Local Economy effects of Large-Scale Agricultural Investments

Paul Hofman[a], Esther Mokuwa[b], Paul Richards[b], Maarten Voors[a,*]

[a] *Wageningen University, Hollandseweg 1, 6706KN Wageningen, The Netherlands*

[b] *Njala University, Sierra Leone*

Abstract

The last decade has seen a surge in land acquisitions in developing countries by foreign companies. To date there has been little rigorous quantitative evidence on the impacts of such investments on local communities. We examine the economic impacts of a large-scale biofuel plantation in Sierra Leone - a major investor target. We conduct a difference in difference analysis using three waves of a large n survey in both communities directly affected by the plantation and those outside the catchment area. We find a large average drop in incomes, mainly driven by lower revenues from agricultural activities. These findings are consistent with a labor demand shock, caused by a clash between the private and commercial agricultural calendar, increasing the local price of labor. A spillover analysis confirms that the impacts are at least partially transmitted by a shock to the local economy. Within land leasing communities, households that are employed at the plantation see their incomes and assets increase. However, as a result, village-level inequality increases.

Keywords: Large-Scale agricultural investments, Inequality, Land Grabs, Labour Market
JEL: F230, J230, O130, O150, Q150

*Corresponding author. Declaration of Conflicting Interests: The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Email addresses: paul.hofman@wur.nl (Paul Hofman), esther_mokuwa_richards@yahoo.com (Esther Mokuwa), paul.richards1945@gmail.com (Paul Richards), maarten.voors@wur.nl (Maarten Voors)

1. Introduction

Foreign investments in African agriculture have increased dramatically. Driven by the 2007-8 price spike of key primary commodities in conjunction with the world financial crisis, commercial investment companies increasingly sought out new investment ventures Arezki et al. (2013); Koning and van Ittersum (2009). The Land Matrix, which documents all transnational land acquisitions, to date has recorded 1694 'concluded' agricultural investments, in total covering about 50 million hectares[1]. In some African countries over 30 percent of arable land is foreign-owned (Landmatrix, 2020; Nolte et al., 2016). These investments often take the form of large scale plantations, with land rights acquired for a long period (typically 49 or 99 years). This trend is likely to increase due to the projected rise in demand for food, animal fodder and energy crops.

Some herald this new wave of investment by commercial parties as an important vehicle to achieve poverty reduction, highlighting the potential benefits of scale economies in agricultural production (Collier and Dercon, 2014; Ellis, 2005), inducing innovation (Borensztein et al., 1998), enabling access to finance (Alfaro et al., 2010) and the organization of production and marketing (Reardon et al., 2003). On the other hand, there are arguments against land consolidation that stress important potential negative impacts on distributional, social and institutional outcomes. First, while large scale investments may create new opportunities for some (through land rents and employment), they exclude others (Peters, 2004). Such effects may be particularly strong in the African context characterized by strong social dependencies (Townsend, 1994). Investments may deepen social divisions, possibly contributing to conflict (Peters, 2013; De Schutter, 2011; Baxter, 2013; Scott, 1998). Second, large-scale land acquisition by foreign companies often amounts to "land grabbing" (Liversage, 2010), generating benefits for foreign investors (and domestic elites). Land rights are impacted as investors obtain leases and clear land for industrial monoculture plantations. For many households this implies a change in access to land (in extreme case even forced migration), and nutritional security, thereby impacting family livelihoods (Liversage, 2010). Recent work shows that foreign investments are greater where property rights regimes are weakest (Alfaro et al., 2010; Arezki et al., 2013). This suggests an important role for institutions as a mediating factor in determining potential development outcomes (Sokoloff and Engerman, 2000; Herbst, 2014; Dorward et al., 2009). Often, land investment deals are made between companies and elites and exclude local people from the negotiations, increasing corruption (Peters, 2013; De

[1] Date of access: November 2019

Schutter, 2011).

Despite the scale of foreign investments in agriculture, local economic impacts have to date failed to receive rigorous quantitative investigation. Exceptions are Herrmann and Grote (2015), who assess a sugarcane plantation and outgrower scheme in Malawi, and find positive economic returns for laborers. The plantation attracts labor from nearby villages typically from the poorest households. Compared to non-laborers, incomes nearly double. A similar paper, by Herrmann (2017) examines rice and sugar plantations in Tanzania. He finds for both sectors an increase in per capita income for plantation laborers compared to other households in the same villages. There is however, no significant effect on agricultural or total household income.[2] A key limitation of these papers is that they rely on post-intervention data, requiring strong assumptions to assess causal effects. Investors typically do not select concession sites at random and take important ecological, political and economic characteristics into account such as agricultural potential, distance to input and output markets, local institutions and labor availability. Failing to adequately control for such variables may severely bias results. Below, we improve on this work and use data from before and after the creation of a large scale agricultural plantation. Baseline data, pre-dating the plantation allows us to control for such selection effects. In addition, the analysis compares those hired by the plantation to those that are not hired but are from the same village. This incorporates household economic impacts through community wide channels such as increased competition over land and labor. It is an open question whether incomes should increase on average in the local village economy. Theoretical work by Kleemann and Thiele (2015) and Dessy et al. (2012) show how the net effects of such investment projects crucially depends on the intermediate impacts on labor and land markets. If labor and land are abundant, increased demand for labor and land should not impact local economies. However, this is rarely the case. For instance in rural Sierra Leone, the country we consider here, where there is severe competition over labor (Mokuwa et al., 2011; Bulte et al., 2018). In such cases increased employment opportunities outside the village may cause a decrease in labor input for private farms, undermining income and food security.

We examine the impact of a large scale agricultural sugarcane investment project in Sierra Leone. The country is an appropriate choice for investigating the impact of foreign agricultural investments. Sierra Leone is a poor country, characterized by rotational fallow agriculture and limited access to financial and output markets. The majority of the population is engaged in the

[2]Other papers include Shete and Rutten (2015) and Jiao et al. (2015). Both use a matching algorithm on post-intervention data, and suffer from a low number of observations/clusters.

agricultural sector. Farms are very small: average farm size is about 0.5 hectares. To a large degree farm output is determined by labor rather than land or capital (fertilizer application and improved seed varieties are rare) (MAFFS, 2011). There has been a surge in commercial investments in agriculture. Since 2000, foreign companies have acquired over 25% of the country's arable land (Baxter, 2013; Landmatrix, 2020).

We use a difference-in-difference approach allowing us to correct for important time-invariant characteristics, such as agricultural potential, distance to input and output markets, local institutions and labor availability, all of which are crucial selection criteria for investors. We assess impacts on several key outcomes: household income (stock and flow), access to land, food security, health and village level inequality. Our data allow us to examine effects over the shorter (2 years) and longer (5 years) periods. We find that average income drops substantially, by about 0.4 standard deviations. We also see a small drop in access to land and some improvements in health outcomes in villages where the company works. We find that the labor demand shock, caused by a clash between the private and commercial agricultural calendar, increases the local price of labor. As a result, average farm productivity and agricultural incomes decrease. In contrast, households that have a member working for the company compensate for this drop with salaried income. As a result village inequality increases. The hypothesis that the main impacts work through local markets is bolstered by a spillover analysis that shows income changes are smaller further away from the investment. As a robustness check, we provide some evidence that the parallel trends assumption holds. We also examine attrition and find that our main findings are robust when examining bounds on the treatment effect.

This fits in a larger literature that aims to move from examining small-scale impacts towards looking a the effects on the local economy (Taylor and Filipski, 2014; Cust and Poelhekke, 2015). One example of this is Aragón and Rud (2013) who examine the impacts of an exogenous expansion of a gold mine in Peru on the local economy. They find that this expansion increases local labor prices and local income, and this effect declines when moving away from the mine. We find similar results: in our case reductions in labor availability reduce household production.

Closest to our work is Bottazzi et al. (2018). They use a matching algorithm to match 592 respondents in 34 villages where the company leased land and compare this to 290 respondents in 21 control villages. They find that on average incomes and food expenditure increase, as well as labor prices. They also see improvements in food and water security. They also note that positive economic effects are mainly for landowners and men that are employed. While we examine the same investment and a similar period, we find an opposite impact for incomes: we find a large and substantial drop. This is likely because our identification strategy allows us to correct for pre-existing differences: we note a strong imbalance in pre-investment data in incomes, villages that end up leasing land to the company are on average richer than comparison villages. As a result, the positive economic results Bottazzi et al. (2018) find may be due different initial conditions, rather than due to the impact of the investment project.

We go further than this work by including pre-investment data, greatly improving the identification strategy and using a substantially larger sample, improving statistical power.

This rest of this article is structured as follows. Section 2 introduces the research context. Section 3 presents our data and empirical strategy and section 4 contains our results. We present some robustness analyses in section 5 and conclude in section 6.

2. Large Scale Investments in Agriculture in Sierra Leone

We focus on Sierra Leone, which has received a lot of attention from land investors. Since 2000, 24 deals have been concluded, covering 1 million hectares (25% of total arable land) (Landmatrix, 2020). The country ranks low on the Human Development Index (UNDP, 2016) and has high poverty levels and low food security. Most Sierra Leoneans are smallholder farmers, especially in rural areas. Farm productivity is low and access to productive inputs, such as fertilizer and high-yielding seeds is minimal. As a result, agricultural production is limited by labor availability. A 2011 survey found that 65% of households experienced a shortage of labor in the agricultural season. Farm production to a large extent relies on family labor. About one third of households hire labor (MAFFS, 2011).

Outside investments can potentially improve this situation by bringing in improved technologies and large-scale production that achieve economies of scale. The Government of Sierra Leone aspires to ‘promote an attractive business environment based on fair and responsible investments in land for both small and large scale businesses’ (GoSL, 2015, pp 7).

We assess the impacts of a large-scale plantation in the north of Sierra Leone. In 2010, a commercial investor acquired 24’000 hectares of land for a 49 year lease. Landowners received 8.90 US$ in compensation per hectare per year, half of which (i.e. 4.45 US$) goes to the landowners and the other half to various local elites. This is according to national standards for land payments. Landowners also receive an additional payment of 3.46 US$, making the total payment for landowners 7.91 US$/Ha/year. In 2014, a peak year the company leased land from 52 villages, amounting to 10-60% of total village land. The investor employed local and international staff to grow sugarcane using center-pivot irrigation. In 2014, the company employed 3500 people, half of whom were on fixed-term contracts. The company aimed to recruit unskilled labor from communities in and around the plantation. The main labor demand of the plantation overlaps with peak periods for smallholder production. Smallholder labor demand is greatest in February-April when land is prepared ('cleared') and planted according to a rotational cycle (Richards, 1986), matching the plantation's peak labor demand. Besides providing benefits in terms of employment for laborers from nearby villages, and surface rents for land owners, the company established a health clinic, provided several farming training programs and had a compensation program for destroyed tree crops. The investment was funded by a consortium of ten Western development banks. This means that besides a business project it was also explicitly aimed to be a development project.

