Spillovers from agricultural processing

Ryan B. Edwards[*]
Dartmouth College

February 19, 2019

Abstract

Can agricultural processing ease the transition out of agriculture? Normally, factory-led industrialization happens in areas that are already growing, have good market linkages, and rising education. It is not clear whether factory-led development can succeed in places without these characteristics, which might be viewed as unprepared to industrialize. This paper uses the proliferation of palm oil factories across Indonesia's outer islands as a natural experiment to study industrial onset and estimate spillovers from agricultural processing. I find that processing increases incomes and non-agricultural employment, with similar effects for locals and recent migrants. Villages near palm factories also have more people, firms, and other economic and social organizations. These patterns can be at least partially explained by economic linkages, more rural infrastructure, and improved local market integration. By focusing on subsistence rural regions in a large developing economy, this paper adds a globally-significant new case to a growing literature emphasizing the importance of agglomeration externalities for understanding the birth of new towns, the spatial distribution of economic activity, and structural transformation.

[*]Contact address: Department of Economics, 6106 Rockefeller Hall, Dartmouth College, Hanover, NH 03755 USA. Email: ryan.b.edwards@dartmouth.edu. Thanks to Eric Edmonds, Treb Allen, Sam Bazzi, Marc Bellemare, Amy Damon, David Eil, Marcel Fafchamps, Wally Falcon, Jim Feyrer, Gordon Hanson, Paul Goldsmith-Pinkham, Will Johnson, Amy Lastuka, Blane Lewis, Paul Novosad, Dietz Vollrath, and seminar participants at the Dartmouth Junior Lunch, LACDEV Middlebury, and the University of Michigan for helpful comments. Grace Hadiwidjaja, Sebastian Kraus, Budy Resosudarmo, Mayang Rizky, and Pablo Pacheco kindly shared data. Lauren Archer, Robert Castle, and Victoria Meyer provided excellent research assistance under the Dartmouth Presidential Scholars program.

1 Introduction

Can growth in export agriculture help ease the transition out of agriculture? Transitioning the majority of workers in developing countries to higher-value, market-based agricultural activities and out of agriculture is essential for breaking rural poverty traps and economic development. Large productivity gaps across sectors persist, despite global agriculture being more integrated than ever (Gollin, Lagakos, and Waugh, 2014). Characterized by giant agro-industrial conglomerates sourcing from both industrial farms and smallholders, modern agricultural value chains commonly reach from cities like Rotterdam, Saint Louis, Seattle, and Singapore to some of the most remote parts of the developing tropics (Byerlee, Falcon, and Naylor, 2016; Bellemare and Bloem, 2018). By imposing industrialization on subsistence agrarian communities, export-oriented processors could dramatically reorganize rural economic life: offering a foothold into global markets and industrialization, or leading to two-tiered labor markets, a loss of land among the poor, and immiserisation of local communities. This paper uses the proliferation of palm oil factories across Indonesia's remote outer islands as a natural experiment to study the onset of industry, test these competing theories, and estimate economic spillovers from agricultural processing.

The main empirical concern is purposive placement. Factories tend to locate in areas already growing, with placement correlated with other factors also affecting local economic outcomes. An ideal natural experiment to study industrial onset would see factories arbitrarily scattered across an undeveloped hinterland, or following some observable placement rule. The unique features of the palm oil supply chain and Indonesia's dramatic expansion provide a useful approximation of this experimental ideal. Indonesia's four fold-increase in palm oil production over the last two decades is the world's largest modern agricultural expansion. Since most Indonesian palm oil is exported, the relevant demand is external to producing communities. On the supply side, expansion into unindustrialized areas reduces concerns about firms colocating to capture agglomeration economies from other firms.

Figure 1 illustrates an archetypical case, the Kerinci area in Riau province, with three satellite images. The first was taken in 1984, when Royal Golden Eagle International started developing plantations in this mostly forested, subsistence agrarian community. The second picture was taken in 2000, as the current expansion commenced. The third image shows the urban center there today, surrounded by a mosaic of commercial estates and small, family farms. [1] The urbanization in Figure 1 is quite close to the factories. The fact that oil palm fruits must be processed within 24 hours means that (a) direct impacts tend to be concentrated near factories, and (b) factory placement closely follows growing conditions. Communities that happen to be in areas most suitable for palm cultivation experience growth in processing because the optimal factory location in a particularly locality is the point that minimizes the amount of land needed to feed the factory, while maximizing feed quality.

My empirical strategies exploit local spatial differences in factory exposure. Using a novel new dataset collecting the universe of palm oil processing firms in Indonesia, I compare villages near factories to unexposed villages slightly farther away. A rich control function captures relative suitability within factory catchments in terms of extremely precisely measured local topographic, hydrological, and other geographic characteristics. I address reverse causality concerns by showing that there are no statistically significant effects on pre-period outcomes, consistent with new factories moving into the hinterland and the time dynamics shown in Figure 1. I address remaining concerns about potential omitted variables by (a) using balance tests on observable characteristics unrelated to suitability to support the conditional independence assumption, (b) showing that my results are robust to the inclusion of demanding set of additional control variables, and (c) presenting consonant estimates from multiple suitability-based instrumental variables strategies.

[1] Figure A1 in the online appendix provides some "street view" images taken from Google Maps. I show Kerinci here because it was the first rural community I visited in Indonesia, but Kerinci's transformation is not dissimilar to that in underway more nascent areas since 2000. Similar time series images from the islands of Kalimantan and Papua are provided in Figures A2 and A3.

The main finding is signs of urbanization and structural change around palm oil factories. Households living closer to factories are more likely to work in non-agricultural sectors. Per capita household expenditures are around ten percent higher near a factory. Effects are most geographically-concentrated and precisely estimated for agricultural households. Households near factories are twenty percent more likely to be employed (c.f., self-employed or dependent). These local labor market estimates paint a picture of villages around palm oil factories as embryonic towns undergoing a process recently dubbed proto-urbanization by Marcel Fafchamps, Michael Koelle, and Forhad Shilpi (2016). Higher population and firm densities around factories confirm this interpretation. The average village 5–10 km from a palm oil factory has one more formal firm and 500 more people than a village 25–100 kilometers away. Effects are twice as large within 5 km, and the labor market results are similar for migrants and locals. Hence, there appear to be significant local benefits from agricultural processing beyond those accruing to farmers.

How could relatively unsophisticated processing facilities lead to such a dramatic reorganization of rural economic activity? Linked industries are a first potential explanation. Villages near the factories have more firms and organizations of many different types, including those which provide inputs to the palm oil sector. However, effects tend to be be quantitatively larger for *un-linked* industries, suggesting that economic linkages are predominantly on the consumption side. Neighboring firms range from micro-enterprises to large firms and are mostly in the retail, maintenance, finance, transport, construction, and processing sectors. Villages near factories also have more cooperatives, agricultural input suppliers, and banks, as observed in the village census. Even though I find some evidence of local backward linkages, the diverse range of economic activities found around factories is more consistent with demand-based consumption linkages and broader agglomeration forces.

A second potential explanation relates to infrastructure and publicly-provided goods, since export-oriented agricultural manufacturing often requires up-front investments not only in the factory but in transport and utilities networks. Villages around factories are more likely to have (a) the main road upgraded from dirt to gravel or asphalt (within 20km), (b) public transport (within 5km), (c) lighting on the main street (within 50km), and (d) fewer households without electricity (within 30km). Improved infrastructure, local market integration, and market access are likely to reinforce the shift out of subsistence food production towards market-oriented agriculture and non-agricultural employment.

A booming economy provides apt opportunities for local governments to raise revenue and increase public good provision. Public goods could in turn generate higher returns with industrial production nearby. Within 10km of a factory, villages' annual budgets are on average 50 percent larger. Villages near factories rely less on inter-governmental transfers, have higher expenditures, and own more land, buildings, and other assets. Publicly-provided goods not closely related to agricultural supply chains—market, education, and health facilities—are also more common. Public good spillovers appear broader than those that would only arise only through the supply chain. To gauge whether village fiscal windfalls from new local industry are driving the improved public good provision, I adjust estimates for local government revenue and expenditure. Results are statistically indistinguishable, suggesting public goods near factories may also be the result of targeted inter-governmental transfers or privately provided.

Improved infrastructure opens up the possibility of falling trade costs as a final channel at work. To test this hypothesis, I introduce a novel new survey of almost 30,000 palm oil farmers and spatially link them to the processors they could plausibly sell their fruit to. Farmers in the same subdistrict as a factory use significantly more inputs, have larger harvests, and achieve higher per hectare yields. This proximity comes with a significant downside. Farmers with factories in the same subdistrict receive substantially lower prices for their fruits. However, this pattern which reverses as the number of potential buyers increases. I interpret these patterns as (a) proximity and better infrastructure (i.e., lower trade costs) leading to higher on-farm productivity, and (b) the decrease in buyer monopsony power as farmer-to-factory markets develop allowing farmers to command better prices for their fruits.

This study contributes to three main streams of economics. I first extend the classic literature on agriculture, industrialization, and economic development, particularly that focused on linkages from agriculture to industrial and service sectors (Clark, 1940; Myrdal, 1957; Hirshman, 1960; Johnson and Mellor, 1961; Chenery and Syrquin, 1975; and Marden, 2018). Emphasizing consumption linkages from rising agricultural incomes and production linkages from locally-sourced inputs for processed agricultural products, much of this work is predicated on the idea that agricultural processing offers an entry point to better jobs with limited skill requirements—a first rung on the ladder of economic development. Despite a rich intellectual history and continued emphasis from policy-makers, empirical support for these theories remains thin, particularly in relation to the global supply chains that characterize the global food system today (Gollin (2010), Dercon and Gollin (2014), and Bellemare and Bloem (2018) provide surveys).

Four recent studies make important headway. Melissa Dell and Benjamin Olken (2018) study the Dutch Cultivation System on the Indonesian island of Java. Due to technology constraints at the time, sugar factories had to be placed along rivers, adjacent to enough sugar-suitable land, and spaced a reasonable distance apart. Using a selection-on-observables design similar to this paper, they find that, despite the overarching extractive colonial system, sugar factories had persistent positive impacts on local economic development. Paula Bustos, Bruno Caprettini, and Jacopo Ponticelli (2016) study the more recent Brazilian soy expansion. Soy is also exported and requires local processing. Comparing sectoral employment, wages, and productivity across regions, they find that Brazil's soy expansion led to local structural change. In a companion paper, I compare regional poverty and consumption trajectories over Indonesia's expansion to find broad consumption gains in producing districts (Edwards, 2019). Kubitza and Gerhke (2018) find significant fertility reductions with a similar approach. Here, I extend the scope of Dell and Olken's analysis to the world's largest modern agricultural expansion and a setting of decentralized democratic institutions, and the other three studies to a much finer level of spatial aggregation with a focus on processing facilities and underlying causal mechanisms.

My study also relates to emerging work at the nexus of trade, spatial development, and economic geography.[2] Normally, factory-led industrialization happens in areas that are already growing, have good market linkages, and rising education. It is not clear whether factory-led development can succeed in places without these characteristics. Indonesia's palm oil expansion offers a rare natural experiment to study the impacts of new factories in places that might otherwise be viewed as unprepared to industrialize. Far from adverse effects, I find that people adapt quickly, that benefits are broad-based, and that consumption linkages, economic and social infrastructure, and local market integration are crucial mechanisms reinforcing agglomerations. My study thus adds a unique and globally-significant new case—the first focused on largely subsistence rural regions in a large developing economy—to a growing literature emphasizing the importance of agglomeration externalities for understanding the birth of new towns, the spatial distribution of economic activity, and structural transformation (Michaels, Rauch, and Redding, 2012; Allen and Arkolakis, 2014; Donaldson, 2018).

Finally, my findings inform a topical debate on sourcing from developing countries (Swinnen, 2007; Maertens and Swinnen, 2009; Dragusanu et al, 2014). The traditional view is that large commercial farms and plantation systems are not only harmful for the environment but also socioeconomic development in producing communities, despite their increasing connectedness to smallholder systems through contract farming and other modalities (Farina and Reardon, 2000; World Bank, 2008, Byerlee, de Janvry, and Sadoulet, 2009, Engerman and Solokoff, 2002; Easterly, 2007). While these concerns are not limited to palm oil, controversy over palm oil has been so prominent that the Indonesian President banned new permits in September 2018 and the European Parliament imports in 2017. The World Bank has had a sector moratorium in place

[2] See Fujita, Krugman, and Venebles (1998), Moretti (2010), and Donaldson (2015) for excellent surveys from the perspectives of new economic geographic, local labor markets, and market integration.

since 2009. An alternative view is that agricultural processing provides a foothold in the ladder of industrial development, offering capital inflows and market access for areas typically lacking both. My findings echo those of Dell and Olken (2018) and Bustos et al (2016), consistent with this second view.

The paper proceeds as follows. The next section introduces the Indonesian palm oil sector and explains factory placement. Section 3 details the empirical strategy. Section 4 presents the main results on rural agglomeration and several key robustness checks. Section 5 explores economic linkages, infrastructure, and local market integration as potential explanations for the main findings. Section 6 offers some concluding remarks.

2 Palm oil processing in Indonesia

Palm oil is the world's most consumed vegetable oil, produced almost entirely in tropical developing countries. Global palm oil production has roughly doubled every decade since the 1960s, increasing from less than 5 million metric tonnes per year in 1970 to over 70 million in 2015. Indonesia supplied 55 per cent of the 65 million metric tons produced globally in 2016–17. Spurred by sweeping decentralization reforms and a devalued rupiah following the Asian financial crisis (Rada, Buccola, and Fuglie, 2010), the area under cultivation for oil palm in Indonesia increased from 2.9 million hectares in 1997 to over 12.5 million today. Decentralization ushered in a parallel regime shift in the palm oil sector from select, centrally-planned plantation developments to liberalized investment and diffuse growth encouraged by newly-empowered local governments (Fitriani, Hofman, and Kaiser, 2005; Burgess, Hansen, Olken, Potapov, and Sieber, 2012; Naylor, Higgins, Edwards, and Falcon, 2018). Since fresh fruit bunches (FFB) from an oil palm must be processed within 24 hours to be of a high enough quality for global markets, cultivation is concentrated around factories. A new factory effectively opens up a new export market for local farmers. From 2008 to 2015 alone the number of palm oil businesses, most of which are processors, increased by 43% —from 1059 to 1511 (BPS, 2017). Smallholder oil palm adoption around new factories tends to be voracious and each factory relies on dozens of surrounding villages for a steady supply of fruit.

I trace the supply chain from the farm gate through the factory to final export markets in Figure 2. Upstream, smallholders manage over 40% of 12 million hectares planted with oil palm. Road networks and transport logistics are needed to transport harvests to factories in a timely manner. In undeveloped rural areas without prior industry, firms or governments need to build the infrastructure. With these networks in place, smallholders usually aggregate and coordinate their activities through a complex network of cooperatives, traders, and other aggregation points (e.g., village loading ramps). After entering the sector, farmers typically report much higher incomes and labor saved (Krishna et al 2017; Kubitza and Gerhke, 2018). Each factory directly employs around 200 highly-skilled workers which together with smallholder adoption could generate consumption linkages through demand for locally-produced goods and services.

Indonesia's approximately 1200 palm oil factories sell mostly to refineries and global commodity markets. According to the 2010 Input-Output Table, nine domestic industries use palm oil and byproducts as inputs: animal and vegetable oil, peeled grain, animal feed, pesticide, drugs, pharmaceutical products, soap and cleaning, industrial cosmetic products, and other chemical goods industries (BPS, 2010). On the input side, the palm oil industry draws from 60 out of 192 sectors (Appendix B provides the complete listing). Hence, economically important production linkages are likely on the input side.

2.1 Factory placement

Understanding why factories are located in particular places is crucial for identifying their impacts. Despite growing smallholder involvement in the post-Suharto era, most palm oil factories still operate as part of a plantation system rather than stand-alone entities. There are three key requirements to develop plantation economies: (a) final product demand, (b) financial capital to build the factory, related supply chain infrastructure, and nearby plantation (traditionally known as the "nucleus" estate in Indonesia), and (c) sufficient suitable cultivation area nearby to operate the factory efficiently (Hayami, 2010).

