Spillover effects of Medium-Scale Farms on Smallholder Behavior and Welfare: Evidence from Nigeria

Lenis Saweda O. Liverpool-Tasie, Michigan State University, lliverp@msu.edu
Ahmed Salim Nuhu, Michigan State University, nuhuahme@msu.edu
Titus Awokuse, Michigan State University, awokuset@msu.edu
Thomas Jayne, Michigan State University, jayne@msu.edu
Milu Muyanga, Michigan State University.
Adebayo Aromolaran, Adekunle Ajasin University, adebayo.aromolaran@gmail.com
Adesoji Adelaja, Michigan State University, adelaja@msu.edu

Selected Paper prepared for presentation at the 2020 Agricultural & Applied Economics Association Annual Meeting, Kansas City, MO
July 26-28, 2020

Spillover effects of Medium-Scale Farms on Smallholder Behavior and Welfare: Evidence from Nigeria

Lenis Saweda O. Liverpool-Tasie,
Ahmed Salim Nuhu,
Titus Awokuse,
Thomas Jayne,
Milu Muyanga,
Adebayo Aromolaran
Adesoji Adelaja

Lenis Saweda O. Liverpool- Tasie is an Associate Professor in the Department of Agricultural, Food and Resource Economics (AFRE); College of Agriculture and Natural Resources (CANR); Michigan State University (MSU), USA . Ahmed Salim Nuhu is a PhD Candidate in AFRE at MSU. Titus Awokuse is Professor and Chair of AFRE at MSU, USA. Thomas Jayne is MSU Foundation Professor in AFRE and Milu Muyanga is Assistant Professor in AFRE at MSU. Adebayo Aromolaran is Professor and Dean of the Faculty of Agriculture, Adekunle Ajasin University, Akungba-Akoko (AAUA), Ondo State, Nigeria and Adesoji Adelaja is the John A. Hannah Distinguished Professor in Land Policy at MSU.

ACKNOWLEDGEMENTS
Funding for this research is provided by Agricultural Policy Research in Africa (APRA), a five-year Research Programme Consortium (RPC) funded by the UK Department for International Development (DFID). The programme, which is based at the Institute of Development Studies (IDS), UK (www.ids.ac.uk), builds on more than a decade of research and policy engagement work by the Future Agricultures Consortium (FAC) (www.future-agricultures.org). This work also benefited from financial support from Guiding Investments in Sustainable Agricultural Intensification in Africa (GISAIA)- funded by the Bill and Melinda Gates Foundation, the Feed the Future Nigeria Agricultural Policy Project funded by USAID-Nigeria and Michigan State University's Agbio Research. We are grateful for this generous support and from the helpful feedback from an internal review. Any views expressed herein are those of the authors and do not necessarily represent the views of the donors.

ACRONYMS

APRA	Agricultural Policy Research in Africa
APS	Agricultural Performance Survey
CF	Control Function
CFA	Control Function Approach
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station
LGA	Local Government Area
LSF	Large- Scale Farms
MSF	Medium-Scale Farms
NAERLS	National Agricultural Extension and Research Liaison Services
OLS	Ordinary Least Squares
SSA	sub-Saharan Africa
SSF	Small-Scale Farms

Table of Contents

Tables

Figures

1. Introduction

A rapid change in farm size distributions is taking place across sub-Saharan Africa (SSA). Many countries are experiencing an increasing share of farmland under "medium-scale" farms between five and 100 ha (Jayne et al., 2016). The share of land owned by these emerging medium-scale farms (MSFs) range from about 20% in Kenya, to 32% in Ghana, 37% in Tanzania, and as high as 53% in Zambia (Jayne et al., 2019). These medium-scale farms co-exist with small-scale farmers (operating on less than 5 hectares), who still constitute the majority of households in rural areas of Africa. While there is a growing literature documenting the drivers of the rise of MSFs (Anseeuw et al., 2016; Jayne et al., 2016) and their characteristics (Jayne et al., 2019; Muyanga et al., 2019; Muyanga and Jayne, 2019) empirical evidence on how this rise in MSFs impacts neighboring small scale farmers (SSFs) is thin.

Compared to large-scale farms (LSFs) operating between five and 100 hectares, MSFs tend to be more socio-culturally similar to SSFs in the communities where they are located (Wineman et al., 2020; Houssou et al. 2016; Chamberlin and Jayne, 2020). Due to their smaller size, they are also more likely than LSFs to be interested in coordinating input purchase or output sales with SSFs. Despite increasing recognition of these potentially stronger spillover effects of MSFs, majority of the existing empirical literature has focused on spillover effects of LSFs (Ali et al., 2019, Burke et al., 2019, Xia and Deininger, 2019, Glover and Jones, 2019, Herrmann, 2017, Lay et al., 2018). Few studies such as Burke et al. (2019) and Wineman et al. (2020) examine the spillover effects of MSFs between 5 and 50 hectares. In these studies (as in most of the literature on large farms), identification of spillover effects relies on changes in SSF behavior due to their proximity to larger farms, conditional on variables likely to be correlated with the location decisions of medium and large farms and farmer behavior. While they are able to speculate on reasons for identified relationships between SSF behavior and the presence of medium-scale farms, they are unable to identify the actual mechanisms that generate these spillover effects. They are also unable to determine if certain potential mechanisms (e.g., improved access to input or output markets versus sales coordination or knowledge transfers) are more important for particular SSF outcomes such as input use or productivity. Finally, we are aware of no studies that have explored the effects of the rise in MSFs on the incomes, productivity and degree of farm commercialization of neighboring SSFs.

Thus, this paper addresses these three observed gaps in the literature. We develop a theoretical model to explain some mechanisms through which spillovers on SSFs can be generated from the existence of larger farms around them. We empirically test this with data from Nigeria, Africa's largest economy and most populous nation. Second, we focus exclusively on MSFs (operating 5-50 hectares) as enterprises that are likely more accessible (than LSFs) to small-scale farmers.[1] Third, by exploring the spillover effects MSFs on a broader set of SSF outcomes including input use, productivity, commercialization and welfare captured via several measures of household income and poverty status, this paper provides a more comprehensive view of spillover effects.

Using comparative statics, our theoretic model yields some important empirical predictions - the effect of proximity to a medium-scale farm is mediated through knowledge spillovers, which we refer to as "learning effects" and the ability of the medium-scale farm to reduce the input related

[1] We find that even when we expand the definition to 5-100 as was done in Jayne et al. (2016) over 95% of our households were between 5 and 50 hectares and thus have used 5-50 so our work is comparable to both the literature using 5-100 (Jayne et al. 2016 and those using 50 such as Anseeuw et al (2016)

costs of the small farm, which we define as a “cost effect”. We characterize the learning effect as the result of receiving productivity-enhancing training from a medium-scale farm while the cost reducing effect emerges from reduced transactions costs from purchasing input from a neighboring medium-scale farm. We also explore the welfare effect of a third channel of selling output to the medium-scale farm since this may generate a combined learning and cost effect. This would occur if for example, MSFs provide training, input access and/or other forms of guidance to small farmers to meet the requirements of their buyers in addition to reducing SSFs cost of finding a market.

Interactions between SSF and MSFs are hardly random. More progressive SSFs may self-select into beneficial relationships with medium-scale farms, which potentially confounds identification and could lead to biased estimates of the spillover effects. To address this we use a two-step control function approach (CFA) as proposed by Wooldridge (2015). The key explanatory variables of interest: learning, cost and sales coordination effects are instrumented with the number of MSFs in the Local Government Area (LGA) of the SSF, conditional on LGA-level socioeconomic and agro-ecological factors (likely to attract medium-scale farms) as well as farmer and plot level characteristics. We argue the appropriateness of excluding the instrument in a falsification like test and examine the robustness of our finding to alternative considerations.

Consistent with theoretical predictions, we find evidence of knowledge and cost spillover effects of MSFs on SSF behavior and welfare. Receiving training and purchasing inputs from a medium-scale farm increases SSF productivity (yields) and welfare via increased incomes and lower poverty incidence and severity of poverty. While receiving training increases the likelihood and intensity of improved seed use, it has no effect on the use of fertilizer or crop protectants. Surprisingly, purchasing inputs from a medium-scale farm has no positive impact on any modern input use. This implies that the increased productivity observed from farmers who purchased inputs from MSFs is likely driven by improved access to higher quality inputs; a big challenge in sub-Saharan Africa (Poku et al., 2018) or the provision of training or other complementary services alongside the sale of inputs. While other studies have found a positive effect of proximity to large farms on yields and/or input use or welfare (Deininger and Xia, 2018, Glover and Jones, 2019, Lay et al., 2018) none that the authors are aware of, has been able to identify how that improvement came about. This study finds that for Nigeria, knowledge spillovers from actual training is driving limited expansion of modern input use and significantly enhancing farmer productivity and income. Direct access to inputs through MSFs increases productivity but this is not through increased likelihood (or quantity) of modern input use.

Compared to all interactions, improved access to output markets via sales to MSFs has the strongest welfare effects for small farmers. The opportunity to sell through MSFs enables SSFs receive a higher price; thereby boosting their crop and total income. This reduces their probability of being in poverty as well as the extent and severity of poverty they experience. Higher yields associated with sales coordination could occur through investments made in agricultural production to take advantage of improved access to a more guaranteed market and/or training offered to support the coordination activities of medium-scale farms.[2] This is consistent with Liverpool-Tasie et al (2020) who find that market outlets (e.g. agro-processors and wholesale traders) in the midstream of food value chains in developing countries are increasingly offering SSFs complementary services such as training and other inputs to ensure that they can get the

[2] Because our input use and crop yield determination occurs before our sales outcome (though through the effect of the number of MS farms in a farmers vicinity in the previous farming year) it is also possible that farmers who use higher inputs and have higher yields sold to MS farms we focus more on the welfare effects of sales coordination.

quantity and quality of products to meet their needs. On exploring this further, we find strong evidence that SSF productivity and welfare are significantly enhanced by more intense interaction with medium-scale farmers.

