Female labor outcomes and large-scale agricultural land investments: Macro-micro evidence from Tanzania

Evans S. Osabuohien, Uchenna R. Efobi, Raoul T. Herrmann, Ciliaka M.W. Gitau

1. Introduction

Africa has been a major destination for large-scale agricultural land investments (LALIs), with previous studies estimating that it accounts for more than 65 percent of the world’s LALIs, more than 80 percent of investments (LALIs), with previous studies estimating that it accounts for more than 65 percent of the world’s LALIs, with previous studies estimating that it accounts for more than 65 percent of the world’s LALIs. LALIs are those investments (mainly in agriculture) that involve the acquisition of land above 200 hectares of land. The availability of cheap labor and the region’s weak institutional framework have all been identified as the principal factors attracting investors to African countries (Cotula et al., 2009; Deininger et al., 2011; Osabuohien et al., 2015; Sulle, 2017; Bluwstein et al., 2018). The upsurge of land investors in Africa can also be linked to the current agricultural transformation agenda in Africa, which aims to advance agriculture from the subsistence level to a more viable commercial level (Osabuohien, 2014).

The significance of LALIs increased in the wake of the new millennium and particularly after the 2007–2008 global financial and economic crises, driven by investors seeking a more stable investment platform and by the desire of many African governments to revitalize their countries’ agricultural sectors (Wolter, 2009; Doss et al., 2014; Osabuohien et al., 2015; Osabuohien, 2018). Employment opportunities, rural development, and provision of social amenities are the promised benefits. At the same time, LALIs have been criticized for potentially displacing smallholders, contributing to land conflicts and negatively affecting access to them, and potentially causing loss of ecosystems (De Schutter, 2011; Massay, 2015). Concerns have also focused on the potential employment effects, which have been estimated to be relatively little (Nolte and Ostermeier, 2017), while providing mainly short-term and of casual jobs. The extent to which LALIs deliver on their promises, particularly for women who are highly vulnerable to socio-economic shocks, thus, remains highly controversial and underexplored.

Similar to agriculture in other developing countries, Tanzania’s agricultural sector plays a prominent role in enhancing food security, income generation and employment creation. The sector employs more...
than 70 percent of the country’s population and over 80 percent of Tanzania’s female population (Tanzanian National Bureau of Statistics -TNBS, 2015), making women particularly vulnerable to negative changes in agricultural and land-related activities. Tanzania has received sizeable LALIs in terms of both number of investments and acreage. Available statistics show that the country has been among the top 20 recipients of LALIs globally, as well as among the top 10 LALIs destinations in Africa (Nolte et al., 2016; Land Matrix Global Observatory-LMOGO, 2017). The pattern of LALIs in Tanzania exhibits some variation across regions (districts), purposes (intentions), sizes, and crops cultivated. Most of the land investments in the country have been targeted at the agricultural sector (Abdallah et al., 2014; Enstrøm and Hajdu, 2018). Therefore, these investments are expected to have significant effect on women, who have higher participation in agricultural activities.

Consequently, the main argument of this study is that although most Tanzanian women eke out their livelihood from the agricultural sector, women’s economic gains from LALIs may not be substantial compared to that of men. This presupposes that when there are issues of land displacement emanating from the occurrence of LALIs, the effects on women might be higher as they are often referred to as vulnerable group (George et al., 2015). We provide empirical insights into three important questions: a) to what extent is the welfare of female-headed households in communities with LALIs affected compared to those living in communities without LALIs?; b) how does the presence of LALIs influence the agricultural wage income of female-headed households?; and c) how does the occurrence of LALIs affect the agricultural labor allocation of female-headed households? We use the Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) dataset to provide macro evidence and complement it with two case studies of LALIs in Kilombero district in Morogoro region to provide the micro evidence.

This inquiry contributes to the growing literature on the consequence of LALIs on households’ welfare outcomes in Africa. These studies include Herrmann (2017) who studied wage and labour outcome of Tanzanian households from LALIs, while Bottazzi et al. (2018) focused on livelihood impacts from a LALI in Sierra Leone. Zaehringer et al. (2018) in addition focused on LALI’s impact on land use and the environment on smallholders in Kenya, and Baumgartner et al. (2015) earlier linked LALI to income, price, and employment on households in Ethiopia. These authors find a significant relationship between LALI and the varying outcomes that were considered, but one important conclusion from the studies, especially as regards labor market effect, is that despite the important effect of LALIs, these investments crowd out smallholder farmers, which is only partially mitigated through the cultivation of labor intensive crops and the application of contract farming schemes (Nolte and Ostermeier, 2017). Noting the importance of these conclusions, our study focuses on female labor participation in Tanzania as an important outcome from LALI, which is a significant contribution for a more specific evaluation of LALI’s presence in low-income countries.

Focusing on female labor outcome is important for the following reasons: first, women constitute the major participants in the rural agricultural sector in Africa (World Bank, 2015; International Labor Organization-ILO, 2016; Kasiwa, 2018), where most LALIs are located. This suggests that LALI’s presence can have an effect on women, which is not yet well established in extant studies. Second, since women participate more in the unpaid household labor – such as home-based activities (Food and Agricultural Organization-FAO, 2011; Adeoti and Oni, 2018) - and their agricultural activities are a source of additional income for household expenditure, the occurrence of LALIs can have an on the households’ welfare.

The remainder of the paper is organized as follows. Section two discusses the stylized facts, where background data is used to situate the agricultural and labor condition of Tanzania in relation to women. The third section describes the research methods. Presentation and discussion of results are contained in the fourth section; implications of findings in section five, while section six concludes.

