



Does farm structure affect rural household incomes? Evidence from Tanzania

Jordan Chamberlin^{a,*}, T.S. Jayne^b

^a International Maize and Wheat Improvement Center (CIMMYT), Kenya

^b Michigan State University, United States



ARTICLE INFO

Keywords:

Household income
Inequality
Spillovers
Farm structure
Land concentration
Farm size distributions
Tanzania, Africa

ABSTRACT

Many African countries have recently experienced rapid growth in the numbers of medium- and large-scale farms. These developments have generated considerable speculation about the impacts of farmland concentration and inequality on smallholder households and communities. This study exploits inter-district variation in farm landholding patterns in Tanzania to determine how differences in localized farmland structure affect rural household incomes using nationally representative household panel survey data. Because farm structure is a multifaceted concept, five alternative indicators of farmland structure are defined for 142 districts in Tanzania: (i) the Gini coefficient; (ii) skewness; (iii) coefficient of variation; (iv) share of controlled farmland under medium-scale farms; and (v) share of controlled farmland under large farms. These alternative farm structure variables are included in models of rural household income to test their effects after controlling for available household and community covariates. The study highlights four main findings. First, most indicators of farmland concentration are positively associated with rural household incomes. Second, household incomes from farm and non-farm sources are positively and significantly associated with the share of land in the district controlled by farms in the 5–10 and 5–20 ha category. Third, these positive spillover benefits are smaller and less statistically significant in districts with a relatively high share of farmland controlled by farms over 20 ha in size. Fourth, poor rural households are least able to capture the positive spillovers generated by medium-scale farms and by concentrated farmland patterns.

1. Introduction

This study is motivated by the need to better understand the impacts of changing farm structure in Africa. As medium- and large-scale farms have acquired substantial amounts of farmland in Africa since 2000 (Deininger and Byerlee, 2011; Schonevald, 2014; Jayne et al., 2014a), important questions arise about the impacts of localized farmland inequality and size distribution patterns on smallholder households and rural communities. This study traces its origins to a longstanding strand of the development economics literature arguing that the scale and distribution of farmland holdings tends to influence local demand patterns, factor markets, and associated multiplier effects of agricultural growth (Johnston and Mellor, 1961; Mellor, 1976; Vollrath, 2007; Eastwood et al., 2010). Most of the evidence undergirding this literature is drawn from Asia and Latin America; to our knowledge there is virtually no applied evidence testing the relationship between localized farm distribution patterns and rural household incomes in

Africa.

To address these issues, this study exploits inter-district variation in farmland distribution patterns in Tanzania to determine its impact on rural household incomes and other characteristics using nationally representative household panel survey data. Our key research question is whether or not the concentration of farmland under medium- and large-scale farms influences the economic outcomes of households within the same areas. Our motivation starts with the longstanding recognition that farm structure may influence the pace of income growth (Johnston and Mellor, 1961; Johnston and Kilby, 1975; Vollrath, 2007) and by evidence that farm structure is changing in many African countries (Jayne et al., 2016). A stylized fact from Asia's agricultural development experience is that relatively uniform land distribution patterns may stimulate rural development more effectively than highly concentrated landholding patterns. Smallholders have high marginal propensities to consume and spend their money in the local rural economy, thereby stimulating growth linkages between farm and non-farm

* Corresponding author at: International Maize and Wheat Improvement Center (CIMMYT), ICRAF House, United Nations Avenue, Gigiri, P.O. Box 1041–00621, Nairobi, Kenya.

E-mail address: j.chamberlin@cgiar.org (J. Chamberlin).

<https://doi.org/10.1016/j.foodpol.2019.101805>

Received 11 March 2019; Received in revised form 24 November 2019; Accepted 27 November 2019

Available online 12 December 2019

0306-9192/© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

sectors (Mellor, 1976). If a few large-scale farmers dominate production and spend their money outside the local rural economy, then local growth multipliers may be weaker than in areas with more egalitarian land distributions (Johnston and Kilby, 1975).

There is also countervailing evidence, much of it more recently from Africa, indicating that large farms may attract public and private investments that provide nearby surrounding smallholders with improved access to markets and services. For example, Sitko et al. (2018) found that large-scale traders have invested in crop buying operations in parts of Africa with a high concentration of medium-scale farms in response to the surplus potential of such farms, thereby improving market access conditions for nearby smallholders too. Other recent studies have found some evidence that smallholder farm households benefit indirectly from being located close to large farms (Deininger and Xia, 2016; Lay et al., 2018).

Identifying the spillover effects of large farms on smallholder farms is complicated because these spillovers may depend on the scale, number, and socio-demographic characteristics of nearby “large” farms. Smallholder households may interact differently with nearby commercialized farms of 5–10 ha than they do with much larger farms for many reasons, not least because they tend to share common social, ethnic or familial connections. Many medium-scale investor farmers go back to their home rural areas to acquire land (Jayne et al., 2016). Large farms in the region are commonly owned and/or operated by individuals from outside the local community. The size and strength of spillover effects between smallholder farms and large farms may therefore depend on the size and characteristics of the large farms, and is an important unresolved empirical question.

To address this question, we examine the case of Tanzania, one of at least several African countries where the numbers of medium- and large-scale farms have grown rapidly in recent years (Schoneveld, 2014, Jayne et al., 2016). We assemble data from the 2009 Tanzanian Agricultural Sample Census (ASC), as well as the 2009, 2011 and 2013 rounds of the Tanzanian National Panel Survey (NPS). We construct several indicators of farmland structure for rural Tanzania, using the ASC, which is statistically representative at the district level. Evidence indicates that certain forms of farm structure, e.g., a high share of farms between 5 and 10 ha, have positive impacts on rural household incomes in the district. Other forms of farm structure, including those with a high concentration of farms larger than 20 ha, have smaller and in some cases even negative impacts on the incomes of smallholder households.

Our primary question – whether or not the local structure of land ownership matters for rural growth – is important for several reasons. First, changes in farm structure are occurring rapidly in many sub-Saharan African countries, with a major trend being one of increasing land concentration driven by growing numbers of medium- and large-scale farms (Jayne et al., 2014, 2016; Sitko and Chamberlin, 2016; Anseeuw et al., 2016). These studies suggest a *de facto* move towards greater land concentration. However, land concentration may be occurring in different ways. The contribution of our study is to emphasize the multi-dimensional nature of farm structure and land concentration, to demonstrate that alternative indicators of land concentration that emphasize different dimensions are often poorly correlated with one another, and most importantly to show how these alternative indicators of farm structure exert varying and in some cases very strong influences on the farm and non-farm incomes of rural households in the vicinity. The study concludes that if land distribution patterns matter for rural transformation, as is strongly indicated by our findings, then researchers and policy makers may need to more accurately understand how farm structure is changing under alternative land tenure systems and more explicitly consider the impact of farm structure on development outcomes and policy objectives.

The remainder of the paper is structured as follows: Section 2 expands on the theoretical relationships between farm structure and economic development outcomes in agrarian areas. Section 3 describes the data used in this study, and Section 4 discusses the challenges of

empirically addressing our research question. Descriptive and econometric results are presented in Sections 5 and 6, respectively, followed by conclusions in Section 7.

2. Conceptual framework

2.1. Definition

“Farm structure” is a multidimensional concept that incorporates both the distribution of farm sizes and the inequality of farm land-holdings (e.g., Stanton, 1991). To illustrate this distinction, consider two districts, one with all farms being 10 ha in size, and the other with all farms being one hectare in size. The farm size distributions of these two districts would be very different, but most measures of land inequality would be the same. In both districts, the Gini coefficient would be zero. We consider that both terms – inequality and the distribution of farm sizes – are subsumed under the broader term “farm structure”, and that both may matter in terms of impacts on rural household incomes.

2.2. Core theoretical perspectives

There are two competing ways in which we might think about the relationship between farm structure and household income growth. The first of these is rooted in the seminal work of Johnston and Kilby (1975) and Mellor (1976), who emphasized the importance of growth multipliers as drivers of rural development. The core idea is that because the propensity to spend additional income on local goods and services is greatest for smallholder households, then virtuous cycles are engendered by broad-based agricultural growth in which income gains by smallholders are re-cycled through local farm and non-farm economies. Broad-based agricultural growth tends to generate greater second-round expenditures in support of local non-tradable goods and services in rural areas and towns. In contrast, if agricultural productivity and household income gains are concentrated within relatively few households (as might be the case in areas where a few large farms have a disproportionate share of land and production), then growth multipliers from agricultural surplus may be more limited, as compared with more egalitarian land distributions. Empirical work by Deininger and Squire (1998) and Vollrath (2007) support this idea, providing evidence that relatively egalitarian national-level land distribution pattern are associated with more broadly based agricultural productivity growth, and higher rates of growth, than more concentrated land distributions.

Other ways of thinking about the influence of land concentration on smallholder welfare posit different channels, but which are consistent with the above framework. For example, Berry and Cline (1979), using national-level data, found that the relative underutilization of agricultural land increases with the degree of inequality in land distribution. Sitko and Jayne (2014) describe similar findings for farm-level data from Zambia: larger farms had lower shares of land being used for cropping or other intensive productive activities than smaller holdings. Due to ethnic and social differences, large-scale and small-scale farmers may have little social interaction, minimizing potential synergies from learning, cooperation, and economic exchange, which could be important avenues by which productivity and income gains and spillovers may be realized.