These findings have important implications for policy makers across Africa as they strive to improve SSF welfare while creating an environment for expanded food production to meet the demands of a rapidly growing populations and changing dietary patterns. This paper finds evidence in support of policies to encourage the beneficial co-existence of medium and small-scale farms. It documents the important role that MSFs are playing in improving SSF productivity and welfare via improved management practices and the opportunity to sell their output at more competitive prices. Finally, this study demonstrates that multiple interactions such as market access alongside training are necessary for positive productivity and welfare effects, which should be encouraged.

The remainder of the paper is organized as follows: Section 2 provides a review of the literature on interactions between small, medium and large-scale farms while Section 3 presents our theoretical model. Section 4 describes the data used while Section 5 presents the empirical strategy. Sections 6 presents the main study results while Section 7 presents our robustness checks. Section 8 concludes.

2. Background: mechanisms of interaction between small and medium and/or large farms

The potential spillovers from larger farms to small farms could be positive or negative. Medium and large-scale farms can enhance small farms' access to improved inputs and new technologies by bringing these resources to the areas in which they operate— making them more readily available to neighboring small farms (Burke et al., 2019; Amanor, 2011). If a high concentration of MSFs attracts private investment in farm input supply and service provision, we may expect small-scale farmers in such areas to face lower transactions costs of acquiring inputs. Moreover, SSFs may purchase inputs or services directly from medium-scale farms, also contributing to lower costs.[3]

Medium and large farms are often hailed as potential mechanisms for knowledge diffusion in rural areas. They can promote technology adoption via demonstration effect or their ability to experiment and discover new crops suited to a particular agroecology (Deininger et al; 2019, Ali et al., 2017; Deininger and Xia, 2016). Larger farms tend to demand specific (often high) input quality standards (Prowse, 2012). With the challenges associated with input quality in many African countries (Bold et al., 2017), larger farms also have more of an incentive (compared to small farms) to verify the quality of these inputs; given their importance in crop productivity and because they purchase inputs in large quantities. Thus, if SSFs are able to procure inputs from or with these medium-scale farms, they can potentially avoid low quality inputs (Ali et al., 2016), yielding higher productivity. This productivity-enhancing spillover may be complemented by knowledge spillovers that occur if larger farms hire and train labor from the local community or offer direct training services to SSFs— resulting in positive learning effects. Separate from learning, the opportunity for wage employment for SSFs is a potentially important source of income and improved livelihood for land-constrained households (Neven et al. 2009; Van den Broek et al., 2017; Wineman et al, 2020).

Conversely, MSFs (similar to LSFs) may induce negative spillovers on small farms. These

[3] In our study, approximately 27% of all smallholder farms reported purchasing inputs from a medium-scale farm

include higher food prices in areas with commercial farms (Schoneveld et al., 2011) as labor shifts from food production on small farms to large single-crop farms (Pryor and Chipeta, 1990). They might also crowd out SSF access to modern inputs where supply in particular geographic locations is limited. In the absence of cooperative interaction with SSFs, the presence of MSFs in an area could divert limited government and/or private extension services to the larger farms. In addition, lands suitable for community/SSF farming could be diverted towards medium and large-scale and commercial farming displacing SSF and or putting upward pressure on land rental and sales price (Jayne et al., 2012, Lundahl, 2015).

These potentially conflicting effects of MSFs on neighboring small farms is borne out in the existing literature (largely on large farms) and has led to a general lack of consensus in the literature on the precise effects of larger farms. For example, while Deininger and Xia (2016) find positive short-term effects of proximity to a large farm on small holder adoption of new practices and job creation in Mozambique, they do not find that LSFs improved the access of small farms to input and output markets. This contrasts with Ali et al. (2016) and Deininger et al (2019) that both find some positive effects of input use and risk-coping among small farms but not employment creation in Ethiopia. In Zambia, while Lay et al. (2018) find evidence of some positive spillovers on the ability of SSFs to expand their acreage, they also find reduced input (fertilizer) use associated with areas with high incidence of large farms. In addition, Deininger and Xia (2016) found that proximity to larger farms decreased the perceived well-being among local people due to disruptions in rural socioeconomic structures (Smalley, 2013, Deininger and Xia, 2016). This negative externality may be reinforced by the acquisition of large areas of lands by real estate firms as they speculate on the land prices in the vicinity of new MSFs (Smalley, 2013) making it harder for poor landless people to obtain lands.

This paper contributes to this ongoing debate with a novel analysis from a largely unexplored (for this topic) but important country in Africa, Nigeria. We develop a theoretic framework to test some of these spillover effects (particularly knowledge and cost spillover effects) and then empirically test for evidence of these in our data from Nigeria. We consider the effects of MSFs on the input use decisions, subsequent yield, sales and ultimate welfare of SSFs around them.

3. Theoretical Framework

We provide a simple framework for understanding how proximity to a medium-scale farm may yield spillover effects on neighboring small-scale farms. Consider a small-scale farming household in the spatial neighborhood of a medium-scale farm that maximizes utility $U(c, l, \alpha)$ where c, l and α refer to consumption, leisure, and a vector of household level covariates respectively. The households maximizes utility subject to its budget constraint given by $c = p * f(x(\omega), L, n, z) - C(\omega(v+t), L, h) - wn^h + wn^o + I \equiv M$ where consumption, c, is bounded by income, M, and v and t refer to transport and other transactions costs respectively. Here, p is output price, $f(\cdot)$ is a twice differentiable concave production function of non-labor inputs $x(\omega)$, proximity to a medium-scale farm, L and labor, n which equals the sum of hired labor n^h and family labor n^f. Let $C(\cdot)$ be cost function associated with production of all non-labor inputs. As standard, we maintain that the utility function is a concave and twice continuously differentiable function of c, l, and α. We will return to the meaning of ω subsequently. Quasi-fixed factors such as agroecological conditions of the farming area that affect farm output are represented by z. In addition to farm output income, the household has exogenous income I, earns a competitive wage,

w from selling labor off-farm, n^0 and hires labor, n^h for the same wage w. For simplicity, we chose to make labor cost additive. The household's time endowment is defined as T. Then it follows that the household utility maximization problem can be summarized as:

$$Max_{c,l}\, U(c,l\,,a) \;\; subject\ to\ c \; = f(x(\omega), L, n, z) - C(\omega(v+t), l, h) - wn^h + wn^o + I \equiv M \tag{1}$$

Where output price is normalized to 1

$$T = l\, + n^f + n^o \text{ and } \; n = n^f + n^h \tag{2}$$

Here, proximity to the medium-scale farm affects the utility of the household through its effect on SSF 's full income $Y(\cdot)$ i.e. profits plus other exogenous nonfarm income,
Where $Y(\cdot) = \; p * f(x(\omega), L, n, z) - C(\omega(v+t), L, h) - wn^h + wn^o + I$

Let the optimal utility from this problem be given by the indirect utility function
$$U^* = v(M^*, w, a) \tag{3}$$

Clearly, U^* is a function of optimal income (and thus farm profits) which in turn depends on proximity to the medium-scale farm, L. Thus, to obtain the effect of proximity to the medium-scale farm on SSF utility we evaluate its effect on the households profit function and hence income. To simplify this analysis, we consider a case where the SSF has to exert some effort, ω, to access inputs. The effort, ω includes transport cost v and other transactions costs t. Total input related costs (beyond market price) can be represented as $\delta = v + t$. Then input use, $x(\cdot)$, is an increasing function of efforts, ω, but is also decreasing in input related costs δ. The input-use function $x(\cdot)$ together with quasi-fixed factors, z, enter the farmer's production function, $f(\cdot)$ to determine the farmer's output. Also, we assume that proximity to the medium-scale farm, L, affects the small farm's input use through knowledge spillovers, i.e. the learning effect. For now, we assume proximity to the medium-scale farm as given. We shall return to the empirical implications of this assumption later.

In addition, the SSF faces a convex cost function, $C(\cdot)$ which is increasing in the input prices h and effort $\omega(\delta)$ (i.e. non-price input related costs). Thus, the small farm's cost function is also affected by proximity to the medium-scale farm through its effect on non-price input related costs. Using the information above, we can summarize the farmer's problem as a choice of effort level, ω, in order to maximize profits, π as given below
$$\pi = p * f(x(\omega), L, n,\; z) - C(\omega(v+t), L, h) \tag{4}$$
Normalizing output price to 1 and maximizing with respect to ω yields the following first-order condition:
$$f_x{}'(x(\omega), L, n,\; z)x_\omega{}'\,(\omega) = C_\omega{}'(\omega(v+t), L, h) \tag{5}$$

For notational simplicity, we can represent equation (5) as follows:
$$f_x{}'(x(\omega), L, n,\; z)x_\omega{}'\,(\omega) = C_\omega{}'(\omega(\delta), L, h) \tag{6}$$

Equation (6) above implies that the farmer chooses a level of efforts ω such that the marginal benefit in terms of output exactly compensates for the marginal input-related costs. However, we are interested in the effect of proximity to a medium-scale farm on the input vector $x(\cdot)$. That is, we are interested in the sign of $dx(\omega)/dL$.

Proposition 1

The effect of proximity to a medium-scale farm on the small farm's input vector $x(\cdot)$ and output is mediated by how the medium-scale farm's activities affect input-related costs and generate knowledge spillovers.

Proof.

To obtain $dx(\omega)/dL$, we will totally differentiate the first order condition obtained in equation (5) with respect to the variable of interest, L. Observe that $x(\omega)$ is an increasing function of ω. Hence to show how $x(\omega)$ changes with respect to L, it is sufficient to show what happens to ω as L changes.

