THE EFFECTS OF MEDIUM AND LARGE-SCALE FARMS ON YOUNG PEOPLE'S EMPLOYMENT IN AGRICULTURE: EVIDENCE FROM TANZANIA

By

Paul C. Samboko

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food and Resource Economics—Master of Science

2022

ABSTRACT

THE EFFECTS OF MEDIUM AND LARGE-SCALE FARMS ON YOUNG PEOPLE'S EMPLOYMENT IN AGRICULTURE: EVIDENCE FROM TANZANIA

By

Paul C. Samboko

There is limited empirical evidence on how the growth of large- and medium-scale farms is affecting employment outcomes across the whole agricultural sector in developing countries, and especially among young people (age 15-35 years). *A priori*, it is impossible to determine how medium- and large-scale farms affect employment for young people in agriculture. Using employment data for young people in Tanzania, this study examines whether increases in the region-level share of cropping households that are medium- and large-scale farms (MLSFs) improve or worsen agricultural employment outcomes for young people. The outcomes include: (i) employment in crop/livestock production on own farm; (ii) self-employment in agribusiness activities and (iv) employment in agriculture via any of the first three categories above. Correlated random effects probit model results suggest that the growth of medium-scale farms is associated with reductions in the participation of young people in the production of crops/livestock on their own or their family's farms. It is also associated with a reduction in the employment of young people in the agricultural sector overall. The growth of large-scale farms is associated with an increase in self-employment in agriculture by young people. The government needs to be cognizant of the effects of different farm sizes on employment. Medium-scale farms may not be an avenue to improve young people's involvement in agriculture. However, large-scale farm expansion may improve young adult's employment in agricultural employment.

To Eustensia, Mirriam, Ireen, and Ryan.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my profound gratitude to the talented people who have helped me reach this point. First and foremost, I thank Dr. Nicole Mason-Wardell and Professor Thom Jayne for the invaluable support rendered throughout my studies and for the guidance rendered in the lead-up to the commencement of my studies. It has been a pleasure knowing and working with you both. Special appreciation goes to Professor Thom Jayne for financial support during my final semester.

My MSU experience would be incomplete without extending gratitude to my assistantship supervisors Dr. Mason-Wardell and Saweda Liverpool-Tasie. The capacity-building series under the PRCI project was a great learning and personal development experience that I will forever cherish. I also wish to thank Dr. Robert Richardson and Dr. Saweda Liverpool-Tasie for agreeing to join my guidance committee. You both provided invaluable advice that will go a long way in shaping how I conceive research ideas and go about writing articles.

Last but not the least, I would like to thank the staff and fellow graduate students in the Department of Agricultural, Food, and Resource Economics at Michigan State University for their contribution to my personal and professional development during my studies.

Funding Acknowledgement

This research is made possible by the generous support of the American people through the United States Agency for International Development (USAID) through funding to the Feed the Future Innovation Lab for Food Security Policy (FSP) under Cooperative Agreement No. AID-OAA-L-13-00001 (Zambia Buy-In) and to the Feed the Future Innovation Lab for Food Security Policy Research, Capacity, and Influence (PRCI) under grant 7200AA19LE00001. The contents

are the responsibility of the study author and do not necessarily reflect the views of USAID or the United States Government.

TABLE OF CONTENTS

LIST OF TABLES vii

LIST OF FIGURES viii

KEY TO ABBREVIATIONS ix

1 INTRODUCTION 1

2 CONCEPTUAL FRAMEWORK 9
2.1 The Spillover Effects of Medium and Large-Scale Farms 9
2.1.1 Effects on Young People's Participation in On-Farm Agricultural Production 9
2.1.2 Effects on Young People's Wage Employment in the Agricultural Sector 14
2.1.3 Effects on Young People's Self Employment in the Agricultural Sector 16
2.2 Summary 17

3 DATA 18
3.1 Data Source 18
3.2 Construction of the Employment Variables 20

4 METHODS 21
4.1 Empirical Strategy 21

5 RESULTS AND DISCUSSION 25
5.1 Descriptive Results 25
5.2 Econometric Results 28
5.2.1 Effects of Changes in the Regional Share of Medium- and Large-Scale Farms on Employment Outcomes of Young People 28
5.2.2 Heterogenous Effects of Changes in the Regional Share of Medium- and Large-Scale Farms on Employment Outcomes of Youth vs. Young Adults 32

6 SUMMARY AND CONCLUDING REMARKS 35
6.1 Summary 35
6.2 Implications for Policy 36
6.3 Study Limitations 37

APPENDICES 38
APPENDIX A: Other Possible Drivers of Young People's Employment Outcomes in The Agricultural Sector 39
APPENDIX B: Diagnostic and Robustness Test Results 44

BIBLIOGRAPHY 51

LIST OF TABLES

Table 1: Descriptive Statistics of Key Variables 27

Table 2: Distribution of Young People's Agricultural Employment Outcomes by Medium and Large-Scale Farm Tercile 28

Table 3: Econometric Results 29

Table 4: Disaggregated Econometric Results for Youth vs. Young Adults 33

Table A1: Industries used to Classify Employment Variables in Agriculture 43

Table B1: Tests for Attrition Bias 45

Table B2: Fixed Effects Model Results (Average Marginal Effects) 46

Table B3: Correlated Random Effects Probit Model Results (Average Marginal Effects) 47

Table B4: Econometric Results- Youths 48

Table B5: Econometric Results- Young Adults 49

LIST OF FIGURES

Figure B1: Relationship between the Prevalence of MLSFs and Young People's Employment Outcomes 44

KEY TO ABBREVIATIONS

CRE	Correlated Random Effects
HH	Household
LSAI	Large-Scale Agricultural Investment
LSF	Large-Scale Farm
LSMS-ISA	Living Standards Measurement Study - Integrated Surveys on Agriculture
MLSFs	Medium and Large-Scale Farms
MSF	Medium-Scale Farm
OLS	Ordinary Least Squares
TZNPS	Tanzania National Panel Survey

1 INTRODUCTION

Since the 2008 food crisis, there has been an increase in large-scale agricultural investments (LSAIs) by foreign-owned firms, mainly in Africa (Land Matrix; Schoneveld 2014). Over this period, there has also been an increase in the share of land controlled by medium-scale farms (MSFs) in less densely populated countries (Jayne et al. 2019; Deininger and Byerlee 2011). For example, the share of total agricultural land controlled by medium-scale farms (i.e., farms cultivating 10-100 hectares) grew from 22.1% in 1992 to 31.6% in 2005 for Ghana. Similar trends are observed in Tanzania (MSF growth from 23.8% to 25.7% between 2008 and 2012) and Zambia (MSF growth from 26.3% to 35.6% between 2008 and 2014) (Jayne et al. 2016). MSFs now control more farmland than large-scale agricultural investors in these countries. For example, in Tanzania, the share of land held by MSFs is 39.0%, while that controlled by large-scale farms at 7% (Ibid).

These observed changes in farm structure prompted several studies on the effects of medium and large-scale farms (MLSFs) on local communities in general and on smallholder farming communities. *A priori*, it is unclear how MLSFs will affect the communities, households, or individuals neighboring them (The Oakland Institute 2011). MLSFs could yield benefits for neighboring communities through employment creation, facilitating infrastructural developments, improved access to markets, use of best practices, and the adoption of improved technologies (Deininger and Xia, 2016). However, MLSFs could also lead to adverse environmental and socioeconomic outcomes such as loss of permanent assets like agricultural land, reduced quality of life, land conflicts, and water shortages (Aabø and Kring, 2012; Messerli

et al., 2015; Williams et al., 2012; Henley, 2017). Despite their overall impacts being unclear, LSAIs have been supported as part of ongoing investment promotion efforts in some African countries such as Zambia.[1] The interactions between MLSFs and their neighboring communities are also complex with variable expected effects, first, because not all MLSFs interact with smallholders in the same manner. This implies that the impacts may depend on the nature of the relationships and business models used (Zhan, Mirza, and Speller 2015; Liverpool-Tasie et al., 2020). Second, medium- vs. large-scale investments differ in scope and scale; as such, their effects are likely to vary. For instance, some crops may be labor intensive with higher employment spillovers than those that are capital intensive, while livestock production may be highly mechanized compared to crop production (Cordes, Östensson & Toledanoy 2016; Deininger & Xia 2016; Baumgartner et al., 2015). Third, the impacts of MLSFs are also likely to depend on the strength of land governance systems in the investment destinations (Nolte, 2014). Where land governance systems strongly protect smallholders, the impact may be less adverse than in areas where this is not the case (Ibid).

This paper contributes to the evidence and debates about the spillover effects of medium and large-scale farms on nearby local communities. Specifically, the paper addresses the question, what is the effect of MLSFs on young people's (age 15-35) employment outcomes in the agricultural sector? The employment outcomes considered are (i) employment in crop production or livestock raising on an individual's own or their family's farm; (ii) self-employment in

[1] Attracting foreign direct investment and producing crops for export have been the primary reasons for supporting LSAIs in recipient countries. The promotional activities are also part of broader policy objectives for agricultural development. The Zambian government through the Seventh National Development plan and the Second National Agricultural Policy seeks to promote medium and large-scale agricultural investments through a farm blocks development model. In addition, the Zambia Development Agency seeks to attract foreign direct investment in agriculture by providing incentives to investors (see Manda, Tallontire and Dougill, 2017; Republic of Zambia Government of., 2016; 2017).

agribusiness activities excluding (i);[2] (iii) wage/salaried employment in agriculture; and (iv) employment in the agricultural sector, defined as employment in any of (i), (ii), or (iii).[3] Data from three waves of the Tanzania National Panel Survey are used to answer these questions using econometric tools. As shown below, previous work on the spillover effects of medium and large-scale farms has not addressed the question of the effects of MLSFs on young people's employment outcomes across the agricultural value chain.

The interest in young people stems from the fact that young people represent the fastest-growing population group in developing countries; this presents opportunities and challenges to these countries' governments (United Nations, 2015). The opportunities lie in the potential for reaping the demographic dividend that comes with a bulging productive population, while the concerns arise from the social and political unrest, stagnant growth, and disillusionment that could come with high levels of unemployment and under-employment of the young (Adelaja & George, 2020; United Nations, 2019; Jayne & Yeboah, 2017). The challenge is how to create employment for this growing youthful population, as enshrined in United Nations Sustainable Development Goal Number 8[4] and Tanzania's National Strategy for Youth Involvement in

[2] This includes activities such as the provision of tillage services, agro-dealerships, and crop and livestock output trading, among others.

[3] More specifically and based on the definitions used by the Tanzania National Panel Surveys (the main dataset used here), an individual is classified as being employed in agricultural production if they engaged in any crop production or livestock raising activities as an unpaid helper on their family's farm at any time in the seven days before the survey date. An individual is classified as having been employed in wage/salaried employment in agriculture if they engaged in relevant paid wage or salaried activities off their own/family farm at any time in the seven days before the survey date. An individual is defined as being self-employed in agriculture if they engaged in self-employment or any self-owned business activity related to agriculture at any time in the seven days before the survey date, or if they engaged in any unpaid work related to the family's agri-business in the seven days before the survey date.

[4] The United Nations Sustainable Development Goal Number 8 seeks to "promote sustained, inclusive, and sustainable economic growth, full and productive employment and decent work for all". Target No. 8.6 under this goal is directly related to youth employment creation and reads, "By 2020, substantially reduce the proportion of youth not in employment, education or training" (United Nations, 2020).

