labour | 3.78 | 2.70 | 4.01 | 3.49 |
| Total labour | 14.60 | 13.48 | 13.92 | 14.05 |

**Figure 4.4:** Composition of Labour in Different Survey Years
*Source:* Author's calculation of the ALCS Data

While household own family labour stays relatively constant across the years, the average of hired labour fluctuates more, particularly in the year 2013/14. This could be due to the year-to-year differences in the geographical coverage of the survey or the outcome of other socio-economic and agro-climatic conditions.

With regard to socio-demographic characteristics, the household size is relatively large in Afghanistan and variable (overall average of 8.4 persons per household with a standard deviation 3.5), rising from 8.19 in 2011/13 to 8.48 in 2016/17. The composition of household shows that the average share of prime male age demographic group (47.4%) and prime female group (47.3%) are equally distributed and consistent and not vary much over the years. The prime age-sex demographic groups are defined as household members between the age of 14-64 years. The average age of the household head is about 44 years in the pooled sample. Almost all households are headed by males, for this reason we excluded the household head gender variable from our analysis.

About 30% of the household heads in the pooled sample are literate meaning that they can read and write, although only 24% of the household heads have attended any formal schooling (i.e. primary, secondary, and tertiary education). The remaining 6% (out of the 30% that are literate) have received literacy training at home (i.e. home-based schooling).

The mean years of education completed by the head of the household is two years, with a very slight increase over time.

The average size of the total land cultivated is considerably low (8 Jeribs equivalent to 1.6 ha), with a slight increase over time, implying small landholdings per capita. The landholding shows high variability as the estimated standard deviation is 21.5 Jeribs, ranging from a minimum of 0.1 Jeribs to a maximum of 1,500 Jeribs per household. About 50% of the farm households reported some involvement in non-agricultural activities. While the overall mean of the off-farm income in the pooled sample is 4.5K Afghani with a standard deviation of 9K Afghani, it is much higher (8K Afghani) for the subsample of the farm households who actually reported a positive off-farm income.

The average number of livestock (including cattle such as cows, sheep, goats, donkeys, and horses, but not oxen) owned by the farm households in the pooled sample is 14 per household. Given the nature of our analysis, we separated oxen from the rest of the livestock and included it as a separate variable in our regression, as ownership of oxen not only proxy for the household wealth but also contributes towards land preparation and ploughing. On average, about 19% of households in the pooled sample reported that they own oxen. Average number of tractors/threshers owned by the household is about 0.036.

Next, we further break down our variables to compare the mean differences in selected household characteristics among users (market participants) and non-users (non-participants). A two-sample mean comparison t-test show that there are significant differences in selected household characteristics among participants and non-participants in all three input markets (Table 4.5). Except for the household head age, the mean differences for all household characteristics reported in Table (4.5) are statistically different among participants and non-participants. Surprisingly, households with smaller landholdings reported higher participation in fertilizer and chemical markets. This could be due to the resources needed to purchase sufficient fertilizers and pesticides to cover a larger area, or households with small landholdings may use more fertilizer to increase per unit production, given that they cannot expand their scale of operation. In addition, it could also be the case if small farmers have access to more manure or grow different types of crops. Households that reported participation in factor markets are generally in communities with better access to roads and markets. Similarly, farmers who participate in the fertilizer and chemical markets and those that hire tractors own less livestock (i.e. implies less application of manure and use of oxen for traction).

**Table 4.5:** Mean Difference in Household Characteristics by Input Use

| | Non-participants | Participants | Two-sample t-test | |
|---|---|---|---|---|
| VARIABLES | Mean | Mean | Difference | t-statistic |
| Fertilizers and chemicals | | | | |
| Distance to road (km) | 3.520 | 2.180 | 1.34*** | 10.39 |
| Total Land (Jeribs) | 9.830 | 7.240 | 2.60*** | 8.43 |
| Off-farm income (10K AFN) | 3.250 | 4.950 | -1.70*** | (15.51) |
| Livestock owned (N) | 18.08 | 12.53 | 5.55*** | 13.17 |
| Tractors/threshers (N) | 0.030 | 0.040 | -0.017 | (0.60) |
| HH size (count) | 7.520 | 8.630 | -1.11*** | (22.55) |
| HH head age (years) | 43.91 | 43.79 | 0.12 | 0.55 |
| HH head education (years) | 1.210 | 2.450 | -1.25*** | (22.75) |
| Farm income (10K AFN) | 4.480 | 5.550 | -1.08*** | (11.62) |
| Total Revenue (10K AFN)[17] | 4.530 | 8.130 | -3.60*** | (13.81) |
| Observations | 5,990 | 15,188 | 21,178 | |
| Tractor rental | | | | |
| Distance to road (km) | 2.460 | 2.640 | -0.18 | (1.81) |
| Total Land (Jeribs) | 6.330 | 9.260 | -2.93*** | (10.48) |
| Off-farm income (10K AFN) | 4.220 | 4.670 | -0.44*** | (3.69) |
| Livestock owned (N) | 17.26 | 11.63 | 5.63*** | 14.52 |
| Tractors/threshers (N) | 0.020 | 0.050 | -0.03*** | (12.03) |
| HH size (count) | 7.850 | 8.680 | -0.83*** | (17.70) |
| HH head age (years) | 44.23 | 43.51 | 0.73*** | 3.85 |
| HH head education (years) | 1.860 | 2.290 | -0.43*** | (7.50) |
| Farm income (10KAFN) | 4.310 | 5.980 | -1.68*** | (18.57) |
| Total Revenue (10K AFN) | 4.740 | 8.970 | -4.22*** | (12.80) |
| Observations | 9,287 | 11,891 | 21,178 | |
| Hired labour | | | | |
| Distance to road (km) | 2.620 | 2.380 | 0.24* | 2.15 |
| Total Land (Jeribs) | 6.720 | 11.660 | -4.94*** | (11.41) |
| Off-farm income (10K AFN) | 4.350 | 4.830 | -0.48*** | (3.36) |
| Livestock owned (N) | 13.98 | 14.47 | -0.49 | (1.11) |
| Tractors/threshers (N) | 0.030 | 0.050 | -0.01*** | (3.94) |
| HH size (count) | 8.370 | 8.170 | 0.19*** | 3.52 |
| HH head age (years) | 43.83 | 43.81 | 0.010 | 0.07 |
| HH head education (years) | 1.930 | 2.620 | -0.69*** | (9.88) |
| Farm income (10KAFN) | 4.930 | 6.190 | -1.26*** | (10.05) |
| Total Revenue (10KAFN) | 6.180 | 9.870 | -3.69*** | (9.38) |
| Observations | 15,816 | 5,362 | 21,178 | |

*Source:* Author's calculation of the ALCS Data

[17] Note that data on farm income is directly collected in the survey. Revenues are calculated for each crop using district price data and then aggregated.

Mean farm income and total revenues for participants is significantly higher than non-participants. Summary statistics on the ownership of tractors do not provide any statistically significant differences among participants and non-participants. Similarly, there are no significant age differences among participants and non-participants. Average years of education completed by the household head is higher among participants than non-participants as expected. Highly educated individuals may have more information on the benefits of inputs and are therefore using more inputs.

For categorical variables, we carried out a Pearson chi-squared test to compare whether the observed differences in selected characteristics associated with a change from 0 to 1 in inputs are significantly different among users and non-users (Table 4.A1 in Appendix 4.A). The mean differences in these variables are all significant among users and non-users for all inputs as expected, however they do not imply a causal relationship.

As for the fixed transaction costs, market participants own significantly more communication and transport assets, such as TV, radio, mobile phones, bicycle, and motorbikes which could indicate that farm households who participate in the market may be facing lower costs to access market information. In general, a higher percentage of the non-participants are located farther away from the nearest market and possess fewer assets such as transport equipment (see Table 4.A1 in Appendix 4.A).

We further illustrate the differences in household expenditures on inputs with respect to the ICT use and plot expenditures against the ICT use status across different years (Figure 4.5). While averages of expenditures on inputs slightly increase over time, there are significant differences in expenditures among farm household who own/have access to the ICT equipment (radio, TV, mobile phones, and internet) and household that don't. This may suggest that households with more information on factor markets and benefits associated with the input use are spending more money on inputs. The average expenditure on fertilizers and chemicals is about 1.95K Afghani for households that do not own ICT, whereas for households that own ICT the average expenditures on chemicals and fertilizers are 4.91K Afghani (about 2.5 times higher). Similarly, tractor rental and expenditures on hired labour is about 1.2K and 0.54K Afghani for farmers that don't own ICT equipment, whereas these expenses are 2.6K and 0.95K Afghani for tractor and labour hire respectively for households with access to ICT equipment (about 2.15 and 1.77 times higher). This is an indication that ownership of ICT equipment enables farm households to participate in the market, perhaps through provision of reliable and timely information on market prices and other services.

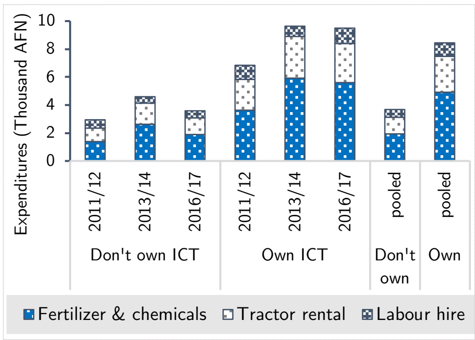



**Figure 4.5:** Input Expenditures (AFN) by the ICT Status Over Years
*Source:* Author's calculation of the ALCS Data

Likewise, there are significant mean differences in input expenditures with respect the ownership of transport assets (Figure 4.6). Farm households that own transport assets spend significantly more on inputs. Ownership of transport equipment facilitates the input delivery from market to the farm and therefore may mitigate transport costs as private transport could be a cheaper option then the public transport. However, since farm household that own transport equipment can avail themselves of cheaper transportation equipment, they are likely to participate more and use larger quantities of input factors as compared to their counterparts who may use other means of transportation which is generally a more expensive option.

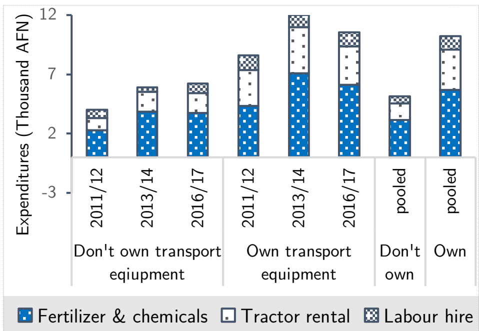



**Figure 4.6:** Input Expenditures (AFN) by the Transport Equipment Status
*Source:* Author's calculation of the ALCS Data

Distance to market is another key variable in our analysis that is used to proxy for proportional transaction costs. The longer the distance to markets, the higher the transportation costs from the market to the farm. It is also possible that households within a close proximity to market may avail of other possible services (e.g. extension services, price information, etc.) available in the district market centres. For this reason, we plotted expenditures on each input against the time taken to the nearest market (Figure 4.7) to illustrate the distribution of inputs with respect to the distance from markets. The statistics reveals that farm households located within a relatively closer proximity to permanent markets spend substantially more to purchase inputs from the market as compare to the households that are located further from the market.

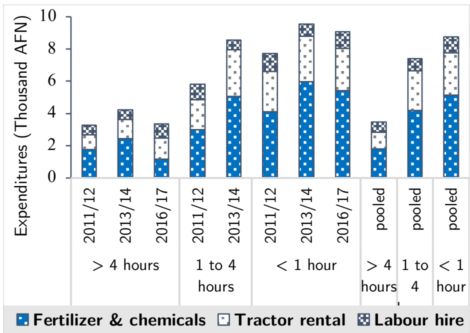



**Figure 4.7:** Input Expenditures (AFN) Over Time Taken to the Market
*Source*: Author's calculation of the ALCS Data

It can be noted that the average expenditures do not fluctuate much throughout time. Note that time taken to reach the nearest market was originally reported as time taken to reach the nearest permanent market by private vehicles/car.

## 4.6 Econometric Results and Discussion

We present empirical results in the following two sections: the estimation of household labour demand and test of separation in section 4.5.1, and the results for assessing the impact of transaction costs on market participation for labour, fertilizer and chemicals, and tractor rental markets in section 4.5.2.

### 4.6.1 Labour Demand and Testing of the Hypothesis of Separation

Before we formally test separability by running a multivariate regression of total labour demand over the household labour endowments (i.e. household size), we present Kernel-weighted regressions to show patterns and direction of linear relationship between household labour demand and labour endowments. Figure 4.8 illustrates the descriptive local polynomial regression of the household labour demand by the type of labour (i.e. household own labour, hired labour, and total labour demand) on the household labour endowments (e.g. household size). Even though the household own labour employed on the farm shows more variability when the household size exceeds 20 persons, with a default Kernel (Epanechikov) distribution and 95% confidence interval bands, the overall trend of the smooth polynomial shows that household own labour employed on the farm significantly increases in household size (Figure 4.8a), implying that larger households supply more labour to the farm work. In contrast, labour hiring decreases in household size (Figure 4.8b). Total labour demand (both own and hired labour) is also highly variable when the total household size is beyond 20 persons (Figure 4.8c), although overall demand for labour increases in household size.

If the separation holds (under perfect or competitive markets), we should not be able to observe a clear relationship between the total labour demand and household size (Dillon and Barrett, 2017). While this relationship does not formally signify rejection of the separation because underlying results are not conditioned on other covariates, it does signify that there exists a strong linear relationship between household endowments and the application of labour on the farm.

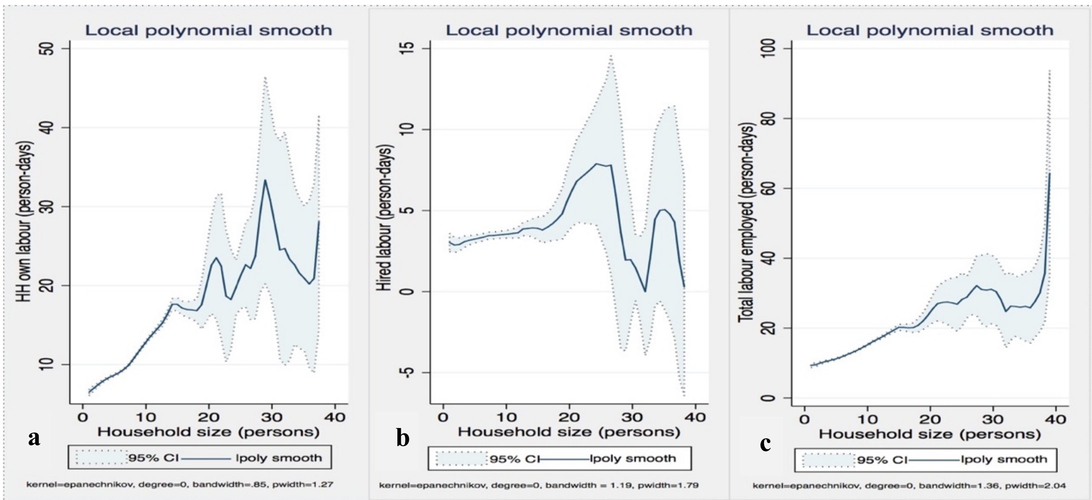



**Figure 4.8:** Local Polynomial Regression of a). HH own, b). Hired, and c). Total Labour (person-days) on HH Size (persons)

*Source:* Author's calculation of the ALCS Data

Table 4.6 reports the result of simplified Benjamin test (Equation 4.18) of the separation hypothesis with different specifications. The first column presents the simplest specification with household preference shifters such as the size of the household, and household composition variables along with the landholding. Columns 2 and 3 of Table 4.6 condition on other covariates such as transaction costs, household socio-demographic and socioeconomic characteristics, and interaction terms to pick up some essential elements of the labour demand and examine their implications on the separation hypothesis. All regressions control for geographical variations by including district fixed effects and standard errors clustered at the district FE level.