That is, we just want to show how the effort exerted by SSF's change as proximity to a medium-scale farm increase. From the first order condition we know that $f_x{}'(x(\omega, L), n, z)x_\omega{}'(\omega) = C_\omega{}'(\omega(m), L, h)$.

Therefore, taking a total derivative with respect to l and solving for $dw/dL \equiv w_L$ gives the following:

$$f_x{}'(x(\omega), L, n, z)x_\omega{}'(\omega) = C_\omega{}'(\omega(m), L, h) \quad (7)$$

$$f_{xx}{}'(x(\omega), L, n, z)[x_\omega{}'(\omega)\omega_L + f_{xl}(x(\omega), L, n, z)]x_{\omega\omega}{}'(\omega) + f_x{}'(x(\omega), L, n, z)x_{\omega\omega}{}''(\omega)\omega_L = C_{\omega\omega(\delta)}{}''(\omega(\delta), L, h)\omega_L + C_{\omega L}{}'(\omega(\delta), L, h) \quad (8)$$

$$\frac{d\omega}{dL} = \omega_L \frac{f_{xL}(x(\omega), L, n, z)x_\omega{}'(\omega) - C_{\omega L}{}'(\omega(v+t), L, h)}{C_{\omega\omega(\delta)}{}''(\omega(\delta), L, h) - f_{xx}{}'(x(\omega), L, n, z)x_\omega{}'(\omega) - f_x{}'(x(\omega), L, n, z)x_{\omega\omega}{}''(\omega)}$$

Given that $f_x, x_\omega > 0$ and $f_{xx}, x_{\omega\omega} < 0$ by concavity and $C_{\omega\omega} > 0$ by convexity, the denominator is positive. Therefore, the sign of ω_L depends on the numerator which implies that it depends on the sign of $C'_{\omega L}$, the cost effect and f_{xL} the learning effect. ■

As the proof shows, the precise direction of the effect depends on the sign of $f_{xL}(x(\omega), l, n, z)x_\omega{}'(\omega) - C_{\omega L}{}'(\omega(v+t), L, h)$ which implies that it depends on the sign of $C'_{\omega L}$, the cost effect and f_{xL}, the learning effect. Even though several possibilities could emerge, we consider three natural cases:

(1) $f_{xl} > 0$ and $C'_{\omega l} < 0$. This is a case of pure positive spillovers where proximity to the medium-scale farm reduces input-related costs and induces knowledge and input quality spillovers unto the small farms.
(2) $f_{xl} < 0$ and $C'_{\omega l} > 0$. This is a pure negative spillover where the presence of the medium-scale farm increases transactions costs and also generates negative learning effect.
(3) $f_{xl} = 0$ and $C'_{\omega l} = 0$. This is the neutral case where proximity to the medium-scale farm has no significant effects on the SSF. This is also possible if the two terms cancel each other out.[4]

Since we assume that proximity to the MSFs impacts the SSF household utility through its effect

[4] These three cases are not exhaustive, as mentioned but give a general sense. Ultimately the point remains that the net effect depends on whether the positives outweigh the negatives and vice versa.

on total income Y, the spillover effects on the profit of SSFs should translate to improved welfare through its effect on productivity, hence income and poverty status. We explore proposition 1 using data from two Nigerian States that have recently experienced rapid growth of medium-scale farms.

4. Data and Study Sample

The two data sources for this paper come from the Agricultural Policy Research in Africa (APRA) 2018 survey for Nigeria. This dataset covers farms in two Nigerian states, Kaduna in North West Nigeria and Ogun in South West Nigeria. These states were purposively selected because of the significant steps they have taken in providing the necessary policy environment for the development of commercial agriculture. In each state, the largest local government area (LGA) based on total LGA land size was selected from each of the State's three senatorial districts. In each LGA, a complete listing of all households controlling (owned, rented in, borrowed, etc.) or operating five hectares and above was collected using a household listing protocol (available upon request). LGAs consist of wards (administrative units within LGAs numbering between 9 and 12), and each ward contains several communities, which may be villages or towns. The listing exercise was carried out across all three selected LGAs in both Kaduna and Ogun states between October 2017 and March 2018. These listing exercises resulted in the listing of 9,361 MFS in Kaduna and 5,848 MSFs in Ogun State (Muyanga et al., 2019). This listing data is our first main data source for all the information on the prevalence of MSFs in the LGA of a SSF.

The second data covers 1078 SSFs and 1031 MSFs randomly selected from sampling frames generated from the listing data. The dataset is a cross-section and contains detailed information on household socioeconomic characteristics including demographics, land holdings, assets and agricultural production and sales over the previous main agricultural season. We define SSFs as farmers who operate a total of less than five hectares of land. Medium-scale farmers were defined as those who operate between 5 and 50 hectares of land across crops. The number of MSFs in the local government area (LGA) of a smallholder farmer is restricted to the number of MSFs in existence prior to the input use and production and sales data used in the empirical section. Since the listing data includes information on the year the medium-scale farm started, we restrict our analysis to the number of MSFs in a given LGA in the year prior to the main agricultural season for which the input and output decisions we study were collected. This guarantees that the study outcomes and interactions between SSFs and MSFs are being related to the prevalence of MSFs in the vicinity prior to those outcomes or interactions.

5. Emprical Analysis

From propostion 1, the main empirical specification to test our learning effect hypothesis is expressed in equation (9), while equation (10) tests our cost effect hypothesis

$$y_{ig} = \alpha + [f_{xl}(x(\omega), l, n, z)x'_{\omega}(\omega)]T_{ig} + \beta_1 X_{1g} + \beta_2 X_{2ig} + \beta_3 X_{3ih} + \mu_{ig} \quad (9)$$

$$y_{ig} = \alpha + [C'_{\omega L}(\omega(v+t), L, h)]P_{ig} + \beta_1 X_{1g} + \beta_2 X_{2ig} + \beta_3 X_{3ih} + \mu_{ig} \quad (10)$$

where y_{ig} is the outcome variable of interest for small plot i in local government area g.[5] The outcomes we consider in our estimations can be grouped into three broad categories: welfare

[5] Apart from the income and poverty outcomes, all study outcomes are at the plot level.

outcomes, input use and output related outcomes. The input use variables we consider are the dichotomous use of improved seeds, inorganic fertilizer and agrochemical crop protectants and the log kilogram of each input used per hectare of land cultivated. The output related outcomes are crop yield, log crop income (in Naira) per hectare and the sale price per kilogram sold and a commercialization index measured by the proportion of harvested output that is sold. The welfare measures we explore are total income and subjective poverty. Subjective poverty is a self-reported measure that asks respondents, "How would you describe your household in general?". Responses that said the household was "struggling" or "unable" to meet household needs was coded as poor while those that said the household was "doing okay and able to meet their needs" was classified as non-poor. In addition, we consider an objective measure of poverty, which is defined as 1 if the household's per capita income is below the international poverty line of \$1.90 per day and 0 otherwise. Using this measure of poverty, we compute a measure of poverty gap, which equals the difference between the household's daily per capita income and the \$1.90 poverty line if the household is poor and 0 otherwise. Poverty severity is then obtained by squaring the poverty gap and is used as an additional outcome variable. The poverty severity measure allows us to examine the severity of poverty among the poor (Foster et al., 1984).

The right-hand side variables include LGA characteristics X_{1g} that affect the decision of MSFs to locate in the in LGA g of plot i, as well as other plot-specific (X_{2ig}) and household-level characteristics (X_{3ih}) while μ_{ig} is the error term. The parameter of interest here is $[f_{xl}(x(\omega), l, n, z)x'_{\omega}(\omega)]$ from proposition 1A, which is the measure of the direct effects of learning from a medium-scale farm while $[C'_{\omega L}(\omega(v+t), L, h)]$ measures the impact on y_{ig} of purchasing input from a medium-scale farm. However, due to unobservable factors such as ability or progressiveness that may influence SSFs to self-select into beneficial interactions with MSFs while simultaneously making those small farmers more likely to use particular inputs or have higher yields, our estimates of $[f_{xl}(x(\omega), l, n, z)x'_{\omega}(\omega)]$ and $[C'_{\omega L}(\omega(v+t), L, h)]$ are likely to be biased .

To be able to identify an unbiased estimate of our parameters of interest, we adopt a control function approach (CFA) as proposed by Imbens and Wooldridge (2007).[6] With the CFA, the generalized residuals from a first stage estimation of the determinants of interacting with MSFs is included in a second stage estimation of the effects of interacting with MSFs on SSF behavior and outcomes. In all second stage estimations, P-values are estimated via bootstrapping at 500 repetitions to account for the fact that the generalized residual came from a first stage regression estimation and the errors are clustered at the household level. As in the traditional instrumental variable (IV)/ two stage least squares (2SLS) approach, the CFA also requires at least one variable that is strongly correlated with a SSFs likelihood of interacting with a medium-scale farm but is uncorrelated with the unobserved factors that affect our outcome variables of interest (μ_{ig}) and thus appropriately excluded from (9) and (10). The estimates from this approach is more efficient although less robust than the IV estimator (Wooldridge, 2015). The excludable instrument used in this analysis is the number of MSFs in a SSFs LGA. Conditional on accounting for factors that influence the emergence of MSFs and interactions with a MSF, the coefficient on the number of MSFs should not be statistically different from zero. Accordingly, we argue that conditional on our rich set of LGA, household and plot control variables, the number of MSFs in a farmer's LGA

[6] It should be noted that the objective of this paper is to isolate the spillover effect that can be attributed to interacting with MSF farms. Thus we acknowledge that this approach does not speak to how things will change subsequent to future changes to access to MSFs

should not affect the farmer's input use behavior and farm outcomes except through the interactions necessary for our hypothesized spillover effects. These LGA characteristics include the mean distance to an all-weather road, total area cultivated in the LGA, total population density and mean labor productivity in the LGA measured as the mean crop yield per labor day as well the mean average rainfall over the last ten years preceding the medium-scale farm listing census. The rainfall data was extracted from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) database.

