Local Economy effects of Large-Scale Agricultural Investments*

Paul Hofman[1], Esther Mokuwa[2], Paul Richards[2], and Maarten Voors[1]

[1]Development Economics Group, Wageningen University, The Netherlands

[2]Njala University, Sierra Leone

May 16, 2019

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. We examine the economic impacts of a large-scale biofuel plantation in Sierra Leone - a major target investments in land. 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 labour demand shock,

*Corresponding author: paul.hofman@wur.nl. We thank conference participants at LandAC, EUDN, World Bank Land and Poverty Conference and numerous seminar participants. We thank our field team who did a tremendous job of collecting data. Thanks also to Sanne Schimmel en Mirella Schrijvers for excellent research assistance. This work was supported by NWO Global Challenges Programme grant # W08.250.104, NWO grant # 451-14-001.

caused by a clash between the private and commercial agricultural calendar, increasing the local price of labour. A spillover analysis confirms that the impacts are at least partially transmitted by a shock to the local economy. Within host communities, households that are employed at the plantation see their incomes and assets increase. However, as a result, village-level inequality increases.

1 Introduction

In recent years, 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, Deininger, & Selod, 2013; Koning & van Ittersum, 2009). The Land Matrix, which documents all transnational land acquisitions, to date has recorded 1591 'concluded' agricultural investments, in total covering about 49 million hectares. In some African countries over 30 percent of arable land is foreign-owned (Landmatrix, 2018; Nolte, Chamberlain, & Giger, 2016). These investments often take the form of large scale plantations, with land rights acquired for a long period (typically 99 years). This trend is likely to increase due to the projected rise in demand for food, animal fodder and energy crops.

To date, the impacts of such land deals have not received rigorous academic attention. 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 & Dercon, 2014; Ellis, 2005), inducing innovation (Borensztein, De Gregorio, & Lee, 1998), enabling access to finance (Alfaro, Chanda, Kalemli-Ozcan, & Sayek, 2010) and the organization of production and marketing (Reardon, Timmer, Barrett, & Berdegue, 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, 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 & Engerman, 2000; Herbst, 2014; Dorward, Kirsten, Omamo, Poulton, & Vink, 2009). Often, land investment deals are made between companies and elites and exclude local people from the negotiations, increasing corruption (Peters, 2013; De Schutter, 2011).

Despite the scale of foreign investments in agriculture, local economic impacts have to date failed to receive rigorous academic investigation. Exceptions are Herrmann and Grote (2015), who assess a sugarcane plantation and outgrower scheme in Malawi, and find positive economic impacts. The plantation attracts labourers from nearby villages. Typically, poorest households opt in to this casual wage labourer opportunity and are substantially better off than non-labourers. Incomes nearly double for plantation labourers. 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 labourers compared to other households in the same villages. There is however, no significant effect on agricultural or total household income.[1]

Methodologically, a key limitation of these papers is that they rely on post-invervention data, requiring strong assumptions to assess causal effects. Investors typically do not select concession sites at random and take important ecoligical, political and economic

[1] Other papers include Shete and Rutten (2015) and Jiao, Smith-hall, and Theilade (2015).

characteristics into account such as agricultural potential, distance to input and output markets, local institutions and labour availability. Failing to adaquately 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, from before the plantation was created allows us to control for such selection effects.

In addition, a key limitation of these papers is that they compare respondents directly hired by the plantations to those not hired but living in the same village. This ignores household economic impacts through other channels such as the due to competition over access to land. 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, Gohou, and Vencatachellum (2012) show how the net effects of such investment project crucially depends on the intermediate impacts on labour and land markets. If labour and land are abundant, increased demand for labour and land should not impact local economies. However, this is rarely the case, such as in rural Sierra Leone, which we consider here where there is severe competition over labour (Mokuwa, Voors, Bulte, & Richards, 2011; Bulte, Richards, & Voors, 2018). In such cases increased employment opportunities outside the village may cause a decrease in labour 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 well suited to investigate the impact of foreign agricultural investments. Sierra Leone is a poor country, characterized by slash and burn agriculture and limited access to financial and output markets. The majority of the population is engaged in the agricultural sector. Farms are very small: average size is 0.5 hectares. To a large degree farm output is determined by labour 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 aquired over 25% of the country's arable land (Baxter, 2013; Landmatrix, 2018).

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 labour 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 allows us to examine effects over the short (2 years) and long run (5 years). We find that average income drops substantially, by about 0.5 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 argue that the labour demand shock, caused by a clash between the private and commercial agricultural calendar, increases the local price of labour. 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 impact is through local market effects is strengthened when we find that changes in incomes over time reduce 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 & Filipski, 2014; Cust & 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 labour prices and local income, and this effect declines when moving away from the mine. We find similar results: in our case redutions in labour availability reduces household production.

The plantation has received considerable attention in the media and has been the focus of several policy reports and journal articles. Table 1 provides an overview. Most reports highlight that the investment was forced through by politicians and the local elite without involving communities, and report decreases in average incomes, some cite improved incomes, especially for specific landowners. Some also point to increases in conflicts due to the plantation creating conflicts over access to land and surfice rents and in other cases exacerbating existing tensions over land claims. A key drawback is that most studies that rely on qualitative small sample case studies alone. While this provides rich insights into relevant local dynamics for selected localities, they fall short in assessing average differences.

Closest to our work is Bottazzi, Crespo, Bangura, and Rist (2018). They use a matching algorithm to match 592 respondents in 34 vilages 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 labour prices. but increased incomes. 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 improve on 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, 2018). Sierra Leone 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. The main agricultural input is labour, with fertilizer and high-yielding seeds virtually unavailable in the country. A 2011 survey found that 65% of households experienced a shortage of labour in the agricultural season. Farm production to a large extent relies on family labour. About one third of households hire labour (MAFFS, 2011). Bottazzi, Goguen, and Rist (2016) also state that land is generally plentiful in the area but labour is limited.

Outside investors can potentially improve this situation by bringing in improved technologies and large-scale production that achieve economies of scale. The Government of Sierra Leonean Government aims to "Make agriculture the 'engine' for socio-economic growth and development through commercial agriculture", (MAFFS, 2009, pp 7). Yet, the societal implications of this development agenda are not well understood. More recently, foreign investment led growth has been adopted as a core stragegy. The government "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).

Outside investment can provide a steady income for subsistence farmers, alleviating cash constraints. Furthermore, the investor can attempt to increase smallholder productivity through farmer training programs. However, decreased access to land and labour might lower local food production causing local shortages. In addition, inequality might increase as land rents are paid to land owners, who are commonly wealthier already. Furthermore, the investment is likely to compete in the market for wage labourers, thus increasing labour prices. This is negative for households that mainly hire in labour, but positive for those who mainly hire out labour.