The estimated elasticity of household size (adults 14 years age or older)[18] with respect to the total labour days is 0.41 in the basic simplified specification, which is significantly different from zero at 1% significance level. A 1% increase in the household size increases total labour employed on the farm by 41%. The significance of household size implies the rejection of the hypothesis of separation that household production and consumption decisions are made independent of each other and that farmers treat all prices as given (under competitive markets). The magnitude of this elasticity can be interpreted as a rough indication of market failures that create dependency on household labour endowments. However these results cannot be interpreted as a test of the labour market failure specifically, as multiple failures are required to generate distortions in factor markets because relative (not absolute) prices determine allocation of resources (Dillon and Barrett, 2017).

Our results on rejecting the hypothesis of separation are in agreement with Dillon and Barrett (2017), LaFave and Thomas (2016), and Grimard (2000) for five Sub-Saharan countries (Ethiopia, Malawi, Niger, Tanzania, and Uganda), Java, and Côte d'Ivoire respectively using the same theoretical and empirical strategy. In contrast, Bowlus and Sicular (2003) for China, and Benjamin (1992) for Java do not reject separation.

The shares of prime female and elderly female (relative to the excluded group of elderly males) in household composition are negatively associated with labour demand, indicating that labour demand is decreasing in the share of females and elderly female. The share of age-sex groups and the log of household size jointly capture the composition and scale effects. As expected, the aged members of the household contribute less (or nothing) to the farm labour requirement (with aging their productivity may depreciate). The significance of household structure or composition variables is further confirmed by the F-

[18] We also regress the total labour demand on the full household size (including 0-14 age-sex group) and redefine the HH sex-age groups; the results presented in Table 4.B1 Appendix 4.B remain largely similar (separation rejected).

test for joint significance. The calculated test statistic is F (4, 385) = 132.83 with a probability of zero (reported at the bottom of Table 4.6) rejecting the null hypothesis that the household size and composition variables are simultaneously zero. Based on the joint significance of the household size and composition variables, the separation hypothesis is still rejected.

**Table 4.6:** Regression Results of the Household Demand for Farm Labour

| | Model (1) | Model(2) | Model (3) |
|---|---|---|---|
| *Dependent variable: Log of total (own and hired) labour in person-days* | | | |
| Log. HH size (adults) [A] | 0.411*** | 0.449*** | 0.517*** |
| | (0.018) | (0.020) | (0.042) |
| Prime HH male share [B] | 0.119* | 0.196** | 0.203** |
| | (0.071) | (0.079) | (0.079) |
| Prime HH female share [C] | -0.161** | -0.110 | -0.105 |
| | (0.081) | (0.091) | (0.091) |
| HH elderly female share [D] | -0.286** | -0.268** | -0.268** |
| | (0.123) | (0.134) | (0.134) |
| Log total land (Jeribs) | 0.231*** | 0.199*** | 0.198*** |
| | (0.014) | (0.013) | (0.013) |
| ICT equipment (1=own/access) | - | 0.080*** | 0.060 |
| | | (0.019) | (0.049) |
| Transport equipment (1=own/access) | - | 0.029** | 0.007 |
| | | (0.015) | (0.040) |
| Time taken to reach nearest market (1-4 hrs) | - | -0.064* | -0.038 |
| | | (0.035) | (0.059) |
| Time taken to reach nearest market (<1 hr) | - | -0.032 | 0.148*** |
| | | (0.030) | (0.052) |
| Log. distance to road | - | -0.021** | -0.029 |
| | | (0.011) | (0.022) |
| Log off-farm income (AFN) | - | -0.027*** | -0.027*** |
| | | (0.002) | (0.002) |
| HH head literacy (1=can read & write) | - | -0.040* | -0.038* |
| | | (0.023) | (0.022) |
| Log. HH head education (years) | - | 0.0001 | -0.0001 |
| | | (0.011) | (0.011) |
| HH head age (years) | - | 0.001 | 0.001 |
| | | (0.001) | (0.001) |
| ICT equipment (access=1) # Log. HH adults | - | - | 0.015 |
| | | | (0.036) |
| Transport equip. (access=1) # Log. HH adults | - | - | 0.017 |
| | | | (0.026) |
| Time taken to market (1-4h) # Log. HH adults | - | - | 0.022 |
| | | | (0.042) |
| Time taken to market (<1h) # Log HH adults | - | - | -0.136*** |
| | | | (0.036) |

| | | | Table 4.6 Continue |
|---|---|---|---|
| Log. Distance to road # Log. HH adults | - | - | 0.007 |
| | | | (0.015) |
| Wave 2 (2013/14) | -0.204*** | -0.183*** | -0.183*** |
| | (0.032) | (0.032) | (0.032) |
| Wave 3 (2016/17) | -0.083*** | -0.085*** | -0.084*** |
| | (0.029) | (0.029) | (0.029) |
| District FE | ✓ | ✓ | ✓ |
| Constant | 1.791*** | 1.901*** | 1.815*** |
| | (0.077) | (0.100) | (0.102) |
| F-test: [A]+[B]+[C]+[D]=0, t-statistic | 132.83 | 127.10 | 51.85 |
| F-test: [A]+[B]+[C]+[D]=0, p-value | 0.000 | 0.000 | 0.000 |
| R-squared | 0.244 | 0.265 | 0.266 |
| Observations | 21,189 | 21,189 | 21,189 |

*Note: Omitted categories: HH elderly male share, no access to ICT & transport equipment, 4 and more hrs for time to market, cannot read & write for literacy, and 2011/12 for wave. SE's clustered at districts FE level (in parentheses). Significance is indicated by *** p<0.01, ** p<0.05, * p<0.1.*

As expected, the impact of land size on household labour demand is positive and highly significant: a 1% increase in the Jeribs farmed would increase the total household labour requirements by about 23%. Other studies that found a significant relationship between the landholding and the total labour demand are Dillon and Barrett, (2017) for Ethiopia, Malawi, Niger, Tanzania, and Uganda, and Benjamin, (1992) for Indonesia.

Columns 2 and 3 include other controls (both independently and interacted with the household size), to assess if these additional controls diminish the magnitude or eliminate the statistical significance of the estimated relationship between log household size and labour demand estimated by the basic model in column 1. The null hypothesis of separation is rejected in all three cases (bottom of Table 4.6). We find that only ownership of ICT by farm households and the log of off-farm income and distance to road were found to have significant positive and negative impact on the total labour days employed on farm respectively. In general, the insignificance of the interaction terms implies that there are no meaningful differences in transaction cost variables (i.e. ICT equipment, transport equipment, and distance to road) in the way that household endowments relate to the labour demand. Therefore, rejection of the separation hypothesis is not driven by the heterogeneities in these variables.

### 4.6.2 Market Participation under Transaction Costs

We present econometric results in two subsections: results from the reduced form equations in subsection 4.6.2.1, and results from the structural equations in the subsection 4.6.2.2.

#### 4.6.2.1 Results from the Estimation of the Reduced Form Equations

Following the discussion in section 4.4, we assume that the ownership or use of ICT and transport equipment is correlated with household unobservable characteristics that influence market participation decisions. We allow ownership of ICT and transport equipment to be endogenous to control for such unobserved heterogeneity using a control function approach. Table 4.7 presents the results of the reduced form regressions from a Probit of the endogenous variables on the instrumental variables (IVs) conditional on other covariates. All estimates in Table 4.7 are Average Partial Effects (APE), with standard errors in parentheses clustered at the FE level. We include provincial fixed effects.
All instrumental variables (indicated by the star sign in Table 4.7) have the expected significant impact on the endogenous variables. They satisfy the orthogonality conditions, implying that IVs are directly and significantly correlated with the endogenous variables but affect dependent variables in the structural models only through the inclusion of the endogenous variables and the computed generalized residuals from the reduced form. Recall that in the Control Function (CF) approach, the analysis involves the estimation of the generalized residuals from the reduced form which are then included in the estimation of structural equations. It is plausible to believe that any leftover endogeneity after using the CF approach will be uncorrelated with the other covariates in the structural model (Ricker-Gilbert et al., 2011).

The APE associated with the access to electricity instrument could be interpreted to mean that a shift of 1 (from having no access to having access to electricity) increases the probability to use ICT equipment by about 6 percentage points. Similarly, an increase of 10,000 Afghani in the off-farm income of other farmers in the community (a proxy for off-farm employment opportunities within the community) is associated with an increase of 0.4 percentage points in ICT use (a very small effect, given that over 80% own ICT), and 0.3 percentage points increase in probability of owning transport equipment.

**Table 4.7:** APE from Reduced form Estimation of Endogenous Variables (Ownership of ICT and Transport Equipment)

| VARIABLES | ICT equipment (binary, 1=own) | Transport equip. (binary, 1=own) |
|---|---|---|
| Electricity (1=access, 0 otherwise)* | 0.057*** | - |
| | (0.017) | - |
| Road/bridge Project (yes=1, 0 otherwise)* | - | 0.033** |
| | - | (0.014) |
| Neighbour off-farm income (10K AFN)* | 0.004** | 0.003** |
| | (0.002) | (0.001) |
| Time taken to reach market (1-4 hours) | 0.004 | 0.005 |
| | (0.011) | (0.020) |
| Time taken to reach market (<1 hour) | 0.032*** | 0.038 |
| | (0.011) | (0.025) |
| Log. distance to road (km) | -0.009* | -0.017 |
| | (0.005) | (0.014) |
| Log. off-farm income (AFN) | 0.004*** | 0.006*** |
| | (0.001) | (0.001) |
| Log. total land (Jeribs) | 0.036*** | 0.061*** |
| | (0.007) | (0.009) |
| Log. HH size (persons) | 0.060*** | 0.128*** |
| | (0.010) | (0.011) |
| HH head literacy (1=can read & write) | 0.061*** | 0.085*** |
| | (0.009) | (0.018) |
| Log. HH head education (years) | 0.018*** | 0.009 |
| | (0.005) | (0.008) |
| Log. HH head age (years) | 0.069** | -0.011 |
| | (0.001) | (0.002) |
| HH head age squared | -2e-5** | -8.4e-06 |
| | (0.000) | (0.000) |
| Log. livestock (number) | 0.007** | 0.025*** |
| | (0.003) | (0.005) |
| Log. electricity cost (AFN)* | 0.007*** | - |
| | (0.002) | - |
| Wave 2 (2013/14) | -0.067*** | -0.005 |
| | (0.016) | (0.031) |
| Wave 3 (2016/17) | -0.018 | 0.019 |
| | (0.018) | (0.031) |
| Province FE | yes | yes |
| Pseudo R-square | 0.210 | 0.234 |
| Observations | 19,042 | 19,042 |

*Note: Omitted categories for factor variables are: no access to electricity, no road/bridge rehabilitation/recondition projects in the community for road/bridge projects, cannot read & write for literacy, more than 4 hours for time taken to reach the nearest permanent market, and 2011/12 for wave. Provincial fixed effects are included in the regression to control for regional variation. Significance is indicated by *** p<0.01, ** p<0.05, * p<0.1 and robust standard errors in parentheses. Variables labelled with stars indicate that they are instrumental variables.*

Households in communities where road or bridge reconstruction and rehabilitation projects were implemented appear more likely to own transportation equipment: construction or rehabilitation projects within the last 12 months in the community increase the probability of owning transport assets by 3.3 percentage points. Improved roads and bridges are likely to improve access and increase the benefits of owning transport equipment. Other standard variables that are statically significant have the expected effect on the ownership or use of ICT and transport equipment.

#### 4.6.2.2 Results from the Estimation of Structural Equations

Results of the structural equations estimated by the Lognormal Double Hurdle (LDH) model, with a Probit of the decision to participate in input markets and a truncated regression for the extent of expenditures, are reported in Tables 4.8 for participation in hired labour markets and Table 4.9 for fertilizers and chemicals, and tractor rental markets. Column 1 of Table 4.8 reports non-instrumental variable estimation of hired labour while column two reports instrumental variable estimation of the hired labour with endogenous transaction costs. Columns The estimated APE of the explanatory variables from the hurdle 1 (probit) of LDH model presents the probability of participation and a marginal change in the extent of expenditures in hurdle 2 (truncated regression).

Endogeneity is detected if the generalized residuals (indicated by GR in the table) are statistically significant in the structural regressions (presented in Table 4.8 and 4.9). The estimates on the generalized residuals for both ICT and transport equipment turn out to be negative and significant for participation in fertilizer and chemical markets, confirming that both are endogenous. The negative sign on the generalized residual implies that error terms of the two models (reduced form and structural) are negatively correlated with each other. This implies that the unobserved factors captured by the generalized residuals increase the probability of owning ICT and transport equipment but reduce the probability of participation in the market. Similarly, with a negative and significant estimate of the generalized residual for ICT, endogeneity of ICT is detected for the probability of hiring tractor. However, endogeneity was not detected in the regression analysing the probability of households participating in the labour market, we therefore treat our non-IV estimates for the hired labour (Table 4.8) as our primary results to avoid concerns that performing IV estimation may inflate the asymptotic variance of the estimator when endogeneity is not detected as stated in Wooldridge (2002) and Tadesse and Bahiigwa (2015).

Since the reduced form Probit is a nonlinear model, we are unaware of any methods to test for the strength of IVs in this context. Following Ricker-Gilbert et al. (2011), Liverpool-Tasie (2014), and Amankwah et al., (2016) we rely on the partial correlation between the IVs and the endogenous variables in our reduced form model. All the IV's are significant at 5% implying they are partially correlated with the endogenous variable of ICT transport equipment. It is, as argued in Section 4.4 (under estimation strategy), unlikely to be directly correlated with our dependent variables in the structural models. Controlling directly for household characteristics and district fixed effects in our structural models, we believe the only remaining pathway that the instruments affect participation is through the channel of IVs. Thus, we feel confident that the instruments are exogenous in the structural model and their strong partial correlation with the endogenous variable reveal their strength.

The results in Table 4.8 and 4.9 reveal that household's possession of the ICT equipment (i.e. mobile phone, TV, radio, and internet services) significantly increases the likelihood of hiring labour, purchasing fertilizers and chemicals, and hiring tractor by 3.3, 22 and 21 percentage points respectively suggesting that the ownership of ICT as a fixed transaction cost helps facilitate entry in markets by providing new and timely information that can reduce search and information costs. Search and information costs are often considered to be fixed transaction costs that influence market entry decisions (Goetz, 1992; Alene et al., 2008; Omiti et al., 2009).

