Income Diversification and Welfare, a Way Out of Rural Poverty?

A Study on Three Commercialized Regions in Ethiopia, Ghana, and Malawi

Andreas Nord

andreasnord93@gmail.com

Abstract: This paper aims to extend the understanding of the connection between household income diversification and welfare in the rural areas of Sub-Saharan Africa. I perform a comparative analysis and investigate three commercialized rural regions in Ethiopia, Ghana, and Malawi, which are countries with diverging economic development over the last decades. By using survey data from the Demographic and Health Surveys, descriptive and econometric results are displayed and analyzed. The findings are that the correlation seems to be dependent on which of the subsamples it regards, but that in general, income diversification to the non-agricultural sector seems to have a positive correlation with wealth in comparison to only engage in agriculture.

Keywords: Income Diversification; Welfare; Rural; Commercialized; Sub-Saharan Africa

Table of Contents

List of Tables

List of Figures

1. Introduction

When we seek economic growth and development, we are essentially looking for a higher living standard. To strive towards a higher living standard happens across all societies irrespective of the given society's affluence in absolute terms. What economic development brings with it, therefore, differs between societies; in richer regions, additional economic growth could allow larger segments of the population access to goods or services that, while not essential to living, increase their enjoyment of life. In poorer regions such as Sub-Saharan Africa (SSA) on the other hand, economic development can entail drastic improvements to the standard of living and give access to necessities such as food, housing, and education. In more general terms, welfare. Especially now with the occurring Covid-19 crisis that may create both health, and an economic crisis reversing the development of SSA (IMF, 2020). Rural residents have higher poverty rates throughout SSA (Beegle et al., 2016), and they earn the majority of their income from agriculture (Davis et al., 2017) which is a sector with relatively low returns. Farmers frequently search for income in other sectors to make a sufficient living, yet often continue to suffer from poverty. *Can perhaps household income diversification, to combine work in the agricultural sector and non-agricultural sector be a way out of poverty, or is income diversification a prerequisite in the current circumstances to find a way out of poverty in rural SSA?* If the non-agricultural sector seems to show potential, and especially in combination with agriculture, formulating policies to assist these sectors should be of the highest concern. As the twin rural development goals state, the ambition is to alleviate poverty and increase agricultural production (UN, 2020). If household income diversification is a way to alleviate poverty, it can be a tool to reach these goals. Besides, previous findings state that the returns from non-agricultural work tend to be invested in agricultural production (Francis, 2000, pp 20; Haggblade, Hazell, and Reardon, 2007, pp 44, 126), an important finding on the path to reach the goals.

This thesis examines the relationship between household income diversification and increased welfare comparing those households who partake in both the agricultural and non-agricultural sectors with those who only engage in agriculture in rural areas of three regions in SSA. The regions have high levels of commercialization and are all based on smallholder farming, with substantially differing country characteristics. The three regions are Tigray in Ethiopia, the Western Region in Ghana, and the Central Region in Malawi. The countries were chosen based on their respective economic development to create a discussion on the economic factors that seem important to explain the diverging findings. Ghana constituting the most affluent case has enjoyed a relatively strong economic development for the last decades. Ethiopia has too enjoyed an economic upswing in past years but remains at a lower level than Ghana. Finally, Malawi exemplifies a country, which has suffered from economic stagnation and a very low standard of living.

The analysis first descriptively shows that there exists a positive association between household income diversification and wealth following two different rural off-farm definitions for the sample. To disentangle it further, the use of econometrics allows the analysis to control this

relationship for education, age, household size, and employment contract, which are found to have different types and strengths of correlations depending on the country, and year analyzed. The general finding is that there exists a positive relationship between income diversification and wealth, with the clearest and strongest relationship for Malawi. However, if diversifying to on-farm agricultural wage work, there is vague evidence on a negative correlation with wealth for Ghana.

1.1 Aim and Research Questions

As shortly explained, this study aims to look further into the relationship between income diversification and welfare in rural SSA as one potential way out of rural poverty. The first research question I will try to answer is specified as:

1. *Is there a positive relationship between household income diversification and an increased amount of wealth relative to solely engage in farming among the rural population in SSA?*

If there is a prevailing positive relationship between household income diversification and wealth in the specific subsamples, then the second question will be reasoned about with the available data related to the same subsamples.

2. *Does income diversification to the non-farm sector[1] have a more positive relationship with wealth compared to diversification to the on-farm sector?[2] If so, which extra job in the non-farm sector seems to have the highest positive relation with wealth?*

The structure of the thesis is that first, a literature review will introduce literature and former findings on the farm- and off-farm sector before laying out income diversification and how these three interacts with welfare. Second, there will be a detailed discussion on the data collection, management, and methodology to investigate the topic, if income diversification has a relationship with increased welfare. Third, the results of the study are evaluated and discussed concerning the three countries different development paths, before concluding the thesis.

---

[1] Defined as all economic activity that is not the production of primary agricultural commodities (Haggblade et al., 2007, pp, 18-20).

[2] Defined as agricultural wage labor on a farm that is not the farmers (Haggblade et al., 2007, pp, 18-20). The definitions will be developed further in the methodology section.

2. Literature Review

In this chapter, I will start by describing the farm-sector and off-farm sector in rural SSA before connecting these two and discuss income diversification and welfare implications with the use of previous research. Subsequently, I will conclude this chapter by shortly discuss why it can be that diversifying to the non-farm sector may correlate with higher welfare than to diversify into the on-farm sector.

2.1 Rural Sub-Saharan Africa and the Farmer

The golden standard of rural development has long been considered the smallholder-farming model (see e.g Alesina & Rodrik 1994; Deininger & Binswanger, 1995; Deininger & Squire, 1998; Griffin, Khan, and Ickowitz, 2002). However, the homogenous subsistence family farmer, once considered a convenient characterization of a typical farmer in contemporary Africa (Oya, 2007), has never really existed. With rural development, there has been a concentration of land ownership, making a large share of rural households landless. This has created an expansion of the rural labor markets since the dependency on wages for the population to survive has increased. Rural labor markets have become an integral and active part, and nowadays, many of the smallholders hire wage workers whereas others depend on selling their labor (Oya, 2010). The changing dynamics together with the emergence of a richer class owning larger farms can be beneficial for the owner of the farm but also the poorest part of the rural population if it allows for the expansion of commercial farming. Higher-income quintiles tend to rely more on commercialized cash crops and off-farm income (Bigsten & Tengstam, 2011) and if the poorer farmers can search for wage employment with a higher wage and leave the smallholder sector (Sender, 2002; Sender, Oya, and Cramer, 2006) it can improve their welfare.

What makes these changing dynamics even more prominent to discuss is that in later decades in SSA, a process of "deagrarianization" ha been taking place, including occupational adjustments, income reorientation, and spatial relocation of rural inhabitants. The rural population has moved away from a livelihood strictly based on agriculture to non-agricultural activities and some households have left agriculture completely (Bryceson, 1996; Bigsten & Tengstam, 2011). The agricultural sector is no longer the main determining factor for welfare in rural SSA, with the off-farm sector becoming increasingly important. This means that we need to understand the off-farm sector to grasp how to alleviate poverty in rural SSA.

2.2 The Off-Farm Sector

Already a few decades ago in Kenya, non-agricultural work was essential and sometimes more important than agriculture for poor rural households (Carlsen, 1980, pp 215). This observation is now generalized to SSA; poor households depend increasingly on the earnings from the off-farm sector to survive (Francis, 2000, pp 55; Barrett, Reardon, and Webb, 2001a; Haggblade et

al., 2007, pp 3, 57). The off-farm sector now constitutes a significant share of the income for farm households in rural Africa, around 40% of the average household income with a variation between 15 and 93% (Bryceson & Jamal, 1997; Little et al., 2001; Haggblade et al., 2007, pp 4).

The rural off-farm sector is heterogeneous in both scale and composition. The variation is due to a combination of factors, such as natural resource endowment and agricultural structure that includes everything from part-time self-employment to large-scale agro-processing, often with seasonal distribution (Carlsen, 1980, pp 94-95; Haggblade et al., 2007, pp 5). This heterogeneity makes labor productivity in the rural off-farm economy to diverge by a factor of 10-20. Economic activities with higher returns than the agricultural sector tend to be inaccessible for the absolute poorest rural Africans due to multiple entry barriers. Thus, the already poor continue to be locked into low-productivity activities, inclined to rely on earnings coming from farming (Barrett et al., 2001a; Haggblade et al., 2007, pp 10, 71; Doss, McPeak, and Barrett, 2008).

The capacity to diversify into the off-farm sector and overcome the entry barriers depends largely on capital assets. For instance, human, social, physical, organizational, and financial capital tend to positively correlate with income and wealth. Further, the most important determinants of off-farm income based on multiple studies on rural Africa seem to be educational attainment (Barrett et al., 2001a; Bigsten & Tengstam, 2011). All these factors are normally scarce in the poorer population which faces limitations of inputs such as land and skilled labor (Haggblade et al., 2007, pp 57-58), physical access to markets (Smith, Gordon, Meadows and Zwick, 2001; Lanjouw, Quizon, and Sparrow, 2001) and public services (Barrett et al., 2001a). Therefore, the only opportunity for the poorest is to leave agriculture and move to the informal off-farm sector (Collier & Lal, 1986, pp 128). Richer households, on the other hand, who often have these capabilities, have been found to partake in off-farm activities with higher wages (Lay, Narloch, and Mahmoud, 2009). Thus, the relationship between the off-farm sector and welfare may not be that clear; it may be negative for some parts of the population and positive for others.

All this considered, it does not mean that the wealthier population necessarily move out of agriculture and only partake in the rural off-farm economy. The decisive point to make is that the wealthier have larger freedom to choose among different options relative to the poor (Barrett et al., 2001a; Haggblade et al., 2007, pp 133-134). The already poor fail to catch up through off-farm earnings and it creates a vicious circle more common in rural Africa than in other regions with an equivalent income level. This suggests that there are processes or characteristics specific to the area (Reardon et al., 1998; Reardon et al., 2000; Barrett et al., 2001a). One process shaping this can be the previously mentioned long-term deagranisation observable across SSA (Bryceson, 1996). It has created the need for diversification of rural livelihood and increased the importance of off-farm activities (Francis, 2000, pp. 62), both as a primary job, and a secondary. Bernstein (2010, pp 106) summarize it in a good way: *"[…] the practices, fortunes, and prospects of farmers are increasingly shaped by their activities outside their farms and the incomes those activities provide […]"*. However, it is still unclear if the relationship

between the off-farm sector and welfare is positive or negative. So far, the discussion seems to point toward that it is highly dependent on which income group the household is in. The discussion so far suggests that diversifying the sources of income may be a good idea no matter what, something that will be investigated further.

2.3 Income Diversification

Income diversification tends to be the norm, but in rural SSA, only 30-50% of the households practice both off-farm activities and farming activities simultaneously (Francis, 2000, pp 55; Barrett et al., 2001a; Haggblade et al., 2007, pp 117). The expectation would be that diversification is negatively correlated with income level and that households participating in both farm and off-farm activities are less efficient, but in this way partly avoid risk and survive. If this were correct, richer households should be more specialized and thus more efficient with a positive relationship between specialization and income, which is reasonable according to basic economic theory. This intuition causes the sign of the relationship between diversification and welfare to be uncertain. If considering evidence from rural Africa more extensive diversification is found among richer households in rural Africa with a bigger share of the total income coming from off-farm income. Diversification increases with income and wealth but with the side note that the individuals within the households still specialize (Reardon, 1997; Francis, 2000, pp 61; Barrett et al., 2001a; Haggblade et al., 2007, pp 117, 121, Bigsten & Tengstam, 2011). Thus, since extensive diversification is found only among richer households, it is still not clear if this is true over a larger sample including the poorer parts. To understand why diversification takes place, and why it is important for the population in rural SSA, we need to investigate the motives for it.

Diversification is essentially about stabilizing flows of income, consumption, and loosen constraints (Barrett et al., 2001a). There are many motives to diversify and the literature typically distinguishes between push- and pull-factors. The push-factors can be a household's way out of poverty (Haggblade et al., 2007, pp 115) and they are normally more complex. Households follow "risk management strategies" by turning to more off-farm activity due to the lack of insurances and credit markets (Haggblade et al., 2007, pp 43). This allows for income smoothing over time, and the poorer parts choose to diversify into economic activities with a low positive covariance with their agricultural returns. It could be a type of "distress diversification" if there is a shortage of land or a low-potential environment for the smallholder. Typically, the diversification is then into low-return work, often agricultural wage work (Haggblade et al., 2007, pp 43; Bigsten & Tengstam, 2011) which the rural poor are highly represented within (Haggblade et al., 2007, pp 57-58). Push-factors are often weather- or climate-related. When the seasonal income from farming drops to a position when it is not enough to survive off-season because of for example price variation. A transitory drop in earnings as a consequence of drought. A permanent drop in earnings from farming because of chronic rainfall, disease, market failures, or land- and labor limitations. A last factor can be large variations in farm earnings and agricultural output because of rainfall instability (Haggblade et al., 2007, pp 127). Thus, the positive relationships that seem to be between

income diversification and welfare are still not certain. If a household is pushed into the off-farm sector, it is not sure that their welfare will increase, since it does not have to be an optional choice. It may as well be that their welfare decrease. If the push-factors show up to be what is explained, we may see a negative correlation between income diversification and welfare.

The pull-factors, on the other hand, are associated with an upward spiral of wealth (Haggblade et al., 2007, pp 115) and it tends to be already high-income households that are "pulled". Farm households in more prosperous agricultural zones want to diversify into off-farm activities that require skilled labor, have higher returns, and a lower risk relative to farm activities. The possible higher returns may make it achievable for farm households to accumulate capital. It becomes a positive feedback loop where credit constraints are loosened and allows for the enhancement of the own farm and reinvestments into their agricultural production and education (Collier & Lal, 1986, pp 258; Evans & Ngau, 1991; Barrett et al., 2001a; Ellis & Bahiigwa, 2003; Ellis & Freeman, 2004; Haggblade et al., 2007, pp 44, 126; Lay & Mahmoud, 2008). The pull-factors point towards that there exists a positive correlation between income diversification and welfare, at least for the richer households. Nevertheless, they tend to make poorer households stuck in short-run recovery strategies (Haggblade et al., 2007, pp 139). Typically, wealthier households diversify because of profit maximization whereas the poorer households point out risk minimization (Haggblade et al., 2007, pp 44, 126). Unfortunately, these are factors and causalities I cannot study here, but they give clues on how potential relations may display. It is still unclear if diversification has a positive relation with welfare, it seems to be dependent on the motives.

As a final discussion, I will review what previous research has found on whether income diversification tends to improve the livelihood of a household. The problem is as mentioned earlier that even though the returns can be much higher than the returns from farming, they are only available for the already richer parts of the population. Reardon (1997) performed a survey of income-diversification literature on Africa and found that the households which already had the highest farm income, also had the highest share and level of income stemming from off-farm activities. On the same track, Barrett, Bezuneh, Clay, and Reardon (2005) find that off-farm earnings, and in particular, when their origin is from other than unskilled labor, have a positive relationship with higher incomes and greater upwards income mobility. This still implies that the relationship between the two may be negative for the poorer parts, and thus the average relation is unclear at this stage and should depend on which part of the income quintiles that are the most eminent.

Barrett et al. (2001) found a positive correlation between off-farm income and household welfare, and in Tanzania, Ivory Coast, Kenya, and Uganda it was found a positive relation between off-farm income and income- and wealth levels. These studies also display that more off-farm income makes the growth in consumption more rapid (Lanjouw et al., 2001; Barrett, Bezuneh, and Aboud, 2001; Canagarajah, Newman, and Bhattamishra, 2001). Studies also find that the possibility to have an external income increase productivity in agriculture and give a higher overall income in Uganda, Tanzania, Malawi, and Kenya (Evans & Ngau, 1991; Ellis & Bahiigwa, 2003; Ellis & Freeman, 2004; Lay & Mahmoud, 2008). Bigsten and Tengstam

(2011) analyze how income diversification affects incomes for smallholder households in rural Zambia, finding that more diversification tends to be associated with higher incomes per worker. The change in a household from only being farmers to a more diversified livelihood can raise the per worker income by 25-100%.

Previous research has also studied other welfare aspects than wealth and income. Berkvens (1997, pp 12) found that households in Zimbabwe who had earnings from more than one type of employment were generally less prone to appeal for food aid. Block and Webb (2001) study Ethiopia and find that income diversification is positively related to higher nutrition. In Malawi, households that were participating in non-agricultural production experienced on average 225% more cash earnings annually relative to those not active in the non-agricultural sector (Tellegen, 1997, pp 152). In crop commercialized areas, this can be expected to be less prominent and important since they already have made a large shift to sales of crops. At the same time, the returns for off-farm activities are higher both near towns and in zones that are more favorable to agriculture (Haggblade et al., 2007, pp 126), thus the phenomenon of diversification should be seen to a higher extent in more commercialized areas, which makes this area more interesting to study since it remains to find empirically. In Kenya, it is found that when combining farming and other livelihoods, farmers can become commercial smallholders naturally (Francis, 2000, pp 22) which may be positive for welfare.