The plantation has received considerable attention in the media and has been the focus of several policy reports and journal articles. We provide a summary in Table 1. Most reports critique the investment and describe how it was forced through by politicians and local elites without involving communities other than through superficial consultation, and conclude average incomes decreased. Some cite improved incomes, especially for specific landowners. Some also point to increases in social disharmony due to the plantation creating conflicts over access to land and surface rents and in other cases exacerbating existing tensions over land claims. A key drawback is that most of these studies rely on qualitative and small sample case studies. While these sometimes provide rich insights into relevant local dynamics for selected localities, they fall short in assessing average differences.

3. Data and Empirical Strategy

3.1. Sample

We use data from multiple rounds survey work. Table 2 summarizes the sample sizes for each round. Baseline data were collected prior to any plantation activities in 2010 by a research team of the University of Cape Town, at company request, to comply with reporting requirements. In total, baseline data encompass 78 villages and 4233 households, comprising a census of all households in these villages. The plantation then started operations in 41 of these villages, creating a natural comparison group. In 2012, a second survey was implemented (again by the University of Cape Town), this time in 118 villages and under 4824 households. In the meantime, the company scaled up operations to 47 villages. Figure 1 shows the locations of all villages where we have access to panel observations for the 2010-2012 survey waves. In 2015, a team from Njala university collected an additional wave of data. Enumerators received extensive training on the survey instrument. We returned to all villages included in the 2010 dataset and interviewed 25 people per village. We spoke to all people interviewed in 2010 and still present. The survey instrument was designed to closely match the earlier rounds of data collection. For the 2015 survey round we have data on 1767 people in 75 villages. In the meantime the company relinquished land from some villages ending with 36 villages from the original pool.

To examine how people are affected over time we ideally rely on data from the same people across each survey wave. Fortunately, the company assigned all households ID codes and identification cards. We used these ID codes to match respondents across waves. In total we have 3155 respondents in both 2010 and 2012, and 628 observations for both 2010 and 2015. Below we estimate model 1 both using the 2010 and 2012 data to assess short run effects and compare across 2010 and 2015 for longer run effects. [3]

3.2. Identification Strategy

Our identification strategy relies on a difference-in-difference approach. We estimate the differences in outcomes over time for both the villages that rented land to the plantation and control locations. This corrects for all time invariant characteristics (observable or not). The main identifying assumption is that in the absence of investment, the villages would have developed in a similar pattern. This assumption is of course fundamentally untestable. However, using data on forest loss and vegetation (EVI) available from satellite images, we can show that deforestation trends were parallel before the investment started. The control group are a set of villages that the company was originally planning to work in but decided not to. This was for various reasons: villages decided not to join, the villages could not provide enough land and most importantly the

[3] If we are more stringent and also match on village name, participant name, years in area and GPS location the number of matched participants drops. In this study we use the match on ID codes, though as a robustness we examine whether the direction of coefficients holds for the more restrictive match. These results are shown in Table A3 and Table A4 and are qualitatively similar to our main results.

distance to the Rokel river (darker in Figure 1) was too great to pump water for the center pivots used by the company for irrigation . Therefore, they are similar in characteristics that are likely to be predictive of yield. Furthermore, since all smallholder agriculture is rain-fed distance to the Rokel river is unlikely to correlate with local agricultural production.

3.3. Empirical Model

To assess impacts of this investment we estimate the average treatment effect on the treated for original households using a standard difference-in-difference specification. Specifically, we estimate:

$$\boldsymbol{Y}_{ij} = \beta_0 + \beta_1 treat_j + \beta_2 post_{ij} + \beta_3 post_{ij} * treat_j + \varepsilon_{ij} \tag{1}$$

Where $\boldsymbol{Y}_{ij}$ refers to our set of outcome variables (such as income, land access, see section 3.4), $treat_j$ refers to the villages where the company leased land and $post_{ij}$ refers to the later time period. β_3 is our coefficient of interest. $_i$ indexes the household level, while $_j$ indexes the village level. We cluster standard errors at the village level.

Furthermore, as a plausibility check to see if labor shortages are driving our results, we examine if our outcomes taper off further away from the plantation. We estimate:

$$\boldsymbol{Y}_{ij} = \gamma_0 + \gamma_1 distance_j + \gamma_2 post_{ij} + \gamma_3 post_{ij} * distance_j + \varepsilon_{ij} \tag{2}$$

γ_3 is our coefficient of interest and we again cluster standard errors at the village level.

Finally, we assess if individuals employed by the company benefit. In the 2015 survey we asked respondents if the had worked for the plantation. We examine the extensive margin and regress our main outcome variables on a dummy indicating if a household member at any time worked for the plantation during the 2010-2015 period. We then estimate a triple differences model:

$$\begin{gathered} \boldsymbol{Y}_{ij} = \eta_0 + \eta_1 laborer_{ij} + \eta_2 treat_j + \eta_3 post_{ij} \\ + \eta_4 post_{ij} * treat_j + \eta_5 treat_j * post_{ij} * laborer_{ij} + \varepsilon_{ij} \end{gathered} \tag{3}$$

Our coefficient of interest is η_5, how laborers differ from non-laborers in treatment villages in the later time period.

3.4. Outcome variables

Our main outcome indicators relate to incomes, land access, food security and health. Our variables are defined in Table A1 and descriptive statistics at baseline for both treatment and control villages are shown in Table 3. Average household monthly income is 60'000 Leones (200'000 in Treated), or 13 USD (36 USD), far below the World Bank international poverty line of 1.25 USD per day [4]. This measure includes only cash incomes and does not account for self-consumption or in-kind contributions. Figure 2 shows the relative components of traditional income[5]. Agricultural income accounts for the majority of income, with 60% for the control group and 80% for the treatment group before treatment. The income differences between treatment and control villages are large. Given our difference-in-difference set up, these drop out. The number of assets in a list of what farmers owned is 4, which might mean a household owned its house, a mosquito net, an iron pot and a bed mattress, but no mobile phone, tv, iron kettle or generator. Housing quality averages 5, which is the rating for a house with a mud floor, wattle and daub walls and thatch or tarpaulin as roof. For the livestock index the value is around 0.25, comprising (for example) 2 goats and 5 chickens. Almost all households have access to arable land for cultivation, though almost all households have faced a seasonal food shortage in the previous year. Half of households had a birth in the previous year (one third in treatment villages). 92% of households had a mosquito net in their house (80% in control villages). The participants are clearly very poor, have few assets and low food security. [6][7]

4. Results

We first estimate model 1, to assess the short-run effects of the large-scale agricultural investment. Table 4 presents the results. Our main variable of interest is the interaction term, which shows the effect of the treatment over time, correcting for initial differences in levels. This shows a big drop in traditional income of over 0.6 standard deviation. For total income this is lower (0.4 SD) but

[4] As income is highly sensitive to outliers we use the inverse hyperbolic sine transformation to correct for this. These numbers are calculated back from the inverse hyperbolic sine transformation.

[5] We split our income into two measures: traditional income and total income. Total income also includes all kinds of payments by the company. See Table A1 for the definition

[6] We have a very low number of observations for child births and deaths in control villages. We have investigated this but cannot find a structural reason for it. There may be a reporting bias, due to diffidence in discussing such events, sometimes suspected to be the result of witchcraft. Any results regarding child deaths and births should be interpreted with this in mind.

[7] We use the following user-written computer programs in preparation of the data, tables and figures: Jann (2005, 2007, 2012, 2016); Van Kerm (2009); Gallup (2012); R Core Team (2017); Hijmans (2017); Wickham et al. (2017); Højsgaard and Halekoh (2018); Müller and Wickham (2018); Gorelick et al. (2017)

still substantial and significant. This drop is largely driven by a large drop in agricultural income (See Table A5 and Table A6 for the effect on the four components of traditional income). We hypothesize that this is caused by an increase in the local labor price which makes it more difficult for households to hire in local labor, reducing agricultural production and thus sales. Our spillover analysis (see Table 6) and laborer analysis (see Table 7) confirm this hypothesis. Furthermore, in 2015 we asked households whether the price of labor had gone up after the company started working. 87% of farmers said that it did. The drop in income (a flow variable) partially translates to a change in stock variables (ie assets). There is a substantial drop (0.11 SD) in housing quality. On the other hand, the TLU score increases by 0.3SD, though neither of these are significant. Access to land goes down 5% more than in the control group which is small but precisely estimated. The often-used narrative that these investments are utilizing unused land and thus not affecting land availability of productive assets does not hold here. When we examine this group that has lost access to land separately using a similar model as 3 we find a large decrease in full income of 0.8SD. Incidence of food shortages drops by 10% in the short run.

Next, we look at three measures of health. We see a large increase in the number of total births, of 0.5 SD. Infant deaths also increase, though by less (0.18 SD) and insignificantly. This might mean that the availability of a local health clinic is increasing the number of total births, without a similarly large increase in deaths. We have already noted that possibly households under-report infant or per-natal mortality. Treated households had a lower rate of bed nets before the investment, and in treated households this has gone up in the short run. This could again be linked to health outreach programs ran by the company.

We dig a little deeper into the drop in income by examining how the proportions of income evolve over time. This is shown in Figure 2. Before treatment the treated group relied more on agricultural income, accounting for almost 80% of total income. In the mid-term data set (2012) this had dropped to around 55%. There is also a drop for the control group, though this is much smaller. This is largely driving the income effect we find. However, it is possible that 2010 was a better than average year for agriculture. Since the treated group relies on agriculture more, they would be more affected when returning to normal harvest levels. Figure 6 suggests, however, that 2010 was a normal year for agriculture.

Next we estimate Model 1 for the longer (5-year) period, shown in Table 5. Generally, results are similar in the longer run, though a much lower number of observations makes our estimates noisier. In the long run there is again a substantial drop in income in treated villages, again signifying a negative income effect of the plantation. Looking at the stock variables there are no significant differences. House quality now has a positive coefficient on the interaction term (opposite to before), but this is not significant. Access to land remains lower in treated villages, and the effect is now slightly larger (7% lower). We again see no effect on food security. For health we still see a substantial increase in number of total births, though it is only marginally significant. We again see higher presence of bed nets, but this is not significant.

The main effects we find for both the longer and shorter run analyses are a drop in income, lower access to land and some health improvements. For the latter two the link to the plantation is clear: the company is using village land and is providing some health services. In terms of the income effect we have hypothesized that this is caused by an increase in the labor price. Whether there is such a local effect can be tested, which we do next.

Within our control group there is substantial variation in distance to the plantation (defined as distance to the closest treated village). The mean is 3.5 km with standard deviation 2.6 km. We can exploit this variation by repeating our previous analysis, but now taking distance to the plantation as the treatment variable and examining only control villages, as in Model 2. The results of this analysis are in Table 6. For both measures of income, we see that before the investment, places further away from the plantation had lower average incomes. The interaction shows that control villages further away increased their income more than control villages closer to the plantation. Being 1 SD further away from the investment results in a 0.39 SD higher full income. If higher labor prices are indeed locally determined and spill over partially to neighboring villages we would expect to find these results. For assets we see an increase in the number of assets further away over time, which also holds for house quality (though the effect is much smaller and only marginally significant). Access to land is higher further away from the investment, but it is not very high (1 SD distance leads to 2% higher access) and only significant at the 10% level. This is unsurprising as the treatment is defined by having land leased. That there is some small effect might indicate that households start farming in neighboring villages. There is again no effect on food security, and the effect for child births is negative, which makes sense as being further away from the investment also increases travel time to the health clinic, making it more costly to use. Overall, these results provide evidence that there are some local market effects (or spillovers) which are driving the effects we found in Tables 4 and 5. This could mean that our results in those tables are biased. However, we consistently find that the treatment effect is weaker the further away from the investment. This means that our main estimates are a lower bound of the actual effect.