Similar findings are observed by Randela et al., (2008) who concluded that the more information on marketing available to households, the lower are transaction costs hence a higher rate of market participation. Chowdhury (2006) finds a strong connection between household use of mobile phones and their marketing decisions and suggests that a reduction in information cost in the form of access to a telephone may change the functioning of markets and market participation. On the contrary, Alene et al., (2008) and Ouma et al., (2010) found that access to communication assets have positive but insignificant effects on market participation in Kenya and Central Africa (Rwanda and Burundi) and argued that communication assets are perhaps less useful in facilitating transactions if there is no viable market information service. In assessing the impact of mobile phones on farmers' marketing decisions in Ethiopia, Tadesse and Bahiigwa (2015)-

**Table 4.8:** Log-normal Hurdle Estimates of HH Participation & Expenditures: Hired Labour

| | Non-IV estimation | | IV estimation | |
|---|---|---|---|---|
| | Hurdle 1 | Hurdle 2 | Hurdle 1 | Hurdle 2 |
| | (Probit,1=use) | Log. Expenditure | (Probit,1=use) | Log. Expenditure |
| VARIABLES | APE | APE (y>0) | APE | APE (y>0) |
| GR (ICT) | - | - | 0.015 | - |
| | | | (0.028) | |
| GR (Transport Equip) | - | - | -0.064 | - |
| | | | (0.044) | |
| ICT equipment (1=own) | 0.033*** | - | 0.007 | - |
| | (0.010) | | (0.050) | |
| Transport assets (1=own) | 0.044*** | - | 0.147** | - |
| | (0.007) | | (0.068) | |
| Log. distance to rd (km) | -0.006 | -0.022 | -0.005 | -0.022 |
| | (0.007) | (0.023) | (0.007) | (0.023) |
| Time to market (1-4h) | 0.020 | -0.069 | 0.021 | -0.069 |
| | (0.018) | (0.066) | (0.018) | (0.066) |
| Time to market (<1h) | 0.037** | 0.013 | 0.034** | 0.013 |
| | (0.016) | (0.049) | (0.017) | (0.049) |
| Log. total land (Jeribs) | 0.066*** | 0.444*** | 0.060*** | 0.444*** |
| | (0.006) | (0.024) | (0.009) | (0.024) |
| Log. off-farm inc (AFN) | 0.004*** | -0.011*** | 0.003*** | -0.011*** |
| | (0.001) | (0.003) | (0.001) | (0.003) |
| Log. household size (N) | -0.031*** | 0.164*** | -0.044*** | 0.164*** |
| | (0.009) | (0.032) | (0.013) | (0.032) |
| HH head literacy(1=yes) | 0.037*** | -0.072 | 0.029** | -0.072 |
| | (0.011) | (0.044) | (0.014) | (0.044) |
| Log. head edu (yrs) | 0.011** | 0.051** | 0.010* | 0.051** |
| | (0.005) | (0.021) | (0.005) | (0.021) |
| Land type (1=all irrigated) | 0.023* | 0.226*** | 0.023* | 0.226*** |
| | (0.012) | (0.043) | (0.012) | (0.043) |
| Landscape 2 (valleys) | -0.025 | -0.063 | -0.024 | -0.063 |
| | (0.019) | (0.056) | (0.019) | (0.056) |
| Landscape 3 (open plain) | 0.008 | -0.052 | 0.008 | -0.052 |
| | (0.014) | (0.049) | (0.014) | (0.049) |
| Log. No of livestock (N) | -0.003 | 0.019 | -0.006 | 0.019 |
| | (0.004) | (0.014) | (0.004) | (0.014) |
| Oxen (binary, 1=own) | -0.011 | 0.046 | -0.011 | 0.046 |
| | (0.012) | (0.033) | (0.012) | (0.033) |
| Log. tractor/threshers (N) | 0.038 | 0.298*** | 0.039 | 0.298*** |
| | (0.027) | (0.103) | (0.027) | (0.103) |
| Wave 2 (2013/14) | -0.129*** | 0.057 | -0.129*** | 0.057 |
| | (0.016) | (0.064) | (0.016) | (0.064) |
| Wave 3 (2016/17) | -0.109*** | 0.356*** | -0.112*** | 0.356*** |
| | (0.018) | (0.058) | (0.018) | (0.058) |
| District FE | ✓ | ✓ | ✓ | ✓ |
| Pseudo R-Square | 0.236 | - | 0.236 | - |
| Observations | 20,436 | 5,334 | 20,436 | 5,334 |

*Notes: Omitted categories: no access to ICT & transport equipment, more than 4 hrs for time to market, irrigated & rain-fed for land quality, cannot read & write for literacy, hills & valleys for landscape, & 2011/12 for wave. SE's (in parenthesis) are clustered in districts FE and significance is indicated by *** p<0.01, ** p<0.05, * p<0.1.*

**Table 4.9:** LDH Estimates of HH Participation & Extent of Expenditures: Fertilizer & Chemicals, and Tractor Rental

| | Fertilizer & chemicals | | Tractor Hire | |
|---|---|---|---|---|
| | Hurdle 1: (Probit, 1=use) | Hurdle 2 Log. Expenditure | Hurdle 1 (Probit,1=use) | Hurdle 2 Log.Expenditure |
| VARIABLES | APE | APE (y>0) | APE | APE (y>0) |
| GR (ICT) | -0.083** | - | -0.110*** | - |
| | (0.033) | | (0.031) | |
| GR (Transport equip) | -0.068* | - | 0.045 | - |
| | (0.035) | | (0.040) | |
| ICT equipment (1=own) | 0.224*** | - | 0.214*** | - |
| | (0.070) | | (0.059) | |
| Transport equip (1=own) | 0.140** | - | -0.031 | - |
| | (0.057) | | (0.065) | |
| Log. distance to Rd (km) | 0.005 | -0.063*** | -0.021*** | 0.003 |
| | (0.007) | (0.017) | (0.006) | (0.014) |
| Time to market (1-4h) | 0.002 | -0.059 | 0.008 | 0.129*** |
| | (0.019) | (0.051) | (0.019) | (0.038) |
| Time to market (<1h) | 0.015 | -0.05 | 0.065*** | 0.056 |
| | (0.015) | (0.049) | (0.018) | (0.037) |
| Log. total land (Jeribs) | 0.036*** | 0.538*** | 0.080*** | 0.600*** |
| | (0.008) | (0.023) | (0.009) | (0.022) |
| Log. off-farm inc (AFN) | -0.003*** | -0.016*** | -0.003*** | -0.008*** |
| | (0.001) | (0.002) | (0.001) | (0.002) |
| Log. HH size (N) | 0.003 | 0.142*** | -0.009 | 0.123*** |
| | (0.012) | (0.023) | (0.011) | (0.021) |
| HH head literacy(1=yes) | -0.003 | 0.044 | 0.004 | 0.065*** |
| | (0.012) | (0.028) | (0.012) | (0.025) |
| Log head education (yrs) | 0.005 | 0.013 | 0.010** | -0.02 |
| | (0.005) | (0.014) | (0.005) | (0.013) |
| Land type (1=all irrigated) | 0.168*** | 0.413*** | 0.045** | 0.096*** |
| | (0.018) | (0.045) | (0.018) | (0.032) |
| Landscape 2 (valleys) | -0.002 | 0.035 | 0.004 | -0.008 |
| | (0.017) | (0.034) | (0.019) | (0.041) |
| Landscape 3 (open plain) | 0.0314* | 0.128*** | 0.103*** | 0.022 |
| | (0.017) | (0.037) | (0.022) | (0.031) |
| Log. No of livestock (N) | 0.007* | 0.002 | 0.003 | 0.003 |
| | (0.004) | (0.010) | (0.005) | (0.010) |
| Oxen (binary,1=own) | 0.036*** | 0.120*** | -0.141*** | -0.090*** |
| | (0.010) | (0.030) | (0.017) | (0.033) |
| Log. tractor/threshers (N) | -0.043 | 0.355*** | -0.192*** | 0.126** |
| | (0.031) | (0.059) | (0.039) | (0.054) |
| Wave 2 (2013/14) | 0.011 | 0.222*** | -0.012 | 0.224*** |
| | (0.019) | (0.045) | (0.015) | (0.047) |
| Wave 3 (2016/17) | 0.031** | 0.278*** | -0.041** | 0.237*** |
| | (0.015) | (0.045) | (0.016) | (0.040) |
| District FE | ✓ | ✓ | ✓ | ✓ |
| Pseudo R-Square | 0.394 | . | 0.401 | . |
| Observations | 19,042 | 13,133 | 19,443 | 11,079 |

*Notes: Omitted categories for factor variables are: no access to ICT & transport equipment, more than 4hrs for time to market, irrigated & rain-fed combined for land quality, cannot read & write for literacy, hills & valleys for landscape, and 2011/12 for wave. SE's (in parenthesis) are clustered in districts and significance is indicated by *** p<0.01, ** p<0.05, * p<0.1.*

-find mixed results; ownership of mobile phones may be useful for certain farmers in making marketing decisions in some circumstances, while in other areas mobile phones do not seem to be an important channel to access price information.

The likelihood of households to hire labour increases with ownership of transport equipment by 4.4 percentage points (Table 4.8). Perhaps own transport provides cheaper opportunity to farm household to pick labourers from the market or district centres and transport them to their farms. Similarly, ownership of transport equipment is associated with a higher propensity of market participation in labour and fertilizer and chemicals markets. Similarly, a shift from no transport equipment to owning transport equipment increases the likelihood of fertilizers and chemicals use by 14 percentage points (Table 4.9). Household endowments of transport assets mitigate transportation, communication and information costs (i.e. own transport equipment is maybe more cost effective compared to public transport) and reduce obstacles to entering the market (Goetz, 1992; Key et al., 2000). Our findings agree with those of Alene et al., (2008) who found significant impact of transport assets on the probability of fertilizer use in Kenya. In contrast, ownership of transport equipment was found to have no significant impact on the probability of hiring a tractor (Table 4.9). This is a plausible conclusion as hiring tractor by households may not require the use of any other means of transport. Hence, we find no significance impact of owning transport equipment on tractor hire.

Even though higher poor accessibility associated with larger distance of farms from the nearest all-season roads do not appear to significantly affect the probability of household's participation in the chemical and fertilizer market, a 1% increase in distance to roads would reduce expenditures on fertilizers and chemicals by 0.06%. An increase of 1 in the log of distance to road (equivalent to an increase of 172%[19]) decreases the probability of households to hire tractors by 2.1 percent points (Table 4.9), holding all other variables constant. Remoteness from the roads appears to have no significant impact on labour hire in both stages (probability and expenditures). Distance from the roads may not largely

[19] A unit increase in log of distance to road (log of distance to road) corresponds to multiplying the land by e ≅ 2.71828, therefore the absolute change in distance to road is 2.71828, thus the net difference simplifies to 1.71828. Putting that in percentage terms, it's is about 172% increase in the distance to road variable.

vary within the districts, so controlling for district fixed effect may possibly render this variable insignificant in the case of some inputs. Most studies assessing farm marketing decisions established a negative relationship between poor road accessibility and market participation and extent of use (Ricker-Gilbert et al., 2011; Winter-Nelson and Temu, 2005).

While time taken to reach the nearest market turns out to have no significant impact on the probability of households to participate in the fertilizer and chemical markets and expenditures, a decrease in time taken to reach the nearest permanent market significantly increases the probability of households to hire labour and tractor for farming activities (by about three and seven percent points respectively for closest households i.e. farms located within one or less than hours to the market). While shorter distance from the market may increase the probability and expenditures on input, there is a possibility that households located in rural areas further from the market may have larger landholdings, and therefore use more inputs. In addition, small local markets are located in the district centres, and since we control for district fixed effects in our regressions, there is a possibility that district level variables (such as time taken to reach market) may become insignificant. These results are consistent with Ouma et al., (2010) and Liverpool-Tasie (2014) who found negative relationship between time taken to reach market and market participation reinforcing the argument that poor market access for households located in remote areas raises costs associated with marketing and information.

Estimated APE's and significance of the landholding size are similar for all inputs, that is larger household land endowment is associated with higher probability of the input use and expenditures for each of the three inputs. An increase of 1 in the log of land (equivalent to an increase of 172%) increases the probability of market participation by 6.6, 3.6, and 7.9 percent points in hired labour, fertilizer and chemicals, and tractor rental markets respectively, holding all other variables constant. The elasticity associated with the land variable in the second hurdle (extent of expenditures) indicate that a 1% increase in land raises expenditures by 0.44%, 0.54%, and 0.60% on hiring labour, chemical and fertilizers, and tractor hire respectively. Our conclusions that land size is associated with the use of higher input usage is consistent with most of the studies in the literature that assed household marketing decisions (Liverpool-Tasie, 2014; Mather et al., 2013; Ricker-Gilbert et al., 2011).

A negative and statistically significant relationship is observed between off-farm income and the probability of market participation and extent of expenditures for all three input markets. A unit increase in the log of off-farm income is associated with a decrease of 0.4, and 0.3 percentage points in the probability of participation labour and fertilizers and chemicals, and tractor rental markets respectively. Similarly, a 1% increase in off-farm income decreases expenditures on hiring labour, chemicals and fertilizers, and renting tractor for farming activities by 1.1%, 1.6%, and 0.8% respectively. Our findings of negative impact of non-farm income on market participation contradicts with those of Alene et al. (2008) in Kenya, however Verkaart et al. (2017) and Makhura (2001) find similar results to ours. Winter-Nelson and Temu (2005) found no significant impact of non-farm income on input use in Tanzania. Households with larger off-farm income may divert resources such as labour to off-farm income to diversify employment risks and thus reduces reliance on farm activities (Ahmadzai, 2017; Mishra et al., 2004). Rao and Qaim (2013) finds that household with employment outside farming sector significantly reduces the quantity of hired labour.

Except for labour hire, household size was found to have no significant effect on the probability of participation in the input markets. A change of one in the log of the household size reduces the likelihood of hiring labour by 3.1 percentage points. The probability to hire labour is decreasing in family size, suggesting that households with a larger endowment rely largely on their own labour. Though household size is insignificant determinant for the entry in fertilizers and tractor markets, larger households significantly spend more on inputs than smaller households (Table 4.8 and 4.9). A 1% increase in household size is associated with 0.16%, 0.14%, and 0.12% increase in expenditures on hiring labour, fertilizer and chemicals, tractor rental, and labour hire respectively. The positive and significant marginal effect of household size may imply that larger families own larger farmers and consume more, thus purchasing more inputs from the markets. Our findings that household size increases the expenditures on inputs are in conformity with Abdullah et al., (2017) but Rao and Qaim (2013) find no significant impact of household size on the likelihood of household to hire, but find a positive and significant relationship between household size and the quantity of hired labour.

Other standard household socio-demographic and socio-economic characteristics were also found to have the expected influence on market participation and input expenditures. While age is not an important factor, household head education and literacy were found to have a significant positive influence on both the propensity of market participation and magnitude of expenditures in the case of some inputs. Our findings that education increases market participation match with those of Randela et al. (2008), Martey et al. (2012) and, Liverpool-Tasie (2014).

