UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2016/120

Can commercial farming promote rural dynamism in sub-Saharan Africa?

Evidence from Mozambique

Steven Glover[1] and Sam Jones[2]

November 2016

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: The impact of large commercial farms on neighbouring smallholders in low-income sub-Saharan Africa remains controversial. Bringing evidence to a largely anecdotal debate, we deploy a dataset covering all commercial farms in Mozambique, linking them to a nationally representative survey of smallholders. We investigate: (i) the degree of selection of commercial farms into more/less favourable locations; and (ii) the effects of these farms on the incomes of neighbouring smallholders. We are unable to find robust evidence of either positive or negative spillovers from commercial farms. However, certain types of commercial agricultural ventures—outgrower schemes, plantations—appear to generate moderate local welfare gains.

Keywords: agriculture, commercial farming, Mozambique, structural transformation
JEL classification: O13, Q12, Q18

Tables: at the end of the paper.

Acknowledgements: We are grateful for helpful comments from Finn Tarp, Vincenzo Salvucci and participants at the 2015 World Bank conference on land and poverty (23–27 March). All errors and omissions are our own.

[1] Independent consultant, Mozambique; [2] University of Copenhagen, Denmark, sam.jones@econ.ku.dk.

This study has been prepared within the UNU-WIDER project on 'Inclusive growth in Mozambique—scaling-up research and capacity'.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-164-2

Typescript prepared by Sophie Richmond.

1 Introduction

The importance of agriculture in the process of social and economic development is established (Christiaensen et al., 2011; Diao et al., 2010; Johnston and Mellor, 1961; World Bank, 2008). However, there remains widespread concern that the agricultural sectors of many low-income sub-Saharan African countries continue to show sluggish growth, particularly relative to the service and natural resource sectors (Gollin et al., 2014; McMillan and Rodrik, 2011). In this context, attention has been given to the potential role of commercial agriculture to stimulate rural economies and contribute to aggregate growth. For instance, in apparent contrast to perspectives that consider smallholders as an essential (but not necessarily exclusive) focus of policy interventions (Chang, 2009; Dorward et al., 2004; Kydd et al., 2004), Collier and Dercon (2014) argue that a narrow focus on smallholders will not stimulate the change required to transform African economies overall. Rather, they advocate a more dynamic agriculture based on new forms of commercial organization involving 'hybrid models' of interaction between smallholders and commercial (farm) enterprises.

Commercial farming means different things to different people. At one extreme, producing for a market as opposed to for home consumption might be considered 'commercial'. At the other extreme we find capital- and input-intensive agribusiness entities that undertake modern industrial forms of agriculture on a large scale, taking advantage of economies of scale in production and distribution (Deininger and Byerlee, 2012). Between these extremes lie intermediate forms, including various types of outgrower schemes, with or without a counterpart corporate hub. These different models imply different forms of interaction between smallholders and any (outside) commercial entity.

Keeping these distinctions in mind, evidence regarding the impact of commercial farming on rural economies (and smallholders in particular) varies according to the particular type in focus. Large literatures address the impacts of increased commercialization (Barrett, 2008; Von Braun and Kennedy, 1994) as well as smallholder outgrower schemes, such as in cotton, coffee, cocoa, and tobacco (Bolwig et al., 2009; Jones and Gibbon, 2011; Oya, 2012). While debate surrounding such impacts remains vigorous, there is evidence that certain interventions in these domains can generate positive direct gains for participating smallholders (Dorward et al., 2004; Glover, 1984; World Bank, 2008). Nevertheless, there remains considerable uncertainty about the magnitude and nature of any dynamic spillovers from commercial activities to non-participating smallholders as well as to the broader local rural economy (Smalley, 2013).

With respect to the local economic impact of large commercial farms in low-income settings, little evidence exists beyond case studies (Cotula et al., 2009). However, it is precisely this form of commercial farming that has been most controversial in recent years. Popular and academic analysts have drawn attention to high levels of foreign investment in farmland across developing countries and the potential for negative effects on local producers (Borras et al., 2011; Cotula, 2013; Oxfam, 2011; White et al., 2012). This resonates with deeper historical concerns regarding inequality in land ownership and the marginalization of poorer rural farmers (Li, 2011). For this reason, such investments are often viewed with suspicion.

The aim of the present paper is to contribute more rigorous quantitative evidence to these debates regarding the impact of large commercial farms on local smallholders. As Deininger (2011) notes, the absence of hard evidence means that this debate has been largely speculative and dominated by ideological positions. Our analysis focuses on the case of Mozambique, which has received significant interest from foreign investors, primarily due to its perceived land abundance as well as active government and donor efforts to attract foreign agricultural investors

(Aabø and Kring, 2012; Cotula, 2013; Deininger and Byerlee, 2012). External data collated in Cotula (2013) suggests that usage rights of at least 2 million (and potentially up to 10 million) hectares (ha) of land was acquired by agricultural investors over the last decade. Additionally, schemes such as the Beira Agricultural Growth Corridor (BAGC) and ProSavana are ongoing, high-profile cases where a significant role has been assigned to large commercial farms owned (and managed) by external investors (Kaarhus, 2011; Nogueira, 2013).

Our analysis combines nationally representative agricultural survey data on over 6,000 smallholders collected in 2012 with a new micro-dataset covering all large farms in the country. These farms, which are by definition commercial, are simply defined as those currently cultivating an area of over 50 ha. We proceed in two steps. First, we investigate whether there is a systematic relationship between characteristics of these large farms and average smallholder or local characteristics. This is necessary to verify the potential role and direction of omitted variables (e.g. selection bias) in the impact analysis. We find evidence that operating commercial farms are more likely to be located near areas with lower than average education levels, urban areas, and farmers using more advanced production techniques.

Second, using the data on smallholders, we investigate the association between the presence of a large farm and various welfare outcomes, such as total household income (controlling for other determinants). We find weak evidence of a moderate positive association between the presence of a commercial farms and the welfare of local smallholders. However, this association does not appear to hold in general. Rather, it appears to relate primarily to commercial farms that produce permanent crops (e.g. fruit trees, sugar cane); also, we find that smallholders who are outgrowers benefit disproportionately from the presence of medium-sized commercial farms. Critically, these findings point to significant heterogeneity in the local economic effects of commercial farms rather than any kind of automatic spillover effects. The implication is that a more nuanced policy position is demanded with respect to promoting rural dynamism through new commercial models. In particular, the creation of suitable incentives for and regulation of outgrower schemes needs to be a priority.

The remainder of the paper is structured as follows: section 2 reviews insights from economic theory concerning the expected benefits for farmers operating in proximity to large commercial ventures. Section 3 introduces the data sources and provides a descriptive snapshot of the changes in large commercial agriculture in Mozambique between 2002 and 2012. Section 4 reports our analysis of patterns in large farm location decisions and outcomes for surrounding smallholder farmers. Section 5 concludes.

2 What do we know about large farms?

Much of the existing literature surrounding large-scale commercial agriculture has investigated the existence of scale economies or lack thereof (Barrett, 1996; Collier and Dercon, 2014; Deininger and Byerlee, 2012; Eastwood et al., 2010). Recent structural changes have challenged the conventional wisdom of a general inverse relationship between farm size and productivity in low-income countries. Technological innovations, improvements in labour monitoring, and greater vertical integration of global value chains appear to have augmented the potential contribution of scale economies in farming (Deininger and Byerlee, 2012), increasing the commercial viability of larger units. For instance, with the growth of mechanization in Indian agriculture following economic reforms of the 1990s, Foster and Rosenzweig (2011) find that larger farms generally are more efficient than smallholders, despite lower unit labour costs of the latter.