2. The agricultural sector and labor market in Tanzania

Agriculture remains a dominant sector in most developing countries, especially in rural areas (Li, 2010, 2011; West and Haug, 2017; Bluwstein et al., 2018; Enstrøm and Hajdu, 2018). In Tanzania, more than 70 percent of women earn their livelihoods from agricultural activities (TNBS, 2015). Over the last decade, Tanzania has recorded high economic growth; the GDP per capita has more than doubled, from US $348 in 2004 to US$955 in 2014. Despite Tanzania’s high growth rates (an average of 7 percent in the last five years); however, it seems that limited benefits are accruable to low-income population.

Tanzania has good climatic conditions for agricultural cultivation. The coastal region is characterized by tropical conditions, while along the highlands temperate conditions are prevalent. In addition, Tanzania is home to three of the major lakes in Africa; this is important, as the country’s agricultural sector is dependent on natural water supplies, with limited mechanized farming or irrigation systems. The main cash crop is coffee, followed by sisal, cashews, cotton, tobacco, tea, and oil seeds (Tanzanian National Bureau of Statistics-TNBS, 2015). Over 80 percent of Tanzania’s poor and extreme poor live in rural areas. These rural areas, which are mainly dependent on the agricultural sector, remain largely underdeveloped with poor infrastructure and limited access to public amenities (Headay and Jayne, 2014), as well as limited established links with productive value chains and higher value crops (Herrmann et al., 2018). Women play an important role in agricultural activities in these areas, who are constrained by family obligations, limited access to finance, and limited ownership of productive resources (Danquah et al., 2017; Bluwstein et al., 2018).

The information in Table 1 shows that the agricultural sector of Tanzania has remained buoyant in the absorption of the working population. In 2006, the agricultural sector employed a total of 75.1 percent of Tanzanian youth and adults. More women worked in the agricultural sector during this period, accounting for 40.8 percent; only 34.3 percent of men worked in the sector. In 2014, the trend was not much different – the agricultural sector employed about 61 percent of the total youth and adult population. This trend reinforces the relevance of the agricultural sector in employment generation in Tanzania.

The employment statistics show that approximately 94 percent of households in rural areas have at least one member working in the agricultural sector (see Table 2). In addition, approximately 93 percent of these households have a household head that is not well-educated (only primary education or less). In Dar es Salaam (the capital city), about 7 percent of households report having at least one person involved in the agricultural sector, and approximately 54 percent of the households have at least one member engaged in wage employment, compared to 15 percent in rural areas.

LALIs are mainly located in rural areas, which mean that they have significant impact on the majority of Tanzania’s poor and on how these

<table>
<thead>
<tr>
<th>Sector</th>
<th>Total Male</th>
<th>Female</th>
<th>Total Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central and Local Government</td>
<td>2.6</td>
<td>1.8</td>
<td>3.2</td>
<td>1.8</td>
</tr>
<tr>
<td>Parastatal</td>
<td>0.4</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>Organization</td>
<td>79.5</td>
<td>43.4</td>
<td>80.9</td>
<td>43.7</td>
</tr>
<tr>
<td>Agriculture</td>
<td>10.1</td>
<td>5.6</td>
<td>15.7</td>
<td>9.7</td>
</tr>
<tr>
<td>Private Informal Sector</td>
<td>8.6</td>
<td>6.1</td>
<td>14.7</td>
<td>9.3</td>
</tr>
<tr>
<td>Private Other Sector</td>
<td>6.6</td>
<td>5.2</td>
<td>11.8</td>
<td>9.5</td>
</tr>
<tr>
<td>Household Duties</td>
<td>3.1</td>
<td>0.5</td>
<td>3.6</td>
<td>0.7</td>
</tr>
</tbody>
</table>

poor populations earn their livelihoods. The trend of LALIs in Tanzania suggests strong interest in land investment in the agricultural sector. More than 95 percent of the investors have agriculture as their main motive for land acquisition in Tanzania (Land Matrix Global Observatory-LMGO, 2017). With this rising interest in LALIs in Tanzania, this study analyses whether these investments translate into improved households’ welfare outcomes in communities where they are situated. We focus on three important issues – consumption, agricultural wages, and the labor hour in the agricultural sector, particularly for women.

3. Data sources and method of analysis

3.1. Data sources

Two main data sources are applied in this study: a national survey (Living Standards Measurement Study-Integrated Surveys on Agriculture, LSMS-ISA) and a field survey carried out in Kilombero district in Morogoro region in 2013. The choice of the case study districts was informed by the importance of the region for ongoing LALIs and as a historical focus of rice and sugarcane plantations. The combination of these two datasets provides a broad perspective as well as a case-specific perspective.

The LSMS – ISA dataset includes data on household- and community-related variables. The LSMS–ISA for Tanzania is currently conducted in three waves, forming a panel dataset. The first wave (2009–2010) constitutes a representative sample of 3265 households. In the second wave (2010–2011), the original households were revisited; all adults who had relocated were tracked down and re-interviewed. The sample size for the second wave expanded to 3924 households, and the Global Positioning System (GPS) coordinates for the households were recorded. The third wave (2012–2013) sample size was 3924 households.

The case study data involve household survey data and qualitative information collected through two surveys between April and June 2013 in villages within a 50 km radius of two case study investments in the Kilombero region of Tanzania, which has attracted substantial interests for LALIs in Tanzania, especially as part of a public-private partnership investment initiative, the Southern Agricultural Growth Corridor of Tanzania (SAGCOT). The Kilombero Valley is situated in low-altitude freshwater wetland, the largest in East Africa (Kangalawe and Liwenga, 2005; Bluwstein et al., 2018). The area is very conducive for crop production as it has both access to water and favorable climatic conditions (Beck, 1964). This region has experienced significant migration in recent decades, with farmers migrating with the aim of cultivating different crops, mostly rice (Kangalawe and Liwenga, 2005; Sulle, 2017; West and Haug, 2017; Bluwstein et al., 2018).