So far, we have been examining these effects across all households in a village. Next, we examine the effect separately for laborers and non-laborers. We lack information on company employment for 2012, thus limiting our analysis comparing over the 2010 - 2015 period. In treated villages, about 40% of the households supplied labor to the plantation, with an average length of employment of 14 months. We examine effects on laborer households in Table 7, using Model 3. With respect to traditional income, ie excluding wage earnings, we see a substantial drop of 0.33 SD. However, this is not the case when examining full income (which includes salaries). When comparing laborers to the other households in the village, there is an increase in income of 0.34 SD. When we move to the stock variables there is a consistent increase of about 0.2 SD, for number of assets, housing quality and livestock. Clearly, laborers were able to transform their additional earnings into tangible assets. Furthermore, laborers do not have lower access to land or better health access compared to others in their village. This is unsurprising, as these are effects enjoyed by all households (the health clinics are accessible to all villagers). There is one small effect, in that laborers are slightly more likely (4%) to have had food shortages in the previous year, though it is not very significant. It might be that laborers now working on the plantation are not producing their own food anymore, causing some domestic shortages. This also implies that there is not enough food available on the local food market.

Finally, as a logical consequence, we can assess whether the investment has affected within-village inequality. In Figure 3, we draw lorenz curves for both traditional and full income for the treated and untreated group separately for both 2010 and 2012. Panels a and b show the results for traditional income. We see that in 2010 the curve for the treated group is closer to the line of unity, indicating higher equality. After the company has started work this is reversed, shown in panel b, suggesting that inequality has increased for traditional sources of income. When we add company payments (panels c and d) this effect is weakened. Figure A1 shows the same analysis for the long run, with qualitatively similar results. We present a more formal analysis in Table A2 where we analyse village level gini coefficient. For both sort and long run, the interaction term is positive, albeit larger for traditional sources of income.

5. Robustness

The previous section showed evidence that this large-scale agricultural investment has had strong effects on local incomes, access to land, health and inequality. In this section we provide a test for our main identifying assumption of parallel trends. Furthermore, we ask whether attrition is systematic, examine some bounds on the treatment effect under alternative attrition assumptions and we provide some evidence that agricultural conditions are similar in treated and control villages.

We first examine the parallel trends assumption. This is fundamentally untestable, but we gain some reassurance from showing that pre-treatment trends are parallel - it is likely that they would be parallel afterwards as well. We do this by examining changes in forest loss for treated and control villages. Agricultural production in Sierra Leone is closely linked to forest loss: most agriculture is rotational bush fallowing, a highly labor-intensive form of production. Under rotational fallowing, forest loss is likely to correlate with increased agricultural production and income in the shorter term - one of our main outcome variables. We examine whether trends in forest loss are parallel using forest loss data from Hansen et al. (2013). Their worldwide dataset contains extremely detailed (30m resolution) data on forest loss for the period 2000-2018. To assess whether forest is lost in a specific year we draw circles with a 1km radius around each village, and then count the number of pixels that were lost in the circle for that village in a certain year (see Figure 4). We convert these pixels to the number of hectares lost per village per year and plot this out in Figure 5. The vertical black line represents the year that the company started their activities. Trends are clearly very similar across treatment and control pre-2010, but diverge after 2010. For most years the amount of forest loss is significantly higher in treated villages, and the amount of forest lost is always higher than in control villages. Pre-2010 they were almost always equal. We surmise that the divergence after 2010 is partially caused by activity by the company, and partially by farmers having to move to new plots after leasing their land to the company. This figure provides some evidence that pre-treatment trends are parallel, and also that examining forest losses in this context is a relevant variable.

Next, we examine attrition. Attrition in the short run (2 years) is somewhat high at 25% (35% in the control group and 20% in the treated group). For the long run comparison it is much higher: 85% (93% for the control group, and 81% for the treated group). This is not surprising as the long run data collection was on a random subsample so it was never the intention to find everyone. In table A7 we examine what pre-treatment variables determine attrition, and crucially, whether this differs between the treatment and control group. We see some differences in dropout in the short run (the non-interacting variables), but these are not worrying as this does not indicate differential dropout. There is one worrying finding: a higher traditional income before treatment leads to a lower chance of dropping out in the short run – but in the treatment group only. This could mean that richer households are overrepresented in the short run in treated villages. This means that the negative effect we find is a lower bound of the actual effect. For the long-run dropout we find no significant predictors that differ between the treated and control group.

To further dig into the effect of attrition on the impacts we find we employ a bounds analysis as suggested by Manski (1990) using the approach by Blattman et al. (2014). For this analysis we make alternative assumptions about those who leave the sample. Values for missing observations are filled in to zero out the treatment effect we find. By doing so we can calculate lower bounds for our treatment effects. We examine four bounds, the Manski worst possible bound, and 3 deviations from the mean. In case of a negative treatment effect, control drop-outs are assigned a low value, while treatment drop-outs are assigned a high value, thus zeroing out the negative treatment effect. For the Manski worst case the high (low) value is that group's maximum (minimum) value. For the SD deviations the high (low) value is the group mean plus (minus) X SD, with X being 0.5, 0.25 and 0.1. These results are shown in table A8 for the short run only. We only show continuous variables as the SD adjustments do not make sense when examining dummies. Column 1 show the original treatment effect of the treat*post interaction as in Table 4. Columns 2-5 show the bounds on this effect. For column 2, the Manski worst possible bound, the treatment effect is opposite to our original effect and highly significant for all outcome variables. This is unsurprising when attrition is high (Blattman et al. (2014) find this also) and is shown here for completeness. When we examine our main effect on income the results from columns 3-5 are reassuring. In most cases the sign of all coefficients are the same, and for the 0.1 and 0.25 SD deviation these effects are significant as well. Note that these deviations represent large, systematic deviations on the characteristics of drop-outs for which we found no evidence in Table A7.

Finally, since our identification strategy partially relies on the distance to a large river, there might be differences in agricultural suitability of the available farmland. We explore this by examining the EVI (Enhanced vegetation index), which is a measure of live green vegetation based on satellite imagery. The EVI ranges from -1 (water bodies) to 0 (desert) to 1 (mature forest). It is often used to examine fertility/crop success in a certain year. We plot trends in EVI in Figure 6 and examine the maximum yearly EVI in the same circles around villages with a 1km radius. By using the maximum we automatically filter out clouds and looking for the maximum within a year means we examine the entire agricultural season. We coarsen the pixels from 30x30m to 150x150m to reduce spatial autocorrelation. The trends are extremely similar. This shows that treated and control villages are subject to very similar agricultural conditions. Interesting to note is that there is no difference between treatment and control after 2010, the company's start year. It appears that while the company did contribute to significant forest loss, live green vegetation was unaffected.

6. Conclusion

This paper is one of the first to provide empirical evidence for the impact of large-scale agricultural investments. This allows us to examine how rural communities respond to land and labor shocks. While there might be positive effects (higher incomes, better infrastructure and access to new farming technologies), most research so far has pointed to negative effects: loss of land, increased marginalization and exploitation by powerful foreign companies (Baxter, 2013; De Schutter, 2011; Liversage, 2010). Our case is a large-scale agricultural investment in Sierra Leone, a country which has received a lot of interest from investors in land during the past decade, consequent on an opening to international capital following decade of civil war. A for-profit company leases 24'000 hectares of land and uses this to grow sugarcane for biofuel. The company pays landowners yearly for the land and employs local labor on the farm.

We use a difference-in-difference analysis to compare outcomes for communities within and outside the catchment area of the plantation investment. We find a large drop in average incomes for treated communities, almost half a standard deviation compared to the control group at baseline. This is mainly driven by lower agricultural income. We surmise that this is because the increased labor demand increases labor price, making it too expensive to hire in labor, the most important factor of production. A spillover analysis confirms this. We see mixed effects on physical assets. It might be that households are holding on to (some) of their assets to weather future shocks. We also see a drop in access to land, which runs counter to the argument that land is plentiful and not a relevant constraint. Lower access to land also likely contributes to lower agricultural income. Food security is largely unaffected, surprisingly. This suggests that the cause of seasonal hunger is related to storage issues and local market failures, We also see some improvements in health, which plausibly can be attributed to company health program. These effects hold in the longer run (5 years) as well. When we examine company laborers specifically, we see that they benefit relative to non-laborers in their village. Their incomes rise and this translates into more tangible assets.

We have hypothesized and given evidence that a portion of the impact is transmitted through local markets, especially the labor market. This is likely to hold for most external agrarian investors; by definition they are looking to acquire land, and often seek to hire local labor to lower transportation costs and to obtain goodwill from the local community. This shows that to examine the full impact of one of these kinds of investment the full village economy should be examined. To improve this, local economy models as in Taylor and Filipski (2014) should be developed to gain more insight into the functioning of local markets and social welfare institutions. We leave this for future work.

Taken together, the results from this paper paint a bleak picture. While an increase in income for laborers is positive, only 40% of households provide laborers, and their gains do not outweigh the losses by households without income from plantation laboring. It is possible that this increased inequality and marginalization increases social conflict, a scenario suggested in previous work (a qualitative assessment of gender-based conflict is by one of the present author's is ongoing). The investment as a whole appears to be a poor deal for recipient communities. Indeed, that over one third of land lease payments go to political elites rather than the land owners themselves is a warning sign that local benefit is not the priority in these kinds of investments.[8][9]

[8] Acknowledgements: We thank conference participants at LandAC, EUDN, World Bank Land and Poverty Conference and numerous seminar participants. We are indebted to our respondents and grateful to our team of excellent enumerators for data collection. Thanks to Sanne Schimmel, Afke Jager and Mirella Schrijvers for research assistance.

[9] Funding: This work was supported by NWO Global Challenges Program grant # W08.250.104 and NWO grant # 451-14-001. Data collection for the endline survey was funded by Addax Biofuel (SL) Ltd under a contract guaranteeing the researchers' independence in carrying out the research and publishing the findings.

Alfaro, L., Chanda, A., Kalemli-Ozcan, S., Sayek, S., 2010. Does foreign direct investment promote growth? Exploring the role of financial markets on linkages. Journal of Development Economics 91 (2), 242–256.

Anane, M., Abiwu, C. Y., 2011. Independent Study Report of the ADDAX Bioenergy Sugarcane-to-Ethanol Project in the Makeni Region in Sierra Leone. Tech. rep., Sierra Leone Network on the Right to Food, Freetown.
URL https://www.brot-fuer-die-welt.de/fileadmin/mediapool/2{_}Downloads/Fachinformationen/Aktuell/aktuell{_}28{_}english.pdf

Aragón, F. M., Rud, J. P., 2013. Natural resources and local communities: Evidence from a peruvian gold mine. American Economic Journal: Economic Policy 5 (2), 1–25.

Arezki, R., Deininger, K., Selod, H., 2013. What drives the global "land rush"? The World Bank Economic Review 29 (2), 207–233.

