

KIEL WORKING PAPER

**Large-Scale Farms and
Smallholders:
Evidence from Zambia**



Nr. 2098 January 2018

*Jann Lay, Kerstin Nolte, Kacana
Sipangule*

Large-Scale Farms and Smallholders: Evidence from Zambia

Jann Lay^{ψζ}, Kerstin Nolte^ψ and Kacana Sipangule^{§ζ*}

Abstract

In light of the surge in large-scale farms in developing countries, concerns have been raised that smallholders may be negatively affected. There is, however, very little evidence beyond case studies to support these claims. Drawing on nationally representative household data sets and an inventory of large-scale farms in Zambia, this study investigates the relationship between large-scale farms and smallholders. First, we analyse the geographical contexts of wards that host large-scale farms and show that large-scale farms are found in wards with good infrastructure and soil quality. Second, we adopt a difference-in-differences approach to estimate the impacts of large-scale farms on smallholders' area cultivated, maize yields, and access to fertiliser. We find that smallholders in wards with large-scale farms increase their area cultivated and maize yields, but have lower fertiliser usage. This hints at positive spillovers at the extensive and intensive margins but not at improved access to agricultural inputs. It is likely that these results are also driven by the emergence of medium-scale farms in these regions.

Keywords: large-scale farms, yields, smallholders, spillovers, Zambia

JEL Classification Codes: Q12, Q15, Q18

^ψ GIGA German Institute of Global and Area Studies

^ζ University of Goettingen

[§] Kiel Institute for the World Economy

* Correspondence: Kacana Sipangule, Kiel Institute for the World Economy, Kiellinie 66, 24105 Kiel, Germany; e-mail: kacana.sipangule@ifw-kiel.de.

1. Introduction

Large-scale agricultural investments (LSAIs) gained international prominence in the early part of this century, when they were depicted as “land grabs” by the media and some civil society organisations.¹ These reports highlighted the opaque acquisition processes and adverse impacts on surrounding local communities (Cotula et al., 2009). Since then, several scholarly works have addressed the impacts and spillovers of LSAIs on neighbouring smallholders through case studies.² While these case studies have provided important insights into individual LSAIs and have thereby furthered the debate on the impacts of LSAIs, they are often too specific and lack external validity beyond their study regions.

More recently, a new literature that goes beyond case studies to provide systematic evidence of the impacts of LSAIs on smallholders at the national or regional level has emerged. For instance, Deininger and Xia (2016) have combined information on the location and start dates of large farms with smallholder surveys to quantitatively assess spillover effects from large land-based investments in Mozambique. Ali et al. (2017) have conducted a similar analysis based on a large-farm census and smallholder surveys in Ethiopia, while Herrmann (2017) has looked at household income and income poverty among out-growers, wage employees, and non-participants of LSAIs in Tanzania. Ahlerup and Tengstam (2015) have used three waves of panel

¹ This work was supported by the German Federal Ministry of Economic Cooperation and Development (BMZ), (project: “Large-Scale Land Acquisitions: Data, Patterns, Impacts, and Policies”) and the African Growth and Development Policy Modeling Consortium (AGRODEP) (2012 GAPS IN RESEARCH GRANT). An earlier version of this paper was circulated under the title: “The Impact of Foreign Large-Scale Land Acquisitions on Smallholder Productivity: Evidence from Zambia.” We are grateful to Viviane Griesinger for her support in creating the maps for this paper and to Rainer Thiele and Johanna Bodewig for their valuable comments. Moreover, we are thankful for feedback received during several presentations, including presentations at the Annual World Bank Conference on Land and Poverty 2016, the PEGNet Conference 2016, the Open Science meeting of the Global Land Programme 2016, and the Young Researchers Seminar at the Kiel Institute for the World Economy.

² For recent literature reviews and meta-analyses of these case studies, refer to Oberlack et al. (2016) and Schoneveld (2016).

data to analyse how large-scale farms affect smallholders' wage incomes at the district level in Zambia.

Our paper provides an innovative contribution to this literature with a context-sensitive quantitative study. We combine a systematic analysis of the geographic characteristics of wards that host large-scale farms with a difference-in-differences and fixed-effects analysis of the impacts of large-scale farms on smallholders' agricultural outcomes for Zambia. Zambia is particularly interesting for our study as it has a long history of large-scale farms that have coexisted alongside smallholder communities (Chu, 2013). However, in recent years, similarly to many other developing countries, Zambia has experienced a sudden increase in the demand for land to be used for large-scale agricultural purposes. The Land Matrix estimates that 26 deals covering an area of 389,774 hectares have been concluded since 2000 (Harding et al., 2016).³

Our paper contributes to the literature in three ways: First, it undertakes a detailed assessment of the large-scale-farm sector in Zambia and the study of the geographical context, thereby providing a comprehensive picture of large-scale agriculture in Zambia and challenging some of the commonly held perceptions surrounding these farms. Second, it assesses the impacts of large-scale farms on the agricultural outcomes of smallholders in Zambia for the first time. In addition, it covers a larger time-span than earlier studies and uses a difference-in-differences analysis with fixed effects for wards, which are the smallest administrative units in Zambia. Finally, it combines the analysis of the geographic context with a quantitative impact assessment, allowing for a meaningful interpretation of the results from the impact study based on a thorough

³ The Land Matrix is a database that provides information on land acquisitions that cover areas larger than 200 hectares and that have been set up since the year 2000. It is a partnership between the International Land Coalition, the Centre de Coopération Internationale en Recherche Agronomique pour le Développement, the Centre of Development and Environment at the University of Bern, the German Institute of Global and Area Studies, and the GIZ (Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH).

understanding of the context in which farms operate. This link between the locations of large-scale farms and the impacts of these farms on surrounding smallholders has been lacking in previous studies.

The results from the geographic analysis nicely highlight the importance of access to infrastructure and the agglomeration of farms in certain regions. The regression results show that smallholders' farm sizes increase in wards with large-scale farms. This could be indicative of the fact that smallholders increase their production at the extensive margin, or it may be a result of the increasing trend towards land consolidation in Zambia. Moreover, we find a reduction in fertiliser usage and an increase in maize yields in wards with large-scale farms. While the results on fertiliser are difficult to interpret in the Zambian context, where government subsidy programmes simultaneously affect smallholder fertiliser use, the increased yields are indicative of increasing smallholder productivity at the intensive margin.

The remainder of this paper is structured as follows. The next section presents the data used for the analysis. In Section 3, we provide a descriptive overview of large-scale farms in Zambia and analyse the geographical contexts of the wards that host large-scale farms. In Section 4, we analyse the impacts of large-scale farms on smallholders. In particular, we discuss the hypotheses, estimation strategy, econometric considerations, results, and robustness checks. The final section discusses our results and concludes.

2. Data

To analyse the impacts of large-scale farms on smallholders' agricultural outcomes over time, we combine several data sets that provide us with information on large-scale farms, spatial characteristics, and smallholders.

The large-scale farm data has been obtained from the Post-Harvest Survey (PHS) for Large-Scale Agricultural Holdings, which is a census on all large-scale farms (defined as farms that are larger than 20 hectares) collected by the Zambian Central Statistical Office (CSO). The large-scale farm data contains information on the crops grown, the area cultivated and harvested, the fertiliser used, the livestock reared, and the value of sales (CSO, 2004). In 2013/14, a total of 1,102 large-scale farms were surveyed, of which we use a subset of 834 large-scale farms that cultivate crops and are owned by non-state actors.⁴ We collaborated with CSO to include additional questions in the 2013/2014 PHS on Large-Scale Agricultural Holdings. The questions gathered information on the year in which the large-scale farms were established (i.e. the year the land was acquired and the year that cultivation started); the development of large-scale farm sizes over time (i.e. the size of the farm upon acquisition, the farm size five years ago, the farm size at the time of the survey); the location of the farm (i.e. which ward the farm is located in); and the countries of origin of the large-scale farm owners.

The spatial data sets, which provide us with information on ward boundaries, land cover, railroad infrastructure, cities, and irrigation data, have been obtained from CSO's Cartography and Mapping Department, the Global Land Cover database, the Digital Chart of the World, the U.S. National Imagery and Mapping Agency, and the Global Map of Irrigation Areas, respectively. Ward-level poverty estimates were obtained from the World Bank (de la Fuente et al., 2015), while the ward population data was obtained from CSO.

The smallholder data has been obtained from the PHS on Small and Medium-Sized Agricultural Holdings, which is collected by the CSO and the Ministry of Agriculture and Livestock. The

⁴ Our main focus is on large-scale farms that engage in crop cultivation. We therefore exclude large-scale farms engaged in animal husbandry from the analysis. We also exclude institutional large-scale farms that are owned or managed by schools, prisons, and other state facilities.

PHS on Small and Medium-Sized Agricultural Holdings is a nationally representative cross-sectional survey that contains information on household-head characteristics, household assets, livestock holdings, and the use of agricultural inputs, as well as the number of crops cultivated, harvested, and sold. The surveys are collected annually between August and September, after the crop harvest period has ended (Megill, 2004).

A two-stage sampling procedure is used to select the households interviewed in the surveys. In the first stage, standard enumeration areas (SEAs) are selected with the probability proportional to size sampling method. In the second stage, the households are stratified by farm-size category, the number of livestock and poultry, and the cultivation of special crops.⁵ Sampling is based on the sampling frame of the current National Census of Housing and Population, and is updated when a new census is collected (Megill, 2004). We use seven cross-sections of the PHS on Small and Medium-Sized Agricultural Holdings collected for the years 2002/2003, 2003/2004, 2004/2005, 2005/2006, 2010/2011, 2011/2012, and 2012/2013.⁶ The first five surveys were collected using the sampling frame of the 2000 National Census of Housing and Population, while the 2011/2012 and 2012/2013 surveys were based on the sampling frame of the 2010 National Census of Housing and Population. This data covers farms with a size of up to 20 hectares. Table A1 in the appendix shows the size distribution of these farms and highlights a clear trend of increases in larger farm sizes.

⁵ The PHS on Small and Medium-Sized Agricultural Holdings identifies eight crops (sorghum, rice, cotton, burley tobacco, Virginia tobacco, sunflower, soybeans, and paprika) that receive special attention in the sample to ensure a representative distribution of crops and to improve the precision of crop area and production estimates (Megill, 2004).

⁶ We take the 2003/2004 data set as the base year for our analysis as ward-level information is not available in the earlier PHS on Small and Medium-Sized Agricultural Holdings. In 2006/2007, the PHS on Small and Medium-Sized Agricultural Holdings did not include a module on the households' asset holdings. In 2007/2008, the PHS was not collected and a few questions on smallholders' harvests were included in the Crop Forecast Survey. The 2008/2009 and 2009/2010 PHS were not collected due to a lack of funding by the government. Thus we exclude the 2006/2007, 2007/2008, 2008/2009, and 2009/2010 PHS on Small and Medium-Sized Holdings from our analysis as the data sets are incomplete or non-existent.

In the sampling frame of the 2000 National Census of Housing and Population, Zambia was subdivided into 1,286 wards; this number increased to 1,421 wards in the 2010 sampling frame. The increase was driven by the partitioning of Northern Province into Muchinga and Northern Province in 2011. This partition resulted in the need to redraw several wards that cut across the boundaries of the new province. In addition, the ward boundaries were also redrawn to facilitate the work of the Electoral Commission of Zambia, which held several local government elections in the period between the two censuses.⁷

The ward level is crucial for our analysis: we investigate the geographical context at the ward level and we link large- and small-scale farms via the location. In order to create a ward panel that accounts for the changes in the ward boundaries over time, we have used the ward shapefiles provided by CSO to create a panel consisting only of wards whose boundaries remained constant across the two sampling frames and wards whose boundaries changed in a consistent and a systematic manner during the same period. Under the latter, we have considered those wards whose boundaries were either split or merged in 2010 and can easily be reconstructed in accordance with the boundaries in the 2000 sampling frame.