As an export-oriented commodity experiencing an unprecedented surge in global demand (over a 300% increase since the 1990s), new factories arise principally due to external rather than local demand conditions. Crucially, the relevant demand is outside producing communities. Large up-front costs see investments in processing infrastructure made by large firms based in capital cities or abroad (c.f., locally). Palm oil expansion has been across Indonesia's relatively undeveloped outer islands and investments are often "greenfield": introducing industry to subsistence agrarian communities with scant preexisting industrial activity. Immediate processing requirements mean direct impacts are then concentrated close by. These two conditions alone set up an ideal natural experiment to study the onset of industrialization and spillovers from new factories.[3]

Whether one relatively undeveloped rural district gets a factory is typically decided by the central government and district heads (i.e., *bupatis*), who grant permission to build factories and adjacent estates.[4] A burgeoning qualitative literature documents how—under both the New Order and in the post-Suharto "reformasi" era—plantation and natural resource projects tend to proceed irrespective of local views and village socioeconomic conditions (Resosudarmo 2008; Cramb and McCarthy, 2016; Robinson and McCarthy, 2016; Gatto et al 2017; Pramyudya et al 2017). Although these power dynamics are concerning from the perspective of indigenous land rights and the environment, they are helpful for identification in the sense that individual villages have limited ability to affect factory placement. Since districts determine whether new factories go ahead in their jurisdiction, my counterfactual comparisons are always against unexposed villages (i.e., not near a factory) close by with the same local institutions, politics, and policy settings.

[3] In most other contexts, the impacts of any one factory are difficult to disentangle because factories locate near existing economic activity to capture agglomeration economies (Greenstone, Hornbeck, and Moretti, 2010).

[4] Industrial estates account for just over half of production, and small family farmers the other half. Around 20 percent of industrial estates were forest within the five years prior to establishment, and proportion likely smaller for smallholder who lack the capital to clear cut (Austin et al 2019).

With factories serving as the processing core for thousands of hectares of surrounding farmland, the primary consideration for investors when placing a new factory is sufficient suitable cultivation area. The optimal factory location in a relatively undifferentiated rural landscape is where yields are highest: maximizing feed quality, while minimizing the area needed operate efficiently (author communications; Corley and Tinker, 2016). Oil palm grows best in the humid low-lying tropics: flat areas with mineral soil, abundant rainfall, and sufficient sunlight. Tree crop estates and transport logistics are also more costly to build and manage in more hilly areas. The agronomic literature and extensive discussions with firms and farmers emphasize the primacy of growing conditions (Corley and Tinker, 2015; Naylor et al, 2019), which firms assess on the basis of local rainfall and humidity, overall terrain, and mineral versus peat soil (author communications). As I show shortly, factory placement closely follows these plausibly exogenous agro-climatic growing conditions.

3 Empirical strategy

Immediate processing implies that local adoption and direct impacts decay over distance. Villages near but not near enough make a reasonable comparison group since they share local institutional, economic, and geographic characteristics. Figure 3 uses the PT Tunas Sawa Erma factory in Papua to illustrate the intuition of my approach. In Panel A, there is just the factory (green dot), village boundaries (light gray lines), and district boundaries (thick black line). Taking the centroids of all the villages, we have the points in Panel B. Every village is assigned an indicator for whether it fits within one of many distance bands, 20 km each in this illustrative example. The red line demarcates 100 km from a factory, beyond which villages are discarded from the estimation sample. Identification is based on spatial heterogeneity and the main identifying assumption is conditional independence: conditional on a comprehensive set of village-specific controls capturing relative within-catchment suitability, this factory could have been placed in any of these distance bands.

3.1 Estimating equation

I relate factories to village outcomes with the specification:

$$y_v = \alpha + \sum_{i=0-5}^{75-80} \gamma dfact_v^i + \sum_{j=1}^{n} fact_v^j + \beta X_v + \epsilon_v \quad (1)$$

y_v is a village outcome of interest, measured in censuses, surveys, and other datasets described at Appendix A. $dfact_v^i$ are treatment indicators equal to one if a village is 0–5 km, 5–10 km, ... , 75–80 km to the nearest palm oil processor. This specification allows me to see how the average value of an outcome of interest varies with proximity to palm factories. Processors are identified in the 2016 Economic Census and exactly matched to village centroids. 80–100 km is the excluded bin, the maximum distance included in the sample to ensure geographic comparability. I exclude villages in cities and on Java, where little oil palm is grown but offices (with the same industry codes as factories) and downstream processors are located. Nearest factory fixed effects $fact_v^j$ restrict comparisons to villages near the same factory and capture spatial unobservables across localities.

X_v is a vector of controls capturing the relative attractiveness of each village for a factory within each factory's catchment. Village elevation, slope, historical precipitation, water flow accumulation (measured as the average number of cells uphill), and distance to the nearest river account for terrain and relative growing conditions. To account for preexisting utilities, industrial activity, and economic development, I include distance to the nearest major road in 2000 and nighttime luminosity in 1993. Since roads and utilities are more expensive to build in more rugged terrain further from existing networks, these two variables also proxy the relative cost of new supply chain infrastructure. Distance to the nearest district capital city and village area account for administrative and economic remoteness and the relative cost and availability of land. An urban dummy allows a different intercept for preexisting rural towns. ϵ_v is a robust error term.

Figure 4 plots processor locations and Table 1 provides summary statistics for the estimation sample and for all villages. In my main estimation sample, the distance to the nearest processor is 38 km and the average village has 2.6 factories within 25 km. Processing is most concentrated in the provinces of Riau and North Sumatra, where the crop was first introduced by the Dutch in the early 20th century.

3.2 Identification issues

The identification assumption is conditional independence: geographic variables related to growing conditions capturing factory placement within each catchment, and other characteristics being uncorrelated. To assess the credibility of this assumption, in Figure 5 conduct "balance tests" by plotting the distance band coefficients from estimating Equation 1 using various geographic characteristics as dependent variables. The balance test reveals that factory placement very closely follows suitability. Villages near factories tend to be flatter (slope), lower (elevation), humid (proxied by flow accumulation and rainfall), and in more remote areas where land tends to be easier to acquire (distance to the nearest main road). By contrast, I find no statistically significant correlation between factory proximity and variables unrelated to suitability (i.e., village area, distance to capital, cost to capital, and distance to river). I conduct a similar exercise using coarser gridded data on crop-specific agro-climatically attainable yields from the Food and Agriculture Organization's Global Agro-Ecological Zones (GAEZ) database (Figure A11).[5] With the exception of coconut, factories closely follow palm suitability and not that of other crops.

[5]The empirical advantage of the FAO-GAEZ data relative to using its more spatially disaggregated inputs is that the model and crop-specific yield functions add an additional dimension of plausible exogeneity beyond that of the climatic and soil inputs. I leverage this advantage for instrumental variable estimates, as an robustness check on my main estimates.

One key identification concern is that factories may be built in areas already growing, with better preexisting infrastructure, skilled workers, and so forth (i.e., reverse causality). Estimates could pick up such underlying differences rather than the effects of factories. Expansion into largely subsistence regions (i.e., without prior industrial activity) makes this unlikely. It is also unclear that we should *a priori* expect *positive* selection. Firms minimize costs and often targete poorer, *less* developed areas (with lower land and labor costs, and possibly more interested or questionable local leaders). This would bias estimates *downwards.* To test for purposive placement more formally, I restrict the sample to villages in districts which began palm oil production after 2000 and conduct "placebo tests" on pre-period outcomes from the 1993 village census. Results are presented in Figure 6, finding no statistically discernible differences near factories.[6]

Another identification concern is that there may be an omitted variable jointly affecting factory placement and outcomes. I provide two types of evidence to address this concern. First, I introduce a host of additional controls to try and explain away the main results. To address concerns over remaining unobservables, I also present results from multiple suitability-based instrumental variables strategies. I present these estimates after the main results, along with results from falsification tests exploiting the reduced form relationships between suitability and key outcomes.

3.3 Proximate adoption and poverty

Using villages slightly farther away from factories as a control group assumes that oil palm adoption and other direct impacts are concentrated around factories and villages slightly farther away are unaffected. Figure 7 presents the estimated coefficients at each distance band using village oil palm acreage and poverty as dependent variables. Each subfigure overlays estimates from (a) the baseline specification, (b) the baseline specification plus a host of geographic controls

[6] 2000 is the first year I have district production data, and roughly when the area and production expansion begin to take off. I also estimate Equation 1 for nightlights in 1993, 2000, and the present. Consistent with limited initial selection (and the presence of some factories in the pre-period), very small initial differences grow larger over time. These estimates are omitted for brevity but available on author request.

from the 2014 Village Census (specifically, travel time to the nearest city (accounting for road quality), travel cost to the nearest city, and indicators for river, coast, plains, valleys, forest proximity (i.e., in, near, and outside), and forest function, e.g., production or conservation), and (c) the baseline specification plus a full polynomial in latitude and longitude to purge remaining geographically-distributed unobservables. The first sub-figure of Figure 7 shows that adoption decays up to 30km, when the estimated coefficient becomes statistically indistinguishable from zero. Within 10km of a factory, villages have on average 100 more hectares planted (50 households at the median plot size of 2 hectares). The second sub-figure finds that poverty is five percent lower within 5km of factories. To the extent that (a) causal chains start with local adoption and (b) other development outcomes are correlated with poverty, the overlapping confidence intervals here suggest the spatial patterns I proceed to document are likely due to the factories.

4 Main results

4.1 Local labor markets

Figure 8 examine villages' primary source of income using 2014 Village Census (PODES). Consistent with the adoption patterns in Figure 7, there is greater dependence on plantation agriculture around factories. However, overall dependence on agriculture declines. Villages near factories are four percent less likely to report agriculture as their primary source of income. Industry as a primary village income source increases slightly, mostly through manufacturing. A large increase in service sector work appears to offset the decline in agriculture. Together, these results suggest greater specialization within agriculture and a movement out of it.

PODES only captures the main source of income in each village as reported by the village head. It is unclear just when these thresholds are crossed (i.e., the main source of income vs. not main source of income). Figure 9 uses pooled household data from the national socioeconomic survey SUSENAS to shine light on other margins of adjustment. SUSENAS is a repeat cross-sectional household survey conducted at least annually by *Badan Pusat Statistik*. I follow Dell and Olken (2018) and pool data over the years with village identifiers to improve coverage. Household size, an urban-rural dummy variable, and survey-year fixed effects are included as controls throughout. Consistent with results from the village census, households in villages closer to factories are less likely to report agriculture as their main source of income and more likely to report industry or services.[7]

A key advantage of incorporating SUSENAS into my analysis is that it allows me to also examine consumption patterns around factories. The second row of Figure 9 examines household expenditures. Per capita expenditures (PCX) for agricultural and service sector households are ten percent higher near factories, decaying up to 20 kilometers away when differences become indiscernible from zero. There is no clear pattern for households employed in industry, consistent with skilled labor being more mobile and wage equalization across locations. The final row turns to work status. SUSENAS defines work status as employee, self-employed, or "income receiver." The latter classifies people on a pension, welfare, getting money from a relative, or similar. Agricultural and informal employment is usually classified as self-employment. Households living in villages near factories are more likely to be employees, and less likely to be self-employed or an income-receiver. Effects on work status are more precisely estimated and larger than those on primary sector of employment and expenditures, perhaps reflecting the fact that people outside formal employment are more likely to have multiple livelihoods. The shift out of self-employment and agriculture is consistent with the spatial division of labor usually observed around nascent urban centers (Fafchamps, 2012).

[7] Within agriculture, most of the effect is coming from a reduction in dependence on crop farming, consistent with a shift to peri-urban agriculture. Manufacturing and construction account for most of the new industrial employment (available on author request).

4.2 Population and firms

Labor reallocation out of agriculture should be mirrored by changes in population and firms. Figure 10 estimates impacts on village population and whether a village is home to a *formal* registered firm, as measured in the 2011 Village Census and the 2016 Economic Census. Firms cluster around factories and nearby villages on average have larger populations. The magnitude of these effects is not trivial. The average village 5-10 km from a palm oil factory has 500 more people and, on the extensive margin, is fifteen percent more likely to have a firm present. On the intensive margin, this equates to one additional firm. Effects are twice as large within 5 km. An alternative explanation for the increased specialization in agriculture and movement out of it is land use change and plantation-related displacement of local populations. Although many more people near factories suggests that selective displacement is unlikely to be driving the results, I examine impacts by migration status shortly.

The main limitation in assessing rural economic activity using the Economic Census is that most micro and small enterprises in rural areas are not registered and thus not counted. These smaller organizations are more common, particularly at early stages of economic development (Hsieh and Olken, 2014; Rothenberg, Gaduh, Burger, Chazali, Tjandraningsih, Radijun, Sutera, and Weilant, 2016). 2014 PODES allows me to probe local economic activity further, as village heads are asked how many of a particular type of business or organization are in their village, regardless of formal status. Villages near factories are more likely to have a bank, and people living in villages moderately close are more likely to have received credit in the last year (see Figure A12 in the online appendix). I find no evidence of more micro-industry or small-scale processing businesses. A village from 5-10 km from a factory has on average ten more small service businesses, implying that (a) industrial employment is mostly coming through the palm factories and (b) economic linkages are mostly on the consumption side. Villages near factories also report having more cooperatives and agricultural kiosks, used to organize agricultural activities and supply technical inputs (e.g., seeds and fertilizer).

The shift out of agriculture into formal employment is complicated by population growth. Skilled in-migration could completely account for the sector and employment effects. I test this conjecture with the 2010 Population Census, examining impacts separately for "locals" and people who lived in a different district in 2005. Effects on employment status are not statistically distinguishable by migration status, although the declines in self employment and housework appear driven by non-migrants (Figure 11). The decline in agricultural employment is concentrated amongst locals, and estate crop workers are more likely to be recent migrants (Figure A15). Hence, we can effectively rule out that increase in employment, decline in self-employment, and decrease in farm work is driven by recent positively-selected migrants.

The results presented thus far suggest a reasonably strong correlation between proximity to palm oil processors and a host of outcomes synonymous with nascent urban centers, and that these urban centers are not reminiscent of economic enclaves. However, these spatial patterns can only be interpreted as the causal effects of factories under the relatively strong assumption of conditional independence. One might still be concerned that these patterns are driven by an unobserved omitted variable determining both factory locations and outcomes of interest. To address this concern, I introduce a host of additional controls and presents results using multiple instrumental variables strategies.

4.3 Robustness—additional controls

To present estimates with additional covariates compactly, I switch to a dichotomous treatment indicator equal to one if villages are less than 10 km from a factory and focus only on population, firms (extensive margin), and whether a village's primary source of income is agriculture. Factory fixed effects and geographic controls remain, with observationally similar villages 10–100 km away the comparison group. Column 1 of Table 2 presents the OLS estimate without conditioning on any controls (i.e., including only nearest factory fixed effects). Column 2 is the binary treatment analogue of Figure 10, with baseline controls. Column 3 adds additional geographic variables reported in the village census (listed at Figure 7) and Column 4 adds a polynomial in latitude and longitude. Estimates up to this point have been adjusted for relative suitability in terms of extremely fine measurements slope, rainfall, elevation, and water flow accumulation. Column 5 adds each village's agro-climatically attainable palm oil yield derived from the Food and Agriculture Organization's Global Agro-Ecological Zones (GAEZ) palm-specific model.[8] Column 6 alternates the approximately 750 nearest-factory fixed effects for around 220 district fixed effects. Column 7 adds 2882 subdistrict fixed effects (the administrative level above a village, usually much smaller than a factory catchment) to exploit only spatial differences in exposure across villages in the same subdistrict. The point estimates remains large, statistically significant, and consistent across specifications.

4.4 Robustness—instrumental variables

An appropriate instrument will predict factory exposure but not affect outcomes through any other channel other than through a factory close by. My instrumental variables all leverage the fact palm factories locate in the most suitable areas. The first strategy is similar in spirit to the approaches used in Duflo and Pande (2007), Dinkelman (2011), and Lipscomb et al (2013). I use a simple probit model to predict the probability that a village will be within 10 km of a palm oil factory. Predictions are made on purely on the basis of exogenous determinants of profitability (described in Section 3) observed at an extremely fine level of spatial disaggregation, and predicted probabilities are used as an instrument for actual exposure. The granularity of my geospatial data is important here, offering more variation within localities than the standard FAO-GAEZ suitability data. However, the FAO-GAEZ data has the additional benefit of incorporating crop-specific yield functions from agronomic models without using any information on observed productivity and other human activity on the ground. Hence, my second instrument is village-specific agro-climatically attainable palm oil yield calculated from

[8]Although GAEZ potential yields lacks the precision of my baseline suitability variables, the model-based yield response function adds additional crop-specific information.

the FAO-GAEZ data. One concern with this instrument is that areas suitable for palm might be suitable for other agriculture, which my final instrument addresses. The third instrument is the normalized difference between potential yields for palm oil and Indonesia's other key cash crops (i.e., coffee, cocoa, and teas).

Table 3 present the causal effects of palm oil factory exposure from my three instrumental variables approaches. Most IV estimates are larger than the corresponding OLS estimates in Table 2. This may be due to a downwards OLS bias—for example due to firms targeting areas with *worse* unobservable *village* characteristics, like corruption (recall that comparisons are within districts)—or the fact that the local average treatment effect related to optimal placement is likely larger than the average treatment effect. Overall, the IV estimates suggest that my main OLS estimates tracing out the spatial decay are more likely to be downward biased and thus conservative, rather than false positives.