Baxter, J., 2011. Understanding Land Investment Deals in Africa - Country Report: Sierra Leone. Tech. rep., The Oakland Institute, Oakland.

Baxter, J., 2013. Who is benefitting? Tech. rep., Action for Large-Scale Land Acquisition Transparency, Freetown.

Blattman, C., Fiala, N., Martinez, S., 2014. Generating skilled self-employment in developing countries: Experimental evidence from Uganda. The Quarterly Journal of Economics 129 (2), 697–752.

Borensztein, E., De Gregorio, J., Lee, J.-W., 1998. How does foreign direct investment affect economic growth? Journal of international Economics 45 (1), 115–135.

Bottazzi, P., Crespo, D., Bangura, L. O., Rist, S., 2018. Evaluating the livelihood impacts of a large-scale agricultural investment: Lessons from the case of a biofuel production company in northern Sierra Leone. Land Use Policy 73 (December), 128–137.
URL https://doi.org/10.1016/j.landusepol.2017.12.016

Bottazzi, P., Goguen, A., Rist, S., 2016. Conflicts of customary land tenure in rural Africa: is large-scale land acquisition a driver of 'institutional innovation'? Journal of Peasant Studies 43 (5), 971–988.

Bulte, E. H., Richards, P., Voors, M., 2018. Institutions and Agrarian Development: A New Approach to West Africa. Springer.

Collier, P., Dercon, S., 2014. African Agriculture in 50 Years: Smallholders in a Rapidly Changing World? World development 63, 92–101.

Cust, J., Poelhekke, S., 2015. The local economic impacts of natural resource extraction. Annu. Rev. Resour. Econ. 7 (1), 251–268.

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

Dessy, S., Gohou, G., Vencatachellum, D., 2012. Foreign Direct Investments in Africa's Farmlands: Threat or Opportunity for Local Populations?

Dorward, A. R., Kirsten, J. F., Omamo, S. W., Poulton, C., Vink, N., 2009. Institutions and the agricultural development challenge in Africa. Institutional economics perspectives on African agricultural development 1, 3–34.

Ellis, F., 2005. Small farms, livelihood diversification, and rural-urban transitions: Strategic issues in Sub-Saharan Africa. In: The future of small farms. Research Workshop Proceedings. pp. 135–149.

Fielding, M., Davis, M., Weitz, N., Cummings-John, I., Hickey, A., Johnson, F. X., Senyagwa, J., Martinez, L., Sun, M., 2015. Agricultural investment and rural transformation: a case study of the Makeni bioenergy project in Sierra Leone. Tech. rep., Stockholm Environment Institute, Stockholm.
URL `http://www.sei-international.org/mediamanager/documents/Publications/Climate/SEI-PR-2015-09-Makeni-Project.pdf`

Gallup, J. L., 2012. A new system for formatting estimation tables. Stata Journal 12 (1), 3.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment.
URL `https://doi.org/10.1016/j.rse.2017.06.031`

GoSL, 2015. National land policy. Tech. rep., Government of Sierra Leone, Freetown.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Others, 2013. High-resolution global maps of 21st-century forest cover change. science 342 (6160), 850–853.

Herbst, J., 2014. States and power in Africa: Comparative lessons in authority and control. Princeton University Press.

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

Herrmann, R., Grote, U., 2015. Large-scale Agro-Industrial Investments and Rural Poverty: Evidence from Sugarcane in Malawi. Journal of African Economies, 645—-676.

Hijmans, R. J., 2017. raster: Geographic Data Analysis and Modeling.
URL https://cran.r-project.org/package=raster

Højsgaard, S., Halekoh, U., 2018. doBy: Groupwise Statistics, LSmeans, Linear Contrasts, Utilities.
URL https://cran.r-project.org/package=doBy

Jahnke, H. E., 1982. Livestock production systems and livestock development in tropical Africa. Kieler Wissenschaftsverlag Vauk, Kiel.

Jann, B., 2005. Making regression tables from stored estimates. The Stata Journal 5 (3), 288–308.

Jann, B., 2007. Making regression tables simplified. The Stata Journal 7 (2), 227–244.

Jann, B., 2012. ESTWRITE: Stata module to store estimation results on disk.
URL https://econpapers.repec.org/RePEc:boc:bocode:s450201

Jann, B., 2016. Estimating Lorenz and concentration curves. Stata Journal 16 (4), 837–866(30).
URL http://www.stata-journal.com/article.html?article=st0457

Jiao, X., Smith-hall, C., Theilade, I., 2015. Land Use Policy Rural household incomes and land grabbing in Cambodia. Land Use Policy 48, 317–328.
URL http://dx.doi.org/10.1016/j.landusepol.2015.06.008

Kleemann, L., Thiele, R., 2015. Rural welfare implications of large-scale land acquisitions in Africa: A theoretical framework. Economic Modelling 51, 269–279.

Koning, N., van Ittersum, M. K., 2009. Will the world have enough to eat? Current Opinion in Environmental Sustainability 1 (1), 77–82.

Landmatrix, 2020. The Online Public Database on Land Deals.

Liversage, H., 2010. Responding to" land grabbing" and promoting responsible investment in agriculture. IFAD Occasional Paper.

MAFFS, 2011. Agricultural Household Tracking Survey (AHTS) Final Report. Tech. rep., Ministry of Agriculture, Forestry and Food Security, Freetown.

Manski, C. F., 1990. Nonparametric bounds on treatment effects. The American Economic Review 80 (2), 319–323.

Marfurt, F., Käser, F., Lustenberger, S., 2016. Local Perceptions and Vertical Perspectives of a Large Scale Land Acquisition Project in Northern Sierra Leone. Homo Oeconomicus 33 (3), 261–279.
URL http://link.springer.com/10.1007/s41412-016-0020-5

Millar, G., 2015a. Investing in peace: foreign direct investment as economic restoration in Sierra Leone? Third World Quarterly 36 (9), 1700–1716.
URL http://dx.doi.org/10.1080/01436597.2015.1044960

Millar, G., 2015b. “We Have No Voice for That”: Land Rights, Power, and Gender in Rural Sierra Leone. Journal of Human Rights 14 (4), 445–462.

Millar, G., 2016a. Knowledge and Control in the Contemporary Land Rush: Making Local Land Legible and Corporate Power Applicable in Rural Sierra Leone. Journal of Agrarian Change 16 (2), 206–224.

Millar, G., 2016b. Local experiences of liberal peace: Marketization and emergent conflict dynamics in Sierra Leone. Journal of Peace Research 53 (4), 569–581.

Millar, G., 2017. For whom do local peace processes function? Maintaining control through conflict management. Cooperation and Conflict 52 (3), 293–308.

Mokuwa, E., Voors, M., Bulte, E., Richards, P., 2011. Peasant grievance and insurgency in Sierra Leone: judicial serfdom as a driver of conflict. African Affairs 110 (440), 339–366.

Müller, K., Wickham, H., 2018. tibble: Simple Data Frames.
URL https://cran.r-project.org/package=tibble

Nolte, K., Chamberlain, W., Giger, M., 2016. International Land Deals for Agriculture Fresh insights from the Land Matrix : Analytical Report II. Tech. rep., Centre for Development and Environment, University of Bern; Centre de coopération internationale en recherche agronomique pour le développement; German Institute of Global and Area Studies; University of Pretoria; Bern Open Publishing, Bern, Montpellier, Hamburg, Pretoria.

Peters, P. E., 2004. Inequality and social conflict over land in Africa. Journal of agrarian change 4 (3), 269–314.

Peters, P. E., 2013. Conflicts over land and threats to customary tenure in Africa. African Affairs 112 (449), 543–562.

R Core Team, 2017. R: A Language and Environment for Statistical Computing.
URL `https://www.r-project.org/`

Reardon, T., Timmer, C. P., Barrett, C. B., Berdegue, J., 2003. The rise of supermarkets in Africa, Asia, and Latin America. American journal of agricultural economics 85 (5), 1140–1146.

Richards, P., 1986. Coping with hunger: Hazard and experiment in a West African farming system. Allen & Unwin, London.

Scott, J. C., 1998. Seeing like a state: How certain schemes to improve the human condition have failed. Yale University Press.

Shete, M., Rutten, M., 2015. Land Use Policy Impacts of large-scale farming on local communities ' food security and income levels – Empirical evidence from Oromia Region , Ethiopia. Land Use Policy 47, 282–292.
URL `http://dx.doi.org/10.1016/j.landusepol.2015.01.034`

SiLNoRF, 2014. Annual Monitoring Report on the operations of Addax Bioenergy. Tech. rep., Sierra Leone Network on the Right to Food, Freetown.
URL `https://www.udot.utah.gov/main/uconowner.gf?n=25260717873807538`

Sokoloff, K. L., Engerman, S. L., 2000. Institutions, factor endowments, and paths of development in the new world. Journal of Economic perspectives 14 (3), 217–232.

Taylor, J. E., Filipski, M. J., 2014. Beyond experiments in development economics: local economy-wide impact evaluation. Oxford University Press.

Townsend, R. M., 1994. Risk and insurance in village India. Econometrica: Journal of the Econometric Society, 539–591.

UNDP, 2016. Human Development Report 2016: Human Development for Everyone. Tech. rep., United Nations Development Programme, New York.

Van Kerm, P., 2009. sgini: – Generalized Gini and Concentration coefficients (with factor decomposition) in Stata, v1.1 (revised February 2010). Tech. rep., CEPS/INSTEAD, Differdange, Luxembourg.

Wickham, H., Francois, R., Henry, L., Müller, K., 2017. dplyr: A Grammar of Data Manipulation. URL https://cran.r-project.org/package=dplyr