Aside from these important debates, less attention has been given to dynamic complementarities (or discords) between existing smallholders and larger farms. Case studies of plantations generally struggle to find consensus on their impacts on the local population, which can often be dependent upon the benevolence of the planation operator (Smalley, 2013) and the varying sets of economic, political, and social relations (McCarthy, 2010). Nonetheless, some theoretical guidance is provided from the wider literature, which can distinguish between direct and indirect impacts due to large farms.

2.1 Direct impacts

Wages from workers on large farms are a principal mechanism through which rural economies can benefit, particularly if they are spent or invested in activities with strong local multipliers. It is fair to expect that areas with many jobs created from commercial agricultural ventures are more likely to be 'dynamic' than otherwise, although the true impact will be felt in accordance with the workers' income elasticity of demand (presumed to be large in poor, rural settings) and the accompanying local supply response. The Food and Agriculture Organization (FAO, 2013) concludes that employment generation appears to be the main type of benefit from plantations, but questions the sustainability of created jobs. However, Smalley (2013) suggests that low wages and the monopolistic tendencies of agribusiness have prevented positive spillovers from taking place in some contexts.

Concerns have been raised that agricultural growth driven by large farms will have a low impact on poverty reduction. Noting that labour-intensive growth tends to be more pro-poor, Lipton (2012) argues that large farms in sub-Saharan Africa are usually capital intensive, even when labour is relatively abundant. Commercial farming operations have also been found to exclude local workers from the production process; rather, the few job opportunities created were often filled by more skilled migrants (Deininger et al., 2011; Li, 2011). Several reports on land acquisitions have argued that investors target the best lands available to them (FAO, 2013; Oxfam, 2011).[1] Combined with the threat of land dispossession (Cotula, 2013), this has generated warnings that the livelihoods of some rural communities may be threatened (Li, 2011).

Case studies from Mozambique reveal very different experiences from large farm investments. The development of sugar plantations has been seen as a catalyst for rural development in surrounding areas (Waterhouse et al., 2010). However, Borras et al. (2011) describe how one 30,000 ha sugarcane investment for a biofuel project disrupted livelihoods, threatened food security, and created intense land conflicts – only for the project to be cancelled before production started.

2.2 Indirect impacts

Commercial activities have several potential spillover effects that may affect the local population. The main mechanisms here include introduction and diffusion of new production techniques, as well as enhanced access to inputs and commercial opportunities. Technical change in smallholder agriculture is driven by learning-by-doing and learning-from-others (Foster and Rosenzweig, 1995), with imperfect knowledge of new and improved technologies being significant barriers to adoption and productivity improvements (Feder et al., 1985). As familiarity with new modes of production increases, barriers to adoption and adaptation diminish.

---

[1] There is not a consensus on this view; among others, Deininger et al. (2011) argue the contrary.

Learning-by-doing can be stimulated by large farms via demonstration effects. For instance, seasonal or temporary workers on commercial farms may learn about new crops or production techniques through their wage work and then apply these to their home production (supported by their wages). In turn, further spillovers may arise through diffusion of this knowledge through the wider community (see Besley and Case, 1993). Additionally, the presence of (large) commercial enterprises may expose communities to commercialization opportunities, including access to new markets and enhanced inputs.

3 Methodology

This section outlines the sources of data and empirical methods employed in the analysis.

3.1 Data sources

Our analysis is based primarily on the 2012 integrated agricultural survey (*Inquérito Agrícola Integrado*, IAI; see MINAG, 2012), which is the latest in a series of nationally representative agricultural surveys undertaken in Mozambique (previously known as the *Trabalhos de Inquérito Agrícola,* TIAs). As described in Kiregyera et al. (2008), these surveys have used a widely accepted sampling methodology; also, they have served as the basis for a range of academic scholarship on Mozambican agriculture (Cunguara and Moder, 2011; Donovan and Massingue, 2007; Mabiso et al., 2014). The principal focus of these surveys has been on small and medium farmers. However, a database of all 'large' farms in Mozambique has been maintained, containing basic characteristics of these farms, including their size, main crops, and location. Specifically, in each district sampled, all active large agricultural entities were identified by government officials and subsequently visited by the data collection team. These encompass privately owned (corporate) farms, large farms run by household units, and a few other organizational forms, such as state-owned farms and cooperatives. This database, from which we retain all farms of over 50 ha, has not been analysed previously.

In order to characterize changes in commercial agriculture in Mozambique over recent years, Table 1 compares the characteristics of large farms operating in 2002 and 2012.[2] The data shows that large farms increased in number over this period, approximately doubling in total size from 235,800 ha in 2002 to 441,500 ha in 2012. Notably, most growth has been driven by land lying fallow and unused. Indeed, 87 per cent of the increase in total area is uncultivated land, of which a considerable majority is in the south of the country. This suggests that land speculation and hoarding has been taking place on a large scale, especially around the capital, Maputo, where the surrounding province contains 79 per cent of the increase in fallow land area. Since the southern provinces have some of the poorest quality soils and least abundant water sources in the country, the commercial draw of Maputo is likely to account for a large part of these trends.

Two further points merit comment. First, the number and size of large farms remains small relative to other actors in the sector. For instance, as of 2012, we estimate there were around 3.9 million small and medium farmers, cultivating a total of around 4.9 million ha. The crop area of large farms accounts for just 2.2 per cent of the 5.3 million ha of cultivated land in Mozambique (MINAG, 2012). Second, our estimates of the expansion of commercial farming are substantially lower than those of other sources. For example, Deininger et al. (2011), Friis and Reenberg

[2] Data from 2002 comes from the equivalent large farms survey, which forms part of the 2002 TIA (see MINAG, 2012).

(2010), and the Land Matrix (2014) all report evidence that more than 2 million ha of land usage rights have been acquired by large agricultural projects in recent years. This disparity is principally due to the fact that our data encompasses operating farms, rather than entities merely granted usage rights. **This focus is important since it is reasonable to expect that spillovers from commercial farming only emerge once some cultivation (operation) takes place.** Additionally, commercial biofuel and forestry plantations, which have been large drivers of land acquisitions, are not included in the survey design.

In order to analyse the local impacts of these large farms, we match them with the IAI micro-data file, which contains detailed information on small and medium-sized farmers. Specifically, all smallholders located in the same *posto administrativo* (henceforth PA) as a large farm(s) are considered as hypothetically 'treated' by the latter (i.e. potentially exposed to spillover effects). The intuition is that the principal diffusion mechanisms of interest (see section 2) should take place in the immediate proximity of a large farm. The final dataset thus contains observations for 6,736 small and medium farmers, matched to 156 large farms. These large farms are found in 68 of the 338 PAs sampled, yielding 1,483 neighbouring ('treated') smallholders. Of the large farms, 22 are found in the North, 51 in the Centre and 83 in the South; also, 44 are larger than 1,000 ha, 8 are larger than 10,000 ha, yet all but 2 of these possess vast areas of fallow land.

3.2 Empirical strategy

Previous literature suggests that commercial agriculture is drawn to areas with better agro-ecological conditions and market access. For Mozambique, Glover et al. (2014) suggest that large farms have generally located towards areas with better infrastructure, higher levels of population, greater wealth, and higher education levels. Others argue that large-scale land acquisitions are increasingly taking place on land with access to water resources, fertile soils, proximity to markets, and in peri-urban areas (Aabø and Kring, 2012; Kaarhus, 2011; Oxfam, 2011). The key point is that if there is geographical self-selection of larger farms (into more commercially productive areas), then we should expect tests of association between the presence of large farms and smallholder outcomes to be confounded. Put differently, when comparing smallholders from areas with and without large farms (which is our principal interest), it may be necessary to control for the determinants of large farm location decision.