Because we are using a cross section dataset and not a panel, we cannot control for household time invariant unobserved factors. We attempt to address this limitation with a rich set of controls to capture the likely time invariant variables that could affect input use and welfare such as education, social capital and wealth. Farmer and household characteristics include age, gender, marital status, years of educational attainment and years of farming experience of the household head, household size and whether the household has any member engaged in non-farm activities, household land and livestock asset holdings (measured by the household's tropical livestock unit). We also control for whether the household has access to the government extension service. For plot characteristics, we control for whether the plot is owned or rented and whether the household has the land title to the plot. We also control for the plot size (in hectares), household distance to the plot (in kilometers), number of household members that worked on the plot in the last agricultural season, the soil type (clay, sandy or loamy) and the level of the parcel slope and terrace (flat, moderate terraced, moderate slope or steep slope).

To show that our exclusion restriction criterion is likely met and confirm that we have plausible reason to believe that we have appropriately accounted for enough factors to expect the coefficient on the number of MSFs to not statistically differ from zero, we also conduct a falsification like test. This test shows that conditional on our LGA and farmer controls, the number of medium scale farms does not significantly affect our study outcomes on input use, productivity and welfare (see section 7 for the full details). In addition to the number of MSFs in a SSF's LGA we also conduct our analysis using the share of the area in the LGA cultivated by MSFs and our results are almost identical.

In theory, the existence of positive spillovers will imply $f_{xl}(x(\omega), l, n, z)x'_\omega(\omega) > 0$ and $C'_{\omega L}(\omega(v+t), L, h) < 0$ while negative spillovers will imply that the converse is true in both cases. As mentioned above, for unbiased and consistent estimates of $f_{xl}(x(\omega), l, n, z)x'_\omega(\omega)$ and $C'_{\omega L}(\omega(v+t), L, h)$ we re-estimate equations (9) and (10) using the two-step control function approach.

To do this, we first estimate a non-linear reduced form model of the endogenous variables (interaction with medium-scale farms) on the instrument (i.e. the number of medium farms in the LGA of the SSF) and a rich set of covariates. Then we estimate the structural equation with the generalized residuals from the first stage non-linear estimation alongside the rich set of covariates included in the first stage. Specifically, obtaining the learning effect involves estimating the following conditional expectation of the outcome y_{ig} in equation (11)

$$E\left(y_{ig} \middle| Z_{ig}, T_{ig}\right) = \alpha + \delta Z_{ig} + \beta_0 T_{ig} + E(u_1 | Z_{ig}, T_{ig}) \quad (11)$$

which implies that we must be able to estimate $E\left(u_{ig} \middle| Z_{ig}, T_{ig}\right)$ where $Z_{ig} = (X_{ig}, \sum_i L_{ig})$ and $\sum_i L_{ig}$ is the number of MSFs in the LGA of farmer i . If equation (10) holds:

$$T_{ig} = 1[\delta_0 \sum_i L_{ig} + \beta_2 X_{2ig} + \beta_3 X_{3ih} + v_{ig} \geq 0] \quad (12)$$

Then $\left(u_{ig}, v_{ig}\right) \perp \sum_i L_{ig}$ and $E\left(u_{ig} \middle| v_{ig}\right) = \rho v_{ig}\ and\ v_{ig} \sim Normal(0,1)$ which implies that by iterated expectations:

$$E(u_{ig}|Z_{ig},T_{ig}) = E[E(u_{ig}|Z_{ig},v_{ig})|Z_{ig},T_{ig}] = \rho E(v_{ig}|Z_{ig},T_{ig}) \quad (13)$$

Which gives:

$$E(u_{ig}|Z_{ig},T_{ig}) = \rho\left[T_{ig}\left[\frac{\phi(\delta_0 \sum_i L_{ig})}{\Phi(\delta_0 \sum_i L_{ig})}\right] - (1-T_{ig})\left[-\frac{\phi(\delta_0 \sum_i L_{ig})}{\Phi(\delta_0 \sum_i L_{ig})}\right]\right] \quad (14)$$

where $\frac{\phi(\cdot)}{\Phi(\cdot)}$ is the inverse mills ratio. The estimate of δ_0 that is $\widehat{\delta_0}$ can then be obtained with a probit estimation. Using $\widehat{\delta_0}$, we can generate the generalized residual as follows:

$$\widehat{v_{ig}} = T_{ig}\left[\frac{\phi(\widehat{\delta_0}\{\sum_i L_{ig})}{\Phi(\widehat{\delta_0}\{\sum_i L_{ig})}\right] - (1-T_{ig})\left[-\frac{\phi(\widehat{\delta_0} \sum_i L_{ig})}{\Phi(\widehat{\delta_0} \sum_i L_{ig})}\right] \quad (15)$$

We then include $\widehat{v_{ig}}$ as a regressor in equation (14). This yields a structural equation of the form:

$$y_{ig} = \alpha + [f_{xl}(x(\omega),l,n,z)x'_{\omega}(\omega)]T_{ig} + \beta_1 X_{1g} + \beta_2 X_{2ig} + \beta_3 X_{3ih} + \rho\hat{v}_{ig} + \mu_{ig} \quad (16)$$

Where the $[f_{xl}(x(\omega),l,n,z)x'_{\omega}(\omega)]$ is the learning effect parameter and $\hat{v}_{ig}$ is the generalized residuals from the first stage without a need to adjust the standard errors in equation(14) for the first stage probit (Wooldridge, 2015) In both equations, X_{1g}, X_{2ig} and X_{3ih} remain as earlier defined (i.e. community characteristics that affect the decision of MSFs to locate in the in LGA g, plot-specific factors and household-level characteristics respectively. A straightforward test of $\rho = 0$ then tells us about endogeneity in the estimated model. We estimate equation (16) separately for the learning and cost effects and for the different outcomes of yield, crop income and sale price as well as input use decisions regarding improved seed use, inorganic and organic fertilizer as well as chemical protectants using linear and non-linear probit techniques as appropriate.

6. Results

We find evidence of significant interaction between small-scale farmers and medium-scale farmers in our sample (Table 1). Approximately 30% of the small-scale farmers reported to have received training on farm activities directly from a medium-scale farm. A similar percentage reported to have purchased inputs from a medium-scale farm and or sold their crop output to a medium-scale farm. These suggest the existence of important channels for knowledge and cost reduction spillovers. Government Agricultural extension in Nigeria is notably weak with a poor extension agent to farmer ratio of over 5,000 farm families to one agent in 2018 (NAERLS, 2018).[7] Studies have shown that extension agents are often not only ill-equipped to reach the many farmers allocated to them but have limited opportunities for training and thus lack correct information about many modern technologies (Ragasa and Mazunda, 2018). This creates ample room for improved productivity through knowledge transfer from commercial MSFs to SSFs around them. Low profitability of modern input use due to high transactions cost has also been documented in Nigeria (Takeshima and Liverpool-Tasie, 2015, Liverpool-Tasie et al., 2017, Liverpool-Tasie et al., 2014, Liverpool-Tasie, 2015). Thus, the opportunity to purchase inputs from MSFs could significantly reduce the transportation costs for SSFs. In addition, if medium-scale farmers have the ability to secure higher quality inputs (e.g. via the ability to test the quality of inputs, incentivize input suppliers to provide good quality inputs for a guaranteed market and/or better storage for inputs) then SSFs can also enjoy an input quality benefit from purchasing inputs from medium-

[7] NAERLS APS 2018 Report https://guardian.ng/saturday-magazine/cover/extension-agents-grossly-inadequate-to-deliver-services-to-farmers/

scale farms.

The mean annual total income in our study sample is N288, 000 (about $800)[8] with about 70% of it accounted for by crop income. About 40% of the study sample are below the income poverty line at $1.09 a day though a smaller share (25%) reported struggling to meet their family needs in the last year. Irrespective of the type of interaction, SSFs in our sample that interacted with a medium-scale farm are less likely to have reported to experience challenges in meeting their households' needs in the last year (Figure 1). Figure 1 also shows farmers that received training from or sold to MSFs tend to have higher incomes compared to those who do not. However, we do not see any difference in the share of farmers that use modern inputs among those that received training, sold to or purchased inputs from MSFs compared to those who did not (Figure 2). This might imply that the mechanism through which SSF welfare is improved through increased interaction with MSFs might lie outside of expanded use of modern inputs. Since these descriptive results do not control for the myriad of other factors that could explain welfare (income and probability of being in poverty) or input use, we confirm this with the empirical results from our CFA results presented in Tables 3-6.