571 # 7. Figures

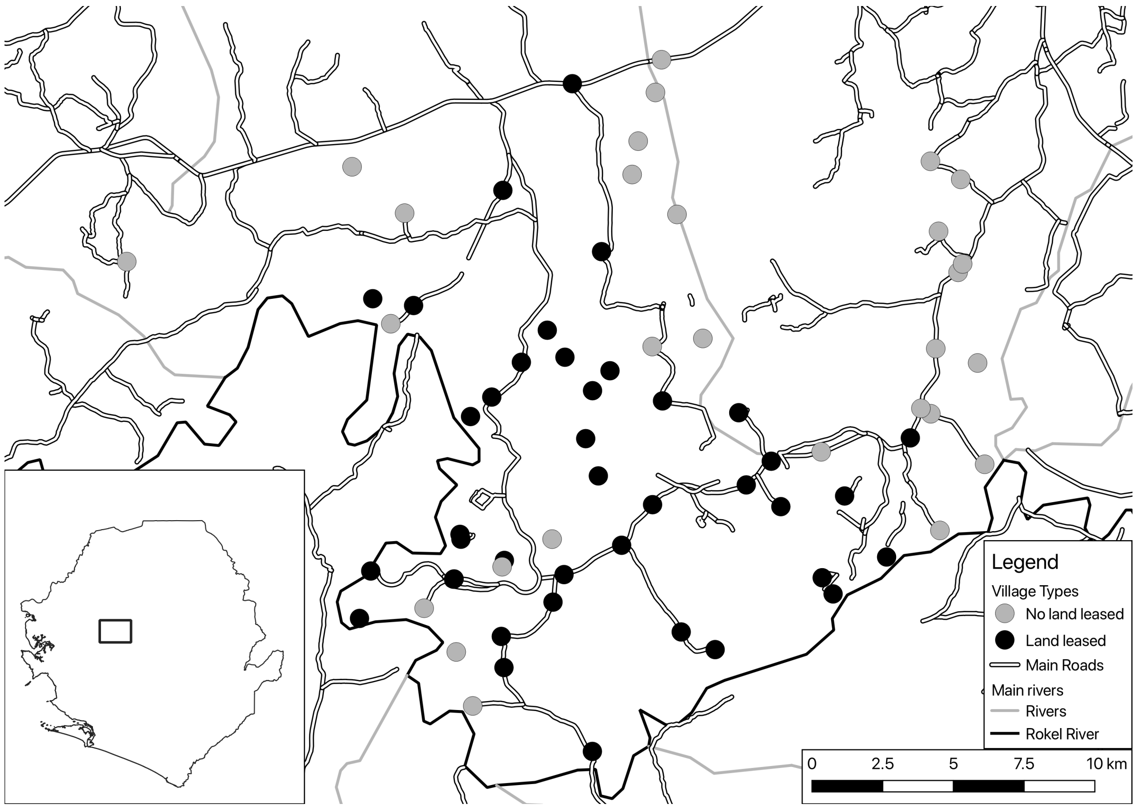

Figure 1: Village Locations

Shows location of all study villages. Source: survey data

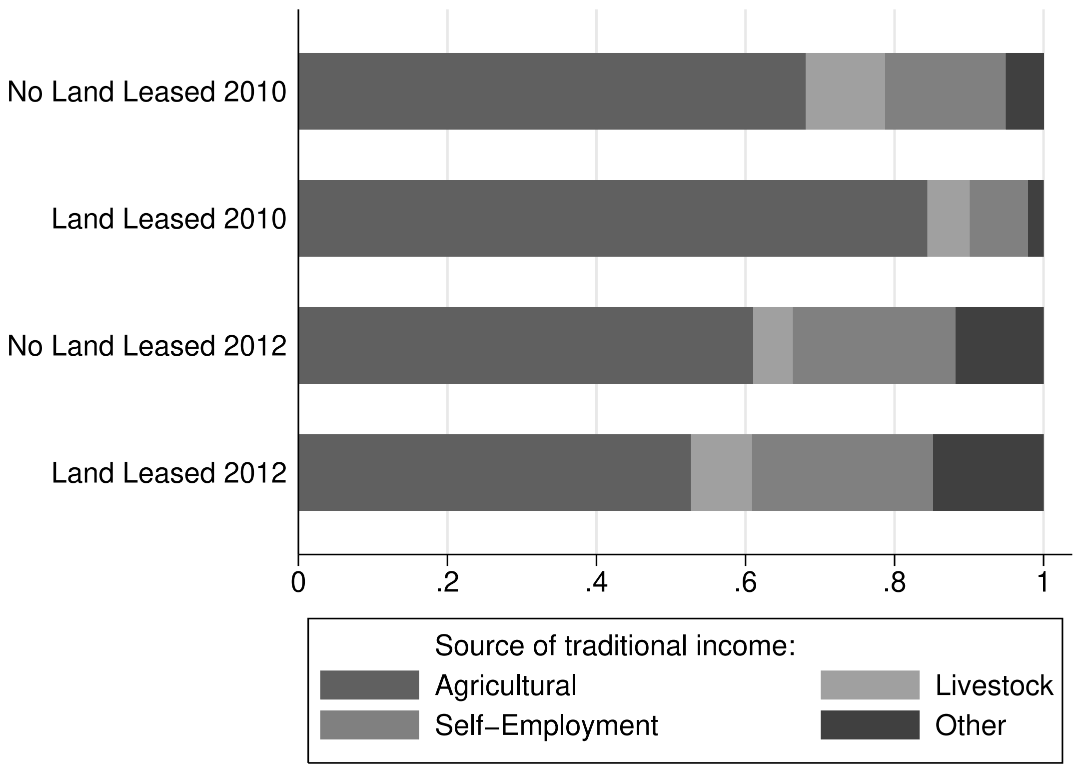

Figure 2: Income Proportions (2010-2012)

Shows proportions of traditional income (that is, excluding 'new' income sources like land lease payments and salaried income). Other income includes remittances, self-declared other revenues and pension income. Source: survey data

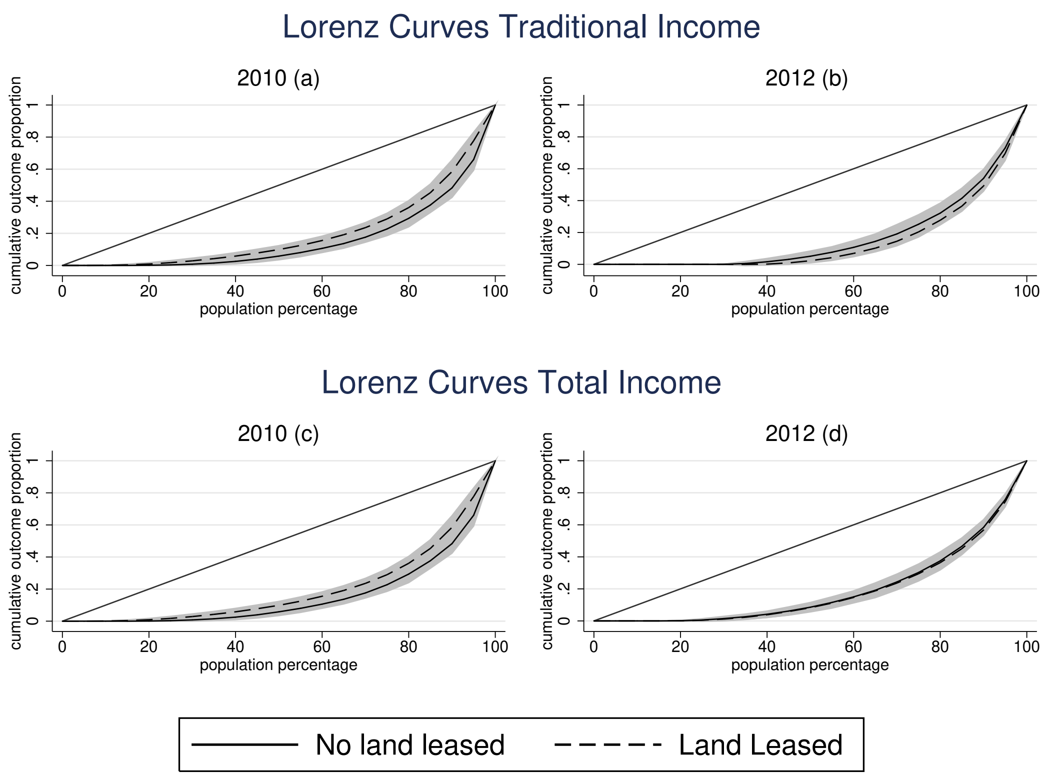

Figure 3: Inequality (2010-2012)

Lorenz curves based on income (not IHS) for panel observations only. Shaded area are confidence intervals, with standard errors clustered at village level. Source: survey data

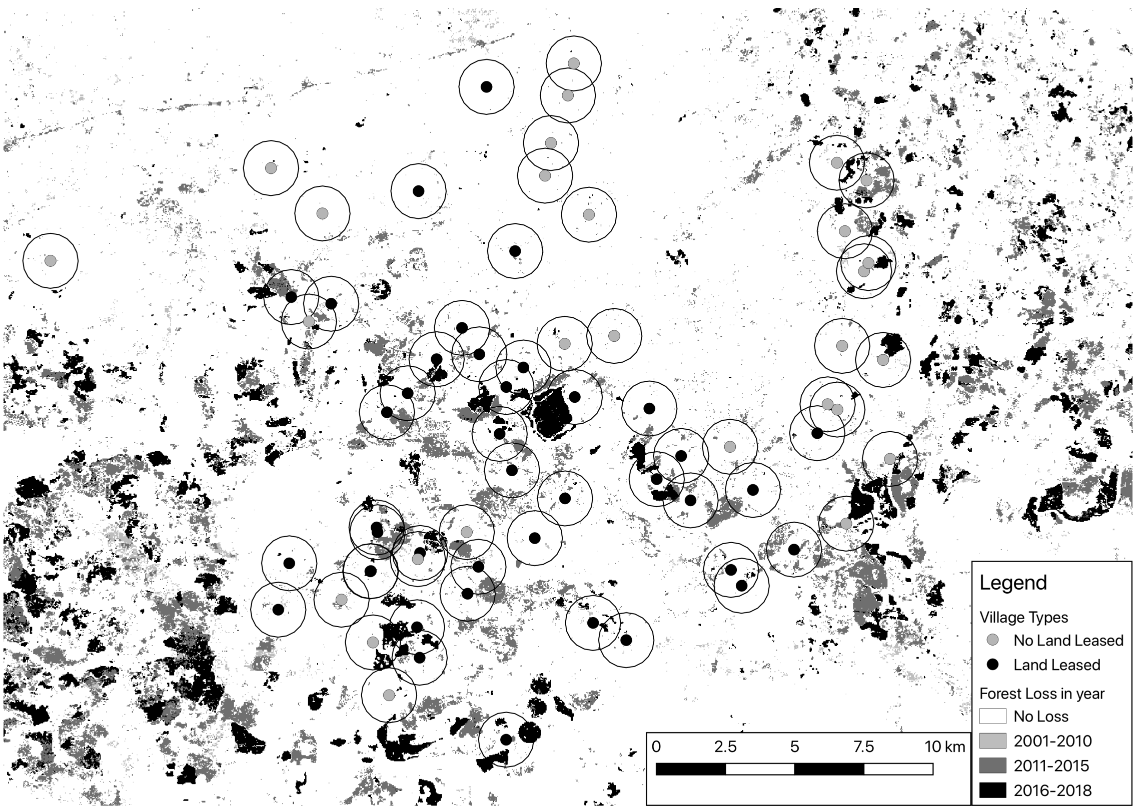

Figure 4: Forest Loss Map

This map shows forest loss from 2001-2018 around the sample villages (Pixel resolution is 30x30m). The circles have a radius of 1km. Forest loss within one of these circles is considered forest loss for that village. Source: Hansen/UMD/Google/USGS/NASA

