USING CACAO TO CATALYZE DEVELOPMENT:

PRODUCTIVITY DRIVERS AND TECHNOLOGY ADOPTION AMONGST SMALLHOLDER FARMERS IN MONTES DE MARIA, COLOMBIA

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Masters of Science in Applied Economics and Management

by

Kalob James Williams

May 2019

ABSTRACT

Smallholder farmers produce a large portion of the world's total food supply, but are often times limited by economic, social or demographic factors that larger farmers find easier to overcome. The body of literature surrounding smallholder farmer crop production is large and addresses a wide range of topics, from gender equality to agronomic considerations. This thesis expands this body of literature by adding a two-step approach that examines what makes some smallholder farmers more productive than others, focusing on the case of cacao. Step one determines which production technologies have the strongest relationship with yields amongst a certain group of farmers, and step two determines which socioeconomic and demographic factors have the largest impact on the adoption of the technologies identified in step one. We use cross-sectional survey data from 277 smallholder cacao producers in the Montes de Maria region of northern Colombia to carry out this process. Based on the findings, we make recommendations that are useful to association leaders and government technicians in the area, who are interested in promoting cacao as an engine for regional economic development. We find that harvest intensity and fertilizer use have strong positive relationships with yields, whereas herbicide use exhibits a strong negative relationship with yields. Our results suggest that association membership status and the number of buyers a farmer sells to have causal relationships with cacao yields, which are mediated positively through an increase in harvest intensity. Finally, we find that formal and informal training are highly associated with the adoption of production technologies, but that formal training seems to be more strongly related to adoption of pruning, grafting, herbicide use and pesticide use, while informal training is more strongly related to increases in fertilizer use and harvest intensity.

BIOGRAPHICAL SKETCH

Kalob Williams first became interested in agriculture in a development context while spending two years as a missionary in the rolling farmland of Central Brazil. After returning from Brazil, he graduated *cum laude* from Brigham Young University in 2014 with a Bachelor of Science in Economics and a minor in Business Management. After graduation, he worked in agriculture as a production foreman, a business analyst and a marketing representative with a multinational, private-sector firm in California, USA that grew almonds, pistachios and olives for oil. In 2017 he enrolled in Cornell University's Charles H. Dyson School of Applied Economics and Management where he has served as both a teacher's assistant for Managerial Economics courses with Dr. Garrick Blalock and a research assistant for Dr. Miguel Gómez in a project focusing on economic analysis of the New York State berry industry. While at Cornell, he also had the opportunity to participate in the Student Multidisciplinary Applied Research Teams (SMART) program, once on a consulting team in Uganda, and again as a team leader on a project in Colombia. The latter was born out of the field research described in this thesis. His current career interests lie in international agricultural development, particularly in the interactions between small- and large-scale farming. Along the way he married and had two children. His family has been his support, motivation and relief through the toughest and most exciting times of his life.

ACKNOWLEDGEMENTS

First, I thank the Cornell Institute for the Social Sciences as their Small Grant provided the primary funding source for this research. I also thank the Cornell College of Agriculture and Life Sciences for the Richard Bradfield Research Award which provided additional funding for this research. Finally, I thank Cornell's Mario Einaudi Center for International Studies for the International Research Travel Grant which helped fund airfare for this research.

Additionally, a very special thanks goes out to Dr. Daniel Toro González and Dr. Pablo Abitbol for the Universidad Tecnológica de Bolívar for their in-country support and expertise on the area of Montes de Maria. I also thank Dr. Ximena Rueda from Universidad de los Andes in Bogotá, Colombia for her collaboration in the project's research design.

I very much appreciate Dr. Deniz Akdemir, Dr. Lynn Johnson, Dr. Erika Mudrak and Stephen Parry from Cornell University's Statistical Consulting Unit, for their help in developing the models we use in this study.

I also thank Eduard Amaya, Leiver Lore, Luis Miguel Ortega and Carlos Perez, the four hard working enumerators that put in long hours to collect the data for this study. I appreciate the aid of Compañía Nacional de Chocolates, S.A. for the use of their quality sampling lab in Valledupar, Cesar, Colombia. This work would not have been possible without the collaboration of Victoria Julio Rocha, Pedro Lora, António Rodriguez, Nays Coheu, Hernando Gonzáles, Oswaldo Garcia, Reuben Quiroz and Geovaldis Jimenez, who are leaders of cacao grower associations in Montes de Maria, and enabled contact with growers and navigation through the region.

Furthermore, I personally thank Dr. Miguel Gómez from Cornell University, my advising professor, for his guidance, direction, encouragement and dedicated support throughout this entire process. He has been the key enabler and catalyst for this project. I also thank Dr. David Lee from Cornell University as a member of my special committee for his advice and input throughout this process.

More than anything, I thank the incredible citizens and cacao producers from Montes de Maria that gave of their time, knowledge, data and ideas to help formulate, motivate and investigate the research questions addressed here on their behalf.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Ha – Hectares

ICCO – International Cocoa Organization

M – Million

MdM – Montes de Maria

MLE – Maximum Likelihood Estimation

NdC – Nacional de Chocolates

OLS – Ordinary Least Squares

SEDG – Socioeconomic and Demographic

SEM – Structural Equation Modeling (Model)

Introduction

Over the last several decades, the advent of advanced communications and surveying systems has made the importance of smallholder farmers to food security and international development goals increasingly clear. Currently, large portions of the food supply around the world are grown by smallholder farmers – e.g. as of 2017, smallholder farmers produced 80% of the food supply in sub-Saharan Africa and Asia (Food and Agriculture Organization of the United Nations, 2017).

These smallholder farmers face not only the typical challenges and risks inherent in any agricultural operation, such as market price fluctuations and severe climate events, but also a set of challenges unique to them due to their relative size to other players on the world food stage. Though the proportion has been declining in recent years, many smallholder farmers still live in poverty (Food and Agriculture Organization of the United Nations, 2017). This poverty has driven numerous former farmers to urban areas in search of better employment and has, to some degree, enabled the consolidation of agricultural lands. Through the mechanisms of urban pull or rural push (Barrett, Christiaensen, Sheahan, & Shimeles, 2017), farmers are leaving or selling their land, many times to larger, more established farmers or corporations. Land consolidation allows larger farmers to reduce costs and increase production through economies of scale, and puts pressure on remaining smallholder farmers to increase performance or sell their land. Thus, the cycle described above perpetuates itself For farmers that decide to stay and compete, the adoption of new technologies becomes a critical factor in their survival, and often times those who are slow to adopt or cannot adopt are left behind or forced out (Food and Agriculture Organization of the United Nations, 2017).

What can and should be done about these challenges is a topic of much debate and research, but the fact remains that smallholder farmers need to become more competitive if they wish to survive. Food production amongst smallholder farmers takes on many forms, from fruits and vegetables to livestock and spices. A few cash crops are produced primarily by smallholder farmers including palm oil, rubber, coffee and cacao (Byerlee & Rueda, 2015). As smallholder farmers that produce these crops are not going toe-to-toe with larger farms with considerably more resources, these crops offer a more level playing field to smallholder farmers, one in which they may find it more profitable to compete. In recent years, with the advent of labeling schemes such as Fairtrade, Rainforest Alliance, etc., smallholder farmers are also well positioned to capitalize on market premiums that come from moving from the mainstream of a market to a niche (Murphy, 2012). Some farmers are even able to combine the previous two effects and distinguish themselves in a niche market for a product that is primarily grown by smallholder farmers, more effectively protecting themselves from global agricultural consolidation.

Another common method of increasing competitiveness of smallholder farmers is to employ resources to boost yields, as measured by volume of production per unit of land. Employing improved production techniques – i.e. agricultural technologies and practices – that increase yields can generally lead to increases in profit margins and make producers better off economically without a need to increase farm size. Developed countries, which commonly house larger family and corporate farms, have easier access to the technologies necessary to increase yields, whereas less-developed countries tend to house higher proportions of smallholder farmers, who find these technologies more elusive and difficult to disseminate (Collier & Dercon, 2014). Many government, non-government and development assistance organizations throughout the world have programs focused on providing these technologies to smallholder farmers with hopes of reducing poverty in a sustainable way (Conservation Alliance International, 2013; Food and Agriculture Organization of the United Nations, 2017; Funcicar, 2015; USAID Colombia, 2017; United Nations Development Programme, ISA Interconexion Electrica S.A. E.S.P, Compania Nacional de Chocolates, 2018). In order to be effective, these programs need to know not only which technologies have the largest effects on yields, but also which factors affect the adoption of those technologies. By combining this information with knowledge of crops and markets in which smallholder farmers can have a competitive edge, these organizations can tailor their efforts to maximize impact.

The purpose of this study is to use cacao production in the Montes de Maria (MdM) region of northern Colombia as a case study for developing a system that first considers which technologies have the greatest effects on yield for a particular crop, and then which socioeconomic and demographic (SEDG) factors most effect the adoption of those technologies. While each of these two aspects have been examined separately in various cases, we lay out a method to consider them simultaneously, and examine the benefits of doing so. To do this, we surveyed 277 smallholder farmers of cacao in MdM during the summer of 2018 about both their production practices and key SEDG factors. Using these data, we developed a production function and a set of technology adoption functions that allowed us to examine these relationships. We then employed structural equation modeling to investigate causality in these estimated relationships.

As we seek to understand the effects of SEDG factors on technology adoption amongst smallholder farmers, cacao is a useful crop for this framework as 80-90% of world cacao production is carried out by smallholder farmers (Gayi & Tsowou, 2016). It is also useful because cacao is a crop that is grown with the primary purpose of increasing cash income to the farmer, rather than being directly consumed by the farmer and his/her family. Likewise, MdM as a region serves this study well since, as will be discussed fully later, cacao was introduced to this region very recently and efforts to train farmers on production practices are in their nascent stages. This has caused significant variability in the adoption of many crucial technologies. Knowledge gained through this study will be beneficial not just to farmers, but also to policy makers and organizations targeting development and poverty reduction in regions of high concentrations of smallholder farmers.

The rest of this thesis will formulate and explain the execution of methods used to answer the research questions identified above. First, it will evaluate literature relevant to the topic at hand and state the contribution to the literature. Next it will review the methodology used in the field research in MdM and subsequent analysis. Then it will discuss the results of the study, as well as some key limitations. Finally, this thesis concludes with recommendations for development practitioners and areas for potential future research.

Literature Review

To provide context for the research questions at hand we will begin with an overview of the global cacao industry. Next, we will review studies focusing on the effects of production practices and SEDG factors on crop yields, with a special emphasis on knowledge relating to cacao. Finally, we will review research related to technology adoption amongst farmers, again with special emphasis on cacao farmers.

The Global Cacao Industry

Theobroma cacao, also known as just cacao or cocoa, originated in the jungles of northern South America, east of the Andes mountain range (Young, 2007). Ancient Mayan and Aztec cultures consumed the fruit in drink form much like coffee is consumed today. With similar caffeine content, this drink gave its consumers energy and thus earned cacao the name “Food of the Gods,” or in Greek “Theobroma.” With time and the development of transatlantic shipping routes, the fruit found its way to Africa where commercial production was developed during the last half of the 19^{th} century. Since then production has expanded to many tropical countries located in the belt around the world that lies 15 degrees north and south of the Equator. Today a few cultures still consume cacao in drink form, but most cacao is processed further into the ever-popular chocolate (Young, 2007).

According to the International Cocoa Organization (ICCO), in 2016, three quarters of the world’s cacao supply came from Africa, with the Ivory Coast and Ghana being the largest producers, with 40% and 19% of total supply, respectively. Ecuador and Brazil lead production for South and Central America, having produced 17% of the world’s cacao in 2016. The remainder came from Indonesia and other Southeast Asian countries. On the processing side, Europe and the Americas accounted for 60% of the world’s cacao grindings in 2016. All but the small portion of cacao that was ground in Brazil was ground in non-cacao producing countries. The remainder was ground in Africa, Asia and Oceania.

Being a globally-traded commodity, cacao futures contracts trade in three locations: ICE Futures in New York City, and ICE Futures Europe and CME Europe, which are both located in London. Global cacao production (supply) over the last four years has averaged 4.40M metric tons annually, while grindings (demand) has averaged only 4.30M metric tons over the same period. Looking even further back, global cacao demand has increased by 1.9% annually over the last 10 years while supply has increased by 2.5% annually over the same period (International Cocoa Organization, 2018). These facts taken in isolation might suggest coming challenges for smallholder cacao producers as supply seems to be outstripping demand, though, as with any global commodity market, effective forecasting can prove difficult. Some of the larger industry players have been cited as projecting a more stable supply-demand relationship in coming years, and possibly even an undersupply in the coming decades. Other parties that benefit less from a drop in global cacao prices, such as the International Cocoa Organization, take more moderate stances (International Cocoa Organization, 2014).

Between 90% and 95% of all cacao produced worldwide is classified by markets as a commodity and known as “bulk” cacao (International Cocoa Organization, 2018). This is the cacao that is traded on futures markets and is what makes it into a majority of chocolate products throughout the world. The remaining 5% to 10% is sold as “fine flavor” cacao which, due to its varietal origins and favorable flavor profiles, can fetch a large premium over bulk cacao. While most of the world’s cacao is produced in Africa, nearly all fine flavor cacao is produced in Central and South America (International Cocoa Organization, 2018). Along with fine flavor premiums, some cacao markets are willing to pay for a more direct connection to the producer through “source verified” supply chains, a model that shows promise for smallholder cacao farmers throughout Central and South America (Byerlee & Rueda, 2015). The “Farm to Fork” equivalent in the cacao industry is known as the “Bean to Bar” movement, and for the last decade this movement has established a strong presence amongst specialty chocolatiers in the

United States and Europe, offering new opportunities for smallholder farmers in Central and South America.

Factors Affecting Cacao Yields

There is a large body of research that strives to isolate the effects of specific factors on yields of various crops. Using the basic production theories common in agricultural economics, researchers examine production practices and SEDG factors to tease out the effects of these factors on yields (Griliches, 1963; Griffin, Montgomery, & Rister, 1987; Taher, 1996; Rosenberg & Marcotte, 2005; Almeida, Chaves, Bonomo, Almeida, & Fernandes, 2014). When examining production practices, studies may look at continuous measures, such as fertilization use, irrigation quantity, agrochemical use, pruning intensity, etc. to determine the effect of a one-unit change in one of these inputs on expected yield for a given crop in a given region (Ramirez et al., 2001; Rosenberg et al., 2005; Almeida et al., 2014). When continuous data are not available for input use due to lack of recordkeeping, such as is common with smallholder farmers, researchers may look at production practices in a binary fashion – e.g. whether or not a farmer irrigates – to determine which production practices are having the largest effects on yields in a given area (Karli & Gul, 2015). This knowledge is useful not only for farmers, who can use it to determine the expected costs and benefits of employing certain production practices at different levels, but is also informative for development practitioners and donors seeking to promote the largest impacts on yields for populations with limited resources.

Other studies focus on the effects that SEDG factors have on yields (Lee, 2005; Knowler & Bradshaw, 2007). Researchers may look at variables such as age, gender, education, wealth or household size to understand each of their effects on yield potential. They may also measure factors such as social capital, training, access to inputs and infrastructure that have less to do with the individual producer and more to do with their environment and fixed effects. Though this information may be interesting to farmers, it is often more directly useful to development practitioners such as grower association leaders, government extension agents and domestic and foreign aid workers in that it can help them understand factors beyond production practices that are affecting productivity and livelihoods. Examining the effects of both production practices and demographic and socioeconomic factors is key to understanding how to increase productivity for a given region. Here we will review literature for both aspects, while focusing on the small body of literature that uses cacao as their crop of interest.