Also at the country-level, Binswanger and Deininger (1997), Engerman and Sokoloff (1997), and Sokoloff and Engerman (2000) discuss ways in which land inequality may be associated with institutional control. In particular, land concentration (inequality) is often associated with an elite class of rural landholders that wields political power; this power often limits the ability of non-elites to participate in political systems or benefit from public institutions such as crop marketing boards, input promotion programs, and education, which may condition household income. These processes may play out at both national, regional and local levels. For example, if large farms dominate in a particular area, they may influence the nature of local supply

chains, such that input providers, commodity traders and other service providers are more oriented towards supporting larger farms in ways that are less accessible to smaller producers.

A countervailing hypothesis is that larger farms (at least under some conditions) may generate important spillover benefits for smallholders operating in their vicinity. The surplus production of relatively large farms may attract private investment in crop buying, storage, transport, input supply and finance into rural areas, providing spillover benefits to all households in the areas (Collier and Dercon, 2014). The political clout of large farmers may also attract state investment in infrastructure development, which would also benefit all farms in an area (von Braun and Meinzen-Dick, 2009; Deininger and Xia, 2016). Introduction of new production technologies may facilitate technological spillovers via knowledge transfers and increased access to agricultural technologies (Kleemann et al., 2013; Rakotoarisoa, 2011).¹ Direct linkages between large and small farmers may also exist, e.g. out-grower schemes, contract farming arrangements, and the generation of wage employment opportunities. Direct or indirect service provision may also characterize linkages across farm size categories; for example, van der Westhuizen et al. (2018) find that small-scale farmers in Tanzania are substantially more likely to rent tractors in areas with a high concentration of medium-scale farms. Sitko et al (2018) also found that large-scale grain traders tend to locate their buying operations in areas with a high concentration of medium and large farms, and that small-scale farms in the vicinity also benefit from the improved market access conditions that these large grain buyers provide. Knowledge transfers from large-scale to nearby small-scale farmers may also be important in some cases (De Schutter, 2011; Mujenja and Wonani, 2012). To the extent that such positive spillovers exist, then land concentration may promote income growth across all households in a shared location.

As it stands, the evidence base for either positive or negative impacts of large farm spillovers on nearby smallholders remains weak. Lay, Nolte and Sipangule (2018) find some evidence for localized positive productivity spillovers of large agricultural investments to nearby smallholders in Zambia.² Deininger and Xia (2016) find that large-scale investments produce short-run positive and negative effects on nearby smallholder farms in Mozambique. These studies identify spillover effects through a smallholder household's physical proximity to the number of large-scale farms within a certain radius or whether there are any large farms within the locality. While this may be a reasonable approach for addressing specific kinds of questions, we believe that the effects of land concentration and spillover effects from particular kinds of farms may be more comprehensively understood by constructing measures that represent the concentration of large, medium, and small farms within in a given area, as reflected in various measures of localized farm size distributions.³

A key premise of this paper should now be clear, i.e. that different aspects of farm structure cannot be fully captured in a single indicator, such as the Gini coefficient, or the number of farms of a certain size in a given area. Conceptually, the pathways by which large farms may influence the behavior and welfare of nearby smallholder households may only incidentally be related to standard land inequality measures. Because farm structure is a multidimensional concept, empirical

¹ Such knowledge transfer from large to small farmers may take place directly, e.g. via technical assistance, formal and informal training and/or service provision, or indirectly, e.g. via learning-by-doing.

² Empirical evidence is somewhat limited. Some literature uses firm level data (Javorcik, 2004; Görg and Greenaway, 2004), and does not focus on agriculture. At least two studies have provided evidence in support of large-scale land-based investments contributing to infrastructural improvements in the investment locations (Mujenja and Wonani, 2012; FAO, 2012).

³ In addition to the several studies cited here, we also found a few studies examining the impacts of large-scale farm investments on local communities using qualitative case study approaches (e.g. Cotula et al., 2009; Anseeuw et al., 2012).

analysis seeking to understand the effects of alternative land distribution patterns on local growth patterns must consider alternative dimensions of farm structure.

2.3. Model of per-full time equivalent gross income determinants

We may generalize the above ideas as follows. Let us start with a farm-level production function:

$$Y_{i,j,t} = \beta X_{i,j,t} + \gamma C_j + \theta G_{j,t-1} + \epsilon_{i,j,t} \quad (1)$$

where Y is gross income per full-time equivalent (FTE) for farmer i in community j at time t ; X is a vector of household-level characteristics, C is a vector of local geographic context characteristics, G is a measure of access to local public and private stocks of capital, information, services and other resources in community j , and ϵ is an idiosyncratic error term which may be heteroskedastic and clustered at the household level. If we accept that (unobservable) access to local public and private resource endowments is conditioned by the (observable) localized distribution of land control, i.e.

$$G_{j,t} = f(I_{j,t}, Z_{j,t}) \quad (2)$$

where I is a measure of farmland structure in community j at time t , and Z is a vector of other factors which influence G , then we may rewrite an estimable production function as:

$$Y_{i,j,t} = \beta X_{i,j,t} + \gamma C_j + \delta I_{j,t-1} + \gamma Z_{j,t-1} + \epsilon_{i,j,t} \quad (3)$$

where δ is an estimable coefficient on observable farm structure. If $\delta < 0$, then the net effect of more concentrated land distribution patterns or the share of farmland controlled by medium or large farms on smallholder household incomes is negative. If $\delta > 0$, then positive spillovers dominate the relationship.

3. Data and variable construction

3.1. Data sources

Data used in this analysis come from two main sources. Data on household per-FTE gross income measures, along with other household- and community-level controls, were constructed from the Tanzanian National Panel Survey (NPS), available for three waves (2009, 2011, and 2013).⁴ Given the nature of our research question, we are interested in estimating impacts on households in rural areas. However, the census definitions of "urban" SEAs in this sample are not urban in the conventional sense of being primarily composed of town dwellers without agricultural land and farming activities. Therefore, we include households in "urban" SEAs which have population densities of less than 500 persons per square kilometer. (As a robustness check, we compare results with models estimated on samples which only include rural SEAs.) After dropping households which only appear in a single wave, we have an unbalanced panel of 7450 observations across the three waves.

While the NPS is considered to be statistically representative of households in rural Tanzania, it may not have sufficient observations to be statistically representative of the full range of farm sizes found in Tanzania (Christiaensen and Demery, 2018). We therefore constructed indicators of farm structure from the Tanzanian Agricultural Sample Census (ASC) for 2009. The ASC contains data on 52,635 rural agricultural households randomly selected from the prior census, as well as all 1006 farms categorized as large scale⁵ that were identified in the

⁴ The Tanzanian NPS is part of the LSMS-ISA program supported by the World Bank.

⁵ Large scale farms were defined on the basis of landholdings (> 20 ha) as well as a number of other criteria, which allowed some smaller holdings to be included (e.g. operating at least 0.5 ha of intensive greenhouse horticultural

country at the time. The ASC enables district-level inference, but the large-farm component only contains regional identifiers. We thus faced a quandary: either construct measures of farm structure at the more localized district level ($n = 142$) and omit the large farm module, or construct measures of farm structure at the much larger region level ($n = 26$) that include the large farm module. To evaluate this decision, we constructed the Gini coefficient and other measures of farm structure at the region level, first including the large farm module and then excluding it. As will be shown below, alternative measures of land concentration have different sensitivity to whether the large-farm module is included or not. Because the district-level indicators enable us to examine the relationship between farm structure/inequality and household incomes at a more disaggregated geographic level, the analysis in this paper focuses on these results. However, to address potential bias, the models which use district-level concentration measures also include a dummy indicator for regions which are sensitive to exclusion of the ASC large-farm component (described further below). We also estimated models using regional-level land concentration measures for comparison; these results are shown in the appendix.

3.2. Variable construction

3.2.1. Variables measuring farm structure and land concentration

Our exogenous variables of main interest pertain to farm structure and land concentration, which we compute for each of Tanzania's 142 districts with a population density of less than 500 persons per square kilometer. Household landholdings are defined in this analysis as all the land controlled by the household (including land rented in), including land that is cultivated, in fallow, undeveloped, under pasture, planted with trees or other permanent crops, or any other land usage. For every household in the sample, the total landholding size is constructed as the sum of plot-level records. There are many alternative possible measures, including (i) the Gini coefficient; (ii) skewness (third standard moment); (iii) coefficient of variation (standard deviation/mean); (iv) share of operated farmland on farms between 5 and 10 ha; and (v) share of operated farmland on farms over 10 ha. Each of these variables measure different aspects of farm structure, with some emphasizing the importance of specific scales of farm operation, while others focus more on the degree of concentration of landholdings. Because of this, we do not necessarily expect these indicators to be highly correlated.

Fig. 1 shows stylized landscapes which represent alternative farm size configurations of a constant total area. Concentration metrics are calculated for each, and shown in the figure. For the most part, these correspond with an intuitive understanding of concentration in that, as we progress from the upper left, through the upper right, lower left and lower right, we have generally increasing values in most metrics.

3.2.2. Outcome variables

Our key outcome variables of interest are per-FTE (full-time equivalent) income measures. These are constructed using household-level earnings divided by the household sum of individual-level full-time equivalent values calculated from labor-allocation data as recorded by the NPS. Four main categories of income are considered: farm, non-farm, agricultural wages, and total income. Farm income includes income from crop production and livestock income (which includes the value of sales of live animals, value of slaughtered animals, and value of production of milk, eggs, honey, hides and skins). Non-farm income includes income from non-farm business activities and off-farm wage employment (excluding agricultural wage labor). Agricultural wage employment income was also included as a separate category. Finally, total income is the sum of farm, agricultural wage, and non-farm income.

(footnote continued)

production). See National Bureau of Statistics (2012: p11) for more details.