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 50 year lease. Landowners received 8.90 US$ in compensation per hectare per year, half of which goes to the landowners and the other half to various local elites. They also receive an additional payment of 3.46 US$, making the total payment for landowners 7.91 US$/Ha/year. In 2014, the company leased land from 52 villages, leasing between 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 which were on fixed-term contracts. The company aimed to recruit labour from communities in and around the plantation. The main labour demand of the plantation overlaps with smallholder production. Peak smallholder labour demand is in February-April when land is prepared ('cleared') using slash-and-burn agriculture (Richards, 1986), matching the plantation's peak labour demand. Besides providing benefits in terms of employment for labourers from nearby villages, and surface rents for land owners, the company created a health clinic, provided several farming training programs and has a compensation program for destroyed tree crops. The investment was funded by a consortium of 10 Western development banks. This means that besides a business project it was also explicitly aimed to be a development project.

3 Data and Empirical Strategy

3.1 Sample

We use three rounds of survey data. Baseline data was collected prior to any plantation activities in 2010 by a research team of the University of Cape Town. For that round we have data on 78 villages and 4233 households, see Table 2. The data comprise a census of all households in these villages. The plantation operations started in 41 of these villages, creating a natural control 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 that we have panel observations for for the 2010-2012 survey waves. In 2015, we collected an additional wave of data. Data collection was by students of Njala University who received extensive training on the survey instrument by the researchers. We returned to all villages included in the 2012 data and interviewed 25 people per village. We first attempted to speak to people that were interviewed in 2012, but if this was not possible we randomly selected other heads of households. The survey instrument was designed to closely match the earlier rounds of data collection but also added some additional questions. In 2015 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 need data on the same people within each survey wave. Fortunately, the company assigned all households ID codes and identification cards. We use 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. Table 2 summarises the sample sizes for the cross-section and the panel. [2]

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 (observeable or not). The main assumption we have to make is that the villages would have had parallel trends had there been no investment. This is fundamentally untesteable. In the robustness section we use data on forest loss and vegetation (EVI) to show that trends were parallel pre-investment, which is reassuiring. 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 distance to the Rokel river was too large (darker in Figure 1), which the company used to pump water for the center pivots. Therefore, they are similar in characteristics that are likely to be predictive of yield. Furthermore, since all smallholder agriculture is rainfed distance to the Rokel river is unlikely to correlate with 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,

[2]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 A2 and Table A3 and are qualitatively similar to our main results.

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 labour shortages are driving our results, we examine if our outcomes are 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:

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

Our coefficient of interest is η_5, how labourers differ from non-labourers 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 (160'000 in Treated), or 13 USD (36 USD), far below the World Bank international poverty line of 1.25 USD per day [3]. 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[4]. 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 that farmers owned is 4, which could mean a household owns their 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 a house with a mud floor, reed and thatch walls and a tarp as roof. For livestock the value is around 0.25, equalling (for example) 2 goats and 5 chickens. Almost all households have access to arable land for cultivation, though almost all households have faced a food shortage in the previous year. Half of households had a birth in the previous year (one third in treated). 92% of households had a mosquito net in their house (80% in control). The participants are clearly very poor, have few assets and low food security. [5]

[3]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.

[4]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

[5]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 this. Any results regarding child deaths and births should be interpreted with this in mind.

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 almost half a standard deviation. For total income this is lower (0.35 SD) but still substantial and significant. This drop is largely driven by a large drop in agricultural income (See Table A4 and Table A5 for the effect on the four components of traditional income). We hypothesise that this is caused by an increase in the local labour price which makes it more difficult for households to hire in local labour, reducing agricultural production and thus sales. Our spillover analysis (see Table 6) and labourer analysis (see Table 7) confirm this hypothesis. Furthermore, in 2015 we asked households whether the price of labour 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 (0.16 SD) in housing quality. 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 does not hold here. Incidence of food shortages drops by 10% in the short run, but this is not different in treated villages: it appears that food security is unaffected in treated villages. Next, we look at three measures of health. We see a very large increase in the number of births, of 0.7 SD. It might be that the availability of a local health clinic is increasing the number of succesful births. 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 outreach programs ran by the company. Finally, we look at village-level inequality. Pre-treatment the gini coefficient was 0.15, which means very high equality. In the short run this increased to 0.3 and in treated villages it even increased to 0.4, showing a substantial increase in inequality because of the plantation. We show in Table 7 that this is caused by labourers' increased incomes.

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 2012 this has 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 an exceptionally good year for agriculture. Since the treated group relies on agriculture more, they would be more affected when returning to normal harvest levels. Figure 5 indicates that 2010 was a normal year for agriculture.

Next we estimate Model 1 for the long run, shown in Table 5. Generally, results are similar in the long 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 signficant 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 even slightly larger now (7% lower). We again see no effect on food security. For health we still see a large increase in number of births, though it is only marginally significant. We again see higher presence of bed nets, but this is not significant. Inequality increases like in the short run. The main effects we find for both the long and short run are a drop in income, lower access to land and some health improvements. For the latter two the link to the investment's work is clear: they are using village land and are providing some health services. For the income effect we have hypothesized that this is caused by an increase in the labour 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 0.25 SD higher income. If higher labour prices are indeed locally determined and spill over partially to neighbouring villages we would expect to find these results. For assets we see an increase in the number of assets further away over time, but not for housing quality or livestock. Access to land is higher further away from the investment, but it is not very high (1 SD distance leads to 1% higher access) and only significant at the 10% level. This is unsurprising as our treatment is defined by having land leased. That there is some small effect might indicate that households start farming in neighbouring 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.

Finally, so far we have been examining these effects across all households in a village. Next, we examine the effect separately for labourers and non-labourers in the long run. In treated villages about 40% of the households has labourers who have worked for the company in the past. Of this group, 40% was still employed when we collected data. Average length of employment was 14 months. We examine effects on labourer households in Table 7, using Model 3. Regarding traditional income, we again see that this drops in the long run, but this happens equally across labourers and non-labourers. However, this is not the case when examining full income (which includes salaries). When comparing labourers to the other households in the village, there is an increase in income of almost half a standard deviation. This explains the increase in inequality in Tables 4 and 5: labourer households gain while others lose out. When we move to the stock variables there is a consistent increase of about 0.2 SD, for number of assets, housing quality and Tropical Livestock Unit. Clearly labourers are able to transform their additional earnings into tangible assets. Furthermore, labourers do not have lower access to land or better health access compared to others in their village. This is unsurprising, as these are effects that affect all households equally (the health clinics are accessible for all households). There is one small effect that labourers are slightly more likely (4%) to have had food shortages in the previous year, though not very significant. It might be that labourers that are now working on the plantation are not producing their own food anymore causing food shortages for them. This also implies that there is not enough food available on the local food market.