The crucial identification assumption for the instrumental variables estimates is that palm suitability is conditionally independent, not affecting outcomes other than though factory exposure. The concern here is that more suitable areas were going to urbanize anyway, for some reason other than agricultural processing. To test whether this is the case, I present results from an intuitive falsification test. The first panel of Table 4 presents the reduced form impacts on population, firms, and agriculture. The second panel presents the same for a restricted sample discarding villages within 25 km of a factory, where the first stage is effectively zero. I find no evidence of any statistically significant effect of palm suitability or relative suitability in unexposed villages, suggesting that the exclusion restriction is likely satisfied.

5 Potential explanations

5.1 Production and consumption linkages

The results presented up to this point suggest that the location of palm oil production affects other economic activity through agglomeration economies and other forms of economic spillovers. I now examine three potential explanations for these patterns. Armed with the knowledge that the average village 5-10 km from a factory has one additional firm present, a natural first question to ask concerns linked industries. Do these firms provide services to palm oil firms and farmers, or use palm-related products as inputs?

I use the industry codes in the 2016 Economic Census to examine the relationship between factory proximity and the presence of different types of firms. Linked industries are identified in the 2010 Input-Output Table. Since the codes change over time and there is not yet a newer input-output table or official concordance, I translate descriptions then hand-match sector codes one-by-one on descriptions (see Appendix B for details). Figure 12 examines the type of firms around near factories. The first panel finds that villages within 20 km are more likely to be home to firms providing inputs to the palm oil sector. The second panel finds no evidence of any significant forward (i.e., output) linkages. Impacts on linked industries overall, in the third panel, are thus driven by backward linkages. The final panel examines whether villages near factories have more firms in sectors without input-output linkages to palm oil. Villages near palm factories are around ten times more likely to have a firm in a non-linked industry than one in a linked industry (note the different axis scales), consistent with the the evidence of more services sector business and banks presented in the previous section.

Figure 13 examines impacts separately for each major sector, again replacing the distance bands with a single 10 km treatment indicator to present results compactly. The dependent variable is a dummy variable indicating the presence of particular type of firm in a village. I find significantly more retail, maintenance, finance, insurance, transport, construction, and other service sectors firms near factories. Collectively, my findings here and in previous sections point towards consumption rather than production linkages being a crucial channel for economic spillovers to other sectors.

5.2 Infrastructure and public goods

As palm expansion has mostly been into remote areas with little commercial infrastructure, firms and governments seeking to develop a local palm industry often need to invest in infrastructure to transport, process, and market the crop. Such infrastructure could reinforce agglomeration and make rural public good provision easier, offering new revenue opportunities while reducing the cost of service provision. In the article appendix, I find significantly greater fiscal capacity in villages near factories. Village own source revenues are over fifty percent higher near factories (Figure A16). Near factories, villages also receive twenty percent less in Alokasi Dana Desa (ADD) intergovernmental grants (which are based partially on need) and have around ten percent higher expenditures. Villages near factories are also more likely to report having land (7 percent more likely), building (12 percent), and other assets (5 percent).

Figure 14 turns to public goods. Across outcomes, impacts vary in magnitude, statistical significance, and the distance at which they are felt. The top left figure shows that villages living within 10km of a palm oil factory tend to have 10 less families living without access to any electricity, despite having larger populations. Note that rural electrification in Indonesia is generally high, and population in 2011 is included as a control variable for this regression. The second and third figures (top row) show that villages within 20km of a factory are more likely to have the main road fitted with a street light, and that main road is less likely to be dirt but rather upgraded to gravel or asphalt. The first figure in the second row shows that villages within 5 km of factories are slightly more likely to have public transport. The middle and right figures (middle row) estimate impacts on the total number marketplaces (of all types) and the number of *permanent* physical marketplaces, which serve as crucial centers of exchange. Up to 20 km from a factory, villages are five percent more likely to have a market. With just 16 percent of villages having markets in 2014, the magnitude of this effect is not trivial.

The bottom row of Figure 14 looks at non-economic infrastructure, more related to human development than agricultural supply chains. I find large and precisely estimated impacts on the number of education, health, and worship facilities in villages near factories, with impacts statistically significant at the 5% level as far as 25 km away.[9] Although new palm oil supply chains could generate mechanical improvements in economic infrastructure, impacts on health and education facilities suggests that indirect benefits are not strictly limited to the supply chain. To gauge whether village fiscal windfalls are the likely channel for improved public good provision, estimates adding local government revenue and expenditure as controls are overlaid in red. The point estimate relating to having a dirt road is the only one with noticeable movement, although the confidence intervals still overlap. Despite the fiscal variables clearly being "bad controls", statistically indistinguishable results suggest that the new public goods may be privately provided or due to targeted intergovernmental transfers.

5.3 Local market integration

Improved infrastructure opens up the possibility of falling trade costs and local market integration as extra channels at work. I introduce a new household survey with detailed information on inputs, outputs, and farm gate prices for 27,710 farmers to test this hypothesis. Specifically, it is the farmer cost survey (SKB) specific to palm oil farmers, conducted by BPS and

[9] I look at places of worship because mosques and churches are often the first thing that communities will invest in (a helpful leading indicator of community financial health), and because they often serve as important centers of social activity, social services (e.g., child care), and risk-sharing.

the Department of Agriculture with the 2013 Agricultural Census. Since the survey identifies farmers at the subdistrict level, my main distance-based comparisons are now too coarse. Instead I calculate two new measures of farmer-to-factory market density: (a) whether a village has a factory in its subdistrict, and (b) the number of factories that the central village in each subdistrict could plausibly sell to based on 50 kilometer "supply shed" catchments.

I relate factories to farm households with the equation:

$$y_i = \alpha + \gamma fact_s + \beta X_i + \delta_d + \epsilon_i \tag{2}$$

y_i is an outcome of interest for farmer i in subdistrict s of district d (input use, total production in tons, per hectare yields, or farm gate prices). $fact_s$ is either a dichotomous indicator for whether subdistrict s has a palm oil factory or a continuous indicator measuring the number of factories that the central village in that subdistrict could potentially sell to. X_i includes farm size, household size, number of household members, sex, age, and education attainment. δ_d are district fixed effects and ϵ_i is a robust error term, clustered on subdistrict. Assuming no problematic omitted variables varying across subdistricts within districts, γ is the treatment effect of having a factory in your subdistrict or the marginal effect of an additional buyer.

Table 5 presents the results. Columns 1 and 2 examine the quantity of urea used by farmers, since it is one of the main fertilizers and my earlier results found more input suppliers (i.e., agricultural kiosks) near factories. Farmers in subdistricts with a factory use on average 30% more urea. Having an additional factory to sell to leads to around 8% percent higher use. Columns 3–6 examine log production in tons and log yields per hectare. Both are higher in subdistricts with factories and when there are more factories to sell to. Indonesia's widely discussed yield gap thus appears to manifest as a function of distance to buyer and market density. Given the prominence of "nucleus-estate" and other smallholder-firm partnerships, these patterns may be partly due to firm extension activities to smallholders nearby.

Columns 7 and 8 of Table 5 turn to the price farmers get for their fruit. Total value received in Indonesian rupiah is divided by production in kilograms. Column 7 finds that having a factory in their subdistrict sees farmers get paid approximately 30% less for their fruit. My data do not allow me to infer as to whether this is due to being tied to one particular buyer close by, having lower quality fruit, or some other factor, but this is an important question to explore in future work. Moreover, this finding is particularly puzzling because we would expect farmers without a factory in their subdistrict to pay higher transport costs and have their fruit deteriorate more with transport. These estimates mostly net out such quality and transport cost deductions. Column 8 reports the impact of local market density on farm gate prices. For each additional factory a farmer can sell to, she gets a five percent higher price. In the context of rural agricultural markets often characterized by monopsony on the part of large firms, this last finding implies that increased competition among factories allows farmers to command a significantly higher price for their fruit because they are no longer price-takers bound to one particular buyer. Through reduced trade costs and a decrease in local monopsony power—or, the scope for price discrimination among different potential suppliers—improved local market integration for palm oil farmers may be an important channel mediating the direct income gains and thus the overally agglomeration findings.

6 Concluding remarks

This paper used Indonesian palm oil factories as a natural experiment to study industrial onset and estimate spillovers from agricultural processing. I found that the proliferation of export-oriented agricultural processing factories is reshaping the Indonesia's rural economic geography, generating economic spillovers well beyond those accruing directly to farmers in their supply chains. Living in a village near a palm oil factory corresponds to higher household incomes, increased specialization in agriculture, and a greater likelihood of non-agricultural employment. Villages near factories also have more people, large firms, and other economic and social organizations. Villages near factories also generate significantly more revenue, hold more assets, and have better roads, more electrification, more marketplaces, and more health and education facilities. Consistent with this improved local infrastructure reducing trade costs, farmers living in areas with a thicker farm-to-factory market report higher technical input adoption, higher yields, and higher prices, suggesting that these patterns may be at least partially explained by improved local market integration.

By focusing on reduced-form comparisons across villages, my analysis precludes any conclusions on whether Indonesia's dramatic palm oil expansion supports or hinders aggregate structural transformation. Rather, this paper shows highly localized patterns of agglomeration, urbanization, and economic development around factories in otherwise very remote, rural areas. Modern agricultural supply chain investments appear to be planting the roots of broader economic development deep in the Indonesian countryside. Future research could structurally estimate the general equilibrium impacts, explore linkages between cities and rural areas in more detail, estimate the impacts of causal impacts of new factories on fire, deforestation, and land use change, and unpack market access, integration, and competition issues further.

The findings in this paper plausibly inform topical policy debates on global trade, food systems, and the impacts of rising palm oil consumption. While there remains significant work to be done to improve the overall environmental sustainability of our global food system, the Indonesian case highlights the promise of export-oriented agricultural manufacturing as an avenue for developing countries to integrate into global trade networks, attract investment, and improve livelihoods in rural regions with otherwise limited economic opportunities.

References

Allen, T., and Arkolakis, C. (2014), “Trade and the Topography of the Spatial Economy”, Quarterly Journal of Economics, 129: 3, pp. 1085–1140.

Bellemare, M., and Bloem, J. (2018), “Contract farming: a review,” World Development, forthcoming.

Burgess, R., Hansen, M., Olken, B. A., Potapov, P., and Sieber, S. (2012), “The Political Economy of Deforestation in the Tropics”, Quarterly Journal of Economics, 122:1, pp. 73–117.

Bustos, P., Caprettini, B., and Ponticelli, J. (2016), “Agricultural Productivity and Structural Transformation: Evidence from Brazil,” American Economic Review, 106:1, pp. 130–1365.

Byerlee, D., de Janvry, A., Sadoulet, E. (2009), “Agriculture for Development: Toward a New Paradigm”, Annual Review of Resource Economics”, 1: 15–31.

Byerlee, D., Falcon, W., and Naylor, R. (2016), “The Tropical Oil Crops’ Revolution: Food, Feed, Fuel, and Forests”, Oxford University Press, Oxford.

Chenery and Syrquin (1975), “Patterns of Development, 1955–1970,” Oxford University Press, London.

Clark, C. (1940), “The Conditions of Economic Progress”, McMillan and Co., London.

Corley, R., and Tinker, R. (2015), “The Oil Palm”, 5th Edition, Blackwell, Oxford.

Cramb, R. and McCarthy, J. (2016), “The Oil Palm Complex: Smallholders, Agribusiness and the State in Indonesia and Malaysia”, National University of Singapore Press.

Dell, M., and Olken, B. (2018), “The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java”, Review of Economic Studies, forthcoming.

Dercon, S. and Gollin, D. (2014), “Agriculture in African Development: Theories and Strategies”, Annual Review of Resource Economics, 6, pp. 471–492

Dinkelman, T. (2011) “The Effects of Rural Electrification on Employment: New Evidence from South Africa”, American Economic Review, 101:7, pp. 3078–3108.

Directorate General of Estate Crops (2017), “Tree Crop Estate Statistics of Indonesia, Palm Oil, 2015–2017”, Department of Agriculture, December 2017, Jakarta.

Donaldson, D. (2015), “The Gains from Market Integration”, Annual Reviews in Economics, 7: 617-47.

Donaldson, D., (2018), “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure, American Economic Review, 108:4-5, pp. 899–934.

Dragusanu, R., Giovannucci, D., and Nunn, N. (2014), “The Economics of Fair Trade,” Journal of Economic Perspectives, 28: 3, pp. 217–236.

Duflo, E., and Pande, R. (2007) “Dams”, Quarterly Journal of Economics, 2:1, pp. 601–646.

Easterly, W. (2007), "Inequality does cause underdevelopment: Insights from a new instrument," Journal of Development Economics, 85, pp. 755–776.

Edwards, R. B (2018), "Export agriculture and rural poverty: evidence from Indonesian palm oil", working paper.

Engerman, S., and Sokoloff, K. (2002) "Factor endowments, inequality, and paths of development among new world economies," NBER Working Paper Series 9259.

Fafchamps (2012), "Development, Agglomeration, and the Organization of Work", Regional Science and Urban Economics, 42, pp. 459-472.

Fafchamps, M. Koelle, M. and Shilpi, F. (2016) "Gold Mining and Proto-Urbanization: Recent Evidence from Ghana", Journal of Economic Geography, forthcoming.

Farina, E., and Reardon, T. (2000), "Agrifood Grades and Standards in the Extended Mercosur: Their Role in the Changing Agrifood System," American Journal of Agricultural Economics, 82:5, pp. 1170–1176.

Fitriani, F., Hofman, B., and Kaiser, K. (2005), "Unity in diversity? The creation of new local governments in decentralising Indonesia", Bulletin of Indonesian Economic Studies, 74: 1, pp. 70–93.

Fujita, M., Krogman, P., and Venables, A.J. (1998), "The Spatial Economy: Cities, Regions, and International Trade," MIT Press, Cambridge MA.

Gatto, M., et al (2015), " Oil palm boom and land-use dynamics in Indonesia: The role of policies and socioeconomic factors", Land Use Policy, 46, pp. 292–303.

Greenstone, M., Hornbeck, R., and Moretti, E. (2010) "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," Journal of Political Economy, 118:3, pp. 536–498.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014), "The agricultural productivity gap," Quarterly Journal of Economics, 129:2, pp. 939–993.

Gollin, D. (2010), "Agricultural Productivity and Economic Growth", in Handbook of Agricultural Economics, pp 3825–3866.

Hayami, T. (2010), "Plantations Agriculture", in Handbook of Agricultural Economics, pp. 3305–3322.

Hirschman, A.O. (1960), "The Strategy of Economic Development," Yale University Press.

Hsieh, C., and Olken, B. A. (2014), "The missing "Missing Middle"", Journal of Economic Perspectives, 28:3, pp. 89–108.

Johnston, B.F., and Mellor, J.W. (1961), "The Role of Agriculture in Economic Development", American Economic Review, 51:4, pp. 566–593.

Krishna, V. et al. (2017) "Differential livelihood impacts of oil palm expansion in Indonesia", Agricultural Economics, 48, 639–653.

Lipscomb, M., Mobarak, A.M., and Barham, T. (2013) "The development effects of electrification: evidence from the topographic placement of hydropower plants in Brazil", American Economic Journal: Applied Economics, 5:2, pp. 200–23.

Maertens, M. and Swinnen, J.F.M. (2009), "Trade, Standards, and Poverty: Evidence from Senegal," World Development, 37:1, pp. 161–178.

Marden, S. (2018) "The Agricultural Roots of Industrial Development: Rural Savings and Industrialisation in Reform Era China," working paper.

Moretti, E. (2010) "Local labor markets," Handbook of Labor Economics, Volume 4b.

Michaels, G., Rauch, F., and Redding, S. J. (2012), "Urbanization and structural transformation," Quarterly Journal of Economics, 127:2, pp. 535–586.

Myrdal, G. (1957) "Economic Theory and Under Developed Regions", G. Duckworth.

Naylor, R. L., Higgins, M. M., Edwards, R. B., and Falcon, W. P. (2019), "Decentralization and the Environment: Assessing Smallholder Oil Palm Development in Indonesia", Ambio, forthcoming.

Oster, E. (2017) " Unobservable Selection and Coefficient Stability: Theory and Evidence," Journal of Business and Economic Statistics, forthcoming.

Pramudya, E. P., Hospes, O., and Termeer, C. (2016), "Governing the palm oil sector through finance: the changing roles of the Indonesian state," Bulletin of Indonesian Economic Studies, forthcoming.

Rada, N., Buccola, S.T., and Fuglie, K.O. (2011), "Government policy and agricultural productivity in Indonesia", American Journal of Agricultural Economics, 93:3, pp. 867–884.

Resosudarmo, B. (2005) "The Politics and Economics of Indonesia's Natural Resources," Indonesia Update Series, Research School of Pacific and Asian Studies, The Australian National University, ISEAS Press, Singapore.

Robinson, K., and McCarthy, J. (2016) "Land and development in Indonesia: searching for the people's sovereignty," Indonesia Update Series, College of Asia and the Pacific, The Australian National University, ISEAS Press, Singapore.