The Kilombero region is currently the largest rice producer in Tanzania, yet with most rice farmers in this region are engaged in small-scale farming (Tanzanian National Bureau of Statistics - TNBS, 2012). The Kilombero Cluster was identified as one of the first development clusters within SAGCOT. The Kilombero Cluster, covering Kilombero District and part of Kilosa District comprises an estimated one million hectares, 312,000 ha of which have been estimated to be arable, yet with only 80,000 ha under cultivation (SAGCOT, 2011). According to the SAGCOT Blue Print (SAGCOT, 2011) the Kilombero Valley Zone is a suitable location for large-scale investments in irrigated sugarcane and rice production/processing. At the time of this study, there had been no new commercial investments implemented in Kilombero under the SAGCOT initiative. Yet the two investments analyzed in this study, the Kilombero Sugar Company Limited (KSCL) and Kilombero Plantation Limited (KPL), while predating the initiative, have been included as part of SAGCOT and showcased as models for “responsible” investments in Tanzania (SAGCOT, 2012).

3.2. LALIs identification and matching

The first step in the data analysis was to track the presence of LALIs using both the Land Matrix Global Observatory (LMGO) data and the LSMS-ISA data. The LMGO data allows for the identification of communities with LALIs at the time that the LSMS-ISA data was collected in 2012-2013. The communities identified to host LALIs are categorized as one and the communities that do not have any LALIs are categorized as zero. Only those LALIs that have been established and are currently operational are considered for our matching process. Apart from the information from the LMGO, we also confirmed the activities of the LALIs from the Tanzania Investment Centre (TIC) during the field work.

The classification of household across communities with LALIs presence is such that we identified community names with LALIs as stated in the land matrix dataset, we then traced such communities in the LSMS-ISA dataset and households in such communities are then classified as one. Those households that are not in such communities are classified as zero. The challenge of with the classification is that we are not able to accurately relate whether households in such communities work in the LALIs investments, or whether such households take on contractual production and sale agreements like out-grower schemes.

3.3. Outcome variables

The three outcome variables of interest include: household consumption, household agricultural wages, and agricultural hour allocation per month. These three outcome variables capture diverse mechanisms through which LALIs can influence household welfare. Household consumption is measured as the sum of all household food and non-food consumption per month. For household agricultural wages, the computation is based on the total wage per month that is received by the household from agricultural activities, such as agricultural labor input. The third outcome variable is measured as the total average hour spent on agricultural activities per month. These three outcome variables are obtained from the LSMS-ISA dataset.

3.4. Estimation technique

To examine the data for comprehensiveness and expected relationships, we employed descriptive statistics and kernel density plots. The

---

Table 2

<table>
<thead>
<tr>
<th>Source: Authors’ computation from TNBS (2015)</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least one Person in Household Engaged in Wage Employment</td>
</tr>
<tr>
<td>At least one Person in Household Engaged in Agricultural Sector</td>
</tr>
<tr>
<td>Household Heads with Primary Education and Less</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

2 Attempt was made to provide more details on the descriptive statistic of the outcome variables before using them in the estimation; however, space consideration constrained us to focus more on the econometric estimation.

3 The Double Difference (DD) approach was equally used as a robustness check. We only report highlights of the result in Section 5 and do not give details regarding the technique due to space limitations.
objective of this empirical approach is to compare the estimated mean effect on households between communities with LALIs and communities without LALIs with respect to our three outcome variables. The main advantage of this empirical strategy is its ability to generate a comparison group of households with a similar distribution of characteristics in communities both with and without LALIs. The comparison captures the effect of LALIs on households in communities with LALIs (the treatment group), controlling for household characteristics. The treatment effect is therefore calculated as the difference of the mean outcomes.

To explain this process in mathematical terms, we assume that there are two groups of female-headed households indexed by their location, such that \( P = 0/1 \), where 1 (0) indicates that a household is (not) located in a community with LALIs. The presence of LALIs is expected to yield gains to households, where:

\[ Y^1_i: \text{the consumption, income, or agricultural labor allocation for a household in communities with LALIs (i.e. } P = 1) \]  
\[ Y^0_i: \text{the consumption, income, or agricultural labor allocation for a household in communities without LALIs (i.e. } P = 0) \]

Therefore, the Average Treatment on the Treated Effect (ATT) will be such that:

\[ \text{ATT} = E(Y^1_i - Y^0_i | P = 1) \]  
Eq. 1 can be simplified as:

\[ \text{ATT} = E[(Y^1_i | P = 1) - (Y^0_i | P = 1)] \]  
(2)

Where: \( E(.) \) represents the average (or the expected value). This equation attempts to capture the consumption, income, or agricultural labor allocation of the household in a community with LALIs compared to what that household would have experienced without LALIs.

The data on \( (Y^1_i, Y^0_i | P = 1) \) is available in the national survey dataset. However, the data on \( (??^0_i, ??^1_i | P = 0) \) is missing. The matching process compares the effect of the presence of LALIs on a household’s main outcomes with those of matched non-participants (those households located in communities without LALIs). The matches are chosen on the basis of similarity in observed household characteristics. The household characteristics considered include: household location, number of individuals in the household, age and education of the household head, household credit access, number of assets owned by the household, whether the household owns a plot of land, and whether the household cultivates that land. Rosenbaum and Rubin (1983) advance this approach by proposing the use of Propensity Score Matching (PSM) as a reliable technique to derive the equivalent non-participant data.

The underlining assumptions guiding the PSM analysis are (i) the conditional independence assumption and (ii) the common support condition. The conditional independence assumption assumes that the potential outcomes for households in communities without LALIs are independent of their status of being located in these communities, given a set of observable covariates “\( X \).

\[ \text{i.e.} \ Y^0_i, Y^1_i | P, X \]  
(3)

Hence, after adjusting for observable differences, the mean of the outcome variable is the same for both households in communities with LALIs and households in communities without LALIs. This condition allows us to use matched households in communities without LALIs as a control group when measuring the effect of LALIs on household welfare. Hence, Eq. (2) can be represented as:

\[ (Y^1_i | P = 1, X) = (Y^0_i | P = 0, X) \]  
(4)

The second assumption, common support condition, is based on the expectation that for each value of “\( X \)”, there is a positive probability of either being located in communities with LALIs or not located in communities with LALIs. This assumption supports the overlap condition such that the proportion of these two categories of households must be greater than “0” for every possible value of “\( X \)”. Hence, it ensures that there is a sufficient overlap in the characteristics of the two groups of households to find adequate matches. Once these two conditions are satisfied, the treatment assignment is said to be efficient (Rosenbaum and Rubin, 1983). This approach has been used in recent studies such as Nkhata et al. (2014).