First, we address what has been found relating to how production practices affect cacao yields or productivity more generally. A study performed in Indonesia shows that mixed agroforestry systems that include wood trees tend to increase the biomass of cacao trees as compared to other systems including fruit trees or monocultures of cacao (Muhardi & Effendy, 2017). This same study posits that monoculture cacao production does not optimize net income, whereas a polyculture system that includes woody trees and/or fruit trees maximizes revenue for growers. Similarly, Ramirez, Somarriba, Ludewigs, & Ferreira (2001) suggest that production and price risk are reduced for cacao farmers in Central America by diversifying away from monoculture production of cacao and towards an agroforestry system that includes timber trees as well. In an article that looks into the agronomic science behind the cacao production function, Almeida et al. (2014) find an optimal yield of 1,026 kg/ha when considering the effect on yields of nitrogen fertilizer use and irrigation applications for a particular region of cacao farmers in Bahia, Brazil. They identify a positive relationship between yields and nitrogen fertilizer use, as well as a quadratic relationship between yields and irrigation application. Furthermore, they were able to break down these associations with yield increases into increases in seeds per cacao pod, weight of the seeds per pod, and number of pods per tree, all of which were positively and statistically significantly correlated with nitrogen fertilizer use and irrigation use. Although these positive associations are to be expected based on basic agronomic science, having quantitative results like these facilitates cost-benefit analysis.

Second, we address what has been discovered in relation to the effects of SEDG factors on cacao production and yields. Navarrete & Rahman (2014) examine productivity management in terms of efficiency and effectiveness, and find that a farmer's years of education, negotiation ability, reliance on cacao income, and certification possession have positive effects on overall productivity amongst cacao producers in Tabasco, Mexico. They also stress the importance of considering a *set* of solutions to increase productivity, rather than focusing on individual practices, as they find that sets of solutions typically yield better results than do measures focusing on one solution. Hes, et al. (2017) focus on the effect that social capital has on cacao production amongst smallholder farmers in the same region of Mexico. They aggregate measures of social capital into four categories – trust, education, social interaction, and network resources – and find that none of these variables significantly affects costs of production and that only social interaction significantly affects the effort farmers put towards tending their cacao plantations. Karli & Gul (2015) identify farm size, farmer age and farmer experience as SEDG factors that have the greatest effects on farmers' yields amongst cacao farmers in the eastern region of Ghana. In this study, they report a mean yield of 345 kg/ha which, when compared to the yields suggested by Almeida et al. (2014), shows the yield spread experienced by cacao farmers throughout the world.

More detailed research that generates quantitative results of the effects of specific cacao production practices would greatly boost the industry. However, these efforts are complicated by the general lack of recordkeeping amongst the smallholder farmers that produce the bulk of cacao worldwide. As a consequence, future research would benefit from experiments collecting data from a specific farm which subjects subplots of land to different production practices.

Factors Affecting Technology Adoption

Those who seek to understand the factors that affect technology adoption amongst smallholder farmers are typically researchers or development practitioners. By understanding why certain smallholder farmers adopt technologies and others do not, practitioners can target their extension or other outreach efforts to have the greatest effect. Much has been done over the last half century to understand technology adoption across a myriad of crops in the developing world.

One important distinction in the agricultural technology adoption literature is the use of metrics of aggregate adoption versus adoption of individual technologies (Feder, Just, & Zilberman, 1985). Each can be useful in its own right, but whereas the use of general adoption metrics, perhaps in the form of an index of adoption, can offer a general perspective, examination of the adoption of individual technologies can allow practitioners to be more “surgical” about their intervention methods. Another key point in the agricultural technology adoption literature is that measuring levels of adoption – e.g., quantity of fertilizer applied – is preferred to binary metrics of adoption – e.g., was fertilizer used: yes or no? – as the former offers much more quantitative insight into development policy (Feder et al., 1985). As with metrics of input use, recordkeeping amongst smallholder farmers often precludes accurate measurement of levels of adoption, whereas binary measurements of technology adoption are much simpler and more easily collected (Lee, 2005; Knowler & Bradshaw, 2007).

Researchers have identified several SEDG factors that have significant effects on technology adoption amongst smallholder farmers. Feder et al., (1985) perform an exhaustive review of literature available at the time relating to farmer technology adoption in developing countries. They suggest that education and access to credit positively affect technology adoption. They also suggested that farm size is not correlated with technology adoption, a claim that has since been refuted, as larger farms have been found to adopt technology at higher rates (Ersado, Amacher, & Alwang, 2003; Ferguson & Olfert, 2015). Sheahan & Barrett (2014) examine input use across sub-Saharan Africa and find a consistent inverse relationship between farm plot size and input use intensity. Additionally, they find that men use and own more inputs or technologies than women. Also, importantly, they find that national level factors explain almost half of the use of fertilizers and agrochemicals in sub-Saharan Africa, suggesting that the factors driving technology adoption amongst a group of farmers in one country may be significantly different than those for a group of farmers in another country, even if the same crop is being considered. Kosarek, Garcia, & Morris (2001) scrutinize diffusion (or adoption) rates of hybrid maize varieties throughout Latin America and the Caribbean and find that perception of profitability and private participation in the supply chain of seeds have positive and significant effects on diffusion rates. They find that the level of commercialization of a product positively affects technology adoption and were able to confirm the logical theory that the price of the seeds (the technology in their case) had a negative marginal effect on technology adoption. Finally, several studies use logistic regression to identify factors affecting agricultural technology adoption amongst smallholder farmers (Feder & Slade, 1984; Akudugu, Guo, & Dadzie, 2012) . Feder and Slade identify a lag in technology adoption, and explain this lag partially by a lack of resources – human and capital – to allocate to the acquisition of knowledge, assuming that a certain critical level of knowledge must be obtained before a farmer will adopt a new technology.

Turning our focus to cacao, we examine literature that addresses which SEDG factors affect adoption of technologies amongst smallholder farmers of cacao throughout the world. Baffoe-Asare, Danquah & Annor-Frempong (2013) use a Tobit model to examine an adoption index of pest and disease control methods and production technology packages amongst smallholder cacao farmers in the central region of Ghana. They find that experience, training, age of household head, household size and social capital have positive effects on technology adoption, while the age of the farm has a negative effect. Also, in Ghana, Aneani, Anchirinah, Owusu-Ansah & Asamoah (2012) use a multinomial logistic regression approach to identify factors influencing the level of adoption of technologies recommended by the Cocoa Research Institute of Ghana such as pesticide and fungicide use, hand weeding, fertilizer use and advanced variety selection. They find that credit, number of cacao farms owned by the farmer, gender, age of the cacao farm, migration status of the farmer, cacao farm size, and cacao yield significantly affected levels of adoption amongst these farmers. Taher (1996) investigates the question of technology adoption amongst smallholder cacao farmers in Indonesia and finds that migrants to an area are less likely to adopt fertilizer and herbicide use than are those indigenous to an area. He also finds that social capital, as measured by number of intimate relationships held by a farmer, has a negative effect on fertilizer use.

In Ecuador, Rueda, et al. (2018) find that those who adopt technologically advanced varieties generally have larger, younger plantations and that a higher percentage of their land is devoted to cacao production. They also find that smallholder farmers use agricultural inputs such as non-organic fertilizers, pesticides and irrigation water at higher rates than non-adopters, suggesting the possibility that farmers who have already adopted one technology are more likely to adopt another. Slightly farther north in Belize, Rosenberg & Marcotte (2005) underline the importance of technology adoption by smallholder cacao farmers in order to generate positive returns to land. They find that the production practice most likely to increase the net present value of a plantation, through increases in yields, is grafting.

Previous studies use a strong combination of metrics of aggregate adoption and metrics of adoption for individual technologies. The literature examines levels of adoption where data are available and binary metrics of adoption where it is not. A more technical literature is present in West Africa than in Latin America or Southeast Asia, which makes sense when considering the prominence of West Africa in global cacao production. It is important to keep in mind, however, that factors affecting technology adoption amongst one group of farmers may or may not have similar effects on another group in a different region due to cultural variables that are difficult to observe.

This study aims to combine several of the factors outlined above in order to offer an additional method for various types of development practitioners to understand which production practices are key for a given group of smallholder farmers, and what may be some of the best ways to go about promoting adoption of those production practices. As outlined above, studies have examined how binary measures of production practices affect cacao yields. They have also examined which SEDG factors affect the adoption of production practices. The main contribution of this study to the literature will be the combination of these two methods and a look at the practical implications of the information gained. It is, we believe, a more holistic approach to development strategy, one that can provide deeper and more insightful understanding than either method can provide individually.

Methodology

Our aim with this thesis is to make our methodology easily replicable and adaptable to various crops and regions throughout the world. Being able to identify the production practices with the largest impact on yields, and then identify which SEDG factors may affect the adoption of those practices is not only locally important to development practitioners and association leaders in MdM as they develop training curriculum and training techniques, but also more broadly to the agricultural development community. Our approach was to first survey smallholder farmers to understand their specific SEDG characteristics and cacao production practices. We then build a model to identify the most important yield-boosting production practices, and then a set of models to identify factors affecting technology adoption. An analysis of the results of these models allows us to map the key factors affecting yields. Finally, we supplement this analysis by estimating all models simultaneously to identify mediation effects. The details of this process follow.

For the purpose of this study, we collected data from 277 small cacao growers in MdM. The data we collected were used to populate one model that examines associations between production practices and SEDG factors and yields, and a second set of models that examines associations between SEDG factors and the adoption of production practices. It is important to note that the production practices considered in the yield question are the same as the technologies considered in the technology adoption question. Table 1 in the Data Overview section of this thesis lays out 1) the expected associations between indicators of production practices and yields, as well as the expected directions of associations between SEDG factors and yields and 2) the expected direction of association between SEDG factors and the adoption of technologies.

Key amongst the SEDG variables identified in the literature are what are often called policy variables, or variables that policy makers, grower association leaders and development practitioners have some control over and can use in the construction and implementation of development strategies. In this analysis we will include association membership, receipt of financing and receipt of training as key policy variables through which interventions may affect change. We will also include as a policy variable a metric of the number of buyers that a farmer sells cacao to, based on the assumption that development practitioners can also affect change by focusing efforts on supply chain development. These variables are described in Table 1, and will be given special consideration in the Results and Discussion section of this thesis.

The remainder of this section will outline the theoretical, contextual and methodological underpinnings of the research performed in MdM and the analysis of the data collected there. First, we provide useful background information of the study area. Second, we lay out the methods of sampling and data collection, followed by an overview of the data and descriptive statistics of key variables. Finally, we outline the analytical frameworks and methods used in the study.

Study Area

MdM is a geographic region divided between the northern states of Sucre and Bolivar in Colombia (see Figure 6 in the Appendix for a map of MdM). It covers 6,466 km^2, and is comprised of 15 municipalities that are between 85% and 95% rural. Topographically, it is a relatively flat region between 32 and 300 ft. (10 and 100 meters) above sea level that is traversed by the San Jacinto mountain range that tops out at 2,100 feet (650 meters) in some parts inside MdM (Funcicar, 2015). The temperature averages 80°F (27°C) with little fluctuation throughout the year due to its proximity to the Equator. MdM receives 46 inches (1,148 mm) of rainfall each year with a notable dry season from January to March and a slight abatement from the otherwise intense rains during July and August (Weather Spark, 2018). These climatic conditions make it ideal for growing tropical crops such as cacao.

The economy in MdM is primarily driven by agriculture, which has traditionally been comprised almost entirely of smallholder farmers, but recently is giving way to some degree to more commercial operations (Funcicar, 2015). The mountainous region produces a large number of crops, of which avocado, yucca root, corn, yams, tobacco and a few fruit crops are the most economically important for smallholders. Palm oil has gained a strong presence, but the drive has been primarily through the commercial sector (Funcicar, 2015).

Armed violence has had a large impact on MdM since the 1950's, with the primary point of conflict being land reform. Guerrilla forces and paramilitary groups have each, at times, had a strong presence in MdM. As it is such a small, remote region, police and military forces often did not dedicate enough resources to MdM to properly address the violence these groups perpetuated. Additionally, the region's mountainous and rural terrain made it difficult for law-enforcement officials to access, and thus provided prime conditions for those seeking to carry out illegal activities. MdM had become a favorite route to the Caribbean for drug traffickers. As a result, from the 1970's to the early 2000's, MdM experienced high rates of homicide, massacres and disappearances leaving many residents living in fear. This fear caused hundreds of thousands of "campesinos" or "peasants", especially those in remote areas, to migrate to larger surrounding cities such as Sincelejo, Sucre or Cartagena, Bolivar in search of a safer life. From 2008 to 2010, however, the violence abated to a great degree in Colombia, but in particular in MdM. Paramilitary and guerilla groups began to leave MdM and campesinos started to return to their land in search of some semblance of the life they knew before (Funcicar, 2015).

According to our interviews with industry players and farmers in MdM, in the early 2000's avocado production, the primary cash source for most smallholder farmers in MdM, was weakened by a strain of mold called *phytophthora*, which greatly reduced yields and made avocados much less profitable to produce. Since that time, producers have been in search of a suitable cash crop replacement for avocados. When government and non-government organizations started to commercially introduce cacao as a potential cash generator at the beginning of the decade, growers paid attention, as it showed promise of financial security. Currently cacao production in MdM is largely focused in the mountainous region and accomplished without exception by smallholder farmers. Most cacao farmers grow various other crops and dedicate an average of only 1.6 hectares – or 27% of their productive land – to cacao production.

When examining profitability drivers for cacao production in the MdM, we found the following: 1) yields average 135 kg/ha, easily one tenth of production in other parts of the country and the world (Conservation Alliance International, 2013); 2) 100% of MdM cacao is sold as *corriente*, which is the lowest cacao grade in Colombia; and 3) 95% of all cacao is sold through intermediary buyers to one company, Nacional de Chocolates (NdC), leaving growers without price negotiation power. "Growing pains" such as these can be expected when introducing a new cash crop to a geographically and commercially secluded region such as MdM. This is why several national organizations such as Fedecacao and Corpoica, as well as international organizations such as the United States Agency for International Development, have been working to strengthen cacao yields, quality, and market opportunities in the region. With the average MdM grower only having seven years of experience growing cacao, there is still much technical training to do to help farmers achieve the full economic potential of cacao.

Sampling and Data Collection

In order to obtain the desired data we developed a survey that collected information from smallholder farmers, and this material was supplemented with personal interviews with key industry players. The survey asked cacao farmers to self-report demographics, socioeconomic characteristics, production practices, post-harvest practices and marketing practices for the calendar year 2017. The calendar year corresponds well with the semi-continuous nature of cacao production in Colombia, which experiences zero production in January and February of each year, effectively making the crop year the same as the calendar year. After an initial visit to MdM in April of 2018 to meet with the leaders of the cacao associations, a draft of the survey was developed based on conversations with them and examination of the literature. The survey was subsequently tested for feasibility with industry experts and the same association leaders. In the beginning of June 2018, a formalized draft was pre-tested twice, once with the group of enumerators that would administer the survey to adjust the language to the local vernacular, and a second time with a sample of growers to determine their ability to understand and answer the questions asked. Four local enumerators administered the finalized survey throughout MdM during seven weeks starting in the middle of June 2018. Figure 1 shows the timeline of the fieldwork. See Table 7 in the Appendix for the finalized version of the primary survey tool.