There are however also negative aspects of diversification. It may deplete the energy of the workers when they have to focus on two different activities instead of giving full attention, time and put all resources into just one activity (Francis, 2000, pp 20), a factor that has not been studied that well. With this caveat in mind, can income diversification lead to an increase in welfare, and subsequently lift people out of poverty? The findings of a positive relationship seem to be often for the richer quintiles, whereas the poorer are mainly theorized to benefit from diversification, yet with little empirical evidence. Furthermore, findings on the consequences of income diversification are rarely differentiated between rural and urban areas. An exception is a study by Bigsten and Tengstam (2011) who studies rural smallholders in Zambia and find the livelihood to be increased for the studied sample. To further build on this validity, this study will look at three rural areas in three African countries, where especially Ethiopia and Ghana have been studied less. Both rich and poor households are included, and in this way, I hope to find how the relation looks on average and if we can say something about the relationship when both poor and rich are included, in the rural area.

2.4 Income Diversification and the Rural Off-Farm Economy

To shortly discuss the second research question, and summarize what has been review already, diversification to the non-farm sector, in theory, seem to have a more positive relationship to wealth than diversifying to on-farm wage work. For a smallholder household with little land compared to labor and no land markets that are well functioning, some labor will work on the own farm whereas some will be hired to wage employment on other farms, which tend to be an informal job with low returns (Collier & Lal, 1986, pp 128; Haggblade et al., 2007, pp 43; Bigsten & Tengstam, 2011). As discussed, this can be distress diversification into low-return

agricultural work due to factors the household cannot influence (Haggblade et al., 2007, pp 43; Bigsten & Tengstam, 2011). The better-remunerated non-agricultural work often has entry barriers the poor cannot pass by. This leaves the poorest parts of the population stuck in a dynamic poverty trap (Barrett et al., 2005; Haggblade et al., 2007, pp 122) since they cannot accumulate enough capital to invest in their farm or in for example education for themselves. This makes it unclear if both types of diversification have a positive correlation with welfare.

To understand this better, the reasons why the off-farm economy can increase welfare are important to consider. First, if the subsector is good in terms of remuneration. Second, if the off-farm activity can act as a safety net for the poor being "pushed" out of their traditional occupation. As discussed earlier, natural disasters and extreme weather occur frequently and can be catastrophic for the poorest parts of the population that might be partly dependent on agriculture. An illness or injury can make these poorer households look for alternative incomes. There are also indirect reasons to consider since the poor are often limited in their direct connection to the off-farm sector. If the sector grows, they may be helped by the income increases of the workers in that sector, which creates a demand for more agricultural products produced by the poor farmers. Likewise, off-farm incomes in the rural area can be the financial capital needed for new investments in agriculture to increase productivity and this leads to an inclined demand for agricultural labor (Francis, 2000, pp 20; Haggblade et al., 2007, pp 62-63). I cannot study the specific mechanisms here but it is important to look at what type of diversification that seems to be the one more positively correlated with wealth, and also which jobs that will bring humans out of poverty. The route to poverty reduction will differ depending on this. If to increase welfare can be considered rural development, which I argue, then the golden standard for rural development might not anymore be the smallholder but rather the "diversifier".

3. Data and Methodology

3.1 Definitions

Before going deeper into the data and methodology used for the analysis, conceptual issues must be taken care of when examining livelihood diversification. One part is the inconsistent terminology used in the literature, and the goal here is to define the concepts I will use extensively to make this analysis as clear as possible.
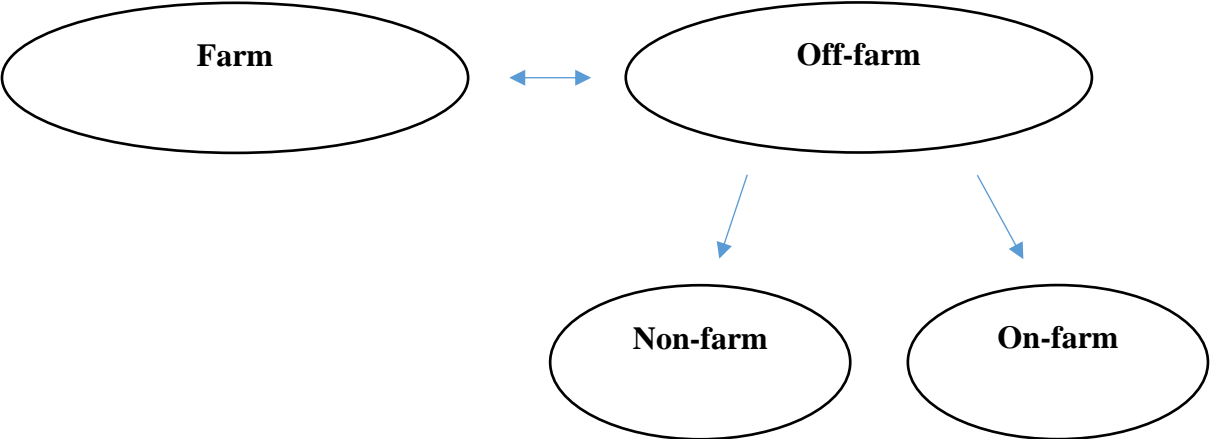
Figure 1: Definition of the off-farm sector

A quite simple way to organize rural household income can be to look at the income that stems from productive assets divided by sector (agricultural against non-agricultural), function (self- and wage-employment), and location it comes from (local versus migratory). This and the nature of the product is what typically matters (Barrett et al., 2001a). To be able to examine the area of the rural off-farm economy and diversification it is critical to make a clear distinction between the four categories in figure 1.[3] Starting on the right side, I will use the off-farm as an umbrella term for the two definitions below and adapt the definitions from Haggblade et al (2007, pp 18-20) to fit my data and make the analysis as clear as possible.

Non-farm use the division by sector and it is all economic activity that is not the production of primary agricultural commodities, thus, it includes all ISIC groups except A (Agriculture; forestry and fishing). One fundamental component of this group is agro-processing which often takes place at the farm were the primary commodities come from, but it is still a part of the non-farm definition. The same is true regarding where it takes place, it can be in factories, or by traders as long as it is not agriculture (Haggblade et al, 2007, pp 18-20). Here, the opposite, farm is defined as ISIC group A, agriculture; forestry, and fishing.

On-farm instead makes a locational and functional division. In this context, it means that the person is working on another farm than their own as labor, and often for a wage, an important perspective when analyzing this topic (Barrett et al., 2001a; Haggblade et al, 2007, pp 18-20). The farm-sector as an opposite is if the person is working on the own farm or the farm of their

---

[3] Some authors as Bigsten and Tengstam (2011) narrow it down even further and include "own business" as its own income-generating activity, but that division will not be done here.

family in subsistence agriculture. An intuitive way of thinking is that the off-farm and farm together should be 100% of the occupations that a person can engage in, so what does not fit in one of them will fall into the other definition. The analysis throughout the rest of the thesis will use these definitions and build the income diversification variables by the use of them, which is why it is important to define them.

Next is to clarify the concept of income diversification and how I will use it further. If a household is diversifying their income, it is when the female and her partner combine their work according to the definitions. To make a simple example, non-farm diversification will be referring to when one person in a household is doing non-farm work and the other farm work, such as subsistence agriculture and selling cell phones. Both definitions will be under scrutiny, but first, the data I use will be discussed.

3.2 Data

For the analysis, data is gathered from one source, the IPUMS Demographic and Health Surveys (DHS) (Heger Boyle, King, and Sobek, 2019). I consist of surveys as a part of the DHS by USAID (ICF, 2020a) and even though this data is already harmonized, the choice to not use DHS data directly is that IPUMS-DHS gives several advantages for my thesis. First, for the use of regions, harmonized geography variables over time (IPUMS International Geography Variables) are introduced to make it possible to compare information between subnational regions over time. Second, a variable is called by what it is, the current work of a woman is "wkcurrwork" instead of as in the original DHS-file "v705" which makes it more efficient to arrive at the analysis of the data. The third reason is easier data management such as that possible comparability problems are highlighted. It is possible to download a complete harmonized dataset directly, and a few variables are bridged between surveys to make it easier to analyze over time. The IPUMS data facilitates multivariable analyses both over time and between countries with the DHS microdata. Additional data would have been preferable, and over time for the same individuals, but this is not permitted with this data set.

The surveys contain answers from women in childbearing age, 15-49 years. Since the data is only on females between 15-49 years, it may be that there are more women within the households important to consider. This should not affect the analysis too much however and is party taken care of by controlling for household size. Furthermore, with an expected average life length of approximately 48-68 years for females in the three countries of interest (Ethiopia, Ghana, and Malawi) over the period for the different waves (The World Bank, 2020), it is not bold to assume that around the age of 49 years, the females start to wind down on their income labor and spend their time with household work. In the data set the females have answered a few questions about their partners, such as where they work, their age, and their education which I will use. The information on these women's partners are sparse in general, and there do exist another data set for only males based on similar surveys. However, the decision to not include this data comes from that the variables that would give extra information than what I already have are as scarce. Second, in the data set that I am using for the analysis, the partners are already connected to the females and the corresponding household, which does not have to

be the case if merging the data set on males. By not merging the data sets, minor errors will not arise in the process of connecting males and females manually since household identifiers are not matching at all times. There are caveats with this choice as well. Only married couples or pronounced partners will be included in the analysis. A few times it may be that more than one couple is in a household because of that this is the living situation in rural SSA, and in particular for the poorest. To only use married couples decrease the sample a little bit since women without a partner are not included. Besides, all couples where none of the individuals are committed to agriculture are dropped to be able to answer the research question if income diversification on a household level, the combination of farm and off-farm labor, leads to higher household welfare compared to only engage in agriculture. To make it clear, the focus is on households where there according to the data is one couple residing (even if in reality the household is bigger). This choice should be including big parts of the active rural labor community.

The main outcome variable to measure the welfare of the household is the wealth-index constructed by DHS. It exists both as an index, but also in quintiles to easily depict different wealth groups within a country. The values go from negative to positive (in the full sample -1.32-4.158), and the many negative values come from that the DHS wealth index accentuates the contrast between urban and rural poverty. It values rural assets negatively in comparison to urban infrastructure by construction (Wittenberg & Leibbrandt, 2017) by not including items that may be valuable in the rural parts or more items that are more common in urban areas (Rutstein, 2008). The index is a composite measure of a household's cumulative living standard and calculated with a household's ownership of specific assets. Examples of them are the material the roof is constructed of, if the household owns a bike or television, what type of stove used, types of sanitation facilities, water access, et cetera, and almost all of them increase the standard of living (ICF, 2020b). The difficulty with valuing assets precisely in rural Africa (Barrett et al., 2001a) makes this wealth index good to use. Another point is that the wealth index considers assets over time, thus it can be a depiction of expenditure data and this is conceptually a good reflection of a household's welfare according to Lipton and Ravallion (1995). It does not fluctuate as incomes tend to do when a household member loses their job. Second, consumption, in general, can be a good indication of the average welfare over time since it partly exposes information about earnings both past earnings such as future earnings. Nevertheless, this type of index can also include noise such as consumption smoothing constraints the poorest experience since they have limited borrowing opportunities (Lipton & Ravallion, 1995). The general meaning of the index is consistent across the different subsamples, but the level of wealth implied by a specific quintile is country-specific. A specific set of assets and services in one household could make it a top quintile household in a poorer country, but in the middle quintile in a richer country. Therefore, the index should be used to illustrate the relative economic position of the household within a particular country and optimally at a specific point in time (Heger Boyle et al., 2019). When I am using quintiles, the poorest are the bottom 20% in the country, and the richest, the richest 20% in the wealth distribution.

The main independent variable is a dummy for household income diversification based on either non-farm or on-farm diversification. It is created from what kind of occupation the female and her partner undertake and how they combine them. Income is the preferable measure to use when investigating household income diversification since it tends to give a more complete picture of the rural off-farm economy, but what is used in comparable studies, and here is employment (Haggblade et al., 2007, pp 4, 13-15). The timing of the collection of data affects the visibility of measurement and thus employment is clearer to use because of the smaller variability, but it is also because of data limitations. There are many thorough steps in constructing this variable, which will be explained extensively step by step in the methodological approach chapter.

Ultimately, there are multiple variables important to control for as determinants of wealth. These variables also come from the IPUMS-DHS surveys and are age, the size of the household excluding children 5 years and younger, the employment contract for the female, and education measured by achievement or female literacy. Data does not exist for male employment contract nor literacy. Age may be important to determine wealth since when life progresses, typically more assets are accumulated. If you are young and a new entrant to the labor market, it is very plausible that this person has less assets. Age is also exogenous, wealth cannot affect the age of a person, thus no reverse causality. As a robustness check, the female age is changed to the male age. Another important variable is the size of the household. With more people living in a household, there should be more potential labor to increase the number of assets the household has. It can also be the other way around, that bigger households are poorer. Poorer households tend to have more children which should cost more money and hence decrease wealth. Since the diversification variable only controls for the couple and their occupations in the data and not the rest of the members in the household, this helps to take care of the possibility of more labor in the household. The likeliness that there are female household members under the age of 15 or males that are not married and working for the account of the household is pretty big, but this is a check the data does not allow me to. Instead, this control helps to get rid of some of the effects on wealth. The household size variable is cleaned of all children of age 5 and lower since they are not expected to be important for determining wealth. The next control is how the employment contract looks for the female since this variable does not exist for males. Women in these areas often perform non-full time labor (Haggblade et al., 2007, pp 4), and thus, this is most likely an important determinant of the wealth of the household and may depend on harvests or linkages to harvests. The choices are either "occasionally", "seasonally", or "all year" coded as one, two, and three. Finally, a variable that most likely has a substantial impact is education. It will be included, but I will not focus on much of the discussion on it since the question always persist if it is the hen or the egg, there are many transmission channels both ways. The main variable of education is measuring educational achievement for the female since this variable has the most observations. The equivalent variable for males is interchanged for a robustness check, and a literacy variable measuring actual reading capability at the time of the survey instead of female educational achievement for a second robustness check.

There are other variables that in an optimal situation would be controlled for since they most likely matter for wealth. Examples of these are parents' education, parents' income and/or

wealth and infrastructure (market access). On a macro level, it could be taxes, natural disasters such as droughts and floods, locusts, and commodity shocks. All these would be interesting to add but are not due to data limitations, and this may create endogeneity problems such as the omitted variable problem.

3.3 Sample

To be able to answer my questions for the analysis I am going to perform a comparative study on one region in each of the three countries: Ethiopia, Ghana, and Malawi. The motivation to pick just these three is based on their different economic development over the last decade in figure 2 and country characteristics explained below. One could argue that I should have picked more countries, but I am confident that this selection will show three very different and multifaceted sides of income diversification and wealth in SSA.
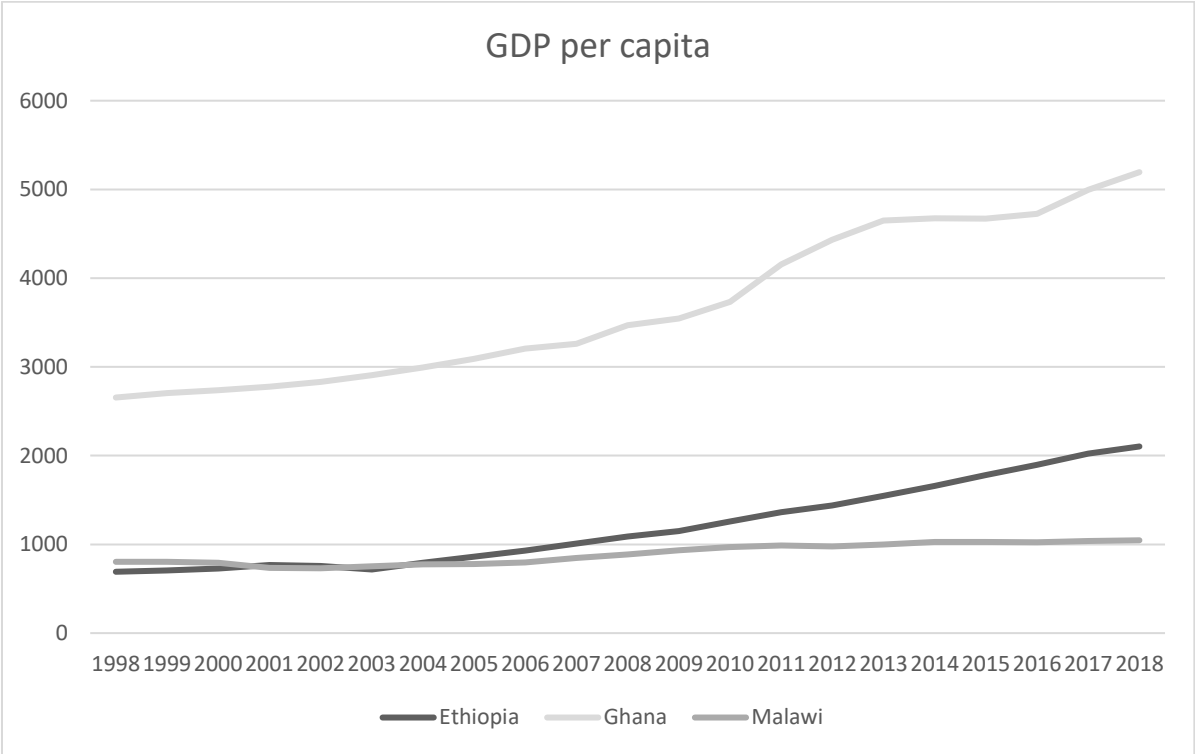
PPP (constant 2017 international $) (The World Bank, 2020).