There are different matching algorithms that are applied when using the PSM approach. This study uses the Nearest Neighbor Matching (NNM) and Kernel Matching (KM) algorithm, as they are deemed sufficient for our analysis. The NNM algorithm compares the outcome of households in communities with LALIs with the closest and most similar households in communities without LALIs, using propensity scores. This approach minimizes the distance between the propensity score of the two groups of households. Thus, it is expressed as:

\[ \text{i.e. } \min \ |P_i - P_j| \]  
(5)

The KM algorithm produces more efficient results and is better suitable for dealing with large, asymmetrically distributed dataset (Baser, 2006). The KM is structured such that each household in locations with LALIs “i” are matched (using propensity scores) with other control observations that have weights that are inversely proportional to the distance between them and those households in locations without LALIs. The weight is computed as:

\[ w_j = \frac{h^2/\Delta_1}{\sum_{i=1}^{n} \Delta_1} \]  
(6)

Where: \( h \) is the bandwidth. Households located in communities with (and without) LALIs are indicated as ‘i’ (and ‘j’).

4. Results and discussions

4.1. Results from the kernel density plots

The kernel density plots are presented for the households’ consumption, income, and labor allocation. The objective is to approximate the density function of the outcome variables and compare the trends for households in communities with and without LALI. The kernel density plot for household consumption is shown in Fig. 1. The plot shows that the total household consumption in communities without LALIs (the comparison group) is tilted to the right compared to the density plot of the households in communities with LALIs. This difference is significant at the one percent level of the Kolmogorov-Smirnov test for equality of distribution, connoting that households in communities without LALIs tend to have higher total consumption expenditure in terms of the total food and non-food expenditure. On average, these households...
households spend about Tsh. 18,000 (equivalent to US$11.4)\(^4\) in a week on household consumption compared to the households located in communities with LALIs, which spend about Tsh. 10,000 (US$6.3).

For the household average monthly wages as shown in Fig. 2, the income disparity plot of households in communities with LALIs is tilted to the left, while that of households in communities without LALIs is rightward sloped. The difference between the two density plots is significant at 1 percent based on the Kolmogorov-Smirnov test. This suggests that households in communities where the LALIs are located possibly earn less than households in communities without the LALIs. On average, while households in communities without LALIs earn a monthly wage that is over Tsh. 98,000, (US$61.8), households in communities with LALI earns about Tsh 51,000 (US$32.2).

Finally, the kernel density plot of households’ agricultural labor allocation is shown in Fig. 3. The results depict that households in communities with LALIs allocate more time to agricultural activities than households in communities without LALIs. The density plot for households in communities with LALIs is rightward biased compared to households in communities without LALIs. This difference is significant at 1 percent, implying that households in communities with LALIs tend to put in more time in agricultural activities than households in communities without LALIs.

The results from the three kernel density plots for the different outcome variables suggest that while households in communities with LALI tend to allocate more time to agricultural activities, they tend to have less household consumption and income. Possibly, compared to the income from farm activities and other non-farm income that households in communities without LALI derive from their activities, the income from working in the farmlands of LALI for households located in such communities with LALI tend to be smaller. More so, households that work for LALI may spend more household time in the investor’s agricultural activities, such that they do not have adequate time to be involved in other non-farm revenue-generating activities. Further, displacements that might occur with LALIs could mean that households in communities with LALIs do not own sufficient land on which they can cultivate sufficiently for own consumption or to sell to generate income. This might affect overall household income and household’s total food and non-food consumption.

There are important differences in the outcome variables when considering the gender dimension of the debate. As shown in Fig. 4 for household consumption by gender of the household head, the density plot of male-headed households tends to be rightward biased compared to that of female-headed households. This implies that for communities without LALIs, the results indicate that on average, male-headed households spend about Tsh 20,000 (US$12.6) in a week on household consumption, while female-headed households spend about Tsh16,000 (US$10.1). A similar trend is observed in communities with LALIs, except for the fact that the difference in total household spending across the two groups is not significant. In essence, male-headed households in communities with LALIs spend about Tsh11,000 (US$6.9) on household consumption, while female-headed households spend similar amount of about Tsh10,000 (US$6.3).

Similarly, male-headed households have higher household income compared to their female-headed counterparts, as shown in Fig. 5. In both communities i.e. with and without LALIs, the kernel density plots for female-headed households tilt more to the left. In communities without LALIs, male-headed households earn about Tsh105,000 (US $66.20) compared to their female counterparts, who have an income of about Tsh 72,000 (US$45.40). This is similar for communities with LALIs, where male-headed households earn about Tsh 55,000 (US $34.70) compared to female-headed households earning Tsh 38,000 (US$24). The persistence of this income disparity across the male- and female-headed households in communities with LALIs may raise the question whether the presence of LALIs may contribute to gendered income disparities.

The kernel density plot for labor allocation to agricultural activities by gender is presented in Fig. 6. For communities both with and without LALIs, the density plot for male-headed households falls to the right compared to the density plot for female-headed households. This suggests that male-headed households tend to allocate more time to agricultural activities compared to female-headed households. The PSM estimations are applied to further analyze the impact of LALIs on household consumption, income, and labor allocation to agricultural activities.