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 examine whether attrition is systematic, we 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 untesteable, but it is reassuring when 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 traditional slash and burn agriculture which is highly labour-intensive. Therefore, forest loss is likely to correlate with increased agricultural production and income - 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 worlwide dataset contains extremely detailed (30m resolution) data on forest loss for the period 2000-2017. 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 3). We convert these pixels to the number of hectares lost per village per year and plot this out in Figure 4. The vertical black line represents the year that the company started their activities. Trends are clearly very similar 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 theorise 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 random subsample so it was never the intention to find everyone. In table A6 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-interacted variables), but these are not worrying as this does not indicate differing dropout. There is one worrying finding:

a higher traditional income before treatment leads to a lower chance of dropping out in the short run - 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, Fiala, and Martinez (2014). For this analysis we make alternative assumptions about the attriters. 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 attriters are assigned a low value, while treatment attriters are assigned a high value, thus zeroing out the negative treatment effect. For the Manski worst case the high (low) value is that groups 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 A7 for the short run only. Column 1 show the original treatment effect of the treatment*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 effects (on income, access to land and bed nets) the results from columns 3-5 are reassuring. In most cases the sign of all coefficients are the same, and for the 0.1 SD deviation these effects are significant as well. Note that these deviations represent large, systematic deviations on the characteristics of attriters for which we found no evidence in Table A6.

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 5 and examine the maximum yearly EVI. 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 labour 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). We study the impact of 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. 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 labour on the farm.

We use a differences-in-differences analysis to compare outcomes for communities within and outside the catchment area of the plantion 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 hypothesise that this is because the increased labour demand increases the labour price, making it too expensive to hire in labour, 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 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. We also see some improvements in health, which can likely be attributed to the health clinics the company has built. These effects hold in the long run (5 years) as well. When we examine labourers specifically, we see that they do benefit compared to non-labourers in their village. Their incomes rise and this translates into more tangible assets. This in turn increases within-village inequality.

We have hypothesised and given evidence that a portion of the impact is transmitted through local markets, especially the labour market. This is likely to hold for most investments: by definition they are looking to acquire land, and often hire local labour to lower transportation costs and to seek goodwill from the local community. This shows that to examine the full impact of an 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 these local markets. We leave this for future work.

Taken together, the results from this paper paint a bleak picture. While an increase in income for labourers is positive, only 40% of households provide labourers, and their gains do not outweigh the losses by non-labourer households. It is likely that this increased inequality and marginalisation increases conflicts (something which we cannot asses), but which earlier qualitative research has indicated as well. The investment as a whole appears to be a bad deal for recipient communities. Indeed, that almost 50% of land lease payments go to local elites rather than the land owners is telling.

References

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.). Freetown: Sierra Leone Network on the Right to Food. Retrieved from `http://www.brotfueralle.ch/fileadmin/deutsch/01{_}Service/Medien{_}Texte/Mediencommuniques/IndependentStudyReportAddaxFinal.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. doi: 10.1257/pol.5.2.1

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.). Oakland: The Oakland Institute.

Baxter, J. (2013). *Who is benefitting?* (Tech. Rep.). Freetown: Action for Large-Scale Land Acquisition Transparency.

Benn, J. (2005). Making regression tables from stored estimates. *The Stata Journal*, *5*(3), 288–308.

Benn, J. (2007). Making regression tables simplified. *The Stata Journal*, *7*(2), 227–244.

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. Retrieved from https://doi.org/10.1016/j.landusepol.2017.12.016 doi: 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. doi: 10.1080/03066150.2015.1119119

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., ... Sun,

M. (2015). *Agricultural investment and rural transformation: a case study of the Makeni bioenergy project in Sierra Leone* (Tech. Rep.). Stockholm: Stockholm Environment Institute. Retrieved from 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*. Retrieved from https://doi.org/10.1016/j.rse.2017.06.031 doi: 10.1016/j.rse.2017.06.031

GoSL. (2015). *National land policy* (Tech. Rep.). Freetown: Government of Sierra Leone.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, A., Tyukavina, A., ... 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. doi: 10.1016/j.worlddev.2016.10.007

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.* Retrieved from https://cran.r-project.org/package=raster

Højsgaard, S., & Halekoh, U. (2018). *doBy: Groupwise Statistics, LSmeans, Linear Contrasts, Utilities.* Retrieved from https://cran.r-project.org/package=doBy

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

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. Retrieved from http://dx.doi.org/10.1016/j.landusepol.2015.06.008 doi: 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. (2018). *The Online Public Database on Land Deals.*

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

MAFFS. (2009). *National Sustainable Agriculture Development Plan 2010-2030* (Tech. Rep.). Freetown: Ministry of Agriculture, Forestry and Food Security. Retrieved from http://maffs.gov.sl/resources/maffs-policies/112-national-sustainable-agriculture-development-plan-2010-2030

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

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. Retrieved from http://link.springer.com/10.1007/s41412-016-0020-5 doi: 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. Retrieved from http://dx.doi.org/10.1080/01436597.2015.1044960 doi: 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. doi: 10.1080/14754835.2015.1032219

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. doi: 10.1111/joac.12102

Millar, G. (2016b). Local experiences of liberal peace: Marketization and emergent conflict dynamics in Sierra Leone. *Journal of Peace Research*, *53*(4), 569–581. doi: 10.1177/0022343316632580

Millar, G. (2017). For whom do local peace processes function? Maintaining control through conflict management. *Cooperation and Conflict*, *52*(3), 293–308. doi: 10.1177/0010836716671757

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.* Retrieved from 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.). Bern, Montpellier, Hamburg, Pretoria: Centre for Development and Environment, University of Bern; Centre de coopération internationale en recherche agronomique pour le développement; German Institute of Global and Area Studies; University of Pretoria; Bern Open Publishing. doi: 10.7892/boris.85304

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.* Vienna, Austria. Retrieved from 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.* London: Allen & Unwin.