Agriculture[5] (United Nations, 2020; Republic of Tanzania, 2016). The agricultural sector is a key employer for many people in developing countries; as such, governments in these countries have promoted LSAIs to increase local employment. Thus, the employment spillovers of MLSFs on young people's employment in the agricultural sector is also an important policy question.

Prior work on the spillover effects of MLSFs can be divided into two strands: one that focuses on the spillover effects of large-scale farms (e.g., Deininger and Xia, 2016; Anti, 2021), and a second that focuses on the spillover effects of both large and medium-scale farms (e.g., Wineman et al., 2021). The first strand emerged directly from the observed increase in LSAIs by multinationals after the global food price crisis of 2008, as well as associated concerns about their potential adverse effects. The first sub-strand of this literature is mainly qualitative and shows that LSAIs contribute to (i) reduced land access and land conflicts among smallholder farmers and marginalized groups in the areas of LSAI investments (e.g., see Messerli et al., 2015; Williams et al., 2012; Zaehringer et al., 2018); (ii) a reduction in access to water and water rights by the communities targeted for LSAIs (Williams et al., 2012); and (iii) an increased incidence of livestock-induced crop damage (Ibid). These studies provided useful insights, but the evidence was largely based on case studies.

A second sub-strand of the literature on the impacts of LSAIs is quantitative and developed in response to the limited and largely case study-based evidence on LSAI impacts that was available up to that point. For example, Deininger and Xia (2016) find evidence of positive spillover effects of LSAIs on smallholder farmers in Mozambique. These spillovers were in the

[5] The National Strategy for Youth Involvement in Agriculture seeks "to create enabling environment for attracting the youth's engagement in agriculture which has the highest potential for assimilating the unemployed youths."

form of increased adoption of agricultural best practices and increased modern input (e.g., inorganic fertilizers and improved/hybrid seeds) use by smallholders located near large-scale farmers. Positive spillover effects on employment were also observed for large-scale crop operations, but not for livestock farms. There were no statistically significant spillover effects of large farms on smallholder commercialization, access to output markets, or crop yields. Also, Lay, Sipangule, and Nolte (2020) use nationally representative household survey data and data on the location and date of commencement of operations by LSAIs to estimate the spillover effects of large-scale farms on smallholder farms in Zambia. They find that large-scale investments locate mainly in areas with infrastructure and existing markets and compete for land with smallholders. They also find evidence of positive spillovers of the presence of large-scale farms on smallholder crop yields. In addition, proximity to LSAIs is associated with a switch to maize cropping at the expense of other staple crops by smallholders. However, they find no evidence of spillovers on smallholder fertilizer use. Other studies within this sub-strand of literature (i.e., quantitative studies on the impacts of LSAIs) focus on modeling changes in smallholder cropping patterns due to LSAIs (e.g., Debonne et al., 2018) or the net direct employment effects of LSAIs (Nolte and Ostermeier, 2017).

More recently, Anti (2021) provides quantitative evidence on the impacts of LSAIs on employment, expenditure, and investments among households in Cambodian villages. The employment outcomes studied include employment in paid non-agricultural employment, employment as an agricultural laborer, and employment in agricultural production at the individual level. Results show that there were no technology spillovers from LSAIs to local farmers in the form of increased expenditure on chemical inputs or hybrid seeds. However, there are positive effects on wage employment. The results also show that proximity to large-scale

farms is associated with the abandonment of agricultural production among individuals. The emergence of the quantitative literature allowed for national-level estimates of the impacts of LSAIs, further complementing the case studies. However, this literature excludes medium-scale farms and led to the emergence of the second strand of the literature, which is discussed next.

The second strand of literature is motivated by Jayne et al., (2016) who observed that, in Africa, LSAIs control less land than medium-scale farms. Therefore, both medium- and large-scale farms need to be considered in studies on the implications of changes in farm structure, particularly on the African continent or other areas that may be experiencing similar trends. For example, Wineman et al. (2021) examine the impact of medium and large-scale farms on the agricultural behavior of smallholder farmers in Tanzania with a focus on technology adoption, agricultural commercialization, farm productivity, and participation in input markets. They find that proximity to medium-scale farms is positively correlated with improved seed usage, the size of cultivated land, and extension access among smallholders.

Similarly, Chamberlin & Jayne (2020) consider the influence of farm structure on smallholder household incomes. They find positive influences of proximity to large farms on wage incomes, nonfarm incomes, and farm income. Burke, Jayne, and Sitko (2020) find evidence of an increased likelihood of selling maize among smallholder households close to medium-scale farms in Zambia. They also show that proximity to large farms is associated with increases in the expected sales and an increased likelihood of selling to large traders. Liverpool-Tasie et al. (2019) study the effect of proximity to medium-scale farms on smallholder farmer outcomes in Nigeria. Outcomes included smallholder household income, the share of output sold, poverty status, productivity, and fertilizer and agrochemical use. In addition, they identify the channels through which medium-scale farms generate spillovers for nearby smallholder farms. They find

that proximity to medium-scale farms is associated with a reduction in the incidence and severity of poverty, higher yields, and increased welfare. Their results indicate that cost reduction and learning effects[6] are the two main channels through which the spillovers to smallholder farmers occur.

Despite the fairly large literature on MLSFs described above, there are several knowledge gaps. First, there is limited empirical evidence on the spillover effects of MSLFs on individual-level outcomes; the primary focus of studies to date has been on household or Meso-level (e.g., village-level) effects. Spillover effects at the individual level could differ from those at the household or meso-level; as such, addressing this gap is important. For example, MLSF presence may lead to mechanization and ultimately improve labor productivity. This could lead young people to exit crop/livestock production. On the other hand, this could translate into cropland expansion at the household level.

Secondly, there are no studies that consider how MLSFs affect young people's employment outcomes broadly or in the agricultural sector. The only study that comes close to this is Anti (2021), which was described above. Studying the employment spillovers of MLSFs is important because the young population is expected to grow rapidly in developing countries, creating employment challenges. Because most developing countries are primarily agrarian, the agricultural sector will, for some time, continue to be crucial for creating employment for young people. Even with economic growth in these countries, agriculture will likely remain important

[6] The cost effects refer to the transaction cost reductions for the small farmer that come with purchasing inputs from or with the medium or large-scale farmers. The learning effect results from trainings or extension messaging from medium and large-scale farms that increase the productivity of smallholder farms (see Liverpool-Tasie et al., 2020).

for job creation since expansions in the agri-food system are likely to create more employment opportunities (Christiaensen, Rutledge & Taylor, 2020).

Third, there are no studies that consider the effect of changes in the prevalence of MLSFs on young people's agricultural employment outcomes across the whole agricultural sector. This study is the first to break down agricultural employment outcomes for young people into agribusiness self-employment (hereafter self-employment), the raising of livestock or production of crops on own/family farms, and wage/salaried employment.

This study seeks to fill the aforementioned gaps in the literature by econometrically testing the relationship between changes in the share of medium and large-scale farms in a region and young people's agricultural sector employment outcomes.

The rest of the thesis is organized as follows. Section 2 discusses the mechanisms through which medium- and large-scale farms may generate spillovers to young people and other individuals in communities located near the MLSFs. In section 3, the data are described in further detail. The methods are discussed in section 4. Section 5 presents the main findings. A conclusion and discussion of the main implications thereafter are presented.

2 CONCEPTUAL FRAMEWORK

2.1 The Spillover Effects of Medium and Large-Scale Farms

Conceptually, the presence of medium and large-scale farm operations could affect nearby communities through two main channels, namely, learning effects and cost effects (Liverpool-Tasie et al. 2020). These two mechanisms work to reduce the transaction costs faced by communities close to MLSFs; increase adoption and availability of improved technologies; and improve their access to better quality inputs, information, and markets offering better prices and services (Ibid.). Other neighborhood effects on local people include direct and indirect employment, land availability and land access, and access to water. Below is a detailed discussion of ways in which young people's agricultural sector employment outcomes may be impacted by their proximity to medium and large-scale farms.

2.1.1 Effects on Young People's Participation in On-Farm Agricultural Production

Participation in the production of crops and livestock depends on access to arable and grazing lands. The investments by medium and large-scale farms may lead to the displacement of local people or may reduce the amount of available land in communities (Chu, Young & Phiri 2015). This may constrain smallholder farmers' participation in agricultural production, with farmers relegated to low-return agriculture on small family farms or pushed into the nonfarm sector. Two scenarios are commonly used to highlight the potential impacts of MLSFs on the neighboring communities' localized land availability and land use. First, medium- and large-scale agricultural investments may be brownfield or greenfield. The brownfield investments take place on previously occupied land or farms, whereas greenfield investments occur on previously

unoccupied land including forested land. When investments are greenfield, the effects on the local population may be more adverse than when they are brownfield (Nolte & Ostermeier, 2017). However, the potential adverse effects of greenfield investments may be mitigated through consultative processes and tend to depend on existing land governance practices and the extent to which land is a constraint in the target destinations (Ibid).

Typically, where investments are on newly acquired land, the local population's access to land and water rights tends to be more adversely affected, all else equal (e.g., see Williams et al., 2012; Meinzen-Dick, 2007), and this tends to be worse for poor and marginalized groups. The poor may be driven onto marginal lands, and access to good pasture may be reduced (Bukari & Kuusana, 2018). Also, under such conditions, livestock-induced crop damage may increase and could be a source of conflict. This is important given that for most African countries, land is becoming more of a constraint due to rapid population growth and the acquisition of large tracts of land by foreign and elite local investors (Land Matrix; Jayne et al. 2016; Jayne et al. 2014). In Zambia, the evidence suggests that large-scale agricultural investments do not affect how much land is cultivated by smallholder farmers (Lay, Nolte, and Sipangule 2020). However, the authors do not show if the presence of large-scale farms reduces available land.

Further, one's access to land in most African countries is still mostly through inheritance, and allocation by traditional leaders While land rental and sales markets have emerged in several African countries, facilitating land transfers (Wineman and Liverpool-Tasie 2017a; Abay, Chamberlin & Berhane 2020; Chamberlin and Ricker-Gilbert 2015; Ricker-Gilbert et al. 2019; Wineman & Liverpool-Tasie 2018) to more productive users, young people's participation in land rental markets is constrained by income as well as by institutional constraints imposed by the government or traditional leaders, as was the case among females in India (e.g., see Sharma,

1982). Moreover, there is evidence showing that land inheritance is being delayed due to an increase in rural life expectancy among the elderly (Brooks, Zorya, and Gautam, 2012). This means that in the presence of land constraints, and the lack of capacity to participate in land rental and sales markets, poor access to land for agriculture may constrain the participation of young people in farming and this is more so for females (e.g., see Wineman and Liverpool-Tasie, 2017b).

Large-scale agricultural investments have also been shown to lead to violent land-related conflicts while pushing the locals into less productive lands and increasing the incidence of crop damage by livestock (see Bukari & Kuusana, 2018). The effects of MLSFs on land availability for the poor also depend on whether locals are consulted, and the prevailing land governance systems. With the weak involvement of locals and weak land governance systems, the inheritance system is being challenged by MLSFs, and land tenure systems are changing more rapidly, with an increased incidence of land under leasehold tenure (Doss, Meinzen-Dick & Bomuhangi, 2013; Nolte, 2016). This threatens land access by vulnerable groups including the poor, young people, and women.

In irrigated production systems, and for livestock rearing, water is a key input. Thus, under conditions of scarcity, LSAIs may further reduce how much water is available for other users or may lead to the loss of water rights (Williams et al., 2012; Meinzen-Dick, 2007). This then could constrain young people's ability to raise livestock and produce crops outside of rain-fed systems of production.