This trend reveals the structural outlook of the Tanzanian labor market, where on the average the females are paid less than their male counterparts; thereby disadvantaging them in different metrics of development, including consumption and income (TNBS, 2015). Despite that women are more engaged in the agricultural sector compared to their male counterparts, the male individuals are more favored in working for LALIs and are remunerated differently than females, which might be due to the patriarchal social systems in relation to land ownership and rights (Mugabi, 2013).

4The average exchange rate used is Tanzania shilling (Tsh,) 1586 to US$ 1. This was the approximate average of exchange rate in 2012 and 2013.

4.2. Descriptive statistics

The household characteristics of interest include household location (i.e. rural or urban), average age of household, educational attainment of household head, number of household members, total assets owned by the household, and households’ credit access, and land ownership. The descriptive statistics are presented in Table 3, which compares these characteristics for households in communities with and without LALIs. These variables are chosen following the determinant of LALIs’ presence in communities (Osabuohien, 2014; Osabuohien et al., 2015).

It is evident from Table 3 that the groups’ means differ across all
variables except for the number of individuals in a household. Focusing on the outcome variables, households also differ across their consumption, the average wage earned, and the time allocated to agricultural activities. Households in communities without LALIs have a higher consumption (Tsh 7171 (US$4.5)) compared to households in communities with LLI, with an average Tsh 3078 (US$1.9). Households in communities with LLI earn wages of about Tsh 57,493 (US$36.3) relative to Tsh 101,223 (US$63.8) for households living in communities without LALIs. Similar to the kernel density plot, households located in communities with LALIs tend to put in more time to agricultural activities than households located in communities without LALIs. This outlook supports the earlier findings from the kernel density plots. The difference in hours spent on agricultural activities across the two groups is significant at 1 percent level of significance. Disaggregating the results by gender, female-headed households have lower household consumption and income, but higher labor hour allocation to agricultural activities. For all the other variables (except household head age) there are significant differences across the gender of the household head in communities with and without LALIs.

4.3. Econometric results

4.3.1. The probit model and balancing tests

To design a set of variables that can match household characteristics in communities with and without LALIs, we apply a probit model. The main intention for estimating the probit regression model is to balance the differences in the observable characteristics that may be present between the groups (i.e. those located in communities with LALIs and those not located in such communities).

Table 4 shows the probit model result, which was used to derive the propensity scores. For all households, household size, age of the household head, whether or not a household is cultivating land or owning and cultivating farm plots (Column 1) are found to be significantly associated with women’s labor participation. The correlation of LALIs and the variables discussed earlier is as expected, since households in rural areas, with larger household size and older household heads, and with plots of land to cultivate have a higher likelihood of being located in communities with LALIs.

In Column 2, a positive coefficient is reported for having a household member who attends school. Therefore, it is possible that as a
result of employment opportunities, households with more educated heads may be located in communities with LALIs. Owning and cultivating farm plots and having more access to credit show a negative association with being located in communities with LALIs.

The results of the balancing quality checks are presented in Fig. 7 and Table 5. The histograms of the predicted propensity scores for both the treated and the control groups are presented in Fig. 7, Panels (a), and (b) for the total sample and the female sub-sample, respectively. From the figure, it is evident that the propensity scores are of a relatively equal distribution and are within a similar range, suggesting comparability of the treatment and control groups. The implication of the figures in the corresponding panels suggests that the common support area includes most of the sampled households. Most of the treated households have an equivalent match on the comparison sample; this is important for the matched sample in the communities with LALIs to be representative of the initial sample.

Table 5 shows that most of the sample in the control group has a match within the treatment group.

The matching quality is also assessed using group mean and median comparisons to check the extent to which the differences in the...
characteristics across groups (treatment and control) are reduced as a result of the matching process. The difference in pre-intervention characteristics is balanced after matching and an appropriate counterfactual outcome derived. In Table 5, these differences are reported as unmatched and matched for both the total sample and the sub-sample of female household members, and both the nearest neighbor with five neighbors and kernel matching algorithm (with the standard 0.06 bandwidth). In both the nearest neighbor and kernel matching technique, the pre-intervention differences between households in communities with LALI and those in communities without LALI are reduced. This is evidenced by the non-significant values of the likelihood ratio (LR) Test. More so, the mean and median absolute bias is also significantly reduced in both techniques

The matching quality for the female-headed household sub-sample in the communities with (and without) LALIs is shown in Table 5. In both the nearest neighbor kernel matching algorithm, the difference in the observable characteristics that exist among the two groups is reduced. For instance, the mean and median absolute bias for the models that comprise total household consumption and household labor allocation are reduced in both matching algorithms. For the total household wages for both the NNM and KM matching algorithm for the female-headed household sub-sample, it is evident that there was no significant reduction in the mean and median differences across the sample in communities with (and without) LALI. However, a closer look at the mean and median values of the total household wages after the matching shows that these differences are somewhat reduced. Therefore, the probit model results indicate that there are comparable households to estimate the Average Treatment Effect.

### 4.3.2. The propensity score matching results

The average treatment effect (ATT) estimation results for the different outcome variables across the matching algorithms are reported in Table 6.

We compare the results in Table 6 with the Ordinary Least Squares (OLS) for sensitivity and robustness checks. The ATT for household consumption and income is negative and significant. The results suggest that the location of households in communities with LALI accounts for between 28 and 39 percent reductions in their total consumption and between 42 and 46 percent reductions in their total wages. However, the presence of LALIs is significantly associated with households’ time allocated to agricultural activities; the ATT was positive for the entire matching technique. Thus, there is a marginal increase in the number of hours per week that households allocate to agricultural activities, yet this increase relatively small from 0.03 to 3 percent.

The female-headed household sub-sample reports different results. For household consumption, it is observed that a negative average difference still exists when considering LALI presence; the reduction in consumption is between 21 and 42 percent for the different matching algorithms. Thus, the presence of LALIs is associated with negative changes for female-headed households (in terms of household consumption). Similarly, for the time allocated to agricultural activities, our model estimates a positive ATT, ranging from 8.9 to 12.4 percent across the different matching algorithms. The analysis is further disaggregated just for agricultural wages, as presented in Table 7. The results do not register any significant change as in Table 6, especially when considering the signs of the coefficient.