Land characteristics such as the types of land and landscape are important determinants of both the decision to purchase inputs and the amount of spending on inputs. Households operating all irrigated land (relative to households that own a combination of both irrigated and rain-fed land) purchase substantially more inputs from the market and spend more money on inputs. Farmers are likely to grow different crops depending on their water requirements (some crops may require more irrigation than others that may not be planted in rain-fed lands) and therefore this limitation may cause the need for input usage to decline. Farmers purchase more inputs if they own flat or plains land as compared to farmers on slopes (i.e. hills and valleys). Altitude and slope of the farm land affect physical conditions of the farm land and therefore may limit the application of some inputs, for instance, it may not be technically feasible to use tractor in farm land with greater slope.

A positive and statistically significant relationship is observed between the ownership of livestock at the farm and the probability of fertilizer and chemical use, consistent with Liverpool-Tasie (2014), but is insignificant for other inputs. While ownership of oxen increases likelihood of use and expenditures on fertilizer and chemicals, it significantly reduces the probability of households to hire tractor by 14 percentage points. In addition, a 1% increase in the number of oxen owned by the farm households reduces the expenditures associated tractor rental by about 9%. This large negative effect of oxen ownership on tractor rental could be due the fact that households use oxen for farm activities such as ploughing as a substitute to tractor and therefore reduces the need to hire tractor.

Ownership of tractors/threshers by household significantly decreases the likelihood of hiring tractors but have no significant impact on the propensity to use fertilizers and chemicals or hire labour. This is not unusual as households that own tractors may not need to hire tractors. However, households that have tractors/threshers spend significantly more on all three inputs. While this is plausible for use of fertilizers/chemicals and hired labour, as such farms may be more commercial, the association with expenditure on tractor hire is difficult to explain other than as a peculiarity in the data.

In general, the time fixed effects in the model reveal lower probability of participation (except for fertilizers and chemicals) and significantly higher spending on inputs in the recent survey years. Given that we use a repeated cross-sectional survey where each farm is observed only once, it is difficult to comment on the magnitude and trend of input use over time, it is plausible to assume that this could simply be as a result of fluctuations in the inflation.

### 4.7 Robustness and Specification Tests

We carried out a number of specification tests to ensure that the statistical model is appropriately chosen to best fit our data. The most widely used statistical models for censored data are double hurdle and the standard tobit models which is nested in the double hurdle model. For this reasons, we first test Cragg type independent double hurdle truncated normal model against standard tobit using a Log-Likelihood Ratio (LR) by Greene (2002):

$$LR\ Test\ Statistic = -2[\ln L_T - (\ln L_P + \ln L_{TR})] \quad 4.29$$

Where $L_T$ is the likelihood for the restrictive tobit model under the null hypothesis, $L_P$ and $L_{TR}$ are the likelihood for the hurdle 1 (probit model analysing participation) and hurdle 2 (the truncated regression model analysing extent of participation) of the Cragg type double hurdle model. With independent error terms, the log-likelihood of the truncated Cragg type double hurdle model is equivalent to the sum of the log-likelihoods of the probit and the truncated regressions (Rao and Qaim, 2013). The computed statistics of the log-likelihood ratio test for each of the three models analysing participation in fertilizer and chemicals, labour, and tractor rental markets reject the null hypothesis at

1% significance indicating that the Cragg type double hurdle model is strictly preferred to the restricted tobit model (Table 4.10).

**Table 4.10:** Robustness/Specification Tests for Model Selection

| **Log likelihood-ratio test for nested models: Truncated-normal DH vs Tobit** | | | | |
|---|---|---|---|---|
| H0: Nested model (Tobit) Specification is valid | | | | |
| H1: Double Hurdle specification is valid | | | | |
| Input market | Test-Stat | Critical value | P-value | Decision |
| Fertilizer & chemicals | 16,996.16 | $\chi2(0.05, 310)=352.06$ | 0.000 | Reject H0, DH is valid |
| Tractor rental | 11,703.96 | $\chi2(0.05, 333)=376.55$ | 0.000 | Reject H0, DH is valid |
| Labour hire | 5,180.13 | $\chi2(0.05, 360)=405.24$ | 0.000 | Reject H0, DH is valid |
| **Vuong's (1989) closeness test for non-nested models: Truncated-normal vs LDH** | | | | |
| H0: Truncated-normal & LDH models offer an equivalent representation of the data | | | | |
| H1: Lognormal double hurdle model is closer | | | | |
| Input market | Ln Ratio | se | P-value | Decision |
| Fertilizer & chemicals | 5.706 | 0.029 | 0.000 | LDH is closer |
| Tractor rental | 4.648 | 0.030 | 0.000 | LDH is closer |
| Labour hire | 1.918 | 0.023 | 0.000 | LDH is closer |
| **Vuong's (1989) closeness test for non-nested models: Sample selection vs LDH** | | | | |
| H0: Sample selection and lognormal hurdle models offer an equivalent representation | | | | |
| H1: Lognormal hurdle model is closer | | | | |
| Fertilizer & chemicals | 0.443 | 0.002 | 0.000 | LDH is closer |
| Tractor rental | 0.561 | 0.004 | 0.000 | LDH is closer |
| Labour hire | 0.149 | 0.002 | 0.000 | LDH is closer |

Next, we use Vuong's closeness test for non-nested or non-overlapping models to distinguish between the truncated normal double hurdle model and log-normal hurdle[20] model. Because truncated normal and lognormal hurdle models are non-nested models (Hsu and Liu, 2008), we use the Vuong (1989) test for non-nested models to evaluate which model provides a closer representation of the data.

Vuong's test is a likelihood-ratio-based test that compares non-nested models in terms of the difference in their respective Kullback-Leibler (KL) distance from the (unknown) "true" model. Suppose the KL distance between two competing models is given by the following equation:

$$LR(A, B) = Log\ L(A) - Log\ L(B) \tag{4.30}$$

[20] Lognormal hurdle does not nest standard tobit model by construction, therefore we can't test lognormal model against standard tobit. We therefore use a likelihood ratio test to first test truncated normal hurdle model against standard tobit.

The null hypothesis suggests that there is no difference between the two models. The test statistic is calculated as:

$$LZ = \frac{LR(A,B)}{\sqrt{n}\ \omega} \qquad 4.31$$

Where ω denotes the variance of pointwise log-likelihood ratio, and $n$ is the sample size. Large positive (negative) values of the computed test statistics are taken as evidence in favour of model A[21].

Based on the non-nested LR test procedure of Vuong, the computed statistics for each of the three inputs reject the null hypothesis that both truncated normal and lognormal hurdle models are an equally good fit for the data and indicates that the lognormal hurdle model is the closest true model. Even though the computed Vuong's statistics are not as large when testing the lognormal hurdle model against the sample selection model, they are significant and indicative[22]. The specification tests therefore reveal that Cragg type truncated normal two-step specification is preferred to the standard tobit, and lognormal hurdle is preferred to the Cragg type truncated normal model and sample selection model (Table 4.10).

### 4.8 Conclusion and Discussion

We test for market failures by testing the hypothesis of separation in the household demand model for the farm labour. Our estimates of the household labour demand show that household's production and consumption decisions are not separable from each other suggesting evidence for the existence of potential market failures in Afghanistan. In theory, under complete and competitive markets separation holds implying that household production and consumption decisions are independent of each other and farm households act as profit maximizers such that households first aim to maximize production, and then make consumption decisions conditional on the profits and income from production. On the contrary, the farm households face imperfections if production and consumptions

[21] When testing lognormal DH against truncated-normal hurdle, our model A is the lognormal and B is the truncated normal hurdle. When testing lognormal hurdle against sample selection model, our model A is the lognormal hurdle and model B is the sample selection specification.

[22] For comparison, we also estimated a sample selection model (results are presented in Table 4.B2 in the Appendix 4.B). The results are largely similar to those of our main results presented in Tables 4.8 and 4.9.

decisions are non-separable (i.e. labour allocations in production are significantly affected by the household preference shifters such as endowments of labour). Hence, the rejection of separation therefore implies that household labour demand is strongly influenced by its endowment of own labour (i.e. household size) which could therefore be interpreted to mean that there exist potential market failures.

Given the evidence for market failures, we then look at whether improving market access by reducing transaction costs would improve farmer's market participation as a potential strategy to address market failures. The results have revealed that transaction costs are important determinants of smallholder participation in the input markets. Ownership of ICT and transport assets by farm households reduces search and information cost and therefore significantly increase the probability of participation in input markets. Locational characteristics were found to be important in explaining market participation. These characteristics depict differences in variable transaction costs to transport inputs from the markets of agricultural crops. Farmers with better access to markets and roads are also more likely to participate in input markets. Moreover, farmers living in communities with better road access and density and within a close radius of markets, were found to spend more on inputs. The significant impact of both fixed and proportional transaction costs in this study reveals that the existence of high transactions costs are responsible for lower market participation, and in some instances could force remote peasant smallholders to opt for self-sufficiency instead of market participation. Identification through a control function approach confirm endogeneity in ownership of communication and transport equipment in the models analysing household market participation in fertilizer and tractor rental markets, but not in the model analysing labour market participation.

Standard factors such as household socio-demographic and socio-economic factors were also observed to have an important influence on household marketing decisions. Household size, literacy and education level, land endowments, off-farm income, and ownership of farming assets such as tractors, oxen and the number of livestock at the farm are among important determinants of household's decisions to participate in market and extent of expenditures. Characteristics of farms such as the type of land and landscape were also found to have implications for market decisions.

This study provides hints on the critical implication of transaction costs on market participation. One area of policy intervention that can be suggested from the findings of this study is that future policies geared towards agriculture commercialization should involve providing viable and timely information on market prices, technical advantages of using modern inputs and other important information through media, so that farmers communication assets are effectively used in accessing market information. This is particularly important as communication assets are maybe less useful in facilitating transactions if there is no viable market information service available through public or private media. Other means such as publishing price information through local newspapers may also help facilitate access to markets and mitigate search costs.

In general, a market-oriented agricultural policy would help improve farmers market participation by improving access to market information, facilitating transportation, addressing institutional weaknesses, and improve public and commercial input distribution systems. Improved access to agriculture extension services particularly in remote areas to assist farmers understand the advantages of using modern agriculture inputs and providing best practices may also enhance factor market participation. Collective action through cooperatives and farmer organizations may also enable farmers, particularly resource-poor farmers to share their resources, achieve economies of scale, and increase efficiencies in accessing local markets. Another possible approach could be contract farming to ensure surplus production is sold to the market and farmers gain sufficient cash money to purchase inputs from the market. Contract farming, in some cases, can also help farmers to trade some of their surplus production for agriculture inputs that can be used in the next planting seasons. Another general recommendation is that future government policy instruments that aim to incentivize Investment in rural infrastructure development such as roads, transportation facilities, and other means to stabilize input supply chain and distribution systems can also improve market participation and avert some of the negative consequences due to market imperfections or failures.

As the findings of this study suggest that there exist potential market failures, it is critical to notice that the theoretically-grounded test carried out in this study to test separation relies on the labour market transactions and therefore it is difficult to conclude which specific input markets may actually fail. In order to better understand and address the perceived market failures, further research is required to identify precisely the drivers and sources of market failures and the specific markets that are failing.

As transaction costs are "hidden" or in many instances not directly observed, and therefore most studies including this study use proxy measures to assess their impact on the household marketing decisions. One potential area to improve this research is to collect better data on transactions to help in quantifying the actual transaction costs incurred such as search, information, transport, other costs related to bargaining and contract enforcement.

# APPENDIX TO CHAPTER IV

## Appendix 4.A: Detailed Descriptive Analysis

**Table 4.A1**: Pearson χ2 Test for Selected HH Characteristics by Input type

| VARIABLE | | Non-users | Users | | Pearson x2 (p-val) |
|---|---|---|---|---|---|
| Fertilizer and chemicals | | | | | |
| ICT equipment | Don't own | 31.74 | 11.63 | 17.32 | 1,215.1***(0.000) |
| | Own | 68.26 | 88.37 | 82.68 | |
| Transport equipment | Don't own | 56.91 | 49.34 | 51.48 | 98.74*** (0.000) |
| | Own | 43.09 | 50.66 | 48.52 | |
| Time taken to reach market | >4 hours | 27.75 | 11.74 | 16.27 | 854.56*** (0.000) |
| | 1-4 hours | 19.92 | 19.65 | 19.72 | |
| | <1 hour | 52.34 | 68.61 | 64.01 | |
| HH head literacy | Cannot read & write | 76.14 | 67.56 | 69.99 | 150.68*** (0.000) |
| | can read & write | 23.86 | 32.44 | 30.01 | |
| Land type | Irrigated & rain-fed | 60.63 | 14.75 | 27.73 | 4,513.60***(0.000) |
| | All irrigated | 39.37 | 85.25 | 72.27 | |
| Landscape | Hills and valleys | 61.49 | 27.35 | 37.01 | 2,148*** (0.000) |
| | Valleys only | 12.52 | 22.45 | 19.64 | |
| | Open plain | 25.99 | 50.20 | 43.35 | |
| Oxen | Don't own | 73.82 | 83.99 | 81.12 | 290.12*** (0.000) |
| | Own | 26.18 | 16.01 | 18.88 | |
| Tractor rental | | | | | |
| ICT equipment | Don't own | 25.54 | 10.90 | 17.32 | 780.67*** (0.000) |
| | Own | 74.46 | 89.10 | 82.68 | |
| Transport equipment | Don't own | 66.20 | 39.98 | 51.48 | 1,435.23***(0.000) |
| | Own | 33.80 | 60.02 | 48.52 | |
| Time taken to reach market | >4 hours | 25.37 | 9.16 | 16.27 | 1,032.54***(0.000) |
| | 1-4 hours | 19.13 | 20.18 | 19.72 | |
| | <1 hour | 55.50 | 70.66 | 64.01 | |
| HH head literacy | Cannot read & write | 71.07 | 69.14 | 66.99 | 9.174*** (0.002) |
| | can read & write | 28.93 | 30.86 | 30.01 | |
| Land type | Irrigated & rainfed | 39.22 | 18.75 | 27.73 | 1,062.70***(0.000) |
| | All irrigated | 60.73 | 81.25 | 72.27 | |
| Landscape | Hills and valleys | 57.46 | 21.03 | 37.01 | 3,844.11***(0.000) |
| | Valleys only | 21.73 | 18.01 | 19.64 | |
| | Open plain | 20.81 | 60.95 | 43.35 | |
| Oxen | Don't own | 66.46 | 92.57 | 81.12 | 2.320.35***(0.000) |
| | Own | 33.54 | 7.43 | 18.88 | |
| Hired labour | | | | | |
| ICT equipment | Don't own | 17.91 | 15.57 | 17.32 | 15.31*** (0.000) |
| | Own | 82.09 | 84.43 | 82.68 | |
| Transport equipment | Don't own | 53.30 | 46.10 | 51.48 | 83.07*** (0.000) |
| | Own | 46.70 | 53.90 | 48.52 | |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | ***Table 4.A1 Continued*** |
| Time taken to reach market | >4 hours | 16.81 | 14.66 | 16.27 | 46.04*** (0.000) |
| | 1-4 hours | 20.48 | 17.49 | 19.72 | |
| | <1 hour | 62.71 | 67.85 | 64.01 | |
| HH head literacy | Can't read & write | 71.38 | 65.89 | 69.99 | 57.42*** (0.000) |
| | can read & write | 28.62 | 34.11 | 30.01 | |
| Land type | Irrigated & rain-fed | 25.35 | 34.73 | 27.73 | 175.51*** (0.000) |
| | All irrigated | 74.65 | 65.27 | 72.27 | |
| Landscape | Hills and valleys | 37.28 | 36.20 | 37.01 | 425.40 *** (0.000) |
| | Valleys only | 22.58 | 10.97 | 19.64 | |
| | Open plain | 40.14 | 52.83 | 43.35 | |
| Oxen | Don't own | 81.39 | 80.32 | 81.12 | 2.95* (0.086) |
| | Own | 18.61 | 19.68 | 18.88 | |