One of the major reasons LSAIs have been promoted by recipient governments is the promise to generate positive spillovers in the economy. The presence of MLSFs in areas tends to attract

financial and mechanized service providers; it also attracts more large-scale traders that pay better prices to areas where they locate (See Sitko et al., 2018; Liverpool-Tasie et al., 2020). The availability of mechanized services, financial services, and better markets for produce could attract young people into a sector they previously viewed as less productive, dirty, and laborious (Kafle, Paliwal & Benfica, 2019).

Public investments in roads and the location of input and service providers near MLSFs may reduce the transaction costs (e.g., transport, marketing, and communication costs) associated with smallholder farmers' participation in crop or livestock production (Liverpool-Tasie et al. 2020 characterize this as the cost effect). However, if there are no infrastructural investments that come with MLSFs, this may not happen. For example, in Zambia, Lay, Nolte, and Sipangule (2020) show that large-scale agricultural investments are in areas with already good market access and infrastructure.

Input suppliers may also locate closer to large-scale agricultural investments and the surrounding communities. Further, more large-scale traders may locate in these areas, thereby increasing the marketing options and generating spillovers for the local population. All this has the potential to address one of the key barriers to smallholder market participation across the agri-food system (i.e., high transaction costs), thereby increasing the marketed surplus (Pingali and Rosegrant 1995; Key, Sadoulet, & de Janvry, 2000). However, in some cases, investments by large and medium-scale farms may follow areas with already developed infrastructure (Lay, Nolte, and Sipangule 2020). The implication of this is that where transaction cost reductions are not large enough, the employment spillovers from MLSFs for young people may not be significant. On the other hand, the lowering of transaction costs may lower barriers to the participation of young people in crop or livestock production.

Another way in which the presence of larger farms may influence young people's decision to participate in on-farm agricultural production is through several avenues that generate positive spillovers (i.e., yield increases) to small farmers by improving the adoption of best agricultural practices, and technological transfers if the MLSFs grow similar crops to nearby populations (Deininger and Xia, 2016). For example, smallholders located near medium and large-scale farms may be able to access better extension and cheaper inputs directly from the MLSFs who enjoy economies of scale from transportation. There is also the possibility of pooling of input purchases by smallholders and MLSFs, thereby lowering transaction costs associated with input purchases.

In Tanzania, Wineman et al. (2021) show that the presence of large farms is associated with improved extension access, increased likelihood of cultivation of cropland, and increased improved seed usage among neighboring farms. For Zambia, Lay, Sipangule and Nolte (2020) also find evidence of positive yield spillovers arising from the proximity of smallholder farms to large-scale farms. Deininger and Xia (2016) find evidence of positive spillovers in the form of increased adoption of agricultural best practices and modern inputs by smallholders near large-scale farmers in Mozambique. However, no statistically significant effects are found for large-scale farm effects on smallholder commercialization, access to output markets, or crop yields. MLSFs could also impact technological investments in processing and packaging by smaller firms.

The surge in large-scale private firms across Southern Africa also comes with increased livestock packaging, processing, access to modern markets, and branding technologies/practices. This creates opportunities for spillovers to smaller farmers as is the case among small and medium entrepreneurs (see Reardon et al., 2021). The rise in modern markets may create an incentive to

increase the marketable surplus among smallholder farmers. With agriculture more productive due to the MLSF-induced gains, the productivity differential could motivate young people to participate in crop and livestock production. Also notable is the possibility of participation in value chain financing arrangements in the form of out-grower schemes or contract farming that facilitate smallholder access to finance and markets from medium and large-scale farms (e.g., see Matenga & Hichaambwa 2017). This may as well attract young people to on-farm agricultural production.

A priori, it is difficult to determine how growth in the incidence of MLSFs could impact the employment of young people in crop and livestock production. This is due to the many opposing forces that work to ultimately determine the direction of the effect. As such, we cannot hypothesize or justify a relationship between changes in the prevalence of MLSFs and young people's employment in crop production and livestock rearing.

2.1.2 Effects on Young People's Wage Employment in the Agricultural Sector

The increase in the incidence of MLSFs could avail more direct wage employment opportunities for communities near the MLSFs. However, the employment creation potential of the MLSFs is likely to differ by the scale of operation and the nature of enterprises the farms engage in. It is expected that among MLSFs, all else equal, the medium-scale farms will relate more than large-scale farms with the local population, employing them in unskilled jobs while training others in the more specialized activities. On the other hand, compared to medium-scale farms, large-scale farms are expected to employ more urban-based skilled machine operators (Baumgartner et al., 2015; Chinigò, 2015). A reason for this is that medium-scale farms are socio-culturally more similar to smallholder farms than are large-scale farms (Liverpool-Tasie et al., 2020). Therefore,

it is expected that the wage employment effects generated by medium-scale farms will be higher than those generated by large-scale farms, all else equal.

Further, where medium and large-scale farms engage in labor-intensive value addition on-farm, or in areas close to their farms, they may generate employment spillovers to the neighboring local community or for migratory labor. In this respect, Deininger and Xia (2016) find evidence of positive spillovers in the form of employment for large-scale crop operations, but not for large-scale livestock farms. This supports the notion that the nature of the enterprise matters for the employment spillovers of MLSFs. Nolte and Ostermeier (2017) extend this literature by studying the net direct employment impacts of large-scale agricultural investments. They conclude that LSAIs lead to negative net employment effects on the nearby communities after replacing smaller farms. However, as with other related literature, this is not done for young people's employment in agriculture.

Because the decision to participate in agriculture is also based on whether there are alternatives to agriculture, where wage employment opportunities from MLSFs do not exist or are limited, the rural poor can be pushed into subsistence farming. However, in some cases, it is unclear *a priori* how young people may be affected by a high prevalence of off-farm wage employment opportunities. Where wage opportunities from MLSFs are better paying, young people may abandon crop and livestock production. In fact, Wineman et al. (2021) show that MLSFs generate greater multipliers that produce options for small farm-households looking to exit agriculture. However, high wage incomes from MLSFs can also facilitate young people's employment in agriculture.

Indirect employment creation also often follows the increased incidence of MLSFs. MLSFs typically attract other value chain actors in agriculture such as input suppliers, traders, service providers, and suppliers of raw materials used by processors. These may employ locals as aggregators or laborers (see Nolte and Ostermeier, 2017).

As earlier indicated, the public investments that may come with an increase in medium and large-scale agricultural investments may reduce the cost of commuting to urban areas for young people. This may generate opportunities for employment in urban areas.

Based on this, we hypothesize that an increase in the regional share of both medium and large-scale farms will be associated with an increase in agricultural sector wage employment of young people. This impact will be greater for medium-scale farms than large-scale farms because medium-scale farms are more likely to be more labor-intensive than large-scale farms.

2.1.3 Effects on Young People's Self Employment in the Agricultural Sector

Self-employment activities in agriculture include business activities by individuals such as the provision of services for tillage, transportation, agro-dealerships, and crop and livestock output trading, among others. As noted by Nolte and Ostermeier (2017), MLSFs could lead to the expansion of input suppliers, traders, or service providers offering products and services to MLSFs into the areas surrounding MLSFs. This additional competition for business opportunities created by a rise in MLSFs may be associated with the abandonment of self-employment by young people. This is especially true for self-employment activities that may be directly in competition with the larger players who enjoy scale economies.

However, if the larger actors work with the local young people, employing them as aggregators, the self-employment activities may be reinforced or improved. However, these are likely to be

outweighed by the negative effects that may come with reductions in self-employment off the farm due to an increased presence of larger agri-business firms. In other cases, self-employment in agriculture may be abandoned by young people because of the wage differential arising from better wage opportunities from MLSFs or agribusinesses located in the area (Barret et al. 2008).

Against this backdrop, we hypothesize that an increase in the incidence of MLSFs will be associated with a reduction in self-employment in the agricultural sector among young people.

2.2 Summary

This section has presented the key research hypotheses, following a review of the theoretical explanations for the linkages between young people's employment outcomes in agriculture and their proximity to medium and large-scale farms. Other factors may explain observed employment outcomes; these are discussed in Appendix A.

3 DATA

3.1 Data Source

The study uses three waves of the Tanzania National Panel Surveys (TZNPS) that were collected in 2008-2009, 2010-2011, and 2012-2013 (hereafter referred to as the 2009, 2011, and 2013 TZNPS). These surveys were implemented by the Tanzania National Bureau of Statistics and are part of the World Bank's Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) that are being implemented in several African countries. The TZNPS surveys contain detailed information for individual members of households and their livelihoods, as well as household-level variables on agricultural production, consumption expenditure, and other socioeconomic variables. The data also contain community-level characteristics.

These surveys are nationally representative, and they are panel surveys at the household and individual levels. The sample size for the 2009 survey was 3,265 households in 409 enumeration areas. From the 2009 sample, 3,168 and 3,071 households were re-interviewed in 2011 and 2013, respectively. The household attrition rate between 2009 and 2013 is 5.9%.

At the individual level, the panel surveys track adult household members (age 15 and above) that left their original household, enabling panel analysis at the individual level. There were 5,225 young people interviewed in the 2009 survey; of these, 4,243 were reinterviewed in 2011. From the 4,243 young people interviewed in the 2011 survey, 4,180 were reinterviewed in 2013. A total of 4,180 young people were consistently interviewed in all three waves. This balanced panel of 4,180 young people is used for the main analysis. For details on the survey design, see Republic of Tanzania (2011).

The use of agricultural census data would be ideal for constructing the localized indicators of the prevalence of medium and large-scale farms since census data are representative of all farm types in Tanzania. The 2009 Agricultural Census (AC) for Tanzania is available but would result in time-constant MSLF prevalence indicators. Nevertheless, Wineman et al. (2020), noted that the AC MLSF prevalence measures are highly correlated with analogous but time-varying measures based on the TZNPS (with a correlation coefficient >0.8).

Given the time-constant nature of the AC data and the high correlation between the MLSF prevalence indicators constructed using the AC and TZNPS data, I opt to use the TZNPS-based indicators here (following Wineman et al. 2020). The specific indicators used are the share of cropping households that are large-scale farms in each region (the number of regions in the data is 26) and the share of cropping households that are medium-scale farms in each region.[7]

Since the prevalence indicators are constructed using the same dataset as the individual-level variables, a Jackknife procedure was used to construct these indicators to avoid endogeneity issues. Specifically, the time-varying MLSF prevalence indicators for individual *i* in household *j* in region *r* were constructed using all other households in the region while omitting the household in question.

[7] Also possible is the computation of the share of land in a region that are medium scale with the prevalence indicator computed for each household by leaving out the household's own observation. By doing so, it is expected that the household's activities are uncorrelated with the prevalence indicator and therefore its own error term. However, if land is already a constraint in a region, then by definition, the household's land area is by definition a residual amount. Thus, the share of MLSFs in a region is preferred as it is likely to suffer less from this problem.

Following Wineman et al. (2021), for Tanzania, a farm is considered medium-scale if it has 5–20 ha of cultivated cropland, 0.5–2 ha under fruits/vegetables, or holds 35–105 tropical livestock units (TLUs).[8] Farms that are larger than this in any dimension are considered large-scale.

3.2 Construction of the Employment Variables

For all employment variables, the industries used to compute the self-employment variable are restricted to those in the agricultural sector. These are listed in Appendix A. While not explicit in the categorization used by the LSMS-ISA surveys, the categories cover agro-dealerships and agricultural service provision for tillage, veterinary, and spraying. It also includes agro-processing, warehousing, and trading. All employment by government agencies including parastatals was not considered in the variable creation.