The above considerations suggest two null hypotheses of interest:

$H0_I$: The location of commercial farms in a given PA is independent of smallholder, community and spatial characteristics.

$H0_{II}$: After controlling for observable determinants of commercial farm location, there are no systematic effects on smallholder outcomes.

The first null hypothesis ($H0_I$) investigates whether the location of large farms is systematic in relation to fixed or given smallholder characteristics (i.e. characteristics unlikely to be endogenous to the large farm). We investigate this via a probit regression. Specifically, information on households and local communities (from the micro-data) is aggregated to the level of the PA to form a set of explanatory covariates (plus regional dummy variables). The regressand is a binary variable that takes a value of one if a large farm is actively operating in the

PA. This procedure is useful as it allows calculation of inverse probability (IP) weights (see below).[3]

The second null hypothesis ($H0_{II}$) considers whether the welfare of smallholders in PAs that contain large farms is systematically different to that of other smallholders. This is the principal question of interest. The feasibility of being able to test $H0_{II}$ depends on our ability to control for variation in the location of large farms. If we suspect that $H0_{I}$ is false, selection bias is likely to be present in naïve estimates of $H0_{II}$. To address this we adopt two strategies. First, we control for systematic variation in the location of large farms, as well as other determinants of rural household incomes, by using an extensive list of control variables (see section 4). Second, we employ IP weights estimated from the analysis of $H0_{I}$ to re-weight the data, placing greater weight on PAs that are observationally similar regardless of their large farm status. Put differently, employing the IP weights means we place less weight on PAs at the extreme ends of the cumulative probability distribution of containing a large farm, which amounts to a doubly robust procedure (e.g. recommended by Imbens and Wooldridge, 2009).

Throughout the analysis we concentrate on three smallholder outcomes: total income (in natural logarithms), crop income (in natural logarithms), and input use.[4] Since the dependent variable is fixed (constant) at the PA level, we begin by running the analysis for $H0_{II}$ at this more aggregated level. We also consider a number of extensions. In particular, use of a single dummy variable to capture the 'treatment' of a proximate large farm is likely to be crude. Thus, for each PA we divide the total land holding of large farms into terciles (50–263 ha, 263–2,000 ha, > 2,000 ha), yielding a set of three dummy variables. In addition to undertaking the analysis at the PA level, we also use the raw micro-data in which smallholders are the unit of analysis. This has the advantage of permitting inclusion of more detailed fixed effects (e.g. at the district level), which absorbs time-invariant, unobserved determinants of the large farm location at the chosen level. It also permits investigation of effect heterogeneity. For instance, to estimate the marginal effect of large farms on smallholders who are outgrowers we include an outgrower interaction term. As such, the most general specification we employ is given by:

$$y_{ij} = \alpha + \beta \text{farm}_{ij} + \gamma \text{outgrower}_{\text{ij}} + \delta \left(\text{farm}_{ij} \times \text{outgrower}_{ij} \right) + \lambda' Z_{ij} + \mu_j + \epsilon_{ij}$$

where y is the dependent variable; 'farm' and 'outgrower' are dummy variables which respectively capture whether farmer i resides in a PA containing a large farm and/or is an outgrower; Z is a vector of control variables; μ_j is a fixed effect pertaining to location *j*; and ϵ_{ij} is a residual error term, assumed to be white noise. In other specifications we replace the single 'farm' dummy with disaggregated versions that capture different characteristics of such farms.

Before proceeding, we explicitly recognize some limitations of the proposed approach. First, the empirical strategy lacks a definitive means to address all potential sources of omitted variables bias. Our inclusion of a wide range of control variables and fixed effects, as well as probability re-weighting, is intended to minimize selection bias. However, this cannot be established with certainty and reflects the nature of the data – that is, the absence of any experimental or quasi-experimental component in the matching of smallholders and larger farms.

---

[3] For an overview of the use of IP weights see Curtis et al. (2007); also Imbens and Wooldridge (2009).

[4] For the latter we employ a simple dummy variable that takes the value of one if the smallholder uses at least one of pesticide, fertilizer, or improved seeds.

Second, our metric of the connection between smallholders and large farms is imprecise due to the absence of geocoding of the large farms or more direct measures of such relations. In practice, we define smallholders as 'neighbours' of large farms if they are co-located in the same PA, which is the lowest official administrative unit at which we have matched location information in both datasets. This measure may be overly expansive (i.e. captures smallholders that cannot reasonably be expected to be affected by a large farm); if so, our results will be conservative (attenuated).

Third, we are unable to account for impacts associated with the (possible) relocation of smallholders driven by commercial investment (i.e. forms of land dispossession). This is pertinent if smallholders have not resettled locally but have moved to another district or urban centre. Fourth, we recognize that positive spillovers from large farms may only materialize slowly. This means it may be unrealistic to find significant effects where commercial farms have only started operations recently. Due to data limitations, we are not able to fully control for commercial farm tenure. This may add further noise to our main explanatory variable (attenuation bias).

4 Results

This section sets out our main results, in which we distinguish between analysis undertaken at the PA level and that at the smallholder level. Throughout, we focus on selected coefficients of interest. Full details are available on request from the authors.

4.1 PA-level analysis

Column I of Table 2 reports average marginal effects from the probit regression model (i.e. tests of $H0_I$), which takes each PA as the unit of analysis (N = 336). Only selected covariates are reported. The full list of controls includes means, calculated from the micro-data and aggregated to the PA level, of the following: the total area of smallholder farms, the area harvested, the number of smallholders; characteristics of the household head (age, gender, years of education, literacy, agricultural training, formal employment status); the size of the household; whether agriculture is the main economic activity of the household; whether the household is located in a peri-urban area; whether the smallholder has a formal land title; use of animal traction and manure; use of crop rotation, intercropping, and line sowing techniques; whether the smallholder received extension advice, hired labour, or is a member of an agricultural association. Village-level controls (again averaged to the PA level) include: whether the village has a mill, access to electricity, a shop selling agricultural inputs, year-long public transport access, distance from a paved road, knowledge of the 1997 land law, and whether land law information has been disseminated in the village.

The results point to a number of important and statistically significant determinants of large farm locations. Areas with lower education levels, a higher share of peri-urban households and lower dependence on agriculture as a main economic activity appear *more likely* to contain a large farm. Such farms also appear to be attracted towards areas where agrarian techniques of smallholders are more developed (e.g. use intercropping, line sowing, and manure application) and where electricity is available. Tests of the joint insignificance of these explanatory variables are easily rejected, implying we can reject $H0_I$ – location decisions by large farms are not random. Rather, they appear systematically related to local smallholder, community and/or spatial characteristics. This finding is consistent with previous literature. It is reasonable to expect that commercial agricultural entities prefer to locate in peri-urban areas for commercial and logistical purposes, as well as to take advantage of a larger pool of (underemployed) labour. We also find that larger farms are more likely to be found where average human capital levels are somewhat lower (*ceteris paribus*).