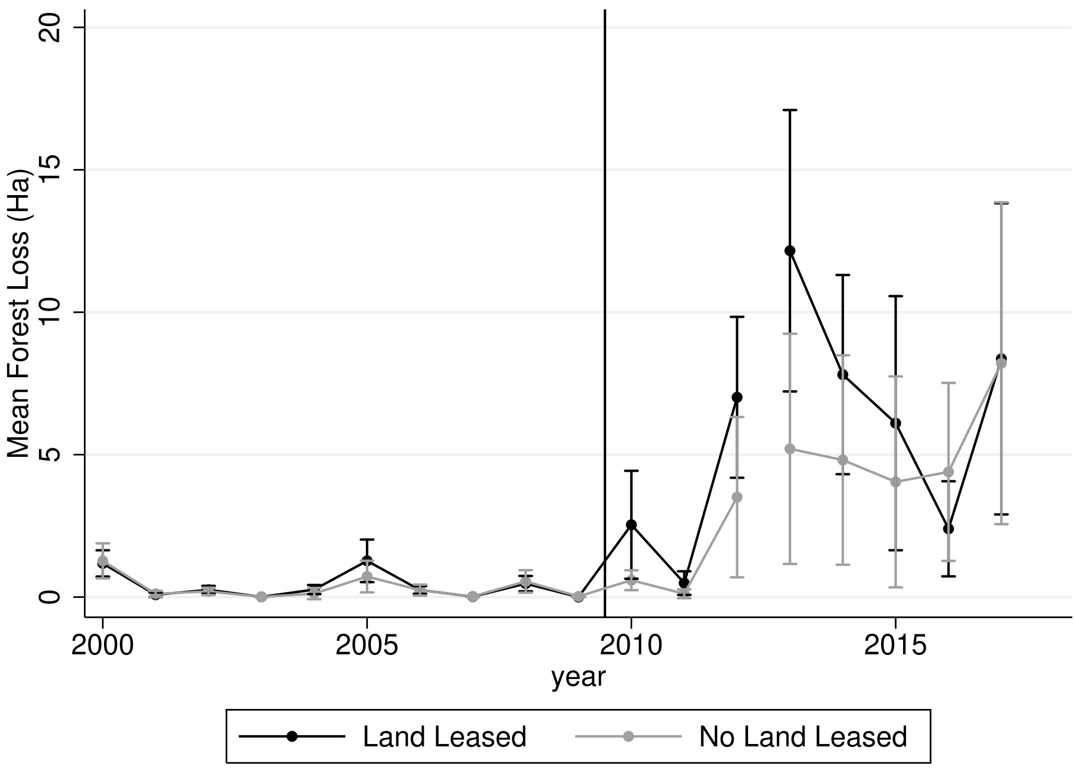

Figure 5: Forest Loss 2001-2017

Graph shows average yearly forest loss in circles with 1km radius around villages. Graph shows 95% confidence intervals. The break in trend lines denote the inclusion of data from a more precise satellite (Landsat 8). Source: Hansen/UMD/Google/USGS/NASA

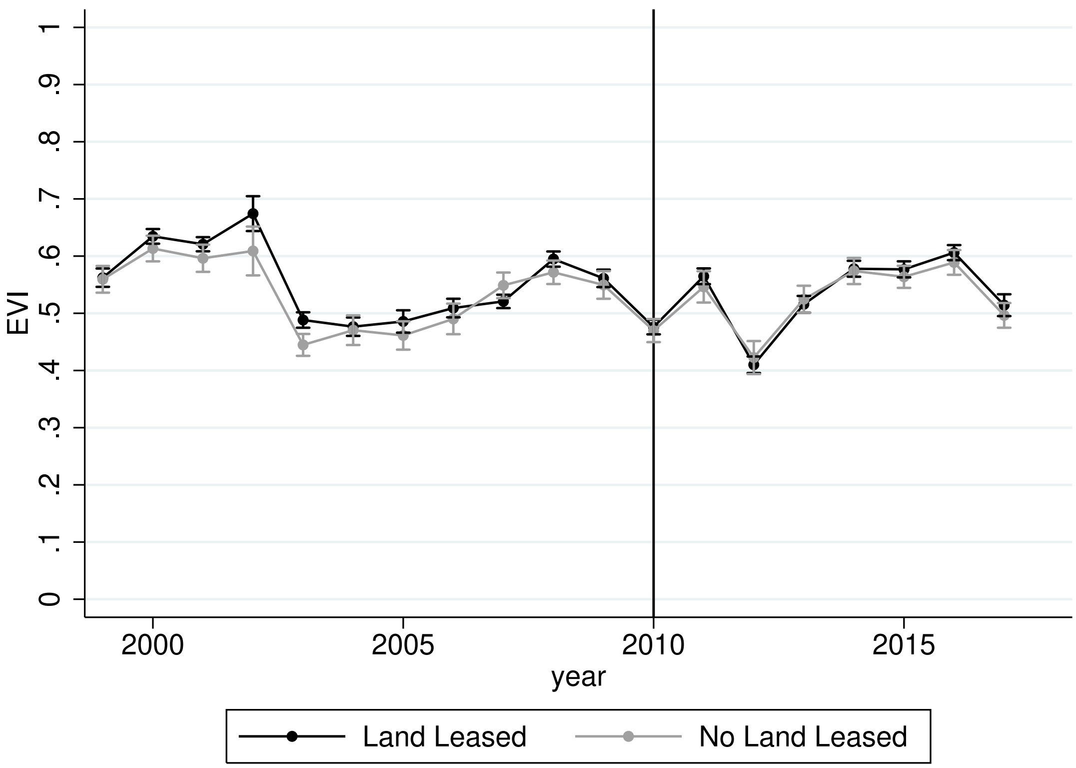

Figure 6: Greenness (EVI) 1999-2017

Graph shows average maximum yearly EVI (greenness or vegetation) in circles with 1km radius around villages. Original pixel size was 30x30m, coarsened to 150x150m to reduce spatial correlation. Graph shows 95% confidence intervals. Source: USGS

8. Tables

Table 1: Studies on investment under study

Author	Type	Cluster N	N	Methods	Findings
Anane and Abiwu (2011)	PR	12	NA	SI, SSI and FG	The development programs were slow to start and did not cause tangible benefits. Food production has gone down because the company is using fertile land. Access to water has gone down. Working conditions for the company are poor: irregular contracts, no safety gear or food provided
Baxter (2011)	PR	NA	NA	SSI	The land leased was under use and fertile, despite contrary claims by the company. Women were not consulted in the decision-making process. Wages for casual labourers are too low to cover daily food needs
Baxter (2013)	PR	10	84	FG and SSI	Food security has gone down, poverty went up. Benefits for job-holders and landowners (though jobs are reported to be low-paying). Higher school dropout, teenage pregnancy, broken marriages, theft, social tensions. Breakdown of traditional social structures
SiLNoRF (2014)	PR	NA	NA	SSI and FG	Increase in income in villages close to the factory. Working conditions for employees are good. The company's development programs are improving local food security. Individuals do not feel that they had a choice in accepting the project. Landowners do not agree with the land rent split (only 50% of rent accrues to them). There are several cases of water shortages because of the company's actions. There were several strikes for higher wages, conditions and discrimination
Fielding et al. (2015)	PR	9	459	SI, SSI and FG	Increased labour scarcity, especially during the growing season. Increased in-migration by individuals looking for work. Improved infrastructure: more roads and houses. Reduced land availability. Lower agricultural productivity (or production). Higher incomes because of wage labor
Millar (2015a)	J	12	55	SSI	Most participants had high hopes for economic improvement because of the investment. Many farmers stopped farming to work for the company. Salaries are lower than income from subsistence farming. Land-lease payments are distributed to three people per village, who do not always distribute further. Economic benefits are concentrated with village elites
Millar (2015b)	J	12	26	SSI	Women were excluded in the decision to accept or not accept the project. Women are rarely employed by the company and have no say in deciding how the land-lease payments are spent. This is in line with persistent disempowering gender norms in Sierra Leone. The company was not aware of these norms and took no measures to correct for this
Bottazzi et al. (2016)	J	NA	54	SSI and FG	Land has become more 'monetized': is now a means to earn money rather than produce food. Migrants do not get any benefits. Monetization of land and 'hard' boundaries create new types of land conflicts. The investment exacerbates existing social cleavages
Marfurt et al. (2016)	J	2	180	SSI, FG and PO	Direct payments do not compensate for the negative effects of the company. Labour contracts are very insecure and wages are low. The company leases fertile land, decreasing agricultural production and income
Millar (2016b)	J	12	115	SSI and PO	Land has become more 'monetized' and families feel they have to defend their claim to it. This requires more formal land titles which causes conflicts over (a.o.) exact borders. Jobs are mainly given to individuals part of landowning families. There are tensions around labour provision: many want work but the company cannot provide. There are also tensions between local (not employed) youth and employed youth from outside the project area. Another source of tension is between generations: youth did not get a say in the decision to accept the company, and do not have control over the land-lease payments
Millar (2016a)	J	12	55	SSI and PO	There is a disconnect between how the company and the inhabitants view land. The company uses technology to 'control' the land, which inhabitants were not able to protest against as this requires literacy. Most land is regularly used, even though it is not under constant cultivation
Millar (2017)	J	NA	NA	SSI and PO	Regional elites, who used to function as conflict-solving institutions are now using their influence to acquiesce the local population to ensure their access to company-provided benefits. This makes it almost impossible for the local population to voice grievances. In the long run this led to feelings of marginalization and increased conflict
Bottazzi et al. (2018)	J	55	882	SI	Farmers around the plantation use less agricultural land, attain lower yields and pay more for labour. In contrast, they also find increased incomes, improved food security and more food expenditures.These improvements were largest for landowners and men.

PR=Policy Report, J=Peer-reviewed journal. SI=Structured interviews, SSI=Semi-structured interviews, FG=Focus Groups, PO=Participatory observation
NA=not specified

Table 2: Sample sizes over time

	2010		2012		2015	
	Control	Land Leased	Control	Land Leased	Control	Land Leased
Cross-section						
Observations	1415	2818	1790	3034	649	1118
Villages	37	41	71	47	39	36
Panel						
Observations			915	2240	99	529
Villages			28	40	15	34

Number of panel observations. Participants were matched based on company-assigned ID code (Matching on names leads to a lower number of observations but similar conclusions). Some subsequent analyses have a lower number of observations and/or clusters. This is because in those cases some participants did not answer that specific question. Source: survey data

Table 3: Descriptive Statistics

	Control			Treatment			
	n	mean	sd	n	mean	sd	Diff
Traditional Income (Leones, IHS)	1004	10.98	4.23	1504	12.21	3.33	1.231**
# Assets	1415	3.94	1.48	2818	3.86	1.52	-0.086
House quality (Score, 1-33)	1098	5.13	2.16	1529	5.28	2.05	0.146
Tropical Livestock Unit	1028	0.27	1.67	2144	0.22	0.38	-0.049
Access to Land (=1)	1351	1.00	0.06	2714	0.99	0.08	-0.003
Food shortage (=1)	1374	0.99	0.11	2700	0.99	0.10	0.003
Live births in last 12 months	537	0.54	0.67	2083	0.31	0.55	-0.231**
Total deaths of infants in last 12 months	537	0.07	0.29	2083	0.04	0.23	-0.023
Bed net in household (=1)	1412	0.92	0.27	2811	0.80	0.40	-0.120***

Table shows averages for 2010 (before any land was leased from communities). The final column shows the coefficient of a simple regression of treatment status on the variable, with clustered standard errors at the village level. Stars indicate whether the treatment - control difference is statistically significant, with $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Short Run (2010-2012) effects of a Large-Scale Agricultural investment