The employment definitions, however, do not capture any employment in agricultural finance, agricultural insurance, or in transportation. This is because these categories likely include substantial employment in sectors other than agriculture, and it is impossible to isolate agriculture's share of the total. This means that for the narrow definition used here, the employment indicators for self-employment and off-farm wage employment may be understated. Thus, estimated results need to be interpreted with this caveat.

[8] A tropical livestock unit is a mature animal unit weighing 250 kgs and is used to obtain standardized herd sizes for different animals (Kassam and Fischer 1991).

4 METHODS

4.1 Empirical Strategy

Equation 1 presents the main equation used to estimate the spillover effects of medium and large-scale farms on young people's agricultural sector employment outcomes.

$$y_{ijrt} = \beta_0 + \beta_1 MSF_{rt} + \beta_2 LSF_{rt} + \boldsymbol{X_{ijrt}\beta}_3 + \psi_{ij} + \tau_t + \varepsilon_{ijrt} \quad (1)$$

y_{ijrt} is a binary indicator for individual *i* in household *j* in region *r* at time *t*, defined separately for the four outcome variables of interest: (i) agricultural production on their or their family's farm, (ii) self-employment in agriculture, (iii) wage/salaried employment off-farm within the agricultural sector and (iv) employment in the agricultural sector - that is employment in (i), (ii), or (iii). y_{ijrt} is set equal to one if the individual participated in that activity and zero otherwise. $MSF_{rt}\ and\ LSF_{rt}$ are indicators of the prevalence of medium and large-scale farms in region *r* at time *t*. $\boldsymbol{X_{ijrt}}$ is a vector of controls (the individual's age, household size, the education level of the household head – not the young individual him/herself – for reasons described below. Also included as covariates are indicators for the survey month to account for the seasonality of the employment variables used. This is because, for the employment variable used, the TZNPS asks about household members' employment status for the last 7 days leading up to the survey and interviews occurred across different months. Summary statistics for the variables used in the analysis are provided in section 5.1. The terms ψ_{ij} and τ_t represent the time-constant unobserved individual heterogeneity, and time-specific heterogeneity, respectively. ε_{ijrt} is the idiosyncratic error term, while $\beta_0, \beta_1, \beta_2$ and $\boldsymbol{\beta_3}$ are parameters to be estimated.

The main source of identification is the behavioral change in individuals due to variations over time in the prevalence of medium and large-scale farms at the regional level, conditional on observed confounding factors and time-constant unobserved heterogeneity. The location of medium and large-scale farms may not be random and interactions between communities and MLSFs may be subject to selection bias. Specifically, conditional on controlling for factors that influence the location of medium and large-scale farms, it is expected that unobservable factors (e.g., unobserved ability, social skills, peer effects, or perceptions about farming) may influence the interaction of individuals with MLSFs in the search for off-farm wage employment in agriculture or in interactions that translate into participation in agricultural production or self-employment in agriculture.

It is necessary to address potential biases arising from correlation between the time-constant unobserved heterogeneity and the MLSF prevalence indicators or other observed covariates. Panel data methods such as first differencing (FD), fixed effects (FE), and the correlated random effects (CRE) approach achieve this (Angrist and Pischke 2008; Lin and Wooldridge, 2019). In this study, equation (1) is estimated via an individual-level linear fixed effects and CRE probit models. However, this approach does not address the endogeneity that may arise from the correlation of the MLSF prevalence indicators or other observed covariates with the idiosyncratic error term (Wooldridge, 2010). Given the obvious diversity in the broader young people grouping, because they are in different stages of life, we also estimate equation (1) for youth (age 15-24 years) and young adults (age 25-35 years) separately. This approach helps us show the heterogeneity between the two groups.

Further, some explanatory variables (usually referred to as bad controls) that influence employment outcomes may also be affected by the changes in the prevalence of MLSFs, leading

to overcontrol bias. For example, the education level of an individual and the accumulation of a household's assets are likely to be influenced by changes in the prevalence of MLSFs in a region. This works through the effect of MLSF prevalence on household incomes and expenditure on education, i.e., with improvements in household incomes, school enrollment and attendance may increase with a positive influence on an individual's education. It is also expected that infrastructural developments that come with MLSFs may reduce the cost of school attendance, improving educational outcomes. Other such variables include an individuals' health outcomes, agricultural prices, and technologies. Because of this, the vector $\boldsymbol{X_{ijrt}}$ excludes all individual- and household-level variables that are likely to be bad controls. Instead of an individual's education, the household head's education level is used (following Anti, 2021). Similarly, the vector $\boldsymbol{X_{ijrt}}$ also includes shocks experienced by the household that may influence employment outcomes. In computing the shocks variable, all shocks that may be due to changes in the prevalence of MLSFs are excluded. Examples of such excluded shocks are the loss of salaried employment or non-payment of salaries, increases in agricultural prices, and severe water shortages.

In panel data analysis, attrition bias is potentially a problem. Specifically, some individuals may not be interviewed across all three waves of the TZNPS, triggering the potential for inconsistent estimates if the attrition is not random. Attrition bias is a problem if there are systematic differences in the characteristics of observations that drop out of the sample thus leading to a correlation between the error term and the MLSF prevalence indicators (or other covariates) (Wooldridge, 2010). However, if the attrition is random, then estimates are consistent.

To test for potential attrition bias, we use the regression-based test recommended by Wooldridge (2010). The test involves the inclusion of a dummy variable for reinterview in equation 1 and estimating the regression involving an unbalanced panel of the first two waves of the TZNPS. The

reinterview dummy variable in wave *t* is equal to 1 if an individual appears in wave *t+1* of the survey, and zero otherwise. This means that for the attrition bias test, the third wave cannot be used since there is no fourth wave to compute the reinterview status for wave 3. With the three waves of data for the TZNPS, a t-test of the coefficient /on the reinterview variable tests the null hypothesis of no attrition bias against the alternative of attrition bias (conditional on the observed covariates). The test results are presented in Appendix B. The test results show that attrition bias is not a problem in this case. For all the equations estimated, the reinterview variable is statistically insignificant at the 10% level. Thus, we fail to reject the null hypothesis of no attrition bias.

5 RESULTS AND DISCUSSION

5.1 Descriptive Results

Table 1 presents descriptive statistics for the key variables used in the econometric modeling. In 2009, 37.9% of the young people in our panel sample engaged in crop or livestock production on their own or family farms; this number dropped to 34.9% in 2011, then rose to 37.5% in 2013. The percentage of young people in self-employment in agriculture increased from 9.7% to 12.4% in 2011, and then to 13.8% in 2013. The percentage of young people in wage employment remains very low, however, it rose from 3.3 to 4.5 between 2009 and 2013. Individuals in self-employment increased by 4.1 percentage points over the period 2009 and 2013—the highest percentage point increase across the three employment outcomes. In addition to the influence of the changes in MLSFs, these changes in employment may be driven by evolving circumstances of young people who may be leaving school into the labor market or into marriage after dropping from school in the subsequent wave. We are unable to isolate the sources of the changes in employment outcomes.

Table 2 shows a summary of the employment outcomes by the medium-scale and large-scale farm prevalence terciles. For the medium-scale farm prevalence terciles, results show that the share of individuals employed in crop/livestock production rises with the medium-scale prevalence terciles (i.e., it is consistently higher in the upper terciles). However, the same cannot be said about the large-scale farm prevalence indicator in that for regions in the top third of the large-scale farm prevalence, the share of young people employed in crop/livestock production is similar to that in the bottom third. It is higher for regions for the middle third of the large-scale farm prevalence

distribution. This suggests a quadratic relationship between the share of medium and large-scale farms and employment outcomes. The computed marginal effects thus consider the squared terms of these variables.

Table 1: Descriptive Statistics of Key Variables

VARIABLES	(1) N	(2) Wave 1	(3) Wave 2	(4) Wave 3	(5) T-test (Wave 1 vs 2)	(6) T-test (Wave 2 vs Wave 3)
Panel A-Household level variables						
Household size (number of members)	2,409	5.8	6.2	6.3	***	***
Share of medium scale farms in a region (%)	2,409	8.2	10.0	10.3	***	**
Share of large-scale farms in a region (%)	2,409	0.90	0.88	1.20	***	***
Household head's education level (years)	2,409	14.1	13.6	13.7	*	ns
Share of households that experienced shocks	2,409	0.78	0.70	0.72	**	*
Panel B-Individual level variables						
Individuals employed in crop/livestock production (%)	4,180	37.9	34.9	37.5	**	**
Individuals in wage employment in agriculture. (%)	4,180	3.3	4.8	4.5	**	**
Individuals in self-employment in agriculture (%)	4,180	9.7	12.4	13.8	**	**
Individuals employed in agriculture (%)	4,180	45.5	45.2	49.3	ns	***
Married individuals (%)	4,180	29.5	30.7	39.8	ns	***
Age of individuals (years)	4,180	21.9	23.9	25.8	***	***
Male individuals (%)	4,180	0.48	0.48	0.48		

Source: Author's calculations using the TZNPS. Notes: ***, **, * indicates statistical significance in the mean differences at the 1%,5% and 10% levels, ns indicates statistical insignificance at the 10% level or lower.

Similarly, self-employment in agriculture among young people is highest for the middle third of the medium and large-scale farm prevalence, and it is lowest for the bottom third for both the medium and large-scale farm prevalence indicators. Young people's off-farm wage employment in agriculture, however, increases with an increase in the share of medium-scale farms in a region. In contrast, while young people's off-farm wage employment in agriculture increases with an increase in the prevalence of large-scale farms up to the second tercile, it declines for regions in the top third of the share of cropping households that are large-scale in a region.

A non-parametric plot of the relationships between employment outcomes and the prevalence of medium and large-scale farms shows quadratic relationships between the prevalence MLSFs and young people's employment outcomes. For this reason, all regressions include the squared values of the prevalence of MLSFs (see Appendix B).

Table 2: Distribution of Young People's Agricultural Employment Outcomes by Medium and Large-Scale Farm Tercile

	Medium-Scale Farm Prevalence Tercile			Large-Scale Farm prevalence Tercile		
	1	2	3	1	2	3
Individuals employed in crop/livestock production, %	26.7	31.2	52.8	35.9	45.1	34.8
Individuals self-employed in agriculture, %	10.6	14.1	11.4	10.3	16.4	12.9
Off-farm wage employment in agriculture, %	2.1	3.7	6.9	3.6	6.4	4.4

Source: Authors' calculations using the 2008, 2010, and 2012 Tanzania National Panel Surveys.

5.2 Econometric Results

5.2.1 Effects of Changes in the Regional Share of Medium- and Large-Scale Farms on Employment Outcomes of Young People

Table 3 contains the estimates of for the variables of interest based on the FE estimates of equation 1 and CRE-probit estimates of a similar model. (The full results are reported in Appendix B Tables B2 and B3.) Overall, the FE and CRE probit results are similar for the MSF

variable but not for the LSF variable. Given the binary nature of the dependent variables, the CRE probit results are preferred and are the focus of the discussion that follows.

Table 3: Econometric Results

	Variables	Share of cropping households that are large-scale (AME)	Share of cropping HHs that are medium-scale (AME)
Fixed Effects model results			
(1)	Employed in own/Family Crop/Livestock Production	0.904	-0.438*
		(0.715)	(0.238)
(2)	Self-employed in agriculture	0.412	-0.178
		(0.493)	(0.153)
(3)	Off-farm wage employment in agriculture	0.246	0.084
		(0.386)	(0.110)
(4)	Employment in Agriculture	1.124	-0.471*
		(0.736)	(0.248)
Correlated Random Effects Probit			
(5)	Employed in own/Family Crop/Livestock Production	-0.842	-0.560**
		(0.725)	(0.232)
(6)	Self-employed in agriculture	0.993**	-0.195
		(0.488)	(0.157)
(7)	Off-farm wage employment in agriculture	-0.072	0.163
		(0.391)	(0.115)
(8)	Employment in agriculture	0.014	-0.575**
		(0.736)	(0.240)

Notes: AME = average marginal effect. Robust standard errors clustered at the individual level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. See Tables B2 and B3 in Appendix B for the full regression results. CRE Probit estimates include individual-level means of all time-varying right-hand side variables. The MLSF variables in the regression models are shares measured on a [0,1] scale.