3.2.3. Full-time equivalents

To calculate FTEs, we add up the hours an individual household member reports allocating to on-farm activities, non-farm business activities, and wage employment activities. For any given individual, the total hours per week spent working cannot exceed 112 ($= 16 \text{ h} * 7 \text{ days}$). If the amount reported across all categories exceeds this amount, we scale hours in each category proportionally, such that 112 h per week is not exceeded. All monetary values were converted to real 2010 USD.

4. Implementation challenges

There are several important challenges in estimating our model of interest. These include data quality constraints, an inherent arbitrariness in defining measures of land concentration, and endogeneity issues in ascertaining impacts of land concentration on growth outcomes.

4.1. Data quality issues

The Tanzanian National Panel Surveys, described above, are limited in how information on per-FTE gross income was collected. First of all, for wage income, the amount worked by an individual over the previous 12 months was only calculated for the last two waves. Furthermore, while the second and third waves asked about primary and secondary jobs, the first wave only asked about the primary job. Thus, we were not able to use the first wave of the NPS (2008) in this analysis.

Even though total time worked for wage income over the previous year was nominally recorded (via three questions: "During the last 12 months, for how many months did you work in this job?", "During the last 12 months, how many weeks per month do you usually work in this job?" and "During the last 12 months, how many hours per week do you usually work in this job?"), the informal nature of much wage employment in Tanzania (perhaps particularly in rural areas) implies a high degree of variability, which may not easily filter through such averaging questions. The consequence of this is that our income data are somewhat noisy, as are our measures of per-FTE gross income based thereon. Measurement error in the dependent variable does not bias coefficient estimates but it may inflate their standard errors (Wooldridge, 2010). To address sensitivity of estimation results to such noise, we also estimate regression models for dependent variables which are not normalized by FTEs (i.e. on household-level income measures). These regression results differ little from our per-FTE measures, and we therefore report results from the per-FTE models.

4.2. Endogeneity concerns

In estimating our model of interest, equation (3), there are several endogeneity concerns. The first of these is that localized farm structure and per-FTE gross income may be jointly driven by unobserved factors. For example, if land concentration is associated with commercial land investments that target areas of favorable agro-ecological or market access conditions, we may get upwardly biased coefficient estimates. We take three steps to control for this. First, we include available agro-ecological and market access controls in our regression models. Agro-ecological zones were identified for survey locations from the HarvestChoice database maintained by IFPRI.⁶ Market access variables, calculated by the World Bank, are based on estimated travel time to the nearest road or market town. Second, we include year and year*zone dummy variables to control for unobserved time-constant and time-varying regional effects that are not otherwise captured by our controls.

While the first and second steps above can control for unobserved

⁶ The dataset is available at <http://harvestchoice.org/production/biophysical/agroecology>.

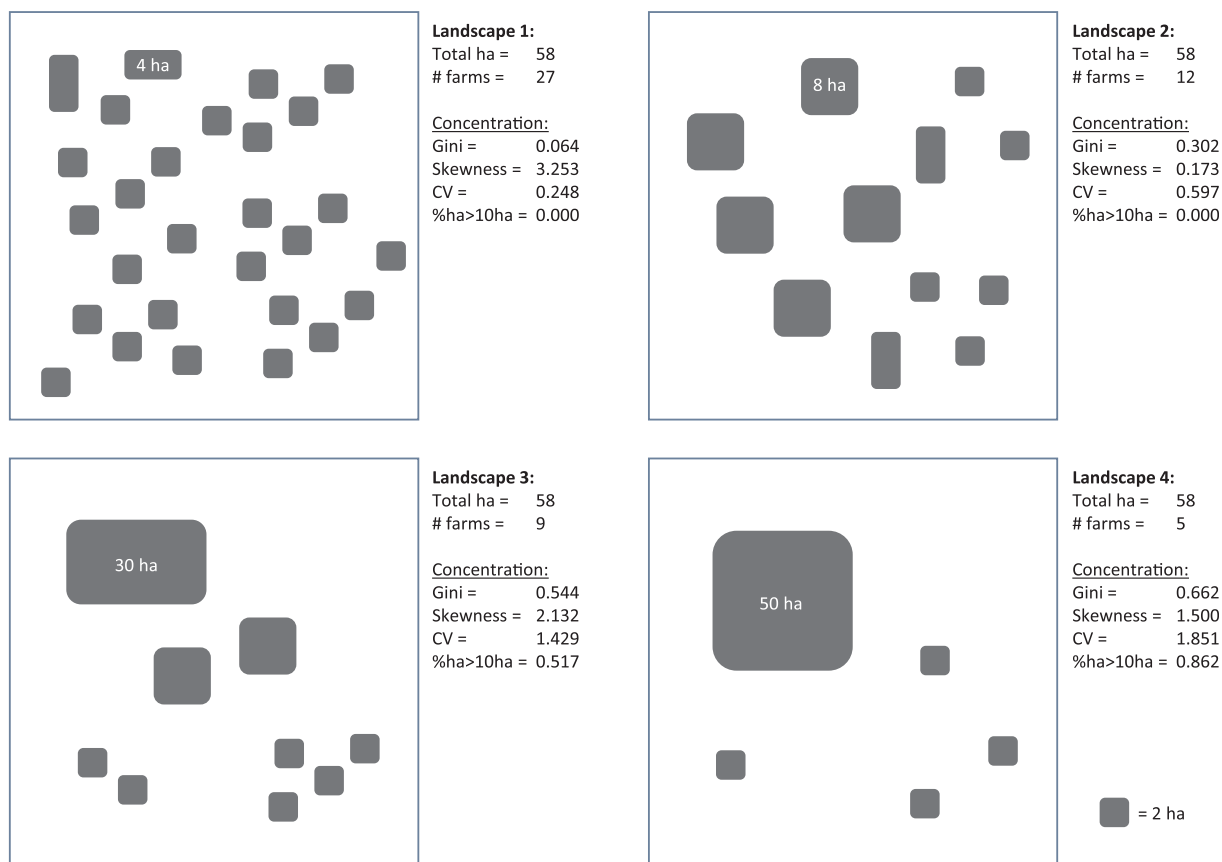


Fig. 1. Stylized landscapes and corresponding land concentration measures.

regional effects, there remains the issue of unobserved heterogeneity at the household level. To address this, our third step is to exploit the panel nature of the data to incorporate the Mundlak-Chamberlain (MC) device (Mundlak, 1978; Chamberlain, 1984) into our models, giving us an estimator that Wooldridge (2010) refers to as the Correlated Random Effects model. The MC device employs household-level averages of all time-varying components of the model in order to control for unobserved time-constant heterogeneity, under the assumption that such heterogeneity is correlated with the time-averages. While we cannot fully eliminate all potential sources of endogeneity, we feel that the three steps described here go as far as possible to do so with the available data.

4.3. Attrition

A third concern is potential attrition bias arising from the use of panel data. We test for this and reject the null hypothesis of no attrition bias in our models. We therefore define and implement attrition weights in all of the regression models, following the methods described in Baulch and Quisumbing (2011).⁷

⁷ In practical terms, we find that whether or not we use attrition weights, our coefficient estimates change very little, and the overall analytical conclusions are the same. Nonetheless, because we reject the null hypothesis of zero attrition bias in many of our model specifications, we report results from the weighted models in this paper. Unweighted model results are also available from the authors upon request.

5. Descriptive results

5.1. Farm size structure

The Tanzanian farm sector is dominated by smallholdings, as elsewhere in the region. However, measures of farm structure are sensitive to choice of dataset. Table 1 shows distributions of farm sizes across the country, using the NPS and ASC for 2009. Three observations are highlighted here. First, the distribution of farm holdings using NPS is sensitive to whether landless households are included in the analysis, at least at very low percentiles of the distribution. Less than one percent of rural Tanzanian households are found to be landless. Second, and more importantly for our analysis, farm sizes start to diverge between NPS and ASC at high ends of the farm size distribution. At the 95th percentile, the ASC shows farm size to be 8.1 ha compared to 6.8 ha for NPS. At the 99th percentile, the ASC shows farm size to be 33.8 or 31.7 percent higher than that of NPS, depending on whether the ASC's large-scale module is included or not. Third, the distribution of farm sizes for ASC up to the 99th percentile is virtually the same regardless of whether the large-scale farm module is included or not. As mentioned earlier, we can define district-level measures of farm structure using ASC only if the large-scale module is excluded, which fortunately has little bearing on most indicators of farm structure. Given the much larger sample size of the ASC and its statistical representativeness at district level, we prefer it to the NPS for constructing indicators of farm structure and land concentration.

5.2. Income trends, by farm size category

Table 2 below shows changes over time in per-FTE income measures for four farm size categories using the balanced panel of NPS households between 2009 and 2013. Overall, real household incomes per FTE

Table 1
Farm structure in Tanzania.

	Hectares per farm holding at the x^{th} percentile of weighted sample distribution								mean
	5th	10th	25th	50th	75th	90th	95th	99th	
controlled land (NPS)	0.1	0.3	0.6	1.3	2.4	4.5	6.7	14.6	2.3
controlled land (NPS), excluding landless HHs	0.3	0.4	0.8	1.4	2.6	4.5	6.8	15.1	2.4
controlled land (ASC), large-scale module included	0.4	0.4	0.8	1.6	2.8	4.9	8.1	20.2	2.7
controlled land (ASC), large-scale module excluded	0.4	0.4	0.8	1.6	2.8	4.9	8.1	19.8	2.5

Note: NPS data for 2008/2009; ASC data for 2009. NPS households in sampling areas designated as “urban” are excluded from the calculation.