Columns II to VII of Table 2 move on to investigate $H0_{II}$, continuing to treat the PA as the unit of analysis (N = 336). Each column represents an OLS (ordinary least squares) regression of the indicated dependent variable on the full list of covariates employed in column I, plus the dummy variable for whether the PA contains a large farm. In other words, they correspond to aggregated versions of equation (1) imposing the preliminary restriction $\gamma = \delta = 0$ and including (ten) agricultural zone and (three) regional fixed effects. Columns II, IV, and VI use sample survey weights in the estimation; columns III, V, and VII deploy IP weights given by:

$$\hat{\pi}_i = \left[\frac{F_i}{\hat{p}_{i,F}}\right] + \left[\frac{(1 - F_i)}{\left(1 - \hat{p}_{i,F}\right)}\right]$$

where F_i takes a value of one if there is a large farm in the PA and zero otherwise; and $\hat{p}_{i,F}$ is the estimated probability that the PA has a large farm, as estimated from column I of Table 2. Intuitively, $\hat{\pi}_i$ is larger where the PA does in fact contain a large farm but takes a low probability from the probit model, or where the PA does not have a large farm but is expected to be more likely to have one according to the model. To reduce the impact of excessively large or small estimated weights, we follow Lee et al. (2011) and trim the estimated weights at the 95 per cent and 5 per cent percentiles. This procedure is not material to the qualitative nature of our results.

Three main findings can be highlighted. First, looking across all outcomes, there is no unambiguous or systematic association between the presence of a large farm and the mean welfare of smallholders in the same PA. None of the estimated β coefficients are significantly different from zero when using sample weights. When IP weights are applied, only total income is significant at the 5 per cent level. However, this result may be spurious since we have run multiple tests. Indeed, a conservative Bonferonni correction would require the null hypothesis to be rejected with a likelihood of under 2 per cent in order to achieve a standard (one-off) 10 per cent confidence level. Calculated on this basis, the previous result is no longer significant.

Second, the same general pattern of results is found when 'large farms' are split into size terciles. These findings are summarized in Table 3, which reports the primary farm coefficients of interest. Of the 18 coefficients shown, just 3 are significant at conventional levels and 2 of these are borderline significant (probability > 1 per cent). The strongest result ($\hat{\beta}$ = 0.33; see column II) indicates a material positive association between the presence of smaller large (commercial) farms and smallholder outcomes – namely, smallholder incomes are around 40 per cent (0.33 log points) larger in PAs with smaller large farms compared to PAs with no commercial farms. This estimate is significant at the 1 per cent level (probability = 0.21 per cent) and therefore is more resilient to the suggested correction for multiple testing. Moreover, the additional tests in the footer of the same column indicate we can reject the hypothesis that the welfare effects associated with the different terciles are equal.

Third, as with the probit results, in all models summarized in Tables 2 and 3 the coefficient estimates for other covariates are of the expected direction and the overall coefficient of variation (adjusted R-squared) is satisfactory. Additionally, estimated coefficients are similar (in direction and magnitude) when either the sample or IP weights are used. This provides comfort that the model is adequate and the IP weights are neither distortionary nor do they place undue weight on an 'unusual' sub-domain of the sample space.

4.2 Smallholder analysis

The previous results focused on average outcomes for each PA and did not take full advantage of the information in the micro-data. Thus we proceed by running the same models for outcomes (incomes and inputs) now taking the smallholder as the unit of analysis (N > 6,200). Table 4 reports results from a range of different estimates for the total and crop income (log.) outcomes. Each cell refers to a different outcome (dependent variable), choice of fixed effects and choice of weights (sample or IP weights, where the latter are taken from the probit estimates in Table 1, as before). All cells uniquely report estimates of β, based on models that continue to impose the restriction $\gamma = \delta = 0$; this corresponds to the overall marginal effect of the presence of a commercial farm. Thus, columns I and II in the first (top) row are the micro-data counterparts to the estimates reported in columns II and III of Table 2; similarly, columns III and IV are counterparts to the estimates reported in columns IV and V of Table 2. The second row of Table 4 replaces agricultural zone and regional fixed effects with agricultural zone and provincial effects; the third row employs district fixed effects. The remaining rows report similar estimates for the small, medium, and larger farm terciles (in each case estimated separately).

The results of Table 4 are broadly similar to those undertaken at the PA level (Tables 2 and 3), as might be expected. A cursory reading would suggest that commercial farms, taken as a whole, are associated with a positive and significant impact on total incomes equal to around 0.12 log points (13 per cent) when IP weights are employed (column II, Table 4). Similarly, small farms appear to be the main contributor to this overall effect. However, two caveats must be placed on these findings. First, the estimated statistical significance of these coefficients is always greater than 1 per cent, meaning that these results turn insignificant after correction for multiple testing. Second, inclusion of fixed effects at lower levels of aggregation (e.g. districts) weakens both the magnitude and significance of the findings, especially for the smaller farms tercile.[5] This implies that the previously noted results could well be driven by unobserved aspects of large farm location and establishment decisions. So, despite the larger numbers of observations available in the micro-data, the evidence for a systematic positive welfare effect of the presence of a commercial farm is weak at best. Similarly, none of the estimates in columns III or IV of the Table 4 (crop income) are robust given the multiple models considered.

Although tentative, the findings reported thus far suggest there may be material differences between commercial farms of different sizes as regards the magnitude (and significance) of effects on local smallholders. That is, the effects associated with commercial farms may be heterogeneous. We explore this hypothesis further in two aspects. First, we consider whether there are differential impacts associated with the *type* of agriculture practised. To do so, we replace the commercial farm dummies with variables capturing whether there is a commercial farm in the PA that cultivates either annual crops, permanent crops, pasture and/or fallow land. These results are summarized in columns I, II, V, and VI of Table 5, in which each column reports the estimated coefficients on these dummy variables from a single regression, including a full set of covariates and district fixed effects (as per Table 4).[6] Different columns correspond to alternative outcomes and choices of weights. The remaining columns consider a modification of these 'treatment variables', employing the share of the area of commercial farms in the PA that are used by each type of agricultural activity. The results clearly support the notion that commercial farms have heterogeneous effects. A consistent finding is a moderate positive effect

[5] The opposite holds for the medium tercile with IP weights *?*; however, this may be spurious.

[6] Note that more than one of these four dummy variables can take a value of one at the same time.

of permanent crops;[7] however, none of the other practices (types) show a robust association with smallholder outcomes. Unsurprisingly, a corollary is that fallow land (perhaps speculative land investment) is not associated with positive local welfare effects.

Second, we consider the outgrower interaction channel indicated in equation (1). The logic is to test a specific but potentially selective channel through which commercial farming operations may impact on smallholder welfare.[8] With reference to equation (1), we focus on the linear combination of estimates for $\beta + \delta$, which captures the marginal effect of the presence of a commercial farm for the sub-set of smallholders in the PA that are outgrowers.[9] To do so, we re-run the same models reported in Table 4 including these additional terms; our results are summarized in Table 6, which repeats the format of Table 4. The main finding is that for both total and crop incomes, there is a positive effect associated with the medium tercile of larger farms only (under both sample and IP weights). While the significance of these estimates is borderline in some instances, the pattern of results is very consistent – outgrowers in PAs that contain a medium-sized larger farms tend to have total and crop incomes that are around 40 per cent larger than those of other smallholders, holding fixed other factors (including district fixed effects). Note that this effect excludes the direct welfare effect of being an outgrower, captured by the parameter λ in equation (1). Interestingly, however, the latter term is consistently positive and significant, which is in line with positive effects associated with outgrower schemes in Mozambique and elsewhere (Bolwig et al., 2009).

Before concluding, two points merit emphasis. First, results for use of improved inputs show no distinct or meaningful patterns; therefore they have been omitted.[10] Second, across all specifications there are no statistically significant findings with respect to the largest tercile of commercial farms (after taking account of multiple testing). As such, for Mozambique, there is no empirical basis for claims that extremely large (mega-) commercial farms generate positive spillovers for local smallholders.