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. Retrieved from http://dx.doi.org/10.1016/j.landusepol.2015.01.034 doi: 10.1016/j.landusepol.2015.01.034

SiLNoRF. (2014). *Annual Monitoring Report on the operations of Addax Bioenergy* (Tech. Rep.). Freetown: Sierra Leone Network on the Right to Food. Retrieved from 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.). New York: United Nations Development Programme.

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

Wickham, H., Francois, R., Henry, L., & Müller, K. (2017). *dplyr: A Grammar of Data Manipulation.* Retrieved from `https://cran.r-project.org/package=dplyr`

7 Figures

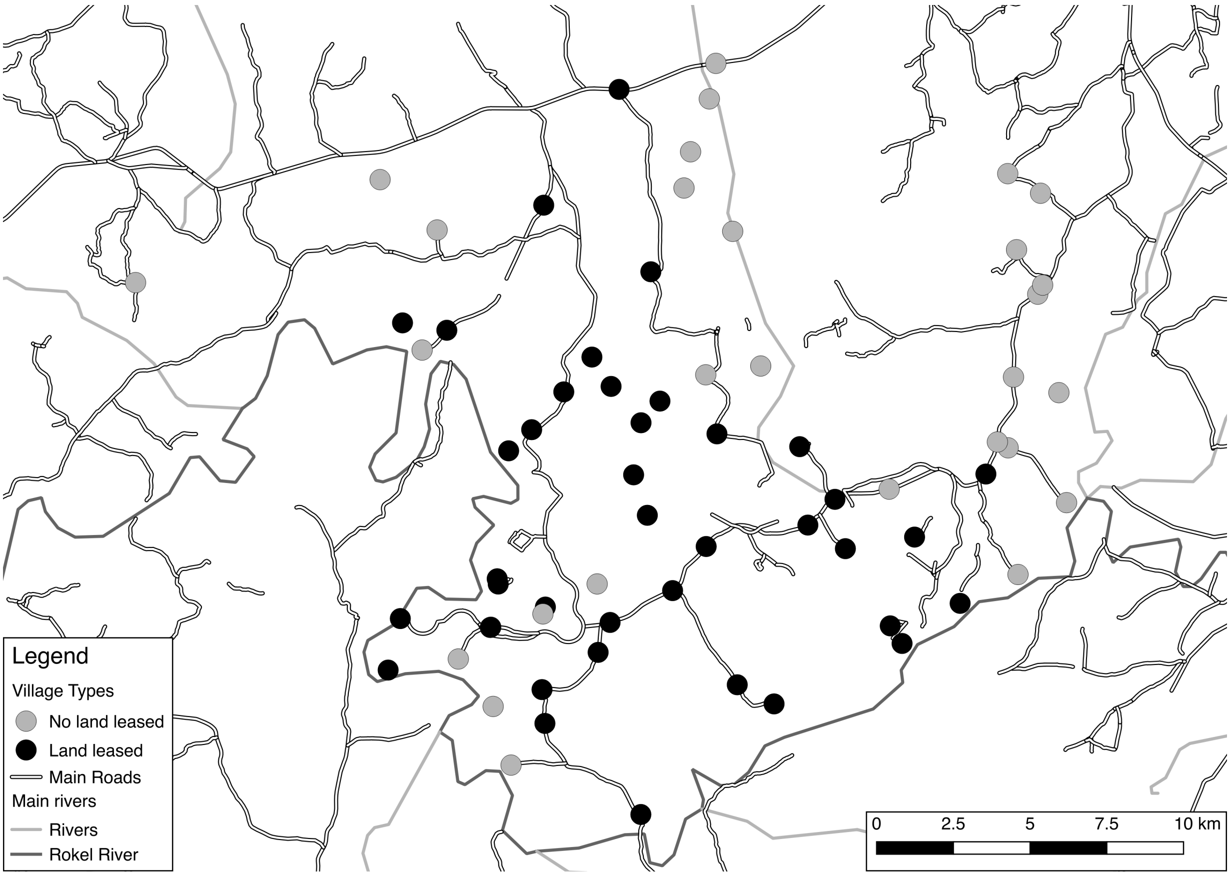

Figure 1: Village Locations

Shows location of all villages examined in this study. Source: survey data

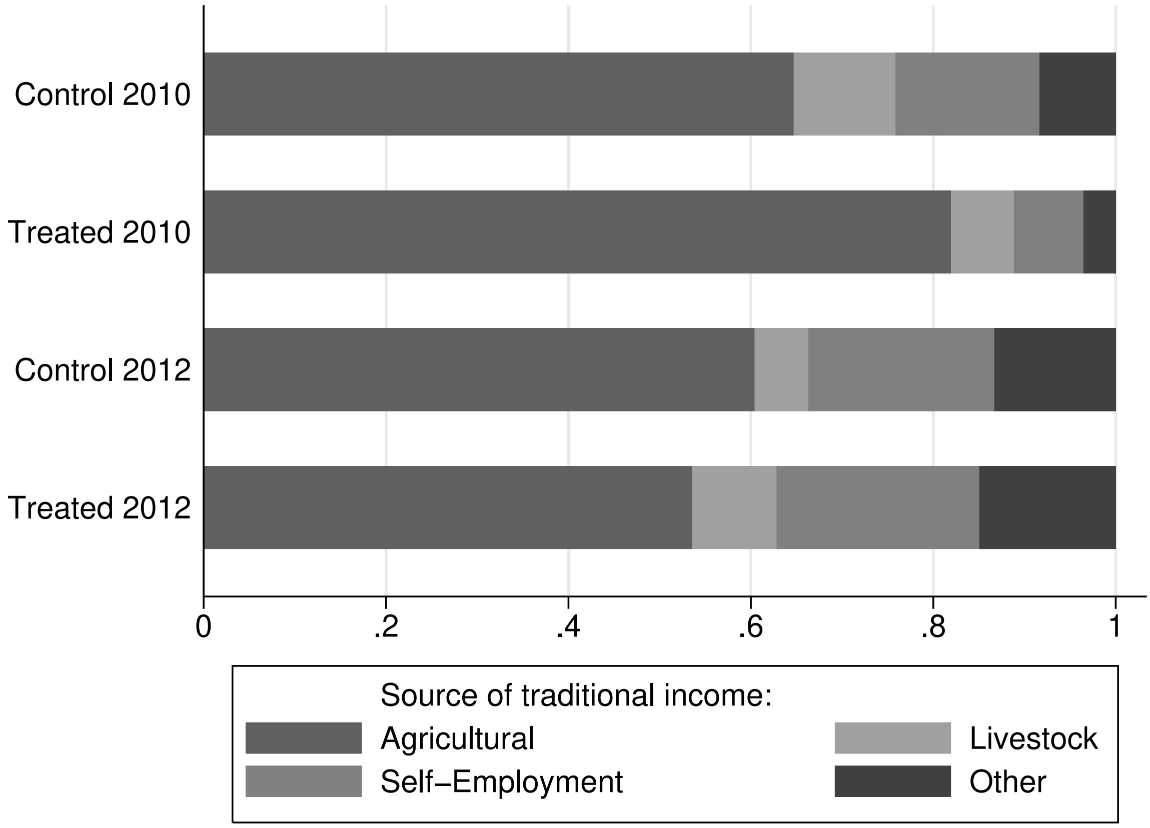

Figure 2: Income Proportions (2010-2012)