From the estimated results, for Tanzania, growth in the share of medium-scale farms in a region is associated with a reduction in the participation of young people in crop and livestock production. The estimated magnitude of this effect (i.e., the average marginal effect) is -0.56. This means that given an increase from no MSF to all MSFs in a region, all else equal, the probability of a young person engaging in crop/livestock production decreases by 56 percentage

points. This finding raises questions about the relative strength of the mechanisms through which medium-scale farms adversely affect the participation of young people in crop and livestock production. While Wineman et al. (2021) find that proximity to medium-scale farms leads to expansion of cultivated land by households in nearby communities, it is unclear whether this expansion translates into the mechanization of agricultural production among existing farms, thereby freeing up young people from working on their family farms into other types of employment within the agricultural sector such as agribusiness and wage employment. Also, the observed relationship may be driven by the freeing up of young people's labor into other sectors due to labor productivity gains arising from improved seed and inorganic fertilizer use (Nolte and Ostermeier, 2017).

Further, it appears that the cost reduction in crop/livestock production (Liverpool-Tasie et al. 2020) that may come with an increase in the incidence of MSFs in a region do not necessarily benefit young people or that other constraints to young people's participation in crop/livestock production may be more as MSFs become more prevalent in regions. Examples of such constraints include the lack of access to agricultural land (see Williams et al., 2012; Meinzen-Dick, 2007; Chu, Young & Phiri, 2015 who show that land shortages arise as MLSFs increase in regions). It is thus expected that if the growth of MSFs translates into land shortages, increased participation of young people arising from learning effects is also unlikely.

The results also show that the growth of medium-scale farms is associated with reduced employment of young people in the whole agricultural sector (i.e., in at least one of crop/livestock production, self-employment or off-farm wage employment). As the MSF effects on self-employment in agriculture and off-farm wage employment in agriculture are not statistically significant, this result (average marginal effect of -0.57) is likely driven by the effect

of an increase in MSFs on employment in crop/livestock production (similar average marginal effect of -0.56).

One possibility is that the linkages between the prevalence of MSF and young people's self-employment or off-farm wage employment in agriculture are not as strong. MSF expansion may not be accompanied with public infrastructure development that reduces transaction costs associated with business activities among young people (see Liverpool-Tasie et al. 2020). Further, medium-scale farms could be underutilizing their land, as such while they are classified as medium-scale, the wage employment that could come with expanding cropland may not happen (Sitko and Chamberlin, 2015). Another possibility is that household labor is in surplus among MSFs and they need not hire additional individuals as they expand cropland or livestock production. As indicated earlier, access to finance may be a barrier towards the establishment of agri-businesses activities by young people.

The growth of large-scale farms in a region is positively and significantly associated with engagement in self-employment in agriculture by young people. The average marginal effect is 0.99, meaning an increase from no LSFs to all LSFs in a region is associated with a 99 percentage points increase in the probability of a young person engaging in self-employment in agriculture. This may arise due to reductions in transaction costs due to public investments that come with LSFs business activities (see Liverpool-Tasie et al. 2020). Another possibility is that the establishment of LSFs in a region creates opportunities for locals to partner with input suppliers, traders, or service providers offering products and services to LSFs (see Nolte and Ostermeier, 2017). However, it is unclear why this is the case for changes in LSF and not for MSFs in these data. One possibility is that the LSFs may be engaged in activities that are quite

different from the MSFs; thus, interactions leading to self-employment in agriculture only hold for changes in LSF prevalence.

The results on the effect of LSFs on young people's employment in crop and livestock production are not statistically significant. A possible explanation is that these may not have a significant effect on land available to young people. This happens if they mostly occur on brownfield investments or if LSFs are mostly in land abundant areas (Nolte & Ostermeier, 2017), or if the process of land acquisition is consultative in areas with strong land governance systems (Nolte 2016). Another possibility is that these LSFs do not operate enterprises similar to what the locals operate. As such, transaction cost reductions or learning (Liverpool-Tasie et al. 2020) that may come with LSF growth may not happen for the young people seeking to engage in crop or livestock production.

Similarly, for the wage employment outcome, the insignificant result may suggest that MLSFs may be highly mechanized and thus not employ young people, or that they employ highly skilled machine operators from outside the region (Cordes, Östensson & Toledanoy 2016; Deininger & Xia 2016; Baumgartner et al. 2015).

5.2.2 Heterogenous Effects of Changes in the Regional Share of Medium- and Large-Scale Farms on Employment Outcomes of Youth vs. Young Adults

Table 4 shows that there is heterogeneity in the effect of changes in the share of medium- and large-scale farms on the employment outcomes of youth vs. young adults. Growth in the share of medium- and large-scale farms has no statistically significant impact on agricultural employment of youth across all employment categories.

Table 4: Disaggregated Econometric Results for Youth vs. Young Adults

	Variables	Share of cropping households that are large-scale (AME)	Share of cropping HHs that are medium-scale (AME)
Young Adults			
(1)	Employed in own/Family Crop/Livestock Production	0.374	-0.922**
		(1.136)	(0.372)
(2)	Self-employed in agriculture	1.124	-0.222
		(0.962)	(0.310)
(3)	Off-farm wage employment in agriculture	0.109	0.233
		(0.695)	(0.226)
(4)	Employment in Agriculture	1.972*	-1.169***
		(1.137)	(0.391)
Youth			
(5)	Employed in own/Family Crop/Livestock Production	-1.302	-0.326
		(0.981)	(0.312)
(6)	Self-employed in agriculture	0.652	-0.142
		(0.498)	(0.161)
(7)	Off-farm wage employment in agriculture	-0.217	0.125)
		(0.465)	(0.118)
(8)	Employment in agriculture	-1.187	-0.19
		(1.021)	(0.317)

Notes: AME = average marginal effect. Robust standard errors clustered at the individual level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. See Tables B2 and B3 in Appendix B for the full regression results. CRE Probit estimates include individual-level means of all time-varying right-hand side variables. The MLSF variables in the regression models are shares measured on a [0,1] scale.

For young adults, the results are mostly consistent with those for young people in Table 3, save for the size of the AME and the self-employment variable. There is a negative association between the growth of medium-scale farms in a region and self-employment in crop/livestock production (significant at the 5% level). There is also a negative association between the growth in the share of medium-scale farms in a region and overall employment in the agricultural sector (significant at the 1% level). However, the growth of large-scale farms is positively associated with young adult's employment in agriculture (significant at the 10% level).

Clearly, there is heterogeneity in the effect of changes in the share of cropping households that are medium- and large-scale on the employment outcomes of youth vs. young adults. This supports the idea that the two groups are at different stages of life, with youth largely in school and under the care of their parents. Thus, one has to be cognizant of the differential effects of changes in the share of MLSFs in a region on different age groups of young people. Policies that promote agricultural employment among young people must be designed to account for these differences.

6 SUMMARY AND CONCLUDING REMARKS

6.1 Summary

Increases in the incidence of large-scale agricultural investments and the share of land controlled by medium-scale farmers have prompted numerous studies across the developing world that aim to understand the spillover effects of these farms on entire households and communities residing near these farms. Empirical evidence on these spillover effects is important because governments often promote large-scale agricultural investments and because the net spillover effects of medium- and large-scale farms (MLSFs) are difficult to predict, *a priori*. However, despite the growing body of literature on MLSF spillover effects, to date, there have been no studies that assess the effects of changes in the prevalence of MLSFs on agri-food system-related employment among young people. Young people currently face high unemployment levels, and these levels are expected to increase further as populations grow, which necessitates such studies.

Against this backdrop, this study used nationally representative panel data on individuals and the incidence of medium- and large-scale farms in a region where the individual lives to examine the effects of the prevalence of MLSFs on agri-food system-related employment outcomes among young people in Tanzania.

Results suggest a negative association between an increase in the prevalence of medium-scale farms in a region and the employment of young people in crop and livestock production on their own or their family's farms. No such effect is found for large-scale farms. The most likely mechanism driving the observed result on young people's crop and livestock production is the

effect that increasing medium-scale farms have on local land availability. However, this paper only provides suggestive evidence; further research is needed to confirm this.

There is also evidence of a positive association between an increase in the prevalence of large-scale farms and young people's self-employment in agriculture. No such effect is found for medium-scale farms. The most likely mechanism responsible for this are transaction cost reductions arising from public investments in infrastructure and potential business partnerships between the local young population and input suppliers, e.g., in service provision and demand aggregation.

Our results also show heterogeneity in the effects of regional changes in the share of MLSFs on youth vs. young adults. Youth are unaffected, while young adults are affected in a similar manner to the broader group of young people, except for the self-employment outcome variable.

6.2 Implications for Policy

The results suggest that medium-scale farms could stifle the participation of young people in crop and livestock production, most likely through their impact on local land availability. If future studies confirm this to be the case, then efforts aimed at encouraging young people to produce crops and livestock may need to improve young people's access to land to be effective. The fact that LSFs are positively associated with self-employment among young people suggests that these may be a potential for creating self-employment for young people. Specifically, the public investments that come with LSF development and the associated location of businesses in areas with LSFs could be key to increasing self-employment among young people in Tanzania. Based on the results presented here, agricultural and employment policies in Tanzania need to be cognizant of the role that medium and large-scale farms could play in affecting employment opportunities for young people in Tanzania.

Given heterogeneity in the impacts, employment policies must be tailored to the needs of youth and young adults, which are likely different. An employment strategy that targets the wider category of young people may not be effective, at least for youth.

6.3 Study Limitations

The study has several limitations. The main limitation is that we fail to account for other employment sources such as agricultural finance and insurance, transportation in agriculture, and agricultural retail trade. It is impossible to isolate these from the data in the TZNPS. In addition, the approach used here does not control for the potential endogeneity arising from the correlation between the idiosyncratic error term and the MLSF prevalence indicators. As such, the results presented are associations and not causal effects. We are also unable to isolate the effects of changes in the share of MLSFs on employment of young people in crop vs. livestock production because the data collects crop/livestock engagement as one variable. Further, the share of LSFs in regions is very low; similarly, wage/salaried employment engagement is quite low. Thus, these results should be interpreted with caution given the likely low statistical power.

APPENDICES

APPENDIX A: Other Possible Drivers of Young People's Employment Outcomes in The Agricultural Sector

In addition to proximity to MLSFs, there are several individual, household, and locational factors that may also influence the employment outcomes of young people in agriculture. This section discusses these factors broadly. A note is that most of the evidence on the correlates of labor supply decisions in agriculture is for all age groups. The literature does not have a special focus on factors that may be associated with young people's labor supply decisions across the whole agri-food system.

Employment in agricultural production on own or family farms

At the individual level, employment outcomes may be influenced by one's age, gender, social networks, peer effects, marital status, education level, and one's beliefs about farming. The literature shows that for young people, the production of crops or livestock is increasingly becoming less attractive because it is perceived as a dirty and laborious job, has low returns, and is for individuals that have lower levels of education (see Tocco, Bailley & Davidova, 2013; Kashi, Paliwal & Benfica 2019).