	(1) Traditional income (IHS)	(2) Full Income (IHS)	(3) # Assets	(4) House Quality (Score)	(5) Tropical Livestock unit	(6) Access to Land (=1)	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Total deaths of infants in last 12 months	(10) Bed net in household (=1)
Treated	0.389***	0.389***	−0.031	0.106	−0.015	−0.003	0.003	−0.576***	−0.164	−0.118***
	(0.143)	(0.143)	(0.094)	(0.079)	(0.072)	(0.003)	(0.005)	(0.180)	(0.105)	(0.040)
Short Run	−0.256	0.124	0.036	0.314***	−0.091	−0.031**	−0.101***	−0.545***	−0.146	−0.082**
	(0.229)	(0.181)	(0.126)	(0.060)	(0.067)	(0.012)	(0.013)	(0.192)	(0.107)	(0.040)
Treated * Short Run	−0.625**	−0.424**	−0.047	−0.110	0.315	−0.049***	−0.006	0.487**	0.175	0.135***
	(0.262)	(0.197)	(0.141)	(0.080)	(0.192)	(0.018)	(0.018)	(0.196)	(0.110)	(0.050)
Constant	0.000	0.000	0.000	0.000	0.000	0.997***	0.988***	0.000	0.000	0.920***
	(0.110)	(0.110)	(0.073)	(0.048)	(0.038)	(0.002)	(0.004)	(0.176)	(0.102)	(0.022)
Observations	3762	3762	6310	3914	3470	6082	6068	3886	3886	6302
# Clusters	67	67	68	68	67	68	68	60	60	68

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 10). Robust standard errors in parentheses clustered at the village level. Gini Coefficient is based on traditional income in villages with at least six observations. IHS is inverse hyperbolic sine transformation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Long Run (2010-2015) effects of a Large-Scale Agricultural investment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Traditional income (IHS)	Full Income (IHS)	# Assets	House Quality (Score)	Tropical Livestock unit	Access to Land (=1)	Food shortage (=1)	Total births in last 12 months	Total deaths of infants in last 12 months	Bed net in household (=1)
Treated	0.221	0.210	−0.070	0.298**	0.214	−0.002	0.002	−0.247	−0.244	−0.042
	(0.204)	(0.200)	(0.169)	(0.144)	(0.144)	(0.002)	(0.011)	(0.230)	(0.344)	(0.068)
Long Run	0.128	0.220	0.279	0.733***	0.361	−0.021	−0.115**	−0.767***	−0.347	−0.061
	(0.240)	(0.220)	(0.196)	(0.196)	(0.280)	(0.013)	(0.046)	(0.259)	(0.341)	(0.070)
Treated * Long Run	−0.615**	−0.427*	0.104	0.333	−0.424	−0.072**	0.072	0.475*	0.339	0.052
	(0.264)	(0.232)	(0.208)	(0.222)	(0.305)	(0.033)	(0.047)	(0.270)	(0.345)	(0.081)
Constant	0.000	0.000	0.000	0.000	0.000	1.000***	0.990***	0.000	0.000	0.847***
	(0.176)	(0.168)	(0.157)	(0.116)	(0.104)	(0.000)	(0.010)	(0.219)	(0.341)	(0.058)
Observations	748	748	1256	796	990	1202	1190	846	846	1242
# Clusters	44	44	49	47	45	48	48	41	41	48

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 10). Robust standard errors in parentheses clustered at the village level. Gini Coefficient is based on traditional income in villages with at least six observations. IHS is inverse hyperbolic sine transformation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Spillover effects (2010-2012)

	(1) Traditional income (IHS)	(2) Full Income (IHS)	(3) # Assets	(4) House Quality (Score)	(5) Tropical Livestock unit	(6) Access to Land (=1)	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Total deaths of infants in last 12 months	(10) Bed net in household (=1)
Distance	-0.225^{***}	-0.225^{***}	-0.106^{*}	0.066	-0.011	0.001	0.003	0.326^{***}	-0.096	0.033^{**}
	(0.075)	(0.075)	(0.056)	(0.039)	(0.027)	(0.001)	(0.003)	(0.111)	(0.088)	(0.014)
Short Run	-0.236	0.140	0.036	0.304^{***}	-0.092	-0.031^{**}	-0.101^{***}	-0.601^{***}	-0.138	-0.082^{**}
	(0.166)	(0.124)	(0.103)	(0.057)	(0.065)	(0.011)	(0.013)	(0.137)	(0.098)	(0.039)
Distance * Short Run	0.482^{***}	0.389^{***}	0.246^{***}	0.072^{*}	-0.059	0.016^{*}	-0.010	-0.332^{**}	0.050	0.020
	(0.138)	(0.102)	(0.088)	(0.041)	(0.044)	(0.008)	(0.014)	(0.118)	(0.091)	(0.031)
Constant	-0.009	-0.009	0.000	-0.009	0.000	0.997^{***}	0.988^{***}	0.055	-0.016	0.920^{***}
	(0.085)	(0.085)	(0.068)	(0.048)	(0.038)	(0.002)	(0.003)	(0.123)	(0.095)	(0.021)
Observations	1288	1288	1830	1444	980	1750	1784	516	516	1828
# Clusters	28	28	28	28	27	28	28	22	22	28

OLS regressions. Only control villages included. Normalized variables (not columns 6, 7 and 10). Distance is euclidean distance (in km) to nearest treated village, standardized. IHS is inverse hyperbolic sine transformation. Robust standard errors in parentheses clustered at the village level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Effects on laborers (2010-2015)

	(1) Traditional income (IHS)	(2) Full Income (IHS)	(3) # Assets	(4) House Quality (Score)	(5) Tropical Livestock unit	(6) Access to Land (=1)	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Total deaths of infants in last 12 months	(10) Bed net in household (=1)
Labourer	0.194	0.186	−0.116	-0.526^{***}	-0.299^{**}	0.003	−0.003	0.198^{*}	−0.052	-0.059^{*}
	(0.141)	(0.146)	(0.069)	(0.116)	(0.117)	(0.003)	(0.009)	(0.114)	(0.056)	(0.032)
Treated	0.146	0.138	−0.027	0.449^{***}	0.330^{**}	−0.003	0.004	−0.332	−0.222	−0.018
	(0.230)	(0.229)	(0.171)	(0.149)	(0.152)	(0.003)	(0.012)	(0.234)	(0.345)	(0.072)
Long Run	0.128	0.220	0.279	0.733^{***}	0.361	−0.021	-0.115^{**}	-0.767^{***}	−0.347	−0.061
	(0.240)	(0.220)	(0.196)	(0.196)	(0.281)	(0.013)	(0.046)	(0.259)	(0.342)	(0.070)
Long Run * Treated	-0.485^{*}	-0.560^{**}	0.028	0.259	-0.555^{*}	-0.054^{*}	0.091^{*}	0.514^{*}	0.291	0.052
	(0.262)	(0.240)	(0.209)	(0.222)	(0.310)	(0.029)	(0.048)	(0.278)	(0.345)	(0.083)
Labourer * Long Run	-0.336^{**}	0.344^{**}	0.231^{**}	0.242^{*}	0.336^{**}	−0.045	-0.044^{*}	−0.090	0.110	0.000
	(0.152)	(0.129)	(0.087)	(0.143)	(0.162)	(0.029)	(0.025)	(0.137)	(0.072)	(0.039)
Constant	0.000	0.000	0.000	0.000	0.000	1.000^{***}	0.990^{***}	0.000	0.000	0.847^{***}
	(0.176)	(0.168)	(0.158)	(0.116)	(0.104)	(0.000)	(0.010)	(0.219)	(0.342)	(0.058)
Observations	748	748	1250	794	990	1202	1190	846	846	1242
# Clusters	44	44	49	47	45	48	48	41	41	48

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 10). Laborers are all households who claimed to work for the company at some point in the 2015 survey. IHS is inverse hyperbolic sine transformation. Robust standard errors in parentheses clustered at the village level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

9. Appendix

Table A1: Variable definitions

Variable	Variable Definition
Traditional Income	Sum of Agricultural and livestock sales, self-employment and other income (including remittances) in January of that year. Winsorized at the 95% level, then transformed with Inverse Hyperbolic Sine
Total Income	Traditional income, plus company's land payments (2012 only) and salaried income (2015 only). Winsorized at the 95% level, then transformed with Inverse Hyperbolic Sine
# Assets	Sum of how many of the following assets they owned: house, car, bicycle, tv, radio, satellite, sewing machine, fridge, iron pots, iron kettle, mobile phone, bed mattress, motorcycle, plastic chairs, mosquito nets, tractor, generator
House Quality	Score based on the average quality of their houses. Floors: No floor 0p, Mud 1p, Cement 5p. Walls: Wattle & Daub 1p, Reeds & Thatch 2p, Mud bricks 3p, Mud bricks and plaster 4p, Wooden 4p, Concrete 5p. Roof: None 0p, Thatch 1p, Tarp 2p, Zinc 5p. Maximum score: 33
Livestock	Tropical livestock unit on number of livestock owned, based on cattle, goats, sheep, pigs, rabbits and chickens. One tropical livestock unit is often equated to a 250 kg animal (Jahnke, 1982).
Access to land	Answer to question 'Do you currently have access to land for cultivation?' (yes/no)
Food Security	Answer to question 'Was there a shortage of food in the household at any time last year?' (yes/no)
Total births	# Total births per household in the previous year
Total deaths	# Total births of infants (<1 year old) in the previous year
Bed net	Whether a bed net is present in the household (yes/no)

Table A2: Inequality Short and Long Run

	(1) Gini Short Run Traditional Income	(2) Gini Short Run Total Income	(3) Gini Long Run Traditional Income	(4) Gini Long Run Total Income
Treated	-0.051	-0.051	-0.051	-0.040
	(0.034)	(0.034)	(0.061)	(0.051)
Post	0.154^{***}	0.038	0.006	-0.004
	(0.042)	(0.040)	(0.071)	(0.062)
Treated * Post	0.114^{**}	0.036	0.162^{**}	0.087
	(0.055)	(0.049)	(0.078)	(0.070)
Constant	0.170^{***}	0.170^{***}	0.149^{**}	0.133^{***}
	(0.029)	(0.029)	(0.057)	(0.046)
Observations	96	96	54	54

OLS regressions. Gini Index is calculated only for villages with at least 5 observations with income data.

* p < 0.10, ** p < 0.05, *** p < 0.01.