*Source:* Author's calculation of the ALCS Data

## Appendix 4.B: FURTHER ECONOMETRIC ANALYSIS

**Table 4.B1:** Estimation of Labour Demand with Different Sex-age Demographic Groups

| VARIABLES | (1) Model 3 | (2) Model 3 | (3) Model 3 |
|---|---|---|---|
| Log. HH size (persons) | 0.547*** | 0.517*** | 0.581*** |
| | (0.022) | (0.025) | (0.054) |
| Share of males (15-19 years) | 1.028*** | 1.045*** | 1.047*** |
| | (0.090) | (0.092) | (0.092) |
| Share of males (20-34 years) | 0.924*** | 0.971*** | 0.974*** |
| | (0.085) | (0.084) | (0.085) |
| Share of males (35-49 years) | 1.013*** | 1.019*** | 1.015*** |
| | (0.126) | (0.122) | (0.123) |
| Share of males (50-64 years) | 1.012*** | 0.935*** | 0.927*** |
| | (0.116) | (0.122) | (0.123) |
| Share of male (65 years & older) | 1.042*** | 0.869*** | 0.857*** |
| | (0.124) | (0.145) | (0.147) |
| Share of females (0-14 years) | -0.078 | -0.066 | -0.069 |
| | (0.053) | (0.050) | (0.050) |
| Share of females (15-19 years) | 0.419*** | 0.439*** | 0.443*** |
| | (0.081) | (0.079) | (0.079) |
| Share of females (20-34 years) | 0.568*** | 0.566*** | 0.568*** |
| | (0.082) | (0.082) | (0.082) |
| Share of females (35-49 years) | 0.680*** | 0.668*** | 0.664*** |
| | (0.094) | (0.094) | (0.094) |
| Share of females (50-64 years) | 0.412*** | 0.390*** | 0.384*** |
| | (0.116) | (0.111) | (0.110) |
| Share of female (65 years & older) | 0.391** | 0.343** | 0.328** |
| | (0.166) | (0.161) | (0.159) |
| ICT equipment (access=1) | - | 0.081*** | 0.043 |
| | | (0.019) | (0.068) |
| Transport equipment (access=1) | - | 0.029* | 0.011 |
| | | (0.016) | (0.062) |
| Time taken to reach nearest market (<1h) | - | -0.066* | -0.067 |
| | | (0.035) | (0.103) |
| Time taken to reach nearest market (1-4h) | - | -0.034 | 0.182** |
| | | (0.031) | (0.084) |
| Log. Distance to road (km) | - | -0.020* | 0.033 |
| | | (0.011) | (0.033) |
| Log. Total land (Jeribs) | - | 0.195*** | 0.195*** |
| | | (0.013) | (0.014) |
| Log off-farm income (AFN) | - | -0.027*** | -0.027*** |
| | | (0.002) | (0.002) |
| HH head literacy (1=can read & write) | - | -0.042* | -0.042* |
| | | (0.023) | (0.022) |
| Log. HH head education (years) | - | 0.002 | 0.002 |

| | | | Table 4.B1 Continue |
|---|---|---|---|
| | | (0.011) | (0.011) |
| HH head age (years) | - | 0.001 | 0.001 |
| | | (0.001) | (0.001) |
| ICT equipment (access=1) # Log. HH size | - | - | 0.019 |
| | | | (0.035) |
| Transport equip. (access=1) # Log. HH size | - | - | 0.009 |
| | | | (0.030) |
| Time taken to market (<1h) # Log. HH size | - | - | -0.0001 |
| | | | (0.051) |
| Time taken to market (1-4h) # Log HH size | - | - | -0.107** |
| | | | (0.042) |
| Log. Distance to road # Log. HH size | - | - | -0.026 |
| | | | (0.017) |
| Wave 2 (2013/14) | - | -0.211*** | -0.210*** |
| | | (0.033) | (0.033) |
| Wave 3 (2016/17) | - | -0.097*** | -0.098*** |
| | | (0.028) | (0.028) |
| District FE | yes | yes | yes |
| Constant | 1.096*** | 1.079*** | 0.966*** |
| | (0.074) | (0.072) | (0.112) |
| *F-test for joint significance of household size and demographic composition (all groups)* | | | |
| F-test: t-statistic | 66.12 | 46.10 | 32.19 |
| F-test: p-value | 0.000 | 0.000 | 0.000 |
| R-squared | 0.217 | 0.265 | 0.265 |
| Observations | 21,189 | 21,189 | 21,189 |

*Note: Dependent variable is the log of total labour days (own and hired) employed by the farm HH. Omitted categories for factor variables are: Share of males between 0-14 years old for the HH composition, no access to ICT and transport equipment, 4 and more hours for time taken to reach market, cannot read and write for HH literacy, and 2011/12 for wave. District fixed effects are included in all regressions. All standard errors are clustered at FE level (in parentheses). Significance is indicated by *** p<0.01, ** p<0.05, * p<0.1.*

**Table 4.B2:** Sample Selection Model-Market Participation & Extent of Expenditures

| Selection Equation: Probit estimates of the input market participation decision | | | | | | |
|---|---|---|---|---|---|---|
| | Fertilizer & chemicals | | Tractor hire | | Labour hire | |
| | (1=use, 0 otherwise) | | (1=hire, 0 otherwise) | | (1=hire, 0 otherwise) | |
| VARIABLES | APE | SE | APE | SE | APE | SE |
| GR (ICT) | -0.064** | (0.029) | -0.092*** | (0.030) | 0.013 | (0.028) |
| GR (Transport Equip) | -0.064* | (0.034) | 0.044 | (0.039) | -0.060 | (0.044) |
| ICT equip (own=1) | 0.184*** | (0.061) | 0.200*** | (0.056) | 0.011 | (0.050) |
| TE equip (own=1) | 0.135** | (0.057) | -0.032 | (0.065) | 0.141** | (0.068) |
| Log. distance rd (km) | 0.005 | (0.006) | -0.021*** | (0.006) | -0.005 | (0.007) |
| Time to market (<1hr) | 0.002 | (0.019) | 0.009 | (0.019) | 0.020 | (0.018) |
| Time to market (1-4h) | 0.015 | (0.015) | 0.068*** | (0.018) | 0.034** | (0.017) |
| Log total land (Jerib) | 0.036*** | (0.008) | 0.079*** | (0.009) | 0.061*** | (0.008) |
| Log off-farm inc. (AFN) | -0.003*** | (0.001) | -0.003*** | (0.001) | 0.003*** | (0.001) |
| Log HH size (count) | 0.005 | (0.012) | -0.007 | (0.011) | -0.042*** | (0.013) |
| Head literacy (1=yes) | -0.004 | (0.012) | 0.005 | (0.012) | 0.028** | (0.014) |
| Log. head edu. (years) | 0.006 | (0.005) | 0.010** | (0.005) | 0.011** | (0.005) |
| Land (1=all irrigated) | 0.167*** | (0.018) | 0.046** | (0.018) | 0.023* | (0.012) |
| Landscape 2 (valleys) | -0.002 | (0.016) | 0.006 | (0.019) | -0.025 | (0.019) |
| Landscape 3 (open plain) | 0.030* | (0.017) | 0.103*** | (0.022) | 0.007 | (0.014) |
| Livestock (N) | 0.008** | (0.004) | 0.003 | (0.005) | -0.005 | (0.004) |
| Oxen(1=own) | 0.034*** | (0.010) | -0.140*** | (0.017) | -0.01 | (0.012) |
| Log. N. tractors (N) | -0.049 | (0.030) | -0.196*** | (0.039) | 0.037 | (0.026) |
| Wave 2 (2013/14) | 0.01 | (0.019) | -0.011 | (0.016) | -0.128*** | (0.016) |
| Wave 3 (2016/17) | 0.031** | (0.015) | -0.042** | (0.016) | -0.111*** | (0.018) |
| District FE | ✓ | | ✓ | | ✓ | |
| Observations | 19,035 | | 19,395 | | 20,415 | |
| Outcome Equation: OLS estimates of the extent of expenditures | | | | | | |
| | ME (y>0) | SE | ME (y>0) | SE | ME (y>0) | SE |
| Log. distance Rd (km) | -0.062*** | (0.017) | -0.001 | (0.014) | -0.021 | (0.023) |
| Time to market (<1hr) | -0.057 | (0.051) | 0.122*** | (0.039) | -0.071 | (0.067) |
| Time to market (1-4h) | -0.047 | (0.049) | 0.064* | (0.036) | 0.011 | (0.049) |
| Log total land (Jerib) | 0.538*** | (0.023) | 0.614*** | (0.022) | 0.442*** | (0.024) |
| Log off-farm inc. (AFN) | -0.016*** | (0.002) | -0.009*** | (0.002) | -0.011*** | (0.003) |
| Log HH size (count) | 0.140*** | (0.023) | 0.115*** | (0.021) | 0.164*** | (0.031) |
| Head literacy (1=yes) | 0.042 | (0.028) | 0.060** | (0.026) | -0.07 | (0.044) |
| Log. head edu. (years) | 0.015 | (0.015) | -0.017 | (0.013) | 0.051** | (0.021) |
| Land (1=all irrigated) | 0.001 | (0.010) | -0.002 | (0.010) | 0.018 | (0.014) |
| Landscape 2 (valleys) | 0.120*** | (0.030) | -0.103*** | (0.034) | 0.045 | (0.033) |
| Landscape 3 (open plain) | 0.362*** | (0.058) | 0.069 | (0.057) | 0.297*** | (0.103) |
| Livestock (N) | 0.035 | (0.035) | -0.017 | (0.042) | -0.063 | (0.056) |
| Oxen(1=own) | 0.130*** | (0.038) | 0.031 | (0.031) | -0.052 | (0.049) |
| Log N. tractors (N) | 0.416*** | (0.045) | 0.102*** | (0.034) | 0.225*** | (0.043) |
| Wave 2 (2013/14) | 0.223*** | (0.045) | 0.233*** | (0.048) | 0.061 | (0.064) |
| Wave 3 (2016/17) | 0.274*** | (0.046) | 0.235*** | (0.040) | 0.357*** | (0.058) |
| District FE | ✓ | | ✓ | | ✓ | |
| Observations | 19,032 | | 19,392 | | 20,411 | |

*Significance is indicated by *** p<0.01, ** p<0.05, * p<0.1*

# 5 CHAPTER V: SUMMARY AND CLOSING REMARKS

This study explored small-scale farmers' decision-making in the context of agrarian transition from subsistence into a diversified and market-led system where production is not only for subsistence and survival but also oriented towards market to maximise cash returns. Improving productivity to increase production to achieve self-sufficiency as well as produce marketable surpluses is an important milestone towards commercialization, particularly as most smallholders are resource-constrained and maximizing productivity is a tough challenge. On-farm diversification via adding high value cash crops to the production portfolio is considered to be a pivotal step towards agriculture commercialization. In this thesis we attempted to address the challenges and opportunity faced by small-scale farmers throughout the agrarian transition to achieve the broader overarching objective of improved welfare through crop production choices, improving farm productivity, and improving market access and marketing decisions.

Household's production and marketing decisions are practically interlinked. Empirical research suggests that when markets are incomplete or missing, farm households tend to produce staple food crops (mostly grains) mainly for home consumption to be food self-sufficient. However, as access to the input and output markets become available farm households aim to diversify production into high value marketable crops to improve their cash income. While diversifying production and gains in productivity are essential for successful integration of small farms into markets, complementary services are required to overcome market related problems and increase the efficiency of marketing systems to sustain and facilitate a diversified market-oriented production portfolio.

The overall picture emerging from the empirical analysis conducted in this thesis is that adopting crop diversification strategies are important for achieving self-sufficiency and rural incomes by improving production efficiency. This finding is particularly critical as the evidence suggests the presence of a relatively low level of diversity in crop production. Household's access to non-farm activities along with other spatial factors were found to significantly reduce on-farm diversity. This is notably interesting as the overall economic growth and revitalization of non-agricultural sectors (i.e. services, industry, and mining) may present employment opportunity to shift reliance from farm to alternative non-farm activities, particularly when farming business experiences high volatility, thus households attempting to spread risk over a diverse profile of both on-and off-fam activities. Meanwhile, in analysing household's marketing decisions, the empirical analysis reveals that there is evidence of potential market failures or incomplete markets (perhaps due to high transaction costs) that constrain farmers market participation. Ownership of communication technologies and transportation equipment by the farm households significantly improves the likelihood of participation in the input markets by reducing transaction costs through the provision of reliable marketing information and cheaper means of transportation.

Increased productivity as a key step towards achieving self-sufficiency and generating market surplus can expedite the agriculture transition process. Our focus in Chapter II is to estimate farm-level Technical Efficiency (TE), identify potential sources of (in)efficiency, and in particular assess the implications of crop diversification strategies on the farm level TE. Our estimates of a translog stochastic frontier production function indicate that substantial technical inefficiency exists, which means there is opportunity to expand farm revenues up to 28% by applying good farm management strategies (e.g. crop diversification) and without having to resort to greater use of production inputs or the introduction of improved production technologies. When correcting for endogeneity bias, the impact of CD is even greater. Our estimates of the transformed Herfindahl Index show the low level of crop diversification in Afghanistan.

Chapter III turns to the microeconomic drivers of diversity in crop production, especially the labour allocation decision among farm and non-farm activities. The major finding is that household's access to off-farm income negatively affects diversity of crop production. Controlling for the endogeneity in off-farm income by employing the Instrumental Variable (IV) techniques revealed that unobserved factors (such as entrepreneurship skills and response to risk) drive farm household decision to diversify into both farm and non-farm activities. As farmers, particularly smallholder subsistence farmers, are risk averse, they do not only rely on crop diversity to sustain their livelihoods and income (or survival), but also diversify their income sources to other non-farm activities.

Beyond production choices, household's market participation is another important component of the agrarian transition from subsistence and Chapter IV addresses market performance and farm households' market participation decisions in marginal but market-oriented conditions. We test for market failures by testing the hypothesis of separability. Data show that factor market participation rates still remain relatively low in Afghanistan. Although some two-thirds of households purchased chemical fertilizers from the market, the most common input purchased, and more than half hired tractor, considerably fewer households (25%) hired labour for their farming activities.

In theory, households would act as if they are profit maximizers when markets function well, where they first make production choices and then make consumption decisions based on the profits from production. Thus, if factor markets are functioning perfectly and separation holds, household labour demand in agricultural production should only depend on relevant prices but not household consumption preferences. Our analysis shows significant dependency of the household labour demand on its endowments. This is not consistent with the hypothesis of separation, and suggests the prevalence of potential market failures or incomplete markets in Afghanistan.