Since the main focus of the study is to analyse the impacts of large-scale farms on smallholders in rural regions, we excluded all wards located in urban areas. We also excluded wards that were included in the PHS on Large-Scale Agricultural Holdings but not in the PHS on Small and Medium-Sized Agricultural Holdings. These restrictions resulted in a final panel of 439 wards, of which 70 wards played host to a large-scale farm in 2003, and 87 in 2013. The ward panel is

⁷ The explanations behind the change in ward boundaries were provided by cartographers and census planners at CSO, who were interviewed by one of the authors in February 2014.

matched with the PHS on Small and Medium-Sized Agricultural Holdings, which contains data on 27,109 households that are unevenly distributed across the seven survey periods.

3. Large-Scale Agricultural Investments and Their Host Regions

In this section, we analyse the characteristics of large-scale farms and the wards that host them.

In providing a descriptive overview of large-scale farms in Zambia and analysing their geographical contexts, we question whether some of the commonly held perceptions about large-scale farms hold for Zambia. More specifically, we use the data introduced in the previous section to shed light on the origin of large-scale farm investors, the duration of large-scale farm investments, and the determinants of their locations.

3.1 Who is Investing in Large-Scale Zambian Farms?

Many studies have linked the recent demand for large-scale agriculture with the aftermath of the 2007–2008 food-price crisis (see for instance, De Schutter, 2011). They argue that the crisis led the governments of industrialised nations to outsource food production to land-abundant developing countries. At the same time, food-importing, resource-constrained countries such as the Gulf countries found it attractive to partner with low-income and land-abundant countries for the production of food (De Schutter, 2011).⁸

Table 1 shows the countries of origin of large-scale farm investors in Zambia, the number of large-scale farms under operation per country, and the total area cultivated by these large-scale farms. We find that 84 per cent (705) of the large-scale farms in Zambia are operated by Zambian investors. This does not match the widely held perception that foreign investors from

⁸ The questionnaire asks for the “country of origin of owner of the farm.” Given that some large-scale farms owned by foreign investors may be registered under a Zambian subsidiary company, the numbers reported for Zambian large-scale holdings may be overestimated.

industrialised nations dominate the large-scale agricultural sector, but instead renders partial support for the growing evidence on emerging medium-scale farmers in Zambia.⁹

Table 1. Countries of Origin of Large-Scale Farm Investors

Country	Freq.	Per Cent	Hectares
Australia	1	0.12	31
China	4	0.48	1,405
Cyprus	1	0.12	800
Denmark	1	0.12	1,080
Germany	3	0.36	32,572
Greece	4	0.48	8,191
India	5	0.59	6,635
Ireland	2	0.24	1,945
Italy	6	0.71	8,665
Kenya	1	0.12	400
Netherlands	7	0.83	10,467
New Zealand	1	0.12	1,371
Nigeria	1	0.12	400
Singapore	3	0.36	4,597
South Africa	25	2.97	35,755
Tanzania	3	0.36	3,220
United Kingdom	26	3.09	39,808
United States	4	0.48	1,698
Zambia	705	83.83	340,232
Zimbabwe	38	4.52	31,287
Total	841	100	530,559

Source: Authors' own compilation based on the 2013/2014 PHS on Large-Scale Agricultural Holdings.

Note: Table 1 reports all countries that operate large-scale farms in Zambia. Seven large-scale farms that are jointly owned by two countries are reported more than once. This raises the total number of foreign-owned farms from 834 to 841.

Large-scale farmers from countries within the southern African region – that is, Zimbabwe and South Africa – also account for a significant share of large-scale farm investments. Not

⁹ In the literature, medium-scale farmers are defined as farmers that cultivate between 5 and 100 hectares (see Jayne et al., 2016). However, CSO classifies all farms above 20 hectares as large-scale farms; the PHS on Large-Scale Agricultural Holdings therefore covers a large share of farmers that would otherwise have been classified as emergent medium-scale farmers in other studies.

surprisingly, Zambia's colonial history attracts a large number of investors from the United Kingdom, which has the third-highest number of large-scale farms in the country.

If we take a look at the total hectares cultivated per country, it is clear that the countries with the highest number of large-scale farms also cultivate the largest amounts of land.¹⁰

3.2 Has the Presence of Large-Scale Farms Increased over Time?

In Figure 1, we distinguish between Zambian-owned and foreign-owned large-scale farms to show the years in which these farms were acquired. The graph provides several interesting insights. First, it confirms that large-scale farms have had a long history in Zambia. The oldest large-scale farms were established in the early 1900s.

Second, it demonstrates that three key political and economic changes have influenced the number of large-scale farms in Zambia. First, in the years following Zambia's attainment of independence from the United Kingdom in 1964, the Zambian government pursued strong nationalisation policies, such as the Land (Conversion of Titles) Act of 1975, which aimed to increase the engagement of indigenous Zambians in the agricultural sector. The Act vested all land in the president, who held it in perpetuity for the people of Zambia (Adams, 2003). This post-independence period coincides with the first spike in Zambian-owned large-scale farms, which occurred in the late 1970s. Second, in the late 1990s, both Zambian and foreign-owned large-scale farms increased significantly. This increase was driven by the adoption of the Land Act of 1995, which allowed foreign investors to acquire land in Zambia via leasehold (Nolte, 2014).

¹⁰ The exception being Germany, which only has three large-scale farms in Zambia but cultivates 32,572 hectares. Amatheon Agri, a German-owned large-scale farm that acquired 30,000 hectares in Mumbwa, accounts for the bulk of land cultivated by large-scale investors from Germany.

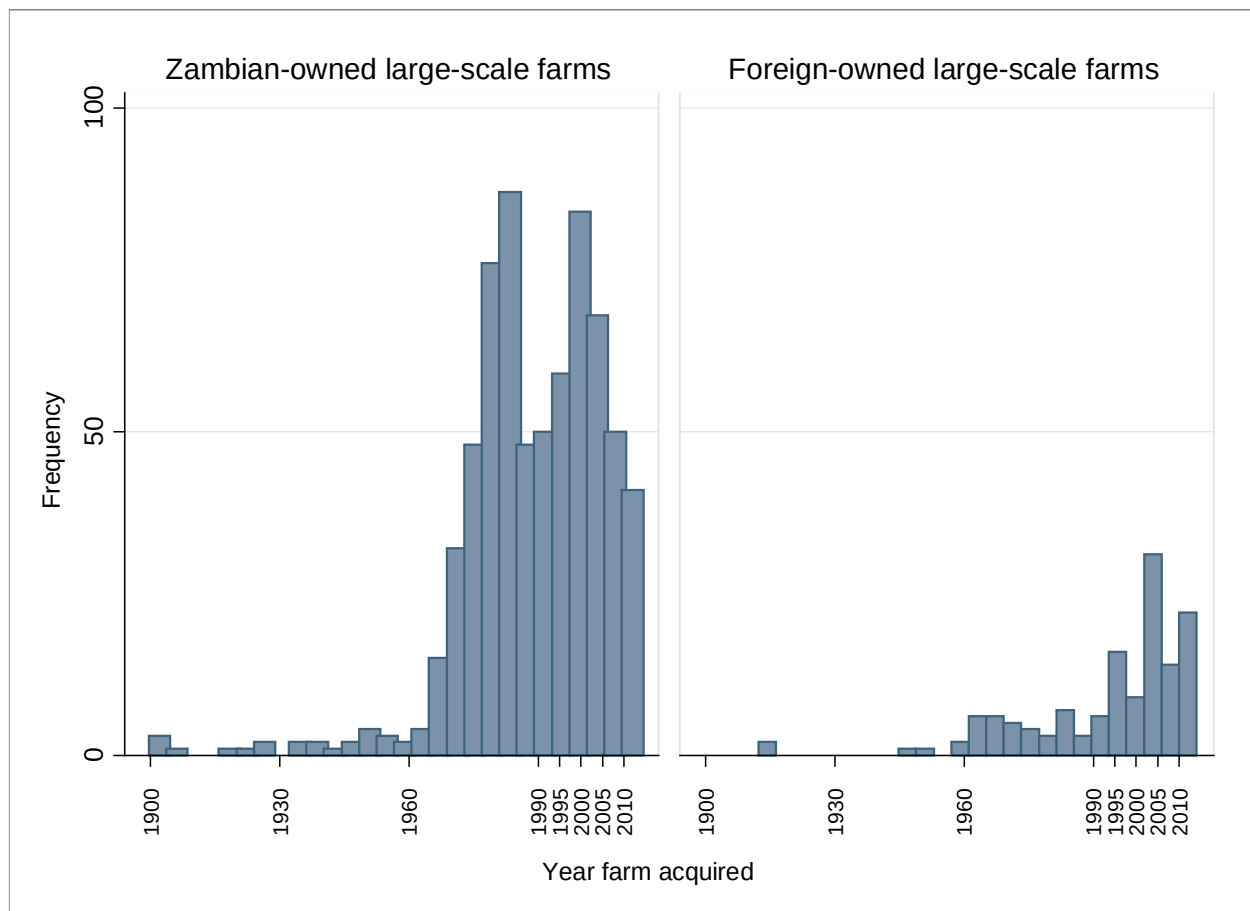


Figure 1. Acquisition of Large-Scale Farms over Time

Source: Authors' own compilation based on the 2013/2014 PHS on Large-Scale Agricultural Holdings.

The last spike in large-scale farms occurred after the year 2010 as a result of more foreign-owned farms. This confirms that the presence of large-scale foreign investments has indeed increased in Zambia in the years since the 2007–2008 food-price crisis.

3.3 Where Are Large-Scale Farms Located?

Figure 2 shows the wards that have hosted large-scale farms since 1995.¹¹ The larger the area under cultivation by large-scale farms, the darker the shading of the ward. We show the

¹¹ The Land Act 1995 was a crucial event in Zambia, representing a new era for commercial farms. We hence consider all farms that came into existence after 1995 to be part of this new wave of commercialisation. See Nolte (2014) for a detailed account of how the Land Act 1995 fostered large-scale foreign investments.

cumulative development of the area cultivated over time, with the first map in the top-left corner covering the period between 1994 and 1999 and the other maps adding five-year intervals until 2014.