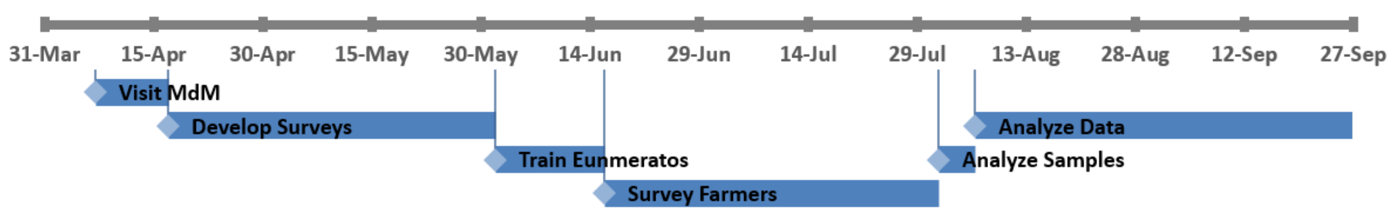


Figure 1: 2018 project timeline.

MdM is home to seven major cacao grower associations ranging from 24 to 180 cacao growers per association. These associations are delineated somewhat geographically, though a boundary overlap is observed between three of the associations. We obtained grower lists from six of the seven associations that totaled 492 growers, and we used these lists for the basis of our sampling pool. For the survey, numerators were assigned geographic regions where they sought out and surveyed any grower that had cacao in 2017 that was at least three years old – a condition necessary to collect the production data upon which much of our analysis is based. Enumerators visited each geographic region where we were aware of the existence of cacao growers and surveyed all growers that met the tree age requirement, independent of association membership status. When visiting a geographic region, they would go to the farms of the growers they were aware of in the region, survey them if they were home, and skip them if they were not. They also asked each of those they surveyed and several people on the road if they knew of other cacao growers in the region. Because of the tight-knit nature of these communities, we feel the enumerators were able to identify a large portion of the growers in each region, even those that were not on our original lists.

Self-reported data were collected from 277 growers, of whom 140 (51%) came from the original grower list and 161 (58%) belonged to a cacao grower association. All farmers were surveyed on their own farm in order to avoid double sampling a farm – i.e. two brothers that claim ownership of the same farm but live in different houses off the farm – and to provide enumerators with the chance to observe certain farm characteristics. We estimate that anywhere from 1,000 to 1,200 smallholder farmers were producing cacao in MdM at the time of this survey, suggesting that this primary survey covered roughly one quarter of the growers in the region.

Data Overview

Table 1 provides descriptions of all variables used in both the yield analysis and analyses of technology adoption. In order to give context to the rest of this thesis, it also provides the mean, standard deviation, and ranges of values for these variables. Hypotheses for the specific relationships between independent variables and dependent variables for both the yield model and the technology adoption models are identified in the "Expected Relationship" columns. If a variable does not have + or - sign in one of these columns, it was not used in that particular model or set of models.

Analytical Frameworks and Empirical Models

This section presents the analytical frameworks and empirical models used to address the research questions, and gives the logic for the inclusion of key variables for each of the empirical

Table 1: Variable names, expected relationships, descriptive statistics and variable descriptions.

	Variable	Short Name	Expected Relationship		Mean	Min	Max	Stdev	Variable Description
			Yield	*Tech Adopt*					
Production Practices/Technologies	Harvest Intensity	Harvest	+++ +		4.60	0	12	3.33	Measure of the number of months a farmer harvested cacao during 2017
	Fertilized	Fertilizer	+++ +		0.44	0	1	0.50	Measure of whether or not the farmer applied fertilizers to their cacao crop in 2017, 1=yes, 0=no
	Applied Herbicides	Herbicides	+		0.16	0	1	0.37	Measure of whether or not the farmer applied herbicides to their cacao crop in 2017, 1=yes, 0=no
	Applied Pesticides	Pesticides	+		0.15	0	1	0.36	Measure of whether or not the farmer applied pesticides to their cacao crop in 2017, 1=yes, 0=no
	Pruned	Pruning	+++		0.79	0	1	0.41	Measure of whether or not the farmer pruned their cacao crop in 2017, 1=yes, 0=no
	Grafted	Grafting	++		0.83	0	1	0.38	Measure of whether or not the farmer grafted their cacao crop in 2017, 1=yes, 0=no
	Hand Weed Frequency	Hand Weed		-	2.67	0	52	4.71	Number of times the farmer weeded their cacao plantation manually in 2017
	Manually Controlled Pests	Manual Pest		-	0.29	0	1	0.45	Measure of whether or not the farmer practiced manual pest control in their cacao crop in 2017, 1=yes, 0=no
Socioeconomic/Demographic	Age	Age	-	-	51.89	20	87	13.43	Measure of farmer's age in years
	Gender	Gender	+	+	0.90	0	1	0.31	Gender of the farmer, 1=male, 0=female
	Years of Education	Education	+	+	5.36	0	23	4.57	Measure of the farmer's formal education in years
	Years of Cacao Experience	Experience	+	+	7.01	1	35	3.98	Number of years the farmer has been growing cacao
	Association Membership	Association	+	+	0.58	0	1	0.42	Measure of whether or not the farmer belonged to a cacao growers association in 2017, 1=yes, 0=no
	Number of Buyers	Buyers	+	+	1.13	0	3	0.61	Measure of the number of buyers the farmer sold their cacao to in 2017
	Ha of Crops Planted	Area Planted	+	+	5.85	0	70	6.24	Total number of hectares that the farmer had planted to any crop on their farm in 2017. Serves as a measure of farmer wealth
	Ha of Cacao Planted	Cacao Planted	+	+	1.56	0	8	1.04	Total number of hectares that the farmer has planted to cacao on their farm in 2017. Serves as a measure of farmer involvement with cacao
	Received Cacao Financing	Financing		+	0.12	0	1	0.39	Measure of whether or not the farmer received financial aid or financing for their cacao crop in the last five years, 1=yes, 0=no
	Formal Skill Training	Formal Training		+	1.52	0	1	1.23	Measure of whether or not the farmer has been trained on the specified production practice by a technician or organization, 1=yes, 0=no
	Informal Skill Training	Informal Training		+	0.31	0	1	0.72	Measure of whether or not the farmer has been trained on the specified production practice by a friend, relative or neighbor, 1=yes, 0=no
	Tech Adoption Index	Tech Adopt			2.37	0	5	1.05	Of fertilizer, herbicide, pesticide, pruning and grafting, a measure of the number of technologies adopted by a farmer in 2017
	Yield per Hectare	Yield			134.89	0	1200	178.6	Kilograms harvested by farmer in 2017 per mature hectare (≥ 3 years)

econometric models. First, we look at the yield question, followed by a discussion of the models used to understand technology adoption, and finally we bring the two together in an examination of potential mediation effects.

Yield per hectare model

Key to our ability to determine the relationship between production practices and SEDG factors and yield per hectare is the selection of an appropriate production function. The simple nature of the data available and the criteria proposed by Griffin et al. (1987), suggest that a linear production function best fits this analysis. Descriptive statistics were calculated and the insights were combined with existing theory to develop the following equation to estimate the production function, which is estimated using Ordinary Least Squares (OLS) regression.

Yield per hectare is estimated as:

$$
\begin{aligned}
lnYield_i = \alpha &+ \beta_1 Fertilizer_i + \beta_2 Herbicides_i + \beta_3 Pesticides_i + \beta_4 Pruning_i + \beta_5 Grafting_i \\
&+ \beta_6 HarvestIntensity_i + \beta_7 Age_i + \beta_8 Gender_i + \beta_9 Education_i + \beta_{10} Experience_i \\
&+ \beta_{11} Association_i + \beta_{12} Buyers_i + \beta_{13} Area\ Planted_i + \beta_{14} Cacao\ Planted_i + \beta_{17} North_i \\
&+ \beta_{18} Central_i + \beta_{19} South_i + \varepsilon_i
\end{aligned}
$$

where α represents a constant, β_1 to β_6 represent parameter estimates for production practices, β_7 to β_{14} represent parameter estimates for SEDG factors, β_{17} to β_{19} represent parameter estimates that control for regional differences, i is a subscript for each farm and ε is a random error term. The natural logarithm of yield is used here as it helps this model meet the assumptions of OLS regression. Details of these variables can be found in Table 1.

Theory and agronomic knowledge support the inclusion of the six production practices, and we would expect to see that the use of any of these practices would have a positive impact on yield, otherwise producers would discontinue the use of the detrimental practice (Conservation Alliance

International, 2013). Pesticide and herbicide use can be seen as more of maintenance practices and would be expected to have the lowest impact on yield, whereas fertilizer use and harvest intensity have more of a direct effect on the productivity of a plant – perceived or real – and thus would be expected to exhibit the largest positive relationship with yields.

Our logic for the inclusion of each SEDG factor is as follows. Since most of these farmers are long-time farmers, having inherited the land from their family and worked on it since they were children, age is highly correlated with general farming experience, and thus, we would likely see a positive association between age and yield. In our field visits, however, we did see that younger farmers tended to be more innovative in their production practices, adopting technologies more readily than older farmers, which could imply a negative association between age and yields. Men are more directly involved with the management of their farms than women in this sample – e.g., we observed that women typically contract out the work on the farm whereas men perform the work themselves – and as such we would expect male farmers to realize higher yields than female farmers. Education increases human capital and the capacity to understand and implement the more technical practices necessary to produce cacao; therefore, we would expect education to positively impact yield. Experience in farming cacao allows farmers to fine-tune practices and is also highly correlated with the age of the orchard in this sample, both considerations that should increase yield. Being a member of an association should provide access to more technical training and social relationships built around cacao, which would certainly be expected to bolster yield. We have included a variable that measures the number of buyers to whom a grower sells, which from empirical experience is very strongly correlated with a buyer's connectivity to the marketplace. By including this variable, we assume that the more connected a farmer is to the marketplace, the more they will perceive that the cacao they grow has value, which will in turn affect the amount of energy and resources they are willing to expend in cacao production. In short, we believe that the number of buyers and yields will be positively correlated. Finally, while the area planted to all crops serves here as our most telling metric of wealth – which ought to make the purchase of inputs easier and yields higher – the area planted to cacao suggests a level of dedication to, and interest in, the crop as well as the possibility for economies of scale in production, leading us to believe that farmers with larger cacao plantations would experience higher yields.

Though the mountainous, cacao-growing region of MdM from which we sampled growers is relatively small and does not vary drastically, we recognize the possibility of slight differences in elevation, soil characteristics and precipitation. To control for this, farmers were divided into one of four geographic regions – north, central, south or southwest – in which the previously mentioned factors representing fixed effects are expected to be similar based on our experience in the field. Though we can break these farmers into smaller geographic regions, the inclusion of more variables in our model is limited by our sample size.

Technology adoption models

To determine the relationship between SEDG variables and the adoption of production technologies, we developed a set of seven models. The first five models employ binary metrics of fertilizer use, herbicide use, pesticide use, pruning and grafting as dependent variables, necessitating the use of binary logistic regression for each. The sixth model uses the metric of harvest intensity – measuring the number of months out of the year that a farmer harvests their cacao – as the dependent variable, while the seventh model looks at factors related to general technology adoption by using as its dependent variable an index of technology adoption. Similar to Baffoe-Asare et al. (2013), this index is simply an equally weighted fraction of the number of technologies adopted out of the total number of binary metrics of adoption we are considering (see Table 1 above for more detail on what is included in this index). These last two models are estimated using binomial logistic count regression.

Each of these models uses Maximum Likelihood Estimation (MLE) to determine parameter estimates using the logistic equation:

$$p_{it} = \frac{1}{1+e^{1(\beta_0+\beta_1 x_{1i}+\beta_2 x_{2i}+\cdots+\beta_p x_{pi})}}$$

where p_{it} is the probability that farmer i adopts technology t, $\beta_0$is a constant, β_1 to β_p are parameters for the relationship between independent variables and the dependent variable, and x_{1i} to x_{pi} are independent variable values for individual i.

More specifically, the general estimation form for our technology adoption models is:

$$\begin{aligned} Adoption of\ Technology_j \\ &= \alpha + \beta_1 Age_i + \beta_2 Gender_i + \beta_3 Education_i + \beta_4 Experience_i + \beta_5 Association_i + \beta_6 Buyers_i \\ &+ \beta_7 Area\ Planted_i + \beta_8 Cacao\ Planted_i + \beta_9 Financing_i + \beta_{10} Formal\ Training_i \\ &+ \beta_{11} Informal\ Training_i + \beta_{12} North_i + \beta_{13} Central_i + \beta_{14} South_i + \varepsilon_i \end{aligned}$$

The fifth and sixth models use a method very similar to simple logistic regression employed in models one through five, but differ slightly in the way they treat an observation. As an example, consider the harvest intensity model, which uses as the dependent variable the number of months out of twelve that a farmer harvested during 2017. Instead of considering each farmer as a single continuous observation, binomial logistic count regression considers each farmer as twelve individual binary observations. If the farmer harvested four months out of twelve, this model would include four observations for which the harvest variable was 1 and eight observations for which the harvest variable was 0, with all twelve observations having the same values for all independent variables. Since the original dataset includes 277 observations, the new data set would include 12 X 277 = 3,324 observations. We employed a statistical software command to generate this larger dataset and perform simple logistic regression on all 3,324 observations.

For continuity and comparison, all seven models include the same basic set of independent variables: age, gender, years of education, years of cacao experience, association membership, number of buyers, hectares of crops planted, hectares of cacao planted, receipt of financing for cacao, formal skill training and informal skill training. We include essentially the same set of SEDG independent variables in the yield model as we do in the technology adoption models to enable the estimation of mediation effects as described in the next section. Consequently, the motivations for the inclusion of these variables in these models have some parallels with the reasoning described for the yield model.

Our logic for the inclusion of each socioeconomic or demographic factor in our basic set of independent variables is as follows. For social and developmental reasons, younger people are typically more open to technology advances, and as such, we would expect age to have a negative effect on technology adoption. Since men in this area are typically more directly involved in the work on the farm, we would expect them to be able to better conceptualize the benefit of a new technology, and thus would expect men to adopt at a higher rate than women. Education would also be expected to increase a farmer's ability to conceptualize the benefits of a new technology and therefore increase adoption rates. As previously mentioned, experience producing cacao is highly correlated with the age of the orchard in this sample, and both of these factors would give a farmer more time to learn how to implement new technologies on their farm in effective ways, thus we would expect more experienced farmers to adopt at higher rates. Being a member of an association would provide farmers more opportunity to hear about technologies in cacao and a setting in which to ask questions about the implementation of technologies, which could reasonably be expected to promote adoption. Having a larger number of buyers to which they can sell strengthens the social network in which a farmer can learn about technologies, but may also increase the perceived value of cacao production as competition is introduced, thus increasing a farmer's willingness and desire to optimize cacao production through the use of production technologies. Similarly, the area a farmer has planted to cacao can represent a farmer's interest in the crop and would be expected to be correlated with higher technology adoption. Both the number of hectares planted to a crop on a farm and the receipt of financing intended for cacao production serve as measures of a farmer's ability to purchase inputs such as agrochemicals and tools needed to implement cacao production, theoretically increasing ability to adopt. Lastly, both formal and informal trainings would be expected to increase technology adoption, as they provide a farmer with a sense of direction and knowhow, the lack of which may have prevented technology adoption before training.