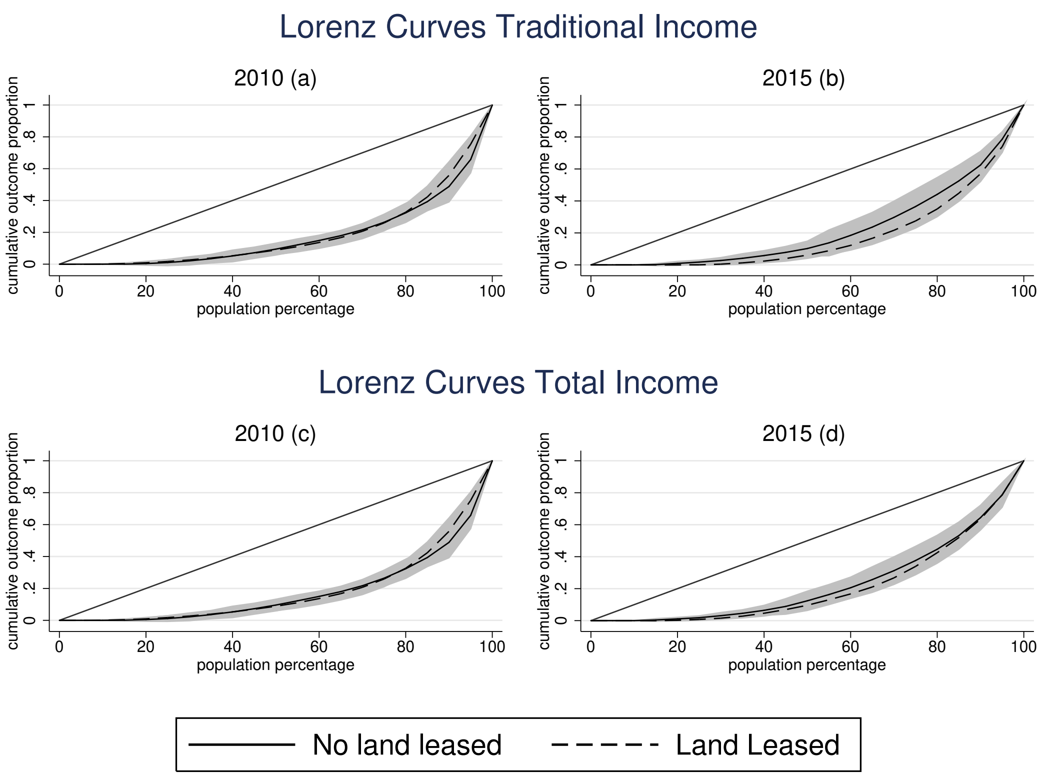

Figure A1: Inequality (2010-2015)

Lorenz curves based on income (not IHS) for panel observations only. Shaded area are confidence intervals, with standard errors clustered at village level. Source: survey data

Table A3: Short Run (2010-2012) effects of a Large-Scale Agricultural investment: stricter merge results

	(1) Traditional income (IHS)	(2) Full Income (IHS)	(3) # Assets	(4) House Quality (Score)	(5) Tropical Livestock unit	(6) Access to Land (=1)	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Total deaths of infants in last 12 months	(10) Bed net in household (=1)
Treated	0.387***	0.387***	−0.028	0.115	0.039	−0.004	0.002	−0.605***	−0.166	−0.120***
	(0.146)	(0.146)	(0.095)	(0.082)	(0.089)	(0.002)	(0.005)	(0.190)	(0.114)	(0.042)
Short Run	−0.241	0.150	0.066	0.326***	−0.089	−0.030**	−0.103***	−0.553**	−0.147	−0.074**
	(0.234)	(0.180)	(0.122)	(0.062)	(0.073)	(0.012)	(0.014)	(0.218)	(0.119)	(0.035)
Treated * Short Run	−0.622**	−0.452**	−0.069	−0.127	0.358	−0.046**	0.000	0.488**	0.167	0.133***
	(0.267)	(0.198)	(0.138)	(0.083)	(0.261)	(0.019)	(0.019)	(0.221)	(0.122)	(0.047)
Constant	0.000	0.000	0.000	0.000	0.000	0.997***	0.989***	0.000	0.000	0.921***
	(0.113)	(0.113)	(0.072)	(0.051)	(0.032)	(0.002)	(0.004)	(0.186)	(0.111)	(0.024)
Observations	3428	3428	5764	3578	3200	5570	5540	3582	3582	5758
# Clusters	65	65	67	67	65	67	67	58	58	67

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 10). IHS is inverse hyperbolic sine transformation. Robust standard errors in parentheses clustered at the village level. Sample is based on a more restrictive merge which also checks name, village name and number of years lived in the area.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Long Run (2010-2015) effects of a Large-Scale Agricultural investment: stricter merge results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Traditional income (IHS)	Full Income (IHS)	# Assets	House Quality (Score)	Tropical Livestock unit	Access to Land (=1)	Food shortage (=1)	Total births in last 12 months	Total deaths of infants in last 12 months	Bed net in household (=1)
Treated	0.243	0.233	-0.085	0.290^{*}	0.175	-0.002	0.002	-0.302	-0.102	-0.075
	(0.205)	(0.200)	(0.193)	(0.148)	(0.144)	(0.002)	(0.012)	(0.262)	(0.123)	(0.075)
Long Run	0.177	0.275	0.273	0.760^{***}	0.321	-0.022	-0.100^{**}	-0.777^{**}	-0.125	-0.087
	(0.259)	(0.233)	(0.205)	(0.189)	(0.290)	(0.014)	(0.045)	(0.297)	(0.123)	(0.075)
Treated * Long Run	-0.658^{**}	-0.482^{*}	0.086	0.222	-0.400	-0.062^{**}	0.055	0.526^{*}	0.131	0.085
	(0.275)	(0.242)	(0.219)	(0.214)	(0.315)	(0.027)	(0.047)	(0.305)	(0.124)	(0.088)
Constant	0.000	0.000	0.000	0.000	0.000	1.000	0.989^{***}	0.000	0.125	0.870^{***}
	(0.179)	(0.170)	(0.181)	(0.123)	(0.103)	(.)	(0.011)	(0.254)	(0.123)	(0.065)
Observations	690	690	1118	716	890	1076	1062	750	750	1108
# Clusters	43	43	47	46	45	47	47	39	39	47

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 10). IHS is inverse hyperbolic sine transformation. Robust standard errors in parentheses clustered at the village level. Sample is based on a more restrictive merge which also checks name and village name.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Short Run (2010-2012) of a Large-Scale Agricultural investment: Income splits

	(1) Agricultural Income (IHS)	(2) Livestock Income (IHS)	(3) Self-Employment Income (IHS)	(4) Other Income (IHS)
Treated	0.470^{**}	-0.184^{**}	-0.192	-0.087
	(0.181)	(0.087)	(0.124)	(0.060)
Short Run	-0.280	-0.068	0.034	0.579^{***}
	(0.223)	(0.097)	(0.072)	(0.097)
Treated * Short Run	-0.695^{***}	0.064	0.167	-0.004
	(0.250)	(0.114)	(0.127)	(0.120)
Constant	0.000	0.000	0.000	0.000
	(0.139)	(0.077)	(0.077)	(0.048)
Observations	3762	3762	3762	3762
# Clusters	67	67	67	67

OLS regressions. Standardized and centered on control group at baseline. IHS is inverse hyperbolic sine transformation. Robust standard errors in parentheses clustered at the village level. Traditional income variable in main tables are sum of these main components of income.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Long Run (2010-2015) of a Large-Scale Agricultural investment: Income splits

	(1) Agricultural Income (IHS)	(2) Livestock Income (IHS)	(3) Self-Employment Income (IHS)	(4) Other Income (IHS)
Treated	0.176	0.174	0.347	0.180
	(0.255)	(0.147)	(0.241)	(0.135)
Long Run	-0.414	0.261^{*}	1.148^{***}	0.909^{***}
	(0.306)	(0.141)	(0.209)	(0.222)
Treated * Long Run	-0.497	-0.128	-0.310	-0.274
	(0.332)	(0.174)	(0.299)	(0.249)
Constant	0.000	0.000	0.000	0.000
	(0.211)	(0.122)	(0.112)	(0.097)
Observations	748	748	748	748
# Clusters	44	44	44	44

OLS regressions. Standardized and centered on control group at baseline. IHS is inverse hyperbolic sine transformation. Robust standard errors in parentheses clustered at the village level. Traditional income variable in main tables are sum of these main components of income.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Attrition

	(1) Dropout SR	(2) Dropout LR
Traditional Income (Leones, IHS)	0.049	0.043
	(0.037)	(0.030)
# Assets	-0.067	0.019
	(0.068)	(0.113)
House quality (Score, 1-33)	-0.124**	-0.025
	(0.063)	(0.063)
Tropical Livestock Unit	0.045**	0.306
	(0.020)	(0.483)
Access to Land (=1)=1	-0.439	0.000
	(0.802)	(.)
Food shortage (=1)=1	-1.136	0.000
	(0.900)	(.)
Live births in last 12 months	-0.525***	0.149
	(0.159)	(0.235)
Total deaths of infants in last 12 months	0.062	-0.119
	(0.319)	(0.379)
Bed net in household (=1)=1	0.279	0.666**
	(0.466)	(0.337)
Treated	-0.314	0.520
	(0.877)	(1.006)
* Traditional Income (Leones, IHS)	-0.124***	-0.025
	(0.047)	(0.035)
* # Assets	0.023	-0.066
	(0.096)	(0.128)
* House quality (score)	0.053	-0.024
	(0.071)	(0.070)
* Tropical Livestock Unit	0.117	-0.338
	(0.168)	(0.495)
* Access to land (=1)	0.000	0.000
	(.)	(.)
* Food Shortage	0.000	0.000
	(.)	(.)
* Total births	0.320	-0.124
	(0.217)	(0.263)
* Total child deaths	0.600	0.585
	(0.422)	(0.447)
* Bed net in household	-0.206	-0.609
	(0.544)	(0.410)
Constant	2.150	0.261
	(1.323)	(0.934)
Observations	666	660
# Clusters	68	68
% Dropped out	0.25	0.85

Probit regressions. Robust standard errors in parentheses clustered at the village level. Data is all 2010 data with indicators for being absent in later rounds. Some dummy variables are dropped from the model because of low variation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Bounds Analysis

	(1) Original	(2) Worst	(3) + (-) 0.5SD	(4) + (-) 0.25SD	(5) + (-) 0.1SD)
Traditional income (IHS)	-0.625**	0.526*	-0.194	-0.377*	-0.488**
	(0.262)	(0.310)	(0.212)	(0.198)	(0.192)
Full Income (IHS)	-0.426**	0.760**	-0.036	-0.182	-0.270*
	(0.197)	(0.298)	(0.160)	(0.145)	(0.139)
# Assets	-0.047	1.946***	0.284**	0.126	0.031
	(0.141)	(0.372)	(0.122)	(0.106)	(0.100)
House Quality (Score)	-0.110	1.753***	0.291**	0.125	0.025
	(0.080)	(0.377)	(0.112)	(0.087)	(0.075)
Tropical Livestock unit	0.315	-3.004***	-0.386**	-0.084	0.098
	(0.192)	(0.753)	(0.185)	(0.155)	(0.144)
Total births in last 12 months	0.487**	-1.036***	-0.034	0.090	0.165
	(0.196)	(0.328)	(0.131)	(0.119)	(0.112)
Total deaths of infants in last 12 months	0.175	-2.891***	-0.137*	-0.022	0.048
	(0.110)	(0.630)	(0.073)	(0.063)	(0.058)

OLS Regressions, standard errors in parentheses clustered at village level. Table reports coefficients of interaction term of treatment and later time period. Column 1 reports the same coefficients as in table 4 and ignores attrited households. Columns 2-5 give alternative values to attrited households, depending on whether the original coefficient is positive (or negative). For column 2 (worst-case) attrited households in the treatment group get the minimum (maximum) in the treatment group and households in the control group get the maximum (minimum) in the control group. Column 3-5 assigns attrited households in the treatment group the treatment mean minus (plus) X SD, and attrited households in the control group the control mean plus (minus) X SD.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.