It is thus expected that higher levels of education among young people open opportunities to engage in economic activities that are more productive than crop or livestock production. Ahaibwe, Mbowa and Lwanga (2013) confirm this, as their results show that youth with secondary education are less likely to engage in farming, compared to those with lower education.

Concerning the age of individuals, young people's access to productive assets is less than for older individuals. Typically, older family members accumulate more savings that facilitate their engagement in farming and are most likely to have acquired land through inheritance than their

younger counterparts. Others retire from other jobs and engage in farming when old (see Bojnec, Dries & Swinnen, 2003). While young people are at a disadvantage with respect to access to productive assets generally, the problem is more pronounced for women (Messerli et al. 2015). Subakanya (2015) finds a negative relationship between being young and participation in farming. Older household members were also more likely to engage in farming activities. Further, young adults (aged 15-24 years) are often disinterested in crop or livestock production (Kafle, Paliwal & Benfica, 2019). In addition, Tocco, Bailley and Davidova, (2013) show that married individuals are more likely to engage in farming; this group is less likely to be mobile given their family responsibilities.

The availability of markets conditions the nature of production that can take place in an area and the productivity of farming. This is true for input, output, finance, and services markets. Agricultural financing is crucial for facilitating productive investments on-farm (e.g., for inputs such as land, labor, seeds, fertilizers, and herbicides, as well as for farm equipment), and can be a barrier to participation in farming if unavailable. A lack of finance has been identified as one of the key challenges being faced by young people (Frost & Sawa, 2017; Ibidapo et al.2017). The financial challenge within households can be met through formal and informal credit. Similarly, output markets incentivize production and shape the nature of production. Where output markets are thin or under-developed, participation in agricultural production is relatively less and usually restricted to subsistence farming. Market access indicators that have been shown to influence the decision to participate in crop and livestock production in general include the density of tarred or feeder roads, and travel time or distance to district headquarters (Yeboah & Jayne, 2017).

In many rural areas, family labor is the main source of labor for agricultural production. Thus, smaller households that cannot afford to hire labor are less likely to participate in on-farm

agricultural production, all else equal. Where possibilities for the utilization of hired labor or mechanization exist, this constraint is less of a concern (Agwu, Nwankwo & Anyanu 2014). On the other hand, individuals in small households may have no choice but to participate in agricultural production. When labor is a constraint, nonland productive assets such as tractors, planters, and ridgers facilitate land preparation and planting. The lack of these may relegate the household to other non-farm enterprises. One may also liquidate these assets to finance production. Also, note that wealth may facilitate the exit out of farming, as discussed below.

Locational factors also determine employment outcomes for young people. Proximity to urban areas restricts the nature of production activities available to capital intensive enterprises given the land constraints in these areas. However, there are more opportunities for off-farm wage employment due to the lower transport costs that make off-farm opportunities more lucrative in these areas (O'Kelly, & Bryan 1996; Gollin & Rogerson 2014; Jin & Deininger 2008; Isgut 2004). This means that young people residing near urban areas are less likely to engage in the raising of livestock and growing field crops. Instead, they are likely to engage in wage employment off the farm. Subject to capital constraints, they may also engage in capital-intensive horticultural production.

Employment in Off-farm Wage Activities in the Agricultural Sector

The education level attained by an individual influences off-farm wage employment in that it imposes limits to what off-farm wage-earning activities one can engage in. Participation in off-farm wage employment is an increasing function of education (Haggblade, Hazell & Reardon 2010; Barrett, Reardon, and Webb 2001; Bojnec & Dries 2005). Thus, we expect more skilled and better-educated individuals to often have access to more wage-earning opportunities off the farm.

Related to one's gender, the nature of enterprises available may influence who gets hired. Some farm operations may prefer women since they have a comparative advantage over men in, for example, the picking of cotton, eggs, or tea, while men may have a comparative advantage in the picking of orchard fruits (Qian, 2008). Further, competing household responsibilities for women such as the raising of children and other household chores may limit their engagement in wage employment off the farm.

Concerning age, some young people's participation may be limited because they may still be in school and thus restricted in what they can participate in. Participation is also dependent on factors outside the control of an individual. If off-farm opportunities do not exist in the area or transaction costs limit participation, individuals are unlikely to participate.

Also, peer effects tend to influence the dynamics of labor supply in a manner akin to how they influence technology adoption in rural areas (see Foster & Rosenzweig 2010). Information exchange from one's social network influences choices about which form of employment to partake in (e.g., see Murendo et al. 2018). Proximity to urban areas and the state of public infrastructure for transport tend to influence the transaction costs associated with participating in off-farm labor supply (O'Kelly & Bryan 1996; Isgut, 2004). The wage differential between farm and off-farm wages may induce movement into the off-farm sector by individuals if the wages are more lucrative (Loughrey et al. 2016). Poorer households are less likely to supply labor off-farm, and this is partly related to the need to overcome transaction cost barriers related to participation (Reddy & Findeis 1988).

Table A1: Industries used to Classify Employment Variables in Agriculture

Crop and animal production, hunting, and related service activities
Fishing and aquaculture
Manufacture of food products
Manufacture of tobacco products
Manufacture of leather and related products
Veterinary activities
Processing and preserving of meat
Processing and preserving of fish, crustaceans, and mollusks
Processing and preserving of fruit and vegetables
Manufacture of vegetable and animal oils and fats
Manufacture of dairy products
Manufacture of grain mill products, starches, and starch products
Manufacture of other food products
Manufacture of prepared animal feeds
Retail sale of food, beverages, and tobacco in specialized stores

Source: Tanzania National Panel Surveys

APPENDIX B: Diagnostic and Robustness Test Results

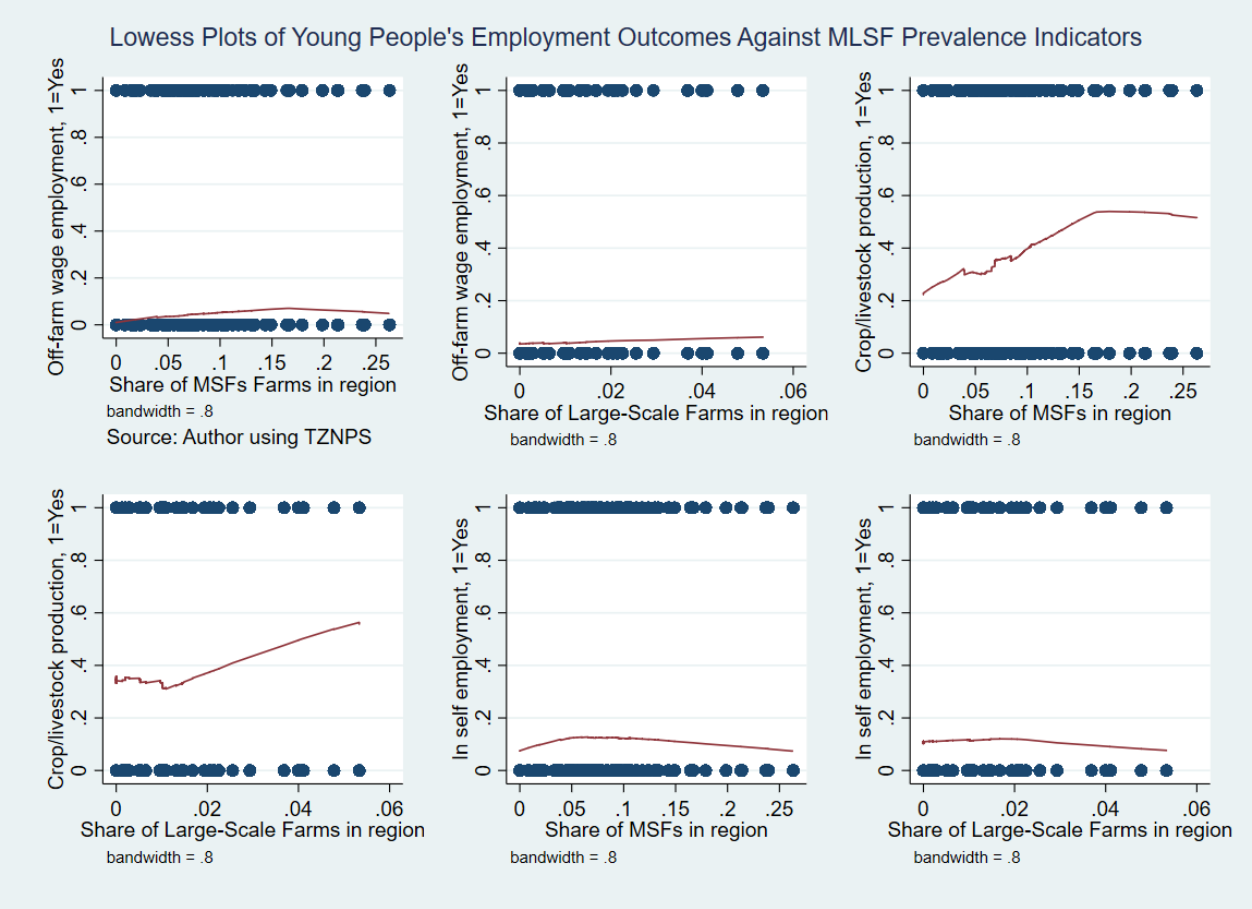

Figure B1: Relationship between the Prevalence of MLSFs and Young People's Employment Outcomes
Source: Author's illustration using the TZNPS

Table B1: Tests for Attrition Bias

VARIABLES	(1) Employment in Agricultural Production	(2) Self-employment in agriculture	(3) Off-farm wage employment in agriculture	(4) Employment in agriculture
1=reinterviewed in period t+1	0.018	0.001	-0.008	0.016
	(0.013)	(0.009)	(0.007)	(0.013)
Proportion of cropping HHs in region that are large-scale	-1.822	2.826***	0.134	0.577
	(1.288)	(0.868)	(0.677)	(1.307)
Squared proportion of cropping HHs in region that are large-scale	-69.564**	-69.127***	-4.148	-108.003***
	(32.554)	(20.800)	(16.463)	(33.193)
Proportion of cropping HHs in region that are medium scale	1.148***	0.184	0.551***	1.452***
	(0.302)	(0.206)	(0.142)	(0.303)
Squared proportion of cropping HHs in region that are medium scale	-0.249	-0.924	-1.666***	-1.967
	(1.219)	(0.813)	(0.598)	(1.240)
Age of Individual, Years	0.003***	0.012***	0.004***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)
1=Individual is married	0.136***	0.044***	-0.013*	0.140***
	(0.014)	(0.011)	(0.008)	(0.014)
1=Individual is male	0.037***	0.010	0.046***	0.058***
	(0.012)	(0.008)	(0.006)	(0.011)
Head's education level, Years	-0.005***	0.002***	-0.001***	-0.003***
	(0.001)	(0.000)	(0.000)	(0.001)
Household Size, Number	0.008***	-0.005***	-0.002***	0.004***
	(0.001)	(0.001)	(0.000)	(0.001)
1=Female headed household	-0.047***	0.021**	-0.008	-0.031**
	(0.014)	(0.009)	(0.008)	(0.015)
1=Household experienced a shock in the last five years	0.015	-0.001	0.016***	0.008
	(0.012)	(0.008)	(0.006)	(0.012)
Household's distance in (Kms) to nearest population center with +20,000 people	0.003***	-0.000**	0.000***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.232***	-0.181***	-0.058***	0.106***
	(0.039)	(0.025)	(0.020)	(0.039)
Observations	10,991	10,991	10,991	10,991
R-squared	0.147	0.090	0.034	0.136
Individual FE	No	No	No	No
Standard error cluster level	Individual	District	Individual	Individual
Sample	Pooled	Pooled	Pooled	Pooled
Year FE	Yes	Yes	Yes	Yes
Survey month dummies	Yes	Yes	Yes	Yes
Estimator	POLS	POLS	POLS	POLS