Figure 2: GDP per capita over time

Ethiopia has had good economic development over the last 20 years starting from a very low level (The World Bank, 2020). Around 85% of the population lives in rural areas with scarce livelihood opportunities. Smallholders in Ethiopia produce between 90 and 95% of the agricultural output of the country, and the agricultural sectors' inclusive growth is a major driver of poverty reduction (IFAD, 2020a) with most smallholders only engaging in subsistence agriculture (Bonaglia, Labella, and Marshall, 2008).

Ghana during the last 2 decades, has had good economic development but it is also a comparatively rich country in the SSA region (The World Bank, 2020). It is very agricultural

and in contrast to many other African countries has a largely smallholder-based agricultural commercialization (Yaro, Teye, and Torvikey, 2017). The smallholders typically cover less than two hectares and are often among the rural poor, which still primarily earn their living through agriculture (IFAD, 2020b). The economic differences between Ethiopia and Ghana should create diverging processes for the two countries, making it an interesting comparison. Both are prosperous starting from different income levels and if the same is true for their rural regions, growth in rural off-farm labor should mean that agricultural incomes are increasing in a demand-led growth in high-return non-farm activities (Haggblade et al., 2007, pp 15).

At last, Malawi is interesting because it has essentially had economic stagnation for the last two decades and is significantly poorer measured by GDP per capita than the other two countries (The World Bank, 2020). The livelihood for more than 30% of the rural households comes only from farming or fishing and approximately 25% combine their farm work with other jobs that are often within agriculture and poorly paid. This leads to a highly seasonal labor market with shortages of labor during the cropping season and underemployment for the rest of the year (IFAD, 2020c). If there would be rural off-farm growth of labor seen in a stagnant rural economy as Malawi, it can be a sign that households with few assets and not many opportunities are pushed into low-paying off-farm activities (Haggblade et al., 2007, pp 15).

There exist data from the surveys by country from 1988 until 2016. However, a few of these are dropped. The year 1988 only exist for Ghana and does not inhibit information on a regional level, thus, it is not included. The years 1992 (Malawi) and 1993 (Ghana) are dropped because of the lack of a survey close in time from Ethiopia. Besides, the relatively large period until the next wave may compromise the study due to the possibility that the methodology of how to perform surveys changed. In conclusion, for Ethiopia, surveys from 2000, 2005, 2011, and 2016 are included, for Ghana, 1998, 2003, 2008, and 2014, and Malawi 2000, 2004, 2010 and 2016. A sample spreading from 1998-2016 should be sufficient to empower the analysis and investigate the research questions.

The analysis is narrowed down to one relatively more commercialized region from each country to reflect characteristics more alike than it would be for the whole countries and make the comparison equal between countries. These regions are interesting because they allow people to search for wage employment more extensively and leave the smallholder sector (Sender, 2002; Sender et al., 2006). Furthermore, returns tend to be higher in zones more favorable to agriculture and where also diversification should be seen to a higher extent, thus commercialized areas (Haggblade et al., 2007, pp 126). Besides this, in commercial farming areas, rural households can earn capital, which they can invest in farming instead of entrusting formal credit, and thus give support to the agricultural growth (Francis, 2000, pp 20).

In Ethiopia, as introduced at the beginning of the thesis, the region for analysis is Tigray. It is located in the north of the country bordering Eritrea where 50% of the commercial farmers in the country live (UNDP Ethiopia, 2013). For Ghana, the region is the Western Region due to the region being Ghana's main producer of cocoa. Ghana as a country is one of the biggest cocoa producers and exporters in the world, and it is mainly smallholder-based agricultural

commercialization (Bray, 2014; Wessel & Foluke Quist-Wessel, 2015). Finally, in Malawi, the more palpable choice of a region may have been to use the Southern Region because it is the most commercialized region, with the commercial hub Blantyre City. However, it also has a very high population density (National Statistical Office of Malawi, 2018) and at the same time the smallest average landholdings in the country (Fisher, Holden, Thierfelder, and Katengeza, 2018). These factors make this region an outlier since the population does not have the same prerequisites as in the other regions, thus, not representative in this comparative study. Instead, the Central Region will be up for scrutiny since it is the most agricultural and still commercialized. Malawi is one of the biggest exporters of groundnuts in Africa, with two major groundnut producing districts in the Central region displaying a good example of the country's commercialization in the agricultural sector (Msofi Mgalamadzi et al., 2018).

3.4 Methodological Approach

Here I will go through step by step how I have used the data and created the variables needed to examine income diversification and welfare. First, since the data is from surveys, caution has to be taken because these may not be 100% representative and correct. It may be that the respondent is not comfortable with being honest because it can unfavorably present them (Wolff, 2015). Nevertheless, I acknowledge this and trust DHS that they have done their best to minimize these biasing factors in the data collection.

The rural area is the focus of the analysis and thus all observations live in the rural area. The definition of this area diverges between countries since the DHS-program uses the national definitions for the urban-rural definition in their surveys. Official definitions are often 5000 people or fewer and the cutoffs vary between countries. Sometimes other attributes are used as well to define the areas such as important infrastructure (Haggblade et al., 2007, pp 18-20). For Ethiopia, until 2007 it essentially defined the urban area as localities with 1000 or more inhabitants that were mainly involved in non-agricultural activities. After 2007, Ethiopia defines urban areas as localities inhibiting 2000 or more people, all administrative capitals, plus smaller localities of 1000 people or more, but only if those people are committed mainly to non-agricultural activities. Ghana's localities with 5000 or more citizens are defined as urban while Malawi defined all townships, town planning areas, and district centers as urban (Heger Boyle et al., 2019). The rural area I Use in all these cases is thus, "the rest", the parts that are not urban. One caveat with this is that the definition of rural area changes in the sample for Ethiopia, which can create problems by altering the composition. A second more general pitfall for this analysis can be that even if a household is in the rural area, the activities that the individuals within it undertake can be a mix of rural and urban (Barrett et al., 2001a), but hopefully, this will not bias the study too much. Rural income is further seen as the income for someone residing in a rural area with the assumption that the person also works there. Since all data is on households in a rural area, this will not be discussed more. Further, only usual residents are included and not those marked as "visitors" in the survey.

To construct the household income diversification variable, the explaining variable, multiple steps, and data is needed. The first step is to consider data on the work the female and her

partner currently engage in. To start and simplify the analysis I merge the alike categories "household and domestic work", "household, domestic, and services" and "service work" into "household, domestic, and services" because these occupations probably have around the same income level. One may suspect that this category can indicate if the person works in their own household. However, this is not the situation since the question asked in the questionnaire is "aside from your own housework, have you done any other work?" (ICF, 2020a). Thus, it is clear that this occupation is not problematic for the study. Considering that it is only rural couples that are usual residents and where at least one of the peers is working in agriculture, there are 10 different occupations left in the sample that the females engage in, of which three are agricultural categorizations (appendix A, table 15). The three agricultural occupations are "self-employed agricultural worker", "agricultural employee" and "agricultural" where the latter is a more general definition including fishers, foresters, breeders, hunters as farmers (ICF, 2020). This can distort the analysis since fishers, foresters, breeders, and hunters are not the focus and due to this composition; the assumption has to be that they are all agricultural which the vast majority should be. If I would drop this category, I would lose 46% of all female observations. The analysis and the comparison should not suffer too much from this inclusion since it should be approximately equal between the three regions. The outlier of the three is the Tigray region of Ethiopia because the vast majority of the population cannot access any fishing waters, and the region consists of many deserts. The Western Region of Ghana and the Central Region of Malawi lies in close connection to big water resources and green forestries (Google Maps, 2020). Further, the equivalent occupational variable for the females' partners is divided into 11 groups equivalent to the females, but with the addition "professional, technical, or managerial" (appendix A, table 15). The same holds here for the agricultural variable, and if these ambivalent observations would be dropped for the countries, the analysis would lose approximately 60% of the observations for the partners' occupation. Again, another problematic feature of the categorization of males can be that since the females respond to the questionnaire, this variable may be less detailed and not as representative as one would wish. However, the variable is kept with these drawbacks in mind.

Second, this data forges one set of non-farm diversifiers and one set of on-farm diversifiers where the rest in each case corresponds to the farm definition (appendix A, table 15). The non-farm variable only includes everything except the three previously mentioned agricultural variables whereas the on-farm variable includes "agricultural employee" since it handles diversification to wage work on other people's farms. Besides, for the females with the help of extra data in the survey, I manage to gather more observations by including females working in "agriculture", if and only if, the person also declared "for someone else". Therefore, two extra additions take place for the opposite farm category as well, the females working in "agriculture" and "works for self" or "for family member". A caveat arises here because of the small sample declaring their occupation to be an agricultural employee, and the impossibility to refine the "agricultural" occupation further for men.

Third, the couple is declared as performing household diversification according to the two definitions if either one of them is assigned to farm and the other non- or on-farm. On-farm diversification is only found in Ghana in 2014 and Malawi 2016. This is due to the low amount

of observations for men being "agricultural employees", 38 in the full sample. This may reflect reality, but it can also be due to that the females responding are not sure of the occupation of their partner and simply say agriculture. To exemplify, if a couple has the female working within agriculture and the husband within sales, then they are performing household diversification according to the non-farm definition. If it would be the case that the female works with (subsistence) agriculture, thus, agriculture at the own farm and the husband with agricultural wage work at another farm, then they are also diversifying, but according to the on-farm definition. If they are doing household diversification, the variable is one, if not it is zero, thus, a dummy variable because of the qualitative characteristic. See the number of households diversifying according to non-farm contra on-farm per subsample in appendix A, table 16.

Finally, to clean the data from "extreme cases" and avoid unnecessary biases and non-representative outcomes, households with the family structure "unrelated adults" are discarded from the data. Besides, a cutoff is if a household consists of eight or more members after I have subtracted the children of age five and lower, then it is also dropped. The consequence of this is that households with many children in working age and not yet married are dropped since there is no more information to find about these peers in the data. However, there may still be some left that the household variable may account for. These considerations decrease the number of observations of approximately 9%. Further, if a household consists of many wives with the same husband based on his age, job, and household identifier, I drop the observation. This is to ensure that the same household and/or couple is not analyzed several times. An extra "eyeball check" has been made in the dataset of the households, to ensure that there are not many possible elder labors living in the household.

3.4.1 Analytical Approach
The approach of the analysis is that the two questions stated in the beginning will be investigated empirically. First, an extensive descriptive analysis with "eyeball regressions" will be made on the research questions. Consequently, the first question will be investigated with the help of econometrics. If the question is not falsified at this point, I will connect the found differences between the countries back to their diverging development paths. To find that income diversification correlates with wealth is important, because otherwise there is no reason to discuss my second research question. Since the sample is not identical every period, and do not follow the same individuals, panel data methods are not used. There are data sources available making it possible to follow the same individual over time, but then the comparative nature of the thesis would be sacrificed due to that this data is from studies made on a single country.

I will try to find a relation between household income diversification based on the two different definitions and wealth by comparing the households only engaging in agriculture versus the ones that diversify, either to non-farm or on-farm. To repeat and clarify, this has led to that all couples where not at least one of the peers are working in agriculture are dropped. One important note is that by analyzing diversification and wealth on a household level, a rough assumption has to be made. We have to make the arbitrary assumption that the prosperity

credited to the household is spread equally (Lipton & Ravallion, 1995). This can be seen as a caveat with the study but will not be discussed further.

The econometric analysis is using Ordinary Least Squares (OLS) and every regression will be executed first for the whole country (pooled over time) and second divided into the survey waves (cross-sectional). This to further identify if there are different traits the countries inhibit that changes the relationship. The three countries will never be present in the same regression, and this is due to the nature of the wealth index and that it is not comparable between countries. With the aggregated country regression, time dummies will be included to get rid of possible correlated error terms between the waves.

One assumption for OLS is that there is a normal distribution for the dependent variable, the wealth index. By using a "normality plot" it is seen that the wealth index does not have normality (appendix A, figure 4). It is further tested with a "sktest" for normality on both an aggregated level and divided by country. However, this test rejects normality on a 1% level for all divisions (appendix A, table 17). Due to the negative values of the wealth index, it is not possible to impose normality by making it logarithmic. However, this does not have to be a big problem. The least-squares estimator is approximately normally distributed when the sample is large "enough" which is here. That the sample is large "enough" often means the number of observations – number of variables = 50. The estimates, therefore, may still be unbiased with minimum variance (Carter Hill, Griffiths, and Lim, 2012, pp 177-178), although with the caveat of distorted p-values. Thus, to be on the safe side I have set the significance level at 5% instead of 10% that is often the choice (Hubbard, 1978). Robust standard errors will be used for all regressions to account for possible heteroscedasticity (Carter Hill et al., 2012, pp 309). One possible problem with endogeneity may come from the omitted variables I mentioned earlier. The exclusion of these variables may interfere with the dependent variable and bias the coefficient. Simultaneity can also be a problem if, for example, wealth and income diversification are decided simultaneously (Longhi & Nandi, 2017). Nevertheless, this should not be a big problem because wealth is an index based on many different physical assets. The chance still exists if for example a natural disaster destroys the house and at the same time push the household out of subsistence agriculture as a consequence of a lost harvest. Finally, to claim causality is hard because as previously stated among other reasons it is not possible to follow an individual over time in this data and this is why the findings will be pure correlations.

3.4.2 Specification

Because of the comparative nature of the study and that the wealth index is not comparable between countries, the specifications will always be performed in three sets divided by country, and further by survey year. The general equation is as follows and will be performed once for non-farm, and once for on-farm diversification:

$$Wealth_{itc} = \beta_1 + \delta * Diversifier_{itc} + \beta_2 * Age_{itc} + \beta_3 * HHsize_{itc} + \beta_4 * Education_{itc} + \beta_5 * Employment_{itc} + \theta_{tc} + \epsilon_{itc}$$

To repeat, diversifier is a dummy variable either according to the non-farm or on-farm definition, they are never included at the same time. $\delta = 1$ if it is a household diversifier, otherwise 0. Age is either for the woman or her partner, neither included at the same time. HHsize is the number of people living in the household minus children under the age of five. Education is either educational achievement by the female or male or the females' literacy, but only one of these is included in a regression simultaneously. Employment is if the female is employed occasionally, seasonally, or all year since the data only exist for females. Θ is a year dummy and only included in the aggregate regression. ϵ is the error term. Finally, the indexes are for individual, time, and country.

To repeat, my first research question is "*Is there a positive relationship between household income diversification and an increased amount of wealth relative to solely engage in farming among the rural population in SSA?*". If it is not falsified in the descriptive analysis I will move on to the second question "*Does income diversification to the non-farm sector have a more positive relationship with wealth compared to diversification to the on-farm sector? If so, which extra job in the non-farm sector seems to have the highest positive relation with wealth?*" and try to interpret this finding descriptively. To further strengthen the finding of a possible relationship in question one, I will analyze with econometrics and test the null-hypothesis that *Household income diversification does not correlate with wealth.* At last, I will conclude with a broader discussion of the findings.

4. Results

The layout of this chapter is as follows: first, a summary of the variables and their correlations will be displayed and discussed where needed. After, I start to analyze the relationship between household income diversification and wealth descriptively, trying to find an indication if this exists and what sign it carries. Second, if some type of relation is found I investigate my second question descriptively; which type of diversification, non-farm, or on-farm seems to be the most highly correlated with welfare, and which particular extra jobs? To strengthen the findings with regards to the first question, I perform econometric analysis, including two robustness checks where I change the educational variables and age. My second question will not be analyzed econometrically due to the lack of data. To finalize, there will be a discussion about the findings to connect them back to the respective country's development path, a more analytical part of the thesis.

4.1 Descriptive Statistics
Table 1: Summary statistics

Variable	Obs	Mean	Std.Dev.	Min	Max
General					
Wealth index	11095	-.42	.43	-1.32	4.16
Non-farm diversifiers	11095	.36	.48	0	1
On-farm diversifiers	11095	.002	.05	0	1
Household size above the age of 5	11095	3.87	1.62	1	7
Female age	11095	30.98	8.56	15	49
Partner age	9851	36.52	10.50	15	97
Employment contract	11086	2.24	.56	1	3
Partner education	10966	1.21	1.10	0	5
Female education	11095	.87	.89	0	5
Literacy	11081	.84	.95	0	2
Ethiopia					
Wealth index	2031	-.581	.21	-.98	1.91
Non-farm diversifiers	2031	.28	.45	0	1
On-farm diversifiers	2031	0	0	0	0
Household size above the age of 5	2031	4.18	1.64	1	7
Female age	2031	31.75	8.75	15	49
Partner age	1661	40.69	11.36	19	95
Employment contract	2028	2.29	.61	1	3
Partner education	2006	.37	.68	0	5
Female education	2031	.2	.49	0	3
Literacy	2031	.20	.55	0	2
Ghana					
Wealth index	623	-.55	.38	-1.32	1.43
Non-farm diversifiers	623	.40	.49	0	1
On-farm diversifiers	623	.01	.08	0	1
Household size above	623	3.64	1.67	1	7

the age of 5					
Female age	623	33.74	8.08	15	49
Partner age	529	40.63	11.24	19	97
Employment contract	623	2.91	.30	1	3
Partner education	571	2.16	1.39	0	5
Female education	623	1.29	1.28	0	4
Literacy	623	.38	.73	0	2
Malawi					
Wealth index	8441	-.38	.46	-1.08	4.16
Non-farm diversifiers	8441	.37	.48	0	1
On-farm diversifiers	8441	.003	.05	0	1
Household size above the age of 5	8441	3.82	1.61	1	7
Female age	8441	30.58	8.49	15	49
Partner age	7661	35.34	9.93	15	95
Employment contract	8435	2.17	.53	1	3
Partner education	8389	1.35	1.04	0	5
Female education	8441	.997	.85	0	5
Literacy	8427	1.02	.96	0	2

In table 1 are the variables that are important for the analysis, in the first panel aggregated and in the following panels divided by country, which is the main interest. As mentioned in the methodology chapter, the wealth index's lowest point is negative and the highest positive. This index should neither be compared between countries since it is relative within one country (ICF, 2020b). The categorical variables are coded with numbers and on display here to give an idea of the spread. The female employment contract is coded with one, two, and three in the order of working occasionally, seasonally, or all year. Here we can see that most women work on seasonal contracts. For educational achievement, the higher the number is, the more educated the person is with zero being no education, three being incomplete secondary, and five being higher education. The variable shows that the males tend to be more educated than the females in the sample, but not by much considering the standard deviation. For female literacy, it is zero if the woman cannot read, one if she reads with difficulties or part of the sentence, and two if she reads easily or the whole sentence. A peculiar sight may be the age of the partner, which is quite high considering the maximum female age of 49. The low average life length makes it possible to identify as an outlier, but since it is likely to reflect reality, these observations stay.