Though this basic set of independent variables is largely the same for all seven models, there are three key exceptions. First, the herbicide use model includes a metric of the frequency with which a farmer weeded their cacao by hand, which would be expected to reduce the need for the use of herbicides but would not necessarily have an effect on the adoption of other technologies. Second, the pesticide model includes a variable that controls for the farmer's use of manual pest control methods – i.e., cutting diseased or infected pods off the tree by hand – which is a common practice amongst these farmers, and would likely reduce the adoption of pesticide use but not necessarily affect the adoption of other technologies. Lastly, each of the models includes a variable for formal training and a variable for informal training. These variables refer to training for the production practice considered in each of the first six models – e.g. formal training in the grafting model asks whether or not a farmer has received formal training in grafting – and in the adoption index model they ask in how many of the five practices considered by the index has the farmer been trained.

Model to test for mediation effects

In our two sets of models we expect to see some of the effects of SEDG factors and policy variables on yields mediated through the technology adoption factors (Hayes, 2013). To understand mediation in the context of the present investigation, consider the following example. We assume that being a member of an association has a direct effect on yield (for multiple reasons), but also that there may be an indirect effect that is specifically channeled through the adoption of fertilizer use, as shown here in Figure 2. In other words, being a member of an association may have a direct relationship with yield, but part of the effect may be realized as association membership promotes or discourages fertilizer use which could indirectly increase or decrease yields.

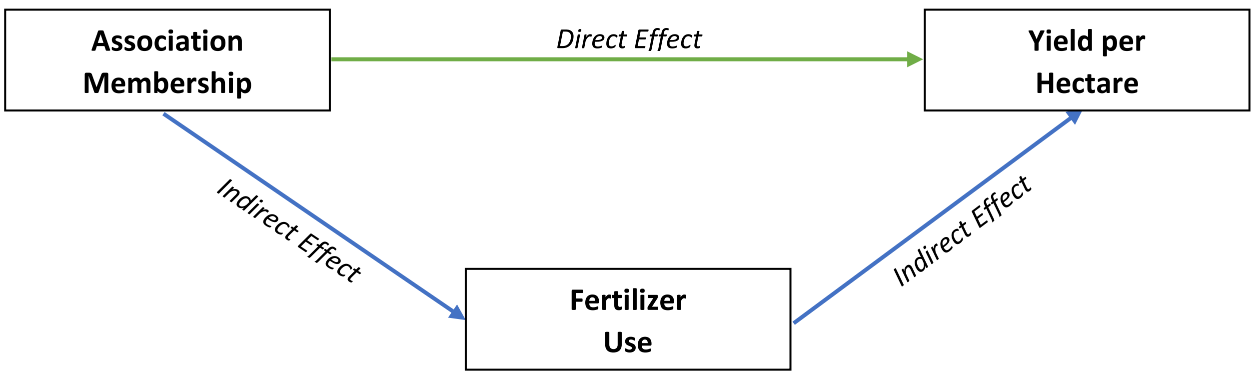


Figure 2: Example of mediation effects.

Understanding and measuring these mediation effects requires not only that we simultaneously estimate the direct and indirect effects identified previously, but also that we estimate all relationships that we expect may exhibit mediation effects at the same time (Hox & Bechger, 1998). When we consider the eleven potential SEDG independent variables identified in our yield and technology adoption models and the six technology adoption factors, the model becomes very complex. Previous research has shown that structural equation modeling (SEM) can be used as a method to address such complexity in the estimation of mediation effects (Hox & Bechger,1998; Hayes, 2013). Using MLE, SEM estimates all parameters between independent, mediation and dependent variables simultaneously, which implies a covariance matrix structure between variables that strengthens the standard error estimates. This model implied covariance matrix can then be compared to that of the observed data to determine a goodness of fit for the model (Hox & Bechger, 1998).

Once parameters are estimated for each of the relationships in the structural model, post-estimation techniques are used to calculate values for the indirect and total effects (Hayes, 2013). Values for indirect effects can be useful on their own, but they are also informative when combined with a measure of the total effect. With these two metrics, we can estimate the percentage of the total effect of an independent variable on a dependent variable that is indirectly due to mediation through another variable. This mediation requires that we perform a non-linear combination of parameter estimates to calculate the indirect effects for the relationship between the independent variable and the mediation variable, and between the mediation variable and the dependent variable.

The example in Figure 3, which is an expansion of the previous example, provides additional detail on the calculation of mediation effects.

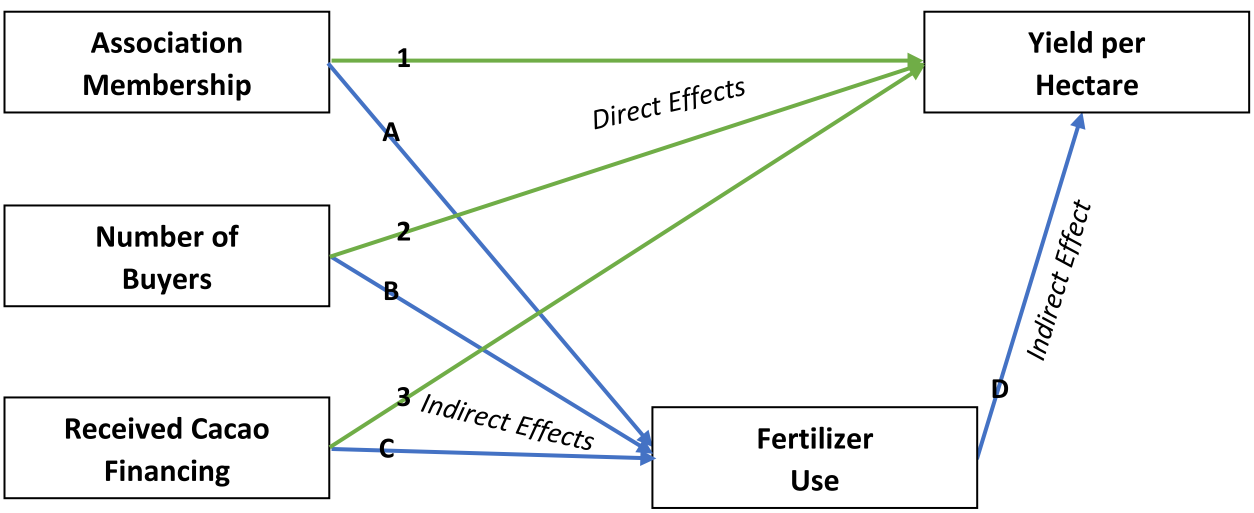


Figure 3: Expanded example of mediation effects.

Lines 1, 2 and 3 represent the parameters or individual direct effects of an independent variable on yield. By performing a non-linear combination of parameters A and D, then of B and D and finally of C and D, we calculate the individual indirect effects of the independent variables that are mediated through fertilizer use. Each of these may have a valuable interpretation in their own right; e.g., we may see that farmers with more buyers have higher yields and that this is partially due to an increase in the use of fertilizers (B-D) and partially due to other factors (2). We can take this a step further, however, and calculate the total direct and indirect effects. By performing a linear combination of parameters 1, 2 and 3, we can estimate a total direct effect of independent variables on yield. Similarly, by performing a non-linear combination of the previously calculated indirect effects – A-D, B-D, and C-D – we arrive at a total indirect effect of independent variables on yield as they are mediated through fertilizer use. We would use the latter step in order to claim causality in a statement such as, "36% of the total effect of independent variables here is mediated through the adoption (or non-adoption) of fertilizer use." In order to include any parameter (A, B, C, D, 1, 2 or 3) in any of these calculations, they must be statistically significant, otherwise they should be omitted from the analysis.

For our model we use STATA Software to perform SEM on a combination of our yield and technology adoption models (see Figure 5 in the Appendix for the path diagram of the full model). As seen in this path diagram, our sample size required that we reduce the number of parameters to be estimated by selecting only five SEDG independent variables from our original group to test for mediation effects. Had the sample size been larger, we would have been able to estimate the exact same models as described in the previous section. The selection of variables was done by considering the statistical significance of each parameter in our original set of models and whether or not an independent variable could be considered a policy variable. The final set of SEDG independent variables to be tested for mediation effects includes gender, hectares of crops planted (as a surrogate of wealth), association membership, number of buyers and receipt of cacao financing.

Results and Discussion

In this section we present and discuss results from all the models described in the Methods section above. General discussion will cover all findings of the models, however, we emphasize the viewpoint of a development practitioner intervening in MdM, striving to influence the yields and well-being of smallholder farmers. For this reason, we will pay special attention to the relationship between yield and the supply chain structure, and yield and policy variables, as those are the mechanisms through which practitioners and development organizations can conceivably affect change. In our discussion these variables include association membership, the number of buyers, the receipt of financing and training.

First, we briefly discuss the results of the yield model and identify those production practices that have the strongest relationships with yields. Then we present results of the technology adoption models to determine how SEDG factors affect the adoption of the production practices with the greatest relationships to yields. Finally, we will discuss the results of the SEM analysis in conjunction with the yield model results to enable a discussion of the direct and indirect effects of production practices and SEDG factors on cacao yields. Note that at first we will only discuss the yield model enough to identify which production practices seem to have the greatest relationships with yield, so that we can focus our discussion on those factors as we review the technology adoption models. A full discussion of the yield model results will come as we discuss the SEM.

Yield Model Results

Table 2 presents the OLS results for the relationship between key production practices and SEDG factors, and yield per hectare amongst cacao farmers in MdM. As the dependent variable is in log form, the coefficients are interpretable as the percentage change in yield related to a one-unit increase in the independent variable.

This model exhibits reasonably good fit with the data with an R-squared value of 0.40 and an adjusted R-squared value of 0.36. This fit would likely had been higher in a study working with larger farmers in a developed country setting where records are kept more meticulously, however, we feel this level of fit permits meaningful discussion of the questions at hand.

Table 2: Regression results - Yield model.

Independent Variable	Coefficient	P-Value	Standard Error
Fertilized	0.374***	(0.00906)	0.142
Applied Herbicides	-0.378*	(0.0914)	0.223
Applied Pesticides	0.268	(0.179)	0.199
Pruned	0.114	(0.507)	0.171
Grafted	-0.0325	(0.851)	0.173
Harvest Intensity	0.160***	(5.33e-09)	0.0265
Age	-0.00132	(0.833)	0.00625
Gender	0.504***	(0.00654)	0.184
Years of Education	0.0207	(0.249)	0.0179
Years of Cacao Experience	0.0595***	(3.74e-06)	0.0126
Association Membership	-0.254*	(0.0971)	0.153
Number of Buyers	0.554***	(7.52e-06)	0.121
Ha of Crops Planted	0.00258	(0.838)	0.0126
Ha of Cacao Planted	-0.234***	(0.000386)	0.0650
North Region	2.109***	(7.39e-08)	0.380
Central Region	1.477***	(1.50e-05)	0.335
South Region	2.000***	(1.25e-08)	0.340
Constant	0.620	(0.258)	0.547
Observations	268		
R-squared	0.40		
Adjusted R-Squared	0.36		
Dependent Variable	ln(Yield)		

Robust p-values in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This model also meets all of the assumptions of the OLS method of regression. The original model exhibited heteroskedasticity in errors, which was corrected for using robust standard errors during subsequent estimations. The original model displayed a non-normal error structure as well. We address this issue by using the natural log of yield per hectare as the dependent variable instead of the original value of yield per hectare. Furthermore, we test for specification error using a link test by regressing predicted values and the square of predicted values on actual values for yield. A significant coefficient on the predicted values and an insignificant coefficient on the square of the predicted values – p-values of 0.000 and 0.105 respectively – suggest a lack of specification error. We also do not see any evidence of multicollinearity as manifest by an examination of the variance inflation factors for each independent variable, all of which are far under the common cut off value of 10. Finally, the Ramsey test for omitted variable bias produces a p-value of 0.09, which fails to reject the null hypothesis that the model has no omitted variables.

Relationships between production practices and yield per hectare.

Three production practices exhibit statistically significant relationships with yield: harvest intensity, fertilizer use and herbicide use. The coefficient of harvest intensity suggests that farmers enjoy a 16% increase in yield for each additional month they harvest during the year. The fertilizer use coefficient suggests that farmers using fertilizers are harvesting 37% more than those that do not. Finally, the coefficient for the use of herbicides exhibits an unexpected relationship both in magnitude and direction. Though only marginally significant, this coefficient suggests that farmers that use herbicides experience yields that are 38% lower than those who do not.

The remaining three production practices – use of pesticides, pruning and grafting – showed statistically insignificant impacts on yield. This does not mean that they are not having an effect on yields, but could instead indicate that they are not highly correlated with yields in this given population. The impacts of all production practices and SEDG variables will be discussed in more detail in the SEM and Direct vs. Indirect Effects subsection of this thesis. For now, we identify harvest intensity, fertilizer use and herbicide use as the factors most significantly related to yields as motivation for our next section.

Technology Adoption Model Results

Table 3 presents the regression results for the seven different technology adoption models considered in this study. Each of these outputs reports odds ratios, meaning that a coefficient value greater than one suggests a positive relationship between the SEDG variable and the metric of adoption (dependent variable), whereas a coefficient value less than one suggests a negative relationship. The goodness of fit for each model as reported by the pseudo R-squared value varies between 0.03 and 0.32. To test for specification error in these models, we regressed the predicted values and the square of the predicted values on the observed values of the dependent variables. These tests indicate that the predicted values are significant predictors of the observed values, and that the squared values are not, suggesting a lack of specification error. We also observed no multicollinearity, as tested for by an examination of the variance inflation factors.

In this section we will first examine the results and implications of the policy variables in these models, and then discuss the relationship between other SEDG variables and the adoption of the various technologies. We will wrap up our discussion with more general findings by examining the technology adoption index model. Special attention is paid in this discussion to results within the fertilizer, herbicide and harvest intensity models, as these are the practices that have shown the most significant associations with yields.

Relationships between policy variables and technology adoption.