Notes: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table B2: Fixed Effects Model Results (Average Marginal Effects)

VARIABLES	(1) Employed in own/Family Crop/Livestock Production	(2) Self-employed in agriculture	(3) Off-farm wage employment in agriculture	(4) Employment in Agriculture
Proportion of cropping HHs in region that are large-scale	0.904	0.412	0.246	1.124
	(0.715)	(0.493)	(0.386)	(0.736)
Proportion of cropping HHs in region that are medium scale	-0.438*	-0.178	0.084	-0.471*
	(0.238)	(0.153)	(0.110)	(0.248)
Age of Individual, Years	-0.008	0.001	-0.002	-0.002
	(0.006)	(0.004)	(0.003)	(0.006)
1=Individual is Married	0.036**	0.017	-0.024**	0.036**
	(0.016)	(0.014)	(0.011)	(0.017)
Household head's education level	0.000	0.001	0.000	0.001
	(0.001)	(0.001)	(0.000)	(0.001)
Household Size, Number	0.002	-0.006***	-0.001	-0.000
	(0.002)	(0.002)	(0.001)	(0.003)
1=Female headed household	-0.027	0.022	0.019	0.021
	(0.023)	(0.017)	(0.014)	(0.025)
1=Household experienced a shock in the last five years	0.022*	-0.007	0.006	0.019
	(0.012)	(0.008)	(0.006)	(0.012)
HH Distance in (KMs) to Nearest Population Center with +20,000	0.002***	-0.001	0.000	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
No. of Observations	12,540	12,540	12,540	12,540
Survey month dummies	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Standard error cluster level	Individual	Individual	Individual	Individual
Sample	Panel	Panel	Panel	Panel
Year fixed effects	Yes	Yes	Yes	Yes
Estimator	OLS	OLS	OLS	OLS

Notes: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table B3: Correlated Random Effects Probit Model Results (Average Marginal Effects)

Variables	(1) Employed in own/Family Crop/Livestock Production	(2) Self-employed in agriculture	(3) Off-farm wage employment in agriculture	(4) Employment in agriculture
Proportion of cropping HHs in region that are large-scale	-0.875	0.993**	-0.088	-0.038
	(0.727)	(0.488)	(0.395)	(0.738)
Proportion of cropping HHs in region that are medium-scale	-0.560**	-0.194	0.154	-0.571**
	(0.232)	(0.158)	(0.116)	(0.240)
Age of Individual, Years	-0.007	0.001	-0.002	-0.002
	(0.006)	(0.004)	(0.003)	(0.006)
1= Individual is married	0.021	0.019	-0.021**	0.025
	(0.016)	(0.012)	(0.009)	(0.017)
Head's education level, Years	-0.005***	0.002***	-0.001***	-0.003***
	(0.001)	(0.000)	(0.000)	(0.001)
Household Size, Number	0.002	-0.008***	-0.001	-0.000
	(0.003)	(0.002)	(0.001)	(0.003)
1=Female headed household	-0.040*	0.021	0.016	0.008
	(0.023)	(0.019)	(0.015)	(0.024)
1=Household experienced a shock in the last five years	0.023**	-0.007	0.004	0.021*
	(0.012)	(0.008)	(0.007)	(0.012)
Household's distance in (KMs) to nearest Population Center with +20,000	0.002***	-0.001	0.000*	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
CRE Means for all RHS variables	Yes	Yes	Yes	Yes
No. of Observations	12,540	12,540	12,540	12,540
Individual fixed effects	No	No	No	No
Survey month dummy variables	Yes	Yes	Yes	Yes
Standard error cluster level	Individual	District	Individual	Individual
Sample	Panel	Panel	Panel	Panel
Year fixed effects	Yes	Yes	Yes	Yes
Estimator used	CRE-Probit	CRE-Probit	CRE-Probit	CRE-Probit

Notes: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table B4: Econometric Results- Youth

VARIABLES	(1) Employed in own/Family Crop/Livestock Production	(2) Self employed in agriculture	(3) Off-farm wage employment in agriculture	(4) Employment in agriculture
Proportion of cropping HHs in region that are large-scale in size	-1.302	0.652	-0.217	-1.187
	(0.981)	(0.498)	(0.465)	(1.021)
Proportion of cropping HHs in region that are medium-scale in size	-0.326	-0.142	0.125	-0.190
	(0.312)	(0.161)	(0.118)	(0.317)
Age of Individual, Years	0.011	0.006	0.001	0.019**
	(0.010)	(0.005)	(0.004)	(0.010)
1=Married	0.011	-0.013	-0.004	0.002
	(0.025)	(0.013)	(0.013)	(0.026)
Head's education level	-0.005***	0.001**	-0.001**	-0.004***
	(0.001)	(0.000)	(0.000)	(0.001)
Household Size, Number	0.005	-0.004**	-0.000	0.001
	(0.003)	(0.002)	(0.002)	(0.003)
1=Female headed household1	-0.065**	0.001	0.009	-0.039
	(0.030)	(0.018)	(0.018)	(0.031)
1=Male	0.152	0.160	-0.000	0.251***
	(0.132)	(0.103)	(0.012)	(0.097)
1=Household experienced a shock in the last five years	0.023	-0.014	-0.004	0.006
	(0.016)	(0.009)	(0.008)	(0.017)
Distance in (Kms) to Nearest Population Center with +20,000	0.002***	-0.000	0.001**	0.001**
	(0.001)	(0.000)	(0.000)	(0.001)
Observations	7,173	7,173	7,173	7,173
Individual FE	No	No	No	No
Survey month dummy variables	Yes	Yes	Yes	Yes
SE Cluster	Individual	Individual	Individual	Individual
CRE means for all RHS variables	Yes	Yes	Yes	Yes
Sample	Balanced Panel	Balanced Panel	Balanced Panel	Balanced Panel
Year Fixed Effects	Yes	Yes	Yes	Yes
Estimator	CRE-Probit	CRE-Probit	CRE-Probit	CRE-Probit

Notes: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1, all parameter estimates are average marginal effects.

Table B5: Econometric Results- Young Adults

VARIABLES	(1) Employed in own/Family Crop/Livestock Production	(2) Self employed in agriculture	(3) Off-farm wage employment in agriculture	(4) Employment in agriculture
Proportion of cropping HHs in region that are large-scale in size	0.329	1.115	0.123	1.977*
	(1.143)	(0.968)	(0.695)	(1.141)
Proportion of cropping HHs in region that are medium-scale in size	-0.891**	-0.218	0.231	-1.147***
	(0.373)	(0.309)	(0.226)	(0.393)
Age of Individual, Years	-0.011	0.005	-0.002	0.002
	(0.008)	(0.007)	(0.005)	(0.009)
1=Married	0.037	0.042**	-0.037**	0.049**
	(0.022)	(0.021)	(0.016)	(0.024)
Head's education level	-0.005***	0.004***	-0.002***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Household Size, Number	-0.004	-0.013***	-0.003	-0.006
	(0.005)	(0.005)	(0.003)	(0.005)
1=Female headed household	0.016	0.051	0.050	0.086**
	(0.040)	(0.040)	(0.033)	(0.037)
1=Male	0.451***	-0.036	-0.125	0.386***
	(0.093)	(0.125)	(0.188)	(0.145)
1=Household experienced a shock in the last five years	0.023	-0.006	0.017	0.033*
	(0.018)	(0.016)	(0.011)	(0.019)
Distance in (Kms) to Nearest Population Center with +20,000	0.002**	-0.001	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Time variable = 2	0.012	0.036*	0.011	0.031
	(0.024)	(0.020)	(0.014)	(0.026)
Time variable = 3	0.046	0.017	0.011	0.053
	(0.035)	(0.029)	(0.019)	(0.037)
Observations	5,367	5,367	5,367	5,367
Individual FE	No	No	No	No

Notes: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1, all parameter estimates are average marginal effects.

Table B5 (cont'd)

VARIABLES	(1) Employed in own/Family Crop/Livestock Production	(2) Self employed in agriculture	(3) Off-farm wage employment in agriculture	(4) Employment in agriculture
SE Cluster level	Individual	District	Individual	Individual
Survey Month	Yes	Yes	Yes	Yes
CRE means for all RHS variables	Yes	Yes	Yes	Yes
Sample	Balanced Panel	Balanced Panel	Balanced Panel	Balanced Panel
Year Fixed Effects	Yes	Yes	Yes	Yes
Estimator	CRE-Probit	CRE-Probit	CRE-Probit	CRE-Probit

Notes: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1, all parameter estimates are average marginal effects.

BIBLIOGRAPHY

BIBLIOGRAPHY

Aabø, E. and Kring, T., 2012. The political economy of large-scale agricultural land acquisitions: Implications for food security and livelihoods/employment creation in rural Mozambique. United Nations Development Programme Working Paper, 4, pp.1-61.

Abay, K.A., Chamberlin, J. and Berhane, G., 2021. Are land rental markets responding to rising population pressures and land scarcity in sub-Saharan Africa?. Land use policy, 101, p.105139.

Adelaja, A. and George, J., 2020. Is Youth Unemployment Related to Domestic Terrorism? *Perspectives on Terrorism*, *14*(5), pp.41-62.

Agwu, N.M., Nwankwo, E.E. and Anyanwu, C.I., 2014. Determinants of agricultural labour participation among youths in Abia State, Nigeria. *International Journal of Food and Agricultural Economics (IJFAEC)*, *2*(1128-2016-92018), pp.157-164.

Angrist, J.D. and Pischke, J.S., 2008. Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Anti, S., 2021. Land grabs and labor in Cambodia. *Journal of Development Economics*, *149*, p.102616.

Barrett, C.B., Reardon, T. and Webb, P., 2001. Nonfarm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications. *Food policy*, *26*(4), pp.315-331.

Barrett, C.B., Sherlund, S.M. and Adesina, A.A., 2008. Shadow wages, allocative inefficiency, and labor supply in smallholder agriculture. *Agricultural Economics*, *38*(1), pp.21-34.

Baumgartner, P., Von Braun, J., Abebaw, D. and Müller, M., 2015. Impacts of large-scale land investments on income, prices, and employment: Empirical analyses in Ethiopia. *World Development*, *72*, pp.175-190.

Bojnec, Š. and Dries, L., 2005. Causes of Changes in Agricultural Employment in Slovenia: Evidence from Micro-data. *Journal of Agricultural Economics*, *56*(3), pp.399-416.

Bojnec, S., Dries, L. and Swinnen, J.F., 2003. *Human capital and labor flows out of the agricultural sector: Evidence from Slovenia* (No. 1002-2016-78294).

Brooks, K., Zorya, S. and Gautam, A., 2012. Jobs for Africa's youth. *International Food Policy Research Institute (IFPRI)*, pp.49-57.

Bukari, K.N. and Kuusaana, E.D., 2018. Impacts of large-scale land holdings on Fulani pastoralists in the Agogo Traditional Area of Ghana. *Land use policy*, *79*, pp.748-758.

Burke, W.J., Jayne, T.S. and Sitko, N.J., 2020. Do medium-scale farms improve market access conditions for Zambian smallholders? *Journal of Agricultural Economics*, *71*(2), pp.517-533.

Chamberlin, J., and Ricker-Gilbert, J.E., 2015. *What are the drivers of rural land rental markets in sub-Saharan Africa, and how do they impact household welfare? Evidence from Malawi and Zambia* (No. 1008-2016-80200).