Shows proportions of traditonal 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

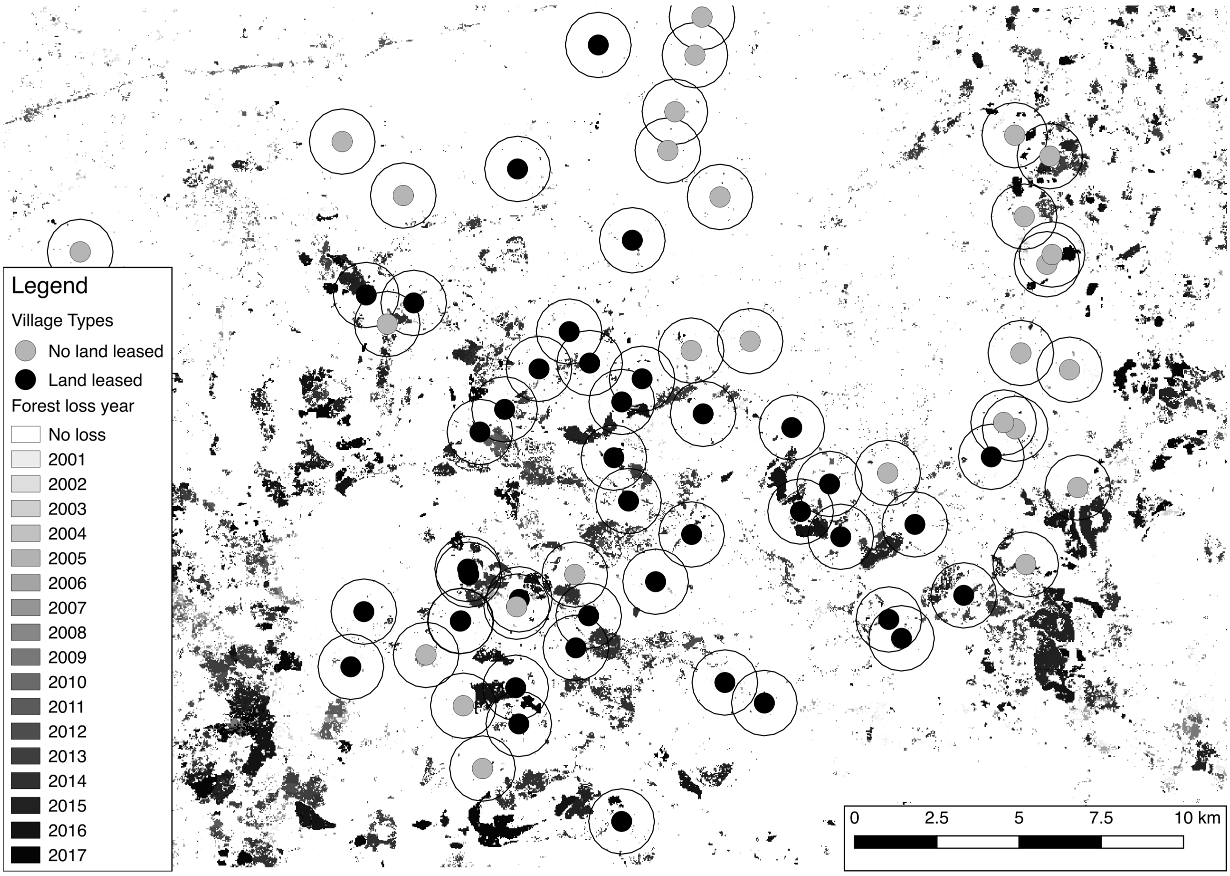

Figure 3: Forest Loss Map

This map shows forest loss over time around the sample villages (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