Association Membership – Results indicate that association members are 2.2 times as likely to apply pesticides and 2.8 times as likely to prune as non-members. Because the measure of harvest intensity is continuous but the binomial logistic count regression treats it as several binary observations and reports an odds ratio, an interpretation of the magnitude of the odds ratio for association membership in the harvest intensity model is not intuitive. It can, however, be said that being

Table 3: Regression results - Technology adoption models (results reported as odds ratios)

Independent Variables	Fertilized	Applied Herbicides	Applied Pesticides	Pruned	Grafted	Harvest Intensity	Technology Adoption Index
Age	1.001 (0.906)	0.992 (0.585)	1.043** (0.0144)	1.007 (0.654)	0.999 (0.939)	0.997 (0.311)	1.003 (0.570)
Gender	0.916 (0.831)	1.107 (0.872)	0.561 (0.307)	0.114** (0.0407)	0.111** (0.0462)	0.764** (0.0286)	0.705* (0.0573)
Years of Education	0.992 (0.802)	1.031 (0.488)	1.107** (0.0274)	1.044 (0.346)	0.996 (0.943)	0.977** (0.0115)	1.014 (0.299)
Years of Cacao Experience	0.964 (0.295)	1.041 (0.325)	0.928 (0.216)	1.036 (0.522)	0.970 (0.489)	1.003 (0.737)	0.988 (0.392)
Association Membership	1.244 (0.460)	0.793 (0.596)	2.191* (0.0961)	2.794** (0.0134)	1.310 (0.574)	1.334*** (0.00142)	1.224 (0.125)
Number of Buyers	1.310 (0.201)	1.036 (0.889)	1.344 (0.383)	1.721* (0.0576)	1.242 (0.489)	1.785*** (0)	1.182* (0.0812)
Ha of Crops Planted	1.012 (0.543)	0.979 (0.579)	0.941 (0.256)	1.089 (0.103)	1.168** (0.0272)	1.010* (0.0976)	1.008 (0.384)
Ha of Cacao Planted	1.107 (0.422)	1.002 (0.994)	1.704*** (0.00541)	1.009 (0.964)	1.101 (0.704)	1.052 (0.181)	1.090 (0.132)
Received Cacao Financing	1.266 (0.548)	1.014 (0.983)	4.450*** (0.00401)	1.266 (0.696)	0.785 (0.712)	0.694*** (0.00342)	1.214 (0.268)
North Region	1.109 (0.852)	-	0.511 (0.560)	1.484 (0.560)	5.940** (0.0169)	0.151*** (0)	1.221 (0.431)
Central Region	1.011 (0.982)	0.193*** (0.00837)	1.488 (0.661)	2.271 (0.160)	19.01*** (1.55e-05)	0.181*** (0)	1.613** (0.0308)
South Region	0.929 (0.876)	1.065 (0.904)	1.300 (0.769)	7.130*** (0.00109)	41.76*** (2.08e-07)	0.151*** (0)	2.272*** (0.000195)
Formal Skill Training	1.161 (0.661)	3.709* (0.0636)	1.572 (0.353)	4.333*** (0.000200)	7.373*** (6.06e-05)	1.350** (0.0144)	1.129** (0.0128)
Informal Skill Training	2.918* (0.0613)	-	1.026 (0.984)	1.706 (0.326)	0.637 (0.487)	1.927* (0.0795)	0.979 (0.787)
Hand Weed Frequency	-	1.036 (0.347)	-	-	-	-	-
Manually Controlled Pests	-	-	0.0939*** (0.000272)	-	-	-	-
Constant	0.500 (0.450)	0.411 (0.503)	0.00652*** (0.00141)	0.458 (0.610)	0.663 (0.802)	1.914** (0.0190)	0.337*** (0.00909)
Observations	277	238	277	277	277	3,324	1,385
Pseudo R-Squared	0.0247	0.139	0.193	0.231	0.316	0.0740	0.0250

Robust p-values in parentheses
*** p<0.01, ** p<0.05, * p<0.1

a member of an association is correlated with more intense harvesting. This could perhaps be because farmers have increased social connectivity which promotes imitation – "I need to harvest because my fellow association members are harvesting" – or because farmers are taught to harvest more intensely at association meetings or in association trainings. The possibility that being a member of an association may be positively effecting yields through an increase in harvest intensity will be further explored in the Mediation Effects section of this thesis.

Number of Buyers – Each additional buyer that a farmer sells to is correlated with an increase in the odds that a farmer prunes by 1.7 times. Similar to the odds ratio for association membership in the harvest intensity model, the magnitude of the odds ratio for number of buyers offers no intuitive interpretation but does allow for a directional interpretation. The observed association between number of buyers and harvest intensity is positive in our model, and may be due to an increase in the perceived value of cacao as a crop that promotes more labor input and therefore more intense harvesting. It may also simply represent a lack of buyers in the area that deters farmers from harvesting, though the latter explanation seems unlikely from our experience in the field. The causal effect of buyers on yield as mediated through an increase in harvest intensity will be further explored in the Mediation Effects section of this thesis.

Receipt of Financing – Recipients of cacao financing are 4.5 times as likely to use pesticides as non-recipients. Interestingly, however, recipients of cacao financing harvest less intensely than non-recipients. Since this measure of financing includes not just traditional debt financing but also many forms of government or non-government grants that need not be repaid, this finding may suggest a free-rider effect. Farmers that receive cacao financing may use the money to support themselves and their families effectively reducing the need to harvest their cacao as a means of income generation, though the data does not allow for confirmation of this possibility. Our attempts to determine if financing is having a negative effect on yields as mediated through harvest intensity yielded no significant results, and therefore we are limited to the previous analysis of the facts.

Training – As would be expected a positive, and relatively significant, association is present between training for a specific production practice and the adoption of that practice. Formal training seems to generally be more highly correlated with adoption than informal training. Those who have been formally trained in herbicide use, pruning and grafting adopt these practices 3.7, 7.1 and 41.8 times more than those who have not been formally trained, respectively. We also see that formal training is positively correlated with harvest intensity. Farmers informally trained use fertilizer 2.9 times as much as those not trained, and there is a positive association between informal training and harvest intensity.

A few key points need to be made from these observations, especially because training is one of the key avenues used by development practitioners to promote change. First, informal training is strongly correlated with fertilizer use – which is strongly and positively correlated with yields – whereas formal training does not seem to be correlated with increased use. This could be because fertilizer use is a common element in growing many crops produced by these farmers, and therefore more farmers are able to train their neighbors on this technique than are able to train on practices like pruning and grafting which are somewhat unique to cacao in these farmers' crop portfolios. Indeed, our experience in the field suggests farmers readily share knowledge on how to fertilize, but await more "qualified" training for more cacao-specific practices such as pruning. Second, development practitioners need to consider the possibility herbicide-use trainings are not being fully understood, as formal training is having a positive effect on herbicide use, which is having a negative effect on yields. This does not mean that they should stop training because herbicide use is bad for yields, but instead that trainers need to understand how herbicides are being applied and identify potential ways in which herbicides could be negatively affecting yield – e.g., overspray on trees or saturation in the soil and uptake by trees – that they can then address in future trainings. Lastly, both formal and informal training are positively correlated with harvest intensity, but that the relationship with informal training seems to be stronger. This may mean that farmers are in fact subject to peer influence when deciding how often to harvest, an important piece of knowledge when trying to influence farmers to harvest their cacao more intensely.

Relationships between other SEDG variables and technology adoption.

For the sake of concision, we will limit our discussion of non-policy SEDG variables to those that have significant results. Contrary to our original hypothesis, men are 0.1 times as likely to prune, 0.1 times as likely to graft and less likely to harvest intensely than are women. Reasoning for this gender difference is not immediately apparent. We also see that an increase in education is correlated with an increase in pesticide use but a decrease in harvest intensity. It is possible that education facilitates pesticide use through an ability to read cacao production manuals or grower pamphlets that suggest the use of pesticides, and that the more educated a person is the more likely they are to have alternative forms of employment, effectively reducing the time they can spend on the farm and the intensity with which they harvest. Furthermore, we find evidence that wealth, as measured by the area of land planted to any crop, is positively correlated with grafting and harvest intensity. This may be because people that are wealthier in terms of land, have greater cash flows from other activities on that land that allow them to pay the increased expenses of grafting and harvesting more frequently. Finally, we note that farmers that control for pests manually are 0.1 times as likely to use pesticides as those that do not. This fact is intuitive and validates the inclusion of this variable in the model.

As we will see with the yield model as well, the region in which a farm is located is significantly correlated to the adoption of certain technologies. The adoption of herbicide use, pruning, grafting and changes in harvest intensity vary to some degree by region. Two interesting observations present themselves as we further scrutinize these coefficients: first, no one region seems to experience the highest adoption across all of these technologies, and second, the adoption of some technologies – e.g. grafting and pruning – show larger variances by region than the adoption of other technologies – e.g. harvest intensity and herbicide application. When we consider this analysis in the context of the other control variables included in the model as outlined previously, we note that something other than the factors we have controlled for is varying by region in a way that affects the adoption of certain technologies. A possible explanation, supported by our experience in the field, is that the "cacao culture" of some areas is more developed than others. Influential farmers from some regions tell local success stories of cacao, thus influencing people's willingness to invest in the development of cacao. Influential farmers from other regions, on the other hand, have a more negative perspective of the potential of cacao, which stems from a few instances of crop failure or failure of the crop to deliver returns promised by government and non-government organizations.

Relationships between SEDG and policy variables, and general propensity to adopt.

The technology adoption index model offers us a higher-level look at what is influencing a farmer's propensity to adopt certain production technologies. A lack of significance in this model, however, limits our discussion to the two following points. First, women are 1.3 times more likely to adopt a given technology than men. As we discussed before, and as evidenced by our observations from field research, this may be due to increased sociality amongst women. And second, the number of buyers that a farmer sells to is positively correlated with their adoption of production technologies. A possible explanation for this may also be related to an increased sociality and therefore exposure to new ideas inherent in interacting with more industry players.

In summary of the findings from our technology adoption models, association membership and an increase in the number of buyers sold to are positively correlated with the adoption of several technologies, including harvest intensity which is positively correlated with yields. Receipt of financing is also positively correlated with the adoption of some technologies, though it is negatively correlated with harvest intensity. And finally, training is strongly and positively correlated with the adoption of technologies: with informal training being more highly correlated with fertilizer use and harvest intensity, and formal training being more highly correlated with herbicide use, pruning and grafting. Other empirical results are less definitive, and therefore we do not discuss them here.

SEM and Direct vs. Indirect Relationships

As described in the Methodology section of this thesis, we use SEM to estimate potential mediation effects between SEDG factors, adoption of production technologies and yield per hectare. Results of the SEM are divided into two tables for ease of review, with Table 4 displaying the coefficient estimates for the yield or independent variable portion of the model and Table 5 displaying coefficient estimates for the production practice or mediation variable portion of the model. Although results are reported in two separate tables, coefficients were estimated simultaneously according to the path diagram laid out in Figure 5. Notice that since SEM does not modify coefficient estimates, the coefficients for the yield model in Table 2 and the coefficients for the independent variable portion of the SEM in Table 4 are the same, but that the p-values are smaller in the SEM model due to the modified covariance matrix structure employed by SEM (Hox & Bechger, 1998). You will also notice different coefficients for the mediation variable models in Table 3 as compared to their SEM counterparts in Table 5. This is due to the need to reduce the SEM model size by excluding some indirect relationship paths as described in the Methodology section of this thesis.

Direct relationships between production practices yield, and SEDG variables and yield, are identified in Table 4 (also Table 2). Indirect relationships between key SEDG variables, production practices, and yield are presented in Table 5. In order to identify the direct relationships between production practices and yield, and SEDG variables and yield, we will now dive deeper into an analysis of the yield model results. After this, we will combine the analysis of direct effects with that of indirect effects from Table 5 to calculate mediation effects. These relationships are illustrated numerically in Table 6 and graphically in Figure 4.

Table 4: SEM Results – Independent variable or yield model portion

Independent Variable	Coefficient
Fertilized	0.374*** (0.00656)
Applied Herbicides	-0.378* (0.0799)
Applied Pesticides	0.268 (0.163)
Pruned	0.114 (0.492)
Grafted	-0.0325 (0.846)
Harvest Intensity	0.160*** (4.10e-10)
Age	-0.00132 (0.827)
Gender	0.504*** (0.00459)
Years of Education	0.0207 (0.232)
Years of Cacao Experience	0.0595*** (1.01e-06)
Association Membership	-0.254* (0.0852)
Number of Buyers	0.554*** (2.27e-06)
Ha of Crops Planted	0.00258 (0.832)
Ha of Cacao Planted	-0.234*** (0.000200)
North Region	2.109*** (9.88e-09)
Central Region	1.477*** (5.04e-06)
South Region	2.000*** (1.16e-09)
Constant	0.620 (0.241)
Observations	277

Table 5: SEM Results – Mediation variables or production practices portion

Independent Variables	Fertilized	Applied Herbicides	Applied Pesticides	Pruned	Grafted	Harvest Intensity
Gender	-0.0248	0.00407	-0.623	-2.229**	-2.189**	-0.630
	(0.953)	(0.995)	(0.229)	(0.0351)	(0.0409)	(0.318)
Association Membership	0.223	-0.281	0.695	0.990**	0.285	0.772*
	(0.445)	(0.503)	(0.148)	(0.0247)	(0.546)	(0.0613)
Number of Buyers	0.245	0.131	0.219	0.585**	0.195	1.440***
	(0.237)	(0.582)	(0.495)	(0.0126)	(0.527)	(0.000271)
Ha of Crops Planted	0.0126	-0.0114	-0.0209	0.0908*	0.158*	0.0247
	(0.573)	(0.733)	(0.534)	(0.0514)	(0.0502)	(0.334)
Received Cacao Financing	0.287	0.0389	1.341***	0.179	-0.225	-0.831***
	(0.457)	(0.947)	(0.00457)	(0.767)	(0.728)	(0.00222)
North Region	0.00221	-16.04***	-0.824	0.607	1.648**	-5.119***
	(0.997)	(0)	(0.552)	(0.359)	(0.0311)	(9.18e-08)
Central Region	-0.0454	-1.471**	0.314	0.913	2.882***	-4.457***
	(0.922)	(0.0149)	(0.722)	(0.130)	(1.01e-05)	(4.85e-06)
South Region	-0.156	0.0665	-0.159	1.947***	3.678***	-4.932***
	(0.734)	(0.893)	(0.857)	(0.000762)	(1.95e-06)	(1.51e-07)
Formal Skill Training	-	1.125*	0.309	1.503***	2.041***	0.813
		(0.0840)	(0.554)	(0.000161)	(4.17e-05)	(0.165)
Informal Skill Training	1.141**	-	-	0.536	-	1.719
	(0.0388)			(0.331)		(0.208)
Manually Controlled Pests	-		-1.856***	-	-	-
			(0.00303)			
Constant	-0.721	-0.919	-1.760	-0.0717	-0.575	7.096***
	(0.228)	(0.246)	(0.101)	(0.952)	(0.672)	(1.57e-09)
Observations	277	277	277	277	277	277

Robust P-values in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Direct relationships between production practices and yield per hectare.

As mentioned previously, the three production practices with significant, direct relationships to yield are harvest intensity, fertilizer use and herbicide use. Harvest intensity has a positive coefficient that suggests a 16% increase in yield per additional month harvested. A positive relationship here is highly logical, but the magnitude of the relationship and the high significance suggest that farmers may be leaving perfectly good fruit on the tree by not harvesting during all possible months. This explanation is corroborated by our experience while collecting the data on the farms, as well as by the feelings of Pedro Lora, the president of the largest grower association in MdM, who said, “these [farmers] are not ‘growers’, they are ‘harvesters’ who go out from time to time to pick what the forest has to offer.” Indeed, the average farmer in MdM only harvested during 4.6 months in 2017, whereas the literature suggests a harvest season of 6 to 10 months for countries similar to Colombia (Gayi & Tsowou, 2016).

Another important finding in this model is that farmers who use fertilizer experience yields that are 37% higher than those that do not. The direction is obvious, but the magnitude of the relationship suggests that this factor merits considerable thought while developing intervention strategies in this region, especially because only 44% of farmers in MdM are using fertilizers currently. Herbicide use by farmers, on the other hand, is associated with a 38% reduction in yields. Our experience in the field indicates this could be due to a lack of understanding of how and when to apply herbicides, or more specifically that farmers are overspraying herbicides and affecting the health of their trees. This possibility is another that ought to be carefully considered in any attempt by to boost yields.

Effective intervention efforts should consider the whole gamut of cacao production practices, but our results suggest that special attention ought to be given to harvest intensity, fertilizer use and herbicide use. These results generally conform to our hypotheses of the magnitudes of relationships in Table 1, with two important exceptions. First, we did not expect herbicide use to have a negative relationship with yields, and second, we would have expected a significant and positive relationship between pruning and yields. This may be due to the perennial nature of cacao plants, which causes a lag in the yield response of certain production practices, e.g. farmers typically only see healthier trees due to the pruning of the year before. Each of these results may be affected by adoption rates (see Table 1), a possibility that is explored in the Technology Adoption Model Results section of this thesis.