### 4.4. Sensitivity analysis and robustness checks

To check the robustness of the ATT, the propensity scores are re-estimated by applying direct NNM before estimating the propensity score equation. This approach estimates the ATT on the outcome variables by using direct nearest-neighbor matching with one match per treatment (in our case, per household in communities with LALI). The rationale for this is to examine whether the ATT results change; if they do not change, this shows that they are reliable.

The results in Table 8 indicate that sign of the ATT coefficient for the different outcome variables remained the same for both the total sample and the sub-sample of female-headed households. For instance, there is a Tsh 217.07 reduction in consumption, a Tsh 8,027.73 reduction in the total household wages, and a 0.054 increase in the number of hours put into agricultural activities. For female-headed households, consumption remains negative, but as usual, household wages and labor input (in terms of hours put into agricultural activities) significantly increase.

#### 4.4.1. LALIs impact at the individual level: further evidence

The estimations in this study are based on the household level (in terms of female-headed households). However, we take a step further to consider the impact of LALIs at the individual level. In this section, the ATT was re-estimated, taking into consideration individual females in the Tanzanian national survey who experienced a higher labor hour input in the agricultural sector, but women’s relative consumption and wages did not change, this shows that they are reliable.

The results presented in Table 9 are similar to the earlier result seen in Table 6. The OLS, NNM, and KM were not different from those in Table 6 in terms of the signs and significance of the coefficients of the ATT; the exception is wages, which is now negative in Table 9. Women in the Tanzanian national survey will experience a higher labor hour input in the agricultural sector, but women’s relative consumption and wages maintain a reduced value in communities with LALIs.

#### 4.4.2. Average treatment effect based on double difference estimate

The Double Difference (DD) is further estimated to examine the consistency of our estimates in terms of the ATT when controlling for the difference between the treatment and control groups. This estimation approach is valuable because it adjusts for the time-varying factors that may affect the outcome variables and because it eliminates bias that arises, assuming that these factors are not controlled for.
the DD approach controls for unobserved heterogeneities that may affect the outcome variables (apart from LALIs) as Gertler et al. (2016) emphasized that the DD should be included in PSM estimates for robustness. The results of the DD for both female-headed households and the female gender are presented in Table 10.

The ATT estimates in Table 10 show similar results to those in Table 6. The household consumption continues to have a significant negative sign, suggesting that the presence of LALIs in communities reduces the consumption of female-headed households. This result is also applicable when considering the female gender in the sample. For household wages, the results are insignificant using the DD approach. The result for the agricultural time input, however, confirms a similar result to that seen in Table 6. The significant value of the outcome variable was consistent both for the female-headed household and the entire female population in the sample. Therefore, we surmise that the consumption and agricultural labor hour input effects could be verified with the DD approach; however, the total wage effect could not be verified. However, this does not raise much concern, as the sign of the variable remains consistent with Table 9.

4.5. Case studies evidence

To provide further evidence, two cases of LALI were observed from a field survey of Kilombero sugar and rice plantations in Tanzania. Kilombero Sugar Company Limited (KSCL), located in the northern part of the Kilombero Valley, was established in the early 1960s as the first major commercial project in Kilombero, with a total concession area of more than 10,000 ha. It was acquired in 1998 by a British-South African consortium that holds 75 percent of shares, with the remaining shares held by the Tanzanian Government (TNBS, 2012, 2015). KSCL produces...
Table 6
Estimated Average Treatment Effect.
Source: Authors’ Computation.

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>%difference</th>
<th>NNM</th>
<th>%difference</th>
<th>KM</th>
<th>%difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Household</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Consumption</td>
<td>−678.74***</td>
<td>−39.42%</td>
<td>−496.40*</td>
<td>−28.80%</td>
<td>−485.59***</td>
<td>−28.20%</td>
</tr>
<tr>
<td>Wages</td>
<td>−25313.2</td>
<td>−44.95%</td>
<td>−23475.9**</td>
<td>−41.70%</td>
<td>−25711.7***</td>
<td>−45.70%</td>
</tr>
<tr>
<td>Time allow</td>
<td>0.105***</td>
<td>3.46%</td>
<td>0.103**</td>
<td>0.03%</td>
<td>0.103**</td>
<td>3.40%</td>
</tr>
<tr>
<td><strong>Female Headed Household</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Consumption</td>
<td>−2933.79***</td>
<td>−20.96%</td>
<td>−2919.83***</td>
<td>−21.34%</td>
<td>−2136.57***</td>
<td>−42.45%</td>
</tr>
<tr>
<td>Wages</td>
<td>19098.96</td>
<td>20.88%</td>
<td>339.06</td>
<td>0.371%</td>
<td>11,373.34</td>
<td>12.432%</td>
</tr>
<tr>
<td>Time allow</td>
<td>0.295** (0.037)</td>
<td>9.473%</td>
<td>0.385** (0.021)</td>
<td>12.364%</td>
<td>0.278*(0.092)</td>
<td>8.927%</td>
</tr>
</tbody>
</table>

Note: Probability values are in parentheses *p < 0.1, **p < 0.05, ***p < 0.01.
OLS: Ordinary Least Squares; NNM: 5 Nearest Neighbor Matching; KM: Kernel Matching.

Table 7
Estimated Average Treatment Effect of only Agric Wages.
Notes and Source: Same as Table 6.