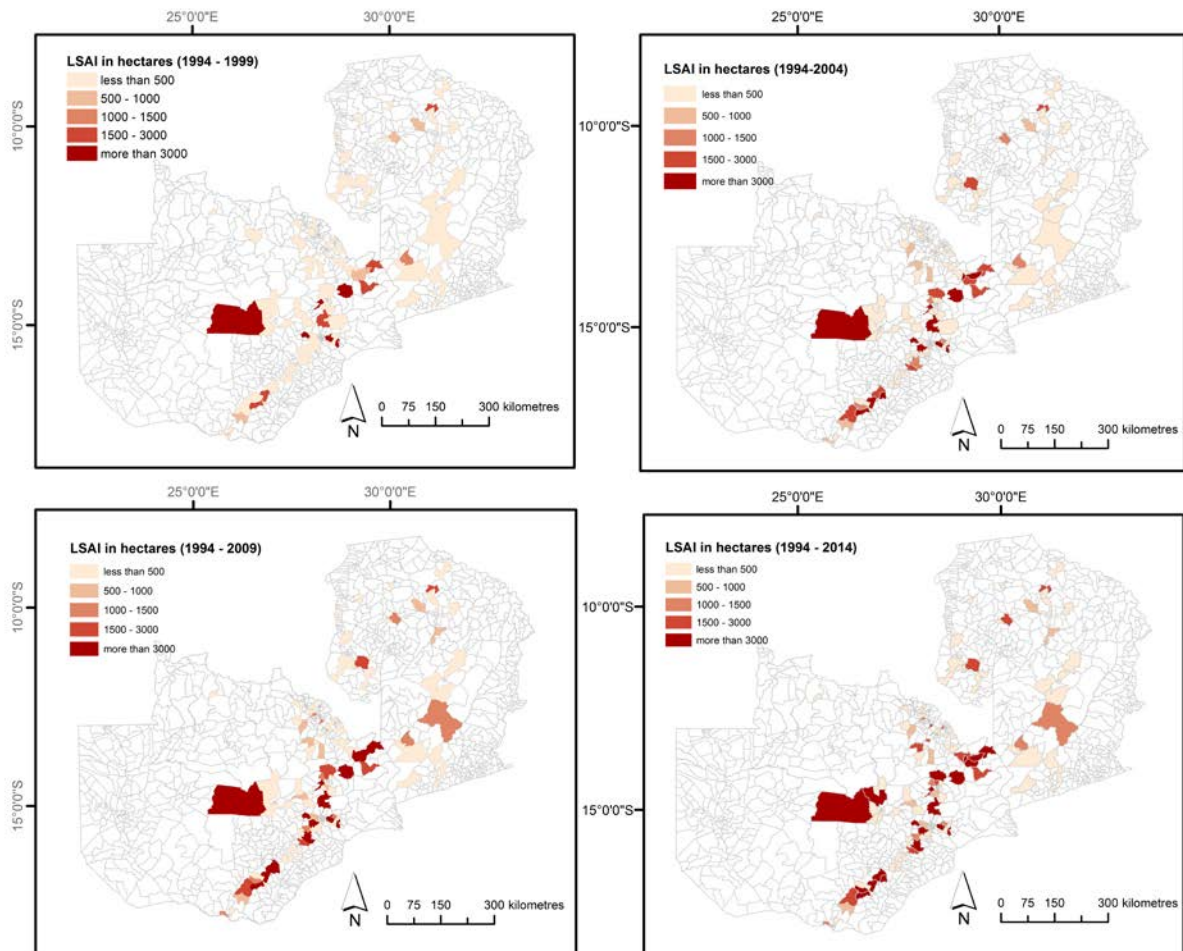


Figure 2. Location of Large-Scale Farms in Zambian Wards according to the Hectare Size of Farms

Source: Authors' own compilation based on data from CSO PHS for large-scale holdings 2013/2014.

As can be seen in the maps in Figure 2, large-scale farms are concentrated in only a few wards in Zambia. Moreover, one can see that most new large-scale farms are located in wards that already had large-scale farms prior to 1995. This hints at the agglomeration of large-scale farms in

certain wards. To better understand this non-random distribution, we use various data sets to undertake an analysis of the wards that host large-scale farms.

3.4 Does Infrastructure Influence the Location of Large-Scale Farms?

The areas shaded grey in Figure 3 illustrate three buffers (10 km, 25 km, and 50 km) between wards with large-scale farms, key transport infrastructure (main roads and railway routes), and urban centres (cities with more than 20,000 inhabitants).¹²

In addition, we calculated the Euclidian distance between the nearest point of a ward with a large-scale farm and (a) a highway (mean: 43.9 km), (b) a railroad (mean: 43.3 km), and (c) a city with more than 20,000 inhabitants (mean: 42.7 km).

If we take into account the shortest distance to one of these infrastructural features only, the distance to a ward that hosts a large-scale farm is, on average, 24.5 km, whereas the distance to wards without large-scale farms is 64 km on average. In addition, we can see from Figure 3 and the average distance that, with a few exceptions, wards targeted by large-scale farms are relatively close to the main transport infrastructure and/or urban centres.

¹² Sitko and Chamberlin (2016) use a travel-time model to estimate market access, which shows similar patterns to the buffer areas. We have opted for this simplified approach as we are looking at wards – i.e. larger polygons – instead of exact coordinates. This map is meant to provide a rough overview of accessibility in Zambia. We use three different buffers (10 km, 25 km, and 50 km) to show different degrees of access.

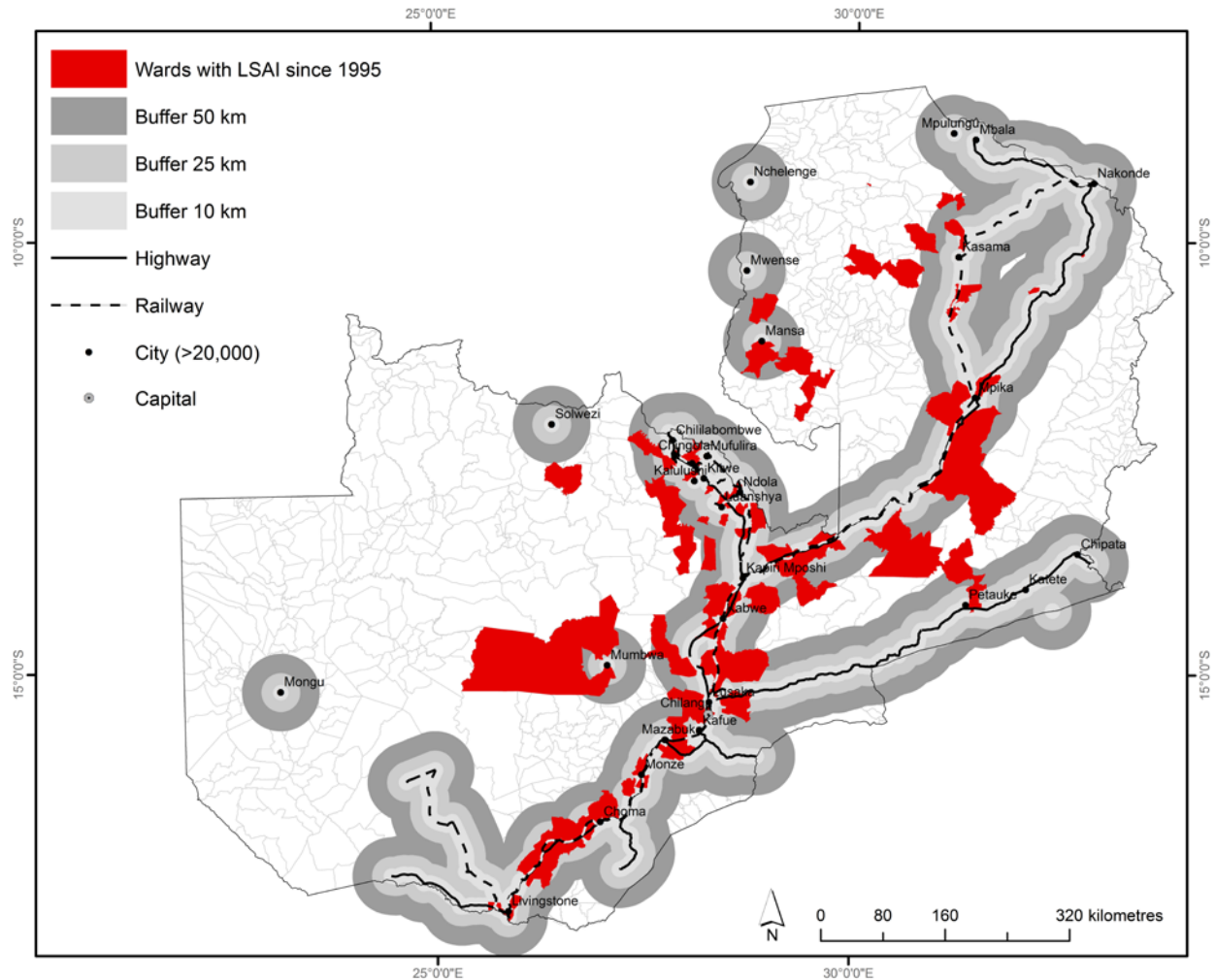


Figure 3. Large-Scale Farms and Infrastructure

Source: Authors' own compilation based on data from the Digital Chart of the World (roads and railroads), and the U.S. National Imagery and Mapping Agency (cities) with cross-checking of recent population sizes at <http://www.citypopulation.de/Zambia-Cities.html>.

3.5 Does the Idle-Land Narrative Hold for Zambia?

A number of studies have established that large-scale farms are not located on “idle land” as was previously reported (Messerli et. al. 2014). We examine whether this growing consensus on

large-scale farm location and idle land is also valid for Zambia by looking at the main land cover of wards with large-scale farms in the year 2000.¹³

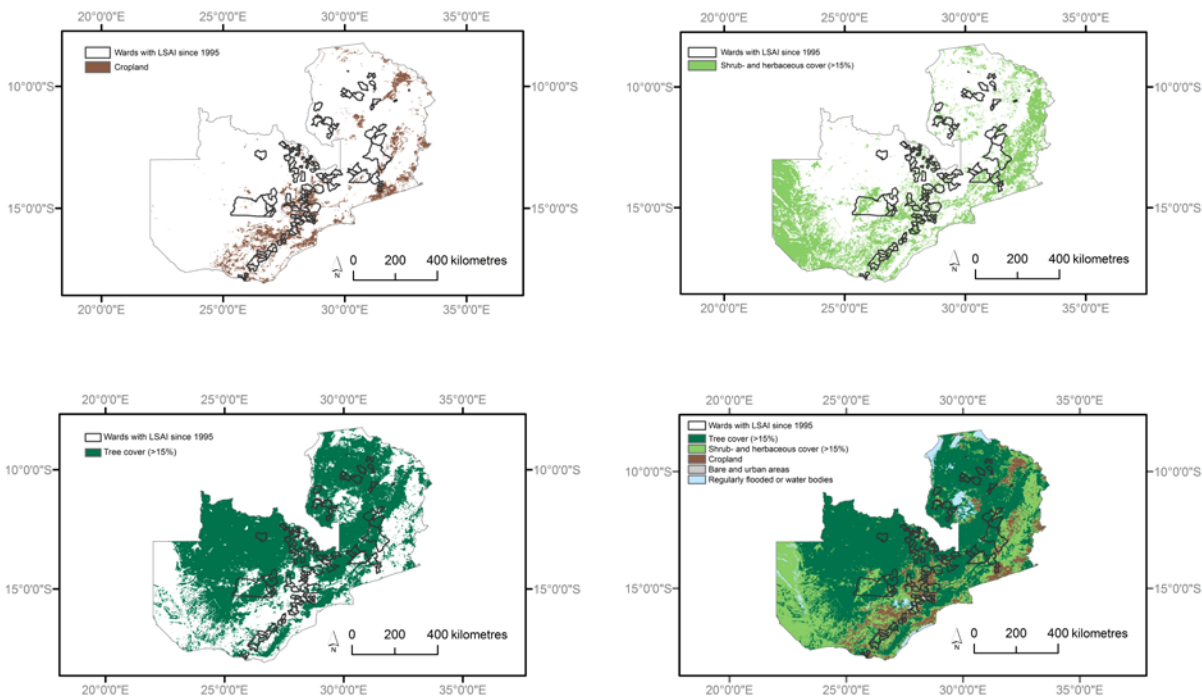


Figure 4. Land Cover and Wards with Large-Scale Farms

Source: Authors' own compilation based on data from GLC 2000 (resampled by DIVA-GIS onto a 30-second grid).

Starting from the top right of Figure 4 and moving counterclockwise, we superimpose the boundaries of wards with large-scale farms (outlined in black) over maps of cropland cover, shrub and herbaceous cover, and tree cover. The map in the bottom-right corner combines all three. One can see that all three types of land cover are present in wards that host large-scale farms. Cropland is most frequent along the “line of rail” and coincides with the presence of large-scale farms.¹⁴ Thus, most wards targeted by large-scale farmers already had cropland in

¹³ The date from year 2000 is selected as the starting data due to availability.