Table 2
Measures of real household income per FTE, by farm size category.

	landholding size category	sample size in 2013	2009	2011	2013	avg annual growth
<i>Values in 1000 s of real 2010 TSh</i>						
agricultural income per-FTE	0 < 2 ha	2100	227	228	217	−0.5%
	2–5 ha	629	324	332	343	0.6%
	5–10 ha	157	373	401	404	0.9%
	> 10 ha	64	488	481	572	1.7%
non-farm income per-FTE	0 < 2 ha	2165	751	889	884	1.8%
	2–5 ha	746	844	890	933	1.1%
	5–10 ha	218	710	808	870	2.2%
	> 10 ha	108	892	998	935	0.5%
agricultural wage income per-FTE	0 < 2 ha	2030	34	60	89	10.7%
	2–5 ha	680	22	80	87	15.0%
	5–10 ha	211	46	89	76	5.3%
	> 10 ha	104	29	42	21	−3.4%
Total per-FTE gross income	0 < 2 ha	2181	770	822	831	0.8%
	2–5 ha	743	937	940	956	0.2%
	5–10 ha	215	978	1091	1180	2.0%
	> 10 ha	101	1622	1382	1308	−2.3%

Source: NPS. Landholding size categories are based on the controlled area, which includes all plots which are reported as cultivated, fallow, virgin, forest and pasture. Size category is based on the household’s average landholding size across the three waves. The sample is restricted to rural areas and households with at least one reported plot. Zero-valued income is included.

rose by a small amount between 2009 and 2013: by 0.8% for the < 2 ha landholding category, which constitutes roughly 56% of the rural population; 0.2% for households in the 2–5 ha category; 2.0% for 5–10 ha farms, but negative for the largest category (although with the caveat that the number of farms in this category in the NPS is relatively small). With the exception of agricultural wage income, incomes grew faster for farms over 5 ha than for the majority of farms below 2 ha. Overall, the non-farm and agricultural wage income categories experienced the most growth. This is consistent with other analyses showing a shift in Tanzania’s employment patterns from farm to off-farm and non-farm sources of income in recent years (Yeboah and Jayne, 2018).

5.3. Land concentration measures

To evaluate sensitivity of land concentration indicators to choice of landholding dataset, we constructed land concentration measures at the national level from both the NPS and ASC for 2009 (Table 3). Comparing measures constructed using the small farm component of the ASC with measures from the NPS, we find that measures differ substantially from one another in some respects and very little in other respects. As expected, when including the large scale farm component of the ASC, some measures are substantially higher, e.g. skewness, CV and the share of land under farms of 10 or more hectares, although other measures are very similar with those based on only the small farm portion (Gini and the share of land in farms of 5–10 ha).

To further explore the comparability of alternative farm structure measures, we construct correlation matrices for alternative indicators at the region level (Table 4). Alternative indicators correlate imperfectly with one another. This should not be considered surprising considering that they emphasize different aspects of farm structure. In any case, the

Table 3
National measures of farm structure from alternative data sources.

measure of land concentration	Dataset			
	NPS	NPS (landless excluded)	ASC (excl. large farm module)	ASC (incl. large farm module)
Gini	0.58	0.56	0.53	0.57
Skewness	25.5	25.1	15.8	512.8
Coefficient of variation	3.19	3.12	1.77	17.95
Share of land held by farms 5–10 ha	0.17	0.17	0.16	0.15
Share of land held by farms > 10 ha	0.24	0.24	0.23	0.38

Note: NPS data for 2008/2009; ASC data for 2009. Land measures calculated on controlled land (including rented land). Landless households are included in the calculations, except in column 2. To be consistent with ASC, NPS households in sampling areas designated as “urban” are excluded from the calculation.

results in Tables 3 and 4 point to the need to evaluate the robustness of our results to the choice of alternative farm structure indicators.

Given the possible distortion of district-level land concentration measures from the ASC which are not able to include the large farm component, we evaluate the correlation of measures constructed at the regional level (which does permit inclusion of the large-farm module). Comparing such region-level measures constructed with and without the large farm sample, we find that most regions do not vary much. As an example, comparisons of Gini coefficients are shown in Fig. 2. Nonetheless, to account for potential biases in our regression work relying on district-level concentration measures, we include a dummy

Table 4
Correlation matrix of farm structure measures, district level.

	Gini	Skewness	CV	% land under farms of 5–10 ha		
Gini	1					
Skewness	0.4171 ***	1				
CV	0.7119 ***	0.8162 ***	1			
% land in farms 5–10 ha	0.3567 ***	0.0728	0.1279	1		
% land in farms > 10 ha	0.7331 ***	0.3725 ***	0.5576 ***	0.5407 ***	1	

Source: ASC data for 2009. Landholding based on land controlled (i.e. includes non-cultivated plots).

*** denotes significance at the 1% level.

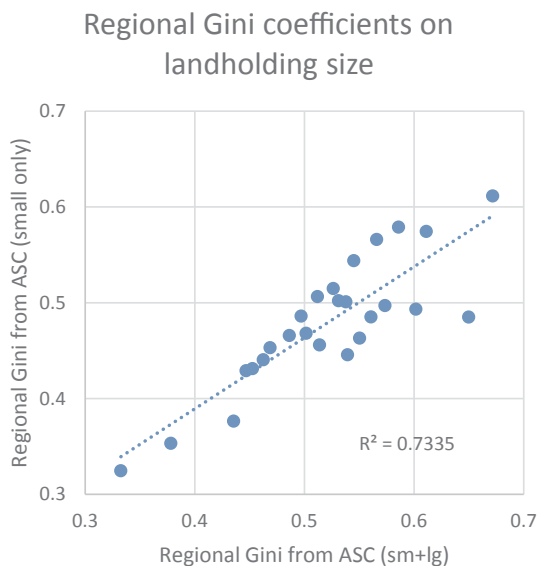


Fig. 2. Scatterplot of Gini coefficients on landholdings from Agricultural Sample Census with and without large farm sample, region level. Source: 2009 Agricultural Sample Census (ASC).

variable to identify regions where the inclusion of the large farm component results in differences in the more sensitive land concentration measures.

6. Econometric analysis

6.1. Baseline specification

Coefficient estimates for land concentration measures from our baseline regression specifications are shown in Table 5 (the full set of results are shown in Appendix Table A1). There are 4 dependent variables: (a) agricultural per-FTE gross income, (b) non-agricultural per-FTE gross income, (c) off-farm agricultural per-FTE wage income, and (d) total per-FTE gross income. All of these dependent variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990). Because most of this function’s domain approximates that of a logarithm, the coefficient estimates can be interpreted as one would in a log-level specification. For each of these dependent variables, the specifications differ only in the choice of farm structure measure: (i) Gini coefficient, (ii) skewness, (iii) coefficient of variation, (iv) share of land in farms of 5–10 ha, (v) share of land in farms of > 10 ha, and (vi) share of land under farms of 5–10 and > 10 ha, respectively. Because initial testing indicated that panel attrition was influenced by variables in our models, these specifications all use inverse probability weights to correct for the probability of household attrition. However, it is noted that weighted and unweighted

regressions differ very little in the resulting coefficient estimates. Finally, we estimate heteroskedasticity-consistent standard errors clustered at the household level to account for potential non-constant variance.

The main analytical conclusion is as follows: while impacts on any particular income type are highly dependent upon which measure of land concentration is used, the overall impacts of more concentrated landholding patterns on farm, non-farm and total per-FTE gross income is positive for most measures. The impact of the share of land in medium-scale farms (between 5 and 10 ha) is particularly pronounced and highly significant. Estimated impacts on agricultural wage income are not significant, however, for any of the specifications. The higher the share of district farmland among farms 5 to 10 ha, the greater the impact on farm, non-farm and total incomes among households within the district. This positive contribution of medium-scale farms does not appear to carry over to farms > 10 ha. In fact, a higher share of district farmland under farms over 10 ha has negative influences after controlling for share of land in 5–10 ha farms, although this result is not significant at conventional levels. This suggests that commercial operations over 10 ha in our sample may not engender the same kinds of positive spillover effects as with medium-scale farms of 5–10 ha.⁸ This finding is consistent with the idea that income multipliers are smaller when very large farms possess a relatively large portion of the land under production in a localized farm-based economy.

As mentioned previously, to address concerns about omission of the large farm component of the ASC in our district-level land concentration measures, we include a dummy indicator for regions where the regional-level landholding coefficient of variation (CV) is sensitive to exclusion of the large-farm component of the ASC. Other land concentration measures are much less sensitive to whether the large-farm component of the ASC is included in their calculation. This dummy takes a value of 1 for all regions with greater than average changes in CV. Interestingly, this dummy is significant and negative in the farm and total income models (panels a and d) and significantly positive in the agricultural wage model (panel c). One interpretation of this result is that while commercial farm operations may generate some kinds of agricultural wage opportunities, they may negatively impact farm and total incomes through reduced multiplier effects (which we would expect of larger farming operations which invest smaller shares of their income in locally traded goods and services). While this result is interesting, we do note that the other coefficient estimates are not particularly sensitive to inclusion of this control. Furthermore, we estimated models using regional-level concentration measures which include the large-farm component (presented in appendix table A4).⁹ In comparing these estimation results, we find little substantive difference, providing some reassurance about robustness of the overall results to alternative data choices. One difference worth noting is that the models with regional-level land concentration measures do find positive impacts of medium-scale farm concentration on agricultural wages, although negative impacts of concentration of 10+ hectare farms and the Gini coefficient. This signals that our assessments of larger farm impacts on agricultural wage income should be taken more cautiously than for other income categories, which generally have more participants and may be observed with greater precision.

We also ran models with alternative definitions of land concentration, e.g. share of land in 5–20 ha farms, 20+ ha farms, etc., in order to evaluate whether any additional structural insights might be gained. For all of the alternative definitions we checked, the same basic results were obtained, i.e. the coefficient on share of land in the “medium-

⁸ For reference, mean and median farm sizes of farms greater than 10 ha in the ASC sample are 15 and 25, respectively (14 and 19, when not including the commercial farm component).