5 Conclusions

Developments in the agricultural sector in Mozambique will be of upmost importance to its future trajectory. Within this, there is little doubt that modern commercial entities will play an important role in addition to smallholders. A recent wave of external investment in Mozambican farmland has indicated the potential for large-scale restructuring of social and economic relations in some rural areas, raising important policy questions around the appropriate role of capital investment in agriculture. This paper empirically investigated the potential of commercial agriculture (large farms) to stimulate broader agricultural development and rural dynamism. Specifically, we considered whether there is a systematic association between smallholder outcomes (i.e. incomes and input use) and the presence of large (commercial) farms, defined as a farm over 50 ha in size, in the same locale.

---

[7] The 'area share' marginal effect estimates can be interpreted as the predicted welfare gain if there was a commercial farm in the PA that exclusively undertook*?* permanent crops.

[8] Note, in this set-up we do not explicitly consider the potential endogeneity of choosing to be an outgrower. This is not critical, however, as one impact of commercial farms is that smallholders switch to producing cash crops. This direct effect is not included in our analysis (see text); thus, these estimates may be conservative.

[9] 'Outgrower' is defined as a dummy variable that takes a value of one if the smallholder cultivates cotton or tobacco, which are the primary crops for which outgrower schemes exist in Mozambique.

[10] Available on request from the authors.

At the outset, we recognized that the location of such large farms is unlikely to be random – theoretically, there are likely to be strong incentives for large farms to locate close to existing infrastructure and available labour, as well as on the most productive land. In turn, this implied that selection bias would likely complicate naïve estimates of the relationship between smallholder outcomes and the presence of a large farm. The first stage of our analysis explored these selection issues directly. We estimated a probit model for the presence of a large farm, taking the unit of analysis as the PA (a low-level administrative unit below that of the district; N = 336). We found evidence that currently operating commercial farms favour areas that are: peri-urban, where the local population has fewer years of education; and where local farmers practise a wider range of agricultural techniques. Most importantly, we were able to reject the hypothesis that the location decisions of commercial farms are random.

The second stage of our analysis considered the association between the presence of a large farm and a range of smallholder outcomes. Selection bias was addressed in three ways – via inclusion of a full set of control variables (including those used in the first stage); deploying IP weights estimated from the probit regression, placing greater weight on observations that are comparable aside from the presence of a large farm; and employing fixed effects (including down to the district level). We recognized that these methods are partial and may not fully address all sources of omitted variables bias (OVB). In addition, our metric of the 'connection' between a smallholder and a large farm is simply based on co-location and therefore does not fully capture distance or other related linkages. Thus, while the former challenge (OVB) might reasonably be expected to bias our estimates upward, the latter would be expected to generate attenuation bias. These work in opposite directions, but the net effect of possible sources of bias remains unknown. Even so, our estimates remain a useful first approximation and provide an important quantitative contribution to the existing literature.

Our results showed no clear, general association between smallholder outcomes and the presence of a large farm. Under some specifications moderate positive effects appeared – for example when using IP weights we found that, on average, smallholders in PAs with a large farm have 13 per cent higher income than smallholders in other PAs (holding fixed other determinants of income). Also, this effect appeared to be driven by the smallest tercile of large farms (those of 50–263 ha). Nonetheless, we noted that these results typically were not robust to appropriate corrections for multiple testing and, therefore, may be fragile.

We found stronger indications of effect heterogeneity. Large farms growing permanent crops (fruit tree plantations) show a more consistent positive association with smallholder outcomes versus all other types of agriculture. As there are few value-chain linkages between plantation agriculture and smallholders (e.g. outgrowing arrangements), employment generation and wages may be important drivers of this result. Also, the presence of a (medium tercile) large farm appears to particularly benefit smallholders who are cash crop producing outgrowers and therefore more integrated into value chains. This is separate from the more general finding that incomes for outgrowers tend to be higher than incomes for non-outgrowers.

Three broader implications follow from these findings. First, although our evidence is not as complete as we might wish, there is no evidence of a material negative effect of larger commercial farming operations on existing smallholder outcomes.[11] At the same time, and second, there is little evidence of a systematic positive impact. Consequently, while commercial farms may well provide an important contribution to the agricultural sector as a whole (i.e. to

[11] As noted in section 3, we recognize that our analysis may exclude certain relocation effects, such as where smallholders migrate out of the PA due to the establishment of a large farm.

aggregate output), hopes that a large number of smallholders will *typically* benefit from the presence of any kind of commercial farm in their neighbourhood cannot be supported for Mozambique. More simply, one cannot expect that direct or indirect positive spillover effects will occur automatically from all forms of commercial farming.

Third, evidence pointing to a positive outgrower channel, as well as differences in effects between types of commercial farms, point to the relevance of heterogeneity in the local economic impacts of such farms. Benefits from commercial farms may well accrue to some local smallholders, but these are likely to be conditional on the specific type of agriculture practised by the commercial farm (or other characteristics), the nature of smallholders in the immediate locale, and/or whether direct production or labour linkages between commercial farms and smallholders are established, such as through outgrower schemes. This indicates that future research would do well to focus on the specific circumstances, including the contribution of public regulation and oversight (Jones and Gibbon, 2011) that determine when and how smallholders benefit from the introduction of new forms of commercial agriculture. It also indicates that policy needs to move beyond simplistic notions of promoting commercial farming and, rather, focus on the incentives for and regulation of outgrower schemes.

References

Aabø, E., and T. Kring (2012). 'The Political Economy of Large-scale Agricultural Land Acquisitions: Implications for Food Security and Livelihoods/Employment Creation in rural Mozambique'. UNDP Working Paper 2012-004. New York: UNDP.

Barrett, C.B. (1996). 'On Price Risk and the Inverse Farm-size Productivity Relationship'. *Journal of Development Economics*, 51(2): 193–215.

Barrett, C.B. (2008). 'Smallholder Market Participation: Concepts and Evidence from Eastern and Southern Africa'. *Food Policy*, 33(4): 299–317.

Besley, T., and A. Case (1993). 'Modeling Technology Adoption in Developing Countries'. *American Economic Review Papers and Proceedings*, 83(2): 396–402.

Bolwig, S., P. Gibbon, P., and S. Jones (2009). 'The Economics of Smallholder Organic Contract Farming in Tropical Africa'. *World Development*, 37(6): 1094–104.

Borras, S.M., D. Fig, and S.M. Suárez (2011). 'The Politics of Agrofuels and Mega-land and Water Deals: Insights from the ProCana Case, Mozambique'. *Review of African Political Economy*, 28(128): 215–34.

Chang, H.J. (2009). 'Rethinking Public Policy in Agriculture: Lessons from History, Distant and Recent'. *Journal of Peasant Studies*, 36(3): 477–515.

Christiaensen, L., L. Demery, and J. Kuhl (2011). 'The (Evolving) Role of Agriculture in Poverty Reduction – An Empirical Perspective'. *Journal of Development Economics*, 96(2): 239–54.

Collier, P., and S. Dercon (2014). 'African Agriculture in 50 Years: Smallholders in a Rapidly Changing World?' *World Development*, 63: 92–101.

Cotula, L. (2013). *The Great African Land Grab? Agricultural Investments and the Global Food System*. London: Zed Books.

Cotula, L., S. Vermeulen, R. Leonard, and J. Keeley (2009). 'Land Grab or Development Opportunity? Agricultural Investment and International Land Deals in Africa'. Technical paper. Rome/London: FAO/IFAD/IIED. Available at: http://pubs.iied.org/12561IIED/ (accessed March 2014).

Cunguara, B., and K. Moder (2011). 'Is Agricultural Extension Helping the Poor? Evidence from rural Mozambique'. *Journal of African Economies*, 20(4): 562–95.