Table 1 Interactions with medium-scale farms

Small scale farmer:	All crops	Cereals	Roots and tubers	Other Crops
Purchased inputs from a medium-scale farm	27%	30%	18%	21%
Received training on farm activities from a medium-scale farm	28%	25%	30%	41%
Sold farm output to medium-scale farm	28%	30%	24%	18%

Source: Author's calculation

[8] We use the exchange rate of N360=$1 that was prevalent in 2017

Table 2: Summary statistics of key outcome variables

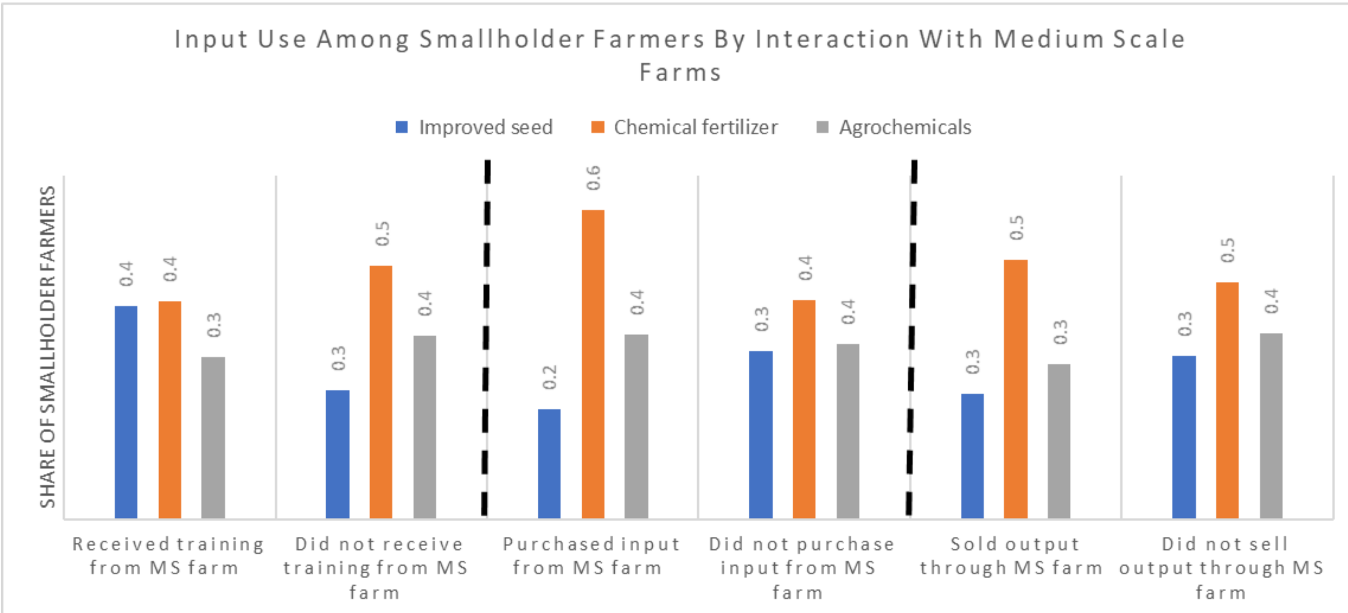
	Mean	Median	SD
Gross total income (Naira)	286,030	150,000	460,284
Gross crop income (Naira)	198,982	100,000	343,538
Share of households below income poverty line	0.39	0	0.49
Share of household that reported to be struggling to meet basic needs (Subjective poverty)	0.25	0	0.43
Poverty gap	0.24	0	0.34
Poverty severity	0.19	0	0.35
Sale price per kg	147.26	90.00	243.45
Share of output sold	0.66	0.80	0.38
Used improved seed (1/0)	0.24	0	0.43
Used chemical fertilizer (1/0)	0.53	1	0.50
Used agrochemicals (1/0)	0.38	0	0.49
Seeding rate for improved seed (kg/ha)	53.25	15.15	126.08
Kilograms of fertilizer used/ha	3.28	2.7	2.48
Kilograms of chemicals used /ha	5.59	4.00	5.43
N= 1, 783			

Source: Authors calculation

Figure 1: Differences in income and subjective poverty by interaction with medium-scale farms

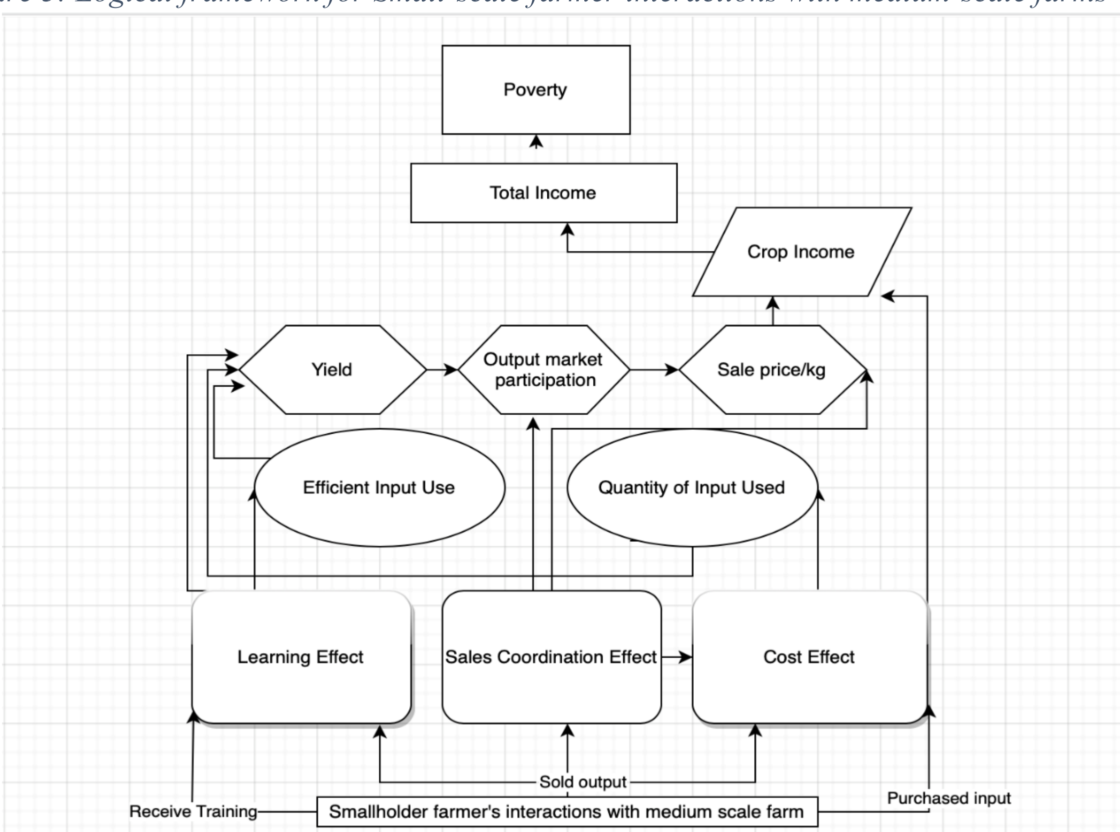
Source: Based on authors calculations

Figure 2: Input use among small-scale farmers by interaction with medium scale farms



Source: Based on authors calculations

Figure 3: Logical framework for Small-scale farmer interactions with medium-scale farms



Source: Authors

In line with the logical framework in Figure 3 we start with the basic hypothesis about whether interacting with medium-scale farmers is welfare enhancing for SSFs. Then we try to identify the mechanisms through which any observed welfare effects materialize, drawing from our theoretical framework. Table 3 presents the first stage regressions on the determinants of SSF interaction with medium-scale farms. These are the marginal effects from the non-linear probit results of the determinants of SSF interaction with medium-scale farms. As expected, it confirms that increased presence of MSFs in a SSFs LGA increases the likelihood of their interaction. The coefficients on the number of MS farms in the local government are highly significant at 1%, 5% and 10%[9] for receiving training, selling output to and purchasing inputs from MSFs respectively.

Table 4 presents the CFA results for our six welfare outcomes; crop income, total income,[10] household income poverty status (1/0), poverty gap (distance of total income per capita from the income poverty line), poverty severity (squared poverty gap) and subjective poverty status equal to 1 if household responded to be struggling or unable to meet household needs in the last year. The results indicate that receiving training from a medium-scale farm is associated with large, statistically significant welfare effects for small farmers. All other things being held constant, total and crop incomes will increase by 94% and 109% respectively for a small farmer who received training from a MS farm than one who did not. This higher income is associated with a statistically significantly lower probability of being income poor (by 11.6 percentage points) as well as a smaller poverty gap and severity for the poor. We find similar results for SSFs who purchased inputs from or sold outputs to medium-scale farms. These interactions are associated with higher crop and total income as well as lower poverty incidence, poverty gap and poverty severity at household level. Apart from the impact of sales to MSFs on income poverty (significant at 10%), all of the welfare impacts are statistically significant at 5% or less and large in magnitude. In each model, the significance of the generalized residuals from the first stage reveals the endogeneity of the training variable, but also correct for it (Rivers and Vuong, 1988, Smith and Blundell, 1986, Vella, 1993).[11]

To identify drivers of the observed welfare gains and test for evidence of cost and knowledge spillover on SSFs, we explore the impact of interacting with MSFs on SSF modern input use, productivity and commercialization. Tables 5 and 6 present these results. The only positive effect of interacting with MSFs on modern input use comes from receiving training. SSFs who received training from a medium-scale farm are 3.9 percentage points more likely to use improved seed and with higher seeding intensity. However, they are no more likely (than SSFs without such interaction) to use fertilizers or crop protectants. This might reflect the role that training can play in encouraging the adoption of modern technologies that are not commonly used (only 24% of the SSF sample are using improved seeds) compared to fertilizer that is currently already being used by 55% of SSFs in the sample (See table 2).

We do not find any evidence of cost spillovers on input use. Rather we find that SSFs who purchase inputs from medium-scale farmers are significantly less likely to purchase chemical fertilizers and use it with lower intensity compared to those who do not. They are no more likely to use improved seed or crop protectants. Though this result is consistent with Sipangule et al.

[9] Technically purchasing input is significant at 6%

[10] Total income in this study is the sum of incomes from all documented sources namely, non-farm income including regular and casual income, remittances and gifts as well as farm income from crop and livestock sales

[11] For all estimations, the coefficients on the generalized residuals for the CFA analysis are presented. However, when we fail to reject exogeneity (the coefficient on the generalized residual is not significantly different from zero), the coefficient on the OLS model is reported.