Direct relationships between policy variables and yield per hectare

Two of the policy variables examined in this model have significant relationships with yields. First, farmers that are selling to more buyers are yielding more, and each additional buyer that a farmer sells to is linked to a 55% increase in yields, compounded over the total number of buyers. It is important to note that the average farmer only sold to 1.13 buyers and the most buyers any farmer sold to was three, so the compounding effects of this result are limited by the range of the data. It is also important to note that not all farmers have access to various buyers. Some are so remote that only one buyer will go to their farm, or they can only afford to take their beans to the closest buyer, while others have several buyers within a reasonable range of their farm. This coefficient does not suggest causality, only association, and that causality will be examined further in the Structural Equation Model and Mediation Effects section of this thesis.

Second, we find that growers that belong to a grower association are yielding 25% less on average than growers that do not. One interpretation of this result that is of concern is that a production practice is being poorly taught or promoted in the grower associations, which is reducing yields. There are, however, two alternate interpretations of this result. First, it is possible that better growers are opting out of grower associations if they believe they already have the skills and market connections that many associations promise in order to attract more members. Second, it is possible that programs focused on the formation of associations are targeting growers with greater production challenges or those with a greater lack of production knowledge. Our field experience suggests that more skilled growers are most likely opting out of grower associations. As with the previous case, causality will be examined further in the Structural Equation Model and Mediation Effects section of this thesis.

Direct relationships between other SEDG factors and yield per hectare

We also see several other factors that are significantly and directly correlated with yield, that are not necessarily policy variables, but that are easy to build a strategy around or act upon. First, male farm owners are producing 50% more per hectare on average than female farm owners. As described previously, this is likely due to the fact men are more directly involved in the operations on their farms, whereas women farm owners tend to contract out labor to hired hands, effectively reducing the available labor input and the amount of cash available to purchase inputs and pay for improvements on the farm. Second, we observe that each additional year of experience farming cacao is correlated with a yield increase of 6% compounded by each additional year, which supports our initial expectation. Lastly, each additional hectare planted to cacao is correlated with a 23% decrease in yield per hectare. This result again is limited by an average cacao plantation of 1.56 hectares with few farmers having more than 3 hectares. This is contrary to the principle of economies of scale but supports a body of literature that suggests that smallholder farmers can realize higher yields per acre than can larger farms given the same set of physical inputs (Barrett, Bellemare, & Hou, 2010).

It is also important to note the statistically significant differences in yield by region. After controlling for the variables described above, the north, central and south regions all experienced greater yields than the omitted Southwest region. More specifically, farmers in the northern region experienced the highest yields, followed closely by those in the southern region, then by those in the central region. Several factors that may be expected to vary geographically – such as years of cacao experience, association membership, hectares of cacao planted and training – were controlled for in this model, suggesting that the remaining relationship we observe between yield and region is likely due to differences in climatic and topographic factors such as elevation, precipitation and soils.

Those SEDG factors in our model that are not significantly correlated with yield are the age of the farmer, their educational attainment, and their wealth as measured by the area of land they have planted to crops. These variables will be considered further in subsequent sections of this thesis that review technology adoption and mediation effects.

In summary, the yield model suggests that harvest intensity and fertilizer use are largely and positively correlated with yield, whereas herbicide use has an equally large but negative association with yield. In the way of policy variables, growers that are selling to more buyers are yielding more, while growers that belong to grower associations are yielding less. With respect to non-policy SEDG variables, being male and having more years of experience growing cacao are correlated with higher yields, whereas having a larger cacao plantation is correlated with lower yields.

Mediation effects

In order to effectively establish mediation effects, our SEM must show statistical significance between the independent variable (policy and other SEDG variables), mediation variable (production practice) and dependent variable (yield). In our model two relationships meet these criteria, and both of them have harvest intensity for their mediation variable. As Table 6 and Figure 4 lay out, the intensity with which a farmer harvests their cacao crop mediates the relationships between association membership and yields, and the relationship between number of buyers and yields. Understanding the implications behind these numbers is important.

First, the direct effect of association membership on yield is negative. However, as we inspect the indirect effect of association membership on yield as mediated through harvest intensity, we find that the direction is flipped and that the magnitude of the indirect effect is roughly half that of the direct effect. In other words, associations appear to be effective in increasing the intensity with which their farmers harvest, but that some other factor is having an even larger but negative effect on their yields. The negative direct effect could be due to self-selection of farmers into these associations – i.e. farmers that need more production help seek out grower associations, whereas those that are more technically advanced in cacao production avoid associations because of their added fees. Second, the number of buyers that a farmer sells to has a positive total effect on yields. This total effect can be decomposed to show that 29% of the total effect of the number of buyers on yields is indirectly mediated through an increase in harvest intensity that is promoted by that increase in buyers. It seems that the more buyers that a farmer has to sell to, the more they are willing to put in the effort of harvesting their cacao, which seems to be going to waste on the trees otherwise.

Table 6: Statistical summary of individual indirect and direct effects of significant factors.

Independent Variable	Mediation Variable	Dependent Variable	Indirect Effect Coefficient (P-value)	Direct Effect Coefficient (P-value)	% of Total Effect that is Indirect
Association Membership	Harvest Intensity	Yield	0.124 (0.071)	-0.25 (0.085)	N/A$^{\Psi}$
Number of Buyers	Harvest Intensity	Yield	0.231 (0.002)	0.55 (0.000)	29%

Ψ: The N/A here represents the fact that the indirect effect is positive and the direct effect is negative, making an interpretation on a percentage basis illogical.

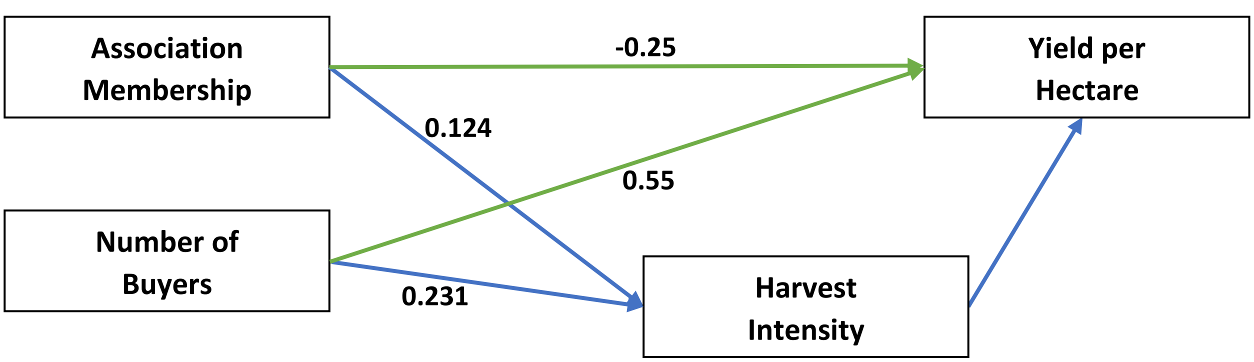


Figure 4: Visual summary of individual indirect and direct effects of significant factors.

Conclusion

In this study we lay out a method that produces results that can be used by grower associations, cooperative leaders, government and non-government association practitioners, and academics that are interested in increasing yields amongst smallholder farmers of any crop. By first examining which production practices are most strongly related to changes in yield for a specific region, and then determining what is affecting the adoption of those technologies and testing for mediation effects, one can build a strong intervention plan that addresses training, supply chains, finance and social needs. Indeed, during this process we were able to identify certain crucial facts that we brought to the attention of the association leaders and government field technicians that service the farmers we worked with.

We find that harvest intensity, fertilizer use and herbicide use have the most significant relationships with yield amongst cacao producers in the Montes de Maria region of Colombia. When examining policy variables that have the strongest relationships with technology adoption of these most significant practices, our results suggest that being a member of an association, selling cacao to a larger number of buyers, and receiving training are all associated with an increase in harvest intensity by a farmer. Our results indicate that at least some portion of the relationships between yield and association membership, and yield and the number of buyers, appears to be causal and mediated through an increase in harvest intensity. Furthermore, training is highly correlated with increases in adoption of several practices and technologies, and that formal training seems to be more strongly related to adoption of pruning, grafting, herbicide use and pesticide use, while informal training is more strongly related to the adoption of fertilizer use and an increase in harvest intensity. Few of the SEDG factors observed had any statistically significant effect on the adoption of fertilizer use and herbicide use besides training.

Overall, this analysis suggests a strategic development plan that could be used by association leaders and government and non-government organizations in MdM that would include several key elements. One very literal "low hanging fruit" in MdM in terms of boosting cacao yields is increasing the harvest intensity amongst farmers. Two effective ways to increase harvest intensity would be first, promoting grower association membership, and second, strengthening supply chains in a way that more growers are exposed to more potential buyers. Other development efforts ought to include increasing trainings focused on harvesting, fertilizer use and herbicide use, all of which seem to be key in boosting yields amongst these particular farmers. To encourage fertilizer use, informal trainings from friends and neighbors seem to be most effective; these trainings could be promoted through association meetings and word-of-mouth campaigns. On the other hand, association leaders and local organizations could promote increased yields by developing formal training regimes focused on proper harvest techniques and intensity as well as the proper use of herbicides in cacao production. The focus on *proper* herbicide use is important for this region as it is currently negatively affecting cacao yields, though when applied correctly, agronomic research suggests that proper herbicide use should increase yields.

Again, we emphasize that the suggestions in the last paragraph must be part of a larger plan. This knowledge is useful, but could be detrimental when considered in isolation. Those who are developing strategies for development must couple the analysis presented here with costs of implementation, knowledge of local and national legal structures, and more in-depth understanding than is presented in this thesis of the effect that the armed violence in Colombia had on the citizens of MdM. They must also keep in mind that though fertilizer use, herbicide use and harvest intensity are the only production practices that show a significant effect on yields in this study, agronomic research has demonstrated the value of other practices such as pruning and irrigation in cacao production, and these must not be ignored.

We recognize many areas that could be improved or strengthened in future research. Researchers would benefit from, where possible, increasing the precision of their data through direct observation or experimentation. Self-reporting by farmers on key variables such as yields likely decreased our estimation strength. Metrics of intensity of adoption of key practices, as opposed to binary metrics of adoption, would provide deeper insights into the questions presented here. Again, here we faced the challenge of self-reporting by smallholder farmers. In this study, our ability to thoroughly examine mediation effects through SEM was limited by a small sample size, and thus for future studies, we recommend a larger sample size and potentially a larger geographic survey area. Finally, deeper scientific and agronomic studies into the specific production functions of cacao, which are relatively sparse compared to some other crops, would greatly aid in future studies of this kind.

References

Almeida, R. L., Chaves, L. H., Bonomo, P., Almeida, R. L., & Fernandes, J. D. (2014). Cacao Production Submitted to Different Water Slides and Nitrogen Doses. *Caatinga*, 171-182.

Aneani, F., Anchirinah, V. M., Owusu-Ansah, F., & Asamoah, M. (2012). Adoption of Some Cocoa Production Technologies by Cocoa Farmers in Ghana. *Sustainable Agriculture Research*, 103-117.

Baffoe-Asare, R., Danquah, J. A., & Annor-Frempong, F. (2013). Socioeconomic Factors Influencing Adoption of Copadec and Cocoa High-tech Technologies among Small Holder Farmers in Central Region of Ghana. *American Journal of Experimental Agriculture*, 277-292.

Barrett, C. B., Bellemare, M. F., & Hou, J. Y. (2010). Reconsidering Conventional Explanations of the Inverse Productivity-Size Relationship. *World Development*, 88-97.

Barrett, C. B., Christiaensen, L., Sheahan, M., & Shimeles, A. (2017). On Strucutural Transformation of Rural Africa. *Journal of African Economies*, i11-i35.

Binham, J. N., Gockowski, J., & Nkamleu, G. B. (2008). Technical Efficiency and Productivity Potential of Cocoa Farmers in West African Countries. *The Developing Economies*, 242-263.

Byerlee, D., & Rueda, X. (2015). From Public to Private Standards for Tropical Commodities: A Century of Global Discourse on Land Governance on the Forest Frontier. *Forests*, 1301-1324.

Collier, P., & Dercon, S. (2014). African Agriculture in 50 Years: Smallhoders in a Rapidly Changing World? *World Development*, 92-101.

Conservation Alliance International. (2013). *Sustainable Cocoa Production.* Accra: Conservation Alliance International.

Feder, G., & Slade, R. (1984). The Acquisition of Information and the Adoption of New Technology. *American Journal of Agricultural Economics*, 312-320.

Feder, G., Just, R. E., & Zilberman, D. (1985). Adoption of Agricultural Innovations in Developing Countires: A Survey. *Economic Development and Cultural Change*, 255-298.

Food and Agriculture Organization of the United Nations. (2017). *The State of Food and Agriculture.* Rome: FAO.

Funcicar. (2015). *Sistematizacion de la Experiencia de Colombia Responde de la Zona de Consolidacion de los Montes de Maria.* Cartgena: Fundacion ProCartagena.

Gayi, S. K., & Tsowou, K. (2016). *Cocoa Industry: Integrating Small Farmers into the Global Value Chain.* New York and Geneva: United Nations.

Gomez, M. I., & Ricketts, K. D. (2013). Food Value Chain Transformations in Developing Countries; Selected hypotheses on Nutritional Implications. *Food Policy*, 139-150.

Griffin, R. C., Montgomery, J. M., & Rister, M. E. (1987). Selecting Funcitonal Form in Production Function Analysis. *Western Journal of Agricultural Economics*, 216-227.

Griliches, Z. (1963). Estimates of Aggregate Agricultural Production Function from Cross-Sectional Data. *Journal of Farm Economics*, 419-428.

Hayes, A. F. (2013). *Introduciton to Mediation, Moderation, and Conditional Process Analysis: A Regression Based Approach.* New york: The Guilford Press.

Haynes, J., Cubbage, F., Mercer, E., & Sills, E. (2012). The Search for Value and Meaning in the Cocoal Supply Chain in Costa Rica. *Sustainability*, 1466-1487.

Hes, T., Sulaiman, H., Mintah, S., Arrieta, J. S., Esquivel Bc, B. J., Saladana, T. M., & Hernandez, J. A. (2017). Weak Levels of Social Capital as One of the Causes of the Fall of Mexican Cacao Production. *Emirates Journal of Food and Agriculture*, 411-419.

Hox, J., & Bechger, T. (1998). An Introduction to Structural Equation Modeling. *Family Science Review*, 354-373.

International Cocoa Organization. (2014). *The Cocoa Market Situation.* London: International Cocoa Organization.

International Cocoa Organization. (2018). *Grindings of Cocoa Beans.* Abidjan: ICCO Quarterly Bulleting of Statistics.

International Cocoa Organization. (2018). *Production of Cocoa Beans.* Abidjan: ICCO Quarterly Bulletin of Cocoa Statistics.

International Cocoa Organization. (2018, October 24). *Trading and Shipping*. Retrieved from International Cocoa Association: https://www.icco.org/about-cocoa/trading-a-shipping.html

International Cocoa Organization. (2018). *World Cocoa Bean Produciton, Grindings and Stocks.* Abidjan: ICCO Quarterly Bulletin of Cocoa Statistics.