If looking closer at the countries divided, Ghana has comparably few observations while Malawi has many. The low number of observations for Ghana may compromise the results, decreasing the efficiency of the estimations given that each year of the survey will be analyzed separately, reducing the number of observations further in each analysis. It may simply be a lack of data, but it can also give interesting insights into the characteristics of the subsample, and the situation in Ghana. The number of observations compared to the approximate population in the regions is 5,174,094 (2015/2016) to 2031 for Tigray in Ethiopia (Central Statistical Agency Ethiopia, 2018). 2,165,241 (2019) for the Western Region in Ghana to 623 (Ghana Statistical Service, 2019). Lastly, 7,526,160 (2018) in the Central Region in Malawi to 8441 (National Statistical Office Malawi, 2018). Nevertheless, the observations here are only

for the studied rural population and aggregated all years together, whereas the population numbers are for the whole regions.

Further, Ghana has a higher prevalence of non-farm income diversification relative to the others, but also with a higher standard deviation which can be due to both it being a smaller subsample, or that the variation is higher. For on-farm diversifiers, Ethiopia lacks households with this characteristic, and for the other two countries, a very small share does this. The consequence of this is that on-farm diversification will only be discussed for Ghana and Malawi throughout the thesis. The education of males and females seem to be much higher in Ghana. There is still a higher standard deviation, but it is interesting in its own right. It suggests that there is a bigger spread in educational achievements in Ghana than in other countries. At the same time, it seems that literacy is higher on average in Malawi even though Ghana has a substantially higher level of educational achievement. For the employment contract, it appears that in Ghana, it is more likely to be employed in all-year contract work, and this variable has a lower standard deviation than in Ethiopia and Malawi. Malawi seems to have the youngest subsample; both the female and male average ages are lower than for the other two countries. All these points, and in particular education and employment contract already give a small indication that there may be a difference in the development paths of the three countries. Richer and more developed countries tend to have higher educational standards and to be less dependent on farming and harvest seasons. Even if this is a small subsample, it may give an indicative picture of the settings. To conclude, most of the averages for Ethiopia are in between the other two countries as one may suspect. As a poor but rapidly growing country, it lands between poor and stagnant Malawi and the richer, growing Ghana.

Table 2: Correlation matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ethiopia										
(1) Wealth	1.00									
(2) Non-farm	0.18	1.00								
(3) On-farm	.	.	.							
(4) Household size	-0.02	-0.12	.	1.00						
(5) Female age	0.03	-0.06	.	0.53	1.00					
(6) Partner age	-0.01	-0.09	.	0.43	0.84	1.00				
(7) Employment	0.03	-0.09	.	-0.07	0.002	0.04	1.00			
(8) Partner education	0.16	0.18	.	-0.09	-0.20	-0.25	-0.07	1.00		
(9) Female education	0.14	0.14	.	-0.15	-0.27	-0.30	-0.12	0.33	1.00	
(10) Literacy	0.14	0.11	.	-0.11	-0.23	-0.25	-0.10	0.30	0.74	1.00
Ghana	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Wealth	1.00									
(2) Non-farm	0.15	1.00								
(3) On-farm	0.02	-0.05	1.00							
(4) Household size	0.05	-0.12	-0.07	1.00						
(5) Female age	0.14	-0.03	-0.08	0.39	1.00					

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(6) Partner age	0.11	-0.01	-0.03	0.36	0.77	1.00				
(7) Employment	0.01	-0.11	0.02	0.05	0.04	0.07	1.00			
(8) Partner education	0.12	0.20	0.04	-0.10	-0.01	-0.05	-0.09	1.00		
(9) Female education	0.18	0.20	-0.02	-0.15	-0.12	-0.12	-0.11	0.27	1.00	
(10) Literacy	0.12	0.14	-0.03	-0.05	-0.03	-0.02	-0.05	0.19	0.64	1.00
Malawi	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Wealth	1.00									
(2) Non-farm	0.16	1.00								
(3) On-farm	-0.01	-0.04	1.00							
(4) Household size	0.14	-0.008	-0.01	1.00						
(5) Female age	0.10	-0.004	-0.01	0.56	1.00					
(6) Partner age	0.11	0.01	-0.01	0.51	0.89	1.00				
(7) Employment	0.01	-0.08	-0.00	0.02	0.04	0.04	1.00			
(8) Partner education	0.28	0.14	0.003	-0.09	-0.16	-0.14	-0.03	1.00		
(9) Female education	0.21	0.10	-0.00	-0.15	-0.23	-0.23	-0.02	0.41	1.00	
(10) Literacy	0.18	0.06	0.01	-0.10	-0.19	-0.20	-0.03	0.29	0.58	1.00

In table 2, the correlations are split by country since this is the highest aggregation that will be used to perform the regressions, with the variable names shortened. The same story reveals in all three subsamples, there seem to be no big issues that may affect the study, but two common tendencies. First, the ages of the males and females correlate 77% to 89% and this can certainly be problematic although it is a mirror of reality. However, the two variables will not be in the same regression. In the main regression, only the female age will be used, and first, in a robustness check the male age will be included instead. A second reason to omit the male age in the main regressions is the lower amount of observations (as seen in table 1). Second, intuitively female literacy has a high correlation with female education. It varies from 58% to 74%, but these two variables will neither be used in the same regression. The female educational achievement will be switched for literacy in a robustness check. Concluded from this, multicollinearity should not be an issue.
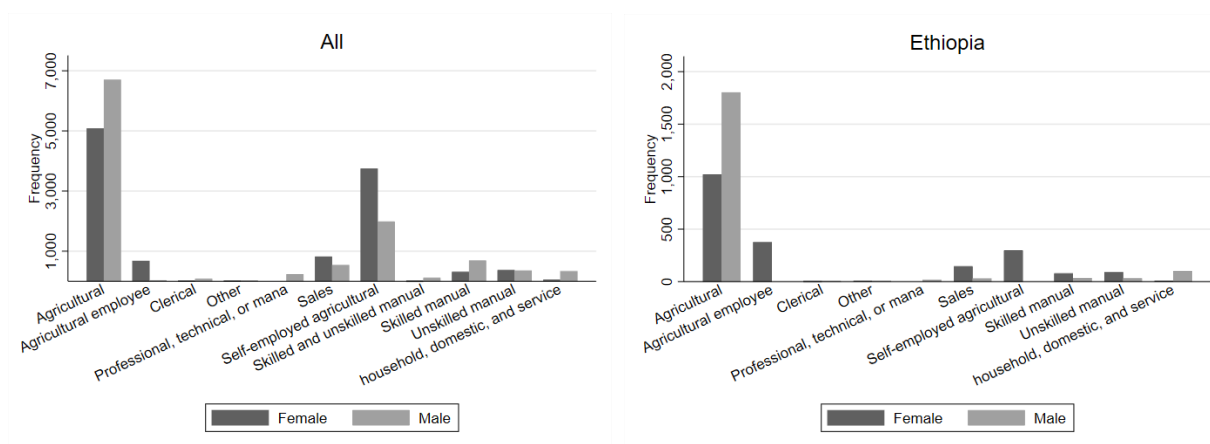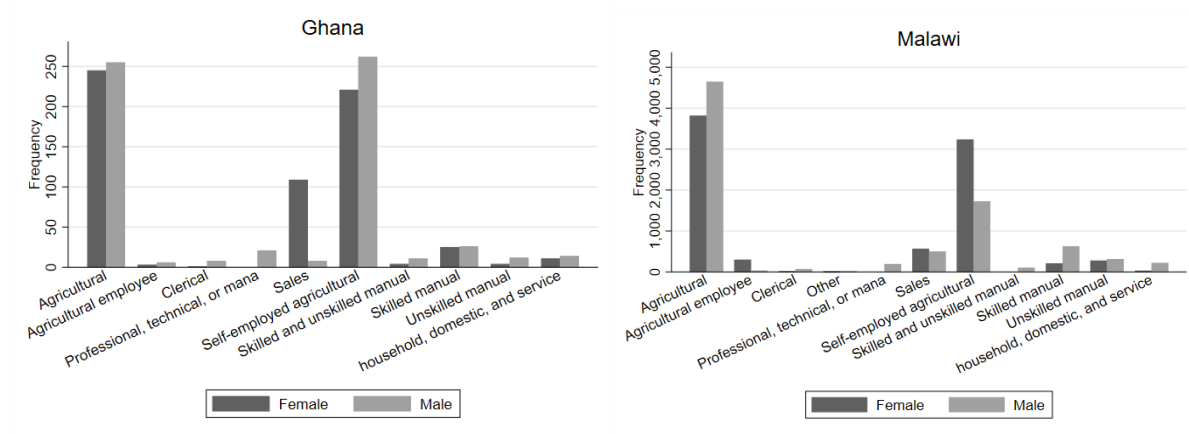
Figure 2: Division of labor between occupations

For a picture of the division between the different occupations for the sample, see figure 2. Most people work in the three agricultural sectors. There are somewhat more men working in the general agricultural sector than women, though this is reversed when it comes to those who are (for certain) self-employed in the agricultural sector. The number of agricultural employees is very small. Another interesting indicator is that females are overrepresented in sales, which can be a very diverse sector. It probably involves everything from working in a call center to standing by the street or going to the local market selling crops. There seems to be an overrepresentation of men among the occupations that presumably are paid higher and for which higher skills are needed such as professional, technical, or managerial or skilled manual labor. At last, the occupational categories skilled and unskilled manual, skilled manual, and unskilled manual are divisions made by the DHS and that is why these categories are not very well defined and cannot be narrowed down further.

For the figures dividing the occupations by country, note the differences in scale on the y-axis. From left to right it is Ethiopia, Ghana, and Malawi. The vast majority of the population in all countries tend to work within agriculture, and for Ghana, self-employed agriculture, which is working on the own farm, is as big as the general occupation group agriculture. The numbers of Malawi lie between Ethiopia and Ghana when it comes to self-employed agriculture, and what is interesting is that there is double the number of females in this sector compared to males. To see the division by country and year, see appendix B, table figure 5.

4.2 Descriptive Analysis

Before I do an econometric analysis, a descriptive analysis is performed. This is to establish indicative patterns for a possible relationship between the prevalence of household income diversification and wealth in comparison to only engage in agriculture in these three rural regions. After the relationship is established, I evaluate what type of diversification, and which jobs that indicate to have a more positive relationship to wealth.

4.2.1 Income Diversification and Welfare

Table 3: Household income diversification and wealth

Ethiopia	Non-farm diversification		
	0	1	Total
poorest	479	138	617
	77.63	22.37	100.00
poorer	419	146	565
	74.16	25.84	100.00
middle	339	144	483
	70.19	29.81	100.00
richer	205	104	309
	66.34	33.66	100.00
richest	27	30	57
	47.37	52.63	100.00
Total	1469	562	2031
	72.33	27.67	100.00

Ghana	Non-farm diversification			On-farm diversification		
	0	1	Total	0	1	Total
poorest	95	55	150	149	1	150
	63.33	36.67	100.00	99.33	0.67	100.00
poorer	178	97	275	272	3	275
	64.73	35.27	100.00	98.91	1.09	100.00
middle	91	80	171	171	0	171
	53.22	46.78	100.00	100.00	0.00	100.00
richer	6	16	22	22	0	22
	27.27	72.73	100.00	100.00	0.00	100.00
richest	2	3	5	5	0	5
	40.00	60.00	100.00	100.00	0.00	100.00
Total	372	251	623	619	4	623
	59.71	40.29	100.00	99.36	0.64	100.00

Malawi	Non-farm diversification			On-farm diversification		
	0	1	Total	0	1	Total
poorest	1801	796	2597	2590	7	2597
	69.35	30.65	100.00	99.73	0.27	100.00
poorer	1518	826	2344	2336	8	2344
	64.76	35.24	100.00	99.66	0.34	100.00
middle	1137	710	1847	1843	4	1847
	61.56	38.44	100.00	99.78	0.22	100.00
richer	697	536	1233	1231	2	1233
	56.53	43.47	100.00	99.84	0.16	100.00
richest	164	256	420	419	1	420
	39.05	60.95	100.00	99.76	0.24	100.00
Total	5317	3124	8441	8419	22	8441
	62.99	37.01	100.00	99.74	0.26	100.00

Table 3 divides diversifiers and non-diversifiers by country and wealth quintiles. By examining the percentage of non-farm diversifying households in the different wealth quintiles, it is clear that there is a high correlation between diversification and being in the top 40% of the wealth distribution. On the other end, the highest prevalence of non-diversifiers is in the bottom 40%

of the wealth distribution. For on-farm diversification, the few that do this have a higher correlation with the bottom quintiles. The caveat is that the tabulations are without using the time perspective and thus four survey years are summarized. This will be developed in the next figure.
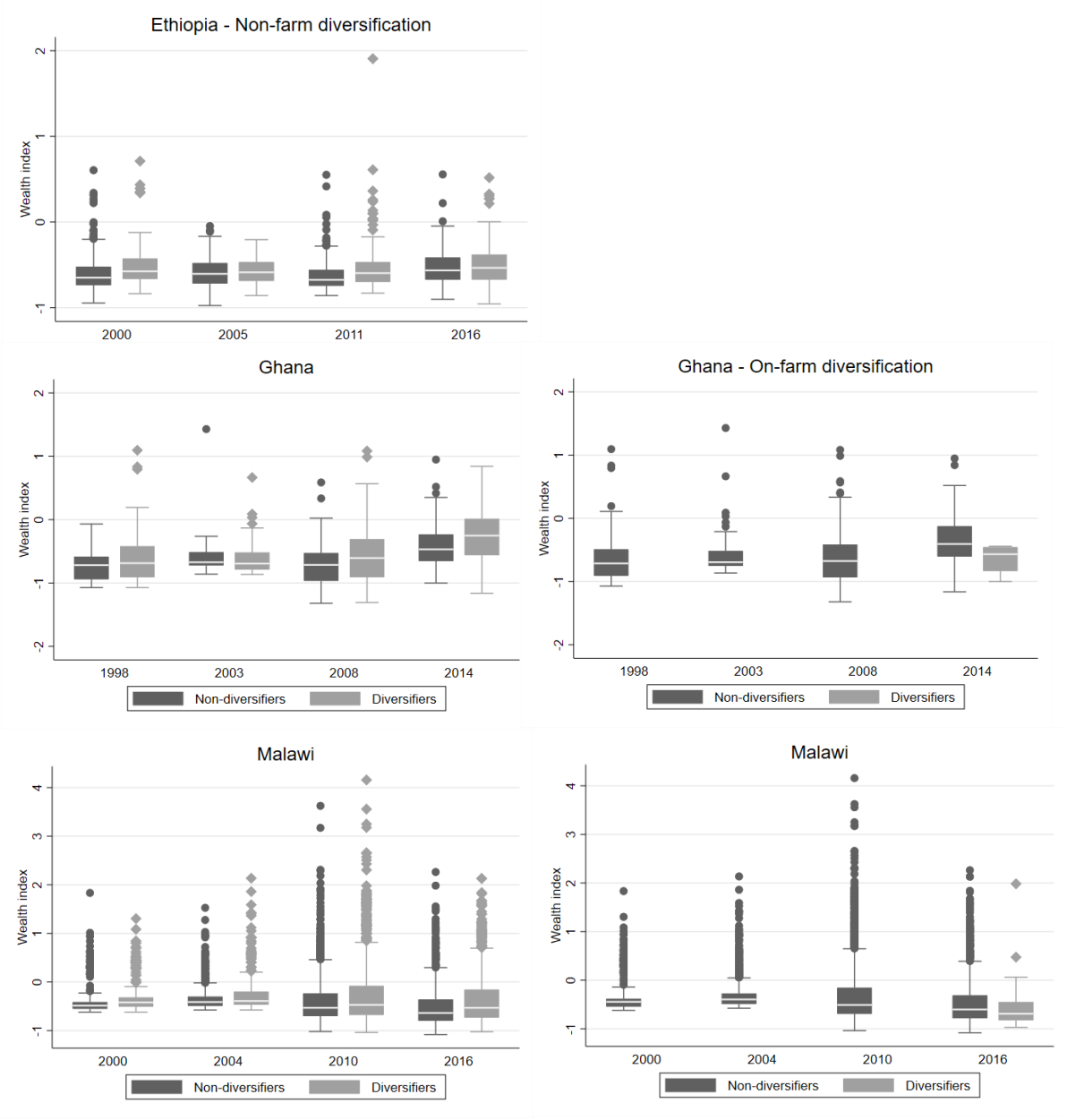
Figure 3: Boxplot distribution of wealth for diversifiers and non-diversifiers

In figure 3, the countries are divided by survey year and by the distribution of the households according to wealth. The diamonds are the households diversifying whereas the circles are the households not diversifying. This is according to the non-farm definition in the three panels to the left and the on-farm definition in the two panels to the right. In 11 out of the 12 subsamples for non-farm, the diversifying households exist in higher parts of the wealth distribution than those who do not diversify. The exception is Ghana 2003. For the on-farm definition, no

households are diversifying their income in Ethiopia and few for the other countries, as previously mentioned. For Ghana (2014) and Malawi (2016), the households that are not diversifying are higher in the wealth distribution than the ones diversifying.