2017 (who find negative spillover effects of large farms on SSF fertilizer use) it is surprising. If purchasing inputs from MS farms guarantees a higher quality for inputs, then SSFs might not have to use excessive amounts of fertilizer to achieve desired yields because of uncertainty about product quality (Khor and Zeller, 2016). This might explain lower fertilizer intensity for these farmers compared to their counterparts purchasing from the open market. With increasing concerns about the overuse of chemicals in agricultural production, the lower probability of chemical fertilizer use might reflect negative messages passed on to SSFs from MSFs about chemical overuse or be a requirements imposed by these farms on SSFs as suggested earlier.[12] Farmers who sell to medium-scale are statistically significantly less likely to use crop protectants and at lower levels compared to those who don't' sell to medium-scale farms. The coefficient on chemical fertilizer use is negative but insignificant.[13]

Though only the provision of training by MSFs seems to promote SSF modern input use (improved seed), both receiving training and input purchase are consistently associated with statistically significantly higher yields (Table 6). This implies that the positive effects of MSFs on SSF productivity and welfare is largely not mediated through cost spillovers that expand modern input use. This yield improvement might occur through improved efficiency of modern input use from higher quality and/or through improved crop management practices through training. We also find that farmers who received training from medium-scale farms, purchased input from or sold output through MSFs receive a higher output price for their crops. Receiving training from a medium-scale farm is associated with receiving a sales price about N1.07 higher per kg sold. Purchasing inputs from and selling output to a medium-scale farm are associated with about N0.09 and N2.31 higher price per kg respectively.

Surprisingly, we find limited evidence of interacting with MS farms on the share of output sold by SSFs. This is consistent for all interactions. The average SSF in our sample sells almost 70% of their output. This high rate of commercialization might explain why we do not see much impact. However the higher price associated with being trained by a medium-scale farm or selling to them, definitely indicates some positive commercialization opportunities from medium-scale farms.

[12] This would particularly be the case if farmers tend to engage in multiple interactions with MS farms such as selling output to MS farms and also buying inputs from them or receiving training from them on input use and or other agricultural practices

[13] We only focus on the sales and welfare effects of selling to medium-scale farms as the input use decision occurs before the sale interaction.

Table 3: First stage results of the determinants of small-scale farmer interactions with MS farms

	Interaction with medium-scale farm:		
VARIABLES	**Purchased input**	**Received training**	**Sold output**
Number of medium-scale farms in LGA	0.004*	0.010***	0.005**
	(0.002)	(0.002)	(0.002)
No. of household members who worked on plot	-0.021**	-0.014	0.031***
	(0.010)	(0.011)	(0.011)
Soil type is Clay (1/0)	0.021	-0.411***	0.157
	(0.108)	(0.113)	(0.098)
Soil type is loamy (1/0)	0.057	-0.278***	0.189***
	(0.063)	(0.089)	(0.054)
Moderate terraced slope (1/0)	0.005	-0.188***	-0.003
	(0.054)	(0.037)	(0.058)
Moderate slope (1/0)	0.013	-0.062*	-0.028
	(0.042)	(0.037)	(0.040)
Steep parcel slope (1/0)	0.470***	0.109	0.444***
	(0.127)	(0.156)	(0.136)
The farmer has land title for this plot	0.021	0.159**	-0.272***
	(0.063)	(0.073)	(0.024)
Total Livestock Unit	0.002	-0.136***	0.076*
	(0.039)	(0.043)	(0.041)
Plot size (hectares)	0.001	-0.004	-0.008
	(0.003)	(0.003)	(0.005)
Distance to plot (kilometers)	-0.004	0.007	-0.025***
	(0.005)	(0.005)	(0.006)
Years of experience in farming	-0.004**	-0.004**	-0.005***
	(0.002)	(0.002)	(0.002)
Head married (1/0)	-0.009	-0.020	-0.011
	(0.016)	(0.017)	(0.018)
Head is male (1/0)	0.215***	0.050	-0.011
	(0.042)	(0.065)	(0.076)
Head education in years	0.005**	0.007***	0.003
	(0.002)	(0.002)	(0.002)
Head age in years	0.003*	0.002	0.005***
	(0.002)	(0.001)	(0.002)
Household size	-0.005	-0.004	-0.009*
	(0.005)	(0.005)	(0.005)
Household has non-ag worker (1/0)	-0.107***	-0.209***	0.068*
	(0.028)	(0.023)	(0.035)
Mean distance to all weather road in LGA (kilometers)	-0.003	-0.055	-0.138*
	(0.074)	(0.071)	(0.076)
Mean LGA Productivity/ha	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Total area cultivated in LGA	0.050	0.431	0.619
	(0.378)	(0.355)	(0.387)
LGA Population density	0.505	2.263	4.033
	(2.418)	(2.227)	(2.465)
Mean annual rainfall in LGA over the period 2007-2017	-0.002	0.001	-0.003*
	(0.002)	(0.002)	(0.002)
State==Ogun	-0.234	0.366**	0.249
	(0.193)	(0.180)	(0.176)
Number of Observations	1,709	1,687	1,687

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4: Impacts of interaction with medium-scale (MS) farms on small-scale farmer welfare outcomes

	Received training from MS farm (1/)	**Residuals**	**Purchased input from MS farm (1/0)**	**Residuals**	**Sold output through MS farm (1/0)**	**Residuals**
Subjective poverty incidence	0.248	-0.167*	-0.123***	-0.103	-0.468**	0.220*
	(0.160)	(0.096)	(0.023)	(0.218)	(0.218)	(0.131)
Household is income Poor (1/0)	-0.116***	0.109	-0.084**	0.001	-0.554*	0.315*
	(0.032)	(0.145)	(0.034)	(0.233)	(0.292)	(0.176)
Poverty gap	-0.068***	-0.002	-0.094***	0.008	-0.420***	0.220***
	(0.025)	(0.084)	(0.026)	(0.104)	(0.127)	(0.077)
Poverty severity	-0.050**	-0.024	-0.100***	0.031	-0.437***	0.228***
	(0.023)	(0.079)	(0.023)	(0.095)	(0.121)	(0.073)
Inverse-hyperbolic sine of total income (naira)	0.663**	0.399	4.756**	-2.005*	7.832***	-4.246***
	(0.271)	(1.360)	(1.961)	(1.121)	(2.479)	(1.464)
Inverse-hyperbolic sine of crop income (naira)	0.739**	0.316	8.773***	-4.054***	11.172***	-5.807***
	(0.323)	(1.310)	(2.207)	(1.254)	(2.976)	(1.794)

Bootstrapped standard errors clustered at household level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All estimations include all the control variables used in the first stage.

Table 5: Impacts of interaction with medium-scale farms on input use among small-scale farmers

	PANEL A					
VARIABLES	**Used Improved Seeds (1/0)**		**Used Fertilizer (1/0)**		**Used Crop Protectants (1/0)**	
	(1)	(2)	(3)	(4)	(5)	(6)
Received training from MS farm (1/0)	0.039*		0.023		0.025	
	(0.024)		(0.022)		(0.024)	
Learning effect residuals	-0.013		0.034		-0.117	
Purchased input from MS farm (1/0)		-0.214		-0.643**		0.289
		(0.236)		(0.276)		(0.359)
Cost effect residuals		0.103		0.459***		-0.204
Observations	1,539	1,561	1,671	1,693	1,669	1,691
	PANEL B					
	Quantity of improved seed used per ha		**Quantity of fertilizer used per ha**		**Quantity of crop protectants used per ha**	
	(1)	(2)	(3)	(4)	(5)	(6)
Received training from MS farm (1/0)	25.434*		-3.252		7.047	
	(14.246)		(5.870)		(4.910)	
Learning effect residuals	-2.640		2.066		-3.637	
	(58.501)		(3.570)			-0.951
Purchased input from MS farm(1/0)		-134.129		-23.975***		(9.311)
		(120.884)		(7.481)		0.230
Cost effect residuals		71.097		15.183***		(5.647)
Observations	1,557	1,579	1,687	1,709	1,684	1,706

Bootstrapped standard errors clustered at household level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
All estimations include all the control variables used in the first stage

Table 6: Yield and commercialization impacts of small-scale farmer interaction with medium-scale (MS) farms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES		Yield/ha			Sale price/kg			Commercialization level	
Received training from MS farm (1/0)	0.156**			1.069***			0.134		
	(0.068)			(0.404)			(0.107)		
Learning effect residuals	-0.080**			-0.501**			-0.081		
	(0.041)			(0.248)			(0.063)		
Purchased input from MS farm (1/0)		0.016**			0.095*			-0.198	
		(0.008)			(0.054)			(0.129)	
Cost effect residuals		-0.028			-0.303			0.116	
		(0.048)			(0.282)			(0.079)	
Sold output through MS farm (1/0)			0.391***			2.307***			0.223
			(0.093)			(0.513)			(0.140)
Coordination effect residuals			-0.232***			-1.327***			-0.134
			(0.057)			(0.311)			(0.086)
Observations	1,660	1,691	1,660	1,364	1,382	1,363	1,382	1,382	1,363

Bootstrapped standard errors clustered at household level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All estimations include all the control variables used in the first stage

1. Robustness checks and additional considerations:

7.1 Testing the exclusion restriction

This study finds strong evidence of positive welfare impacts for SSFs that engage with MSFs in their communities. The significance of the generalized results in some of the CFA results in tables 4-6 reveals that the interactions between small-scale farmers and MSFs are endogeneous for many of the outcome variables. As indicated earlier, our identification strategy is based on the number of MSFs in a SSF's local government area being an appropriate instrument. While this instrument is strongly correlated with our endogenous variable (interaction with a medium-scale farm), it is not usually possible to test if an instrument satisfies the exclusion restriction. In our model, we argue that our instrument satisfies the exclusion restriction. Conditional on the rich set of farmer and LGA characteristics (such as higher agricultural potential, better input markets, infrastructure, market access) that might be correlated with both the choice of location of the medium-scale farm (and farmers interaction with them) and SSF input use, productivity and welfare, the number of MSFs in the LGA of a smallholder farm shouldn't matter for input use decisions and farm outcomes. Thus, leaving learning and cost-reduction channels as the only paths via which MSFs can affect SSF outcomes and behavior.