Karli, B., & Gul, M. (2015). Economic Analysis of Cocoa Production in Ghana: The case of eastern region. *Custos e Agronegocio*, 336-352.

Knowler, D., & Bradshaw, B. (2007). Farmers' Adoption of Conservation Agriculture: A review and synthesis of recent research . *Food Policy*, 25-48.

Kosarek, J. L., Garcia, P., & Morris, M. L. (2001). Factors Explaining the Diffusion of Hybrid Maize in Latin America and the Caribbean Region. *Agricultural Economics*, 267-280.

Lee, D. R. (2005). Agricultural Sustainability and Technology Adoption: Issues and Policies for Developing Countries. *American Journal of Agricultural Economics*, 1325-1334.

Lima, L. J., Almeida, M. H., Nout, M. J., & Zwietering, M. H. (2011). Theobroma Cacao L., "The Food of the Gods": QUality Determinants of Commercial Cocoa Beans, with Particular Reference to the Impact of Fermentation. *Critical Reviews in Food Science and Nutrition*, 731-761.

Muhardi, & Effendy. (2017). Cocoa Farming Patters for Sustainability of Indonesia Lore Lindu National Park (LLNP). *Australian Journal of Crop Science*, 917-924.

Murphy, S. (2012). *Changing Perspectives: Small-Scale Farmers, Markets and Globalization.* London/The Hague: IIED/Hivos.

Navarrete, C. J., & Rahman, S. (2014). Productivity Managemnet Analysis of Cacao Agro-Food System in Tabasco, Mexico: An Application of the 'Fitness' Approach. *International Journal of Food and Agricultural Economics*, 49-64.

Ramirez, O. A., Somarriba, E., Ludewigs, T., & Ferreira, P. (2001). Financial Returns, Stability, and Risk of Cocoa-plantain-timber Agroforestry Systems in Central America. *Agroforestry Systems*, 141-154.

Rosenberg, D. E., & Marcotte, T. P. (2005). Land-use System Modeling and Analysis of Shaded Cacao Production in Belize. *Agroforestry Systems*, 117-129.

Rueda, X., Paz, A., Gibbs-Plessl, T., Leon, R., Moyano, B., & Lambin, E. F. (2018). Smallholders at a Crossroad: Intensify or Fall Behind? Exploring Alternative Livelihood Strategies in a Globalized World. *Business, Strategy and the Environment*, 215-229.

Sheahan, M., & Barret, C. B. (2014). *Understanding the Agricultural Input Landscape in Sub-Saharan Africa.* World Bank Groun, Office of the Chief Economist.

Taher, S. (1996). *Factors Influencing Smallholder Cocoa Production.* Wageningen.

UCLA IDRE. (2018, 11 8). *How Can I Do Meidation Analysis with a Categroical IV in STATA?* Retrieved from UCLA Institute for Digital Research and Educaiton: https://stats.idre.ucla.edu/stata/faq/how-can-i-do-mediation-analysis-with-a-categorical-iv-in-stata/.

United Nations Development Programme, ISA Interconexion Electrica S.A. E.S.P, Compania Nacional de Chocolates. (2018). *Guia para el Cultivo de Cacao.* Retrieved from ISSU: https://issuu.com/pnudcol/docs/cartilla-final.

USAID Colombia. (2017). *Analysis of the Supply Chain of Cacao in Colombia.* USAID.

Wageningen, S. T. (1996). *Factors Influencing Smallholder Cocoa Production.* The Netherlands.

Weather Spark. (2018, October). *Weather Spark*. Retrieved from https://weatherspark.com/y/22602/Average-Weather-in-El-Carmen-de-Bol%C3%ADvar-Colombia-Year-Round.

Young, A. M. (2007). *The Chocoalate Tree: A Natural History of Cacao.* Gainesville: University Press of Florida.

Appendix

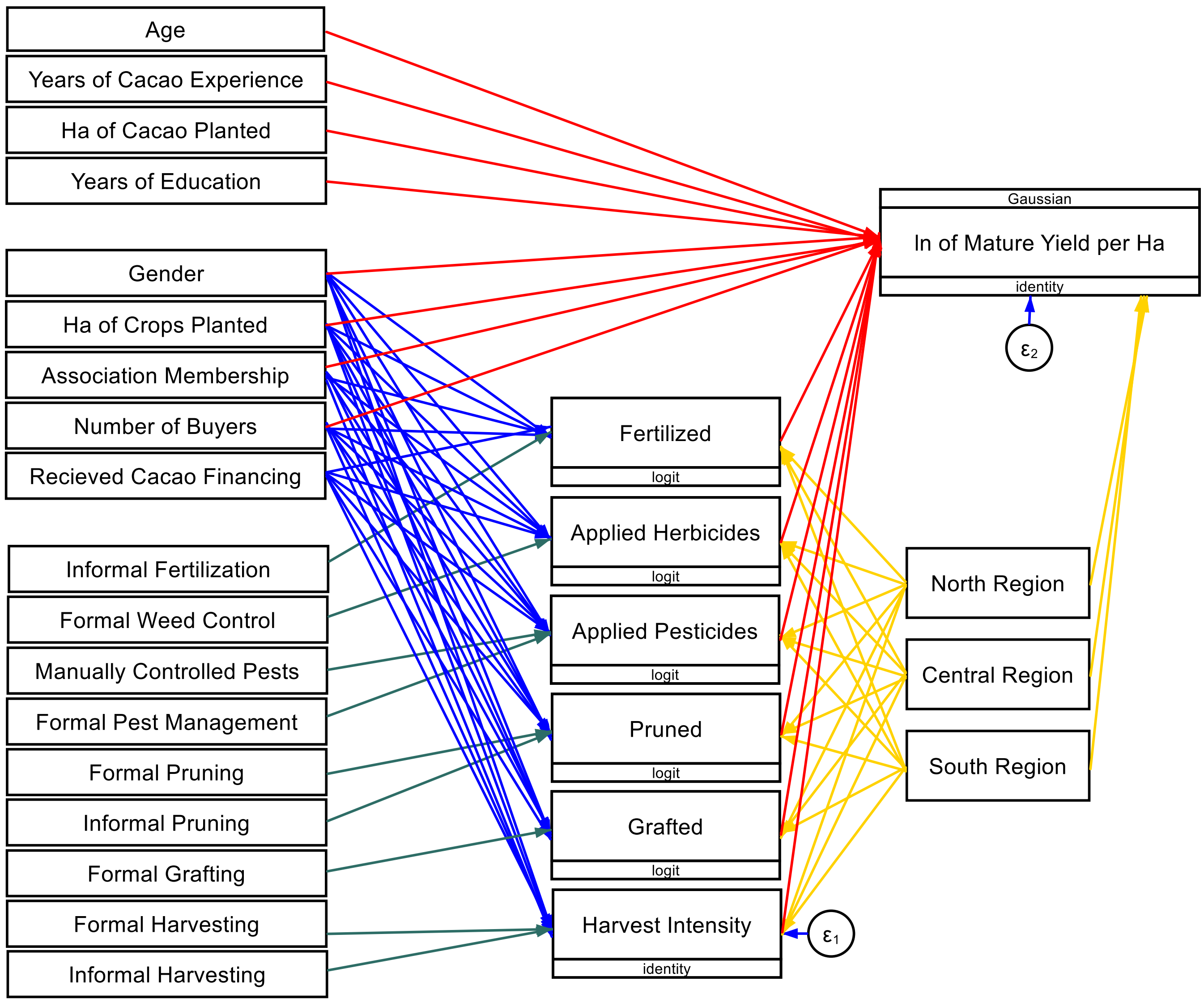


Figure 5: Structural equation model path diagram.

Table 7: Primary survey tool in Spanish and English (logic omitted).

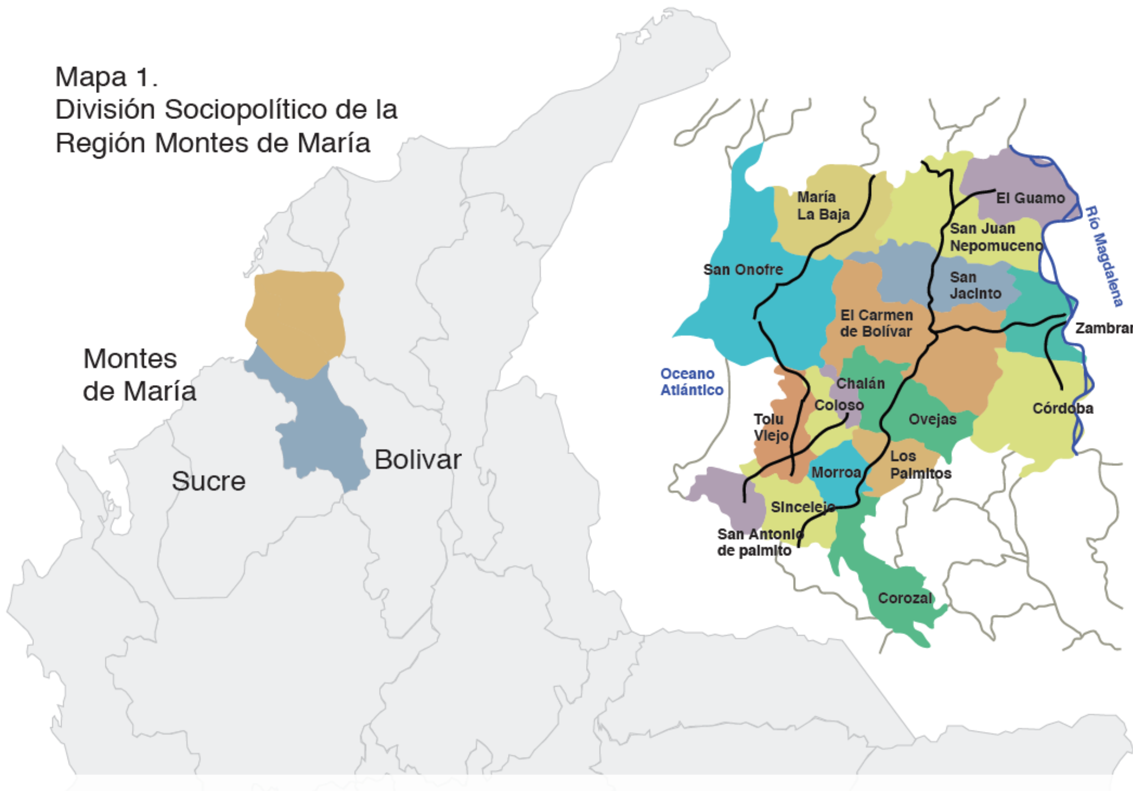
Q#	Question in Spanish	Question in English
1	Nombre del encuestador	Enumerators name
2	Como se llama?	What is your name?
3	Cuantos años tiene?	How old are you?
4	Género	Gender
5	Cuantos años de educación ha completado?	How many years of school have you completed?
6	Cuantas personas dependen de usted para su apoyo principal?	How many people depend on you for their primary support?
7	Cual porcentaje de la comida que comen viene de su finca?	What percentage of the food that you eat comes from your farm?
8	En cual municipio está ubicada su finca?	In which municipality is your farm located?
9	En cual corregimiento está ubicada su finca?	In which corregimiento is your farm located?
10	En cual vereda está ubicada su finca?	In which vereda is your farm located?
11	Su casa tiene electricidad?	Does your house have electricity?
12	El piso de su casa esta hecho de que?	What is the floor of your house made of?
13	Posee una bicicleta?	Do you own a bicycle?
14	Posee un vehículo?	Do you own a vehicle?
15	Posee una motocicleta?	Do you own a motorcycle?
16	Por cuantos años ha trabajado en agricultura?	For how many years have you worked in agriculture?
17	Por cuantos años ha cultivado cacao?	For how many years have you grown cacao?
18	De cuales asociaciones o grupos de productores es miembro?	Which associations or producer groups are you a member of?
19	Cuales certificaciones tiene para su cacao?	Which certifications do you have for your cacao?
20	Cuantas veces ha recibido asistencia técnica de una **organización o asociación** para su cultivo de cacao?	How many times have you received technical assistance from an organization or association for cacao cultivation?
21	Cuantas veces ha provisto asistencia técnica para el cacao **a otros productores** en su área?	How many times have you provided technical assistance to other producers in your area?
22	Cuantas veces ha recibido asistencia técnica para el cacao **de otros productores** en su área?	How many times have you received technical assistance from other producers in your area?
23	Con cuales prácticas de cultivo o procesamiento de cacao ha recibido capacitación formal? (de una asociación o organización)	For which production or processing practices have you received technical training? (from an association or organization)
24	Con cuales prácticas de cultivo o procesamiento de cacao ha recibido capacitación informal?	For which production or processing practices have you received informal training?
25	Cuales son los nombres de las organizaciones en su área que proporcionan asistencia técnica para el cacao?	What are the names of the organizations in your area that provide technical assistance for cacao?
26	Por cuantos años ha cultivado la tierra de su finca actual?	How long in years have you cultivated your current farm?
27	El numero de hectáreas en su finca ha aumentado, disminuido, o permanecido igual en los últimos 5 años?	In the last five years, has the number of hectares on your farm increased, decreased or stayed the same?
28	El numero de hectáreas de cacao en su finca ha aumentado, disminuido, o permanecido igual en los últimos 5 años?	In the last five years, has the number of hectares of cacao on your farm increased, decreased or stayed the same?

Q#	Question in Spanish	Question in English
29	Planea aumentar, disminuir, o mantener igual el número de hectáreas de cacao plantados en su finca en los próximos 5 años?	Do you plan to increase, decrease or maintain the number of hectares of cacao on your farm over the next 5 years?
30	Al nombre de quien está esta tierra?	In whose name is this land?
31	Por cuantos años ha poseído la tierra?	How long in years have you owned this land?
32	Tiene documentos para la tierra?	Do you have documents for the land?
33	Como mantiene el derecho de cultivar la tierra?	How do you maintain the right to cultivate the land?
34	Tiene miedo de perder la tierra que cultiva actualmente?	Are you afraid of losing the land that you currently cultivate?
35	Ha sido victima del conflicto armado?	Have you been a victim of the armed conflict?
36	En los últimos 5 años, ha recibido alguna asistencia financiera por fines agrícolas?	In the last five years have you received any financial assistance intended for agricultural purposes?
37	En que forma vino esta asistencia?	In what form did this assistance come?
38	Para que usó el dinero?	What did you use this money for?
39	Usó una porción del dinero para la producción de cacao?	Did you use a portion of this money for cacao production?
40	Ha buscado asistencia financiera pero no lo ha conseguido?	Have you sought financial assistance but been unable to receive it?
41	Que tan dispuesto está a tomar riesgos?	How willing are you to take risks?
42	Como se llama su finca?	What is your farm's name?
43	Cuantas hectareas tenia en su finca en 2017?	How many hectares did you have on your farm in 2017?
44	Cuantas hectáreas estaban plantadas con algún cultivo en 2017?	How many hectares were planted to some crop in 2017?
45	Cuantas hectáreas de cacao estaban plantadas en su finca en 2017?	How many hectares were planted to cacao in 2017?
46	Cuantos años y que tamaño tenia cada bloque de cacao en 2017?	How old and large was each block of cacao in 2017?
47	Conoce los nombres de las variedades de cacao plantadas en su finca?	Do you know the names of the varieties of cacao planted on your farm?
48	Cuantas hectáreas de cuales variedades de cacao estaban plantadas en su finca en 2017?	How many hectares of each variety were planted on your farm in 2017?
49	Cuales fuentes de agua están presentes en su finca?	Which sources of water were present on your farm in 2017?
50	Irrigó su cultivo de cacao en 2017?	Did you irrigate your cacao crop in 2017?
51	Como irrigó su cultivo de cacao en 2017?	How did you irrigate your cacao crop in 2017?
52	En semanas, con que frecuencia irrigó su cultivo de cacao en 2017 ?	In weeks, how frequently did you irrigate your cacao crop in 2017?
53	En cuales meses irrigó su cacao en 2017?	In which months did you irrigate your cacao in 2017?
54	Por que no irrigó su cacao?	Why didn't you irrigate your cacao?
55	Utilizó abono en su cultivo de cacao en 2017?	Did you use fertilizer in your cacao in 2017?
56	Cuantos abonos diferentes utilizó en su cacao en 2017?	How many different types of fertilizer did you use in your cacao in 2017?
57	Cuanto de cada tipo de abono utilizó en su cacao en 2017?	How much of each type of fertilizer did you use in your cacao in 2017?
58	Por que no utilizó abono?	Why didn't you use fertilizer?