There is an outlier in the subsample 2011 from Ethiopia, a non-farm diversifier. This observation may compromise the boxplots and bias of the distribution. Nevertheless, when removing it, the findings do not change (appendix b, figure 6). This split between countries and years further adds to the insight that there may be a positive correlation between household income diversification and welfare for the non-farm definition, and negative for the on-farm diversifiers.

4.2.2 Diversification, Extra-Jobs, and Welfare
Table 4: Non-farm diversifiers' contra on-farm diversifiers by wealth

Ghana 2014

Non-farm div.	Freq.	Percent	Cum.	*On-farm div.*	Freq.	Percent	Cum.
poorest	55	21.91	21.91	poorest	1	25.00	25.00
poorer	97	38.65	60.56	poorer	3	75.00	100.00
middle	80	31.87	92.43				
richer	16	6.37	98.80				
richest	3	1.20	100.00				

Malawi 2016

Non-farm div.	Freq.	Percent	Cum.	*On-farm div.*	Freq.	Percent	Cum.
poorest	796	25.48	25.48	poorest	7	31.82	31.82
poorer	826	26.44	51.92	poorer	8	36.36	68.18
middle	710	22.73	74.65	middle	4	18.18	86.36
richer	536	17.16	91.81	richer	2	9.09	95.45
richest	256	8.19	100.00	richest	1	4.55	100.00

To be able to reason about my second question, which type of diversification that seems to correlate with higher wealth, non-farm or on-farm, a comparison between these two is made in table 4. The left panels are for non-farm diversification and the right panels for on-farm diversification. Since on-farm only exist for one survey per country, the indications here are insecure. Nevertheless, based on this comparison, it does seem for both countries that the households diversifying completely out of agriculture seem to have a higher correlation with wealth, thus, the non-farm diversifiers. For Ghana, there are no on-farm diversifiers among the

60% richest, and for Malawi, the percentage of on-farm diversifiers is much lower for the top 40% compared to for the non-farm diversifiers. Since non-farm diversification tends to correlate with higher wealth, it becomes interesting to know which specific jobs that are the most beneficial in this sample and seem to correlate the most positive with wealth.

Table 5: Wealth and the relationship with extra jobs among non-farm diversifiers

Ethiopia	Household wealth index					
	poorest	poorer	middle	richer	richest	Total
Professional, technical, or managerial	4	6	4	1	4	19
	21.05	31.58	21.05	5.26	21.05	100.00
Clerical	0	1	2	0	0	3
	0.00	33.33	66.67	0.00	0.00	100.00
Sales	37	47	40	38	14	176
	21.02	26.70	22.73	21.59	7.95	100.00
Household, domestic, and services	19	28	35	27	3	112
	16.96	25.00	31.25	24.11	2.68	100.00
Skilled manual	29	23	29	24	9	114
	25.44	20.18	25.44	21.05	7.89	100.00
Unskilled manual	41	38	29	14	0	122
	33.61	31.15	23.77	11.48	0.00	100.00
Other	8	3	5	0	0	16
	50.00	18.75	31.25	0.00	0.00	100.00
Total	138	146	144	104	30	562
	24.56	25.98	25.62	18.51	5.34	100.00
Ghana	poorest	poorer	middle	richer	richest	Total
Professional, technical, or managerial	1	12	6	1	1	21
	4.76	57.14	28.57	4.76	4.76	100.00
Clerical	1	1	4	3	0	9
	11.11	11.11	44.44	33.33	0.00	100.00
Sales	28	49	35	4	1	117
	23.93	41.88	29.91	3.42	0.85	100.00
Household, domestic, and services	5	10	6	3	1	25
	20.00	40.00	24.00	12.00	4.00	100.00
Skilled and unskilled manual	5	3	6	1	0	15
	33.33	20.00	40.00	6.67	0.00	100.00
Skilled manual	14	17	15	4	0	50
	28.00	34.00	30.00	8.00	0.00	100.00
Unskilled manual	1	5	8	0	0	14
	7.14	35.71	57.14	0.00	0.00	100.00
Total	55	97	80	16	3	251
	21.91	38.65	31.87	6.37	1.20	100.00
Malawi	poorest	poorer	middle	richer	richest	Total
Professional, technical, or managerial	23	28	28	39	76	194
	11.86	14.43	14.43	20.10	39.18	100.00
Clerical	16	16	17	18	10	77
	20.78	20.78	22.08	23.38	12.99	100.00
Sales	247	274	264	204	79	1068
	23.13	25.66	24.72	19.10	7.40	100.00
Household, domestic, and services	67	61	50	46	24	248
	27.02	24.60	20.16	18.55	9.68	100.00
Skilled and unskilled manual	24	32	25	20	2	103
	23.30	31.07	24.27	19.42	1.94	100.00
Skilled manual	224	225	207	140	40	836
	26.79	26.91	24.76	16.75	4.78	100.00
Unskilled manual	187	189	118	69	23	586

	31.91	32.25	20.14	11.77	3.92	100.00
Other	8	1	1	0	2	12
	66.67	8.33	8.33	0.00	16.67	100.00
Total	796	826	710	536	256	3124
	25.48	26.44	22.73	17.16	8.19	100.00

When investigating how the distribution of wealth looks over the extra job next to farming in table 5, the expected pattern appears. It is easy to assume that the highest-paid occupations are professional, technical, or managerial, or the occupation skilled manual worker since these types of labor normally require the highest set of skills. In all countries, the highest share of any work among the richest quintile is well-paid occupations. The incidence of skilled manual labor is not as clear, however. Another finding is that in Ethiopia and Malawi, the biggest share of the unskilled manual workers, a typical low-wage job, are among the poorest. To be clear, on-farm diversification is not included since it only includes agricultural wage labor.

4.2.3 Summary

To clarify the indications from the descriptive analysis, I will shortly summarize what was seen above. It is suggested that there is a positive correlation between non-farm household income diversification and wealth for all the countries at all years but one. Thus, there is no reason to falsify this statement, yet. For on-farm diversification, it seems to be the opposite, a negative correlation. The analysis further suggests that diversification to the non-farm sector tends to have a more positive correlation to wealth than if diversifying to the on-farm sector, with the caveat of a scarce number of observations. The more positive relationship with wealth is found from the extra jobs professional, technical, or managerial as can be assumed since it should require more skills. On the other side, for the poorest in Ethiopia and Malawi, there seems to be a strong correlation with unskilled manual workers.

The results so far are from simple eyeball regressions and there are surely other factors affecting the relationships. In the next step, an econometric analysis will be the task to further investigate the robustness of research question one with the inclusion of possibly important determinants as controls for the relation. I will also divide the countries by survey year, and in this way investigate if there are differences between them. A caveat, as previously stated, is that the data is too scarce to be able to perform a further investigation of question two, and thus the focus will be on examining the question if there is a relationship between household income diversification and welfare.

4.3 Econometric Analysis

To give strength to the indications above an econometric analysis is performed and the regressions will be divided by survey years. Throughout year dummies are included for the aggregated sample, but the coefficients are not reported. The interpretation of the regressions will only be regarding the sign of the finding, due to that the coefficient does not add any special information since the dependent variable is an index. The discussion on all the findings will be joint at the end of the chapter to draw further conclusions.

The r-squared is low throughout all regressions, but this does not have to be problematic. Most likely, there are important variables not included. For instance, parents' education, parents' income, and infrastructure (market access) are all mentioned in the methodology chapter. As explained, the wealth index starts negative, which gives a reason for the peculiar constant. Although not reported, for all regressions, the F-test on a 5% level (except for one specification at 10%) tests the null hypothesis of joint insignificance and rejects it.

4.3.1 Income Diversification and Welfare

Table 6: Non-farm diversification and wealth in Ethiopia

VARIABLES	(1) Wealth	(2) Wealth	(3) Wealth	(4) Wealth	(5) Wealth
Non-farm diversification	0.068***	0.072***	0.029	0.101***	0.045
	(0.011)	(0.021)	(0.019)	(0.023)	(0.028)
Female age	0.002***	0.002*	0.0005	0.001	0.007***
	(0.0006)	(0.001)	(0.001)	(0.001)	(0.002)
Household members over the age of 5	-0.006	-0.018**	0.003	-0.002	-0.009
	(0.003)	(0.006)	(0.006)	(0.005)	(0.011)
Female education	0.051***	0.085	0.029	0.025	0.093***
	(0.012)	(0.065)	(0.022)	(0.013)	(0.018)
Employment contract	0.021	-0.022	0.005	0.038	0.089**
	(0.012)	(0.017)	(0.018)	(0.023)	(0.033)
Constant	-0.704***	-0.55***	-0.634***	-0.744***	-0.939***
	(0.040)	(0.052)	(0.051)	(0.065)	(0.093)
Observations	2,028	708	352	665	303
Adjusted R-squared	0.059	0.049	-0.001	0.056	0.110
Year	All	2000	2005	2011	2016
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

In table 6, household income diversification to wealth is regressed for Ethiopia's non-farm diversifiers. Diversifiers are significant in three out of five regressions and positive in all. In a few subsamples, age and education seem to matter, with a positive correlation to wealth. Household size and employment contract only tend to matter in one subsample with the former having a negative sign. In the subsample from 2005, the adjusted r-square becomes negative, which should be interpreted as being zero, and thus the model, at this point, does not explain anything. For on-farm, no households are diversifying for Ethiopia, and thus, no econometric regressions are to report. Concluded, non-farm diversifying seems to increase wealth in Ethiopia compared to only engaging in agriculture, even when controlling for other determinants, but it depends on the subsample.

Table 7: Non-farm diversification and wealth in Ghana

VARIABLES	(1) Wealth	(2) Wealth	(3) Wealth	(4) Wealth	(5) Wealth

Non-farm diversification	0.112***	0.135*	0.015	0.189*	0.117*
	(0.031)	(0.064)	(0.047)	(0.086)	(0.055)
Female age	0.006**	0.006	0.004	-0.002	0.011**
	(0.002)	(0.004)	(0.003)	(0.005)	(0.003)
Household members over the age of 5	-0.002	0.005	-0.0003	0.036	-0.034
	(0.010)	(0.021)	(0.015)	(0.027)	(0.018)
Female education	0.030*	0.037	0.018	0.043	0.039
	(0.012)	(0.021)	(0.020)	(0.037)	(0.022)
Employment contract	0.057	0.025	0.102*	-0.017	0.045
	(0.037)	(0.073)	(0.041)	(0.120)	(0.064)
Constant	-1.076***	-1.020***	-1.060***	-0.786	-0.853***
	(0.129)	(0.283)	(0.154)	(0.409)	(0.238)
Observations	623	153	157	134	179
Adjusted R-squared	0.155	0.041	0.009	0.031	0.083
Year	All	1998	2003	2008	2014
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

In table 7 for Ghana, four out of five regressions have non-farm diversification statistically significant, and with a positive sign. The female age is statistically significant in two subsamples with a positive relation, but the household size is not statistically significant at any point. Education and employment contracts are only significant in one subsample each, both positive. For on-farm, there is only data for 2014, and it is statistically insignificant and negative in its correlation with wealth. For the other controls, there are no big changes (appendix B, table 18). Concluded, household non-farm income diversification seems to have a strong positive correlation with wealth in Ghana even when controlling for other determinants.

Table 8: Non-farm diversification and wealth in Malawi

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Wealth	Wealth	Wealth	Wealth	Wealth
Non-farm diversification	0.111***	0.090***	0.077***	0.100***	0.167***
	(0.011)	(0.015)	(0.019)	(0.019)	(0.022)
Female age	0.004***	0.0004	0.002	0.004**	0.008***
	(0.0007)	(0.0009)	(0.001)	(0.001)	(0.001)
Household members over the age of 5	0.040***	0.019***	0.022***	0.051***	0.050***
	(0.004)	(0.005)	(0.006)	(0.007)	(0.008)
Female education	0.142***	0.057***	0.092***	0.186***	0.146***
	(0.008)	(0.012)	(0.014)	(0.016)	(0.012)
Employment contract	0.033***	0.008	0.013	0.052**	0.025
	(0.009)	(0.013)	(0.017)	(0.017)	(0.018)
Constant	-0.883***	-0.575***	-0.581***	-1.006***	-1.167***
	(0.031)	(0.040)	(0.051)	(0.060)	(0.056)
Observations	8,435	1,444	1,417	3,177	2,397

Adjusted R-squared	0.116	0.060	0.086	0.104	0.146
Year	All	2000	2004	2010	2016
Year FE	Yes	No	No	No	No

For Malawi, in table 8, non-farm diversification is statistically significant in its positive relationship to wealth in all specifications. The same is true for household size and women's education. The females' age is positive and statistically significant in three of the five subsamples and employment contracts in two subsamples with a positive relation. For on-farm, there is only data for 2016, where diversifiers are statistically insignificant. The other variables are the same (appendix B, table 19). In conclusion, non-farm diversification seems to have a positive correlation with wealth in Malawi even when controlling for other determinants, and the results are highly valid no matter the subsample.

4.3.2 Robustness Checks

For the two robustness checks, I first change the age and education of the female for the equivalent observations for males. After this, I exchange the female education variable from measuring educational achievement to literacy and perform the same regressions again.

4.3.2.2 Male Age and Education

Table 9: Non-farm diversification and wealth in Ethiopia – Male education and age

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Wealth	Wealth	Wealth	Wealth	Wealth
Non-farm diversification	0.061***	0.063**	0.042	0.085***	0.043
	(0.012)	(0.022)	(0.025)	(0.022)	(0.028)
Age of partner	0.0007	0.001	-0.001	-0.0004	0.003*
	(0.0005)	(0.001)	(0.001)	(0.0008)	(0.001)
Household members over the age of 5	-0.001	-0.008	0.002	0.004	-0.0001
	(0.004)	(0.006)	(0.007)	(0.006)	(0.01)
Partner education	0.037***	0.040**	0.014	0.024	0.058***
	(0.008)	(0.015)	(0.019)	(0.015)	(0.016)
Employment contract	0.037***	0.0009	0.020	0.050*	0.095**
	(0.011)	(0.014)	(0.027)	(0.020)	(0.035)
Constant	-0.748***	-0.646***	-0.606***	-0.742***	-0.865***
	(0.038)	(0.058)	(0.070)	(0.055)	(0.093)
Observations	1,646	566	269	510	301
Adjusted R-squared	0.069	0.027	-0.000	0.062	0.078
Year	All	2000	2005	2011	2016
Year FE	Yes	No	No	No	No

When using age and educational achievement of the partner, the relationship between diversification and wealth in table 9 remains the same. However, age as a determinant has lost

its significance in two subsamples and is now only significant in one. The education of the partner has gained significance and household size has lost its significance. Employment contract seems to matter more now from only being significant in one subsample.

Table 10: Non-farm diversification and wealth in Ghana – Male education and age

VARIABLES	(1) Wealth	(2) Wealth	(3) Wealth	(4) Wealth	(5) Wealth
Non-farm diversification	0.078*	0.101	-0.009	0.101	0.101
	(0.036)	(0.084)	(0.054)	(0.085)	(0.059)
Age of partner	0.003	0.004	0.002	-0.001	0.003
	(0.002)	(0.004)	(0.002)	(0.004)	(0.003)
Household members over the age of 5	0.00874	0.004	0.001	0.047	-0.012
	(0.012)	(0.031)	(0.019)	(0.025)	(0.021)
Partner education	0.052***	0.030	0.027	0.111***	0.039
	(0.013)	(0.029)	(0.023)	(0.026)	(0.021)
Employment contract	0.073	0.075	0.091	0.070	0.050
	(0.044)	(0.102)	(0.050)	(0.125)	(0.079)
Constant	-1.131***	-1.136**	-1.017***	-1.280**	-0.717*
	(0.159)	(0.339)	(0.202)	(0.441)	(0.294)
Observations	486	118	115	109	144
Adjusted R-squared	0.164	0.014	-0.007	0.139	0.033
Year	All	1998	2003	2008	2014
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

In table 10 for Ghana, the relationship between diversification and wealth is barely significant anymore. Age has lost its significance whereas education has gained significance in another subsample. The employment contract has lost its only significant subsample. For on-farm, the relationship between diversification and wealth only exists for 2014 and have become significant and negative, which is very interesting. Other than that, there are no big changes (appendix B, table 20).