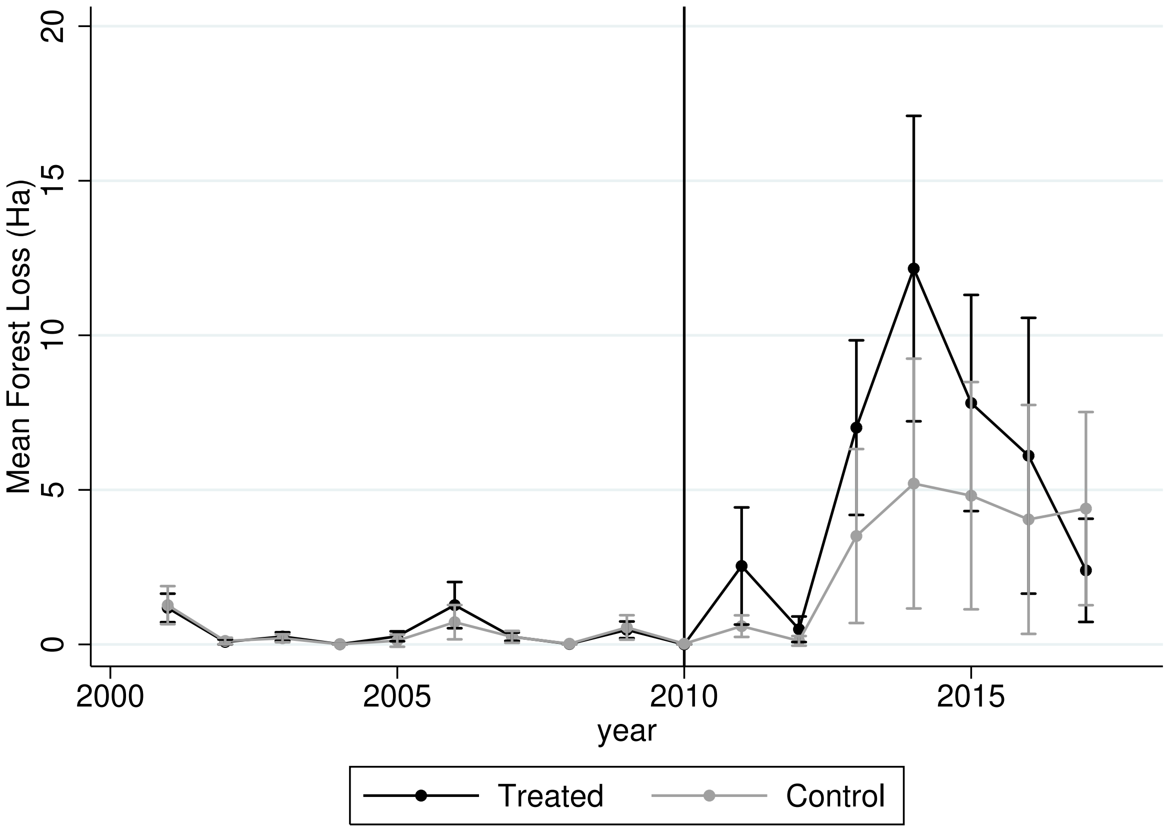

Figure 4: Forest Loss 2001-2017

Graph shows 95% confidence intervals. Graph represent average forest loss in circles with 1km radius around villages. Source: Hansen/UMD/Google/USGS/NASA

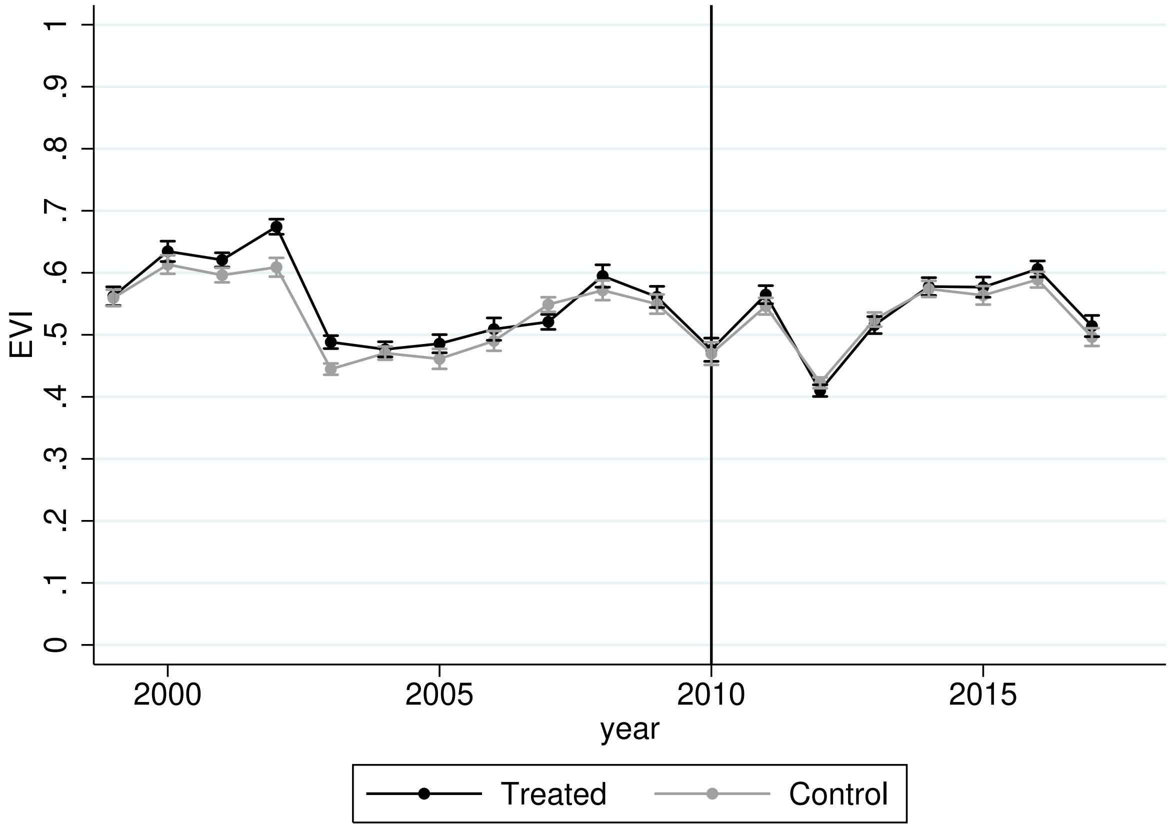

Figure 5: Greenness (EVI) 1999-2017

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

8 Tables

Table 1: Qualitative studies overview

Author	Type	Cluster N	N	Methods	Findings
Anane and Abiwu (2011)	PR	12	NA	SI, SSI and FG	The CSR 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		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 down. Increased poverty. 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 CSR 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	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 exisiting social cleavages
Marfurt, Käser, and Lustenberger (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 insitutions 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

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						
Respondents	1415	2818	1790	3034	649	1118
Villages	37	41	71	47	39	36
Panel						
Respondents			915	2240	99	529
Villages			28	40	15	34

Panel observations based on ID match. Source: survey data

Table 3: Descriptive Statistics

		Control			Treatment		
	n	mean	sd	n	mean	sd	Diff
Traditional Income (Leones, IHS)	1004	11.51	3.87	1504	12.62	2.82	1.110**
# 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	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
Total births in last 12 months	537	0.54	0.67	2083	0.31	0.55	-0.231**
Total child deaths 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, 1-33)	(5) Tropical Livestock unit	(6) Access to land	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Bed net in household (=1)	(10) Inequality (gini)
Treated	0.301***	0.345***	−0.028	0.087	−0.007	−0.003	0.003	−0.777***	−0.118***	−0.047*
	(0.090)	(0.105)	(0.087)	(0.066)	(0.033)	(0.003)	(0.005)	(0.243)	(0.040)	(0.027)
Short Run	−0.198	0.051	0.033	0.259***	−0.042	−0.031**	−0.101***	−0.735***	−0.082**	0.147***
	(0.160)	(0.150)	(0.116)	(0.050)	(0.031)	(0.012)	(0.013)	(0.260)	(0.040)	(0.036)
Treated * Short Run	−0.477**	−0.349**	−0.043	−0.091	0.146	−0.049***	−0.006	0.657**	0.135***	0.102**
	(0.180)	(0.172)	(0.130)	(0.066)	(0.089)	(0.018)	(0.018)	(0.264)	(0.050)	(0.048)
Constant	0.000	0.000	0.000	0.000	0.030*	0.997***	0.988***	0.000	0.920***	0.150***
	(0.075)	(0.085)	(0.068)	(0.040)	(0.018)	(0.002)	(0.004)	(0.237)	(0.022)	(0.022)
Observations	3762	3762	6310	3914	3470	6082	6068	3886	6302	96
# Clusters	67	67	68	68	67	68	68	60	68	