To confirm this, we estimate equation (17)

$$y_{ig} = \alpha + \beta_0 \sum_i L_{ig} + \beta_1 X_{1g} + \beta_2 X_{2ig} + \beta_3 X_{3ih} + \mu_{ig} \tag{17}$$

Where y_{ig}, X_{1ig}, X_{2ig} and X_{3ih} are all as earlier defined and $\sum_i L_{ig}$ is the number of MSFs in LGA g. Conditional on the LGA-specific characteristics, X_{1ig}, that might affect the number of MSFs $\sum_i L_{ig}$ in the in LGA g of plot i, as well as other plot and household-level characteristics, we would expect $\beta_0 = 0$. Thus, by estimating equation (17) we argue that the absence of a direct effect of the number of MSFs on farmer behavior is a likely indication that our exclusion restriction for the instrumental variable $\sum_i L_{ig}$ is met. Table 7 presents the results of our estimation of equation (17). The estimated coefficients from equation (17) are consistently statistically zero. These findings validate the assumption that exclusion restriction likely holds for the number of MSFs conditional on the LGA variables that may affect the number of medium-scale farms located there. Although we include total cultivated area in the LGA as a control, we also explore an alternative instrument, using the number of MSFs as a share of the total area of land cultivated in the LGA.[14]

Table 7: Exclusion restriction plausibility test results for all study outcomes

VARIABLES	(1) Subjective poverty incidence	(2) Household is Income poor	(3) Poverty Gap	(4) Poverty Severity
Number of MS farms	-0.001	0.000	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)

[14] These results are available upon request.

Observations	1,754	1,765	1,765	1,765
	(5)	(6)	(7)	(8)
	Crop Income	Total Income	Non-farm Income	Yield/ha
Number of MS farms	-0.004	0.017	0.007	0.001
	(0.006)	(0.020)	(0.053)	(0.001)
Observations	1,622	1,747	1,747	1,622
	(9)	(10)	(11)	(12)
	Sale Price	Commercialization	Fertilizer Use	Improved Seed (0/1)
Number of MS farms	0.010	-0.000	-0.001	-0.001
	(0.007)	(0.002)	(0.002)	(0.003)
Observations	1,442	1,747	1,626	1,618
	(13)	(14)	(15)	(16)
	Used Agrochemical	Improved seed (Kg/ha)	Fertilizer (Kg/ha)	Agrochemical (Kg/ha)
Number of MS farms	-0.001	-1.029	-0.048	-0.081
	(0.005)	(1.502)	(0.079)	(0.109)
Observations	1,772	1,629	1,774	1,771

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All estimations include all the control variables used in the first stage.

7.2 Extent of interaction

We consistently find significant positive productivity and welfare impacts of small farmer interaction with medium-scale farms. While the higher yields associated with interacting with MSFs is consistent across interactions, it is not clear if this is driven solely by the particular interaction in question or if it is partly driven by other interactions that those SSFs might be simultaneously engaged in with medium-scale farms. In a systematic evidence synthesis, Liverpool-Tasie et al (2020) finds that in addition to serving as a marketing channel for small farmers, economic agents in the midstream of food value chains across Asia and Africa (such as wholesalers and agro-processors) are increasingly offering SSFs complementary services such as training and inputs. Providing these complementary services are presumably mutually beneficial as they ensure that they get the quantity and quality of products to meet their needs for processing and/or sale further down the value chain. They also find that the provision of these additional services is positively correlated with the probability that an interaction between a small farmer and an SME in the midstream of the food value chain yielded a positive outcome for the small holder. To explore the extent to which MSFs might be playing similar roles as these SMEs and to confirm if the multiplicity of interactions is important for the observed welfare gains to small farmers, we explore the extent to which MSFs simultaneously provide training, input purchase and output sale opportunities to SSFs. Then we check if this increased intensity of interaction is necessary for consistent positive welfare and productivity gains. Table 8 shows that while majority of the small-scale farms that interact with medium-scale farms tend to engage with them in only one way (either purchasing an input from them or being trained by them or selling output to them), a significant share engage with MS farms in more than one way. Forty-three percent engage in at least two different interactions while about 15% engage in all the three activities we explored. This indicates that there may be complementary service provisions by MSFs to small farms around them and/or that there are opportunities for the combined effect of access to different inputs or complementary services and output market access.

Table 8: Extent of interaction between small farmers and medium-scale farms

Small scale farmer reported	Total Sample
Only one interaction (conditional on at least one)	58%
Only two interactions (conditional on at least one)	27%
At least two interactions (conditional on at least one)	43%
All three interactions (conditional on at least one)	15%
Average number of interactions (conditional on at least one)	1.6

Source: Authors calculations

Tables 9 and 10 present the results of the productivity and welfare impacts of more intense interactions with MS farms. Again, we apply the CFA and consider three measures of intensity; first the case when a SSF has only one interaction, second, two or more interactions and third, the number of interactions with a medium-scale farm that a small farmer has. Table 9 presents the first stage results of the CFA conducted via a probit model for the probability of having only one interaction or two or more interactions. The first stage for the number of interactions is a Poisson model to account for the fact that our outcome is a count variable with a few numbers of potential outcomes (maximum of 3). We confirm (table 9) that the number of interactions and having at least 2 interactions are all statistically significantly correlated (at 5% or less) with the number of MSFs in a SSF 's LGA. The first stage regression reveals that the relationship between only having one kind of interaction with a medium-scale farm is not significantly correlated with the number of medium-scale farms. While this precludes us making any causal claims, we still explore the correlations between having only one interaction with a medium-scale farm and our study outcomes and see if that differs from the impact of those who have multiple interactions.[15]

Table 10 clearly reveals that higher number of interactions between a small farmer and a MS farm is more consistently associated with positive productivity and welfare impacts. Having one more interaction reduces a SSFs probability of being in income poverty by about 9 percentage points. A farmer who has at least two types of interaction with MSFs is 62 percentage points less likely to have reported having struggled to meet their household needs.

Similarly, while a household with only one interaction does not record having higher yield or receiving a higher sales price nor recording a higher crop or total income, farmers with more interactions tend to have higher crop and total incomes, sales price and yields. These positive impacts are all statistically significant and large in magnitude. We do not find any evidence of expanded modern input use or levels from engaging with MS farms in multiple ways. The limited evidence of multiple interactions on input use is to reduce the probability and/or intensity of modern input use. This confirms the earlier finding that the productivity impacts from engaging with SSFs is likely mediated through improved management practices and access to better quality inputs rather than promoting more modern input use. The welfare impacts occur through improved yields and sales price enabling small-scale farmers to enjoy higher crop incomes and lower

[15] We expect that the endogeneity of the interaction variable to cause our estimates on input use and welfare outcomes to be upwardly biased and thus likely the upper bounds of any effect. Our results are largely insignificant indicating that they are an appropriate baseline confirming the broader findings of limited effects.

probability of struggling to meet household needs.

Table 9: First stage results of the determinants of the intensity of small-scale farmer interactions with MS farms

	(1)	(2)	(3)
VARIABLES	Number of interactions	Two or more interactions	Only one interaction
Number of MS farms	0.012***	0.005***	0.002
	(0.004)	(0.002)	(0.321)
Other controls	Y	Y	Y
Observations	1,774	1,743	1,774

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All estimations include all the control variables used in the first stage

Table 10: Welfare impacts of multiple interactions between medium-scale farms and small-scale farmers

PANEL A	Household is income poor (=1)			**Poverty gap**			**Poverty severity**		
Only one interaction	-0.008			-0.004			0.004		
	(0.032)			(0.023)			(0.021)		
Two or more interactions		-0.088**			-0.461**			-0.468**	
		(0.037)			(0.218)			(0.184)	
Number of interactions			-0.048***			-0.045***			-0.043***
			(0.016)			(0.012)			(0.011)
Observations	1,774	1774	1,743	1,774	1,743	1,774	1,774	1,774	1,743
	Total income			**Crop income**			**Subjective poverty incidence**		
Only one interaction	-0.125			-0.206			0.048		
	(0.256)			(0.274)			(0.02)		
Two or more Interactions		0.587***			7.037**			-0.620***	
		(0.130)			(2.975)			(0.20)	
Number of interactions			6.558***			9.374***			-0.057***
			(1.763)			(1.948)			(0.01)
Observations	1,774	1,743	1,774	1,774	1,743	1,774	1,763	1,732	1,732

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All estimations include all the control variables used in the first stage.

Table 11: Yield, commercialization and input use impacts of multiple interactions between medium-scale farms and small-scale farmers

VARIABLES	**Yield/ha**			**Sale price/kg**			**Commercialization**		
Only one interaction	-0.013			-0.050			0.016		
	(0.009)			(0.057)			(0.013)		
Number of interactions		0.164*			1.289**			0.001	
		(0.092)			(0.556)			(0.005)	
Two or more Interactions			0.196***			1.203***			-0.194
			(0.068)			(0.434)			(0.125)
Observations	1,747	1,747	1,747	1,430	1,430	1,422		1,430	1,422
VARIABLES	**Improved seed use (1/0)**			**Fertilizer use (1/0)**			**Crop protectant use (1/0)**		
Only one interaction	0.053***			-0.016			-0.030		
	(0.019)			(0.019)			(0.024)		
Number of interactions		0.009			0.015			-0.020*	
		(0.009)			(0.010)			(0.011)	
Two or more Interactions			-0.076			-0.435**			0.152
			(0.164)			(0.178)			(0.194)
Observations	1,809	1,747	1,725	1,820	1,758	1,727	1,818	1,756	1,725
VARIABLES	**Kg of improved seed used/ha**			**Kg of fertilizer used/ha**			**Kg of crop protectant used/ha**		
Only one interaction	13.591			0.427			-1.291**		
	(10.641)			(0.952)			(0.568)		
Number of interactions		-259.057*			6.765			1.594	
		(156.359)			(8.187)			(5.513)	
Two or more Interactions			-8.390			-22.624***			-7.127
			(14.771)			(6.810)			(5.929)
Observations	1,598	1,629	1,619	1,743	1,774	1,763	1,740	1,771	1,760

Note: Bootstrapped standard errors clustered at household level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All estimations include all the control variables used in the first stage.