Table 11: Non-farm diversification and wealth in Malawi – Male education and age

VARIABLES	(1) Wealth	(2) Wealth	(3) Wealth	(4) Wealth	(5) Wealth
Non-farm diversification	0.106***	0.093***	0.063**	0.108***	0.142***
	(0.011)	(0.016)	(0.019)	(0.020)	(0.022)
Age of partner	0.004***	0.001	0.002*	0.005***	0.006***
	(0.0006)	(0.0008)	(0.0010)	(0.001)	(0.001)
Household members over the age of 5	0.037***	0.015**	0.020**	0.046***	0.050***
	(0.004)	(0.005)	(0.006)	(0.008)	(0.007)
Partner education	0.136***	0.060***	0.098***	0.204***	0.129***
	(0.007)	(0.011)	(0.012)	(0.016)	(0.009)
Employment contract	0.033***	0.006	0.029	0.042*	0.036*
	(0.010)	(0.014)	(0.018)	(0.018)	(0.019)
Constant	-0.931***	-0.612***	-0.653***	-1.055***	-1.189***
	(0.031)	(0.041)	(0.055)	(0.063)	(0.054)
Observations	7,624	1,282	1,262	2,703	2,377
Adjusted R-squared	0.155	0.086	0.141	0.149	0.169

Year	All	2000	2004	2010	2016
Year FE	Yes	No	No	No	No

For Malawi in table 11, household income diversification and wealth still has a statistical significance in all subsamples whereas age and employment contract has gained significance in one more. Everything else is the same. For on-farm, diversification has no significance, and the other variables are more or less the same (appendix B, table 21).

4.3.2.2 Literacy
Table 12: Non-farm diversification and wealth in Ethiopia - Literacy

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Wealth	Wealth	Wealth	Wealth	Wealth
Non-farm diversification	0.069***	0.074***	0.030	0.102***	0.054
	(0.012)	(0.021)	(0.019)	(0.023)	(0.028)
Female age	0.002***	0.002*	0.0005	0.0008	0.006***
	(0.0006)	(0.0010)	(0.001)	(0.001)	(0.002)
Household members over the age of 5	-0.007*	-0.017**	0.002	-0.002	-0.009
	(0.003)	(0.006)	(0.006)	(0.005)	(0.010)
Literacy	0.045***	0.083*	0.029	0.008	0.096***
	(0.010)	(0.040)	(0.019)	(0.011)	(0.021)
Employment contract	0.021	-0.024	0.005	0.040	0.088**
	(0.012)	(0.018)	(0.018)	(0.023)	(0.032)
Constant	-0.701***	-0.547***	-0.634***	-0.731***	-0.910***
	(0.041)	(0.056)	(0.052)	(0.065)	(0.090)
Observations	2,028	708	352	665	303
Adjusted R-squared	0.059	0.053	0.000	0.053	0.127
Year	All	2000	2005	2011	2016
Year FE	Yes	No	No	No	No

When using female literacy as opposed to educational attainment as seen in table 12, diversification stays the same as in the original regression table. The only other interesting change is that literacy is significant in three out of five subsamples instead of only two. For on-farm, there are no observations for Ethiopia. Thus, with the use of literacy, the relationship between diversification and wealth stays the same.

Table 13: Non-farm diversification and wealth in Ghana - Literacy

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Wealth	Wealth	Wealth	Wealth	Wealth
Non-farm diversification	0.114***	0.136*	0.018	0.194*	0.122*
	(0.031)	(0.064)	(0.046)	(0.085)	(0.056)
Female age	0.005**	0.006	0.004	-0.002	0.010**
	(0.002)	(0.004)	(0.003)	(0.005)	(0.003)
Household members over the age of 5	-0.002	0.003	0.0002	0.029	-0.031
	(0.010)	(0.021)	(0.015)	(0.024)	(0.018)
Literacy	0.059**	0.064*	0.028	0.059	0.088

	(0.023)	(0.030)	(0.053)	(0.057)	(0.053)
Employment contract	0.057	0.027	0.097*	-0.018	0.053
	(0.037)	(0.069)	(0.042)	(0.125)	(0.066)
Constant	-1.057***	-1.019***	-1.019***	-0.697	-0.819***
	(0.128)	(0.273)	(0.142)	(0.425)	(0.239)
Observations	623	153	157	134	179
Adjusted R-squared	0.158	0.050	0.007	0.026	0.089
Year	All	1998	2003	2008	2014
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

In table 13, nothing has changed in the relation between income diversifying households and wealth, nor for any other variable except literacy. It is now in comparison to education significant in one more subsample. For on-farm, diversification is significant for 2014 with a negative sign, as it did when including literacy. Nothing else has changed besides literacy that follows the same pattern as the non-farm diversification regressions here (appendix B, table 22).

Table 14: Non-farm diversification and wealth in Malawi - Literacy

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Wealth	Wealth	Wealth	Wealth	Wealth
Non-farm diversification	0.122***	0.092***	0.081***	0.112***	0.189***
	(0.011)	(0.015)	(0.019)	(0.019)	(0.022)
Female age	0.003***	0.0003	0.001	0.003*	0.007***
	(0.0007)	(0.0009)	(0.001)	(0.001)	(0.001)
Household members over the age of 5	0.037***	0.019***	0.020***	0.047***	0.045***
	(0.004)	(0.005)	(0.006)	(0.007)	(0.008)
Literacy	0.099***	0.035***	0.058***	0.147***	0.101***
	(0.005)	(0.008)	(0.009)	(0.010)	(0.010)
Employment contract	0.036***	0.008	0.011	0.068***	0.018
	(0.009)	(0.013)	(0.017)	(0.017)	(0.019)
Constant	-0.834***	-0.560***	-0.527***	-0.954***	-1.063***
	(0.030)	(0.039)	(0.049)	(0.056)	(0.056)
Observations	8,421	1,444	1,417	3,164	2,396
Adjusted R-squared	0.094	0.057	0.062	0.092	0.109
Year	All	2000	2004	2010	2016
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

For Malawi in table 14, nothing changes when using literacy, and the findings stay the same although the r-squared decreases. For on-farm, diversification loses its significance, and other than that, all the results are the same (appendix B, table 23).

The lessons we can draw from the econometric analysis are that for all three regions of the countries, the null-hypothesis stating *household income diversification does not correlate with wealth* is rejected in the majority of the cases. The positive relationship between income diversification and wealth seems to hold and more subsamples become significant in the robustness checks. When it comes to the statistical significance of the control variables, they

change a lot more and seem to be highly dependent on the setting. For on-farm diversifiers, for Malawi, diversification is not statistically significant, whereas for Ghana in both robustness checks, it is significant, and with a negative sign. Based on this quite simple econometric analysis, the results discussed after the descriptive analysis are still valid and further strengthened. Nevertheless, this is simply an analysis of correlations and no causation is warranted in any way. These results are now be discussed in detail before I start to conclude this thesis.

4.4 Discussion

I will discuss these differences in the findings of the separate countries and roughly compare them. Generally, there seems to exist a robust positive correlation between non-farm household diversification and wealth when comparing it to only working in agriculture in the rural commercialized area. This type of diversification, to non-agricultural sectors, also seems to correlate more positively with wealth compared to diversifying into on-farm labor, i.e. agricultural wage work on other farms. This may be explained by the previously discussed pull-factors. Households with high incomes are pulled into diversification to find higher remuneration (Haggblade et al., 2007, pp 115). Since what is studied is simply a correlation, the causality can go both ways, where one of them is from wealth to diversification, and then possibly this is one explanation. The occupation professional, technical, or managerial has the highest positive correlation with wealth as expected, but for skilled manual labor, it does not seem to be a clear case, which one would suspect.

There seem to be two different patterns depending on the type of diversification where non-farm diversification correlates with increased wealth and on-farm diversification vaguely with decreased wealth. This finding follows the discussion by Haggblade et al. (2007, pp 385) about the impact rural off-farm activity have on equity is typically mixed and completely dependent on the exact type of activities.

Starting with Ethiopia, there is a clear preponderance of unskilled labor in the poorest parts of the population, as expected. The relationship between non-farm diversification and wealth is positive at times, but it depends a lot on the specific subsample, 2000, and 2011, which means that this finding is not very robust over time. To find a positive correlation in Ethiopia 2011 although there was the "2011 East Africa drought" spreading food insecurity and famine in the region is peculiar. I see three potential reasons why this is the case. First, the region Tigray that is analyzed was only affected up to grade two on a five-grade scale, which means that food security was not affected or "just" stressed (OCHA, 2011). It can also be by a theoretical reason, that this drought created distress diversification, and that the push-factors pushed poor people into diversification, which positively affected their wealth in comparison to purely agricultural households. Although it may have been into bad paying sectors, it was still better than the part that lost their harvests and had nothing (Haggblade et al., 2007, pp 43; Bigsten & Tengstam, 2011). A third reason can simply be that the surveys took place before the drought commenced, and thus, the consequences of the drought are not reported. To conclude, age and education also tend to correlate positively with wealth, but it depends on the subsample.

For Ghana, non-farm household income diversification positively correlates with wealth generally, whereas age nor education seems to correlate with wealth typically. This can be because Ghana has a higher education than the other countries in general, and thus already an advantageous starting point. What is interesting about Ghana is the negative correlation found between on-farm diversification and wealth. It is negative throughout and becomes statistically significant in the two robustness checks. This may imply that in Ghana, on-farm diversification, i.e. agricultural wage labor on another farm, decrease wealth, or that wealth decreases your diversification to this type of job. This is expected if relating to the literature review, which states that this type of job often is informal with low returns (Collier & Lal, 1986, pp 128; Haggblade et al., 2007, pp 43; Bigsten & Tengstam, 2011). It can also be related to the idea that higher returns are not accessible for the poorest, and if they partake more in on-farm diversification, the relation to welfare from income diversification may be negative for some parts and positive for some, as the literature review suggests.

At last, for Malawi, for non-farm household income diversification, there is a clear and robust positive correlation with wealth, no matter the subsample. The same is true for household size, education, age, and employment contract that all seem to be positively correlating with wealth. Furthermore, as for Ethiopia, there is also an indication of a correlation between unskilled manual workers and being poor. Malawi as a poor and agriculturally reliant country, with the regional characteristic of being commercialized, is dependent on farming. Farming has a seasonal distribution leading to a seasonal labor market as discussed earlier. The consequence of this is underemployment when it is not harvesting season and shortages of labor when it is cropping season (IFAD, 2020c). With this in mind, it makes sense that the contracts and wealth have a high positive correlation. If a household manages to break from the seasonal dependent work on the fields, it is also expected to correlate with a more positive wealth. According to Monga and Yifu Lin (2015, pp 49), it is especially common for males in these countries to move from agriculture to among others, unskilled manual jobs as a part of the structural transformation. With a structural transformation, together with a premature de-industrialization (Rodrik, 2016), the consequence may be income diversification, if we assume that only half of the household leave agriculture. Besides this, education and wealth have a positive correlation. That this is not found in Ghana can be both due to their different development stages, but also simply a limitation of the data of Ghana. It can be that the wealthier population in the rural areas of Ethiopia and Malawi have a significantly different level of education relative to the poorer parts and that this shows up in the education level in the study. Thus, the causality may go from wealth, and to education. Finally, the negative relationship between on-farm diversification and wealth found in Ghana does not appear for Malawi. Can it be that in Malawi, being much poorer, it may still be better to diversify to on-farm instead of just doing subsistence agriculture? There is no statistically significant correlation of this, but economically speaking, these scarce regressions show positive signs.

To finalize the analysis, I am going to answer the research questions I stated at the beginning of the thesis:

1. *Is there a positive relationship between household income diversification and an increased amount of wealth relative to solely engage in farming among the rural population in SSA?*

This question has not been falsified by my analysis and the positive correlation between household income diversification and wealth stand strong. Therefore, I would say that yes, there is a positive relationship. The rural off-farm sector is not irrelevant, as many have believed for a long time (Haggblade et al., 2007, pp 281). With this in mind, moving on to the second research question:

2. *Does income diversification to the non-farm sector have a more positive relationship with wealth compared to diversification to the on-farm sector? If so, which extra job in the non-farm sector seems to have the highest positive relation with wealth?*

Based on a scarce descriptive analysis, the suggestion is that diversification to the non-farm sector tends to be more positively correlating with wealth, with the occupation professional, technical, or managerial having the highest correlation with wealth. Thus, the answer here is also yes, with the disclaimer that it is simply a descriptive analysis of quite scarce data.

To conclude, there seem to be differences between the countries, but generally, non-farm diversification at least seems to correlate positively with wealth in these commercialized regions rural parts compared to being a household only engaging in agriculture. From this thesis, it is not absolute to say that the poor should get increased opportunities to diversify since the way of causality is unclear. However, there should nonetheless be efforts to increase their opportunities to access growing markets where they can find more income opportunities. A key to link rural poor to the growth of the non-farm economy is labor markets with good communications and transport systems (Haggblade, Hazell, and Reardon, 2010).

5. Conclusion

In this thesis, I have analyzed household income diversification and wealth both descriptively and with econometrics. The question I wanted to answer was first if there is a higher positive relationship between income diversification and wealth compared to only engage in farming in the rural population of SSA. The answer I found was yes, there is a positive difference in favor of income diversification. To further disseminate this finding I investigated if diversification to the non-farm sector or diversification to the on-farm sector tends to have a more positive correlation to wealth, and which extra job that tends to have the most positive correlation to wealth. I concluded that the non-farm sector seems to be more beneficial in comparison to the on-farm sector when diversifying on a household level. Further, it seems to be, as one would assume that the occupation group professional, technical, or managerial have a higher correlation to wealth, and unskilled manual labor to lower wealth as expected. These findings add to the previous literature and add validity for the rural area, these three countries, and a sample of the whole income distribution.

These questions were answered by looking at rural individuals in three more commercialized regions in three economically very different countries in SSA. Tigray in Ethiopia, Western Region in Ghana, and the Central Region in Malawi. The countries can be said to quite well reflect the different developmental states many countries all over SSA inhibit, but necessarily not all. The urge to investigate wealth and income diversification connects back to the necessity to improve welfare for the poor rural population in big parts of SSA.

There are a few topics I have avoided in my analysis and just discussed shortly in the literature review. These should be investigated further in coming analyses on the area, because of their probable relevance for income diversification and in the long run poverty alleviation. Land distribution, entry barriers such as credit limitations or education, and population growth are all very relevant for these studies, and the latter may be especially due to the discussions on the future population boom in Africa. Another potential area for future research is to try to infer causality between income diversification and wealth in the rural area where the poorest people tend to live and what special needs they might have. In my analysis, data limitations made it impossible for me to follow individuals over time, and hence I could just analyze pure correlations.

To finalize, can household income diversification to the off-farm sector possibly be one way out of poverty in rural SSA? My study cannot answer if this is the case, but the findings have at least not decreased the chances that this may be the way. Income diversification may be one tool to a life with a higher living standard in rural SSA than what is the case now.

References

Alesina, A. & Rodrik, D. (1994). Distributive Politics and Economic Growth, *The Quarterly Journal of Economics*, vol. 109, no. 2, pp.465–490.

Barrett, C. B., Reardon, T. & Webb, P. (2001a). Nonfarm Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics, and Policy Implications, *Food Policy*, Vol. 26, Available Online: www.elsevier.com/locate/foodpol.

Barrett, C. B., Bezuneh, M. & Aboud, A. (2001b). Income Diversification, Poverty Traps and Policy Shocks in Côte d'Ivoire and Kenya, *Food Policy*, vol. 26, no. 4, pp.367–384.

Barrett, C. B., Bezuneh, M., Clay, D. C. & Reardon, T. (2005). Heterogeneous Constraints, Incentives and Income Diversification Strategies in Rural Africa, *Quarterly Journal of International Agriculture*, Vol. 44, Cornell University Press.

Beegle, K., Christiaensen, L., Dabalen, A. & Gaddis, I. (2016). Poverty in a Rising Africa, [e-book] Washington, DC: World Bank, Available Online: https://openknowledge.worldbank.org/handle/10986/22575.

Berkvens, R. (1997). Backing Two Horses: Interaction of Agricultural and Nonagricultural Household Activities in a Zimbabwean Communal Area, 24, Leiden, Available Online: https://openaccess.leidenuniv.nl/bitstream/handle/1887/427/01PUB0000001139.pdf?sequence=1.

Bernstein, H. (2010). Class Dynamics of Agrarian Change, Fernwood Publishing.

Bigsten, A. & Tengstam, S. (2011). Smallholder Diversification and Income Growth in Zambia, *Journal of African Economies*, vol. 20, no. 5, pp.781–822.

Block, S. & Webb, P. (2001). The Dynamics of Livelihood Diversification in Post-Famine Ethiopia, *Food Policy*, vol. 26, no. 4, pp.333–350.

Bonaglia, F., Labella, P. & Marshal, J. (2008). Promoting Commercial Agriculture in Africa, Paris, France, Available Online: http://www.atdforum.org/IMG/pdf_Promoting_Commercial_Agriculture.pdf.

Bray, J. (2014). Cocoa Is Ghana, Ghana Is Cocoa..., Available Online: https://onthecocoatrail.com/2014/06/12/cocoa-is-ghana-ghana-is-cocoa/ [Accessed 19 May 2020].

Bryceson, D. F. (1996). Deagrarianization and Rural Employment in Sub-Saharan Africa: A Sectoral Perspective, *World Development*, vol. 24, no. 1, pp.97–111.

Bryceson, D. & Jamal, V. (1997). Farewell to Farms: De-Agrarianisation and Employment in Africa, Aldershot, UK: Ashgate Publishing Ltd.

Canagarajah, S., Newman, C. & Bhattamishra, R. (2001). Non-Farm Income, Gender, and Inequality: Evidence from Rural Ghana and Uganda, *Food Policy*, vol. 26, no. 4, pp.405–420.

Carlsen, J. (1980). Economic and Social Transformation in Rural Kenya.