Rothenberg, A., Gaduh, A., Burger, N., Chazali, C., Tjandraningsih, I., Radijun, R., Sutera, C., and Weilant, S. (2016), "Rethinking Indonesia's Informal Sector," World Development, 80, 96–113.

Swinnen, J.F.M. (2007), "Global supply chains, standards, and the poor," CABI Publishing, Oxford, UK.

World Bank (2008), World Development Report: Agriculture for Development, Washington, D.C.

Tables and Figures

Table 1: Summary Statistics—Palm Oil Processors

Mean	Mean	SD	Min	Median	Max
Estimation sample					
Distance to nearest factory (km)	37.7	27.2	0	31.8	100
Factories within 50km (n)	7.7	16	0	2	106
Factories within 25km (n)	2.6	8.3	0	0	82
N = 59,708 villages					
Full sample					
Distance to nearest factory (km)	93.7	126.4	0	48.5	809.9
Factories within 50km (n)	5.6	14.1	0	1	106
Factories with 25km (n)	1.9	7.2	0	0	82
N = 81,695 villages					

Notes: The dataset for these summary statistics was prepared by the author in geographic information systems (GIS) using 2016 village boundaries and the 2016 Economic Census. All distance calculations use the Asia South Albers Equal Area projection, with minor adjustments. The estimation sample drops villages in cities and on Java.

Table 2: OLS estimates—Factory Proximity, Agglomeration, and Agricultural Employment

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Population (n)							
Near factory treatment (=1)	502.163***	447.850***	428.245***	416.085***	419.011***	532.581***	313.284***
	(36.423)	(34.662)	(35.073)	(35.194)	(35.359)	(40.464)	(42.971)
FE clusters	711	711	709	709	709	217	2882
Village observations	32,203	30,296	29,991	29,991	29,759	29,746	29,762
Panel B: Firms present (=1)							
Near factory treatment (=1)	0.145***	0.138***	0.129***	0.127***	0.127***	0.129***	0.119***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)
FE clusters	760	753	751	751	751	231	2882
Village observations	37,022	34,236	33,880	33,880	33,602	31,112	33,595
Panel C: Primary source of village income is agriculture (=1)							
10 km treatment (=1)	-0.033***	-0.025***	-0.023***	-0.023***	-0.023***	-0.021***	-0.016***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
FE clusters	753	753	751	751	751	231	2882
Village observations	36,238	34,236	33,880	33,880	33,602	31,112	33,595
Baseline geographic controls	N	Y	Y	Y	Y	Y	Y
Additional PODES controls	N	N	Y	Y	Y	Y	Y
Latitude-longitude polynomial	N	N	N	Y	Y	Y	Y
Nearest factory FEs	Y	Y	Y	Y	Y	N	N
District FEs	N	N	N	N	N	Y	N
Subdistrict FEs	N	N	N	N	N	N	Y

Notes: Sample is a cross-section of all Indonesian villages within 100 km of a palm oil processor, excluding villages in cities and on Java. The treatment variable is set to 1 if a village is within 10 km of a palm factory and 0 otherwise. Changes in sample sizes are due to imperfect village matches across datasets. Baseline geographic controls include elevation, slope, historical precipitation, flow accumulation, distance to river, distance to major road in 2000, nighttime luminosity in 1993, distance to city, village area, and an urban dummy. Additional PODES controls include travel time and travel cost to the nearest city, travel cost to the nearest city; river, coast, plain, and valley dummies; in, near, and outside forest dummies, and conservation and production forest dummies. A polynomial in latitude and longitude refers to the latitude and longitude of village centroids and their squared terms. Village, subdistrict, and district identifiers are official 2016 definitions in the 2016 BPS shapefile, and the nearest factory to every village is calculated in GIS using the 2016 Economic Census. Population is taken from PODES 2011, firms from the 2016 Economic Census, and primary village income from PODES 2014. Robust standard errors are in parentheses and stars denote statistical significance at the 10 (*), 5(**), and 1(***) percent levels.

Table 3: IV Estimates—Factory Proximity, Agglomeration, and Agricultural Employment

Instrumental variable	Pr (treatment=1)		GAEZ palm suitability		GAEZ relative suitability	
Column	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Population (n)						
Near factory treatment (=1)	1903.676***	1912.715***	836.744***	822.942***	962.602***	958.383***
	(143.695)	(142.926)	(154.404)	(156.095)	(132.876)	(135.331)
Excluded F	827	832	875	827	1,073	1,018
Village observations	30,288	30,288	30,178	30,178	30,178	30,178
Panel B: Firms present (=1)						
Near factory treatment (=1)	0.393***	0.400***	0.195***	0.207***	0.232***	0.248***
	(0.040)	(0.039)	(0.054)	(0.053)	(0.045)	(0.044)
Excluded F	931	951	950	927	1,172	1,145
Village observations	34,226	34,226	34,080	34,080	34,080	34080
Panel C: Primary source of village income is agriculture (=1)						
Near factory treatment (=1)	-0.114***	-0.109***	-0.041*	-0.034	-0.042**	-0.035*
	(0.020)	(0.019)	(0.024)	(0.024)	(0.021)	(0.020)
Excluded F	931	951	950	927	1,172	1,145
Village observations	34,226	34,226	34,080	34,080	34,080	34080
Baseline geographic controls	Y	Y	Y	Y	Y	Y
Nearest factory FEs	Y	Y	Y	Y	Y	Y
Latitude-longitude polynomial	N	Y	N	Y	N	Y

Notes: Sample is a cross-section of all Indonesian villages within 100 km of a palm oil processor, excluding villages in cities and on Java. The treatment variable is set to 1 if a village is within 10 km of a palm factory and 0 otherwise. Changes in sample sizes are due to imperfect village matches across datasets. Baseline geographic controls include elevation, slope, historical precipitation, flow accumulation, distance to river, distance to major road in 2000, nighttime luminosity in 1993, distance to city, village area, and an urban dummy. A polynomial in latitude and longitude refers to the latitude and longitude of village centroids and their squared terms. Village definitions follow the 2016 BPS shapefile, and the nearest factory to every village is calculated in GIS using the 2016 Economic Census. Population is taken from PODES 2011, firms from the 2016 Economic Census, and primary village income from PODES 2014. Predicted probability is the predicted value for each village obtained from a probit model regressing the main suitability controls on the treatment. GAEZ palm suitability is the normalized value of the grid cell containing the village centroid. Relative suitability is the difference between the normalized potential palm oil yield and the average normalized yield of coffee, cocoa, and teas. Excluded F refers to the Kleibergen-Paap first-stage excluded F statistic; Montiel-Pflueger effective F statistics are similar. Robust standard errors are in parentheses and stars denote statistical significance at the 10 (*), 5(**), and 1(***) percent levels.

Table 4: Reduced-form Falsification Tests

Outcome	Population (n)		Firms present (=1)		Agriculture (=1)	
Column	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Main estimation sample						
GAEZ palm suitability	160.557***	161.838***	0.036***	0.039***	-0.008*	-0.006
	(29.451)	(30.518)	(0.010)	(0.010)	(0.004)	(0.004)
Village observations	30,178	30,178	34,080	34,080	34,080	34080
Panel B: Restricted sample—zero first stage						
GAEZ palm suitability	14.876	9.213	0.013	0.018	-0.005	-0.004
	(31.714)	(33.657)	(0.011)	(0.011)	(0.005)	(0.005)
Village observations	15,516	15,516	17,934	17,934	17,934	17,934
Baseline geographic controls	Y	Y	Y	Y	Y	Y
Nearest factory FEs	Y	Y	Y	Y	Y	Y
Latitude-longitude polynomial	N	Y	N	Y	N	Y

Notes: Full sample is a cross-section of all Indonesian villages within 100 km of a palm oil processor, excluding villages in cities and on Java. Restricted sample additionally excludes villages within 25 km of a factory. Baseline geographic controls include elevation, slope, historical precipitation, flow accumulation, distance to river, distance to major road in 2000, nighttime luminosity in 1993, distance to city, village area, and an urban dummy. A polynomial in latitude and longitude refers to the latitude and longitude of village centroids and their squared terms. Village definitions follow the 2016 BPS shapefile. Population is taken from PODES 2011, firms from the 2016 Economic Census, and primary village income from PODES 2014. GAEZ palm suitability is the normalized value of the grid cell containing the village centroid. Robust standard errors are in parentheses and stars denote statistical significance at the 10 (*), 5(**), and 1(***) percent levels.

Table 5: Farmers, Factory Proximity, and Local Market Density

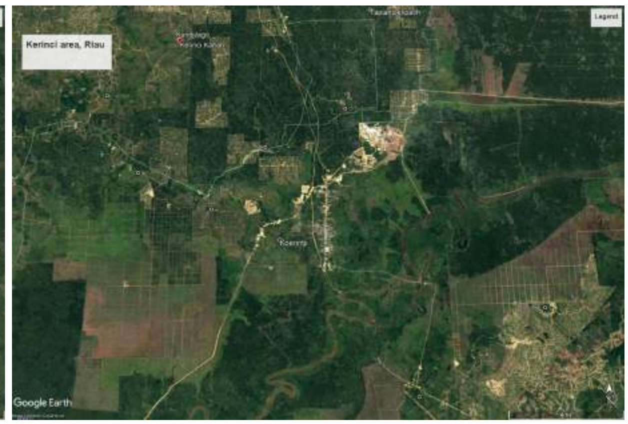
Log dep. var.	Urea (kg)		Production (kg)		Yield/ha (kg)		Price/kg	
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Near (=1)	0.296***		0.888***		1.010***		-0.285**	
	(0.084)		(0.151)		(0.146)		(0.144)	
Number of buyers		0.077***		0.081***		0.018**		0.054***
		(0.005)		(0.008)		(0.007)		(0.007)
Household controls	Y	Y	Y	Y	Y	Y	Y	Y
District FEs	Y	Y	Y	Y	Y	Y	Y	Y
Farmer observations	12,115	16,378	20,359	27,655	20,359	27,655	20,359	27,655

Notes: Sample is a cross section of Indonesian palm oil farmers observed in the farmer cost survey fielded with the 2013 Agricultural Census. The presence of a factory in a subdistrict and the number of potential buyers at a 50 kilometer radius is calculated in GIS, and merged to the household survey. Changes in sample size are due to imperfect merges, missing data, farmers not using urea, and logging zero. District fixed effects and additional controls for farm size, household size, number of household members, sex, age, and education attainment included throughout. Cluster-robust standard errors are in parentheses and stars denote statistical significance at the 10 (*), 5(**), and 1(***) percent levels.

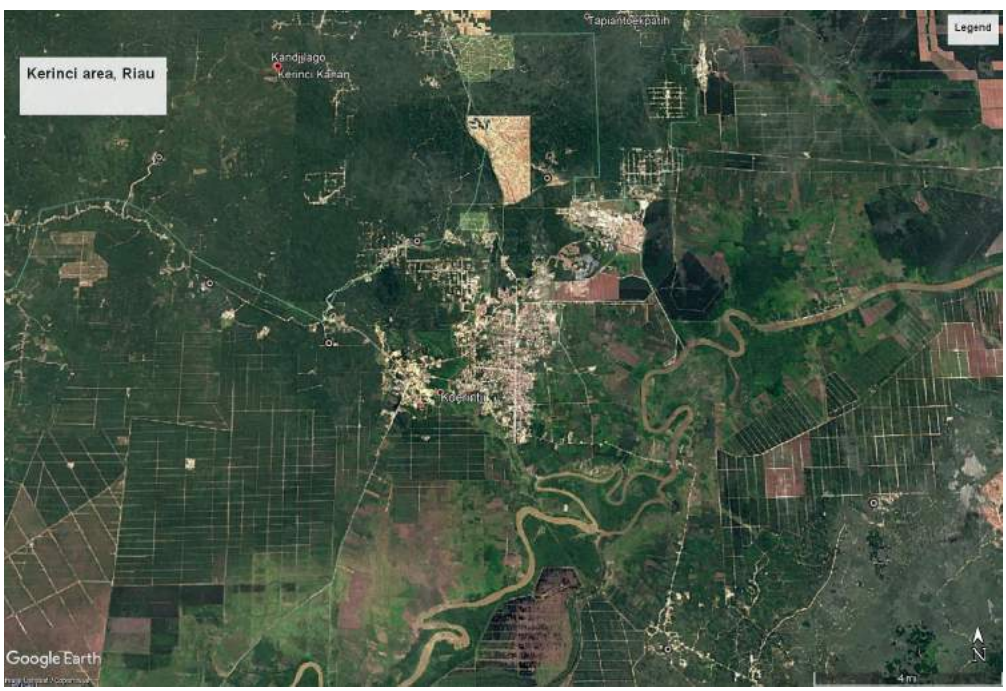
Figure 1: Kerinici Area, Riau—LandSat since 1984