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 9 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) Traditional income (IHS)	(2) Full Income (IHS)	(3) # Assets	(4) House Quality (Score, 1-33)	(5) Tropical Livestock unit	(6) Access to land	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Bed net in household (=1)	(10) Inequality (gini)
Treated	0.196	0.211	−0.065	0.221**	0.130	−0.002	0.002	−0.312	−0.042	−0.042
	(0.134)	(0.140)	(0.156)	(0.107)	(0.087)	(0.002)	(0.011)	(0.291)	(0.068)	(0.055)
Long Run	0.099	0.211	0.259	0.545***	0.220	−0.021	−0.115**	−0.969***	−0.061	0.016
	(0.177)	(0.171)	(0.181)	(0.146)	(0.171)	(0.013)	(0.046)	(0.327)	(0.070)	(0.065)
Treated * Long Run	−0.532***	−0.437**	0.096	0.247	−0.259	−0.072**	0.072	0.600*	0.052	0.150**
	(0.192)	(0.185)	(0.192)	(0.165)	(0.186)	(0.033)	(0.047)	(0.341)	(0.081)	(0.073)
Constant	−0.006	−0.006	0.000	0.005	0.001	1.000***	0.990***	0.000	0.847***	0.131**
	(0.118)	(0.118)	(0.146)	(0.086)	(0.063)	(0.000)	(0.010)	(0.277)	(0.058)	(0.051)
Observations	748	748	1256	796	990	1202	1190	846	1242	54
# Clusters	44	44	49	47	45	48	48	41	48	

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 9 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, 1-33)	(5) Tropical Livestock unit	(6) Access to land	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Bed net in household (=1)
Distance	−0.119*	−0.142**	−0.097*	0.055	−0.005	0.001	0.003	0.439***	0.033**
	(0.059)	(0.066)	(0.051)	(0.032)	(0.012)	(0.001)	(0.003)	(0.150)	(0.014)
Short Run	−0.186	0.063	0.033	0.251***	−0.043	−0.031**	−0.101***	−0.811***	−0.082**
	(0.126)	(0.111)	(0.095)	(0.047)	(0.030)	(0.011)	(0.013)	(0.185)	(0.039)
Distance * Short Run	0.302***	0.290***	0.226***	0.059*	−0.027	0.016*	−0.010	−0.448**	0.020
	(0.107)	(0.090)	(0.081)	(0.034)	(0.020)	(0.008)	(0.014)	(0.160)	(0.031)
Constant	−0.005	−0.006	0.000	−0.008	0.030	0.997***	0.988***	0.074	0.920***
	(0.065)	(0.072)	(0.062)	(0.040)	(0.018)	(0.002)	(0.003)	(0.166)	(0.021)
Observations	1288	1288	1830	1444	980	1750	1784	516	1828
# Clusters	28	28	28	28	27	28	28	22	28

OLS regressions. Only control villages included. Normalized variables (not columns 6, 7 and 9). 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 labourers (2010-2015)

	(1) Traditional income (IHS)	(2) Full Income (IHS)	(3) # Assets	(4) House Quality (Score, 1-33)	(5) Tropical Livestock unit	(6) Access to land	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Bed net in household (=1)
Labourer	0.051	0.043	−0.107	−0.391***	−0.182**	0.003	−0.003	0.250*	−0.059*
	(0.088)	(0.100)	(0.064)	(0.086)	(0.071)	(0.003)	(0.009)	(0.144)	(0.032)
Treated	0.177	0.194	−0.025	0.334***	0.201**	−0.003	0.004	−0.420	−0.018
	(0.149)	(0.158)	(0.159)	(0.111)	(0.093)	(0.003)	(0.012)	(0.296)	(0.072)
Long Run	0.099	0.211	0.259	0.545***	0.220	−0.021	−0.115**	−0.969***	−0.061
	(0.177)	(0.171)	(0.181)	(0.146)	(0.171)	(0.013)	(0.046)	(0.328)	(0.070)
Long Run * Treated	−0.461**	−0.608***	0.026	0.192	−0.338*	−0.054*	0.091*	0.649*	0.052
	(0.195)	(0.192)	(0.194)	(0.165)	(0.189)	(0.029)	(0.048)	(0.351)	(0.083)
Labourer * Long Run	−0.183	0.444***	0.214**	0.180*	0.205**	−0.045	−0.044*	−0.114	0.000
	(0.145)	(0.144)	(0.080)	(0.106)	(0.099)	(0.029)	(0.025)	(0.174)	(0.039)
Constant	−0.006	−0.006	0.000	0.005	0.001	1.000***	0.990***	0.000	0.847***
	(0.119)	(0.118)	(0.146)	(0.087)	(0.063)	(0.000)	(0.010)	(0.277)	(0.058)
Observations	748	748	1250	794	990	1202	1190	846	1242
# Clusters	44	44	49	47	45	48	48	41	48

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 9). Labourers 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

Family	Variable Definition
Traditional Income	Sum of Agricultural and livestock sales, self-employment and other income (including remittances) in January of that year
Total Income	Traditional income, but also with company's land payments (2012 only) and salaried income (2015 only)
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. Also, a 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.
Livestock	Tropical livestock unit on number of livestock owned, based on cattle, goats, sheep, pigs, rabbits and chickens. The tropical livestock unit is often equated to a 250 kg animal (Jahnke, 1982).
Land Access	Answer to question 'Do you currently have access to land for cultivation?' (yes/no)
Food Security	Answer to the question 'Was there a shortage of food in the household at any time last year?' (yes/no)
Health	# births and whether a bed net is present in the household (yes/no)
Inequality	Gini coefficient based on traditional income of villages with at least 6 observations

Table A2: 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, 1-33)	(5) Tropical Livestock unit	(6) Access to land	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Bed net in household (=1)
Treated	0.296***	0.342***	−0.026	0.095	0.014	−0.004	0.002	−0.803***	−0.120***
	(0.093)	(0.110)	(0.087)	(0.068)	(0.031)	(0.002)	(0.005)	(0.252)	(0.042)
Short Run	−0.188	0.083	0.060	0.268***	−0.031	−0.030**	−0.103***	−0.733**	−0.074**
	(0.166)	(0.150)	(0.111)	(0.051)	(0.026)	(0.012)	(0.014)	(0.289)	(0.035)
Treated * Short Run	−0.481**	−0.384**	−0.063	−0.105	0.125	−0.046**	0.000	0.647**	0.133***
	(0.189)	(0.175)	(0.126)	(0.068)	(0.091)	(0.019)	(0.019)	(0.294)	(0.047)
Constant	0.000	0.000	0.000	0.000	0.027**	0.997***	0.989***	0.000	0.921***
	(0.076)	(0.088)	(0.065)	(0.042)	(0.011)	(0.002)	(0.004)	(0.247)	(0.024)
Observations	3428	3428	5764	3578	3200	5570	5540	3582	5758
# Clusters	65	65	67	67	65	67	67	58	67