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>%difference</th>
<th>NNM</th>
<th>%difference</th>
<th>KM</th>
<th>%difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Household</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wages</td>
<td>−4180.507 (0.768)</td>
<td>−13.67%</td>
<td>−9489.846 (0.532)</td>
<td>−31.03%</td>
<td>−943.768 (0.950)</td>
<td>−3.09%</td>
</tr>
<tr>
<td><strong>Female Headed Household</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wages</td>
<td>8927.392 (0.837)</td>
<td>23.40%</td>
<td>16,555.56 (0.500)</td>
<td>43.93%</td>
<td>18,921.97 (0.332)</td>
<td>49.60%</td>
</tr>
</tbody>
</table>

Table 8
Sensitivity Check using the Direct nearest Neighbor Matching.
Source: Authors’ Computation.

<table>
<thead>
<tr>
<th></th>
<th>Matched Pairs</th>
<th>Consumption ATT (Sig)</th>
<th>Wages ATT (Sig)</th>
<th>Labour Input ATT (Sig)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sample</td>
<td>1</td>
<td>−217.072 (0.350)</td>
<td>−8027.733** (0.031)</td>
<td>0.054** (0.031)</td>
</tr>
<tr>
<td>Female Hh</td>
<td>1</td>
<td>−82.404 (0.882)</td>
<td>6446.250* (0.093)</td>
<td>0.341* (0.093)</td>
</tr>
</tbody>
</table>

Note: Probability values are in parentheses *p < 0.1, **p < 0.05, ***p < 0.01. Female Hh: Female headed households; ATT: Average Treatment on the Treated Effect.

more than 60 percent of Tanzania’s sugar, and comprises of two adjacent agricultural estates and sugar factories as well as an alcohol distillery. KSCL employs 870 permanent staff and 2073 seasonal workers at peak periods (Illovo Sugar Ltd, 2015; Herrmann, 2017). In addition, more than 10,000 small and large sugarcane farmers are integrated in an out-grower scheme. The out-growers are integrated in the supply chain throughout grower associations, which have sugarcane supply agreements with KSCL. Harvesting and transportation is organized by these associations.

The second case - Kilombero Plantation Limited (KPL) - is a rice investment, which is a joint venture between Agrica Tanzania Limited (ATL), a Tanzanian subsidy of British Agrica Limited, and the Rufiji Basin Development Authority (Rubada), a parastatal mandated to promote agricultural investments in the Rufiji Basin, with ATL owning 97.54% of the shares (SAGCOT, 2011). In 2008, ATL purchased the 5,818 ha Mngeta rice farm from Rubada, which had initially started as a Tanzanian-North Korean government joint venture (KOTACO) in 1986. An industrial rice mill, a large warehouse, and an automated cleaning and drying facility have since been established by KPL (SAGCOT, 2012; Herrmann, 2017). At full operation, the investment produces rice on 5,000 ha of estate land and creates employment for about 180 full-time workers and up to 300 part-time workers (KPL, 2009). The employment on the estate and within the factory is opened mainly to the residents of the neighbouring villages to the estate. As the estate processes are highly mechanized, agricultural jobs mainly involve weeding. Other jobs available are in the processing facilities, warehouses, and support services.

Focusing on the experience of the households whose members work or do not work in both LALIs, Fig. 8 shows that the households working at the LALIs have a smaller agricultural wage income as a ratio of total household income compared to those that are not working in the LALIs. The agricultural wages of this category of households are about half of that of households in the out-grower scheme and those who are not working in the LALIs. This is also applicable to female-headed households that earn less from agricultural activities compared to those who not working in the LALIs. The differences in agricultural incomes are likely to be due to the differences in times allocated to agricultural activities. Workers spend more of their time working on the LALIs; thus, while generating high wage incomes, they trade-off their agricultural income. Nonetheless, the analysis by Herrmann (2017) on these investments shows that for the KSCL households working at the LALIs generate higher incomes than non-participants, where the differences between workers and non-participants is negligible for the case of KPL.

Overall, the out-growers in the two plantations benefit the most, and those households with land yet more so than land-poor households (Herrmann, 2017). This finding from Fig. 8 reflects the potential for LALI to translate to rural development in their host community if the investments consider schemes that support the local individuals.
engaging in their specific agricultural activities, rather than absorbing them to work for the LALI, which risk neglecting their respective agricultural activities. This shows why the out-grower scheme appears more lucrative for the individuals in the host communities (e.g. Deininger et al., 2011; Herrmann, 2017; Sulle, 2017; West and Haug, 2017; Bluwstein et al., 2018). To be more realistic, this scheme is confronted with some institutional challenges that can limit its potential of being a significant intervention in reducing household poverty in rural Tanzania (Herrmann, 2017). For instance, issues of power imbalances between sellers and buyers have been identified in the broad literature (e.g., Little and Watts, 1994) as a factor that could have undermined the potential of this scheme. More so, Herrmann (2017) identified that other governance issues within the out-grower associations, such as elite capture that is such that larger out-growers being able to influence service provision – including, for instance, harvesting and transport – thereby neglecting resource-poorer out-growers with less access. Addressing these concerns will improve the potentials of the scheme to deliver its expected outcomes in addressing rural poverty.

5. Implication of findings

The results from the propensity score matching (PSM), double difference (DD), and the case study suggest that large-scale agricultural land investments (LALIs) has a negative association with the household consumption and household income, but positively associated with the number of agricultural hours put in by the household individuals. This result can be considered from three perspectives and implications. First, institutionally, most low-income countries (including Tanzania) do not have strong institutions that protect disadvantaged population from being taken advantage of by investors. For instance, as at the time of this research, the land policies in Tanzania is lacking in affirming the rights in landholding, especially customary rights of small holders in rural areas (Mugabi, 2013; Massay, 2015). This means that there is the possibility of land expropriation for them to work for the LALI, which risk neglecting their respective productive but labor-displacing ventures.