(a) 1984

(b) 2000

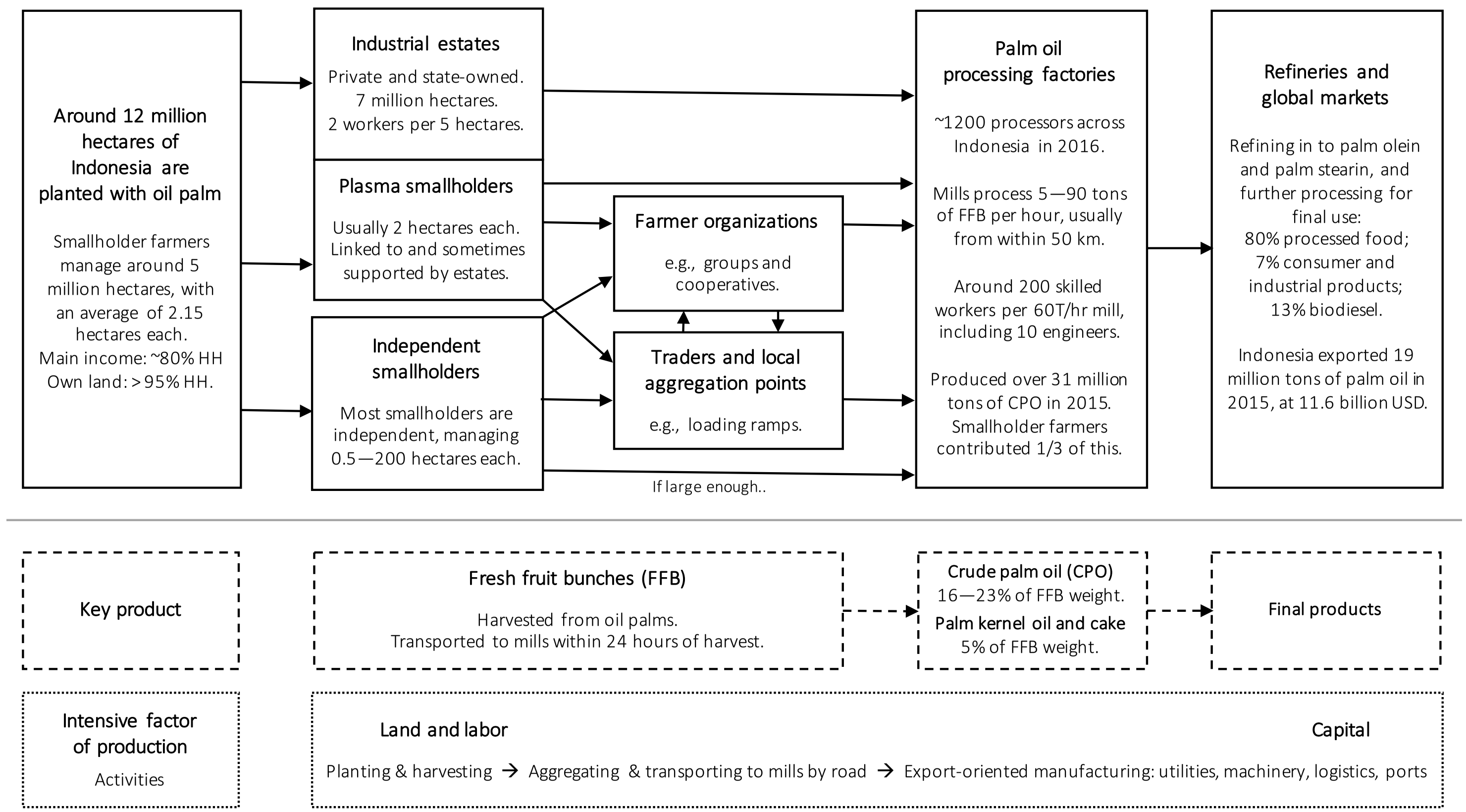


(c) 2016



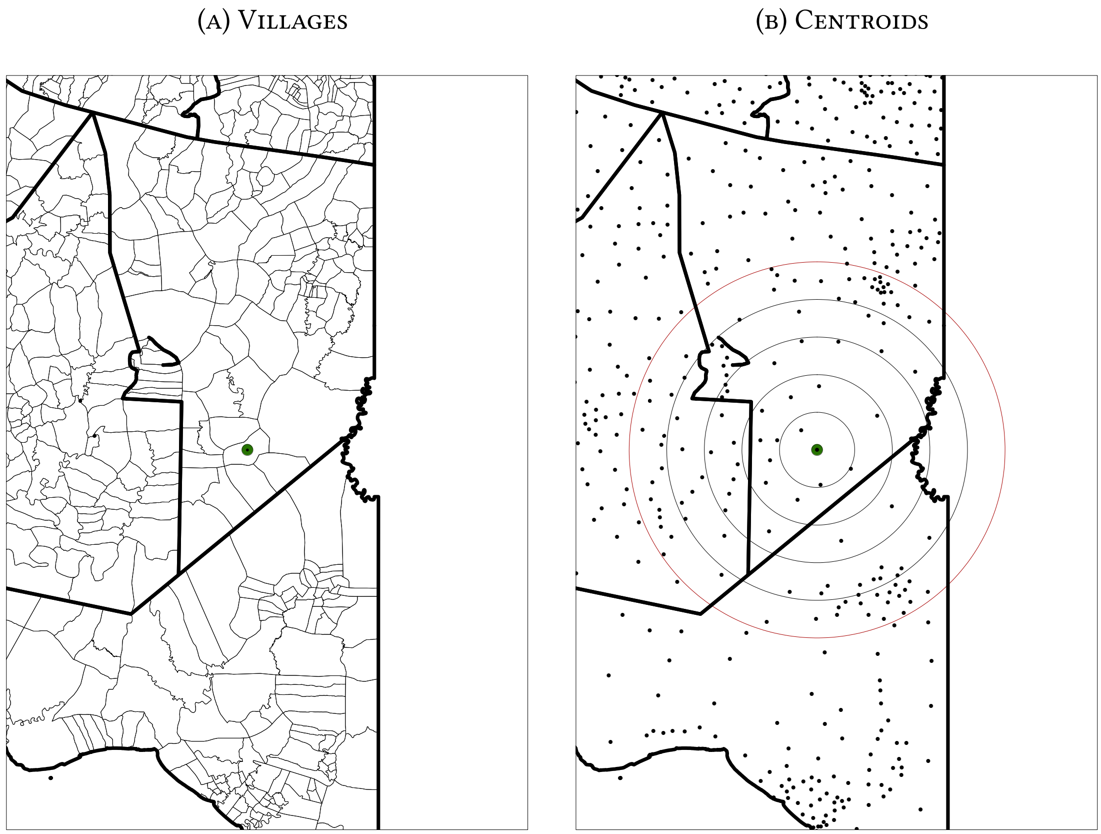
Notes: LandSat imagery is directly as screenshots from Google Earth Pro's desktop historical slider. Modern palm oil processors in 2016 are the black dots in each image.

Figure 2: The Farm-to-Factory Supply Chain



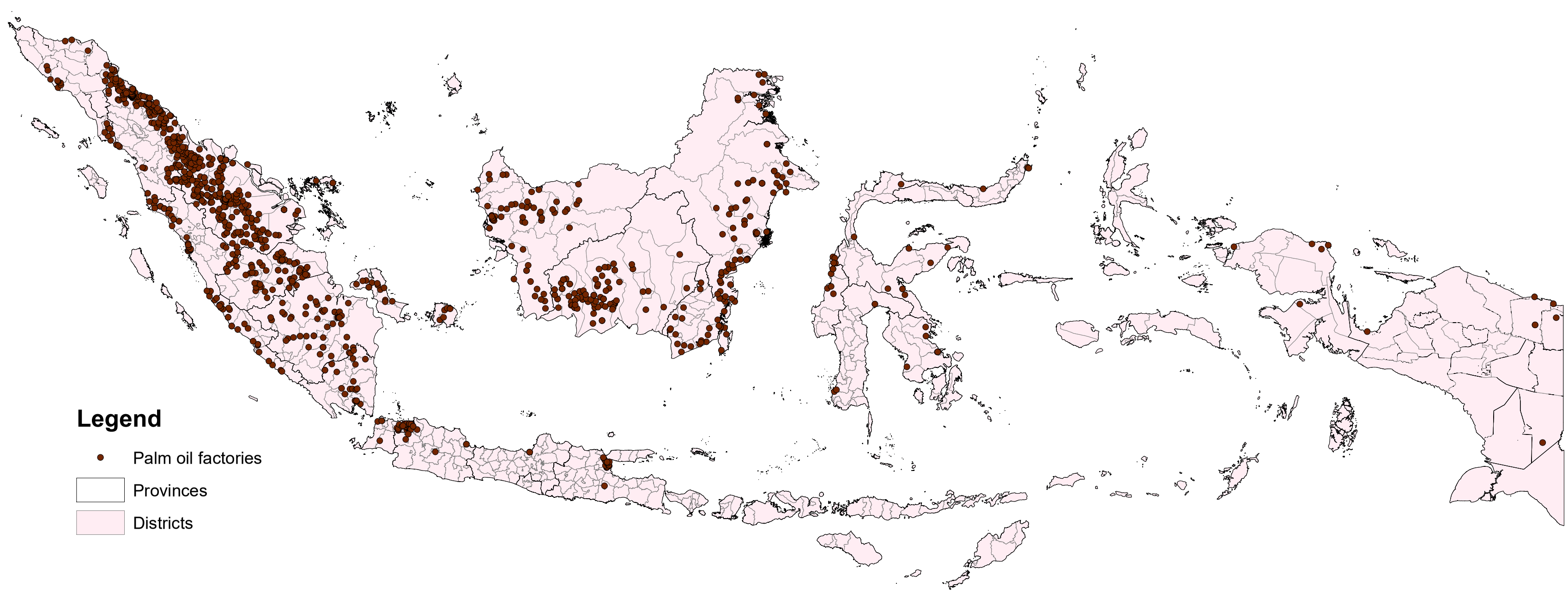
Notes: Author's own depiction. Figures are for Indonesia, from no earlier than 2013, and sourced from official government statistics, site visits, personal discussions, and correspondence.

Figure 3: Research Design Intuition



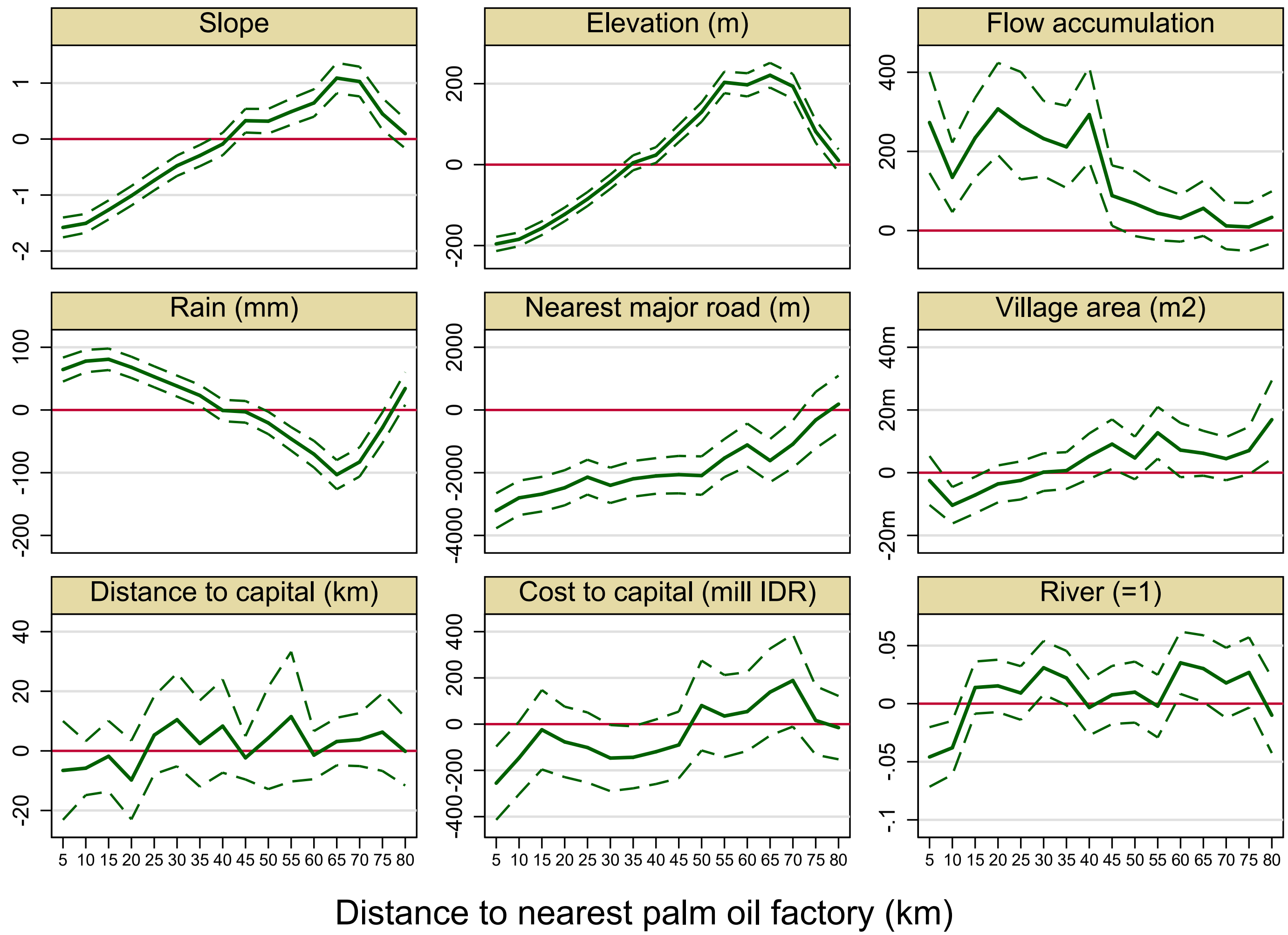
Notes: This figures show the intuition of my research design with the mapping of village polygons (grey lines in A) to village centroids (black dots in B), and the overlay of buffers around a factory (green dot) in Papua province. Dark lines are district boundaries. Each buffer is 20 km farther away from the factory. Beyond the red 100 km buffer village observations are discarded. Please note that (a) all estimates in the article use 5 km intervals, and (b) Papua is a region where villages tend to be larger and factories sparse, allowing a clearer illustration here.

Figure 4: Indonesian Palm Oil Processors, 2016



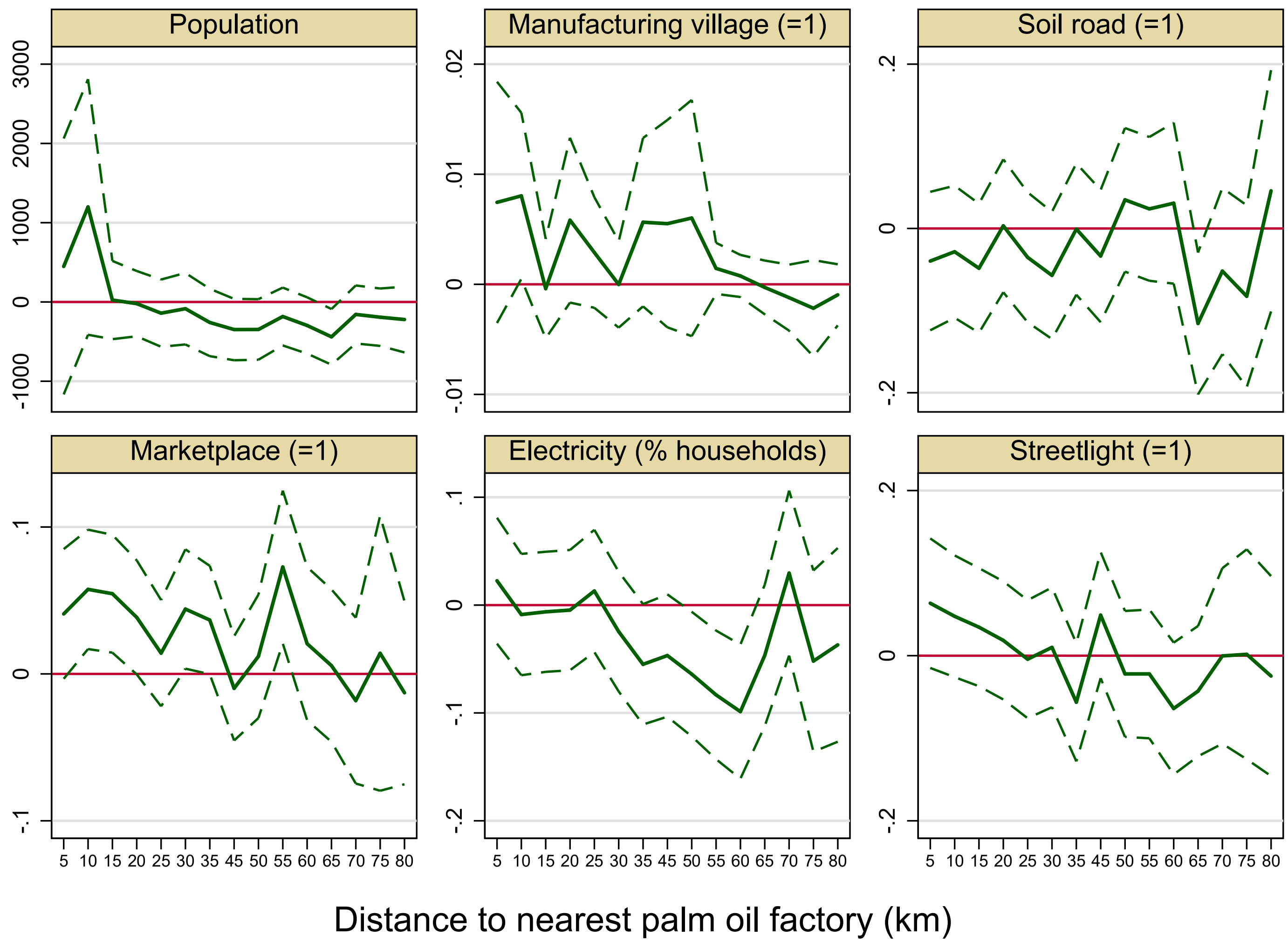
Notes: This figure plots the locations of all palm processors reported in the 2016 Economic Census of Indonesia. Grey lines are districts and dark lines are provinces, as defined in the 2010 Population Census shapefile.

Figure 5: Geographic Selection and Balance Tests



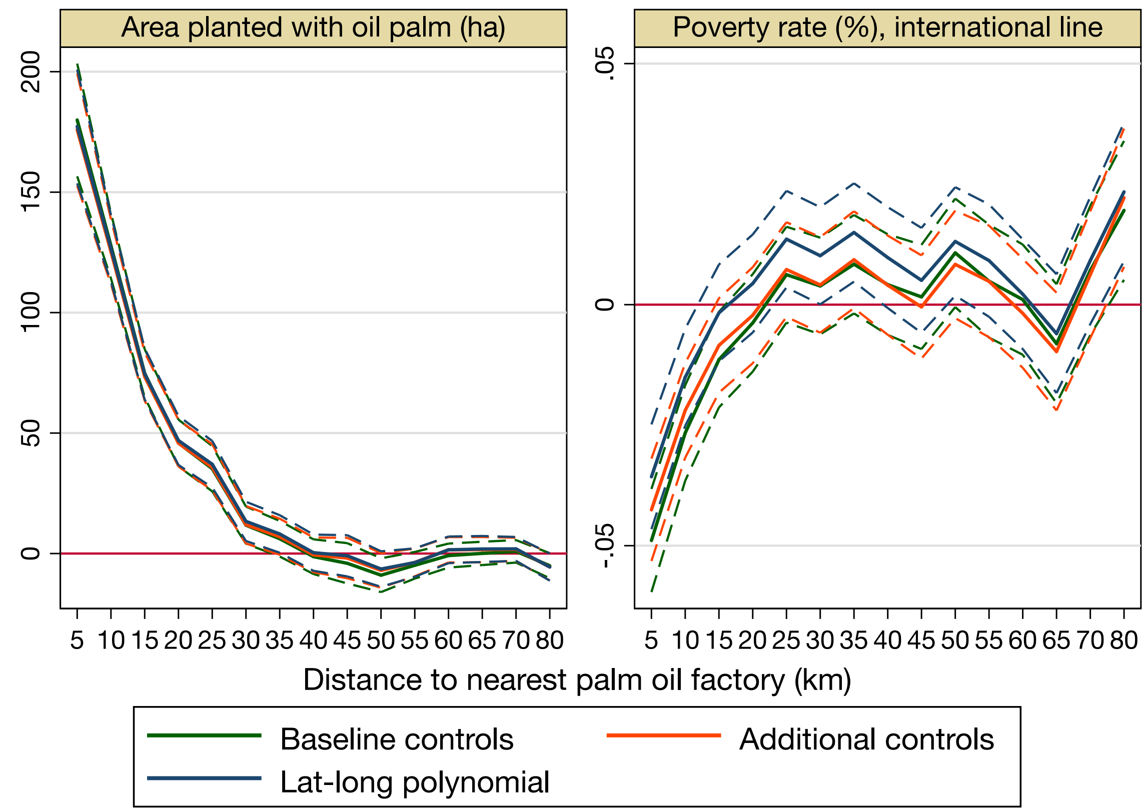
Notes: The figures plot the coefficients from estimating Equation 1 at 5km bins using geographic variables calculated in GIS or taken from PODES 2014 as the dependent variable. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.95% confidence intervals are represented by the dotted lines.