Some of the most limiting constraints and barriers to market participation and even market failures is the existence of high transaction costs due to lack of access to cheaper transportation and market information, high search costs, and remoteness from markets. In analysing households' marketing decisions, we give specific attention to the impact of transaction costs on the likelihood of market participation and extent of expenditures. Household ownership of communication and transportation equipment (as proxies for fixed transaction costs) and distance to local markets and roads (measures for proportional transaction costs) together with standard explanatory variables (household socio-economic, farm and geographical characteristics) are included in the econometric analysis. The major conclusion from the market participation analysis is that market imperfections and high transaction costs are responsible for hindering participation in the local agriculture markets. Households that have improved access to communication technologies and transport assets are likely to participate more in markets. Farm households living within radius to local markets and with improved access to roads are also more likely to participate in markets and spend more on purchasing inputs.

## 5.1 Recommendations and Implications for Policy

When markets are incomplete or missing, farmers are pushed to produce staple crops that are mainly for home consumption. Since the analysis in this thesis clearly suggests that more diversified farmers are more efficient and achieve higher revenues, over-reliance on food grain-based production systems will no longer be adequate to effectively respond to the changing market demand. Farm policies should therefore focus on promoting crop diversification strategies to enhance production gains as well as to be able to meet changing market demand. Other options to improve production obviously lies in shifting production frontiers by promoting R&D in agriculture and adopting improved farming technology.

Another recommendation that can be directly generated from the analysis carried out in this thesis is that enhanced production performance and efficiency levels may also be pursued through promotion of integrated agricultural extension services with emphasis on transforming knowledge about proper combination of crops and other technical aspects of farming to local farmers, particularly farmers that live in remote areas and have limited access to information. Meanwhile, extension services could increase famer's familiarity with local markets by providing reliable and timely information on input and output prices, market research and technology and innovation.

Given that the separation of production and consumption decisions does not hold, suggesting the presence of incomplete markets, future policies to tackle potential market failures or missing market should not rely on the assumption of efficient or complete markets to devise policies encouraging agrarian transformation. Meanwhile, if household's agriculture marketing decisions are linked to other decisions involving non-agriculture sectors, policies targeting individual sectors distinctly may not be viable or effective. Thus, a holistic policy approach will be required to address cross-cutting issues that affect farming while also improving household's engagement in off-farm sources of income to spread risk over a diverse portfolio of activities.

The need for and type of design of interventions to tackle down market failures would largely depend on the type of the situation confronted or the factors that cause markets to fail. For instance, policy instruments to target completely missing markets may involve removal of legal restrictions or imposing property rights, whereas situations where markets exist but do not function efficiently may require other interventions aiming at increasing investment in public infrastructure to reduce transaction costs (roads, access to telecommunication, etc.), termination of collusions and formation of oligopolistic situation, education and provision of extension services, and possibly government subsidies.

This study provides hints on the critical implication of transaction costs on market participation and missing markets. One area of policy intervention that can be suggested from the findings of this study is that future policies geared towards agriculture commercialization should involve providing viable and timely information on market prices, technical advantages of using modern inputs and other important information through local media, so that farmers communication assets are effectively used in accessing market information. This is particularly important as communication assets are maybe less useful in facilitating transactions if there is no viable market information service available through public or private media. Other means such as publishing price information through local newspapers and bulletins may also help facilitate access to markets and mitigate search costs.

In general, a holistic market-oriented agricultural policy would augment productivity gains, improve adoption of crop diversification, and subsequently enhance market orientation by improving access to market information, facilitating low-cost transportation, addressing institutional weaknesses, and improving public and commercial input distribution systems. Enhanced access to market-oriented agricultural extension services particularly in remote areas to assist farmers to recognise the benefits of using modern agriculture inputs and utilizing best practices is unquestionably essential for enhancing market orientation and market participation. Collective action through cooperatives and farmer organizations may also enable farmers, particularly resource-poor farmers, to share their resources, achieve economies of scale, and increase efficiencies in accessing local markets. Another potential strategy could be contract farming to guarantee surplus production is effectively sold to the market and farmers gain sufficient cash to purchase inputs from the market. Contract farming, in some cases, can also aid farmers to trade some of their surplus production for agriculture inputs that can be used in the next planting season. Another general recommendation is that future government policy instruments that aim to incentivize investment in rural infrastructure development such as roads, transportation facilities, and other means to stabilize input and output supply chain and distribution systems could also promote market orientation and avert some of the negative consequences due to market imperfections or failures.

## 5.2 Future Research

The ALCS, as the only household level data available for Afghanistan comes with some limitations that may pose a number of shortcomings that might affect our results. As an example, the input data could not be disaggregated by crop or plot level, thus restricting the analysis to the estimation of aggregate production functions. The current data do not allow to model household's participation in the output markets as there were not data collected on the quantities or value of crop sales in the survey. It is plausible to assume that household output marketing decisions are linked with their input utilization decisions. For instance, farmers who participate in input markets, may also purchase inputs from markets or trade some of their surplus output for inputs for the next planting season. Thus, further research may incorporate this property to analyse market orientation from the output point of view and identify what drives participation in output markets and how output marketing decisions are related to factor market decisions.

Future research could extend the current analysis if better production and cost data for individual crops become available to estimate potential gains from allocative efficiency and cost minimization associated with the optimal allocation or combination of production inputs. In addition, further data will make it feasible to carry out cost and benefit analysis for different combinations of crops and their implications for household income and employment. Since there were no data to measure actual costs associated with transactions or exchange of goods or production inputs, we resorted to using observed factors as proxy measures such as access to communication and transport assets, and distance to markets and roads. Thus, detailed data on actual transaction costs may help to properly quantify transaction costs and analyse their implications on market participation and amount of transactions.

Based on a theoretically-grounded test, we rejected the hypothesis of separation (i.e. the assumption that there is a separation between household's production and consumption decisions is not consistent with the data) which could mean that markets are incomplete or failing. However, our analysis fails to identify and address potential sources of market failures. Perhaps a market fails because of high transportation costs, lack of access to information, or other contract enforcement issues. The tests for separation rely on the labour market transactions and therefore it is difficult to conclude which specific input markets may actually fail. Meanwhile, it is important to underscore that if farm household behaviour is consistent with separation or recursive case, it does not necessarily mean that markets are complete and fully functional. Rather, one interpretation of failing to reject separation is that perhaps farm households allocate resources in a manner that makes up for missing markets and, thereby, their choices can be modelled as if all markets exist. Thus, future research could further build on the analysis carried out in this thesis to identify which markets are failing and why, so as to develop policy recommendations.

# 6 References


Abay, F., Bjørnstad, A., Smale, M., 2009. Measuring on farm diversity and determinants of barley diversity in Tigray, northern Ethiopia. *Momona Ethiopian Journal of Science* 1, 44–66.

Abdalla, S., Leonhäuser, I., Siegfried, S.B., Elamin, E., 2013. Factors influencing crop diversity in dry land sector of Sudan. *Sky Journal of Agricultural Research* 2, 88–97.

Abdulai, A., Regmi, P.P., 2000. Estimating labor supply of farm households under nonseparability: empirical evidence from Nepal. *Agricultural Economics* 22, 309–320.

Abdulai, A., CroleRees, A., 2001. Determinants of income diversification amongst rural households in Southern Mali. *Food Policy* 26, 437–452.

Abdullah, Rabbi, F., Ahamad, R., Ali, S., Chandio, A.A., Ahmad, W., Ilyas, A., Din, I.U., 2017. Determinants of commercialization and its impact on the welfare of smallholder rice farmers by using Heckman's two-stage approach. *Journal of the Saudi Society of Agricultural Sciences.*

Acharya, S.P., Basavaraja, H., Mahajanashetti, S.B., Bhat, A.R.., 2011. Crop Diversification in Karnataka: An Economic Analysis. *Agricultural Economics Research Review* 24, 351–357.

Adhikari, C.B., Bjorndal, T., 2012. Analyses of technical efficiency using SDF and DEA models: evidence from Nepalese agriculture. *Applied Economics* 44, 3297–3308.

Ahmadzai, H., 2017. Crop Diversification and Technical Efficiency in Afghanistan: Stochastic Frontier Analysis. *School of Economics*, UoN, CRP 17/04.

Aigner, D., Lovell, C.K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6, 21–37.

Alam, M.F., Khan, M.A., Huq, A.S.M.A., 2012. Technical efficiency in tilapia farming of Bangladesh: a stochastic frontier production approach. *Aquaculture International* 20, 619–634.

Afghanistan Living Condition Survey (ALCS) Report 2013-14, Central Statistic Organization (CSO), Available at: http://cso.gov.af/Content/files/ALCS/ALCS%20ENGLISH%20REPORT%202014.pdf

Afghanistan Living Condition Survey (ALCS) 2016-17, Central Statistics Organization (CSO). Available at: http://cso.gov.af/Content/files/ALCS/ALCS%20-%202016-17%20Analysis%20report%20-%20pre-print%20for%20web_rev.pdf

Alene, A.D., Manyong, V.M., Omanya, G., Mignouna, H.D., Bokanga, M., Odhiambo, G., 2008. Smallholder market participation under transactions costs: Maize supply and fertilizer demand in Kenya. *Food Policy* 33, 318–328.

Amankwah, A., Quagrainie, K.K., Preckel, P.V., 2016. Demand for improved fish feed in the presence of a subsidy: a double hurdle application in Kenya. *Agricultural Economics* 47, 633–643.

Amaza, P.S., Bila, Y., Iheanacho, A.C., 2006. Identification of factors that influence technical efficiency of food crop production in West Africa: Empirical evidence from Borno State, Nigeria. *Journal of Agriculture and Rural Development in the Tropics and Subtropics (JARTS)* 107, 139–147.

Amsler, C., Prokhorov, A., Schmidt, P., 2016. Endogeneity in stochastic frontier models. Journal of Econometrics 190, 280–288.

Aneani, F., Anchirinah, V.M., Owusu-Ansah, F., Asamoah, M., 2011. An analysis of the extent and determinants of crop diversification by cocoa (Theobroma cacao) farmers in Ghana. *African Journal of Agricultural Research* 6, 4277–4287.

Arnberg, S., Hansen, L.G., 2012. Short-run and long-run dynamics of farm land allocation: panel data evidence from Denmark1. *Agricultural Economics* 43, 179–190.

Asfaw, S., Pallante, G., Palma, A., 2018. Diversification Strategies and Adaptation Deficit: Evidence from Rural Communities in Niger. *World Development* 101, 219–234.

Asunka, S., Shumway, C.R., 1996. Allocatable Fixed Inputs and Jointness in Agricultural Production: More Implications. *Agricultural and Resource Economics Review* 25, 143–148.

Ayieko, M.W., 2015. Patterns, determinants, and welfare effects of agricultural and livelihood diversification among smallholder farmers in rural Kenya. Michigan State University. Available at: https://d.lib.msu.edu/etd/3522

Babatunde, R.O., Qaim, M., 2009. Pattern of income diversification in rural Nigeria: determinants and impact.pdf. *Quarterly Journal of International Agriculture*, 48, 305–329.

Barrett, C.B., 2008. Smallholder market participation: Concepts and evidence from eastern and southern Africa. *Food Policy* 33, 299–317.

Barrett, C.B., Bezuneh, M., Aboud, A., 2001. Income diversification, poverty traps and policy shocks in Coˆte d'Ivoire and Kenya. Food Policy 18, 367–384.

Barrett, C.B., Bezuneh, M., Clay, D.C., Reardon, T., 2005. Heterogeneous Constraints, Incentives and Income Diversification Strategies in Rural Africa. *Quarterly Journal of International Agriculture,* 44, 37–60.

Barrett, C.B., Reardon, T., 2000. Asset, activity, and income diversification among African agriculturalists: Some practical issues. Dept. of Agricultural, Resource and Managerial Economics, Cornell University, Ithaca, NY. Available at: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=73B84C85864EDA6813EF43C0B21D23C4?doi=10.1.1.14.6975&rep=rep1&type=pdf

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

Bartolini, F., Andreoli, M., Brunori, G., 2014. Explaining determinants of the on-farm diversification: empirical evidence from Tuscany region. *Bio-based and Applied Economics* 3, 137.

Battese, G.E., Coelli, T.J., 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical economics* 20, 325–332.

Battese, G.E., Corra, G.S., 1977. Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia. *Australian journal of Agricultural Economics* 21, 169–179.

Behera, U.K., Sharma, A.R., Mahapatra, I.C., 2007. Crop Diversification for Efficient Resource Management in India: Problems, Prospects and Policy. *Journal of Sustainable Agriculture* 30, 97–127.

Benin, S., Smale, M., Pender, J., Gebremedhin, B., Ehui, S., 2004. The economic determinants of cereal crop diversity on farms in the Ethiopian highlands. *Agricultural Economics* 31, 197–208.

Benjamin, D., 1992. Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models. *Econometrica* 60, 287.

Bezabih, M., Sarr, M., 2012. Risk Preferences and Environmental Uncertainty: Implications for Crop Diversification Decisions in Ethiopia. *Environmental and Resource Economics* 53, 483–505.

Bhatt, M.S., Bhat, S.A., 2014. Technical Efficiency and Farm Size Productivity-Micro Level Evidence from Jammu & Kashmir. *International Journal of Food and Agricultural Economics* 2, 27.

Birthal, P.S., Joshi, P.K., Roy, D., Thorat, A., 2013. Diversification in Indian Agriculture toward High-Value Crops: The Role of Small Farmers. *Canadian Journal of Agricultural Economics/Revue, canadienne d'agroeconomie* 61, 61–91.

Bowlus, A.J., Sicular, T., 2003. Moving toward markets? Labor allocation in rural China. *Journal of Development Economics* 71, 561–583.

Bowman, M.S., Zilberman, D., 2013. Economic Factors Affecting Diversified Farming Systems. *Ecology and Society* 18, 33.

Bozoğlu, M., Ceyhan, V., 2007. Measuring the technical efficiency and exploring the inefficiency determinants of vegetable farms in Samsun province, Turkey. *Agricultural Systems* 94, 649–656.

Cavatassi, R., Lipper, L., Winters, P., 2012. Sowing the seeds of social relations: social capital and agricultural diversity in Hararghe Ethiopia. *Environment and Development Economics* 17, 547–578.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2, 429–444.

Chavas, J.-P., Petrie, R., Roth, M., 2005. Farm Household Production Efficiency: Evidence from The Gambia. *American Journal of Agricultural Economics* 87, 160–179.

Chhatre, A., Devalkar, S., Seshadri, S., 2016. Crop diversification and risk management in Indian agriculture. *DECISION* 43, 167–179.

Chowdhury, S.K., 2006. Access to a telephone and factor market participation of rural households in Bangladesh. *Journal of Agricultural Economics* 57, 563–576.

Coelli, T., 1995. Estimators and hypothesis tests for a stochastic frontier function: A Monte Carlo analysis. *Journal of Productivity Analysis* 6, 247–268.

Coelli, T., Fleming, E., 2004. Diversification economies and specialisation efficiencies in a mixed food and coffee smallholder farming system in Papua New Guinea. *Agricultural Economics* 31, 229–239.

Coelli, T., Rao, O'Donnell, Battese (Eds.), 2005. An introduction to efficiency and productivity analysis, 2nd ed. ed. Springer, New York.

Coelli, T.J., 1995. Recent developments in frontier modelling and efficiency measurement. *Australian Journal of Agricultural Economics* 39, 219–245.