2. Conclusions:

The last decade has seen a rapid rise in the number of medium-scale farms (and share of land cultivated by them) in sub-Saharan Africa. While these commercialized farms are a potential mechanism to increase food production to meet Africa's rapidly growing population, there is limited empirical evidence on the myriad of ways through which this could happen. Beyond their own production, MSFs could support expanded food production and other dimensions of the structural transformation process if they also increase the productivity, commercialization and ultimate income and welfare of small-scale farmers around them. This could occur through increased access to input and output markets, knowledge and employment. On the other hand, their presence could compete with SSFs for land, limited modern inputs and poor government extension services. Ultimately, the empirical evidence on the impacts of the recent rapid growth of MSFs on small producers around them remains extremely limited.

Consequently, this paper examines if MSFs in Nigeria have an impact on the farming behavior and welfare of SSFs in their vicinity. We find strong evidence of positive welfare impacts for SSFs that engage with MSFs in their communities. For Nigeria, knowledge spillovers from actual training is a key driver of farmer productivity and ultimate welfare. This appears to be partly through some impacts on modern input use (largely improved seed), but likely more through improved agricultural practices. While purchasing inputs from MS farms does not increase modern input use, it is still associated with higher yields, crop income and lower probability of income and subjective poverty. We find that the opportunity to sell to MSFs is a very important source of improved welfare in our study sample. It enables SSFs receive higher prices, crop and total incomes and thus experience lower probability of being poor (and lower poverty gap and severity). We also find that having more than one interaction with MSFs (e.g. the ability to sell to them while also receiving training or purchasing inputs) is important to guarantee the improved welfare for small-scale farmers.

Our findings suggest that in areas where significant interactions between SSFs and MSFs can take place (to link these SSFs to training as well as output markets and high-quality inputs); there are likely benefits from government and/or donor support of these larger farms. With such poorly functioning government extension services and longstanding issues about input quality, leveraging on MSFs to facilitate the diffusion of new technologies could be extremely beneficial.

References:

ALI, D., DEININGER, K. & HARRIS, A. 2016. *Large farm establishment, smallholder productivity, labor market participation, and resilience: evidence from Ethiopia*, The World Bank.

ALI, D., DEININGER, K. & HARRIS, A. 2019. Does large farm establishment create benefits for neighboring smallholders? Evidence from Ethiopia. *Land Economics,* 95**,** 71-90.

AMANOR, K. S. Global landgrabs, agribusiness and the commercial smallholder: a West African perspective. International Conference on Global Land Grabbing, Brighton, 2011. 6-8.

ANSEEUW, W., JAYNE, T., KACHULE, R. & KOTSOPOULOS, J. 2016. The quiet rise of medium-scale farms in Malawi. *Land,* 5**,** 19.include page numbers

BOLD, T., KAIZZI, K. C., SVENSSON, J. & YANAGIZAWA-DROTT, D. 2017. Lemon technologies and adoption: measurement, theory and evidence from agricultural markets in Uganda. *The Quarterly Journal of Economics,* 132**,** 1055-1100.

BURKE, W. J., JAYNE, T. S. & SITKO, N. J. 2019. Do Medium-scale Farms Improve Market Access Conditions for Zambian Smallholders? *Journal of Agricultural Economics*.

CHAMBERLIN, J. and JAYNE, T.S. 2020. Does farm structure affect rural household incomes? Evidence from Tanzania *Food Policy* (90), 101865.

DEININGER, K. & XIA, F. 2016. Quantifying spillover effects from large land-based investment: the case of Mozambique. *World Development,* 87**,** 227-241.

DEININGER, K. & XIA, F. 2018. Assessing the long-term performance of large-scale land transfers: Challenges and opportunities in Malawi's estate sector. *World Development,* 104**,** 281-296.

FOSTER, J., GREER, J. & THORBECKE, E. 1984. A class of decomposable poverty measures. *Econometrica: journal of the econometric society***,** 761-766.

GLOVER, S. & JONES, S. 2019. Can commercial farming promote rural dynamism in sub-Saharan Africa? Evidence from Mozambique. *World Development,* 114**,** 110-121.

HERRMANN, R. T. 2017. Large-scale agricultural investments and smallholder welfare: A comparison of wage labor and outgrower channels in Tanzania. *World Development,* 90**,** 294-310.

IMBENS, G. & WOOLDRIDGE, J. 2007. Control function and related methods. What's new in Econometrics. *Lecture Notes,* 6.

JAYNE, T., CHAMBERLIN, J. & MUYANGA, M. Emerging land Issues in African agriculture: implications for food security and poverty reduction strategies. paper presented as part of Stanford University's Global Food Policy and Food Security Symposium Series, sponsored by the Center for Food Security and the Environment and the Freeman Spogli Institute for International Studies, Stanford, California, 2012.

JAYNE, T., MUYANGA, M., WINEMAN, A., GHEBRU, H., STEVENS, C., STICKLER, M., CHAPOTO, A., ANSEEUW, W., VAN DER WESTHUIZEN, D. & NYANGE, D. 2019. Are medium-scale farms driving agricultural transformation in sub-Saharan Africa? *Agricultural Economics*.

JAYNE, T. S., CHAMBERLIN, J., TRAUB, L., SITKO, N., MUYANGA, M., YEBOAH, F. K., ANSEEUW, W., CHAPOTO, A., WINEMAN, A. & NKONDE, C. 2016. Africa's changing farm size distribution patterns: the rise of medium-scale farms. *Agricultural Economics,* 47**,** 197-214.

JAYNE, T. S., CHAPOTO, A., SITKO, N., NKONDE, C., MUYANGA, M. & CHAMBERLIN, J. 2014. Is the scramble for land in Africa foreclosing a smallholder agricultural expansion strategy? *Journal of International Affairs***,** 35-53.

KHOR, L. Y. & ZELLER, M. 2016. The Effect of Household Wealth on Fertiliser Use in the Presence of Uncertainty. *The Journal of Development Studies,* 52**,** 1034-1045.

LAY, J., NOLTE, K. & SIPANGULE, K. 2018. Large-scale farms and smallholders: Evidence from Zambia. GIGA Working Papers.

LIVERPOOL-TASIE, L. 2016. Is fertilizer use really suboptimnal in sub-Saharan Africa? The case of rice in Nigeria.*Journal of Agricultural Economics* Volume and number and pages missing.

LIVERPOOL-TASIE, L. S. O., OMONONA, B. T., SANOU, A. & OGUNLEYE, W. O. 2017. Is increasing inorganic fertilizer use for maize production in SSA a profitable proposition? Evidence from Nigeria. *Food policy,* 67**,** 41-51.

LUNDAHL, M. 2015. *Incentives and Agriculture in East Africa (Routledge Revivals)*, Routledge.

MUYANGA, M., AROMOLARAN, A., JAYNE, T., LIVERPOOL-TASIE, S., AWOKUSE, T. & ADELAJA, A. 2019. Changing Farm Structure and Agricultural Commercialisation in Nigeria.

MUYANGA, M. & JAYNE, T. 2019. Revisiting the Farm Size-Productivity Relationship Based on a Relatively Wide Range of Farm Sizes: Evidence from Kenya. *American Journal of Agricultural Economics*.

POKU, A.-G., BIRNER, R. & GUPTA, S. 2018. Why do maize farmers in Ghana have a limited choice of improved seed varieties? An assessment of the governance challenges in seed supply. *Food security,* 10**,** 27-46.
PROWSE, M. 2012. Contract farming in developing countries: a review.
PRYOR, F. L. & CHIPETA, C. 1990. Economic Development through Estate Agriculture: The Case of Malaŵi. *Canadian Journal of African Studies/La Revue canadienne des études africaines,* 24**,** 50-74.
RAGASA, C. & MAZUNDA, J. 2018. The impact of agricultural extension services in the context of a heavily subsidized input system: The case of Malawi. *World development,* 105**,** 25-47.
RIVERS, D. & VUONG, Q. H. 1988. Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of econometrics,* 39**,** 347-366.
SCHONEVELD, G. C., GERMAN, L. A. & NUTAKOR, E. 2011. Land-based investments for rural development? A grounded analysis of the local impacts of biofuel feedstock plantations in Ghana. *Ecology and Society,* 16.
SMALLEY, R. 2013. Plantations, contract farming and commercial farming areas in Africa: A comparative review.
SMITH, R. J. & BLUNDELL, R. W. 1986. An exogeneity test for a simultaneous equation Tobit model with an application to labor supply. *Econometrica: Journal of the Econometric Society***,** 679-685.
TAKESHIMA, H. & LIVERPOOL-TASIE, L. S. O. 2015. Fertilizer subsidies, political influence and local food prices in sub-Saharan Africa: Evidence from Nigeria. *Food Policy,* 54**,** 11-24.
VELLA, F. 1993. A simple estimator for simultaneous models with censored endogenous regressors. *International Economic Review***,** 441-457.
WINEMAN, A., THOMAS, S. JAYNE, EMMA ISINIKA MODAMBA AND HOLGER KRAY.2020. Characteristics and Spillover Effects of Medium-Scale Farms in Tanzania. Working Paper
WOOLDRIDGE, J. M. 2015. Control function methods in applied econometrics. *Journal of Human Resources,* 50**,** 420-445.
XIA, F. & DEININGER, K. 2019. Spillover Effects of Tobacco Farms on the Labor Supply, Education, and Health of Children: Evidence from Malawi. *American Journal of Agricultural Economics*.