¹⁴ The “line of rail” is a term used to refer to the regions surrounding the main infrastructure networks depicted in Figure 3. These regions extend from the south-west of the country in Livingstone to the north-east border with Tanzania in Nakonde.

2000. One striking observation is that many of the wards that host large-scale farms were largely covered by trees in 2000. This provides a hint of the scale of deforestation that occurs when land is prepared for farming activities and simultaneously suggests a loss of income for local communities since forest lands are a source of firewood and non-timber forest products.

3.6 Do Agro-Ecological Conditions Determine the Location of Large-Scale Farms?

Arezki et. al. (2015) show that agro-ecological potential is one of the major determinants of land-based investment. In Figure 5, we investigate whether wards that host large-scale farms have better soil quality than wards without such farms.

To construct a measure of soil quality we combine the seven soil characteristics that have been identified by the Harmonized World Soil Database as being good for crop production.¹⁵ The map shows the mean soil quality values, with the green cells indicating high soil quality and the red cells indicating low soil quality. From Figure 5, it is easy to see that the majority of wards with large-scale farms are located in regions with high soil quality. Sitko et al. (2015: pp. 12) show that the southern part of Zambia has the highest variation in intra- and inter-seasonal rainfall and is also the driest part of the country.

¹⁵ The seven soil characteristics are (1) nutrient availability, (2) nutrient retention capacity, (3) rooting conditions, (4) oxygen availability to roots, (5) excess salts, (6) toxicity, and (7) workability. For further information, please refer to: <http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/SoilQuality.html?sb=10>. All soil classes are scaled between 1 and 7. We add up all the values for the seven individual soil classes and divide this by seven so that our result is equally scaled between 1 and 7. A value of 7 indicates the best soil quality and 1 very low soil quality.

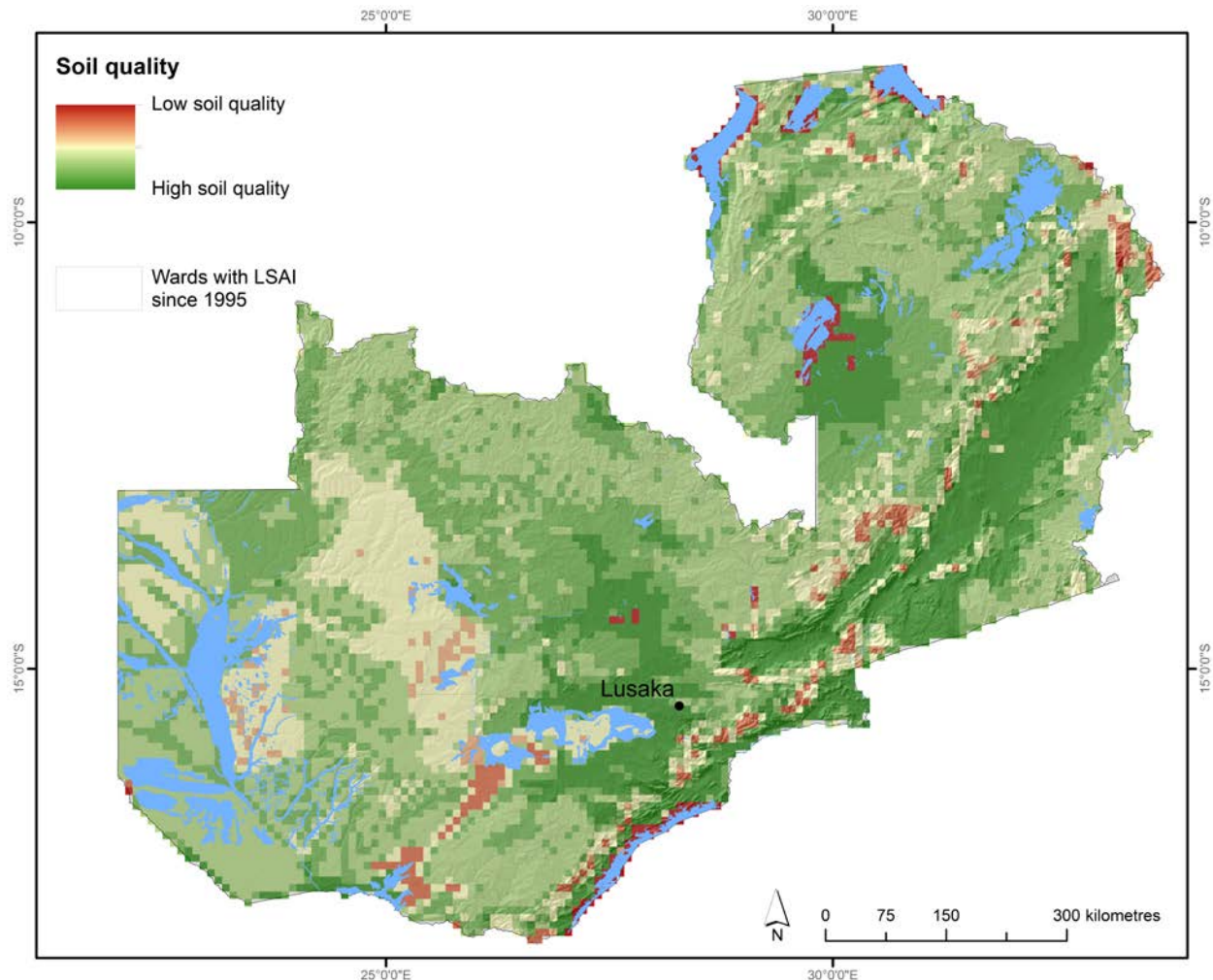


Figure 5: Soil Quality

Source: Authors' own compilation based on data from the Harmonized World Soil Database from IIASA and FAO (Fischer et al., 2008).

Interestingly, our data shows that wards targeted by large-scale farms are located in both the northern and southern parts of Zambia. We therefore assume that a historical legacy is being played out: European settler agriculture occurred in the southern parts of Zambia (Sitko et al., 2015:12) and thus investors today might prefer regions with a tradition of large-scale farming, despite the rainfall patterns not being ideal.

3.7 Do Smallholders' Socio-Economic Characteristics Determine the Location of Large-Scale Farms?

Differences in the socio-economic characteristics of smallholders living in wards with large-scale farms may determine the location of large-scale farms. For instance, if a large-scale farm is labour intensive, the large-scale farm owners may decide to set up their farm in densely populated areas where they are assured that labour is readily available.

Table 2 compares the socio-economic characteristics of wards that host large-scale farms with those that do not host such farms. The population in wards with large-scale farms is significantly higher than in wards without large-scale farms for both the years 2000 and 2010, when national census data was collected. The poverty head count is also larger in wards that host large-scale farms. However, there are no significant differences in population density and the incidence of poverty across these wards.

Comparing smallholder households in wards with and without large-scale farms for the years 2002/2003, 2003/2004, 2004/2005, 2005/2006 and the years 2010/2011, 2011/2012, 2012/ 2013, we observe that there are many statistically significant differences in the socio-economic characteristics of smallholders. These results for the years 2003–2006 and 2011–2013 are reported in the lower panels of Table 2.¹⁶

¹⁶ For the sake of brevity, these two periods will henceforth be referred to as 2003–2006 and 2011–2013. Survey data from the years 2006/2007, 2008/2009, and 2009/2010, which coincide with the global food-price crisis, are not reported, for the reasons outlined in Section 2.

Table 2. Ward Characteristics and Socio-Economic Characteristics of Smallholders

	Ward socio-economic characteristics				
	Wards with large-scale farms		Wards without large-scale farms		p-value
	Mean	SD	Mean	SD	
Population 2000	10386.64	5866.03	6821.23	4116.14	0.00
Population 2010	12709.86	7181.67	8367.11	5643.20	0.00
	28.04	24.28	31.02	62.89	0.32
Number of poor 2010	10593.89	6088.28	6739.76	4059.90	0.00
Poverty headcount 2010 ¹⁷	0.78	0.11	0.81	0.12	0.09
Observations	28		323		
	Smallholder socio-economic characteristics (2003–2006)				
	Wards with large-scale farms		Wards without large-scale farms		p-value
	Mean	SD	Mean	SD	
Age of household head	45.69	14.87	44.95	14.88	0.07
Years of schooling of household head	8.27	5.01	8.62	5.29	0.17
Number of household members	6.88	3.49	6.52	3.13	0.01
Household grows cash crops (dummy=1 if yes)	0.11	0.31	0.07	0.25	0.00
Total fertiliser used per hectare (in kg)	54.12	104.31	48.82	128.23	0.01
Maize harvest (kg)	2003.31	4578.58	1492.63	3053.57	0.00
Maize yield	1310.81	918.18	1455.65	990.61	0.00
Household asset index ¹⁸	0.17	0.18	0.13	0.14	0.00
Hectares cultivated by household	2.42	2.60	1.91	1.87	0.00
Hectares cultivated by household for maize	1.22	1.45	0.95	1.16	0.00
Hectares cultivated by household for staple crops	0.18	0.43	0.27	0.64	0.00
Hectares cultivated by household for cash crops	0.65	1.24	0.31	0.76	0.00
Share of small farms (<1.42 hectares)	0.45	0.50	0.52	0.50	0.00
Share of big farms (1.42 to 20 hectares)	0.55	0.50	0.48	0.50	0.00
Observations	1,107		9,158		

¹⁷ The poverty head count is the proportion of the population that lives below the national poverty line, which is valued at the cost of the national food basket in 2010 (ZMW 96,366) (de la Fuente et al., 2015) .

¹⁸ The asset index is constructed from a linear index of households' physical assets, whose weights have been obtained using a principal components analysis. It includes the household's ownership of assets such as ploughs, harrows, tractors, ox carts, vehicles, water pumps, cattle, and livestock.

Smallholder socio-economic characteristics (2011–2013)					
	Wards with large-scale farms		Wards without large-scale farms		
	Mean	SD	Mean	SD	p-value
Age of household head	44.95	14.89	44.80	14.51	0.73
Years of schooling of household head	4.77	3.97	4.51	3.89	0.04
Number of household members	7.15	3.66	6.56	2.99	0.02
Household grows cash crops (dummy=1 if yes)	0.20	0.40	0.09	0.29	0.00
Total fertiliser used per hectare (in kg)	120.33	140.36	125.51	155.46	0.59
Maize harvest (kg)	4528.63	7635.83	2900.69	5369.45	0.70
Maize yield	1758.47	1148.64	1804.21	1214.61	0.21
Household asset index ¹⁹	0.21	0.20	0.14	0.14	0.00
Hectares cultivated by household	3.30	3.02	2.16	2.23	0.00
Hectares cultivated by household for maize	2.04	2.20	1.33	1.58	0.00
Hectares cultivated by household for staple crops	0.03	0.24	0.13	0.52	0.00
Hectares cultivated by household for cash crops	0.72	1.06	0.26	0.67	0.00
Share of small farms (<1.42 hectares)	0.30	0.46	0.48	0.50	0.00
Share of big farms (1.42 to 20 hectares)	0.70	0.46	0.52	0.50	0.00
Observations	1,321		11,015		

Sources: Population and population density from CSO ward shapefiles, poverty headcount from de la Fuente (2015). All other data is sourced from the PHS on Small and Medium-Sized Agricultural Holdings.

¹⁹ The asset index is constructed from a linear index of households' physical assets, the weights of which have been obtained using a principal components analysis. It includes the household's ownership of assets such as ploughs, harrows, tractors, ox-carts, vehicles, water pumps, cattle and livestock.