Chamberlin, J., and Jayne, T.S., 2020. Does farm structure affect rural household incomes? Evidence from Tanzania. *Food Policy*, *90*, p.101805.

Chinigò, D., 2015. The new harvest agrarian policies and rural transformation in Southern Africa. *Afriche e Orienti*.

Christiaensen, L., Rutledge, Z., and Taylor, J.E., 2020. *The future of work in agriculture: Some reflections*. The World Bank.

Chu, J., Young, K. and Phiri, D., 2015. Large-scale land acquisitions, displacement and resettlement in Zambia.

Cordes, K.Y., Östensson, O. and Toledano, P., 2016. Employment from mining and agricultural investments: how much myth, how much reality?

Deininger, K. and Byerlee, D., 2011. *Rising global interest in farmland: can it yield sustainable and equitable benefits?*. World Bank Publications.

Deininger, K., & Xia, F. 2016. Quantifying Spillover Effects from Large Land-based Investment: The Case of Mozambique. *World Development*, *87*, 227–241. https://doi.org/10.1016/j.worlddev.2016.06.016

Doss, C., Meinzen-Dick, R. and Bomuhangi, A., 2014. Who owns the land? Perspectives from rural Ugandans and implications for large-scale land acquisitions. *Feminist economics*, *20*(1), pp.76-100.

de Brauw, A., 2019. Rural youth: determinants of migration throughout the world. *IFAD Research Series*, *55*.

Debonne, N., van Vliet, J., Heinimann, A. and Verburg, P., 2018. Representing large-scale land acquisitions in land use change scenarios for the Lao PDR. *Regional environmental change*, *18*(6), pp.1857-1869.

Foster, A.D. and Rosenzweig, M.R., 2010. Microeconomics of technology adoption. *Annu. Rev. Econ.*, *2*(1), pp.395-424.

Frost, O. and Sawa, P., 2017. Young people: energising African agriculture. *Spore*, (184), pp.4-7.

Gollin, D. and Rogerson, R., 2014. Productivity, transport costs and subsistence agriculture. *Journal of Development Economics*, *107*, pp.38-48.

Haggblade, S., Hazell, P. and Reardon, T., 2010. The rural non-farm economy: Prospects for growth and poverty reduction. *World development*, *38*(10), pp.1429-1441.

Henley, G., 2017. *Review of social issues for large-scale land investment in Zambia* (No. 2017/42). WIDER Working Paper.

Himelein, K., 2014. Weight calculations for panel surveys with subsampling and split-off tracking. *Statistics and Public Policy*, *1*(1), pp.40-45.

Ibidapo, I., Faleye, O.M., Akintade, T.F., Oso, O.P. and Owasoyo, E.O., 2017. Determinants of youth participation in dry season vegetables cultivation in urban areas of Ondo State, Nigeria. *Journal of Agricultural Science and Practice*, *2*, pp.109-114.

Isgut, A.E., 2004. Non-farm income and employment in rural Honduras: assessing the role of locational factors. *Journal of Development Studies*, *40*(3), pp.59-86.

Jayne, T.S., Chamberlin, J., Traub, L., Sitko, N., Muyanga, M., Yeboah, F.K., Anseeuw, W., Chapoto, A., Wineman, A., Nkonde, C. and Kachule, R., 2016. Africa's changing farm size distribution patterns: the rise of medium-scale farms. *Agricultural Economics*, *47*(S1), pp.197-214.

Jayne, T.S., Chamberlin, J., and Headey, D.D., 2014. Land pressures, the evolution of farming systems, and development strategies in Africa: A synthesis. *Food policy*, *48*, pp.1-17.

Jayne, T.S., Chapoto, A., Sitko, N., Nkonde, C., Muyanga, M. and Chamberlin, J., 2014. Is the scramble for land in Africa foreclosing a smallholder agricultural expansion strategy? *Journal of International Affairs*, pp.35-53.

Jayne, T.S., Muyanga, M., Wineman, A., Ghebru, H., Stevens, C., Stickler, M., Chapoto, A., Anseeuw, W., Van der Westhuizen, D. and Nyange, D., 2019. Are medium-scale farms driving agricultural transformation in sub-Saharan Africa? *Agricultural Economics*, *50*, pp.75-95.

Jin, S. and Deininger, K., 2009. Key constraints for rural non-farm activity in Tanzania: Combining investment climate and household surveys. *Journal of African Economies*, *18*(2), pp.319-361.

Kassam, A.H. and Fischer, G.W., 1991. *Agro-Ecological Land Resources Assessment for Agricultural Development Planning: A Case Study of Kenya: Resources Data Base and Land Productivity*. Food & Agriculture Org.

Kafle, K., Paliwal, N. and Benfica, R., 2019. IFAD Research Series No. 36-Who Works in Agriculture? Exploring the Dynamics of Youth Involvement in the Agri-Food Systems of Tanzania and Malawi. *Exploring the Dynamics of Youth Involvement in the Agri-Food Systems of Tanzania and Malawi (April 4, 2019). IFAD Research Series*, *36*.

Land Matrix. (n.d.). Africa. Retrieved May 2021, from Land Matrix: https://landmatrix.org/observatory/africa/

Lanjouw, P. and Shariff, A., 2004. Rural non-farm employment in India: Access, incomes and poverty impact. *Economic and Political Weekly*, pp.4429-4446.

Lay, J., Nolte, K. and Sipangule, K., 2020. Large-scale farms in Zambia: Locational patterns and spillovers to smallholder agriculture. *World Development*, p.105277.

Lin, W., and Wooldridge, J.M., 2019. Testing and correcting for endogeneity in nonlinear unobserved effects models. In *Panel Data Econometrics* (pp. 21-43). Academic Press.

Liverpool-Tasie, L.S.O., Nuhu, A.S., Awokuse, T., Jayne, T., Muyanga, M., Aromolaran, A. and Adelaja, A., 2020. Spillover Effects of Medium-Scale Farms on Smallholder Behaviour and Welfare: Evidence from Nigeria.

Loughrey, J., Hennessy, T., Hanrahan, K., Donnellan, T., Raimondi, V. and Olper, A., 2013. *Determinants of Farm Labour Use: A Comparison between Ireland and Italy* (No. 545-2016-38731).

Manda, S., Tallontire, A.M. and Dougill, A.J., 2017. Large-scale Agricultural Investments and Institutions in Zambia: Patterns, Possibilities and Barriers. *SRI Paper*, *106*, pp.1753-1330.

Meinzen-Dick, R., 2007. Beyond panaceas in water institutions. *Proceedings of the national Academy of sciences*, *104*(39), pp.15200-15205.

Murendo, C., Wollni, M., de Brauw, A. and Mugabi, N., 2018. Social network effects on mobile money adoption in Uganda. *The Journal of Development Studies*, *54*(2), pp.327-342.

Matenga, C.R. and Hichaambwa, M., 2017. Impacts of land and agricultural commercialisation on local livelihoods in Zambia: Evidence from three models. *The Journal of Peasant Studies*, *44*(3), pp.574-593.

Messerli, P., Peeters, A., Schoenweger, O., Nanhthavong, V. and Heinimann, A., 2015. Marginal lands or marginal people? Analysing key processes determining the outcomes of large-scale land acquisitions in Lao PDR and Cambodia. *International Development Policy| Revue internationale de politique de dévelopement*, *6*(6).

Nolte, K., 2014. Large-scale agricultural investments under poor land governance in Zambia. *Land use policy*, *38*, pp.698-706.

Nolte, K. and Ostermeier, M., 2017. Labour market effects of large-scale agricultural investment: conceptual considerations and estimated employment effects. *World Development*, *98*, pp.430-446.

O'Kelly, M. and Bryan, D., 1996. Agricultural location theory: von Thunen's contribution to economic geography. *Progress in Human geography*, *20*(4), pp.457-475.

Qian, N., 2008. Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance. *The Quarterly Journal of Economics*, *123*(3), pp.1251-1285.

Reddy, V.K. and Findeis, J.L., 1988. Determinants of Off-Farm Labor Force Participation: Implications for Low Income Farm Families. *Applied Economic Perspectives and Policy*, *10*(1), pp.91-102.

Republic of Tanzania. (2016). National Strategy for Youth Involvement in Agriculture (2016-2021). Ministry of Agriculture, Livestock and Fisheries.

Republic of Zambia, Government of. 2017. 7 National Development Plan:2017-2021. Ministry of Finance and National Development Planning, Lusaka.

Republic of Zambia, Government of.2016. Second National Agricultural Policy. Ministry of Agriculture and Ministry of Fisheries and Livestock, Lusaka.

Ricker-Gilbert, J., Chamberlin, J., Kanyamuka, J., Jumbe, C.B., Lunduka, R. and Kaiyatsa, S., 2019. How do informal farmland rental markets affect smallholders' well-being? Evidence from a matched tenant–landlord survey in Malawi. Agricultural Economics, 50(5), pp.595-613.

Sharma, U.M., 1982. Women's participation in agriculture in India. *Current Anthropology*, *23*(2), pp.194-195.

Sitko, N. and Chamberlin, J., 2015. The anatomy of medium-scale farm growth in Zambia: What are the implications for the future of smallholder agriculture?. *Land*, *4*(3), pp.869-887.

Schoneveld, G.C., 2014. The geographic and sectoral patterns of large-scale farmland investments in sub-Saharan Africa. *Food Policy*, *48*, pp.34-50.

Subakanya, M., 2015. *Changes in the age and gender composition of agricultural participation in Zambia: implications for economic policy* (Doctoral dissertation, University of Pretoria).

The Oakland Institute. (2011). Agrisol Energy and Pharos global agriculture fund's land deal in Tanzania. Understanding land investment deals in Africa. Land Deal Brief. Oakland, June 2011.

Tocco, B., Bailey, A. and Davidova, S., 2013. *Determinants to leave agriculture and change occupational sector: Evidence from an enlarged EU* (No. 545-2016-38697).

United Nations., 2015. Youth population trends and sustainable development. United Nations, Department of Economic and Social Affairs. *Population Division*.

Williams, T.O., Gyampoh, B., Kizito, F. and Namara, R., 2012. Water implications of large-scale land acquisitions in Ghana. *Water Alternatives*, *5*(2).

Wineman, A. and Liverpool-Tasie, L.S., 2017a. Land markets and the distribution of land in northwestern Tanzania. *Land Use Policy*, *69*, pp.550-563.

Wineman, A. and Liverpool-Tasie, L.S.O., 2017b. Land markets and land access among female-headed households in northwestern Tanzania. *World Development*, *100*, pp.108-122.

Wineman, A. and Liverpool-Tasie, L.S., 2018. Land markets and migration trends in Tanzania: A qualitative-quantitative analysis. *Development Policy Review*, *36*, pp.O831-O856.

Wineman, A., Jayne, T.S., Modamba, E.I. and Kray, H., 2021. Characteristics and spillover effects of medium-scale farms in Tanzania. *The European Journal of Development Research*, *33*(6), pp.1877-1898.

Wooldridge, J.M., 2010. *Econometric analysis of cross section and panel data*. MIT press.

Zhan, J., Mirza, H. and Speller, W., 2015. The Impact of Larger-Scale Agricultural Investments on Communities in South-East Asia: A First Assessment. *International Development Policy*, *6*, pp.81-107.

Zaehringer, J.G., Wambugu, G., Kiteme, B. and Eckert, S., 2018. How do large-scale agricultural investments affect land use and the environment on the western slopes of Mount Kenya? Empirical evidence based on small-scale farmers' perceptions and remote sensing. *Journal of Environmental Management*, *213*, pp.79-89.