Second, the presence of LALI can crowd out smallholder farmers (Nolte and Ostermeier, 2017; Sulle, 2017), thereby reducing the income and consumption of these vulnerable groups due to the loss of their main economic activities. To be fair, these LALIs extend the opportunity of contract farming and out-grower schemes to households in the communities (e.g. Herrmann, 2017; Herrmann and Grote, 2015). However, as seen from the analyses, LALIs not significantly increase the income and consumption of the individuals in such communities. This further supports the fact that the current the presence of LALIs and the opportunities that they bring cannot substitute the economic inflow that households generate from their main activities. This implicates that the compensation and economic benefit that are handed out by these LALIs should be closely monitored for fairness and adequacy for the livelihood of these individuals in the communities where they are sited. The submission of Li (2011) lend credence to this finding that poverty reduction in the presence of LALIs is highly unlikely unless vast numbers of jobs are created from these investments and/or a global basic income grant is devised to redistribute the wealth generated in highly productive but labor-displacing ventures.

Third, the negative income and consumption association for female headed households compared to their male counterparts could reflect the social system existence in Tanzania, which disadvantages the female population compared to their male counterparts. LALIs’ presence has not addressed this as seen from the results, instead there seems to be a lower welfare outcome for female individuals compared to the males. Thus, there is the need for policies to better protect female dwellers in rural communities noting that they may be more disadvantaged. More so, the female group engages more in unpaid household works, use the resources for household consumption and childcare; hence, the urgency for them to be protected is further emphasized.

In summary, main policy recommendation that emanates from this study is that the agricultural sector still holds great potential for economic transformation in Tanzania. However, the drive for LALIs may not result in the desired labor outcome, particularly with regards to women. In essence, the effects of LALIs differ markedly between male- and female-headed households, which signal the need for targeted strategies to integrate small farmers into the LALI process. Based on a supplementary analysis of supplementary analysis that was conducted to establish the channel through which LALIs affect female individuals and female household

### Table 9

Estimated Average Treatment Effect for Entire Female Sample.

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>%difference</th>
<th>NNM</th>
<th>%difference</th>
<th>KM</th>
<th>%difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Household</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Consumption</td>
<td>−667.25***</td>
<td>−35.59%</td>
<td>−1007.78***</td>
<td>−46.25%</td>
<td>−1119.25***</td>
<td>−40.30%</td>
</tr>
<tr>
<td>Wages</td>
<td>0.015</td>
<td>0.010</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Time allow</td>
<td>0.241</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>0.073**</td>
<td>0.108</td>
<td>0.060****</td>
<td>2.014%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 10

Double Difference Estimation of the ATT.

Notes and Source: Same as Table 8.

<table>
<thead>
<tr>
<th></th>
<th>Household Consumption</th>
<th>Wages</th>
<th>Time allow</th>
<th>Gender = Female</th>
<th>Household Consumption</th>
<th>Wages</th>
<th>Time allow</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT</td>
<td>−4508.28***</td>
<td>−1987.09</td>
<td>3.035*</td>
<td>−2526.81*</td>
<td>−118376</td>
<td>2.980*</td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>(0.045)</td>
<td>(0.981)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>F-Stat</td>
<td>14.89</td>
<td>0.2679</td>
<td>0.8228</td>
<td>0.1042</td>
<td>0.0463</td>
<td>0.7699</td>
<td></td>
</tr>
<tr>
<td>Prob &gt; F</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
</tbody>
</table>

7 Results are available on request.
heads, it is evident that formal education, land ownership, and women’s age are all important factors that should be considered in Tanzania in relation to the effect of LALIs on women. Our findings suggest that although the presence of LALIs results in negative association with consumption and agricultural wages for female-headed households, these results may not hold if the women are better educated, and/or own land. Thus, policies should consider education and land access as channels through which to improve female outcomes in the case of rising land investors.

6. Conclusion

This study contributes to the understanding of the labor and welfare implications of the presence of large-scale agricultural land investments (LALIs) for female-headed households in Tanzania. We examine household consumption, agricultural income, and labor allocation to agricultural activities for those living in communities with LALIs and those living in communities without LALIs. The analysis was based on ex-post non-experimental analysis that considers the implication of the presence of LALIs after the investments are established and functional. The propensity score matching (PSM) and regression-based approach, including the treatment-effects model, were applied; this combination of different estimations corrects for potential biases.

The study finds that households located in communities with LALIs have lower consumption compared to households located in communities without LALIs. The results suggest that female-headed households are better off living in communities without LALIs. Considering the income effect, our estimation reveals a negative and significant agricultural income (wage) effect for the entire sample; for female-headed households, a non-significant effect was observed. The female-headed households living in communities with LALIs could have a higher agricultural income (between 0.4 percent and 21 percent) compared to those living in communities without LALIs. For the entire sample and the female-headed households, the analysis reveals a positive effect on agricultural labor hour input. These results, however, should be seen only as suggestive evidence of the implications of the presence of LALIs in Tanzanian communities because of some other unobservable factors that are peculiar to the communities being observed and that are not accounted for in our empirical model, which can be investigated in future studies.

Acknowledgments

This paper draws from the research project funded by the African Growth and Development Policy Modelling Consortium (AGRODEP)/International Food Policy Research Institute (IFPRI), under the Innovative Research Grant [2015X192.GIT]. The useful comments from the resource persons are appreciated. Earlier drafts of the paper have been presented at: PEGNet Conference (October 2015, Berlin, Germany); IFPRI Workshop on Impact Evaluation (March 2016, Washington DC, USA) and 17th Annual World Bank Conference on Land and Poverty (March 2016, Washington DC, USA). Thus, support from the organizers and comments from discussants and participants are appreciated. The first author acknowledges the Fellowship Award from Alexander von Humboldt Foundation (AvH), during which period the research project was carried out [3.4-RKS-NRI/1147508]. The assistance received during the fieldwork from individuals and organizations (too numerous to itemize), are acknowledged. This research also benefited from the collaboration between Covenant University, Ota, Nigeria and German Development Institute, Bonn, Germany. We are grateful for the helpful comments from the anonymous reviewers. The views expressed are those of the authors.

References


E.S. Osabuohien et al.