For both time periods, smallholder households in wards with large-scale farms tended to be larger than those in wards without large-scale farms. In addition, households in wards with large-scale farms in the period 2003–2006 had significantly larger areas under cultivation, grew more cash crops, applied more fertiliser, and had higher maize yields

Moreover, these households owned more assets and had a higher share of farms larger than the median smallholder farm of 1.42 hectares. Interestingly, we observe that smallholders in wards without large-scale farms not only caught up with regard to fertiliser use but even exceeded the amount of fertiliser used by smallholders in wards with large-scale farms in the period 2011–2013. This high level of fertiliser usage is not commensurate with maize yields, as maize yields tend to be higher in wards without large-scale farms.

Examining the trends over time, we observe a drastic increase in fertiliser use and maize harvest. We further observe that both smallholders in wards with large-scale farms and smallholders in wards without large-scale farms increased the area they cultivated. A striking insight concerns the share of small and big farms (we compare smallholders that are smaller and larger than the median farm size of 1.42 hectares): while the share of big farms increased for both sets of smallholders between the two time periods, the increase in wards with large-scale farms is particularly striking as only 30 per cent of farms remained smaller than 1.42 hectares.

This initial descriptive overview of large-scale farms in Zambia and the analysis of local geographical contexts shows that some of the commonly held perceptions do not hold for Zambia: we observe that large-scale farms are not located on idle land but in wards that are well connected to infrastructure, that already have some land under large-scale cultivation, and that have significant amounts of tree cover and good soil quality. In addition, we discover that

smallholders in wards hosting large-scale farms tend to have larger farms, focus more on cash crops, and have higher yields than their counterparts in wards without large-scale farms. Over time, we see huge increases in fertiliser use and yields in both types of wards. Strikingly, the share of small farms decreases at a rapid rate in wards hosting large farms.

4. Impacts of Large-Scale Farms on Smallholder Households

4.1. *Hypotheses*

In this section we develop three main hypotheses based on the literature on the impacts of large-scale farms on surrounding smallholder communities. First, we are interested in examining how the presence of large-scale farms in a ward affects the area cultivated by smallholders. If smallholders do not have legally recognised land rights, the arrival of large-scale farms within their community may be accompanied by a heightened sense of uncertainty, land scarcity, or tenure insecurity (Cotula, 2011; Sipangule, 2017). The area cultivated by smallholders may be negatively affected by the presence of large-scale farms if large-scale farms displace smallholders and/or if they heighten land scarcity and tenure insecurity. We would therefore expect a negative relationship between large-scale farms and smallholders' area cultivated (H1:a).

For the case of Mkushi in Central Zambia, Chu (2013) shows that large-scale farm investors prefer to acquire titled state land within already established commercial farming areas over displacing smallholders from communal land. This suggests that in many cases, large-scale farms do not expand to smallholder land. This, in turn, limits the negative effects on smallholder area cultivated as outlined in hypothesis H1:a. Taking these considerations into account, we posit that if large-scale farms acquire land from markets that are not accessible to smallholders, the

expansion of large-scale farms in a ward will not affect the area cultivated by smallholders. Thus we do not expect an effect on smallholders' area cultivated (H1:b).

Lastly, large-scale farms may speed up the current trend of land consolidation (Jayne et al., 2014; Jayne et al., 2016) and contribute to an expansion of the area cultivated by smallholders. Matenga and Hichaambwa (2017) provide evidence of this in a study where they compare a large-scale plantation, a medium- to large-scale commercial farming area, and an out-grower scheme that supplies sugar cane to a large-scale farm for processing. They find that the out-grower scheme leads to the agglomeration or pooling of family land into consolidated land blocks. For the case at hand, we expect that if some smallholders reduce their area cultivated and medium-scale farmers expand, the average area cultivated by smallholders will increase. Furthermore, if the presence of large-scale farms in a ward results in positive spillovers for smallholders, smallholders may respond to these spillovers by increasing their area cultivated (H1:c).

The effects outlined in our threefold hypothesis could be occurring simultaneously, thus making the direction of the net effect on smallholders' area cultivated dependent on which effect (H1:a, H1:b, or H1:c) is more dominant.

Second, we test whether the presence of large-scale farms increases smallholders' access to fertiliser. Studies conducted in other sub-Saharan African countries find that smallholders benefit from the increased access to agricultural infrastructure and inputs provided by large-scale farms. For instance, Deininger and Xia (2016) show that smallholders living within a 50 km radius of large-scale farms in Mozambique have increased access to agricultural technologies in the short

term. Similar results that hint at positive spillovers on fertiliser, yields, and improved seed use in Ethiopia have been observed by Ali et al. (2016).

In Zambia, smallholder access to fertiliser is largely determined by the Farmer Input Support Programme (FISP). The FISP – formerly known as the Farmer Support Programme (FSP) – is a government subsidy that was introduced in 2002 with the initial goal of increasing private sector participation in agricultural input markets. A second goal of increasing household food security and incomes was adopted in 2009/2010, when the FSP was reformed to become the FISP (Resnick and Mason, 2016). In spite of its growth-enhancing and poverty-reducing objectives, the FISP has resulted in the crowding out of private-sector fertiliser supplies and in reductions in the total amount of fertiliser available for smallholders in some regions of Zambia (Xu et al., 2009).

Taking the FISP into account, we posit a second twofold hypothesis: First, large-scale farms may increase smallholders' access to fertiliser, if no distortionary effects are caused by the FISP (H2:a). Second, if the subsidy programme leads to a crowding out of the private sector, we do not expect smallholders' access to fertiliser in wards with large-scale farms to increase (H2:b).

Lastly, we investigate how smallholders' maize yields are affected by the increasing presence of large-scale farms in wards. We select maize yields as our outcome variable as maize is a staple crop grown by all smallholders in our sample. We expect that learning effects, increased access to infrastructure, and agricultural technologies that arise from the presence of a large-scale farms in a ward increase smallholders' investments at the intensive margin and cause them to increase their maize yields (H3). However it is likely that these yield-enhancing effects may be undermined if large-scale farms have adverse environmental effects on smallholders. For

instance, Mujenja and Wonani (2012) show that large-scale farms are responsible for the emission of toxic substances into the air, water, and soil. In addition, they find that large-scale farms contribute to the contamination of groundwater through the excessive use of chemical fertilisers and aerial pesticide sprays. Johansson et al. (2016) find that land acquisitions heavily draw on freshwater resources and thereby overconsume surface and groundwater.

4.2 Estimation Strategy

We adopt a difference-in-differences approach that compares the three agricultural outcomes outlined above for smallholders in wards with large-scale farms with those of smallholders in wards with no such farms. More specifically, we compare the difference in the change in the hectares cultivated, in maize yields, and in access to fertiliser between the periods 2003–2006 and 2011–2013. Our decision to examine changes between these two periods is motivated by the occurrence of the 2007–2008 food-price crisis as well as by the availability of PHS data. As pointed out by several scholars and confirmed by Figure 1, the period directly after the 2007–2008 food-price crisis was exceptional in that it led to an increase in foreign-owned large-scale farms. Thus it is reasonable to study the periods directly before and after this shock. We pool the data into two time periods to ease the interpretability of the results. This approach has also been adopted by others using multiple cross-sectional data (see for instance Abramitzky and Lavy, 2014). All wards that have hosted large-scale farms since the year 2003 are considered for the analysis. The year 2003 is selected as the start date because ward-level data is not available for earlier PHS data sets.

We thus estimate:

$$Y_{iwt} = \alpha + \beta_0 T_{wt}t + \beta_1 T_{wt} + \beta_2 t + \beta_3 X_{iwt} + \beta_4 y_t + \beta_5 w_w + \varepsilon_{iwt}$$

where Y_{iwt} represents the logs of hectares first cultivated for all crops and later for maize only, maize yields, or access to fertiliser by a smallholder household i in a ward w at time t . The time dummy t is equal to 1 for the years 2011–2013 and 0 for the earlier years. $T_{wt}t$, which is the main explanatory variable of interest, is an interaction term between the treatment dummy T_{wt} and t . X_{iwt} is a vector of household and ward-level control variables, y_t are year dummies, and w_w represents ward-level fixed effects. ε_{iwt} is the error term.

4.3. Econometric Considerations

The difference-in-differences strategy we adopt does not account for selection bias that may arise due to the non-random location of the large-scale farms. If the location of the large-scale farms is partially determined by smallholders' agricultural outcomes, this would bias the results. One way to correct this bias would be through the use of propensity score matching that would match the observed pre-treatment characteristics of similar smallholders across wards with and without large-scale farms. However, the data set at hand does not allow us to perform such a matching strategy as it does not contain sufficient pre-treatment variables that simultaneously influence the location of large-scale farms and smallholders' agricultural outcomes (Caliendo and Kopeinig, 2008). Since the analysis of the local geographical contexts of large-scale farms in the previous section reveals that infrastructure and soil quality are more important determinants of large-scale farm locations than smallholders' agricultural outcomes, we are confident that selection bias is not a major problem for our analysis.

Furthermore, endogeneity bias may arise because smallholders in wards with large-scale farms are on different trajectories than those in wards without large-scale farms. This would violate the parallel trends assumption associated with difference-in-differences analyses (Angrist and

Pischke, 2008). A common approach adopted in the literature is to graph the trajectories of the treatment and control group to confirm that they followed a common underlying trend prior to the intervention (Hastings, 2004). Taking this approach, figures A1–A3 in the appendix show that the maize harvested, fertiliser used, and hectares cultivated by smallholders in the treatment and control groups were on parallel trajectories prior to the 2007–2008 food-price crisis.

Moreover, the use of ward-year fixed effects enables us to eliminate any time-invariant variables that would otherwise bias the results.

4.4. Results

We first examine whether the increasing presence of large-scale farms in a ward reduces the area cultivated by smallholders (*H1*). Table 3 reports the results of the difference-in-differences estimation. The first two columns show the impacts of large-scale farms on the total area cultivated by smallholders, while columns 3 and 4 show how the area cultivated for maize by smallholders is affected by large-scale farms. The main explanatory variable of interest – the interaction between the treatment and time variables (*Treat_time*) – shows that large-scale farms have positively and significantly affected the area cultivated by smallholders. In the full regression models where we control for the household heads' characteristics (columns 2 and 4), we observe that the total area cultivated and the area cultivated for maize increase by 15 and 13 per cent respectively. This result is statistically significant for all smallholders at the 1 per cent level. Over time, we can see that the area cultivated significantly increases for all smallholders. The total area cultivated increases by 6 per cent, while the area cultivated for maize increases by 33 per cent (columns 2 and 4). This is in line with the descriptive statistics in Table 2, which also indicate an increase in the area cultivated.