Q#	Question in Spanish	Question in English
59	En semanas, con qué frecuencia quitó las hierbas manualmente de su cultivo de cacao en 2017?	In weeks, how frequently did you manually remove weeds from your cacao crop in 2017?
60	Utilizó herbicidas en su cultivo de cacao en 2017?	Did you use herbicides in your cacao crop in 2017?
61	Cuantas herbicidas diferentes utilizó en su cacao en 2017?	How many different herbicides did you use in your cacao in 2017?
62	Cuanto de cada tipo de herbicida aplicó a su cacao en 2017?	How much of each type of herbicide did you apply to your cacao in 2017?
63	Por que no utilizó herbicida?	Why didn't you use herbicides?
64	Cuales plagas y pestes estaban presentes en su cultivo de cacao en 2017?	Which diseases and pests were present in your cacao crop in 2017?
65	Como controló plagas y pestes en 2017?	How did you control diseases and pests in 2017?
66	Cuantos tipos de pesticida y fungicida utilizó en 2017?	How many types of pesticides and fungicides did you use in 2017?
67	Cuanto de cada tipo de pesticida y fungicida utilizó en su cacao en 2017?	How much of each type of pesticide or fungicide did you use in your cacao in 2017?
68	Por que no utilizó fungicidas o pesticidas?	Why didn't you use fungicides or pesticides?
69	En meses, con que frecuencia poda su cultivo de cacao?	In months, how frequently do you prune your cacao crop?
70	Que porcentaje del árbol quita cuando esta podando?	What percentage of the tree do you remove when you are pruning?
71	Hace algo para mejorar polinización?	Do you do anything to improve pollination?
72	Sus arboles de cacao estan injertados? Quien los injertó?	Are your cacao trees grafted? Who grafted them?
73	En los últimos 5 años, ha realizado una prueba de su suelo en los bloques donde cultiva cacao?	In the last five years, have you had a soil test done on the soil in the blocks where you produce cacao?
74	En los últimos 2 años, ha realizado una prueba de su suelo en los bloques donde cultiva cacao?	In the last two years, have you had a soil test done on the soil in the blocks where you produce cacao?
75	La erosión del suelo es una problema en su finca?	Is soil erosion a problem on your farm?
76	Hace algo para prevenir la erosión del suelo? Que hace?	Do you do anything to prevent soil erosion? What do you do?
77	En 2017, cuales otros cultivos estaban plantados dentro de su cultivo de cacao?	In 2017, what other crops were planted in your cacao crop?
78	Cuales variedades de árbol estaban sombreando su cultivo de cacao en 2017?	Which tree varieties were shading your cacao crop in 2017?
79	En metros, que era la distancia entre sus arboles de cacao en 2017?	In meters, what was the distance between your cacao trees in 2017?
80	En 2017, cuales eran los 5 cultivos mas rentables a su finca?	In 2017, what were the five most profitable crops on your farm?
81	Tiene ganado?	Do you have cattle?
82	Tiene otros animales de coral?	Do you have other livestock?
83	De donde viene sus nuevos arbolitos de cacao?	Where do your cacao seedlings come from?
84	En la semana promedia de 2017, cuantas horas usted pasó personalmente trabajando en su cultivo de cacao?	In the average week of 2017, how many hours did you personally spend working in your cacao plantation?
85	En la semana promedia de 2017, cuantas horas miembros de su familia pasaron trabajando en su cultivo de cacao?	In the average week of 2017, how many hours did a member of your family spend working in your cacao plantation?

Q#	Question in Spanish	Question in English
86	En la semana promedia de 2017, cuantas horas otras personas pasaron trabajando en su cultivo de cacao?	In the average week of 2017, how many hours did other people spend working in your cacao plantation?
87	Que porcentaje del trabajo en el cultivo de cacao se debió al trabajo de fermentación y secado?	What percentage of the work in your cacao crop was due to fermentation and drying?
88	En días, con que frecuencia cosechó cacao en 2017?	In days, how frequently did you harvest your cacao in 2017?
89	En cuales meses cosechó cacao en 2017?	In which months did you harvest cacao in 2017?
90	En 2017, cual era la máxima cantidad de kilos cosechadas en una sola cosecha?	In 2017, what was the maximum quantity of kilos you harvested at one time?
91	En cual mes cosechó este maximo?	In which month did you harvest this maximum?
92	Cuantos kilos de cacao cosechó de su finca en el año de 2017?	How many kilos of cacao did you harvest from your farm in 2017?
93	Cuantos kilos de cacao de su finca consumió or regaló en 2017?	How many kilos of cacao from your farm did you consume or gift in 2017?
94	Usted personalmente fermentó su cacao en 2017?	Did you personally ferment your cacao in 2017?
95	Usted personalmente secó su cacao en 2017?	Did you personally dry your cacao in 2017?
96	Que hizo con su cacao después de fermentarlo?	What did you do with your cacao after fermenting it?
97	A cuantos compradores diferentes vendió su cacao en 2017?	How many different buyers did you sell your cacao to in 2017?
98	Cual tipo de comprador [nombre] es?	What type of buyer is [name]?
99	Que [nombre] hizo con el cacao después de comprar lo de usted?	What did [name] do with the cacao after purchasing it from you?
100	Vendió su caco a [nombre] en mazorca, en baba o seco en 2017?	Did you sell your cacao to [name] in the pod, wet, or dry in 2017?
101	Como su cacao llegó a [nombre]?	How did you cacao arrive to [name]?
102	Cuantos KG de cacao vendió a [nombre] en 2017?	How many kilos of cacao did you sell to [name] in 2017?
103	Que calidad [nombre]requirió de usted en 2017?	What quality of cacao did [name] require of you in 2017?
104	Recibió una prima por calidad de [nombre] en 2017?	Did you recieve a quality premium from [name] in 2017?
105	En COP/KG, que era el precio promedio que recibió de [nombre] por un kilo de cacao en 2017?	In Colombian Pesos per KG, what was the average price you received from [name] in 2017?
106	En kilos, cuanto cacao vendio a todos los otros compradores combinados en 2017?	In kilos, how much cacao did you sell to all other buyers combined in 2017?
107	En COP/Kilo, cuanto recibió como precio promedio de todos los otros compradores combinados en 2017?	In Colombian Pesos per KG, what price did you receive per kilo on average from all other buyers in 2017?
108	Cuantas otras personas o compañías ofrecieron comprar su cacao en 2017?	How many other people or companies offered to buy your cacao in 2017?
109	Cuales tipos de compradores son?	What types of buyers are they?
110	Por que no vendió a ellos?	Why didn't you sell to them?
111	Verificaba preció con varios contactos antes de vender cacao en 2017? Con cuantos contactos en total?	Did you verify prices with various contacts before selling your cacao in 2017? With how many contacts in total?
112	En kilómetros, que tan lejos es su finca del comprador mas cercano?	In kilometers, how far is your farm from the closest buyer?

Q#	***Question in Spanish***	***Question in English***
113	En kilómetros, que tan lejos es su finca del comprador segundo mas cercano?	In kilometers, how far is your farm from the second closest buyer?
114	En kilómetros, que tan lejos es su finca del vendedor de insumos agrícolas mas cercano?	In kilometers, how far is your farm from the closest supplier of agricultural inputs?
115	En kilómetros, que tan lejos es su finca del mercado mas cercano?	In kilometers, how far is your farm from the closest market?
116	Nos gustaría retornar los resultados de esta encuesta a ustedes. Si le interese recibir estos resultados, como seria la mejor manera de mandarlos a usted?	We would like to return the results of this survey to you. If you are interested in receiving the results, what would be the best way to send them to you?
117	Podemos sacar una foto de una mazorca madura de cada variedad de cacao?	Can we take a photo of a mature cacao pod from each variety on your farm?
118	Ya pidieron de usted una muestra de cacao en el centro de agregacion para su asociacion como parte de esta investigacion?	Has anyone in an aggregation center asked you for a sample of cacao as part of this study?
119	Usted tiene acutalmente en su finca cacao que ya fue fermentado y secado de cual podríamos comprar una pequeña muestra?	Do you currently have cacao that you fermented and dried from which we could purchase a small sample?
120	SOLAMENTE PARA EL ENCUESTADOR. Vio usted alguna evidencia de erosión del suelo?	ONLY FOR THE INVESTIGATOR. Did you see evidence of soil erosion?
121	SOLAMENTE PARA EL ENCUESTADOR. Con que estaba cubierto la mayoria del suelo?	ONLY FOR THE INVESTIGATOR. What was the majority of the soil covered with?



Source: (Funcicar, 2015)

Figure 6: Sociopolitical map of Northern Colombia highlighting the Montes de Maria Region.

Table 8: Correlation matrix of variables included in models

	Age	Gender	Education	Cacao Experience	Association Membership	Number of Buyers	Ha of Crops Planted	Ha of Cacao Planted	Harvest Intensity	Fertilized	Pruned	Grafted	Applied Pesticides	Applied Herbicides	North Region	Central Region	South Region	Fertilizer Formal	Fertilizer Informal	Weeds Formal	Weeds Informa	Pests Formal	Pests Informal	Pruning Formal	Pruning Informal	Grafting Formal	Grafting Informal	Harvesting Formal	Harvesting Informal
Age	1.00																												
Gender	0.03	1.00																											
Education	-0.40	-0.10	1.00																										
Cacao Experience	0.13	-0.09	0.01	1.00																									
Association Membership	0.03	0.00	-0.03	0.13	1.00																								
Number of Buyers	-0.06	0.00	0.01	0.13	-0.04	1.00																							
Ha of Crops Planted	-0.08	0.07	0.13	0.06	0.13	-0.03	1.00																						
Ha of Cacao Planted	0.02	0.00	-0.01	0.16	0.16	0.00	0.13	1.00																					
Harvest Intensity	0.02	0.01	-0.07	0.02	-0.05	0.27	0.01	0.06	1.00																				
Fertilized	0.00	0.00	-0.02	-0.04	0.07	0.06	0.03	0.06	0.13	1.00																			
Pruned	-0.04	-0.15	0.08	0.11	0.26	0.15	0.12	0.08	0.04	0.08	1.00																		
Grafted	-0.09	-0.12	0.04	0.03	0.20	0.10	0.14	0.12	-0.08	-0.04	0.33	1.00																	
Applied Pesticides	0.04	-0.09	0.06	-0.02	0.11	0.03	-0.03	0.16	0.07	0.20	0.16	0.03	1.00																
Applied Herbicides	-0.08	0.02	0.04	-0.01	-0.17	0.10	-0.05	-0.07	0.22	0.19	-0.09	-0.12	0.15	1.00															
North Region	0.14	-0.01	0.10	0.02	0.04	-0.13	0.01	-0.09	-0.14	0.01	-0.04	-0.09	-0.10	-0.17	1.00														
Central Region	-0.03	-0.03	-0.03	0.18	0.41	-0.07	0.09	0.32	-0.07	0.02	0.14	0.20	0.13	-0.22	-0.38	1.00													
South Region	-0.13	-0.04	-0.01	-0.12	-0.21	0.18	-0.05	-0.25	-0.11	-0.02	0.08	0.14	-0.02	0.26	-0.23	-0.59	1.00												
Fertilizer Formal	-0.01	0.06	0.06	0.04	0.01	0.06	0.03	0.06	0.02	0.05	0.06	-0.06	0.06	0.04	0.03	0.00	-0.09	1.00											
Fertilizer Informal	-0.03	-0.02	-0.04	-0.09	0.05	-0.03	-0.10	-0.04	0.05	0.12	0.05	-0.01	-0.02	0.02	0.05	-0.12	0.10	0.10	1.00										
Weeds Formal	-0.05	0.08	0.04	0.02	0.00	0.09	-0.02	-0.07	0.08	-0.01	0.04	0.02	0.09	0.13	0.01	-0.07	0.05	0.30	0.08	1.00									
Weeds Informa	0.00	0.05	-0.03	-0.02	-0.05	-0.07	-0.03	-0.07	-0.01	-0.12	-0.06	-0.08	-0.06	-0.06	0.03	-0.08	0.04	-0.06	0.20	0.09	1.00								
Pests Formal	-0.02	0.03	-0.04	-0.01	0.06	0.13	0.01	0.02	0.08	0.03	0.04	0.05	0.03	0.02	-0.11	0.10	-0.09	0.27	-0.03	0.12	-0.06	1.00							
Pests Informal	-0.01	-0.03	-0.07	0.04	0.08	-0.03	-0.02	0.06	0.10	0.02	0.02	0.07	0.01	0.00	-0.06	0.00	0.02	0.07	0.28	-0.03	0.17	0.07	1.00						
Pruning Formal	-0.05	0.09	0.05	0.05	0.22	0.05	0.03	0.16	0.10	0.05	0.24	0.18	0.12	-0.11	0.01	0.23	-0.26	0.23	0.03	0.14	-0.01	0.21	0.01	1.00					
Pruning Informal	-0.04	-0.10	0.13	-0.04	0.00	-0.03	0.05	-0.05	-0.02	0.09	0.02	-0.16	-0.05	0.03	0.04	0.06	-0.02	0.02	0.22	0.10	0.18	-0.09	0.23	-0.22	1.00				
Grafting Formal	-0.01	0.15	-0.01	0.06	0.16	0.08	-0.03	0.17	0.15	-0.05	0.09	0.20	0.01	-0.08	0.01	0.18	-0.26	0.16	0.06	0.10	-0.03	0.12	0.05	0.51	-0.17	1.00			
Grafting Informal	-0.04	-0.03	0.11	0.01	0.09	-0.02	-0.04	0.00	0.00	0.01	0.02	-0.01	-0.08	-0.05	0.05	0.00	0.01	-0.02	0.16	0.05	0.16	-0.06	0.23	-0.14	0.39	-0.03	1.00		
Harvesting Formal	0.04	0.09	0.04	0.07	-0.06	0.11	-0.02	0.06	0.20	0.02	0.02	0.10	0.01	0.12	-0.04	-0.09	-0.01	0.26	0.10	0.38	0.12	0.11	0.10	0.18	0.02	0.18	0.14	1.00	
Harvesting Informal	-0.07	0.04	0.06	-0.01	0.02	-0.02	-0.02	-0.02	0.10	-0.09	-0.03	0.05	-0.04	-0.05	-0.04	-0.04	0.02	-0.05	0.12	0.14	0.51	-0.05	0.22	0.08	0.16	0.11	0.36	0.18	1.00