Curtis, L.H., B.G. Hammill, E.L. Eisenstein, J.M. Kramer, and K.J. Anstrom (2007). 'Using Inverse Probability-weighted Estimators in Comparative Effectiveness Analyses with Observational Databases'. *Medical Care*, 45(10): S103–S107.

Deininger, K. (2011). 'Challenges Posed by the New Wave of Farmland Investments'. *Journal of Peasant Studies*, 38(2): 217–47.

Deininger, K., and D. Byerlee (2012). 'The Rise of Large Farms in Land Abundant Countries: Do They Have a Future?' *World Development*, 40(4): 701–14.

Deininger, K., D. Byerlee, J. Lindsay, A. Norton, H. Selod, and M. Stickler (2011). 'Rising Global Interest in Farmland: Can It Yield Sustainable and Equitable Benefits?' Technical paper. Washington, DC: World Bank. Available at: http://siteresources.worldbank.org/DEC/Resources/Rising-Global-Interest-in-Farmland.pdf (accessed February 2014).

Diao, X., P. Hazell, and J. Thurlow (2010). 'The Role of Agriculture in African Development'. *World Development*, 38(10): 1375–83.

Donovan, C., and J. Massingue (2007). 'Illness, Death, and Macronutrients: Adequacy of Rural Mozambican Household Production of Macronutrients in the Face of HIV/AIDS'. *Food & Nutrition Bulletin*, 28(2 Suppl.): S331–S338.

Dorward, A., J. Kydd, J. Morrison, and I. Urey (2004). 'A Policy Agenda for Pro-poor Agricultural Growth'. *World Development*, 32(1): 73–89.

Eastwood, R., M. Lipton, and A. Newell (2010). 'Farm Size'. In R. Evenson and P. Pingali (eds), *Handbook of Agricultural Economics*, vol. IV, pp. 3323–97. Amsterdam: North Holland.

FAO (2013). 'Trends and Impacts of Foreign Investment in Developing Country Agriculture: Evidence from Case Studies'. Technical paper. Rome: FAO.

Feder, G., R.E. Just, and D. Zilberman (1985). 'Adoption of Agricultural Innovations in Developing Countries: A Survey'. *Economic Development and Cultural Change*, 33(2): 255–98.

Foster, A.D., and M. Rosenzweig (1995). 'Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture'. *Journal of Political Economy*, 103(6): 1176–209.

Foster, A.D., and M. Rosenzweig (2011). 'Are Indian Farms Too Small? Mechanization, Agency Costs, and Farm Efficiency'. Unpublished manuscript. Available at: www.bu.edu/econ/files/2009/09/Rosenzweig_Seminar-Paper.pdf (accessed May 2014).

Friis, C., and A. Reenberg (2010). *Land Grab in Africa: Emerging Land System Drivers in a Teleconnected World.* Copenhagen: GLP-IPO. Available at: http://www.globallandproject.org/arquivos/GLP_report_01.pdf (accessed January 2014).

Glover, D.J. (1984). 'Contract Farming and Smallholder Outgrower Schemes in Less-developed Countries'. *World Development*, 12(11–12): 1143–57.

Glover, S., V. Salvucci, and S. Jones (2014). 'Where Is Commercial Farming Expanding in Mozambique? Evidence from Agricultural Surveys'. DNEAP, Discussion Paper 82. Maputo, Mozambique: National Directorate of Studies and Policy Analysis (DNEAP).

Gollin, D., D. Lagakos, M.E. Waugh (2014). 'The Agricultural Productivity Gap'. *Quarterly Journal of Economics*, 129(2): 939–93.

Imbens, G.W., and J.M. Wooldridge (2009). 'Recent Developments in the Econometrics of Program Evaluation'. *Journal of Economic Literature* 47(1): 5–86.

Johnston, B.F., and J.W. Mellor (1961). 'The Role of Agriculture in Economic Development'. *American Economic Review*, 51(4): 566–93.

Jones, S., and P. Gibbon (2011). 'Developing Agricultural Markets in Sub-Saharan Africa: Organic Cocoa in Rural Uganda'. *Journal of Development Studies*, 47(10): 1595–618.

Kaarhus, R. (2011). 'Agricultural Growth Corridors Equals Land-grabbing? Models, Roles and Accountabilities in a Mozambican Case'. Paper presented at the International Conference of Global Land Grabbing, 6–8 April. Brighton, UK. Available at: http://www.future-agricultures.org/publications/search-publications/global-land-grab/conference-papers-2/1126-agricultural-development-corridors-equals-land-grabbing-models-roles-and-accountabilities-in-a-moz/file (accessed February 2014).

Kiregyera, B., D. Megill, D. Eding, and B. José (2008). 'A Review of National Agricultural Information Systems in Mozambique'. DNEAP Discussion Paper 64E. Maputo, Mozambique: National Directorate of Studies and Policy Analysis (DNEAP).

Kydd, J., A. Dorward, J. Morrison, and G. Cadisch (2004). 'Agricultural Development and Pro-poor Economic Growth in Sub-Saharan Africa: Potential and Policy'. *Oxford Development Studies*, 32(1): 37–57.

Land Matrix (2014). Land Matrix Mozambique. Available at: http://landmatrix.org/en/get-the-detail/by-target-country/mozambique (accessed February 2014).

Lee, B.K., J. Lessler, and E.A. Stuart (2011). 'Weight trimming and Propensity Score Weighting'. *PloS One*, 6(3): e18174.

Li, T.M. (2011). 'Centering Labor in the Land Grab Debate'. *Journal of Peasant Studies*, 38(2): 281–98.

Lipton, M. (2012). 'Learning from Others: Increasing Agricultural Productivity for Human Development in Sub-Saharan Africa'. UNDP Working Paper 2012-007. New York: UNDP.

Mabiso, A., B. Cunguara, and R. Benfica (2014). 'Food (In)Security and Its Drivers: Insights from Trends and Opportunities in Rural Mozambique'. *Food Security,* 6(5): 649–70.

McCarthy, J.F. (2010). 'Processes of Inclusion and Adverse Incorporation: Oil Palm and Agrarian Change in Sumatra, Indonesia'. *Journal of Peasant Studies*, 37(4): 821–50.

McMillan, M., and D. Rodrik (2011). 'Globalization, Structural Change and Productivity Growth'. NBER Working Paper 17143. Cambridge, MA: National Bureau of Economic Research.

MINAG (Ministry of Agriculture of the Republic of Mozambique) (2012). *Inquérito Agrícola Integrado* (IAI 2002 and 2012 datasets). Maputo, Mozambique.

Nogueira, I. (2013). 'Agricultural Systems with Pro-poor Orientation in Mozambique? ProSAVANA and the Forgotten Risks of Contract Farming'. Paper presented at UNU-WIDER Conference on Inclusive Growth in Africa: Measurements, Causes, and Consequences, 20–21 September, Helsinki, Finland. Available at: http://www1.wider.unu.edu/inclusivegrowth/sites/default/files/IGA/Nogueira_0.pdf (accessed October 2013).

Oxfam (2011). *Land and Power: The Growing Scandal Surrounding the New Wave of Investments in Land.* Oxford: Oxfam. Available at: http://policy-practice.oxfam.org.uk/publications/land-and-power-the-growing-scandal-surrounding-the-new-wave-of-investments-in-l-142858 (accessed December 2011).

Oya, C. (2012). 'Contract Farming in Sub-Saharan Africa: A Survey of Approaches, Debates and Issues'. *Journal of Agrarian Change*, 12(1): 1–33.

Smalley, R. (2013). 'Plantations, Contract Farming and Commercial Farming Areas in Africa: A Comparative Review'. Future Agricultures Working Paper 055. Brighton, UK: Future Agricultures Consortium.