Table 3. The Impact of Large-Scale Farms on the Area Cultivated by Smallholders

VARIABLES	(1) Log all hectares	(2) Log all hectares	(3) Log maize hectares	(4) Log maize hectares
Treat_time	0.231*** (0.041)	0.154*** (0.040)	0.164*** (0.041)	0.131*** (0.041)
Time	0.097*** (0.023)	0.056*** (0.021)	0.343*** (0.023)	0.332*** (0.022)
Age of household head		0.042*** (0.002)		0.038*** (0.002)
Age of household head squared		-0.000*** (0.000)		-0.000*** (0.000)
Female household head		-0.303*** (0.014)		-0.261*** (0.015)
Years of schooling		0.008*** (0.001)		0.013*** (0.001)
Asset index based on pca		2.236*** (0.042)		2.100*** (0.044)
Household grows cash crop		0.164*** (0.020)		-0.050** (0.020)
Constant	0.344*** (0.017)	-1.005*** (0.051)	-0.450*** (0.017)	-1.733*** (0.053)
Observations	22,601	19,960	22,601	19,960
R-squared	0.011	0.201	0.030	0.180
Number of ward_id_2010	349	349	349	349
Year FE	Yes	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes	Yes

Standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4 further examines whether this result also holds for heterogeneous groups of smallholders – smallholders that cultivate small farms (below the median farm size of 1.42 hectares) and smallholders that cultivate big farms (between 1.42 and 20 hectares) – and for the area cultivated for crops other than maize. We observe that the presence of large-scale farms in a ward significantly increases the total area cultivated for both small and large farms, by 9 and 8 per cent (columns 1 and 2). Examining how the area cultivated for different crops is affected, we observe that smallholders cultivating smaller farms in wards with large-scale farms increase the area of land cultivated for maize by 10 per cent (column 3) but decrease the area cultivated for other staples (millet, cassava, rice, and sorghum) and cash crops (tobacco, cotton, sunflower) (columns

5 and 7). This suggests that smallholders with smaller farms favour the cultivation of maize. In contrast, we do not observe any significant crop-specific effects for smallholders cultivating bigger farms but find positive signs for maize and cash crops.

Taken together, we find evidence of an increase in the area cultivated by smallholders, supporting hypothesis H1:c. The area cultivated for all crops and for maize is larger for smallholders that are located in wards with large-scale farms. Moreover, we observe that the presence of large-scale farms in a ward leads smallholders to switch the allocation of land from the cultivation of traditional staples and cash crops to maize.

Next, we examine whether the presence of large-scale farms in a ward over time increases smallholders' access to inorganic fertiliser, or whether it has no effect on fertiliser use (H2:a and H2:b). Table 5 reports the results. The variable of interest (Treat_time) has a negative sign for all smallholders as well as for those smallholders with smaller farms and those with bigger farms. Only the coefficient for smallholders with small farms is statistically significant.²⁰ This negative finding is in line with the descriptive statistics presented in Table 2, which hint at smallholders in wards without large-scale farms overtaking those in wards with large-scale farms with regard to fertiliser use.

²⁰ Regressions were also conducted separately for the two most commonly used inorganic fertilisers in Zambia: basal and top-dressing fertiliser (Burke et al., 2016). However, similar negative and insignificant results were obtained for the impact of large-scale farms on smallholders' access to these fertiliser types.

Table 4. Heterogeneous Impacts of Large-Scale Farms on Area Cultivated by Smallholders for Different Crop Types

VARIABLES	(1) Log all hectares (small farms)	(2) Log all hectares (big farms)	(3) Log maize hectares (small farms)	(4) Log maize hectares (big farms)	(5) Log staple hectares (small farms)	(6) Log staple hectares (big farms)	(7) Log cash crop hectares (small farms)	(8) Log cash crop hectares (big farms)
Treat_time	0.090* (0.053)	0.080** (0.033)	0.096* (0.055)	0.003 (0.045)	-0.200* (0.117)	-0.073 (0.088)	-0.198** (0.079)	0.030 (0.100)
Time	-0.002 (0.024)	0.063*** (0.019)	0.188*** (0.025)	0.413*** (0.026)	-1.074*** (0.052)	-1.148*** (0.051)	0.190*** (0.036)	0.101* -0.058
Age of household head	0.025*** (0.002)	0.012*** (0.002)	0.020*** (0.002)	0.015*** (0.003)	0.010** (0.005)	0.008 (0.005)	0.002 (0.003)	0.017*** (0.006)
Age of household head squared	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000* (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000*** (0.000)
Female household head	-0.170*** (0.014)	-0.085*** (0.014)	-0.152*** (0.014)	-0.073*** (0.020)	-0.005 (0.030)	-0.049 (0.039)	-0.158*** (0.021)	-0.252*** (0.044)
Years of schooling	0.003** (0.001)	0.006*** (0.001)	0.008*** (0.001)	0.011*** (0.002)	-0.014*** (0.003)	-0.017*** (0.003)	-0.005*** (0.002)	-0.007** (0.003)
Asset index based on pca	0.944*** (0.084)	1.236*** (0.032)	0.738*** (0.087)	1.313*** (0.044)	0.110 (0.184)	-0.058 (0.087)	0.907*** (0.125)	1.587*** (0.098)
Household grows cash crop	0.198*** (0.031)	0.020 (0.014)	-0.142*** (0.032)	-0.133*** (0.020)	0.069 (0.069)	0.042 (0.039)	2.313*** (0.047)	1.274*** (0.044)
Constant	-1.011*** (0.053)	0.447*** (0.049)	-1.572*** (0.055)	-0.628*** (0.067)	-3.125*** (0.116)	-2.855*** (0.132)	-4.426*** (0.078)	-3.340*** (0.150)
Observations	9,761	10,199	9,761	10,199	9,761	10,199	9,761	10,199
R-squared	0.054	0.162	0.043	0.148	0.062	0.071	0.239	0.129
Number of ward_id_2010	349	344	349	344	349	344	349	344
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5. The Impacts of Large-Scale Farms on Smallholders' Access to Fertiliser

VARIABLES	(1) Fertiliser=1 (all farms)	(2) Fertiliser=1 (small farms)	(3) Fertiliser=1 (big farms)
Treat_time	-0.016 (0.020)	-0.062* (0.036)	-0.026 (0.025)
Time	0.377*** (0.011)	0.353*** (0.016)	0.400*** (0.014)
Household received FSP subsidy in 2006	0.584*** (0.020)	0.638*** (0.032)	0.521*** (0.026)
Female household head	-0.042*** (0.007)	-0.022** (0.009)	-0.017 (0.011)
Total hectares cultivated per household	0.028*** (0.001)	0.180*** (0.011)	0.015*** (0.002)
Age of household head	0.004*** (0.001)	0.001 (0.001)	0.000 (0.001)
Age of household head squared	-0.000*** (0.000)	-0.000 (0.000)	-0.000 (0.000)
Years of schooling	0.008*** (0.001)	0.008*** (0.001)	0.008*** (0.001)
Asset index based on pca	0.271*** (0.024)	0.309*** (0.057)	0.278*** (0.026)
Household grows cash crop	-0.033*** (0.010)	-0.052** (0.021)	-0.051*** (0.011)
Constant	0.073*** (0.026)	-0.014 (0.036)	0.217*** (0.037)
Observations			
R-squared	19,960	9,761	10,199
Number of ward_id_2010	0.307	0.343	0.258
Year FE	349	349	344
Ward FE	Yes	Yes	Yes

Standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The negative result could also be indicative of the crowding-out effect of government fertiliser provision in regions that have large-scale farms. As discussed in the previous section, the FISP is a major determinant of smallholders' fertiliser access. Households that received fertiliser from the FSP in 2006 had approximately 52 to 64 per cent more access to fertiliser, as shown by a dummy introduced in our regressions.²¹

Lastly, we examine whether smallholders' maize yields are affected by the presence of large-scale farms. Table 6 shows that maize yields for all smallholders rise significantly, by 23 per

²¹ The data set only provides information on the distribution of the FSP across smallholders for the year 2006 and not for the whole study period.

cent, in those wards with a large-scale farm. Columns 2 and 3 split the sample into smaller and bigger smallholder farms. We find positive signs for both groups of smallholders, but only the results for bigger farms are statistically significant at the 5 per cent level.

Table 6. The Impacts of Large-Scale Farms on Smallholders' Maize Yields

VARIABLES	(1) Log maize yield (all farms)	(2) Log maize yield (small farms)	(3) Log maize yield (big farms)
Treat_time	0.229*** (0.071)	0.215 (0.136)	0.170** (0.080)
Time	-0.177*** (0.039)	-0.153** (0.063)	-0.172*** (0.048)
Age of household head	0.005 (0.004)	0.010* (0.006)	0.001 (0.005)
Age of household head squared	-0.000** (0.000)	-0.000** (0.000)	-0.000 (0.000)
Female household head	-0.116*** (0.025)	-0.149*** (0.036)	-0.083** (0.036)
Years of schooling	0.008*** (0.002)	0.004 (0.003)	0.011*** (0.003)
Access to fertiliser (1=yes)	0.601*** (0.025)	0.668*** (0.039)	0.509*** (0.032)
Asset index based on pca	1.119*** (0.084)	1.367*** (0.227)	1.168*** (0.087)
Household grows cash crop	0.029 (0.035)	-0.058 (0.082)	0.055 (0.036)
Total hectares cultivated per household	-0.040*** (0.005)	-0.064 (0.045)	-0.030*** (0.005)
Constant	6.649*** (0.090)	6.579*** (0.137)	6.664*** (0.121)
Observations	18,872	9,294	9,578
R-squared	0.073	0.059	0.091
Number of ward_id_2010	349	348	344
Year FE	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes

Standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

4.5. Robustness Checks

We conduct several robustness checks to test the validity of our results. First, we check the specification: rather than pooling the data sets into a before and after period as in the previous set of regressions, we now interact the treatment variable with each year following the 2007–2008

food-price crisis. The three new outcome variables of interest are reported in Table A2 as Treat_2011, Treat_2012, and Treat_2013. The first two columns show the total area cultivated by smallholders and the total area cultivated for maize, while the third column shows the results for maize yields. As can be seen in Table A2, using interactions between the individual years and the treatment variable does not considerably change the results presented for hectares cultivated and maize yields in tables 3 and 6. For instance, smallholders in wards with large-scale farms experience an increase in total areas cultivated of 12 to 17 per cent over the three years, which does not diverge from the 15 per cent increment reported in column 2 of Table 3.

Second, we check our sample: as described in the previous sections, Zambia has a long history of large-scale agriculture that dates back to before 2003. In Table A3, we check whether the results obtained also hold if we account for the fact that large-scale farms existed prior to 2003 by introducing a dummy for wards that hosted large-scale farms prior to 2003, which we interact with the time variable to obtain a new outcome variable (All_treat_time). The results show that large-scale farms still affect smallholders' area cultivated and maize yields in a similar way, although the effect sizes are smaller.²²

Lastly, we examine whether analysing the impacts of large-scale farms on the quantity of fertiliser applied per hectare yields different results compared to the fertiliser dummy used in Table 5. Again, all the signs reported in Table A4 are negative, suggesting a decrease in fertiliser usage in those wards hosting large-scale farms. For smallholders cultivating big farms, we even obtain statistically significant results (at the 5 per cent level, column 3).

²² We also run separate regressions with these specifications for smallholders cultivating small and big farms and find results similar to those reported in tables 5, 6, and 7, albeit with different effect sizes.

5. Discussion and Conclusion

The emergence of LSAIs in the Global South has been widely discussed, and Zambia is a sought-after target country. Based on a census of large-scale farms that incorporates locational information at the ward level, we have been able to paint a comprehensive picture of large-scale farms in Zambia. This unique data set has allowed us to analyse the local geographical contexts of wards targeted by large-scale farms. We have investigated these areas and found sufficient evidence to dismiss the “idle land” narrative: We have found that land targeted by investors is close to infrastructure and markets. Moreover, large-scale farms are often established in areas with a tradition of large-scale farming, as evidenced by the agglomeration of large-scale farms in certain wards. This confirms that large-scale farms are typically set up in close proximity to smallholders, and that the question of how small and large farms interact is crucial. To better understand the impacts of large-scale farms on smallholders, we have derived several hypotheses, which we have tested using a difference-in-differences estimation with ward and year fixed effects.

This analysis yields three main results: First, we find that smallholders located in wards with large-scale farms tend to increase the area cultivated. Interestingly, we also observe that smallholders who cultivate small areas of land and are located in wards that host large-scale farms instead reduce the amount of land dedicated to growing staples and cash crops in order to expand their maize production.