⁹ As mentioned earlier, geographic unit identifiers for the large-scale farm component is available only at the region level.

Table 5
Selected coefficients from baseline regression models.

<i>Panel a: Dependent variable: IHS-transformed household farm per-FTE gross income</i>						
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Land concentration</i>						
Gini	2.615*** (0.798)					
skewness		0.00568 (0.0187)				
CV			0.205 (0.137)			
share land: farms 5–10 ha				1.489** (0.759)		1.717** (0.822)
share land: farms > 10 ha					0.0741 (0.378)	–0.244 (0.411)
1 = region CV estimate sensitive to ASC lg. farms	–0.219** (0.109)	–0.308*** (0.102)	–0.286*** (0.103)	–0.288*** (0.103)	–0.353*** (0.108)	–0.363*** (0.108)
<i>Panel b: Dependent variable: IHS-transformed household non-farm per-FTE gross income</i>						
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Land concentration</i>						
Gini	1.672 (1.611)					
skewness		0.0405 (0.0437)				
CV			0.279 (0.287)			
share land: farms 5–10 ha				5.535*** (1.672)		6.462*** (1.827)
share land: farms > 10 ha					0.200 (0.792)	–1.006 (0.859)
1 = region CV estimate sensitive to ASC lg. farms	0.173 (0.257)	0.0896 (0.240)	0.139 (0.248)	0.158 (0.240)	0.115 (0.261)	0.0771 (0.261)
<i>Panel c: Dependent variable: IHS-transformed household agricultural wage per-FTE gross income</i>						
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Land concentration</i>						
Gini	–2.422 (1.487)					
skewness		–0.0273 (0.0372)				
CV			–0.221 (0.255)			
share land: farms 5–10 ha				2.260 (1.478)		2.653 (1.637)
share land: farms > 10 ha					0.0440 (0.746)	–0.460 (0.824)
1 = region CV estimate sensitive to ASC lg. farms	0.327 (0.220)	0.424** (0.210)	0.394* (0.215)	0.473** (0.210)	0.440** (0.224)	0.428* (0.224)
<i>Panel d: Dependent variable: IHS-transformed total household per-FTE gross income</i>						
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Land concentration</i>						
Gini	1.525*** (0.570)					
skewness		0.00467 (0.0160)				
CV			0.153 (0.103)			
share land: farms 5–10 ha				1.834*** (0.593)		2.106*** (0.632)
share land: farms > 10 ha					0.0831 (0.288)	–0.321 (0.307)
1 = region CV estimate sensitive to ASC lg. farms	–0.167** (0.0842)	–0.235*** (0.0813)	–0.205** (0.0812)	–0.208** (0.0818)	–0.228*** (0.0840)	–0.240*** (0.0841)

Notes: Dependent variables are inverse hyperbolic sine transformed per-FTE gross income measured in 2010 constant Tanzanian shillings. District-level land concentration measures from 2009 Ag. Sample Census. Dependent variables and other independent control variables are from the NPS. All models include the Mundlak-Chamberlain device. Full model results shown in Appendix Table A1. Robust pval in parentheses, with significance indicated by asterisks:

* p < 0.1.

** p < 0.05.

*** p < 0.01.

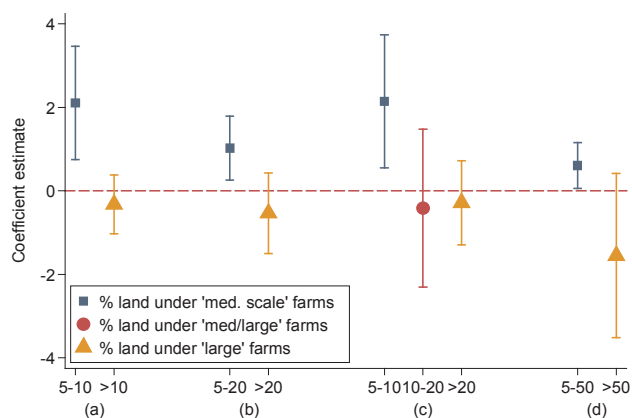


Fig. 3. Sensitivity of results to alternative farmland concentration measures using different farm size categories. Note: Graph shows coefficient estimates of the percentage of district farmland under “medium” and “large” farms, for alternative farm size definitions of these terms. Vertical bars indicate the 95% confidence interval. The dependent variable in all specifications is the inverse hyperbolic sine transformation of total per-FTE household income. The full regression results are shown in Appendix Table A4 (with groupings a-d in the figure above corresponding to the specifications in the table). The dependent variable in these models is the inverse hyperbolic sine transformed per-FTE gross total income measured in 2010 Tanzanian shillings.

scale” category was significant and positive, but decreasing in magnitude with the width of the category (i.e. the positive impact of the coefficient on the share of land under 5–10 ha farms is smaller in magnitude than the coefficient on the share of land under 10–20 ha farms); the coefficient on share of land in the largest category (whether > 10, > 20 or > 50 ha) was invariably insignificant. A graphical presentation of coefficient estimates from these alternative models is presented in Fig. 3 (with full estimation results reported in appendix Table A4). These results suggest that we should be wary of being overly dogmatic in our definition of “medium scale” farms using any particular farm size definition: we would obtain broadly similar results using a 5–10 ha categorization, although some precision is lost as we expand the range of farm sizes that we include in this category.

Given the predominance of gendered differences in land access in sub-Saharan Africa (Doss et al., 2013), we considered the possibility that land concentration may be correlated with poorer access to land by women (in which case gendered land access would constitute an omitted variable that is correlated with and biasing the coefficient on land concentration). To address this concern, we first note that all specifications include a female-head dummy, which is only significant in the total income models. Furthermore, when we specify a model which interacts the female-head dummy with the land concentration measures, the coefficient estimate for the interaction term is not significantly different from that of the non-interacted term. This provides some reassurance that a gendered access story is not driving our results.

To test the robustness of these findings, we estimated a large number of alternative specifications. To begin with, to allay concerns about potentially noisy per-FTE income measures, we also estimate models using farm-level income measures (i.e. not normalized by household FTEs for any given income category), and with per-capita income measures, as dependent variables. Coefficient estimates from these models are extremely similar to the per-FTE measures. We also estimated models using a sample restricted to households in enumeration areas classified as “rural” by the Tanzanian National Bureau of Statistics, to address the possibility that our results are influenced by households which are not truly rural; results are very similar in this case as well.

Table 6
Selected coefficients from regression models of household farm income per FTE including interaction terms between land concentration and asset terciles.

Dep. var: farm per-FTE gross income	(1)	(2)	(3)	(4)	(5)
<i>Land concentration</i>					
Gini	0.000672 (1.349)				
Gini * medium	1.745* (0.971)				
Gini * wealthiest	2.599** (1.050)				
Skewness		-0.170** (0.0824)			
Skewness * medium		0.155* (0.0800)			
Skewness * wealthiest		0.201** (0.0831)			
CV			-0.412 (0.357)		
CV * medium			0.453 (0.330)		
CV * wealthiest			0.669* (0.352)		
share land: farms 5–10 ha				-4.521 (2.905)	
share * medium				4.915* (2.844)	
share * wealthiest				6.738** (2.932)	
share land: farms > 10 ha					-1.307* (0.779)
share * medium					1.706** (0.841)
share * wealthiest					1.238 (0.806)

Notes: Dependent variables are inverse hyperbolic sine transformed per-FTE gross income measured in 2010 Tanzanian shillings. District-level land concentration measures from 2009 Ag. Sample Census. Dependent variables and other independent control variables are from the NPS. All models include the Mundlak-Chamberlain device. Full model results shown in Appendix Table A2 panel a. Robust p-values in parentheses, with significance indicated by asterisks: ***p < 0.01.

* p < 0.1.
** p < 0.05.

6.2. Interacting concentration with household asset dummies

Table 6 shows results from model specifications which interact district land concentration measures with household-specific productive asset categories, defined as terciles of the sample distribution. For conciseness, we only show model results where the dependent variable is agricultural per-FTE income. Full estimation results are shown in Appendix Table A2. While impact estimates vary across land concentration measures and their interactions, a general pattern may be observed in which the positive impacts of land concentration are larger for the top two asset terciles. Relatively poor households benefit the least from localized land concentration (and in some cases are worse off from it), whereas the majority of households and especially the wealthiest one-third of households tend to benefit significantly.

These differential wealth-related impacts on farm income indicate that spillover effects from land concentration are not equally accessible to all farms within an area. Even though Sitko, Burke and Jayne (2018) found that medium-scale farms tend to attract private traders and improve market access conditions for nearby smallholders, relatively poor households may not be able to benefit from this if they cannot produce a farm surplus in the first place. While van der Westhuizen et al., (2018) found that medium-scale farmers tend to rent out their tractors to small-scale farms in their areas, relatively poor households may not be able to afford to take advantage of such services. While we feel these

interpretations are quite plausible in light of other evidence, our data do not allow us to do more than surmise what specific channels of spillover may be taking place. These results are similar to results of models (not shown here) that use farm size categories in place of household asset categories, which is not surprising given the high correlation between asset wealth and land holding size.

6.3. Exploring spillover pathways

Given the indicators of some positive spillovers in the previous results, it is of interest to know whether any particular spillover pathways are identifiable in our data. As a simple way to explore this, we test the individual and joint impact of augmented versions of the baseline model which include, as additional controls, measures of (a) the district share of households with non-farm wages; (b) the district share of households with agricultural wages; (c) the district share of households with rented or borrowed farming equipment (e.g. tractor, oxen, harrow); (d) the district share of households with purchased fertilizer; (e) the district share of households which received extension advice; and (f) the district share of households which received market information from large farmers. To the extent that such indicators are significant, and the extent to which their inclusion in the model diminishes the size and significance of the medium-scale farm concentration variable, we may gain insights into the corresponding pathways of spillover. Results are summarized in Table 7 below (with full estimation results presented in the Appendix table A3).