Von Braun, J., and E.T. Kennedy (1994). *Agricultural Commercialization, Economic Development, and Nutrition*. Baltimore, MD: Johns Hopkins University Press.

Waterhouse, R., G. Lauriciano, and S. Norfolk (2010). 'Social Analysis of Selected Projects—Issues Note and Case Studies: Large-scale Land Acquisition for Agricultural Production'. Technical paper. Available at: http://www.open.ac.uk/technology/mozambique/sites/www.open.ac.uk.technology.mozambique/files/pics/d128185.pdf (accessed February 2014).

White, B., S.M. Borras, R. Hall, I. Scoones, and W. Wolford (2012). 'The New Enclosures: Critical Perspectives on Corporate Land Deals'. *Journal of Peasant Studies,* 39(3–4): 619–49.

World Bank (2008). *World Development Report 2008: Agriculture for Development.* Washington, DC: World Bank. Available at: http://siteresources.worldbank.org/INTWDR2008/Resources/WDR_00_book.pdf (accessed October 2016).

Table 1: Summary of 'large farms' in Mozambique, in thousands of ha (2002 and 2012)

	North		Centre		South		Total		
	2012	2002	2012	2002	2012	2002	2012	2002	Δ'12-'02
No. large farms	22	39	51	34	83	44	156	117	39
Land use in 000s of ha									
Annual crops	2.4	5.0	33.0	20.9	21.9	28.7	57.3	54.7	2.6
Perm. crops	1.4	26.5	51.9	15.2	6.6	2.2	59.9	43.9	16.0
Pastures	8.1	2.1	39.7	47.7	22.0	19.2	69.9	69.0	0.8
Fallow land	57.1	10.2	39.0	53.7	158.4	3.3	254.5	67.2	187.3
Total	69.0	43.9	163.7	137.5	208.8	53.4	441.5	234.8	206.7
as % total									
Annual crops	*3.5*	*11.5*	*20.2*	*15.2*	*10.5*	*53.8*	*13.0*	*23.3*	*-10.3*
Perm. crops	*2.1*	*60.5*	*31.7*	*11.0*	*3.1*	*4.1*	*13.6*	*18.7*	*-5.1*
Pastures	*11.8*	*4.8*	*24.3*	*34.7*	*10.5*	*35.9*	*15.8*	*29.4*	*-13.6*
Fallow land	*82.7*	*23.3*	*23.8*	*39.0*	*75.9*	*6.1*	*57.6*	*28.6*	*29.0*

Notes: 'Large farms' refers to all farms with some operational activity over 50 ha in size; total farm size disaggregated by land use (individual farms can cover multiple categories); first row is the simple number of such farms.

Source: Authors' estimates from 'large farms' database TIA2002, IAI2012 (MINAG, 2012).

Table 2: Aggregate regression results

	Probit	Total income (ln)		Crop income (ln)		Input use (dummy)	
	Survey	Survey	IP	Survey	IP	Survey	IP
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
PA contains a large farm	-	-0.03	0.13**	0.01	0.08	-0.03	-0.03
		(0.07)	(0.06)	(0.07)	(0.08)	(0.02)	(0.02)
Smallholder area cultiv., 1st harvest (ha)	0.07	0.43	0.42**	0.07	0.10	0.08	0.08
	(0.24)	(0.26)	(0.19)	(0.30)	(0.29)	(0.11)	(0.07)
Smallholder area cultiv., 2nd harvest (ha)	0.00	0.39***	0.22*	0.49***	0.44***	-0.16***	-0.09***
	(0.10)	(0.14)	(0.12)	(0.16)	(0.13)	(0.04)	(0.03)
Household head education	-0.10**	0.02	0.13**	-0.08	-0.03	0.01	0.01
	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)	(0.01)	(0.02)
Agriculture is household activity (%)	-0.49*	-0.16	0.12	0.24	0.79**	-0.02	0.06
	(0.27)	(0.27)	(0.25)	(0.36)	(0.32)	(0.10)	(0.07)
Member of farming association (%)	0.10	1.38***	1.16***	0.56	1.39**	-0.12	-0.11
	(0.41)	(0.41)	(0.38)	(0.48)	(0.60)	(0.13)	(0.14)
Uses intercropping techniques (%)	0.33*	-0.48***	-0.49***	0.10	-0.06	-0.15*	-0.13**
	(0.17)	(0.16)	(0.18)	(0.21)	(0.27)	(0.07)	(0.05)
Uses line sowing techniques (%)	0.27**	0.07	0.11	0.13	0.16	0.24***	0.19***
	(0.10)	(0.13)	(0.13)	(0.13)	(0.15)	(0.05)	(0.04)
Uses animal traction (%)	-0.13	-1.04***	-0.79***	-0.92***	-1.00***	-0.11	0.10
	(0.23)	(0.26)	(0.23)	(0.30)	(0.27)	(0.14)	(0.06)
Uses manure (%)	1.09***	1.34***	0.67**	0.86**	1.01***	0.70***	0.65***
	(0.32)	(0.42)	(0.31)	(0.41)	(0.35)	(0.15)	(0.10)
Household head has other employ. (%)	-0.28*	0.82***	0.92***	-0.02	-0.04	-0.10	-0.07
	(0.17)	(0.23)	(0.19)	(0.29)	(0.25)	(0.07)	(0.06)
Smallholders in urban zone (%)	0.25**	0.21*	-0.14	-0.17	-0.05	-0.05	0.01
	(0.11)	(0.12)	(0.11)	(0.15)	(0.16)	(0.04)	(0.03)
Village is connected to electricity (%)	0.26***	-0.01	0.06	-0.24	-0.20	0.02	0.02
	(0.09)	(0.13)	(0.12)	(0.15)	(0.17)	(0.04)	(0.03)
Village has a mill (%)	-0.13*	0.15	0.04	0.13	-0.03	0.03	0.01
	(0.08)	(0.09)	(0.08)	(0.11)	(0.12)	(0.04)	(0.03)
Village has a farming inputs shop (%)	0.13	0.17	0.17	0.52***	0.42***	0.15**	0.04
	(0.13)	(0.14)	(0.11)	(0.16)	(0.15)	(0.07)	(0.04)
No. of observations	336	336	336	336	336	336	336
Chi2 / F statistic	77.36	16.39	12.44	8.26	4.76	7.42	22.58
R-squared (pseudo / adjusted)	0.42	0.67	0.44	0.42	0.46	0.56	0.51

significance: * 10%; ** 5%; *** 1%

Notes: Unit of observation is the PA, implying that variables are local averages of smallholder/village characteristics; selected coefficients shown only; column (I) reports average marginal effects of the determinants of location of large farms; columns (II) through (VII) are OLS regressions for the indicated outcome; regression weights are either survey design or IP weights, derived from column (I); all models include regional fixed effects (North, Centre, South); OLS regressions include agricultural zone fixed effects; robust standard errors, clustered at the district level, reported in parentheses.

Source: Authors' calculations from IAI2012 dataset (MINAG, 2012).