Covarrubias, K.A., 2015. The role of crop diversity in household production and food security in Uganda: A gender-differentiated analysis. FOODSECURE Working paper no. 32. Available at: https://ideas.repec.org/p/fsc/fspubl/32.html

Christensen, L.R., Jorgenson, D.W., Lau, L.J., 1973. Transcendental Logarithmic Production Frontiers. *The Review of Economics and Statistics* 55, 28.

Cutforth, L.B., Francis, C.A., Lynne, G.D., Mortensen, D.A., Eskridge, K.M., 2001. Factors affecting farmers' crop diversity decisions: An integrated approach. *American Journal of Alternative Agriculture* 16, 168.

Davis, B., Winters, P., Carletto, G., Covarrubias, K., Quiñones, E.J., Zezza, A., Stamoulis, K., Azzarri, C., DiGiuseppe, S., 2010. A Cross-Country Comparison of Rural Income Generating Activities. *World Development* 38, 48–63.

de Janvry, A., Fafchamps, M., Sadoulet, E., 1991a. Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. *The Economic Journal* 101, 1400.

de Janvry, A., Fafchamps, M., Sadoulet, E., 1991b. Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. *The Economic Journal* 101, 1400.

De Janvry, A., Sadoulet, E., 2006. Progress in the modelling of rural households' behavior under market failures, in: Poverty, Inequality and Development. Springer, pp. 155–181.

De, U.K., Chattopadhyay, M., 2010. Crop diversification by poor peasants and role of infrastructure: Evidence from West Bengal. *Journal of Development and Agricultural Economics* 2, 340–350.

Diiro, G.M., Sam, A.G., 2015. Agricultural technology adoption and Nonfarm earnings in Uganda: a Semiparametric analysis. *The Journal of Developing Areas* 49, 145–162.

Dillon, B., Barrett, C.B., 2017. Agricultural factor markets in Sub-Saharan Africa: An updated view with formal tests for market failure. *Food Policy* 67, 64–77.

Donovan, J., Poole, N., 2014. Changing asset endowments and smallholder participation in higher value markets: Evidence from certified coffee producers in Nicaragua. *Food Policy* 44, 1–13.

Dube, L., 2016. Factors Influencing Smallholder Crop Diversification: A Case Study of Manicaland and Masvingo Provinces in Zimbabwe. *International Journal of Regional Development* 3, 1.

Duflo, E., Kremer, M., Robinson, J., 2008. How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya. *American Economic Review* 98, 482–488.

Dupree, L., 1973. Afghanistan, Oxford Pakistan paperbacks. Oxford University Press.

Elias, A., Nohmi, M., Yasunobu, K., Ishida, A., 2013. Effect of Agricultural Extension Program on Smallholders' Farm Productivity: Evidence from Three Peasant Associations in the Highlands of Ethiopia. *Journal of Agricultural Science* 5.

Ellis, F., 2008. The Determinants of Rural Livelihood Diversification in Developing Countries. *Journal of Agricultural Economics* 51, 289–302.

Ellis, F., 2000. The determinants of rural livelihood diversification in developing countries. *Journal of Agricultural Economics* 51, 289–302.

Ellis, F., 1998. Household strategies and rural livelihood diversification. *Journal of Development Studies* 35, 1–38.

Farrell, M.J., 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)* 120, 253.

Felix Larry Essilfie, 2011. Estimation of farm level technical efficiency in small scale maize production in the Mfantseman Municipality in the Central Region of Ghana: A stochastic frontier approach. *Journal of Development and Agricultural Economics* 3.

Finlay, K., Magnusson, L.M., 2009. Implementing weak-instrument robust tests for a general class of instrumental-variables models. *The Stata Journal* 9, 398–421.

Fischer, E., Qaim, M., 2012. Linking Smallholders to Markets: Determinants and Impacts of Farmer Collective Action in Kenya. *World Development* 40, 1255–1268.

Gebregziabher, K., Mathijs, E., Maertens, M., Deckers, J., Bauer, H., Gebrehiwot, K., 2012. Is Non-Farm Income Relaxing Farm Investment Liquidity Constraints for Marginal Farms? An Instrumental Variable Approach. *International Journal of Economics and Finance Studies* 4, 123–32.

Girish, M., Mehta, P., 2003. Crop diversification: an empirical analysis on Kangra farms of Himachal Pradesh, India. *Asia-Pacific Journal of Rural Development* 13, 1–18.

Goetz, S.J., 1992. A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa. *American Journal of Agricultural Economics* 74, 444–452.

Goodwin, B.K., Mishra, A.K., 2004. Farming efficiency and the determinants of multiple job holding by farm operators. A*merican Journal of Agricultural Economics* 86, 722–729.

Greene, W.H., 2012. Econometric analysis, 7th ed. ed. Prentice Hall, Boston.

Greene, W.H., 2002. Econometric Analysis, 5th Edition. ed. Prentice Hall.

Greene, W.H., 1980. Maximum likelihood estimation of econometric frontier functions. *Journal of Econometrics* 13, 27–56.

Grimard, F., 2000. Rural labor markets, household composition, and rainfall in Côte d'Ivoire. *Review of Development Economics* 4, 70–86.

Gronberg, T.J., Jansen, D.W., Karakaplan, M.U., Taylor, L.L., 2015. School district consolidation: Market concentration and the scale-efficiency tradeoff: School Consolidation: Scale & Efficiency. *Southern Economic Journal* 82, 580–

Guan, Z., Kumbhakar, S.C., Myers, R.J., Lansink, A.O., 2009. Measuring Excess Capital Capacity in Agricultural Production. American Journal of Agricultural Economics 91, 765–776. Guvele, C.A., 2001. Gains from crop diversification in the Sudan Gezira scheme. *Agricultural Systems* 70, 319–333.

Haji, J., 2007. Production Efficiency of Smallholders' Vegetable-dominated Mixed Farming System in Eastern Ethiopia: A Non-Parametric Approach. *Journal of African Economies* 16, 1–27.

Hazell, P.B.R., Poulton, C., Wiggins, S., Dorward, A. (Eds.), 2007. The future of small farms for poverty reduction and growth. International Food Policy Research Institute (IFRI), 2020 Discussion Paper 42. Available at: http://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/125263/filename/125264.pdf

Helfand, S., Levine, E., 2004. Farm size and the determinants of productive efficiency in the Brazilian Center-West. *Agricultural Economics* 31, 241–249.

Heltberg, R., Tarp, F., 2002. Agricultural supply response and poverty in Mozambique. *Food policy* 27, 103–124.

Herath, H.M.G., 1980. Crop mixtures under peasant conditions: A decision theoretic interpretation. *The Journal of Development Studies* 16, 322–331.

Hitayezu, P., Zegeye, E.W., Ortmann, G.F., 2016. Farm-level crop diversification in the Midlands region of Kwazulu-Natal, South Africa: patterns, microeconomic drivers, and policy implications. *Agroecology and Sustainable Food Systems* 40, 553–582.

Holloway, G., Nicholson, C., Delgado, C., Staal, S., Ehui, S., 2000. Agro-industrialization through institutional innovation Transaction costs, cooperatives and milk-market development in the east-African highlands. *Agricultural Economics* 23, 279–288.

Hsu, A.-C., Liu, S.-C., 2008. The hurdle models choice between truncated normal and lognormal. *Applied Economics* 40, 201–207.

Ibrahim, H., Rahman, S.A., Envulus, E.E., Oyewole, S.O., others, 2009. Income and crop diversification among farming households in a rural area of north central Nigeria. *Journal of Tropical Agriculture, Food, Environment and Extension* 8, 84–89.

Jacoby, H.G., 1993. Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra. *The Review of Economic Studies* 60, 903–921.

Jadhav, S.K., Deshmukh, K.V., 2014. Crop Diversification in Marathwada Region of Maharashtra: An Economic Analysis. *Agricultural Economics Research Review* 27.

Jagwe, J.N., 2011. The impact of transaction costs on the participation of smallholder farmers and intermediaries in the banana markets of Burundi, Democratic Republic of Congo and Rwanda. Available at: https://repository.up.ac.za/handle/2263/25567

Ji, Y., Yu, X., Zhong, F., 2012. Machinery investment decision and off-farm employment in rural China. *China Economic Review* 23, 71–80.

Jondrow, J., Lovell, C.K., Materov, I.S., Schmidt, P., 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics* 19, 233–238.

Joshi, P.K., Gulati, A., Cummings, R.W., 2007. Agricultural Diversification in South Asia: Beyond Food Security. *New Delhi: Academic Foundation* 47–82.

Joshi, P.K., Joshi, L., Birthal, P.S., 2006. Diversification and Its Impact on Smallholders: Evidence from a Study on Vegetable Production. *Agricultural Economics Research Review* 19, 219–236.

Just, R.E., Zilberman, D., Hochman, E., 1983. Estimation of Multicrop Production Functions. *American Journal of Agricultural Economics* 65, 770.

K. Mwajombe, K., R. S. Mlozi, M., 2015. Measuring Farm-level Technical Efficiency of Urban Agriculture in Tanzanian Towns: The Policy Implications. *World Journal of Social Science* 2.

Kanyua, M., Ithinji, G., Muluvi, A., Gido, O., Waluse, S., 2013. Pattern, Trend and Determinants of Crop Diversification: Empirical Evidence from Smallholders in Eastern Ethiopia. *Current Research Journal of Social Sciences* 5, 103–111.

Karakaplan, U.M., 2017. Fitting endogenous stochastic frontier models in Stata. *The Stata Journal* 17, 39–55.

Karakaplan, U.M., Kutlu, L., 2017. Handling Endogeneity in Stochastic Frontier Analysis. *Economics Bulletin* 37, 889–901.

Kasem, S., Thapa, G.B., 2011. Crop diversification in Thailand: Status, determinants, and effects on income and use of inputs. *Land Use Policy* 28, 618–628.

Key, N., Sadoulet, E., Janvry, A.D., 2000. Transactions costs and agricultural household supply response. *American Journal of Agricultural Economics* 82, 245–259.

Kilic, T., Carletto, C., Miluka, J., Savastano, S., 2009. Rural nonfarm income and its impact on agriculture: evidence from Albania. *Agricultural Economics* 40, 139–160.

Kimhi, A., Chiwele, D., 2000. Barriers for Development in Zambian Small- and Medium-Size Farms: Evidence from Micro-Data. American Agricultural Economics Association (AAEA). Available at: https://ideas.repec.org/p/ags/aaea00/21877.html

Koundouri, P., Laukkanen, M., Myyra, S., Nauges, C., 2009. The effects of EU agricultural policy changes on farmers' risk attitudes. E*uropean Review of Agricultural Economics* 36, 53–77.

Krupinsky, J.M., Bailey, K.L., McMullen, M.P., Gossen, B.D., Turkington, T.K., 2002. Managing plant disease risk in diversified cropping systems. *Agronomy Journal* 94, 198–209.

Kudaligama, V.P., Yanagida, J.F., others, 2000. A comparison of intercountry agricultural production functions: A frontier function approach. *Journal of Economic Development* 25, 57–74.

Kumar, A., Kumar, P., Sharma, A.N., 2012. Crop diversification in Eastern India: Status and determinants. *Indian Journal of Agricultural Economics* 67, 600–616.

Kumar, S., Gupta, S., 2015. Crop diversification in India: Emerging trends, determinants and policy implications. *International Journal of Current Research* 7, 17188–17195.

Kumara, Rathnasekar, Samarakon, Wijesinghe, Mowjood, Galagedara, 2016. Factors Effecting to Participation and Cultivation Extent of Other Field Crops: Heckman Two Stage Model Approach in Awlegama, wariyapola, Sri Lanka. Available at: http://www.worldresearchlibrary.org/up_proc/pdf/249-146113480111-17.pdf

Kumbhakar, S., Wang, H.-J., 2015. A practitioner's guide to stochastic frontier analysis using Stata. Cambridge University Press, New York, NY.

Kumbhakar, S.C., Lovell, C.A.K., 2000. Stochastic Frontier Analysis. Cambridge University Press.

Kurosaki, T., 2003. Specialization and Diversification in Agricultural Transformation: The Case of WestPunjab, 1903-92. *American Journal of Agricultural Economics* 85, 372–386.

LaFave, D., Thomas, D., 2016. Farms, Families, and Markets: New Evidence on Completeness of Markets in Agricultural Settings. *Econometrica* 84, 1917–1960.

Lakner, S., Kirchweger, S., Hoop, D., Brümmer, B., Kantelhardt, J., 2015. Impact of diversification on technical efficiency of organic farming in Switzerland, Austria and Southern Germany. Diskussionspapiere, Department für Agrarökonomie und Rurale Entwicklung. Available at:
https://ideas.repec.org/p/zbw/daredp/1508.html

Lapar, M.L., Holloway, G., Ehui, S., 2003. Policy options promoting market participation among smallholder livestock producers: a case study from the Philippines. *Food Policy* 28, 187–211.

Leathers, H.D., 1991. Allocable Fixed Inputs as a Cause of Joint Production: A Cost Function Approach. *American Journal of Agricultural Economics* 73, 1083–1090.

Le, K.T., 2010. Separation Hypothesis Tests in the Agricultural Household Model. *American Journal of Agricultural Economics* 92, 1420–1431.

Lee, L.-F., 1983. A test for distributional assumptions for the stochastic frontier functions. *Journal of Econometrics* 22, 245–267.

Lence, S.H., Miller, D.J., 1998a. Estimation of multi-output production functions with incomplete data: A generalised maximum entropy approach. *European Review of Agricultural Economics* 25, 188–209.

Lence, S.H., Miller, D.J., 1998b. Recovering Output-Specific Inputs from Aggregate Input Data: A Generalized Cross-Entropy Approach. *American Journal of Agricultural Economics* 80, 852–867.

Lin, B.B., 2011. Resilience in Agriculture through Crop Diversification: Adaptive Management for Environmental Change. BioScience 61, 183–193.

Lin, T.-F., Schmidt, P., 1984. A Test of the Tobit Specification Against an Alternative Suggested by Cragg. *The Review of Economics and Statistics* 66, 174.

Liverpool-Tasie, L.S.O., 2014. Fertilizer subsidies and private market participation: the case of Kano State, Nigeria. *Agricultural Economics* 45, 663–678.

Llewelyn, R.V., Williams, J.R., 1996. Nonparametric analysis of technical, pure technical, and scale efficiencies for food crop production in East Java, Indonesia. *Agricultural Economics* 15, 113–126.

Lovo, S., 2011. Pension Transfers and farm household technical efficiency: Evidence from South Africa. *American Journal of Agricultural Economics* 93, 1391–1405.

Makate, C., Wang, R., Makate, M., Mango, N., 2016. Crop diversification and livelihoods of smallholder farmers in Zimbabwe: adaptive management for environmental change. SpringerPlus 5.

Makhura, M.T., 2001. Overcoming Transaction Costs Barriars to Market Participation of Smallholder Farmers in the Northern Province of South Africa 207. Available at: https://repository.up.ac.za/bitstream/handle/2263/27659/Complete.pdf?sequence=10

Maletta, H., Favre, R., 2003. Agroecological zones. Food and Agriculture Organization of the United Nations. Available at: http://www.fao.org/tempref/docrep/fao/007/ae407e/ae407e00.pdf

Malik, D.P., Singh, I.J., 2002. Crop Diversification - An Economic Analysis. *Indian Journal of Agricultural Research* 36, 61–64.