Figure 6: Pre-Period Outcome Placebo Tests



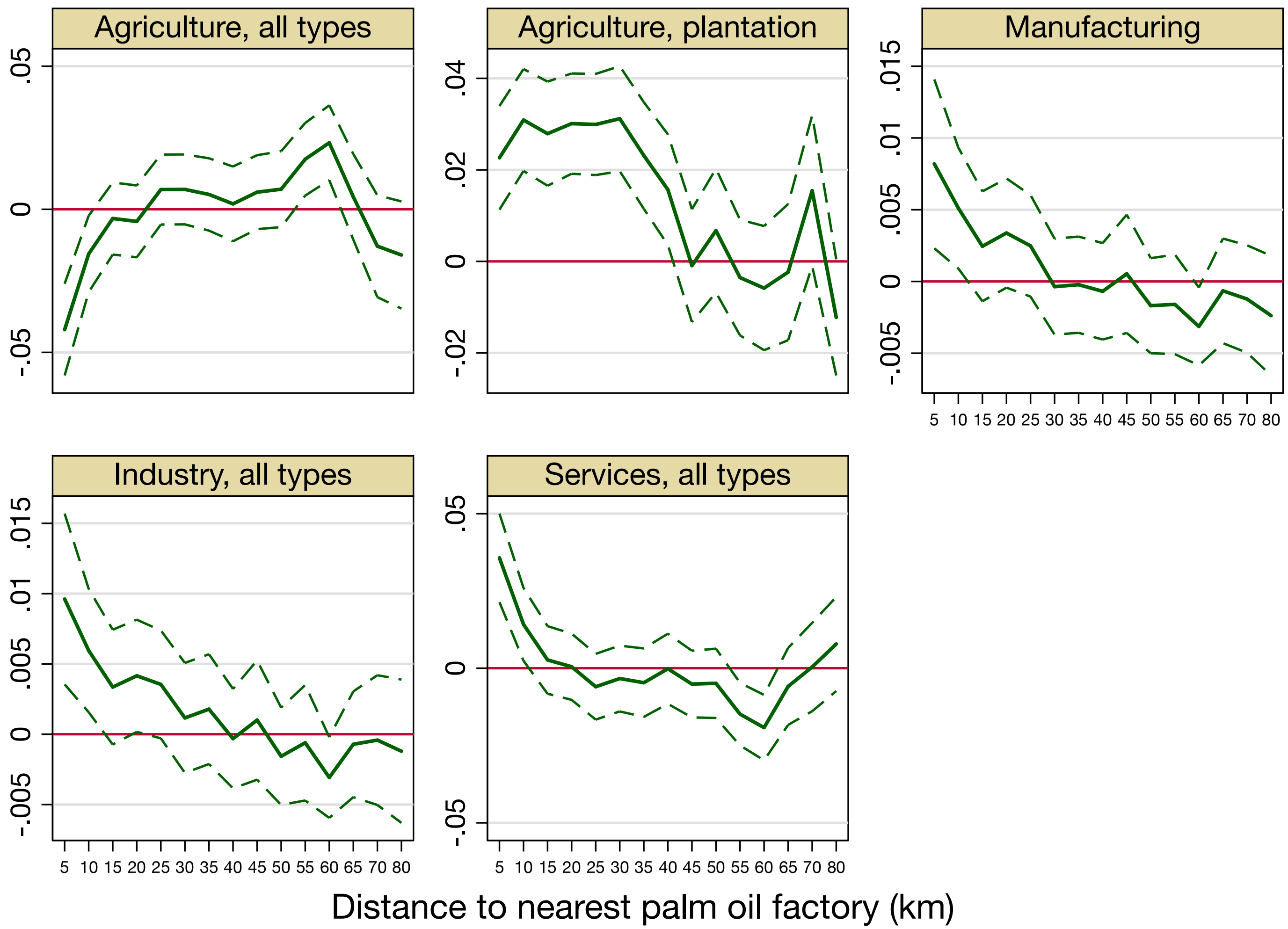
Notes: The figures plot the coefficients from estimating Equation 1 at 5km bins using as a dependent variable pre-period outcomes observed in PODES 1993. The sample excludes villages in districts which produced palm oil before 2000, to focus on new factories. 95% confidence intervals are represented by the dotted lines.

Figure 7: Proximate Adoption and Poverty



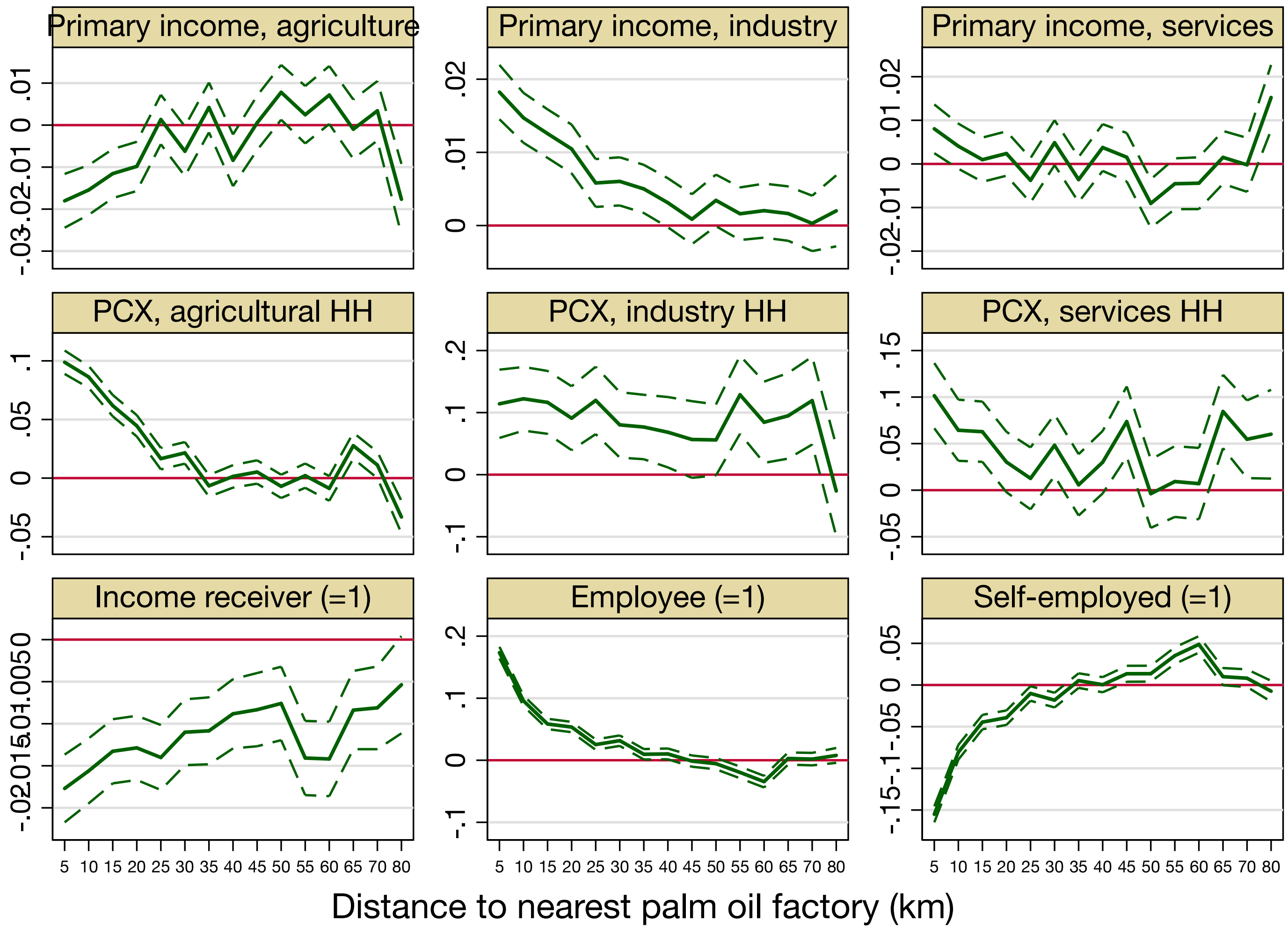
Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using the village oil palm acreage and poverty rates measured at the international line as dependent variables. 95% confidence intervals are represented by the dotted lines. Additional controls add to the baseline controls travel time to the nearest city, travel cost to the nearest city, and indicators for river, coast, plains, valleys, forest proximity (in, near, and outside) and forest function. Lat-long polynomial further adds a polynomial in latitude and longitude to the baseline and additional and baseline controls. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

Figure 8: Primary Sector of Village Income



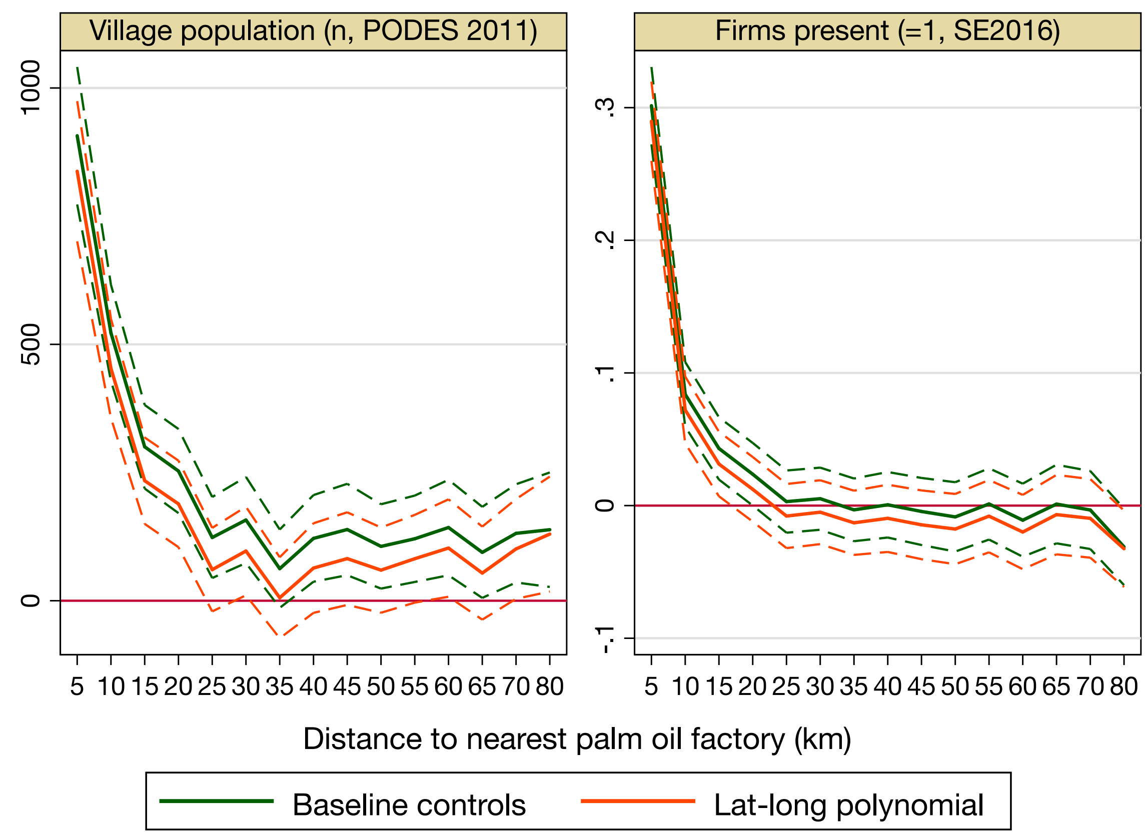
Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using as a dependent variable the main source of income for the village, as measured in PODES 2014. 95% confidence intervals are represented by the dotted lines. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

Figure 9: Employment, Expenditure, and Work Status



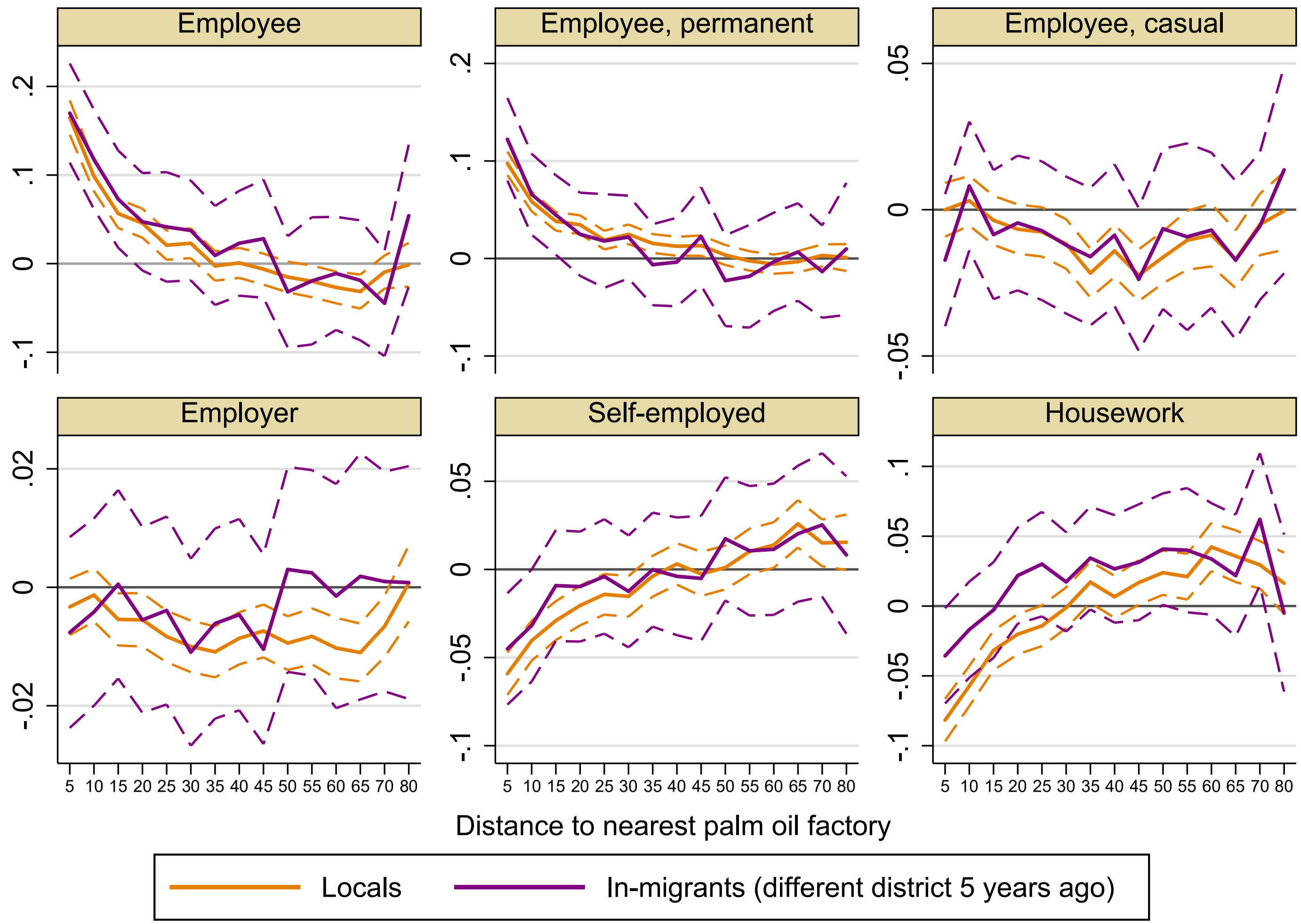
Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using as a dependent variable per capita household expenditures, work status, and the probability of deriving most of your household income from a particular sector, as measured in SUSENAS pooled from 2006–2011. Pooling increases village coverage as the survey is designed to only be representative at the district level. 95% confidence intervals are represented by the dotted lines. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