Given our threefold hypothesis, we expect multiple dynamics to simultaneously affect the relationship between large-scale farms and smallholders’ area cultivated. We find a positive relationship which can be explained in two ways: First, it is likely that the rise of medium-scale

farms may be driving our result. Assuming that medium-scale farms are likely to target the same regions as large-scale farms, we cannot rule out that the increase in area cultivated by smallholders is partially driven by a process of land consolidation during which less efficient smallholders sell or rent land to medium-scale farmers (Jayne et al., 2016). In fact, our descriptive statistics confirm a general trend of a growing medium-scale agricultural sector and a shrinking number of land-poor smallholders that has been identified in the literature (for instance Anseeuw, 2016; Jayne et al., 2016). As we do not have smallholder panel data, we cannot follow changes in smallholders' landholdings over time. Thus, in this case of a positive relationship between large-scale farms and smallholders' area cultivated, we cannot say with certainty what drives our results. Second, we do not find any support for the hypothesis that large-scale farms increase smallholders' access to fertiliser, as has been found to be the case for neighbouring Mozambique (Deininger and Xia, 2016). However, one should note that the results on fertiliser are to be interpreted with caution as the FISP is a major determinant of smallholders' fertiliser use. Albeit insignificant, our finding that fertiliser access decreases in wards where large-scale farms are active is unexpected. This could be a result of private fertiliser suppliers being crowded out by the FISP and other government fertiliser suppliers in these regions; however, we lack sufficient data to say this with certainty. Further research is required to better understand these mechanisms.

Third, we find yield increases in wards with large-scale farms. Based on the literature, we assume that learning effects, increased access to infrastructure, and agricultural technologies outweigh any negative impacts associated with large-scale farms. The positive results may also be explained by the fact that smallholders cultivating larger areas of land rent or purchase land

cultivated by smaller smallholders and produce more efficiently – especially given that the result is only statistically significant for bigger smallholders.

We cannot capture the trade-offs between these different effects in our analysis. It is highly likely that both positive and negative effects go hand in hand. For instance, an increase in smallholder productivity as evidenced by higher yields might be accompanied by adverse environmental effects. Moreover, it is not clear whether this effect would last in the long term as negative environmental impacts such as soil degradation may compromise smallholders' productivity. Hence, we argue that the effects have to be considered and evaluated more comprehensively. Policymakers and investors should be transparent about the fact that investment projects are likely to be accompanied by both favourable and less favourable effects.

We close this analysis with some ideas for future research. First, more comprehensive and longitudinal smallholder data would allow us to include more outcome variables that highlight the linkages between large-scale farms and smallholders. Data that allows us to separately estimate the impacts of medium- and large-scale farms would be especially useful in disentangling the mechanisms through which smallholder outcomes are affected. Such data would also provide us with more insights on the trend of land consolidation that is simultaneously taking place. Furthermore, detailed data on large-scale farm labour would enable us to analyse the spatial and temporal dimensions of the employment effects that large-scale farms have on smallholders in their vicinities. For instance, we would expect that smallholders are employed in the early project stages but as the years of operation of large-scale farms increase, mechanised methods reduce the associated employment effects. Employment effects are also expected to be stronger for smallholders living in the direct proximity of large-scale farms compared to those

further away. In addition, our data does not enable us to capture any environmental effects that may accompany the clearing of land for agriculture, the increased application of inorganic fertilisers and pesticides, and the increased pressure on groundwater levels from large-scale farms. Moreover, better data on the investment projects would allow for a clearer distinction between the crops cultivated by investors as well as the influence of their countries of origin and other investor traits. Finally, better spatial data that includes the exact geographic coordinates of smallholders and large-scale farms would be helpful in measuring how the effects on smallholders vary with the distance to large-scale farms.

Bibliography

- Abramitzky, R. and Lavy, V. (2014). How Responsive is Investment in Schooling to Changes in Redistributive Policies and in Returns? *Econometrica*, Vol. 82(4):1241-1272.
- Ahlerup, P. and Tengstam, S. (2015). Do the land-poor gain from agricultural investments? Empirical evidence from Zambia using panel data.
- Adams, M. (2003). Land Tenure Policy and Practice in Zambia: Issues Relating to the Development of the Agricultural Sector. Oxford, UK: Mokoro Ltd.
- Ali, D., Deininger, K. and Harris, A. (2016). Large Farm Establishment, Smallholder Productivity, Labor Market Participation, and Resilience. Evidence from Ethiopia. World Bank Policy Research Working Paper, Nr. 7576.
- Angrist, J., D. and Pischke, J., S. (2008). *Mostly harmless econometrics: An empiricist's companion*. New Jersey, United States: Princeton University Press.
- Anseeuw, W., Jayne, T., Kachule, R. and Kotsopoulos, J. (2016). The Quiet Rise of Medium-Scale Farms in Malawi. *Land* Vol. 5 (3):19.
- Arezki, R., Deininger, K., and Selod, H. (2015). What drives the global “Land Rush”? *The World Bank Economic Review*, Vol.29(2): 207-233.
- Burke, J.W., Frossard, E., Kabwe, S. and Jayne, T.S. (2016). Understanding Fertilizer Effectiveness and Adoption on Maize in Zambia. MSU International Development Working Paper 147.
- Chu, J. M. (2013). Creating a Zambian Breadbasket: ‘Land grabs’ and foreign investments in agriculture in Mkushi District, Zambia. LDPI Working Paper 33. Rotterdam: The Land Deal Politics Initiative (LDPI).
- Caliendo, M. and Kopeinig, S. (2008). Some Practical Guidance for the implementation of propensity score matching. *Journal of Economic Surveys*. Vol 22(1): 31-72
- Cotula, L, Vermeulen, S., Leonard., R, and Keeley, J. (2009). *Land grab or development opportunity? Agricultural investment and international land deals in Africa*. London and Rome: International

Institute for Environment and Development, Food and Agricultural Organization of the United Nations and the International Fund for Agricultural Development.

Cotula, L.(2011). Land Tenure Issues in Agricultural Investment. SOLAW Background Thematic Report—TR05B. Rome: Food and Agricultural Organization of the United Nations

Deininger, K., and Xia, F. (2016). Quantifying Spillover Effects from Large Land-based Investment: The Case of Mozambique. *World Development*, Vol.87, 227-241.

de la Fuente, A., A. Murr and Rascón, E. (2015). *Mapping Subnational Poverty in Zambia*. Report for the World Bank and Zambian Central Statistical Office. Washington DC and Lusaka: The World Bank Group, Republic of Zambia: Central Statistical Office.

De Schutter, O. (2011). How not to think of land-grabbing: three critiques of large-scale investments in farmland. *The Journal of Peasant Studies*, Vol.38 (2): 249-279.

Fischer, G., F. Nachtergaele, S. Prieler, H.T. van Velthuisen, L. Verelst, and Wiberg, D. (2008). Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Harding, A., W. Chamberlain, I. Maluleke, S. Niassy, W Anseeuw and Manco, G. (2016). Large-scale land acquisitions profile. Zambia. Accessed on 28 August 2016 from <http://www.landmatrix.org/>

Hastings, J. S. (2004). Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California .*American Economic Review*, Vol. 94(1):317-328.

Herrmann, R. T. (2017). Large-Scale Agricultural Investments and Smallholder Welfare: A Comparison of Wage Labor and Outgrower Channels in Tanzania. *World Development*, Vol. 90: 294–310.

Jayne, T. S., Chamberlin, J., Traub, L., Sitko, N., Muyanga, M., Yeboah. F. K., Anseew, 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*. Vol.47:197-214.

- Johansson, E. L., Fader, M., Seaquist, J. W., and Nicholas, K. A. (2016). Green and blue water demand from large-scale land acquisitions in Africa. *Proceedings of the National Academy of Sciences*, 113(41), 11471–11476.
- Matenga, C. R., and Hichaambwa, M. (2017). Impacts of land and agricultural commercialization on local livelihoods in Zambia: evidence from three models. *The Journal of Peasant Studies*. Vol. 44(3): 574-593.
- Megill, J., D. (2004). Recommendations on sample design for post-harvest surveys in Zambia based on the 2000 census. Working Paper No. 11, Food Security Research Project. Lusaka: FSRP
- Messerli, P., Giger, M., Dwyer, M. B., Breu, T., and Eckert, S. (2014). The geography of large-scale land acquisitions: Analysing socio-ecological patterns of target contexts in the global South. *Applied Geography*, Vol. 53: 449–459.
- Mujenja, F., and Wonani, C. (2012). Long-term outcomes of agricultural investments: Lessons from Zambia. International Institute for Environment and Development. London, United Kingdom
- Nolte, Kerstin (2014), Large-scale agricultural investments under poor land governance in Zambia. *Land Use Policy*, Vol.38: 698-706
- Nolte, K., Chamberlain, W. and Giger, M. (2016). International Land Deals for Agriculture. Fresh insights from the Land Matrix: Analytical Report II. Bern, Montpellier, Hamburg, Pretoria: Centre for Development and Environment, University of Bern; Centre de coopération internationale en recherche agronomique pour le développement; German Institute of Global and Area Studies; University of Pretoria; Bern Open Publishing.
- Oberlack, C., Tejada, L., Messerli, P., Rist, S., and Giger, M. (2016). Sustainable livelihoods in the global land rush? Archetypes of livelihood vulnerability and sustainability potentials. *Global Environmental Change*, Vol. 41: 153–171.
- Schoneveld, G. C. (2016). Host country governance and the African land rush: 7 reasons why large-scale farmland investments fail to contribute to sustainable development. *Geoforum*, Vol.83:119-132.

- Sipangule, K. (2017). Agribusinesses, smallholder tenure security, and plot-level investments: Evidence from rural Tanzania. *African Development Review*. Vol.29:179-197.
- Sitko, N. J., Chamberlin, J., and Hichaambwa, M. (2015). The Geography of Customary Land in Zambia: Is Development Strategy Engaging With The Facts? IAPRI Working Paper 98. Lusaka: Indaba Agricultural Policy Research Institute.

Appendix

Figure A1. Mean Trends in Hectares Cultivated for Smallholders in Wards with and without Large-Scale Farms

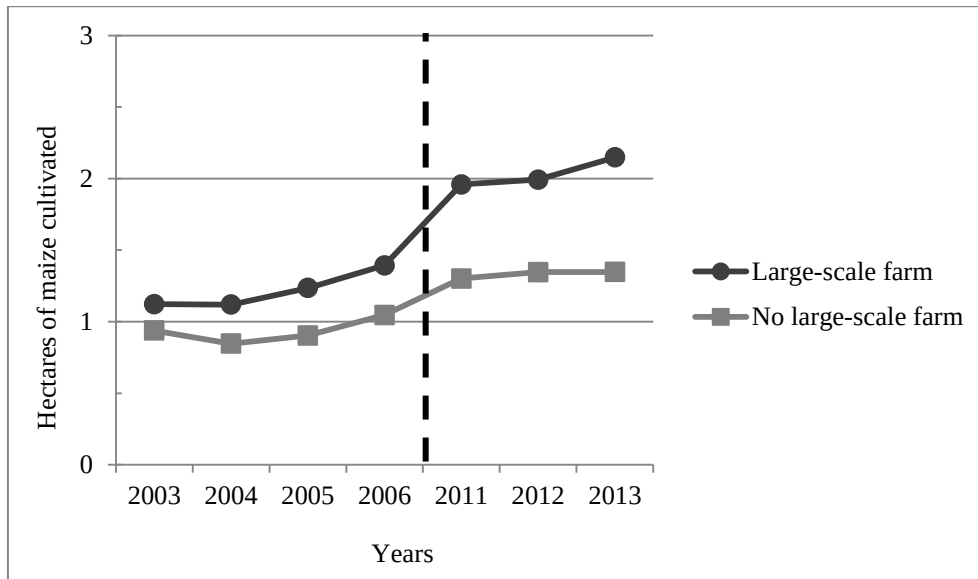


Figure A2. Mean Trends in Kilograms of Fertiliser Applied by Smallholders in Wards with and without Large-Scale Farms