Central Statistical Agency Ethiopia. (2018). THE 2015/16 ETHIOPIAN HOUSEHOLD CONSUMPTION – EXPENDITURE (HCE) SURVEY, Addis Ababa, Available Online: www.csa.gov.et/ehioinfo-internal?start=10.

Collier, P. & Lal, D. (1986). Labour and Poverty in Kenya: 1900–1980, Vol. 28.

Davis, B., Di Giuseppe, S. & Zezza, A. (2014). Income Diversification Patterns in Rural Sub-Saharan Africa : Reassessing the Evidence, *Policy Research Working Paper*, [e-journal] vol. 7108, no. November, pp.1–38, Available Online: http://documents.worldbank.org/curated/en/2014/11/20378029/income-diversification-patterns-rural-sub-saharan-africa-reassessing-evidence.

Deininger, K. & Binswanger, H. P. (1995). Rent Seeking and the Development of Large-Scale Agriculture in Kenya, South Africa, and Zimbabwe, *Economic Development & Cultural Change*, vol. 43, no. 3, pp.493–522.

Deininger, K. & Squire, L. (1998). New Ways of Looking at Old Issues: Inequality and Growth, *Journal of Development Economics*, vol. 57, no. 2, pp.259–287.

Doss, C., McPeak, J. & Barrett, C. B. (2008). Interpersonal, Intertemporal and Spatial Variation in Risk Perceptions: Evidence from East Africa, *World Development*, vol. 36, no. 8, pp.1453–1468.

Ellis, F. & Bahiigwa, G. (2003). Livelihoods and Rural Poverty Reduction in Uganda, *World Development*, vol. 31, no. 6, pp.997–1013.

Ellis, F. & Freeman, H. A. (2004). Rural Livelihoods and Poverty Reduction Strategies in Four African Countries, *Journal of Development Studies*, vol. 40, no. 4, pp.1–30.

Evans, H. E. & Ngau, P. (1991). Rural-Urban Relations, Household Income Diversification and Agricultural Productivity, *Development and Change*, vol. 22, no. 3, pp.519–545.

Fisher, M., Holden, S. T., Thierfelder, C. & Katengeza, S. P. (2018). Awareness and Adoption of Conservation Agriculture in Malawi: What Difference Can Farmer-to-Farmer Extension Make?, *International Journal of Agricultural Sustainability*, vol. 16, no. 3, pp.310–325.

Francis, E. (2000). Making a Living: Changing Livelihoods in Rural Africa, London: Routledge.

Ghana Statistical Service. (2019). Population by Regions - Western, Available Online: https://statsghana.gov.gh/regionalpopulation.php?population=MTQ1MTUyODEyMC43 MDc1&&Western&regid=7 [Accessed 20 May 2020].

Google Maps. (2020). Google Maps, Available Online: https://www.google.se/maps.

Griffin, K., Khan, A. R. & Ickowitz, A. (2002). Poverty and the Distribution of Land, *Journal of Agrarian Change*, vol. 2, no. 3, pp.279–330.

Haggblade, S., Hazell, P. B. R. & Reardon, T. (2007). Transforming the Rural Nonfarm Economy, *Opportunities and Threats in the Developing World. Baltimore*, [e-book], Available Online: http://environmentportal.in/files/ib58.pdf.

Haggblade, S., Hazell, P. & Reardon, T. (2010). The Rural Non-Farm Economy: Prospects for Growth and Poverty Reduction, *World Development*, [e-journal] vol. 38, no. 10, pp.1429–1441, Available Online: http://dx.doi.org/10.1016/j.worlddev.2009.06.008.

Heger Boyle, E., King, M. & Sobek, M. (2019). IPUMS-Demographic and Health Surveys: Version 7 [Dataset].

Hubbard, R. (1978). The Probable Consequences of Violating the Normality Assumption in Parametric Statistical Analysis, *Area*, [e-journal] vol. 10, no. 5, pp.393–398, Available Online: https://www.jstor.org/stable/pdf/20001404.pdf.

ICF. (2020a). Demographic and Health Surveys, *The DHS Program Website. Funded by USAID*, Available Online: http://www.dhsprogram.com. [Accessed 6 May 2020].

ICF. (2020b). Wealth Index Construction, Available Online: https://dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm [Accessed 3 May 2020].

IFAD. (2020a). Ethiopia, Available Online: https://www.ifad.org/en/web/operations/country/id/ethiopia [Accessed 7 May 2020].

IFAD. (2020b). Ghana, Available Online: https://www.ifad.org/en/web/operations/country/id/ghana [Accessed 5 May 2020].

IFAD. (2020c). Malawi, Available Online: https://www.ifad.org/en/web/operations/country/id/malawi [Accessed 7 May 2020].

IMF. (2020). Sub-Saharan Africa: COVID-19 A Threat to Development, Available Online: https://www.imf.org/en/News/Podcasts/All-Podcasts/2020/05/12/papa-ndiaye-afr-reo-covid [Accessed 14 May 2020].

Lanjouw, P., Quizon, J. & Sparrow, R. (2001). Non-Agricultural Earnings in Peri-Urban Areas of Tanzania: Evidence from Household Survey Data, *Food Policy*, vol. 26, no. 4, pp.385–403.

Lay, J. & Mahmoud, T. O. (2008). Few Opportunities, Much Desperation: The Dichotomy of Non-Agricultural Activities and Inequality in Western Kenya, *World Development*, vol. 36, no. 12, pp.2713–2732.

Lay, J., Narloch, U. & Mahmoud, T. O. (2009). Shocks, Structural Change, and the Patterns of Income Diversification in Burkina Faso, *African Development Review*, vol. 21, no. 1, pp.36–58.

Lipton, M. & Ravallion, M. (1995). Chapter 41: Poverty and Policy, in *Handbook of Development Economics*, Vol. 3, pp.2551–2657.

Little, P. D., Smith, K., Cellarius, B. A., Coppock, D. L. & Barrett, C. (2001). Avoiding Disaster: Diversification and Risk Management among East African Herders, *Development and Change*, vol. 32, no. 3, pp.401–433.

Longhi, S. & Nandi, A. (2017). A Practical Guide to Using Panel Data, Croydon: Sage Publications.

Monga, C. & Yifu Lin, J. (2015). The Oxford Handbook of Africa and Economics. Vol. 2., Oxford: Oxford University Press.

Msofi Mgalamadzi, L., Chimombo, M., Chinsinga, B., Kaiyatsa, S., Matita, M., Mazalale, J. & Chirwa, E. (2018). Groundnut Commercialisation Trends in Malawi, *Future Agricultures*, Available Online: https://www.future-agricultures.org/blog/groundnut-commercialisation-trends-in-malawi/ [Accessed 5 May 2020].

National Statistical Office Malawi. (2018). 2018 Malawi Population & Housing Census, Zomba, Available Online: https://malawi.unfpa.org/sites/default/files/resource-pdf/2018 Census Preliminary Report.pdf.

OCHA. (2011). Horn of Africa Crisis, Available Online: https://reliefweb.int/sites/reliefweb.int/files/resources/KuwaitOCHAHornofAfricaCrisis1 20911_BenParker_Kuwait.pdf [Accessed 24 May 2020].

Oya, C. (2010). Rural Inequality, Wage Employment and Labour Market Formation in Africa : Historical and Micro-Level Evidence, International Labour Organization.

Oya, C. (2007). Stories of Rural Accumulation in Africa: Trajectories and Transitions among Rural Capitalists in Senegal, *Journal of Agrarian Change*, vol. 7, no. 4, pp.453–493.

Reardon, T. (1997). Using Evidence of Household Income Diversification to Inform Study of the Rural Nonfarm Labor Market in Africa, *World Development*, vol. 25, no. 5, pp.735–747.

Reardon, T., Stamoulis, K., Balisacan, A., Cruz, M. E., Berdegué, J. A. & Banks, B. (1998). Rural Nonfarm Income in Developing Countries, in *The State of Food and Agriculture 1998*, Rome: Food and Agricultural Organization of the United Nations.

Reardon, T., Taylor, J. E., Stamoulis, K., Lanjouw, P. & Balisacan, A. (2000). Effects of Non-Farm Employment on Rural Income Inequality in Developing Countries: An Investment Perspective, *Journal of Agricultural Economics*, vol. 51, no. 2, pp.266–288.

Rodrik, D. (2016). Premature Deindustrialization, *Journal of Economic Growth*, vol. 21, no. 1, pp.1–33.

Rutstein, S. O. (2008). The DHS Wealth Index: Approaches for Rural and Urban Areas, Calverton, Maryland, USA, Available Online: https://dhsprogram.com/publications/publication-wp60-working-papers.cfm.

Sender, J. (2002). Women's Struggle to Escape Rural Poverty in South Africa, *Journal of Agrarian Change*, vol. 2, no. 1, pp.1–49.

Sender, J., Oya, C. & Cramer, C. (2006). Women Working for Wages: Putting Flesh on the Bones of a Rural Labour Market Survey in Mozambique, *Journal of Southern African Studies*, vol. 32, no. 2, pp.313–333.

Smith, D. R., Gordon, A., Meadows, K. & Zwick, K. (2001). Livelihood Diversification in Uganda: Patterns and Determinants of Change across Two Rural Districts, *Food Policy*, vol. 26, no. 4, pp.421–435.

Tellegen, N. (1997). Rural Enterprises in Malawi: Necessity or Opportunity?, Aldershot: Ashgate Publishing Ltd.

The World Bank. (2020). World Bank Data.

UN. (2020). Sustainable Development Knowledge Platform, Available Online: https://sustainabledevelopment.un.org/ [Accessed 10 May 2020].

UNDP Ethiopia. (2013). An Assessment of Operation and Performance of Commercial Farmers in Ethiopia.

Wessel, M. & Quist-Wessel, P. M. F. (2015). Cocoa Production in West Africa, a Review and Analysis of Recent Developments, *NJAS - Wageningen Journal of Life Sciences*, vol. 74–75, pp.1–7.

Wittenberg, M. & Leibbrandt, M. (2017). Measuring Inequality by Asset Indices: A General Approach with Application to South Africa, *Review of Income and Wealth*, vol. 63, no. 4, pp.706–730.

Wolff, E. N. (2015). Wealth Distribution, in *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*, Elsevier Inc., pp.450–455.

Young, K. H. & Theil, H. (1973). Principles of Econometrics., *Journal of the American Statistical Association*, Fourth., Vol. 68, Hoboken: Johan Wiley & Sons.

Appendix A

Table 15: Jobs by gender and income diversification definition

Female occupation	Non-farm			Farm		
	0	1	Total	0	1	Total
Clerical	0	9	9	9	0	9
Sales	0	823	823	823	0	823
Agricultural	5084	0	5084	0	5084	5084
Self-employed agricultural	3751	0	3751	0	3751	3751
Agricultural employee	679	0	679	0	679	679
Household, domestic, and services	0	51	51	51	0	51
Skilled and unskilled manual	0	4	4	4	0	4
Skilled manual	0	312	312	312	0	312
Unskilled manual	0	372	372	372	0	372
Other	0	10	10	10	0	10
Total	9514	1581	11095	1581	9514	11095

Female occupation	On-farm			Farm		
	0	1	Total	0	1	Total
Clerical	9	0	9	9	0	9
Sales	823	0	823	823	0	823
Agricultural	4757	327	5084	329	4755	5084
Self-employed agricultural	3751	0	3751	0	3751	3751
Agricultural employee	0	679	679	679	0	679
Household, domestic, and services	51	0	51	51	0	51
Skilled and unskilled manual	4	0	4	4	0	4
Skilled manual	312	0	312	312	0	312
Unskilled manual	372	0	372	372	0	372
Other	10	0	10	10	0	10
Total	10089	1006	11095	2589	8506	11095

Partner's occupation	Non-farm			Farm		
	0	1	Total	0	1	Total
Professional, technical, or managerial	0	234	234	234	0	234
Clerical	0	80	80	80	0	80
Sales	0	539	539	539	0	539
Agricultural	6702	0	6702	0	6702	6702
Self-employed agricultural	1987	0	1987	0	1987	1987
Agricultural employee	38	0	38	38	0	38
Household, domestic, and services	0	335	335	335	0	335
Skilled and unskilled manual	0	114	114	114	0	114
Skilled manual	0	689	689	689	0	689
Unskilled manual	0	359	359	359	0	359
Other	0	18	18	18	0	18
Total	8727	2368	11095	2406	8689	11095

Partner's occupation	On-farm			Farm		
	0	1	Total	0	1	Total

Professional, technical, or managerial	234	0	234	234	0	234
Clerical	80	0	80	80	0	80
Sales	539	0	539	539	0	539
Agricultural	6702	0	6702	6702	0	6702
Self-employed agricultural	1987	0	1987	0	1987	1987
Agricultural employee	0	38	38	38	0	38
Household, domestic, and services	335	0	335	335	0	335
Skilled and unskilled manual	114	0	114	114	0	114
Skilled manual	689	0	689	689	0	689
Unskilled manual	359	0	359	359	0	359
Other	18	0	18	18	0	18
Total	11057	38	11095	9108	1987	11095

Table 16: Number of diversifying households per sample

Sample	Non-farm Diversification			On-farm diversification		
	0	1	Total	0	1	Total
Ethiopia 2000	570	140	710	710	0	710
	80.28	19.72	100.00	100.00	0.00	100.00
Ethiopia 2005	246	107	353	353	0	353
	69.69	30.31	100.00	100.00	0.00	100.00
Ethiopia 2011	487	178	665	665	0	665
	73.23	26.77	100.00	100.00	0.00	100.00
Ethiopia 2016	166	137	303	303	0	303
	54.79	45.21	100.00	100.00	0.00	100.00
Ghana 1998	92	61	153	153	0	153
	60.13	39.87	100.00	100.00	0.00	100.00
Ghana 2003	98	59	157	157	0	157
	62.42	37.58	100.00	100.00	0.00	100.00
Ghana 2008	83	51	134	134	0	134
	61.94	38.06	100.00	100.00	0.00	100.00
Ghana 2014	99	80	179	175	4	179
	55.31	44.69	100.00	97.77	2.23	100.00
Malawi 2000	941	503	1444	1444	0	1444
	65.17	34.83	100.00	100.00	0.00	100.00
Malawi 2004	969	451	1420	1420	0	1420
	68.24	31.76	100.00	100.00	0.00	100.00
Malawi 2010	1760	1420	3180	3180	0	3180
	55.35	44.65	100.00	100.00	0.00	100.00
Malawi 2016	1647	750	2397	2375	22	2397
	68.71	31.29	100.00	99.08	0.92	100.00
Total	7158	3937	11095	11069	26	11095
	64.52	35.48	100.00	99.77	0.23	100.00

Figure 4: Qplot on normality

Table 17: Skewness/Kurtosis test for normality

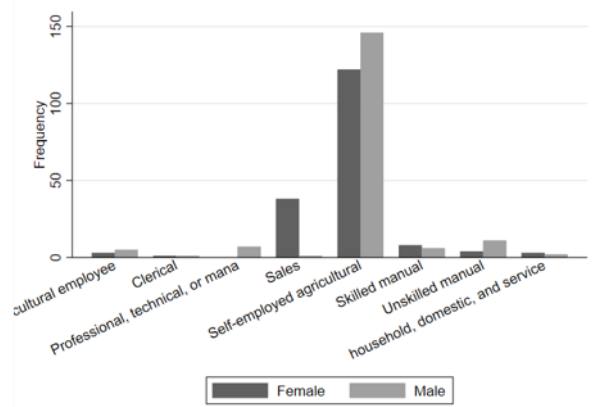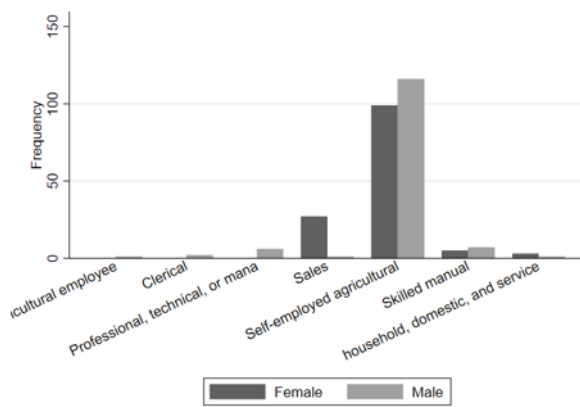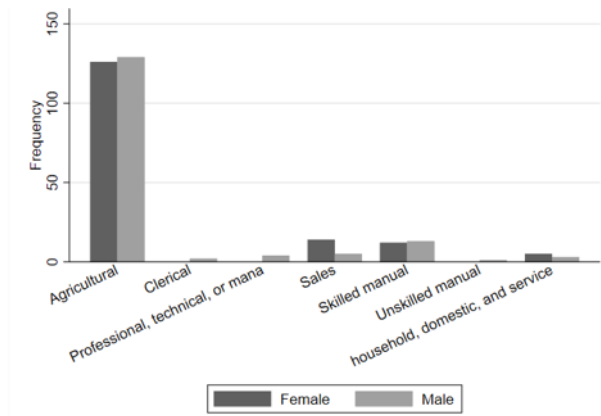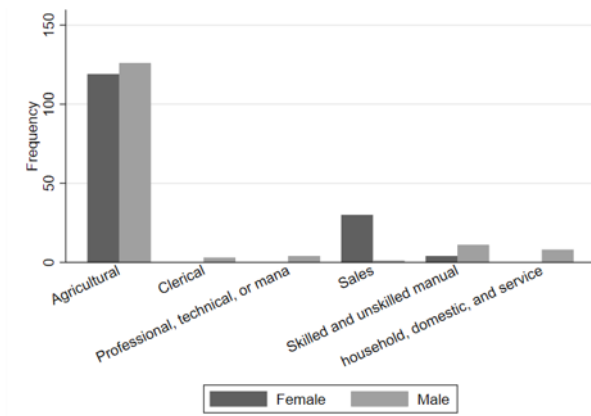| Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj_chi2(2)	Prob>chi2
Wealth (aggregate)	11,095	0.000	0.000	.	.
Wealth (Ethiopia)	2,031	0.000	0.000	.	0.000
Wealth (Ghana)	623	0.000	0.000	.	0.000
Wealth (Malawi)	8,441	0.000	0.000	.	.