For farm income (panel a), we see that district agricultural wage market participation is negatively associated with farm income (possibly reflecting impacts on own-farm labor supply), and that the share of households purchasing fertilizer is positively associated with farm income. However, the coefficient estimate on the share of land in 5–10 ha farms is reduced by only a small amount upon inclusion of these controls, and is fairly consistent across specifications, suggesting that there are other channels of positive spillovers (or that the indicators included only capture some of the salient variation).

Unsurprisingly, the district share of households with non-farm income (which may represent non-farm economic vibrancy) is a positive and significant determinant of non-farm per capita income levels (panel b), while agricultural wage market participation is a positive and significant determinant of agricultural wage per-FTE income (panel c). Otherwise, however, the coefficient estimate for 5–10 ha farms is robust to these additional controls. For total income (panel d), the district share of households with non-farm and farm income are significant regressors, but again, the estimated spillover impact of 5–10 ha farms is largely unaffected. In sum, the processes by which medium-scale farms may be driving income improvements in surrounding areas appear to be diffuse, and extend beyond the simple indicators of input, service and labor markets that we control for here.

6.4. Simulated impacts of changes in farmland distributions

To understand the magnitude of these impacts, we simulate the impacts of a change in the farm structure variables on total per-FTE income, using the baseline results in Table 5. We consider a change in land concentration, moving from the 25th percentile to the 75th percentile. For example the 25th and 75th percentiles of the farmland Gini coefficient in our sample are 0.41 and 0.53, respectively.¹⁰ Results, shown in Table 8 below, indicate that such an increase in land concentration, as measured by the Gini coefficient, is associated with an

¹⁰ Here, we measure percentiles in terms of the sample-weighted households in our dataset, i.e. we compare the land concentration experienced by the sample household at the 25th percentile, when households are ranked by that land concentration measure, to that experienced by the household at the 75th percentile.

average gain of 140,000 TSh in total per-FTE gross income (about 100 USD), *ceteris paribus*, which is equivalent increasing the mean level of total per-FTE gross income in our dataset by 15%. The estimated impact of increasing the share of land under medium-scale farms from the 25th to the 75th percentile is even larger: a 217,000 TSh gain, 23% of the sample mean. We reiterate that the model contains numerous granular controls for agro-ecological, rainfall, population density, and market access conditions, and their interactions with year dummies. While we cannot rule out that the magnitude of these simulation results are affected to some degree by unobserved time-varying heterogeneity correlated with district-level farm structure, the inclusion of these granular controls with year interaction terms, with CRE estimation, provides fairly strong evidence that differences in localized farm structure independently exert major effects on economic activity in an area, judging from the very strong estimated effects on rural households' farm and non-farm incomes. We find these findings to be very plausible in light of growing evidence that medium-sized farms tend to be more commercialized than small-scale farms in both input market participation and output market sales, and are attracting private investments in a variety of ways that are improving market access conditions for all rural residents in an area. Anecdotal evidence also suggests that medium-scale farms are a major source of cash expenditures in the local non-farm economy, potentially increasing the demand for a wide variety of goods and services locally (e.g., Poulton, 2018).

7. Policy implications

Our findings suggest that localized heterogeneity in farm structure may be conducive to rural development in smallholder-dominated agricultural production systems. This appears to be particularly true when such heterogeneity includes medium-scale farms in the 5–10 ha range. Such farms may have stronger multiplier effects than larger commercial farms, and may also have greater direct linkages (e.g. through wage employment, equipment rental, service provision) than either larger or small-scale farms.

Our findings for Tanzania may be viewed as being somewhat in contrast to what has become a widely accepted stylized fact from the Asian literature on agricultural development and structural transformation, i.e., that relatively unconcentrated unimodal patterns of farmland distribution tend to generate the greatest localized multiplier effects, non-farm employment growth, and economic transformation resulting from agricultural growth. We do not wish to challenge this literature or assert that it does not apply to Africa, especially not based on one case study. We do raise the following points that may reduce the apparent inconsistency between our results and the Asian literature stressing the developmental superiority of a unimodal, smallholder-led development approach. First, the Asian literature is largely derived from areas with irrigation or water control that allow for two or three production seasons from the same plots, which clearly generates much greater farm commercialization and associated multiplier effects from relatively small farms compared to farms totally reliant on rainfed moisture with one growing season per year, which is the case in much of Tanzania and Africa more generally. If we accept that differences in crop productivity and surplus output per land unit – driven by fundamental differences in soil, rainfall and water control conditions – may influence the scale of farm production capable of driving economic transformation, then it may be plausible that the farm scales and farm size distributions required to achieve these outcomes in sub-Saharan Africa may be different than in much of Asia. Second, we find that it is farms in the 5–10 ha range that generate the strongest indirect effects on the incomes of households around them, and these farms are not very much bigger than the “commercialized smallholder farms” that Mellor (2014) and others refer to as the dynamic change agents in Ethiopian agriculture. As district-level shares of medium-scale farms in the 5–10 ha range are positively correlated with measures of land inequality, such as the Gini and CV of landholdings, this may explain why

Table 7
Baseline regression augmented with diagnostic controls.

<i>Panel (a): farm per-FTE gross income</i>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Land share: 5–10 ha farms	1.920** (0.802)	1.848** (0.803)	2.183*** (0.785)	1.869** (0.799)	1.487* (0.834)	1.878** (0.816)	1.976** (0.810)	1.686** (0.821)
Land share: 10 + ha farms	-0.781* (0.422)	-0.786* (0.421)	-0.813* (0.421)	-0.908** (0.433)	-0.669 (0.432)	-0.797* (0.421)	-0.759* (0.421)	-0.781* (0.439)
District share HHs w/ non-farm wages		-0.478 (0.396)						-0.710* (0.407)
District share HHs w/ agric. wages			-0.891** (0.358)					-1.029*** (0.377)
District share HHs w/ rented equipment				0.562 (0.349)				0.526 (0.350)
District share HHs w/ purchased fert.					0.646** (0.280)			0.717*** (0.277)
District share HHs rcv'd extn. advice						0.249 (0.460)		-0.132 (0.467)
District share HHs rcv'd market info.							-1.661 (2.073)	-2.217 (2.104)
<i>Panel (b): non-farm per-FTE gross income</i>								
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Land share: 5–10 ha farms	6.456*** (1.828)	7.638*** (1.819)	6.040*** (1.838)	6.441*** (1.825)	6.290*** (1.900)	6.287*** (1.846)	6.566*** (1.835)	6.874*** (1.895)
Land share: 10+ ha farms	-0.765 (0.926)	-0.684 (0.936)	-0.715 (0.923)	-0.802 (0.943)	-0.723 (0.943)	-0.830 (0.924)	-0.721 (0.927)	-0.904 (0.959)
District share HHs w/ non-farm wages		7.880*** (0.848)						8.707*** (0.854)
District share HHs w/ agric. wages			1.407 (0.905)					3.161*** (0.915)
District share HHs w/ rented equipment				0.163 (0.795)				0.828 (0.799)
District share HHs w/ purchased fert.					0.247 (0.698)			-0.469 (0.716)
District share HHs rcv'd extn. advice						1.006 (1.016)		1.617 (1.050)
District share HHs rcv'd market info.							-3.288 (3.653)	-2.320 (3.607)
<i>Panel (c): agricultural wage per-FTE income</i>								
	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
Land share: 5–10 ha farms	2.644 (1.630)	2.564 (1.630)	-1.632 (1.554)	2.679* (1.626)	2.282 (1.688)	2.510 (1.639)	2.729* (1.636)	-1.871 (1.612)
Land share: 10+ ha farms	-0.457 (0.824)	-0.463 (0.824)	0.0609 (0.791)	-0.369 (0.830)	-0.364 (0.837)	-0.509 (0.826)	-0.423 (0.823)	0.385 (0.812)
District share HHs w/ non-farm wages		-0.535 (0.764)						2.340*** (0.748)
District share HHs w/ agric. wages			14.46*** (0.784)					15.08*** (0.781)
District share HHs w/ rented equipment				-0.390 (0.677)				-0.182 (0.654)
District share HHs w/ purchased fert.					0.539 (0.591)			0.927 (0.624)
District share HHs rcv'd extn. advice						0.801 (0.851)		-1.084 (0.894)
District share HHs rcv'd market info.							-2.529 (3.346)	-0.471 (3.236)
<i>Panel (d): total per-FTE gross income</i>								
	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)
Land share: 5–10 ha farms	2.133*** (0.613)	2.364*** (0.620)	1.991*** (0.608)	2.106*** (0.613)	1.981*** (0.651)	2.167*** (0.629)	2.236*** (0.621)	2.064*** (0.653)
Land share: 10+ ha farms	-0.318 (0.301)	-0.302 (0.303)	-0.301 (0.299)	-0.386 (0.309)	-0.279 (0.311)	-0.305 (0.298)	-0.277 (0.297)	-0.281 (0.311)
District share HHs w/ non-farm wages		1.543*** (0.317)						1.714*** (0.335)
District share HHs w/ agric. wages			0.479* (0.268)					0.863*** (0.290)
District share HHs w/ rented equipment				0.301 (0.250)				0.490* (0.255)

(continued on next page)

Table 7 (continued)

Panel (d): total per-FTE gross income								
	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)
District share HHS w/ purchased fert.					0.227 (0.226)			0.249 (0.212)
District share HHS rcv'd extn. advice						-0.202 (0.424)		-0.279 (0.427)
District share HHS rcv'd market info.							-3.083* (1.826)	-2.797 (1.859)

Notes: Full estimation results provided in Appendix table A3.