Mandal, R., Bezbaruah, M.P., 2013. Diversification of cropping pattern: Its determinants and role in flood affected agriculture of Assam Plains. *Indian Journal of Agricultural Economics* 68, 170–181.

Mango, N., Makate, C., Hanyani-Mlambo, B., Siziba, S., Lundy, M., 2015. A stochastic frontier analysis of technical efficiency in smallholder maize production in Zimbabwe: The post-fast-track land reform outlook. *Cogent Economics & Finance* 3.

Manjunatha, A.V., Anik, A.R., Speelman, S., Nuppenau, E.A., 2013. Impact of land fragmentation, farm size, land ownership and crop diversity on profit and efficiency of irrigated farms in India. *Land Use Policy* 31, 397–405.

Martey, E., Al-Hassan, R.M., Kuwornu, J.K.M., 2012. Commercialization of smallholder agriculture in Ghana: A Tobit regression analysis. *African Journal of Agricultural Research* 7, 11.

Mather, D., Boughton, D., Jayne, T.S., 2013. Explaining smallholder maize marketing in southern and eastern Africa: The roles of market access, technology and household resource endowments. *Food Policy* 43, 248–266.

Mburu, S., Ackello-Ogutu, C., Mulwa, R., 2014. Analysis of Economic Efficiency and Farm Size: A Case Study of Wheat Farmers in Nakuru District, Kenya. *Economics Research International* 2014, 1–10.

McDonald, J.F., Moffitt, R.A., 1980. The Uses of Tobit Analysis. *The Review of Economics and Statistics* 62, 318–321.

McNamara, K., Weiss, C., 2005. Farm Household Income and On- and Off-Farm Diversification. *Journal of Agricultural and Applied Economics* 37, 37–48.

Meeusen, W., Van den Broeck, 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* 435–444.

Mesfin, W., Fufa, B., Haji, J., 2011. Pattern, trend and determinants of crop diversification: empirical evidence from smallholders in eastern Ethiopia. *Journal of Economics and Sustainable Development* 2, 78–89.

Minot, N., Epprecht, M., Anh, T.T.T., Trung, L.Q., 2006. Income Diversification and Poverty in the Northern Uplands of Vietnam. Washington, DC: International Food Policy Research Institute (IFPRI). Available at: http://www.ifpri.org/sites/default/files/publications/rr145.pdf

Mishra, A.K., El-Osta, H.S., Sandretto, C.L., 2004. Factors affecting farm enterprise diversification. *Agricultural Finance Review* 64, 151–166.

Mishra, A.K., Goodwin, B.K., 1997. Farm income variability and the supply of off-farm labor. *American Journal of Agricultural Economics* 79, 880–887.

Mofya-Mukuka, R., Hichaambwa, M., 2016. Factors Influencing Smallholder Crop Diversification in Zambia and the Implications for Policy. IAPRI Working Paper. Available at: http://www.iapri.org.zm/images/WorkingPapers/wp112_for_pdf.pdf

Mottaleb, K.A., Mohanty, S., Nelson, A., 2014. Strengthening market linkages of farm households in developing countries. *Applied Economic Perspectives and Policy* 37, 226–242.

Moschini, G., 1989. Normal Inputs and Joint Production with Allocatable Fixed Factors. *American Journal of Agricultural Economics* 71, 1021.

Mundlak, Y., 1996. Production Function Estimation: Reviving the Primal. *Econometrica* 64, 431.

Narala, A., Zala, Y.C., 2010. Technical Efficiency of Rice Farms under Irrigated Conditions in Central Gujarat. *Agricultural Economics Research Review* 23, pp 375-381.

Narrod, C., Roy, D., Okello, J., Avendaño, B., Rich, K., Thorat, A., 2009. Public–private partnerships and collective action in high value fruit and vegetable supply chains. *Food Policy* 34, 8–15.

Newey, W.K., 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. *Journal of Econometrics* 36, 231–250.

Nguyen, H.Q., 2014. Crop Diversification, Economic Performance and Household's Behaviours: Evidence from Vietnam. Available at: https://mpra.ub.uni-muenchen.de/59168/1/MPRA_paper_59168.pdf

Ogundari, K., 2013. Crop diversification and technical efficiency in food crop production: A study of peasant farmers in Nigeria. *International Journal of Social Economics* 40, 267–287.

Oladeebo, J.O., Oyetunde, O.T., 2013. Relationship Between Plot Size and Technical Efficiency of Cassava Farms in Oyo State of Nigeria: A Stochastic Frontier Analysis. *International Journal of Current Research* 5, 2782–2786.

Omiti, J., Otieno, D., Nyanamba, T., McCullough, E., 2009. Factors influencing the intensity of market participation by smallholder farmers: A case study of rural and peri-urban areas of Kenya. *African Journal of Agricultural and Resource Economics* 3, 57–82.

O'Neill, S., Matthews, A., others, 2001. Technical change and efficiency in Irish agriculture. *Economic and Social Review* 32, 263–284.

Ouma, E., Jagwe, J., Obare, G.A., Abele, S., 2010. Determinants of smallholder farmers' participation in banana markets in Central Africa: the role of transaction costs. *Agricultural Economics* 41, 111–122.

Oushy, H., 2010. Farm Resource Management in Afghanistan–Training Course. United States Agency for International Development (USAID) / New Mexico State University (NMSU) / Afghanistan Water, Agriculture, and Technology Transfer (AWATT) project. Available at: https://aces.nmsu.edu/pubs/docs/

Peterson, J.M., 2002. Environmental policies for a multifunctional agricultural sector in open economies. *European Review of Agriculture Economics* 29, 423–443.

Petrin, A., Train, K., 2010. A control Function Approach to Endogeneity in Consumer Choice Models. J*ournal of Marketing Research* 47, 3–13.

Pingali, P., Khwaja, Y., Meijer, M., 2005. Commercializing small farms: Reducing transaction costs. *The Future of Small Farms* 61, 5–8.

Pingali, P.L., 1997. From Subsistence to Commercial Production Systems: The Transformation of Asian Agriculture. *American Journal of Agricultural Economics* 79, 628.

Pingali, P.L., Rosegrant, M.W., 1995. Agricultural commercialization and diversification: processes and policies. *Food Policy* 20, 171–185.

Pope, R.D., Prescott, R., 1980. Diversification in Relation to Farm Size and Other Socioeconomic Characteristics. *American Journal of Agricultural Economics* 62, 554.

Poulton, C., Dorward, A., Kydd, J., 2010. The Future of Small Farms: New Directions for Services, Institutions, and Intermediation. *World Development* 38, 1413–1428.

Ragasa, C., Mazunda, J., 2018. The impact of agricultural extension services in the context of a heavily subsidized input system: The case of Malawi. *World Development* 105, 25–47.

Rahman, S., 2009. Whether crop diversification is a desired strategy for agricultural growth in Bangladesh? *Food Policy* 34, 340–349.

Randela, R., Alemu, Z.G., Groenewald, J.A., 2008. Factors enhancing market participation by small-scale cotton farmers. *Agrekon* 47, 451–469.

Rao, E.J.O., Qaim, M., 2013. Supermarkets and agricultural labor demand in Kenya: A gendered perspective. *Food Policy* 38, 165–176.

Rao, P.P., Joshi, P.K., Kumar, S., Ganguley, K., 2008. Agriculture Diversification in Andhra Pradesh, India: Patterns, Determinants and Implications (No. 2). ICRISAT and IFPRI, New Delhi. Available at: http://oar.icrisat.org/1360/1/533-2008_Research_Report_No_2.pdf

Reardon, T., Berdegué, J., Escobar, G., 2001. Rural nonfarm employment and incomes in Latin America: overview and policy implications. *World Development* 29, 395–409.

Rehima, M., Belay, K., Dawit, A., Rashid, S., 2013. Factors affecting farmers' crops diversification: Evidence from SNNPR, Ethiopia. *International Journal of Agricultural Sciences* 3, 558–565.

Ricker-Gilbert, J., Jayne, T.S., Chirwa, E., 2011. Subsidies and Crowding Out: A Double-Hurdle Model of Fertilizer Demand in Malawi. *American Journal of Agricultural Economics* 93, 26–42.

Satyasai, K.J.S., Viswanathan, K.U., 1997. Commercialisation and Diversification of Indian Agriculture. *Economic and Political Weekly* 31, 3027–3028.

Seng, K., 2014. Determinants of Farmers? Agricultural Diversification: The Case of Cambodia. *Asian Journal of Agriculture and Rural Development* 4, 414.

Shumway, C.R., Pope, R.D., Nash, E.K., 1984. Allocatable Fixed Inputs and Jointness in Agricultural Production: Implications for Economic Modeling. *American Journal of Agricultural Economics* 66, 72.

Sichoongwe, K., Mapemba, L., Tembo, G., Ng'ong'ola, D., 2014. The Determinants and Extent of Crop Diversification Among Smallholder Farmers: A Case Study of Southern Province Zambia. *Journal of Agricultural Science* 6.

Singh, I., Squire, L., Strauss, J., World Bank (Eds.), 1986. Agricultural household models: extensions, applications, and policy, A World Bank research publication. Johns Hopkins University Press, Baltimore.

Smale, M., Kusunose, Y., Mathenge, M.K., Alia, D., 2016. Destination or Distraction? Querying the Linkage Between Off-Farm Work and Food Crop Investments in Kenya. *Journal of African Economies* 25, 388–417.

Stampini, M., Davis, B., 2009. Does nonagricultural labor relax farmers' credit constraints? Evidence from longitudinal data for Vietnam. *Agricultural Economics* 40, 177–188.

Stoeffler, Q., 2016. Crop portfolio choices in Burkina Faso. *The Journal of Developing Areas* 50, 261–285.

Tadesse, G., Bahiigwa, G., 2015. Mobile Phones and Farmers' Marketing Decisions in Ethiopia. *World Development* 68, 296–307.

Tavva, S., Aw-Hassan, A., Rizvi, J., Saharawat, Y.S., 2017. Technical efficiency of wheat farmers and options for minimizing yield gaps in Afghanistan. *Outlook on Agriculture* 46, 13–19.

Thieme, O., 2006. Country Pasture/Forage Resource Profiles Afghanistan. Food and Agriculture Organization. Rome, Italy. Available at: http://www.fao.org/ag/AGP/agpc/doc/Counprof/afgan/afgan.htm

Tobin, J., 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica* 26, 24–36.

Turner, E.C., 2014. Determinants of crop diversification among Mozambican smallholders: Evidence from household panel data (M.S. Thesis). Michigan State University. Available at: https://pdfs.semanticscholar.org/672b/2d12b642473be59f8929a5f53d69d1715b7e.pdf

Vakis, R., Sadoulet, E., de Janvry, A., Cafiero, C., 2004. Testing for separability in household models with heterogeneous behavior: a mixture model approach. Department of Agricultural & Resource Economics, UCB. Available at: https://escholarship.org/uc/item/4hs3g5dj

Van den Berg, M.M., Hengsdijk, H., Wolf, J., Van Ittersum, M.K., Guanghuo, W., Roetter, R.P., 2007. The impact of increasing farm size and mechanization on rural income and rice production in Zhejiang province, China. *Agricultural Systems* 94, 841–850.

Van Dusen, M.E., 2000. In situ conservation of crop genetic resources in the Mexican milpa system (PhD Dissertation). University of California Davis. Available at: http://ageconsearch.umn.edu/bitstream/21889/1/sp00va01.pdf

Van Dusen, M.E., Taylor, J.E., 2005. Missing markets and crop diversity: evidence from Mexico. *Environment and Development Economics* 10, 513–531.

Verkaart, S., Munyua, B.G., Mausch, K., Michler, J.D., 2017. Welfare impacts of improved chickpea adoption: A pathway for rural development in Ethiopia? *Food Policy* 66, 50–61.

Vik, J., McElwee, G., 2011. Diversification and the entrepreneurial motivations of farmers in Norway. *Journal of Small Business Management* 49, 390–410.

Vincent, D.P., Dixon, P.B., Powell, A.A., 1980. The Estimation of Supply Response in Australian Agriculture: The Cresh/ Creth Production System. *International Economic Review* 21, 221.

Vuong, Q.H., 1989. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica* 57, 307.

Wadud, A., White, B., 2000. Farm household efficiency in Bangladesh: a comparison of stochastic frontier and DEA methods. *Applied Economics* 32, 1665–1673.

Wang, H.-J., Schmidt, P., 2002. One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis* 18, 129–144.

Weinberger, K., Lumpkin, T.A., 2007. Diversification into Horticulture and Poverty Reduction: A Research Agenda. *World Development* 35, 1464–1480.

Weiss, C., Thiele, H., others, 2002. Diversifikation und Wachstum landwirtschaftlicher Unternehmen. *Agrarwirtschaft* 51, 156–163.

Weiss, C.R., Briglauer, W., 2000. Determinants and dynamics of farm diversification. FE Workingpaper/Universität Kiel, Department of Food Economics and Consumption Studies. Available at: https://www.econstor.eu/bitstream/10419/23578/1/Ewp0002.pdf

Weltin, M., Zasada, I., Franke, C., Piorr, A., Raggi, M., Viaggi, D., 2017. Analysing behavioural differences of farm households: An example of income diversification strategies based on European farm survey data. *Land Use Policy* 62, 172–184.

Winter-Nelson, A., Temu, A., 2005. Impacts of prices and transactions costs on input usage in a liberalizing economy: evidence from Tanzanian coffee growers. *Agricultural Economics* 33, 243–253.

Woldenhanna, T., Oskam, A., 2001. Income diversification and entry barriers: evidence from the Tigray region of northern Ethiopia. *Food Policy* 26, 351–365.

Wooldridge, J.M., 2015. Control function methods in applied econometrics. *Journal of Human Resources* 50, 420–445.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data, Second Edition. ed. The MIT Press.

Wooldridge, J.M., 2007. What's new in econometrics? Control functions and related methods.

Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data. Cambridge. MA: MIT Press.

World Bank, 2017. Afghanistan Development Update, November 2017. World Bank, Washington, DC. © World Bank. Available at: https://openknowledge.worldbank.org/bitstream/handle/10986/28928. License: CC BY 3.0 IGO

World Bank, 2014. Islamic Republic of Afghanistan Agricultural Sector Review : Revitalizing Agriculture for Economic Growth, Job Creation and Food Security. World Bank, Washington, DC. © World Bank. Available at: http://documents.worldbank.org/handle/10986/21733. License: CC BY 3.0 IGO

Yang, J., Wang, H., Jin, S., Chen, K., Riedinger, J., Peng, C., 2016. Migration, local off-farm employment, and agricultural production efficiency: evidence from China. *Journal of Productivity Analysis* 45, 247–259.

Zereyesus, Y.A., Embaye, W.T., Tsiboe, F., Amanor-Boadu, V., 2017. Implications of Non-Farm Work to Vulnerability to Food Poverty-Recent Evidence From Northern Ghana. *World Development* 91, 113–124.

Zhang, L., Su, W., Eriksson, T., Liu, C., 2016. How Off-farm Employment Affects Technical Efficiency of China's Farms: The Case of Jiangsu. *China & World Economy* 24, 37–51.