Appendix B

From left to right, four figures per country, Ethiopia (2000, 2005, 2011, 2016), Ghana (1998, 2003, 2008, 2014), and Malawi (2000, 2004, 2010, 2016).
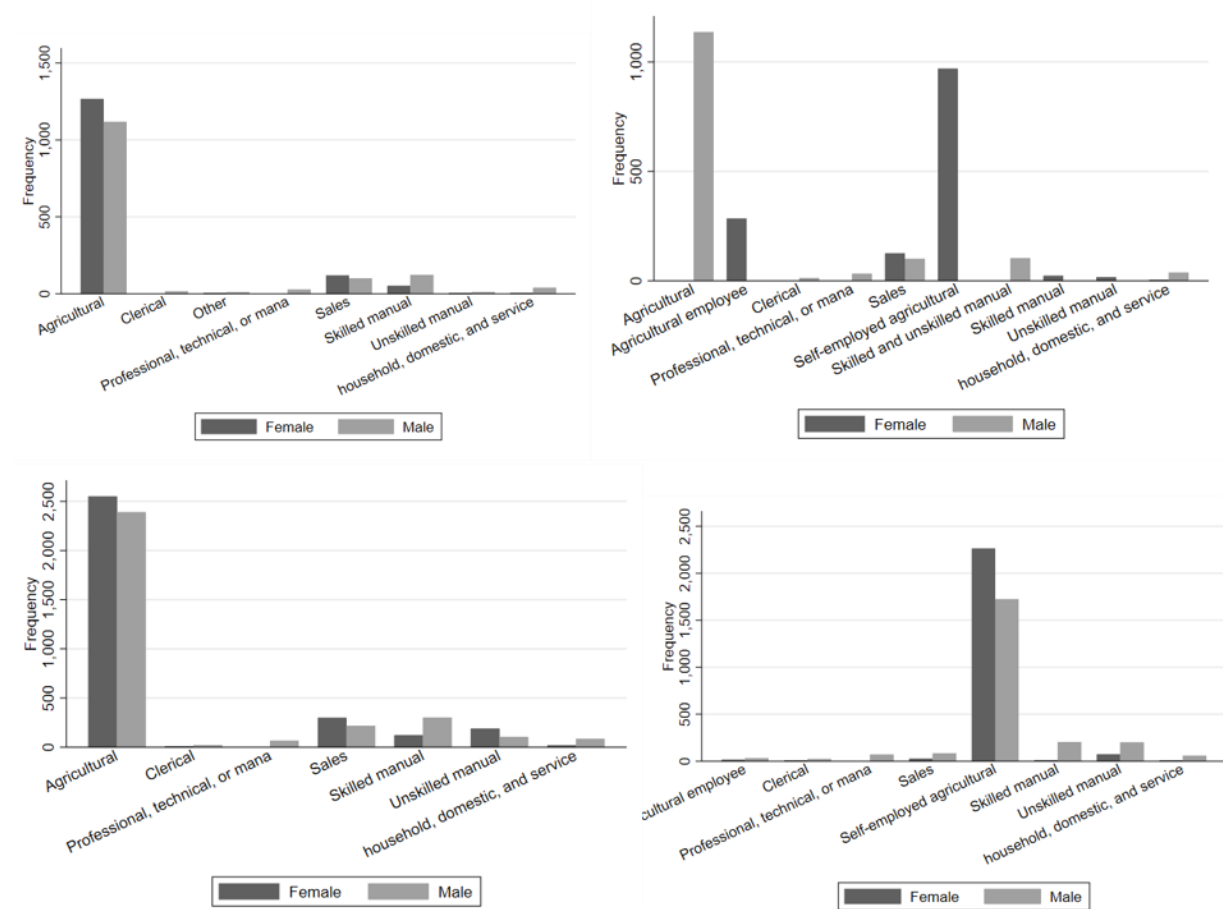
Ethiopia
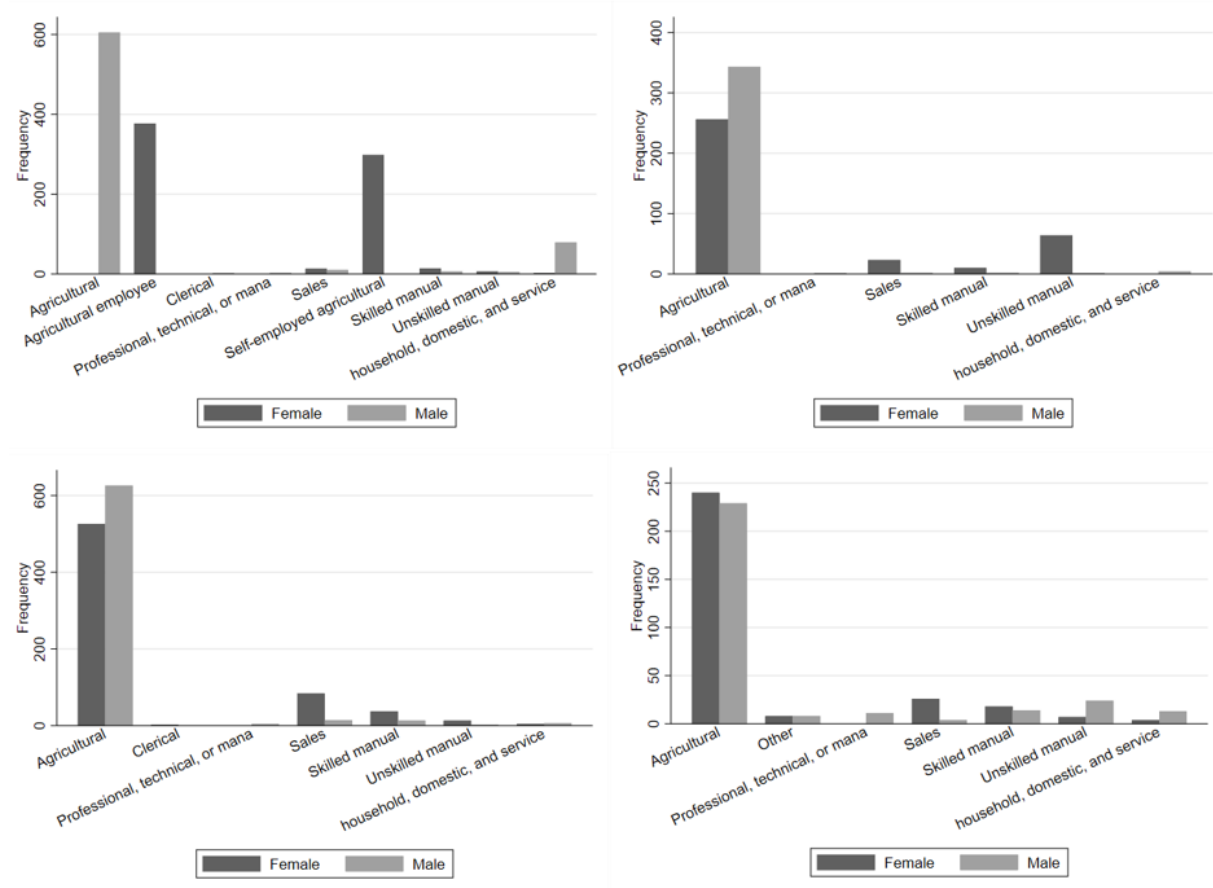
Ghana

Malawi

Figure 5: Division of labor between occupations by year
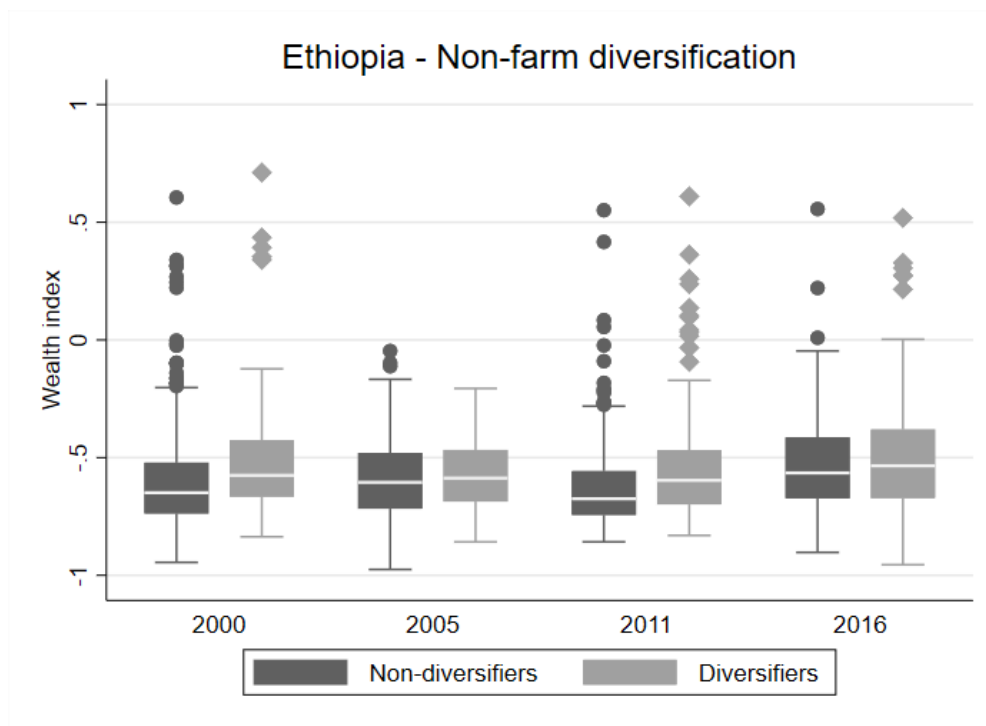
Figure 6: Boxplot distribution Ethiopia non-farm excluding outlier

Table 18: On-farm diversification and wealth in Ghana

VARIABLES	(1) Wealth	(2) Wealth	(3) Wealth	(4) Wealth	(5) Wealth
On-farm diversification	-0.248				-0.236
	(0.132)				(0.121)
Female age	0.005**	0.006	0.004	-0.002	0.010**
	(0.002)	(0.004)	(0.003)	(0.005)	(0.004)
Household members over the age of 5	-0.006	0.002	-0.001	0.029	-0.040*
	(0.010)	(0.021)	(0.017)	(0.030)	(0.018)
Female education	0.037**	0.040	0.019	0.056	0.048*
	(0.012)	(0.022)	(0.018)	(0.037)	(0.022)
Employment contract	0.041	-0.025	0.100*	-0.030	0.048
	(0.036)	(0.063)	(0.039)	(0.128)	(0.062)
Constant	-0.971***	-0.824***	-1.046***	-0.670	-0.751**
	(0.128)	(0.240)	(0.151)	(0.421)	(0.246)
Observations	623	153	157	134	179
Adjusted R-squared	0.138	0.011	0.015	-0.004	0.068
Year	All	1998	2003	2008	2014
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table 19: On-farm diversification and wealth in Malawi

VARIABLES	(1) Wealth	(2) Wealth	(3) Wealth	(4) Wealth	(5) Wealth
On-farm diversification	0.033				0.042
	(0.115)				(0.113)
Female Age	0.004***	0.0006	0.002	0.004**	0.008***
	(0.0007)	(0.0009)	(0.001)	(0.001)	(0.001)
Household members over the age of 5	0.039***	0.018***	0.022***	0.050***	0.048***
	(0.004)	(0.005)	(0.006)	(0.007)	(0.008)
Female education	0.148***	0.061***	0.095***	0.193***	0.154***
	(0.008)	(0.012)	(0.015)	(0.016)	(0.013)
Employment contract	0.026**	0.003	0.010	0.045**	0.016
	(0.009)	(0.013)	(0.017)	(0.016)	(0.019)
Constant	-0.838***	-0.539***	-0.563***	-0.958***	-1.098***
	(0.030)	(0.040)	(0.051)	(0.058)	(0.057)
Observations	8,435	1,444	1,417	3,177	2,397
Adjusted R-squared	0.103	0.033	0.073	0.097	0.120
Year	All	2000	2004	2010	2016
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table 20: On-farm diversification and wealth in Ghana - Male education and age

VARIABLES	(1) Wealth	(2) Wealth	(3) Wealth	(4) Wealth	(5) Wealth
On-farm diversification	-0.130***				-0.166***
	(0.039)				(0.047)
Age of partner	0.003	0.005	0.002	-0.001	0.003
	(0.002)	(0.003)	(0.002)	(0.004)	(0.003)
Household members over the age of 5	0.006	-1.63e-05	0.002	0.045	-0.018
	(0.011)	(0.031)	(0.020)	(0.025)	(0.021)
Partner education	0.057***	0.036	0.026	0.116***	0.050*
	(0.013)	(0.029)	(0.021)	(0.027)	(0.021)
Employment contract	0.063	0.026	0.093	0.070	0.051
	(0.044)	(0.085)	(0.049)	(0.128)	(0.079)
Constant	-1.079***	-0.981***	-1.024***	-1.249**	-0.656*
	(0.158)	(0.275)	(0.213)	(0.443)	(0.298)
Observations	486	118	115	109	144
Adjusted R-squared	0.155	0.005	0.002	0.135	0.017
Year	All	1998	2003	2008	2014
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table 21: On-farm diversification and wealth in Malawi - Male education and age

VARIABLES	(1) Wealth	(2) Wealth	(3) Wealth	(4) Wealth	(5) Wealth
On-farm diversification	0.028				0.035
	(0.108)				(0.107)

	(1)	(2)	(3)	(4)	(5)
Age of partner	0.004***	0.002*	0.002*	0.005***	0.006***
	(0.0006)	(0.0008)	(0.0010)	(0.001)	(0.001)
Household members over the age of 5	0.036***	0.015**	0.021**	0.046***	0.048***
	(0.004)	(0.005)	(0.006)	(0.008)	(0.007)
Partner education	0.143***	0.065***	0.102***	0.211***	0.139***
	(0.007)	(0.011)	(0.012)	(0.016)	(0.010)
Employment contract	0.027**	0.0006	0.027	0.035*	0.029
	(0.010)	(0.014)	(0.017)	(0.018)	(0.019)
Constant	-0.896***	-0.584***	-0.643***	-1.006***	-1.145***
	(0.031)	(0.041)	(0.055)	(0.061)	(0.055)
Observations	7,624	1,282	1,262	2,703	2,377
Adjusted R-squared	0.143	0.060	0.133	0.141	0.151
Year	All	2000	2004	2010	2016
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table 22: On-farm diversification and wealth in Ghana - Literacy

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Wealth	Wealth	Wealth	Wealth	Wealth
On-farm diversification	-0.241*				-0.223*
	(0.116)				(0.106)
Female age	0.005*	0.006	0.004	-0.003	0.008*
	(0.002)	(0.004)	(0.003)	(0.005)	(0.004)
Household members over the age of 5	-0.007	-0.0001	-0.0007	0.020	-0.037*
	(0.010)	(0.021)	(0.016)	(0.026)	(0.018)
Literacy	0.067**	0.067*	0.030	0.081	0.099
	(0.022)	(0.031)	(0.051)	(0.057)	(0.052)
Employment contract	0.040	-0.024	0.093*	-0.031	0.057
	(0.036)	(0.058)	(0.039)	(0.136)	(0.065)
Constant	-0.939***	-0.819***	-0.998***	-0.551	-0.696**
	(0.127)	(0.226)	(0.131)	(0.442)	(0.253)
Observations	623	153	157	134	179
Adjusted R-squared	0.139	0.020	0.012	-0.012	0.070
Year	All	1998	2003	2008	2014
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table 23: On-farm diversification and wealth in Malawi - Literacy

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Wealth	Wealth	Wealth	Wealth	Wealth
On-farm diversification	0.008				0.014
	(0.123)				(0.122)
Female age	0.003***	0.0004	0.002	0.003*	0.007***
	(0.0007)	(0.0009)	(0.001)	(0.001)	(0.001)
Household members over the age of 5	0.036***	0.018***	0.021***	0.045***	0.042***
	(0.004)	(0.005)	(0.006)	(0.007)	(0.008)
Literacy	0.102***	0.037***	0.061***	0.152***	0.104***
	(0.005)	(0.008)	(0.009)	(0.010)	(0.010)
Employment contract	0.028**	0.003	0.008	0.060***	0.007

	(0.009)	(0.013)	(0.017)	(0.017)	(0.019)
Constant	-0.780***	-0.521***	-0.507***	-0.895***	-0.968***
	(0.030)	(0.049)	(0.048)	(0.054)	(0.057)
Observations	8,421	1,444	1,417	3,164	2,396
Adjusted R-squared	0.078	0.029	0.047	0.083	0.075
Year	All	2000	2004	2010	2016
Year FE	Yes	No	No	No	No

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05