Table 3: Modified version of aggregate regressions, by large farm terciles

	Total income (ln)		Crop income (ln)		Input use (dummy)	
	Survey	IP	Survey	IP	Survey	IP
Smallest tercile (50–263 ha)	0.10	0.33***	-0.00	0.19	0.02	-0.01
	(0.10)	(0.11)	(0.11)	(0.14)	(0.04)	(0.03)
Medium tercile (263–2,000 ha)	-0.29**	-0.07	-0.17	-0.20	-0.07*	-0.06
	(0.13)	(0.10)	(0.11)	(0.16)	(0.04)	(0.04)
Largest tercile (> 2,000 ha)	0.04	0.10	0.12	0.20	-0.03	-0.04
	(0.08)	(0.08)	(0.08)	(0.12)	(0.02)	(0.03)
No. of observations	336	336	336	336	336	336
R-squared (adjusted)	0.67	0.46	0.43	0.47	0.56	0.51
Farm vars. all equal (prob.)	0.05	0.01	0.07	0.11	0.17	0.49
Farm vars. all zero (prob.)	0.10	0.00	0.14	0.12	0.12	0.33

significance: * 10%; ** 5%; *** 1%

Notes: Table reports extended estimates of the models reported in columns (II)–(IV) of Table 2 using large farm terciles (as opposed to a single dummy variable); terciles are included simultaneously; all other aspects of the models are as per Table 2; only selected coefficients shown; robust standard errors, clustered at the district level, reported in parentheses.

Source: Authors' calculations from IAI2012 dataset (MINAG, 2012).

Table 4: Micro-data regression results, aggregated and by size tercile

		Total income (log. × 100)				Crop income (log. × 100)			
		Survey (I)		IP (II)		Survey (III)		IP (IV)	
		Beta	Prob.	Beta	Prob.	Beta	Prob.	Beta	Prob.
a. All large farms	AgZ + Region	1.88	(0.80)	13.72	(0.04)	4.79	(0.50)	10.70	(0.16)
	AgZ + Prov.	4.99	(0.53)	10.63	(0.12)	10.92	(0.15)	9.86	(0.21)
	District	5.10	(0.41)	12.72	(0.04)	11.99	(0.07)	17.55	(0.03)
b. Smallest tercile	AgZ + Region	22.64	(0.03)	23.79	(0.04)	11.21	(0.21)	16.34	(0.18)
	AgZ + Prov.	27.33	(0.01)	20.53	(0.10)	19.90	(0.05)	11.62	(0.34)
	District	14.32	(0.09)	0.90	(0.93)	10.57	(0.28)	-3.93	(0.73)
c. Medium tercile	AgZ + Region	-14.65	(0.20)	7.70	(0.39)	-10.45	(0.32)	5.05	(0.73)
	AgZ + Prov.	-11.35	(0.34)	8.64	(0.34)	-3.06	(0.77)	9.07	(0.51)
	District	6.45	(0.34)	19.73	(0.01)	9.77	(0.20)	18.69	(0.11)
d. Largest tercile	AgZ + Region	-4.13	(0.67)	-2.20	(0.79)	9.18	(0.32)	1.12	(0.92)
	AgZ + Prov.	-2.40	(0.80)	-7.19	(0.33)	10.25	(0.24)	-1.25	(0.91)
	District	-7.37	(0.47)	8.59	(0.38)	8.75	(0.41)	25.18	(0.01)
No. observations	-	6,266		6,266		6,282		6,282	

Notes: Table reports regression estimates on the micro-data (i.e. households are the unit of analysis) from different models, incorporating alternative fixed effects (as indicated) and for both aggregate and disaggregate versions of the large farm variable (i.e. all vs. terciles); dependent variable indicated in the column header; each row/column combination corresponds to a different model; row (a) of columns (I) and (II) are the micro-data analogues to estimates reported in columns (II)–(V) of Table 2; rows (b)–(d) are the micro-data analogues to estimates reported in Table 3; only selected coefficients shown; 'Prob.' is the likelihood that the beta coefficient is equal to zero in the population, calculated using robust standard errors, clustered at the PA level.

Source: Authors' calculations from IAI2012 dataset (MINAG, 2012).

Table 5: Micro-data regression results, by farming type

	Total income (log.)				Crop income (log.)			
	Binary		Area share		Binary		Area share	
	Survey	IP	Survey	IP	Survey	IP	Survey	IP
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Annual	-0.19*	-0.23***	-0.02	-0.04	-0.18	-0.36***	-0.12	-0.17
	(0.10)	(0.10)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
Permanent	0.63***	0.53***	0.40**	0.36**	0.49***	0.48***	0.59**	0.78***
	(0.10)	(0.10)	(0.15)	(0.18)	(0.11)	(0.14)	(0.23)	(0.24)
Pasture	0.21	0.15	-0.24	0.06	0.36**	0.30*	-0.17	0.32
	(0.15)	(0.13)	(0.16)	(0.17)	(0.16)	(0.18)	(0.22)	(0.20)
Fallow	-0.40**	-0.19	0.08	0.28**	-0.35	-0.65**	0.17	0.28**
	(0.20)	(0.22)	(0.11)	(0.13)	(0.22)	(0.31)	(0.12)	(0.14)
All equal (prob.)	0.00	0.00	0.03	0.20	0.00	0.00	0.04	0.01
All zero (prob.)	0.00	0.00	0.05	0.08	0.00	0.00	0.06	0.00

significance: * 10%; ** 5%; *** 1%

Notes: Table reports coefficient estimates from alternative models where the aggregate large farm ‘treatment’ variable is disaggregated by farm type (overlapping); dependent variable as indicated in column header; ‘binary’ indicates farm type variables are dummy variables; ‘area share’ indicates types are the share of the large farm area in the PA allocated to the given activity; all models include district fixed effects; other aspects of the models are as per Table 4; only selected coefficients shown; robust standard errors, clustered at the district level, reported in parentheses.

Source: Authors’ calculations from IAI2012 dataset (MINAG, 2012).

Table 6: Micro-data regression results, outgrower marginal effects

		Total income (log. × 100)				Crop income (log. × 100)			
		Survey (I)		IP (II)		Survey (III)		IP (IV)	
		Beta	Prob.	Beta	Prob.	Beta	Prob.	Beta	Prob.
a. All large farms	AgZ + Region	20.53	(0.27)	18.86	(0.31)	29.39	(0.17)	9.42	(0.63)
	AgZ + Prov.	26.81	(0.18)	20.00	(0.30)	37.67	(0.10)	15.59	(0.46)
	District	22.00	(0.20)	17.69	(0.30)	33.66	(0.07)	19.82	(0.31)
b. Smallest tercile	AgZ + Region	-5.77	(0.81)	-9.58	(0.72)	-9.71	(0.67)	-24.33	(0.41)
	AgZ + Prov.	-1.17	(0.97)	-12.00	(0.71)	-2.79	(0.92)	-23.40	(0.49)
	District	-14.55	(0.52)	-16.65	(0.54)	-10.17	(0.57)	-18.38	(0.51)
c. Medium tercile	AgZ + Region	38.06	(0.11)	59.16	(0.00)	47.86	(0.02)	50.48	(0.00)
	AgZ + Prov.	47.59	(0.06)	64.46	(0.00)	61.00	(0.01)	61.37	(0.00)
	District	37.88	(0.05)	47.46	(0.04)	44.74	(0.08)	40.55	(0.23)
d. Largest tercile	AgZ + Region	37.41	(0.26)	-7.70	(0.54)	63.91	(0.20)	-1.38	(0.93)
	AgZ + Prov.	39.15	(0.26)	-10.65	(0.45)	60.43	(0.23)	-4.33	(0.80)
	District	63.94	(0.06)	14.86	(0.48)	81.51	(0.04)	28.54	(0.23)
No. observations	-	6,266		6,266		6,282		6,282	

Notes: Table reports results from the full model (equation 1) including interaction terms; effect shown is the marginal impact on total/crop income (as indicated) for smallholders who are outgrowers of the PA containing a large farm; each row/column combination corresponds to a different model as per Table 4; only selected coefficients shown; 'Prob.' is the likelihood that the beta coefficient is equal to zero in the population, calculated using robust standard errors, clustered at the PA level.

Source: Authors' calculations from IAI2012 dataset (MINAG, 2012).