Figure 10: Population and Firms



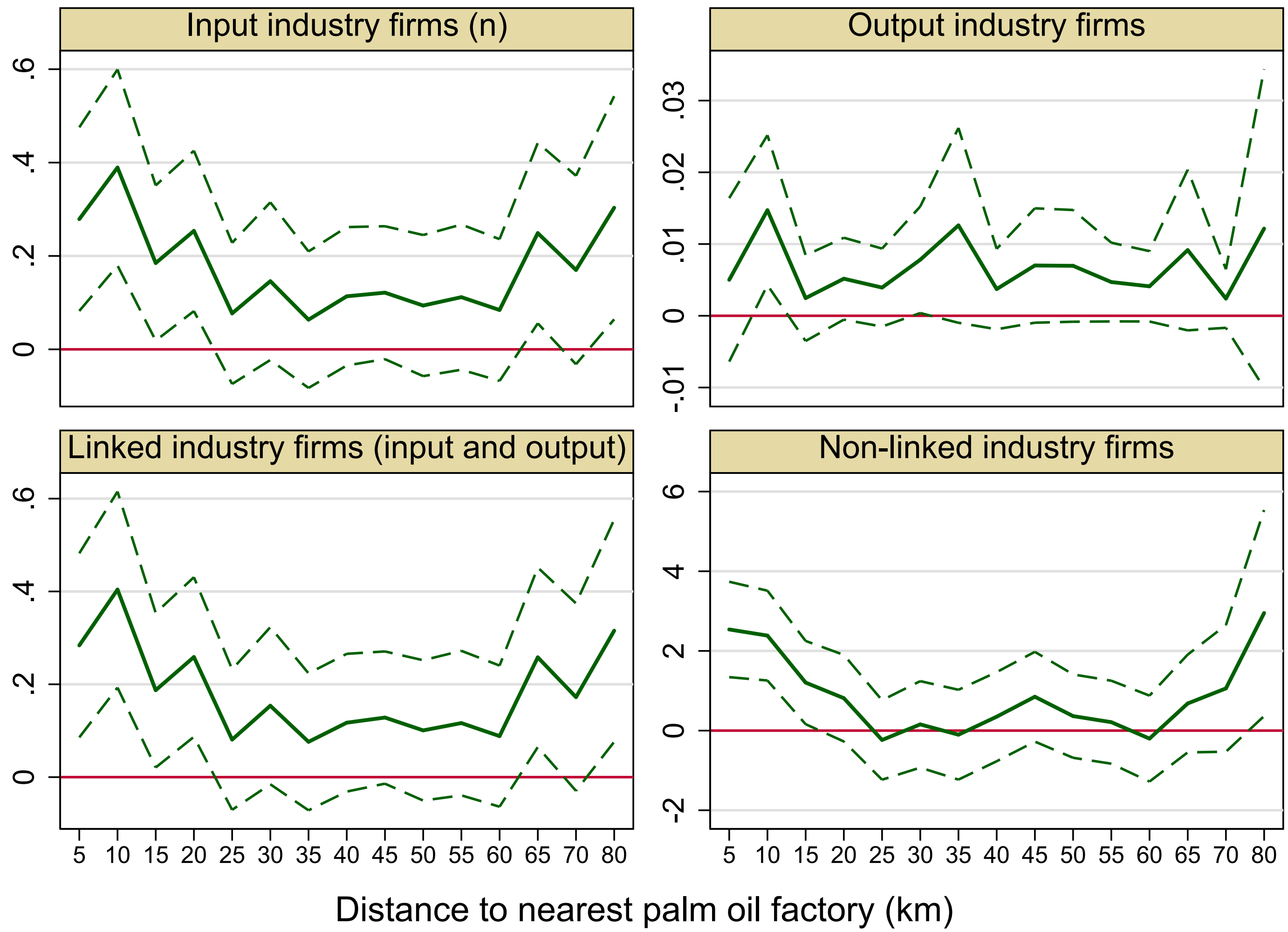
Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using as a dependent variable village population measured in PODES 2011 and the whether a village has a firm present according to the 2016 Economic Census. 95% confidence intervals are represented by the dotted lines. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

Figure 11: Employment, By Migration Status



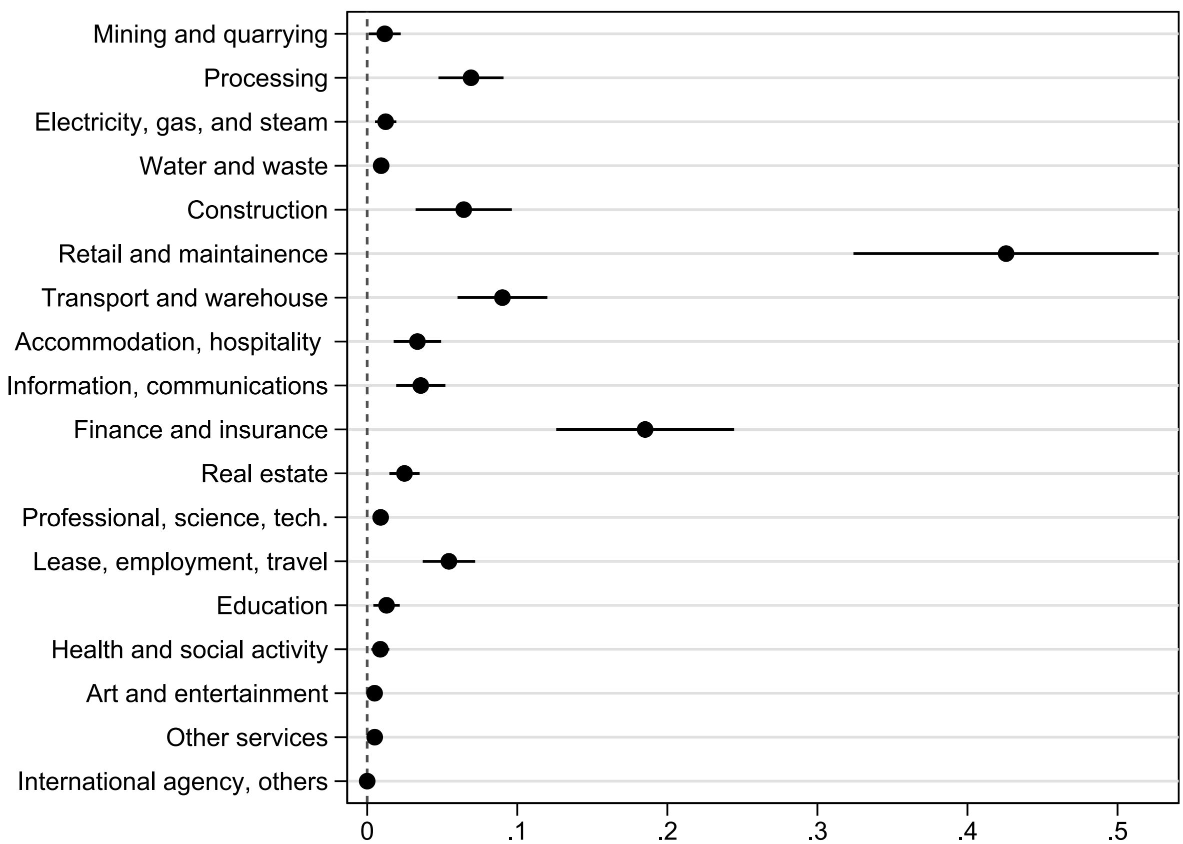
Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using as a dependent variable employment status, as measured in the 2010 Population Census. Samples are split by "locals" (orange) and people who lived in a different district 5 years ago (purple). 95% confidence intervals are represented by the dotted lines. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

Figure 12: Linked Industries



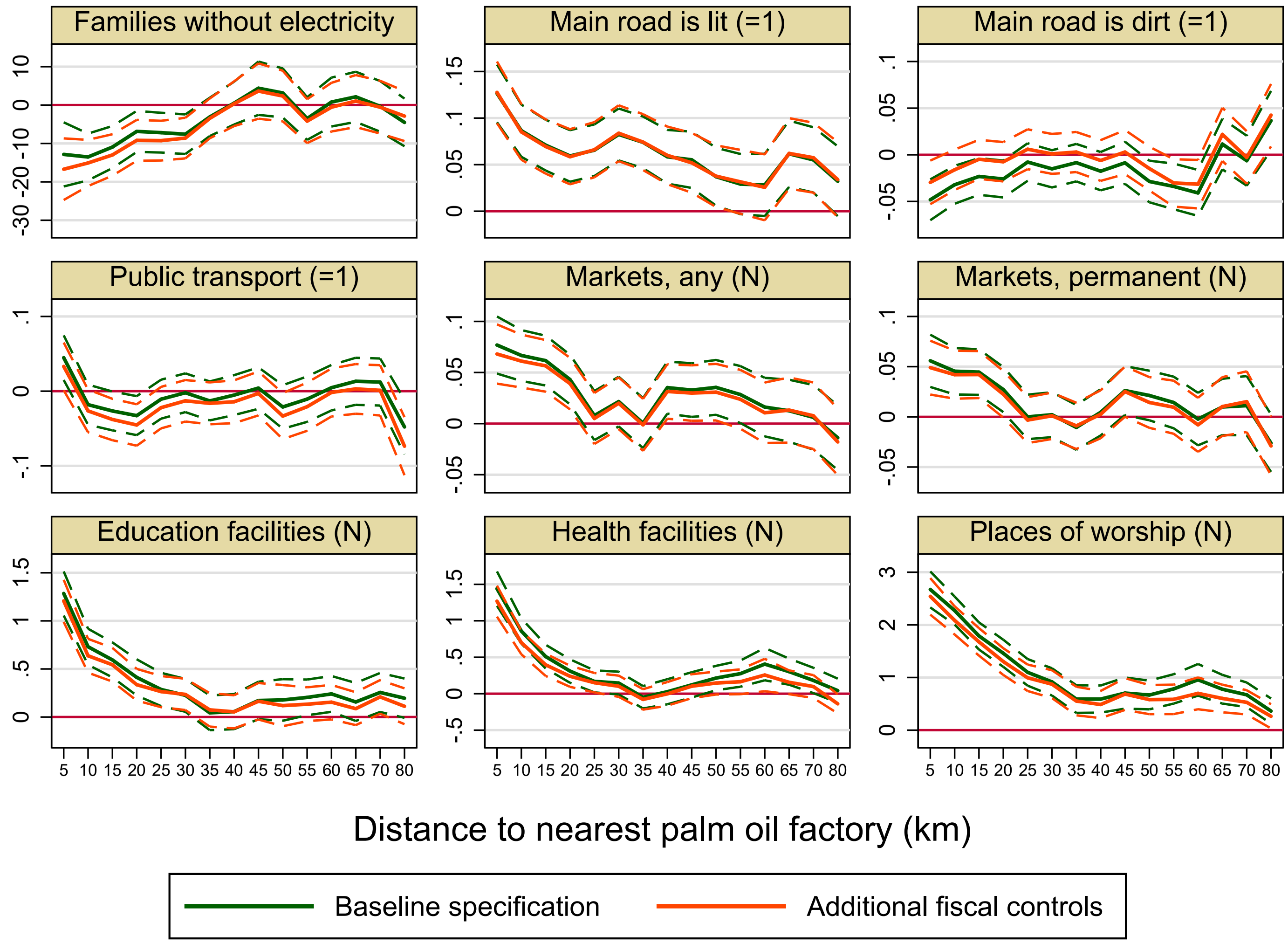
Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using as a dependent variable the number of firms in each village, as measured in the 2016 Economic Census. Input-output linkages are identified through the 2010 Input-Output Table of BPS. 95% confidence intervals are represented by the dotted lines. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

Figure 13: Proximate Firms, By Sector



Notes: This figure plots treatment effects from estimating equation one using as the treatment a 10 km from a factory dichotomous treatment indicator, and using as dependent variables the number of firms by industry. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

Figure 14: Village Public Goods



Notes: These figures plot the coefficients from estimating Equation 1 at 5km bins using as a dependent variable different public good outcomes reported in PODES 2014. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java. The green lines indicate distance band coefficients from my baseline specification, while the red lines indicate those obtained from the same regressions adding village revenue and expenditures as control variables.

Online Appendix—Not For Publication

Appendix A—Data

Overview

Dataset	Years	Source	Variables
Agricultural Census	2013	BPS	area planted by crop
Economic Census	2016	BPS	number of firms by type
Population Census	2010	BPS	employment, status, migration
Village Census (PODES)	1993–2014	BPS	population, organizations, distance to city & capital, other village characteristics
SUSENAS	2005–2013	BPS	sector of employment, household consumption
Poverty map	2015	SMERU	village poverty estimates
Hydrosheds–WWF	N/A	USGS	slope, flow accumulation, distance to river, elevation
Road map	2000	GoI	distance to road
World Port Index	2015	WPI	distance to port
DMSP	1992	NOAA	night time luminosity
Global Agroecological Zones	N/A	FAO	agro-climatic suitability

Dataset and variables description

Agricultural Census, 2013. *Badan Pusat Statistik (BPS).* A census of all agricultural households, conducted every 10 years, which I use to calculate the area planted in each village by crop.

Economic Census, 2016. *BPS.* Census of all firms, conducted every 10 years, which I use to calculate the number of firms by size and sector in each village. The 2016 Economic Census is also used to identify all palm oil processing firms, which are then geolocated to the centroid of the village in which they are domiciled, as observed in the official 2016 BPS village map, using an exact matching procedure on village codes. The nearest processors, the distance to the nearest processor, and the number of possible buyers for fruit for each village are then calculated in ArcMap. Together with the BPS 2010 Input-Output Table, the KBLI 2015 codes assigned to observations in the Economic Census are also used to count the number of forward and backward linked firms in each village. This procedure is described and a list of linked industries provided in Appendix B.

Population Census, 2010. *BPS.* The Population Census is the principal way to measure Indonesia's population, demographics, and migration extent. Migration is assessed through questions on whether individuals live in different district or provinces to that which they lived in 5 years ago or at birth. The census also reports employment by sector and status, which I combine with migration information to examine employment effects by migration status. I draw a ten percent random sample from the complete census, stratified on village, to ease computation.

Village Census, 1993–2014. *BPS.* The census of village heads, Potensi Desa (PODES), is conducted roughly every three years for most all villages (desa) and urban wards (kelurahan) throughout Indonesia. I draw several key outcomes and controls from the village census, for example related to population, village organizations, public goods (schools, health clinics, and churches/mosques), and the primary source of village income. I also draw several key controls, including the distance and travel time to the nearest city and a host of geographical identifiers, for example related to coast proximity, forest proximity, whether a villages is on a slope, plain, or valley, whether a village has a river, and more.

National Socioeconomic Survey, 2005–2013. *BPS.* The National Socioeconomic Survey SUSENAS is Indonesia's main at-least-annual, district-representative household survey, covering over 2 million households in its core wave each year. Because SUSENAS is only representative at the district level, I pool observations over the years for which I have village identifiers to improve village coverage. The primary goal of incorporating SUSENAS is to move beyond the sector of employment (which SUSENAS has) examine expenditures (i.e., consumption).

Indonesia Poverty Map, 2015 *SMERU Research Institute.* The SMERU Research Institute produces a poverty map of Indonesia, which using populations censuses, SUSENAS, and standard poverty mapping techniques (e.g., like those used at the World Bank) to estimate poverty (P0, P1, P2) at Indonesia's lowest administrative level, the village. The SMERU Poverty Map is available from SMERU's website and on request from their research team.

Hydrosheds-WWF *USGS.* The Hydrosheds-WWF online data portal provides extremely fine geospatial data related to geography. Using their various raster and shapefile products and the 2016 BPS village map, I calculate in ArcMap for every village its slope, water flow accumulation, the distance to the nearest river, and elevation.

Major road map, 2000. *Government of Indonesia.* I use a map of Indonesia's major roads in 2000 to calculate in ArcMap the distance to the nearest major road from each 2016 village.

World Port Index, 2015. *WPI.* The World Port Index provides a shapefile geolocating ports all around the world. Using ArcMap and the 2016 village map, I calculate the distance from each village to its nearest port.

Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS), 1992–2015. *National Oceanic and Atmospheric Administration (NOAA), U.S. Government.* I calculate from the Nighttime Lights Time Series Version 4, Defense Meteorological Program Operational Linescan System the average nighttime luminosity for each village. Calculations are done in ArcMap by assigning village centroids the values of the centroids of their nearest raster grid cell. The goal of bringing this data into my analysis is to allow me to capture pre-period local economic development and industry. Using the gridded data and 2016 village maps allows me to preserve all villages in my dataset. By contrast, linking villages back across administrative datasets, which also contain such proxies, is more challenging to match villages over time.

Global Agro-Ecological Zones. *Food and Agriculture Organization of the United Nations.* GAEZ uses agronomic models and high resolution geographic and climatic data to predict attainable yields for different crops on each piece of land regardless of whether the land is cultivated. It does not rely on actual cultivation in its estimates, nor does it involve estimating any sort of statistical relationship between observed inputs, outputs, and agro-climatic conditions. By excluding investments in productivity that might be considered outcomes, GAEZ offers plausibly exogenous variation in crop-specific geographic endowments by construction. The benefit of GAEZ data over using the plausibly exogenous model inputs (e.g., rain, soil type, temperature) is the additional crop-specific yield responses which, without the agronomic model, these variables do directly speak to. Fischer, van Nelthuizen, Shah, and Nachtergaele (2002) detail GAEZ construction. Costinot, Donaldson, and Smith (2016) and Nunn and Qian (2011) discuss additional benefits of GAEZ for identification. Village potential agro-climatically attainable yields for different crops are calculated in ArcMap by assigning village centroids the values of the centroids of their nearest raster grid cell. Note that the level spatial aggregation in the GAEZ data is large than many villages, giving limited variation within a particular locality.