Table 8

Simulated impacts of changes in land concentration on total income.

	(a)	(b)	(c)	(d)	(e)
Land concentration measure	Average per-FTE income predicted for land concentration at 25th percentile (1000s of 2010 TSh)	Average per-FTE income predicted for land concentration at 75th percentile (1000s of 2010 TSh)	Difference(b) - (a) (1000s of 2010 TSh)	difference as % of mean per-FTE income	difference as % of median per-FTE income
<i>Total per-FTE income:</i>					
Gini	801.7	940.1	138.4	15%	8%
CV	828.9	890.5	61.7	7%	4%
Share of land: 5–10 ha farms	760.8	978.5	217.7	23%	12%

Note: Values in columns (a), (b) and (c) are thousands of real 2010 Tanzanian shillings (TSh). In 2010, 1 USD \approx 1450 TSh. These simulation results based on the baseline regression specifications shown in Table 5 and Appendix Table A1.

we find a positive effect of land concentration on rural household incomes in the area. However, this does not mean that extreme forms of land inequality are favorable for economic development. Consistent with the Asian unimodal vs. bimodal literature, we find some evidence that a higher concentration of district-level farmland under farms holding over 20 ha adversely affects rural household incomes in a given district. Therefore, we do not view the results of this study from Tanzania as challenging the longstanding evidence from Asia in favor of an inclusive smallholder-led approach to development, but rather as provoking the need for a more nuanced discussion in Africa about the farm scales associated with agricultural commercialization, factor market participation, and economic transformation in light of fundamental differences in farm output per land unit per year. Additional evidence from sub-Saharan Africa on this issue will certainly contribute to a deeper understanding of the role of alternative farm size distributions in promoting transformation dynamics in Africa, and whether the size of farms and distribution of farmland that tended to generate the greatest indirect benefits to surrounding rural households in Asia may be somewhat different in Africa, where over 95% of agricultural land remains under rainfed production.

While the findings of this study indicate that localized differences in farmland inequality and size distributions are associated with major differences in rural household incomes, this does not imply that land redistribution programs are warranted to alter farm structure in a particular direction. As stressed above, much more evidence from other sub-Saharan African countries is needed to better understand whether robust cross-country evidence emerges in support of spillover effects resulting from particular farm scales or farmland structure variables. Moreover, it will be important to take account of evidence in related literatures on land tenure security, investment and land productivity, and broader livelihood effects resulting from struggles over the institutions governing access to land (e.g., Boone, 2014; Lawry et al., 2014; Melesse and Bulte, 2015), which identify important rural development effects resulting from land tenure programs that may alter farm structure in ways not considered in this analysis.

8. Conclusions

This paper is motivated by the need to better understand the impacts of changing farm structure in Africa. Recent research has documented the rapid pace of new land acquisitions by foreign large-scale interests (Deininger and Byerlee, 2011) and by medium-scale African farmers (Jayne et al., 2016). Tanzania in particular has experienced a rapid increase in land controlled by medium- and large-scale farms in recent years (Schoneveld, 2014; Jayne et al., 2016). These changes in farm structure and composition have generated much speculation and debate about the impacts on smallholder households and rural communities, unfortunately with very little hard evidence to guide land policies and programs. Recent analysis has addressed the potential for spillover effects of large farms on technology adoption and yields of nearby smallholder farms, but the broader impacts on household incomes, disaggregated by farm and non-farm sources, and how these impacts may differ according to gender and wealth characteristics of the household have yet to be explored. Moreover, the few studies on this topic tend to identify the effects of large farms based only on proximity to a given smallholder household, without considering potential differences in impacts between large-scale and medium-scale farms or the degree of concentration of such farms within a given locality. This study addresses these issues by characterizing the structure of farm operations at the district level in Tanzania, both in terms of the relative importance of small, medium, and large farms, as well as by various indicators of farmland concentration in the locality. We estimate whether different kinds of farm structure are beneficial or detrimental to rural development as measured by rural household incomes, disaggregating the effects on income from own farming, agricultural wages, and non-farm sources. Farm structure indicators are derived from a unique Agricultural Sample Census carried out in Tanzania in 2009, which is statistically representative for the country's 142 districts.¹¹

¹¹ The number of districts has since increased. As of the 2012 census, there were 169 districts in Tanzania.

Our first observation is that alternative indicators of farm structure chosen for this study differ considerably from one another. Some indicators emphasize the relative importance of different scales of farm operations in the locality, while others focus on the degree to which farmland operations are concentrated or unequally distributed. These observations reinforce a point that should already be well accepted – that farm or “agrarian structure” is a multifaceted concept and that specific indicators of agrarian structure may not be highly correlated with one another. For this reason, the impacts of different farm structures are likely to be highly sensitive to the choice of indicator.

Guided by these findings, we estimate several models of household income per full-time labor equivalent using panel estimation techniques, based on five alternative indicators of farm structure, including (i) the Gini coefficient; (ii) skewness; (iii) coefficient of variation; (iv) share of controlled farmland under 5–10 ha farms; and (v) share of controlled farmland under large farms (defined here as farms of 10 or more hectares).

The study highlights four main findings. First, farmland concentration is generally positively associated with rural household incomes, after controlling for other geographical and household-level factors. Second, household incomes from farm and non-farm sources (excluding agricultural wages) are particularly positively and significantly associated with the share of land in the district under farms of 5–10 ha. Third, these positive spillover benefits are smaller and less statistically significant in districts with a relatively high share of farmland under farms over 10 ha in size. While our econometric results do not identify the reason why medium-scale farms appear to generate greater spillover effects to local communities than relatively larger farms, recent published studies give us some clues – most medium-scale farmers come from the same social and ethnic backgrounds as small-scale farmers, and tend to have more extensive social interactions with the local community than do most large-scale farms (Sitko and Jayne, 2014). Fourth, poor rural households are least able to capture the positive spillovers generated by medium-scale farms and by concentrated farmland patterns. The greatest benefits to household income were enjoyed by households in the upper two-thirds of the wealth distribution, which still includes the majority of rural households. We speculate that poor rural households are less able to afford taking advantage of the improved access to markets and services that medium-scale farms tend to provide. This aligns with our finding that agricultural wages do not seem to be a primary channel of influence. However, more detailed research is needed on the pathways by which medium- and large-scale farms affect household and broader local community welfare.¹²

The analysis presented in this paper contributes to our understanding of how various dimensions of farm structure may uniquely influence rural development trajectories, both conceptually and empirically in a particular African context. Our analysis has demonstrated the value of structurally interpretable measures of farmland distribution – e.g. local share of land under 5–10 ha farms – as compared with more generic measures of distribution, such as the Gini coefficient.

More generally, our study underscores the importance of good data on land distributions in developing countries (see also Lowder et al., 2016). The recent bounty of nationally representative data available through such initiatives as the World Bank’s LSMS-ISA program is certainly to be applauded. Nonetheless, we would advocate for even

¹² While it is beyond the scope of this study to address spillover mechanisms, we do note that medium scale farms are more productive, more engaged with input and output markets, more likely to own traction equipment, and more likely to hire in labor (and to do so at greater intensities). These patterns indicate both a greater capacity to engage with local economies through both upstream and downstream linkages, which may confer advantages to neighboring smallholders, as well as greater likelihood to directly engage with smallholder neighbors by hiring in labor and hiring out services. Appendix Figure A1 summarizes several descriptive indicators by farm size category, which illustrate these patterns.

greater investments in expanding the sampling frame – both to ensure adequate representation of larger farms and to enable more spatially disaggregated measures of farm structure and concentration. Furthermore, more detailed data collection on localized tenure arrangements, settlement histories and the modalities of land acquisition should help clarify policy linkages with land distribution outcomes. Although this will require increased investments in data collection, the analytical payoffs stand to be substantial, particularly in countries where medium- and large-scale land acquisitions are known to be taking place. Our analysis suggests that small-to-medium-scale local investors, rather than large-scale foreign investors, are key players in the farmland dynamics being experienced in Tanzania, in ways which are consistent with expected patterns of structural transformation. But much empirical work remains to be done to flesh out our understanding of such dynamics – and their impacts – in Tanzania and elsewhere in the region. Ultimately, the scope for research to inform and guide African states will depend upon how well researchers and policymakers are able to accurately monitor and evaluate these changes taking place on the ground.

Acknowledgements

This study was supported by (i) the CGIAR Research Program on Policies, Institutions, and Markets (PIM), which is led by the International Food Policy Research Institute (IFPRI); (ii) the CGIAR Research Program on Maize (MAIZE), which is led by the International Maize and Wheat Improvement Center (CIMMYT); and (iii) the USAID Bureau for Food Security, through the Food Security Policy Innovation Lab led by Michigan State University. The authors acknowledge valuable comments received from Karen Brooks, Andrew Dillon, Milu Muyanga, Robert Myers, Frank Place, Ayala Wineman and Felix Kwame Yeboah. Any errors are those of the authors.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodpol.2019.101805>.