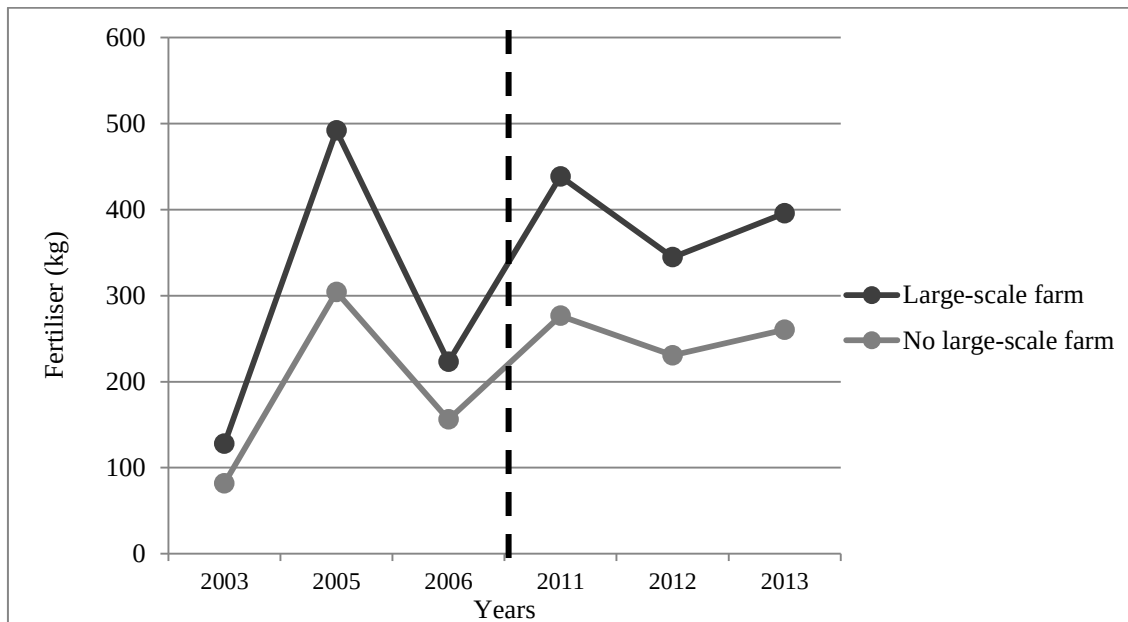


Figure A3. Mean Trends in Maize Yields for Smallholders in Wards with and without Large-Scale Farms

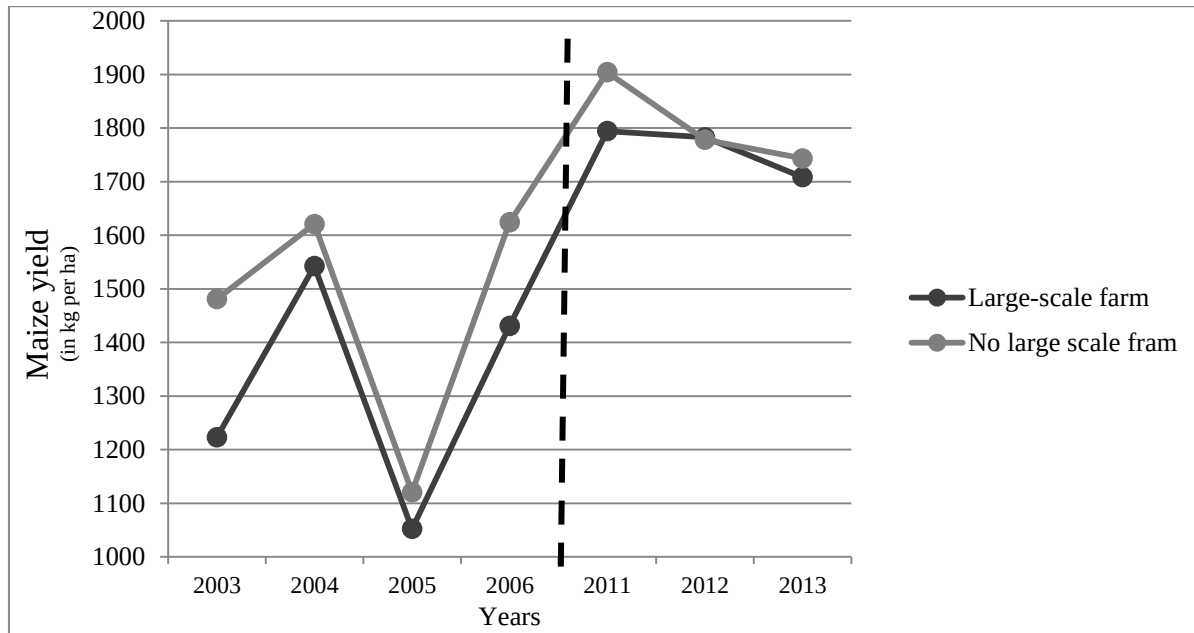


Table A1. Distribution of Hectares Cultivated by Smallholders

2002/2003				
	Mean	Standard Deviation	Median	Share
0–5 hectares	1.64	1.07	1.4	93.06
5–10 hectares	6.7	1.17	6.48	5.91
10–15 hectares	12.44	1.34	12.76	0.79
15–20 hectares	18.05	1.65	18.32	0.24
2003/2004				
	Mean	Standard Deviation	Median	Share
0–5 hectares	1.62	1.04	1.42	93.92
5–10 hectares	6.44	1.18	6.07	5.41
10–15 hectares	11.36	0.98	11	0.52
15–20 hectares	17.5	1.35	18	0.16
2004/2005				
	Mean	Standard Deviation	Median	Share
0–5 hectares	1.54	1.06	1.25	94.69
5–10 hectares	6.61	1.23	6.2	4.59
10–15 hectares	11.55	1.39	11.07	0.57
15–20 hectares	17.24	1.68	16.52	0.15
2005/2006				
	Mean	Standard Deviation	Median	Share
0–5 hectares	1.41	1.03	1.13	92.83
5–10 hectares	6.8	1.4	6.34	5.36
10–15 hectares	11.89	1.26	11.87	1.58
15–20 hectares	16.99	1.43	17	0.24
2010/2011				
	Mean	Standard Deviation	Median	Share
0–5 hectares	1.57	1.11	1.25	92.21
5–10 hectares	6.68	1.34	6.48	6.44
10–15 hectares	11.86	1.43	11.4	0.97
15–20 hectares	16.56	1.35	16.06	0.38
2011/2012				
	Mean	Standard Deviation	Median	Share
0–5 hectares	1.66	1.15	1.38	89.35
5–10 hectares	6.7	1.18	6.48	8.73
10–15 hectares	11.87	1.3	11.56	1.52
15–20 hectares	16.69	1.3	16.5	0.40
2012/2013				
	Mean	Standard Deviation	Median	Share
0–5 hectares	1.71	1.16	1.42	90.46
5–10 hectares	6.64	1.27	6.25	7.70
10–15 hectares	12.35	1.31	12.18	1.44
15–20 hectares	17.54	1.54	17.81	0.40

Table A2. Impacts of Large-Scale Farms on Smallholders' Total Area Cultivated and Maize Yields (Yearly Treatment)

VARIABLES	(1) Log all hectares	(2) Log maize hectares	(3) Log maize yield
Treat_2011	0.168*** (0.055)	0.107* (0.057)	0.221** (0.098)
Treat_2012	0.176*** (0.050)	0.138*** (0.052)	0.219** (0.089)
Treat_2013	0.123** (0.049)	0.143*** (0.051)	0.245*** (0.086)
Age of household head	0.042*** (0.002)	0.038*** (0.002)	0.005 (0.004)
Age of household head squared	-0.000*** (0.000)	-0.000*** (0.000)	-0.000** (0.000)
Female household head	-0.303*** (0.014)	-0.261*** (0.015)	-0.116*** (0.025)
Years of schooling	0.008*** (0.001)	0.013*** (0.001)	0.008*** (0.002)
Access to fertiliser (1=yes)			0.601*** (0.025)
Asset index based on pca	2.237*** (0.042)	2.100*** (0.044)	1.119*** (0.085)
Household grows cash crop	0.163*** (0.020)	-0.050** (0.020)	0.030 (0.035)
Total hectares cultivated per household			-0.040*** (0.005)
Constant	-1.005*** (0.051)	-1.733*** (0.053)	6.648*** (0.090)
Observations	19,960	19,960	18,872
R-squared	0.201	0.180	0.073
Number of ward_id_2010	349	349	349
Year FE	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes

Standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A3. Impacts of all Large-Scale Farms on Smallholders' Total Area Cultivated and Maize Yields

VARIABLES	(1) Log all hectares	(2) Log maize hectares	(3) Log maize yield
All_treat_time	0.106*** (0.034)	0.074** (0.036)	0.125** (0.061)
Time	0.057*** (0.022)	0.336*** (0.022)	-0.169*** (0.039)
Age of household head	0.042*** (0.002)	0.038*** (0.002)	0.005 (0.004)
Age of household head squared	-0.000*** (0.000)	-0.000*** (0.000)	-0.000* (0.000)
Female household head	-0.303*** (0.014)	-0.261*** (0.015)	-0.116*** (0.025)
Years of schooling	0.008*** (0.001)	0.013*** (0.001)	0.008*** (0.002)
Access to fertiliser (1=yes)			0.601*** (0.025)
Asset index based on pca	2.235*** (0.042)	2.100*** (0.044)	1.118*** (0.085)
Household grows cash crop	0.164*** (0.020)	-0.049** (0.020)	0.031 (0.035)
Total hectares cultivated per household			-0.039*** (0.005)
Constant	-1.007*** (0.051)	-1.735*** (0.053)	6.646*** (0.090)
Observations	19,960	19,960	18,872
R-squared	0.201	0.180	0.072
Number of ward_id_2010	349	349	349
Year FE	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes

Standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A4. Impacts of Large-Scale Farms on Smallholders' Application of Fertiliser per Hectare

VARIABLES	(1) Log fertiliser/ha (all farms)	(2) Log fertiliser/ha (small farms)	(3) Log fertiliser/ha (big farms)
Treat_time	-0.343 (0.212)	-0.503 (0.403)	-0.493** (0.248)
Time	3.753*** (0.108)	3.549*** (0.169)	3.938*** (0.138)
Household received FSP subsidy in 2006	5.426*** (0.202)	6.161*** (0.332)	4.634*** (0.249)
Female household head	-0.460*** (0.074)	-0.323*** (0.104)	-0.144 (0.108)
Total hectares cultivated per household	0.210*** (0.015)	1.725*** (0.125)	0.105*** (0.016)
Age of household head	0.038*** (0.011)	0.012 (0.016)	0.011 (0.015)
Age of household head squared	-0.000*** (0.000)	-0.000 (0.000)	-0.000 (0.000)
Years of schooling	0.098*** (0.007)	0.098*** (0.010)	0.091*** (0.009)
Asset index based on pca	3.324*** (0.240)	3.900*** (0.619)	3.446*** (0.254)
Household grows cash crop	-0.429*** (0.100)	-0.597*** (0.227)	-0.578*** (0.106)
Constant	-4.137*** (0.264)	-4.919*** (0.390)	-2.957*** (0.364)
Observations	18,146	8,649	9,497
R-squared	0.202	0.176	0.202
Number of ward_id_2010	348	348	343
Year FE	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes

Standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1