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 9). 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 A3: Long Run (2010-2015) effects of a Large-Scale Agricultural investment: stricter merge results

	(1) Traditional income (IHS)	(2) Full Income (IHS)	(3) # Assets	(4) House Quality (Score, 1-33)	(5) Tropical Livestock unit	(6) Access to land	(7) Food shortage (=1)	(8) Total births in last 12 months	(9) Bed net in household (=1)
Treated	0.215	0.233	−0.079	0.228*	0.110	−0.002	0.002	−0.397	−0.075
	(0.139)	(0.144)	(0.178)	(0.116)	(0.091)	(0.002)	(0.012)	(0.345)	(0.075)
Long Run	0.115	0.234	0.252	0.597***	0.202	−0.022	−0.100**	−1.023**	−0.087
	(0.189)	(0.182)	(0.189)	(0.149)	(0.182)	(0.014)	(0.045)	(0.390)	(0.075)
Treated * Long Run	−0.546***	−0.462**	0.079	0.175	−0.251	−0.062**	0.055	0.692*	0.085
	(0.199)	(0.197)	(0.202)	(0.168)	(0.198)	(0.027)	(0.047)	(0.401)	(0.088)
Constant	−0.003	−0.004	0.000	0.000	0.001	1.000	0.989***	0.000	0.870***
	(0.123)	(0.123)	(0.167)	(0.097)	(0.065)	(.)	(0.011)	(0.334)	(0.065)
Observations	690	690	1118	716	890	1076	1062	750	1108
# Clusters	43	43	47	46	45	47	47	39	47

OLS regressions. Standardized and centered on control group at baseline (not columns 6, 7 and 9). 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 A4: 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	0.409***	−0.123	−0.183	−0.135*
	(0.152)	(0.097)	(0.118)	(0.074)
Short Run	−0.224	−0.100	0.034	0.310***
	(0.186)	(0.103)	(0.069)	(0.068)
Treated * Short Run	−0.561***	0.003	0.160	0.074
	(0.206)	(0.125)	(0.122)	(0.093)
Constant	0.000	0.000	0.000	0.000
	(0.117)	(0.080)	(0.073)	(0.062)
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 A5: 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	0.142	0.164	0.176	−0.008
	(0.201)	(0.114)	(0.123)	(0.166)
Long Run	−0.323	0.089	0.596***	0.481***
	(0.240)	(0.117)	(0.108)	(0.142)
Treated * Long Run	−0.388	−0.103	−0.163	−0.148
	(0.260)	(0.143)	(0.155)	(0.169)
Constant	−0.008	−0.021	0.003	0.006
	(0.166)	(0.085)	(0.058)	(0.151)
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 A6: Attrition

	(1) Dropout SR	(2) Dropout LR
Traditional Income (Leones, IHS)	0.057*	0.052*
	(0.030)	(0.031)
# Assets	-0.070	0.015
	(0.069)	(0.114)
House quality (Score, 1-33)	-0.127**	-0.024
	(0.065)	(0.061)
Tropical Livestock Unit	0.043**	0.292
	(0.020)	(0.480)
Access to land=1	-0.422	0.000
	(0.805)	(.)
Food shortage (=1)=1	-1.178	0.000
	(0.917)	(.)
Total births in last 12 months	-0.523***	0.156
	(0.164)	(0.239)
Total child deaths in last 12 months	0.082	-0.094
	(0.305)	(0.385)
Bed net in household (=1)=1	0.267	0.661*
	(0.475)	(0.338)
Treated	-0.001	0.735
	(0.891)	(1.005)
* Traditional Income (Leones, IHS)	-0.152***	-0.043
	(0.045)	(0.039)
* # Assets	0.029	-0.061
	(0.098)	(0.130)
* House quality (score)	0.057	-0.025
	(0.074)	(0.068)
* Tropical Livestock Unit	0.124	-0.321
	(0.175)	(0.492)
* Access to land	0.000	0.000
	(.)	(.)
* Food Shortage	0.000	0.000
	(.)	(.)
* Total births	0.327	-0.144
	(0.220)	(0.266)
* Total child deaths	0.605	0.560
	(0.414)	(0.451)
* Bed net in household	-0.174	-0.616
	(0.553)	(0.409)
Constant	2.090	0.154
	(1.303)	(0.920)
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 A7: Bounds Analysis

	(1) Original	(2) Worst	(3) + (-) 0.5SD	(4) + (-) 0.25SD	(5) + (-) 0.1SD)
Traditional income (IHS)	-0.477**	0.514**	-0.132	-0.280**	-0.369***
	(0.180)	(0.244)	(0.152)	(0.139)	(0.134)
Full Income (IHS)	-0.349**	0.945***	0.022	-0.128	-0.218*
	(0.172)	(0.304)	(0.146)	(0.131)	(0.125)
# Assets	-0.043	1.788***	0.261**	0.116	0.028
	(0.130)	(0.342)	(0.112)	(0.098)	(0.092)
House Quality (Score, 1-33)	-0.091	1.448***	0.241**	0.103	0.021
	(0.066)	(0.312)	(0.092)	(0.072)	(0.062)
Tropical Livestock unit	0.146	-1.393***	-0.179**	-0.039	0.045
	(0.089)	(0.349)	(0.086)	(0.072)	(0.067)
Access to land	-0.049***	0.308***	0.018	-0.013	-0.031**
	(0.018)	(0.085)	(0.018)	(0.015)	(0.014)
Food shortage (=1)	-0.006	0.328***	0.070***	0.025	-0.002
	(0.018)	(0.075)	(0.018)	(0.016)	(0.017)
Total births in last 12 months	0.657**	-1.399***	-0.046	0.122	0.223
	(0.264)	(0.443)	(0.177)	(0.160)	(0.152)
Bed net in household (=1)	0.135***	-0.094*	-0.005	0.045	0.075**
	(0.050)	(0.055)	(0.041)	(0.037)	(0.036)

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