References

- Anseu, Ward, Boche, Mathieu, Breu, Thomas, Giger, Markus, Lay, Jann, Messerli, Peter, Nolte, Kerstin, 2012. Transnational Land Deals for Agriculture in the Global South: Analytical Report based on the Land Matrix Database. CDE/CIRAD/GIGA, Bern/Montpellier/Hamburg.
- Anseu, Ward, Jayne, T.S., Kachule, R., Kotsopoulos, J., 2016. The quiet rise of medium-scale farms in Malawi. Land. <https://doi.org/10.3390/land5030019>.
- Baulch, Bob, Quisumbing, Agnes, 2011. Testing and Adjusting for Attrition in Household Panel Data. CPRC Toolkit Note. Chronic Poverty Research Centre. Mimeo. Revised: 22 September 2011.
- Berry, A.R., Cline, W.R., 1979. *Agrarian Structure and Productivity and Developing Countries*. The Johns Hopkins University Press, Baltimore.
- Binswanger, H.P., Deininger, K., 1997. Explaining agricultural and agrarian policies in developing countries. *J. Econ. Lit.* 35, 1958–2005.
- Boone, Catherine, 2014. *Property and Political Order in Africa*. Land Rights and the Structure of Politics. Cambridge University Press, New York.
- Chamberlain, G., 1984. Panel Data. In: In: Griliches, Z., Intriligator, M.D. (Eds.), *Handbook of Econometrics*, vol. 2. North-Holland Press, Amsterdam, pp. 1248–1318.
- Christiansen, L., Demery, L., (Eds.), 2018. *Agriculture in Africa: Telling Myths from Facts*. Directions in Development. World Bank, Washington, DC.
- Collier, Paul, Dercon, Stephan, 2014. African Agriculture in 50 Years: Smallholders in a Rapidly Changing World? *World Development*, vol. 63, pp. 92–101.
- Cotula, L., Vermeulen, S., Leonard, R., Keeley, J., 2009. Land grab or development opportunity? Agricultural investment and international land deals in Africa. Food and Agriculture Organization of the UN (FAO); /International Fund for Agricultural Development (IFAD); and /International Institute for Environment and Development (IIED), Rome; London. Retrieved from <http://www.iied.org/pubs/display.php?o=12561IIED>.
- De Schutter, O., 2011. How not to think of land-grabbing: three critiques of large-scale investments in farmland. *J. Peasant Stud.* 38 (2), 249–279.
- Deininger, K., Squire, L., 1998. New ways of looking at old issues: inequality and growth. *J. Dev. Econ.* 57, 259–287.
- Deininger, Klaus, Byerlee, Derek, Lindsay, Jonathan, Norton, Andrew, Selod, Harris, Stickler, Mercedes, 2011. *Rising Global Interest in Farmland: Can it Yield Sustainable*

- and Equitable Benefits? World Bank Publications, Washington, DC.
- Deininger, K., Xia, F., 2016. Quantifying spillover effects from large land-based investment: the case of Mozambique. *World Develop.* 87, 227–241.
- Doss, Cheryl, Kovarik, Chiara, Peterman, Amber, Quisumbing, Agnes R., van den Bold, Mara, 2013. “Gender Inequalities in Ownership and Control of Land in Africa: Myths versus Reality” IFPRI Discussion Paper 01308. December 2013. Poverty, Health, and Nutrition Division, International Food Policy Research Institute, Washington DC.
- Eastwood, R., Lipton, M., Newell, A., 2010. Farm size. In: Pingali, P.L., Evenson, R.E. (Eds.), *Handbook of Agricultural Economics*, North Holland, Elsevier, pp. 3323–3397 (Chapter 65).
- Engerman, S.L., Sokoloff, K.L., 1997. Factor endowments, institutions and differential paths of growth among new world economies. In: Haber, S.H. (Ed.), *How Latin America Fell Behind*. Stanford University Press, Stanford CA, pp. 260–304.
- FAO, 2012. *Trends and Impacts of Foreign Investment in Developing Country Agriculture*. FAO, Rome.
- Görg, H., Greenaway, D., 2004. Much ado about nothing? Do domestic firms really benefit from foreign direct investment? *World Bank Res. Observer* 19 (2), 171–197.
- Javorcik, B.S., 2004. Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *Am. Econ. Rev.* 94 (3), 605627.
- Jayne, T.S., Chapoto, A., Sitko, N., Nkonde, C., Muyanga, M., Chamberlin, J., 2014a. Is the scramble for land in Africa foreclosing a smallholder agricultural expansion strategy? *J. Int. Affairs* 67 (2), 35–53.
- Jayne, T.S., Chamberlin, J., Headey, D., 2014b. Land pressures, the evolution of farming systems, and development strategies in Africa: a synthesis. *Food Policy* 48, 1–17.
- Jayne, T.S., Chamberlin, J., Traub, L., Sitko, N., Muyanga, M., Yeboah, F., Anseeuw, W., Chapoto, A., Wineman, A., Nkonde, C., Kachule, R., 2016. Africa’s changing farm size distribution patterns: the rise of medium-scale farms. *Agric. Econ.* 47s, 197–214.
- Johnston, Bruce F., Mellor, John, 1961. The role of agriculture in economic development. *Am. Econ. Rev.* 51 (4), 566–593.
- Johnston, B.F., Kilby, P., 1975. *Agriculture and Structural Transformation: Economic Strategies in Late Developing Countries*. Oxford University Press, New York.
- Kleemann, Linda, Jann Lay, Kerstin Nolte, Konrad Ott, Rainer Thiele, and Lieske Voegt-Kleschin. 2013. Economic and Ethical Challenges of “Land Grabs” In Sub-Saharan Africa. Kiel Policy Brief, No. 67.
- Lawry, S., Samli, C., Hall, R., Leopold, A., Hornby, D., Mtero, F., 2014. The impact of land property rights interventions on investment and agricultural productivity in developing countries: a systematic review. *Campbell Syst. Rev.*
- Lay, Jann, Nolte, Kerstin, Sipangule, Kacana, 2018. Large-Scale Farms and Smallholders: Evidence from Zambia. Working Paper 2098. Kiel Institute for the World Economy. ISSN 2195-7525.
- Lowder, S., Skoet, J., Raney, T., 2016. The number, size and distribution of farms, smallholder farms and family farms worldwide. *World Dev.* <https://doi.org/10.1016/j.worlddev.2015.10.041>.
- Mackinnon, J.G., Magee, L., 1990. Transforming the dependent variable in regression models. *Int. Econ. Rev.* 31 (2), 315–339.
- Melesse, M.B., Bulte, E., 2015. Does land registration and certification boost farm productivity? Evidence from Ethiopia. *Agric. Econ.* 46 (6), 757–768.
- Mellor, J., 1976. *The New Economics of growth: A Strategy for India and the Developing World*. Cornell University Press, Ithaca.
- Mujenja, F., Wonani, C., 2012. *Long-Term Outcomes of Agricultural Investments: Lessons from Zambia*. International Institute for Environment and Development, IIED, London.
- Mundlak, Y., 1978. On the pooling of time series and cross section data. *Econometrica* 46 (1), 69–85.
- National Bureau of Statistics. 2012. Large Scale Farms: Volume IV of Reports from the National Sample Census of Agriculture 2007/08. The Tanzanian National Bureau of Statistics and the Office of the Chief Government Statistician, Zanzibar, in collaboration with the Ministry of Agriculture, Food Security and Cooperatives, Ministry of Livestock Development and Fisheries, Ministry of Water and Irrigation, Ministry of Agriculture, Livestock and Environment, Zanzibar, Prime Minister’s Office, Regional Administration and Local Governments, Ministry of Industries, Trade and Marketing. Dar es Salaam. June, 2012.
- Poulton, C., 2018. Kilombero Stories. *Futures Agric.*
- Rakotoarisoa, M.A., 2011. A Contribution to the Analysis of the Effects of Foreign Agricultural Investment on the Food Sector and Trade in Sub-Saharan Africa. Paper presented at the 14th Annual Conference on Global Economic Analysis. Venice, Italy, June 16–18.
- Schoneveld, G., 2014. The geographic and sectoral patterns of large-scale farmland investments in Sub-Saharan Africa. *Food Policy* 48, 34–50.
- Sitko, Nicholas, Burke, W.J., Jayne, T.S., 2018. The quiet rise of large-scale trading firms in East and Southern Africa. *J. Develop. Stud.* 54 (5), 895–914.
- Sitko, Nicholas, Jayne, T.S., 2014. Structural transformation or elite land capture? The growth of “emergent” farmers in Zambia. *Food Policy* 48, 194–202.
- Sitko, Nicholas, Chamberlin, Jordan, 2016. The geography of Zambia’s customary land: assessing the prospects for smallholder development. *Land Use Policy* 55, 49–60.
- Sokoloff, K.L., Engerman, S.L., 2000. Institutions, factor endowments, and paths of development in the new world. *J. Econ. Perspect.* 14 (1), 217–232.
- Stanton, B.F., 1991. Farm Structure: Concept and Definition. Cornell Agricultural Economics Staff Paper, 91-6. Department of Agricultural Economics, Cornell University, Ithaca.
- van der Westhuizen, D., Jayne, T.S., Meyer, F.H., van Niekerk, J., 2018. Causes and Consequences of Rising Tractor Use in Sub-Saharan Africa: Evidence from Tanzania. Paper for presentation at the International Conference of Agricultural Economists, Vancouver, Canada, July 30, 2018.
- Vollrath, D., 2007. Land distribution and international agricultural productivity. *Am. J. Agric. Econ.* 89 (1), 202–216.
- von Braun, J., Meinzen-Dick, R.S., 2009. Land Grabbing by Foreign Investors in Developing Countries: Risks and Opportunities. International Food Policy Research Institute, Policy Brief. Washington, DC: IFPRI.
- Wooldridge, J.M., 2010. *Econometric Analysis of Cross Section and Panel Data*, second ed. MIT Press, Cambridge, MA.
- Yeboah, F.K., Jayne, T.S., 2018. Africa’s evolving employment trends. *J. Develop. Stud.* 54 (5), 803–832.