Appendix B—Industries Linked to Palm Oil

Input—51/181 sectors (37th in value)

Oil Palm, Other Seasonal Plantations, Other Annual Plantations, Biopharmaca plants, Livestock and their results except fresh milk, Wood, Agricultural Services, Forestry Support Services, Fisheries Support Services, Tapestry, Rope and Other Textiles, Industrial Preservation and Tanning of Leather, Paper and Cardboard, Industrial Paper and Cardboard Goods, Manufacture of Printed Goods, Basic Chemical Industries Except Fertilizers, Fertilizers, Pesticides, Manufacture of Oil Refinery Products, Plastic Products, Magnetic Media and Optical Media, Glass Industry and Glass Products, Industrial Kitchen Tools, Carpentry and Agriculture from Metals, Hand-Mobilized Hand Tool Industry, Other Metal Goods Industry, Machinery and Equipment Industry, Other Equipment Repair Services, Other Industrial Items, Electricity, Natural and Artificial Gas, Steam / Hot Water Supply, Cold Air and Ice Production, Procurement of Water, Waste and Recycling Processing, Building Construction, Special Construction, Agricultural Infrastructure, Roads, Bridges and Ports, Wholesale and Retail Trade in addition to Cars and Motorcycles, Car and Motorcycle Trade, Car and Motorcycle Repair, Restaurants, Hospitality, Railway Transport, Highway Transportation, Sea Freight, River and Lake Transport, Air Freight, Transportation Support Services, Information Services, Communication, banks, Insurance and Pension Funds, Financial Support Services, and Company Services.

Output—9/181 sectors

Oil Palm, Animal Oil and Vegetable Oil , Peeled Grain, Animal Feed, Pesticide, Drugs, Pharmaceutical Products, Soap and Cleaning Products, Industrial Cosmetic Products, and Other Chemical Goods.

Notes

Linked industries are identified through BPS' 2010 Input-Output Table. Concordance to the KBLI 2015 codes in the 2016 Economics Census was done for each 2010 linked sector by hand, translating then matching codes on descriptions. All output sectors were matched successfully. 9/51 input sectors were not: (1) other seasonal plantations, (2) agricultural services, (3) industrial paper and cardboard goods, (4) industrial kitchen tools, carpentry, and agriculture, (5) hand mobilized hand tool industry, (6) other industrial items, (7) agricultural infrastructure, (8) financial support services, and (9) company services. Agricultural services, tools, agricultural infrastructure, and financial and company services are likely relatively important; estimates without them should be interpreted as such.

Appendix C—Supplementary Figures

Figure A1: Kerinici area, Riau—Google Maps Today

Figure A2: Factory Onset—Kayung Agro area, West Kalimantan

(a) 2000, No Factory

(b) 2010, Factory Construction

(c) 2016, Production

(d) Nearest Road to the East

(e) Road on Google Maps

Figure A3: Aiske area, Papua—LandSat Imagery since 1984

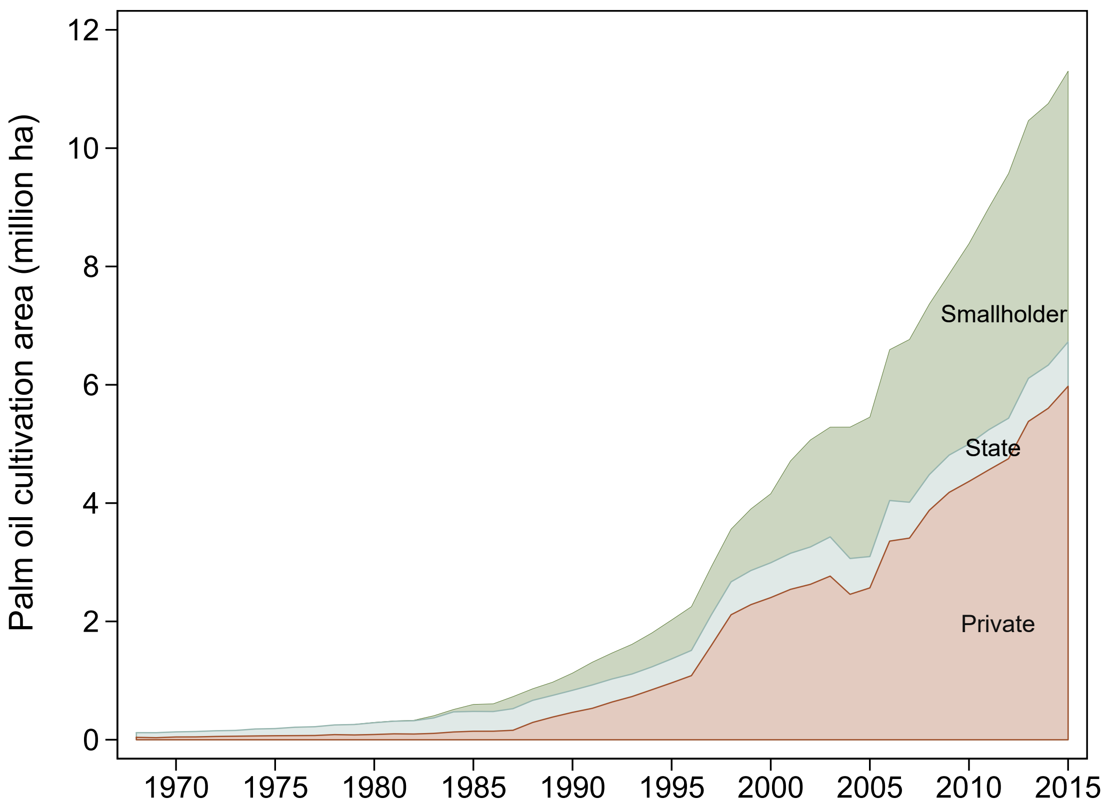
(a) 2000

(b) 2010

(c) 2016

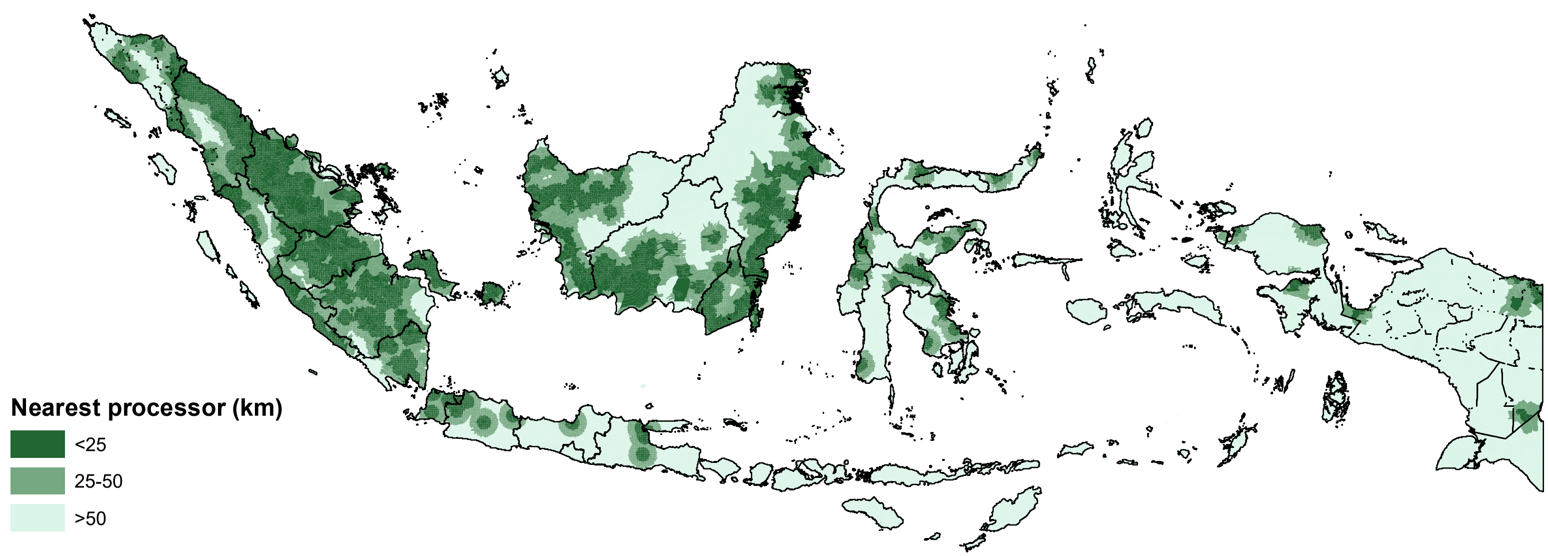
Figure A4: Recent Peri-Urbanization in Bengkalis, Riau—2006-16

Figure A5: Indonesian Palm Oil Area, 1970–2015



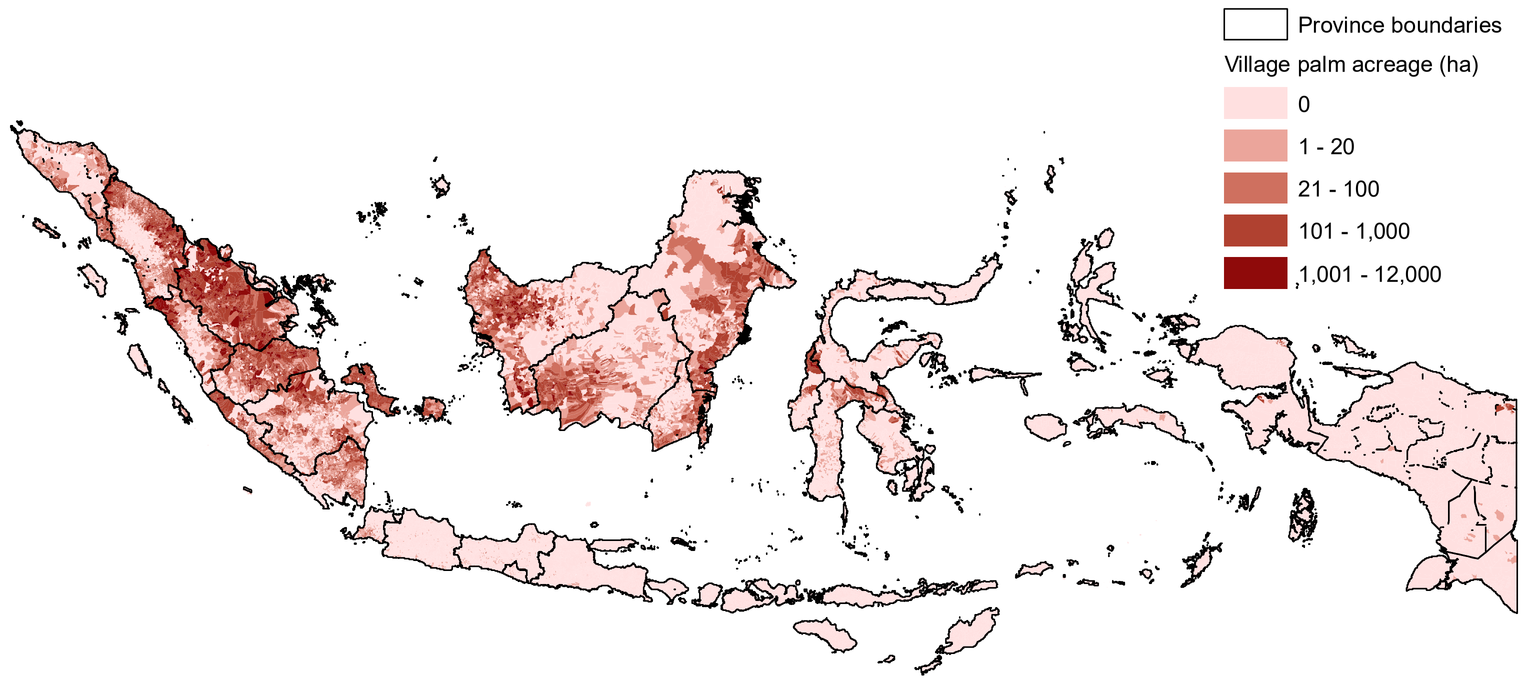
Notes: Data are taken from the Tree Crop Statistics of Indonesia for Oil Palm yearbooks, produced annual by Badan Pusat Statistik (BPS) and the Department of Agriculture of the Government of Indonesia and digitized by the author.

Figure A6: Distance to Nearest Palm Oil Processor From Every Village



Notes: Source: Author's own calculations from the 2016 Economic Census.

Figure A7: Village Oil Palm Acreage



Notes: Calculated from 2013 Agricultural Census.

Figure A8: Processors and Village Oil Palm Acreage—Sumatra

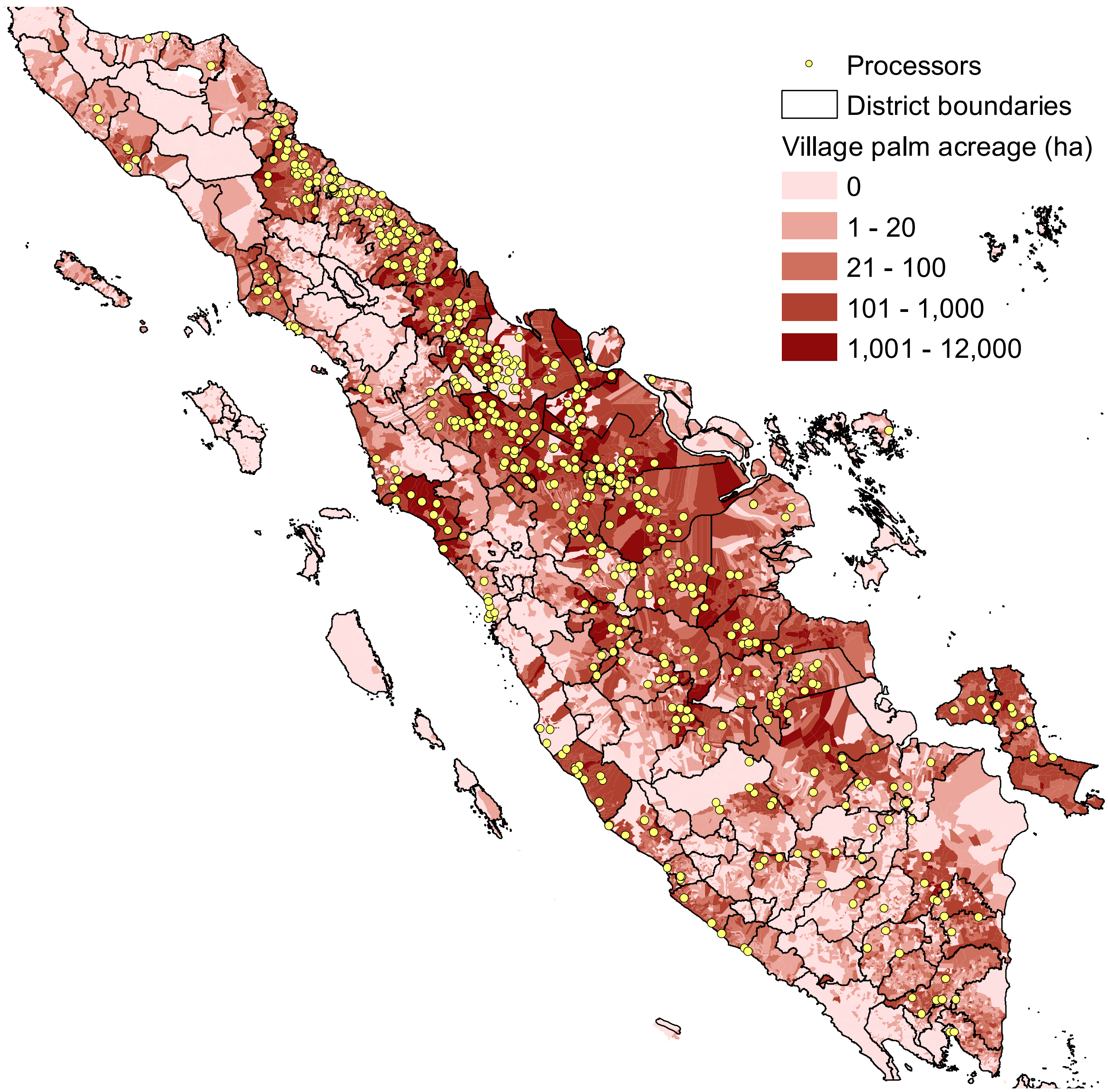


Figure A9: Processors and Village Oil Palm Acreage—Kalimantan

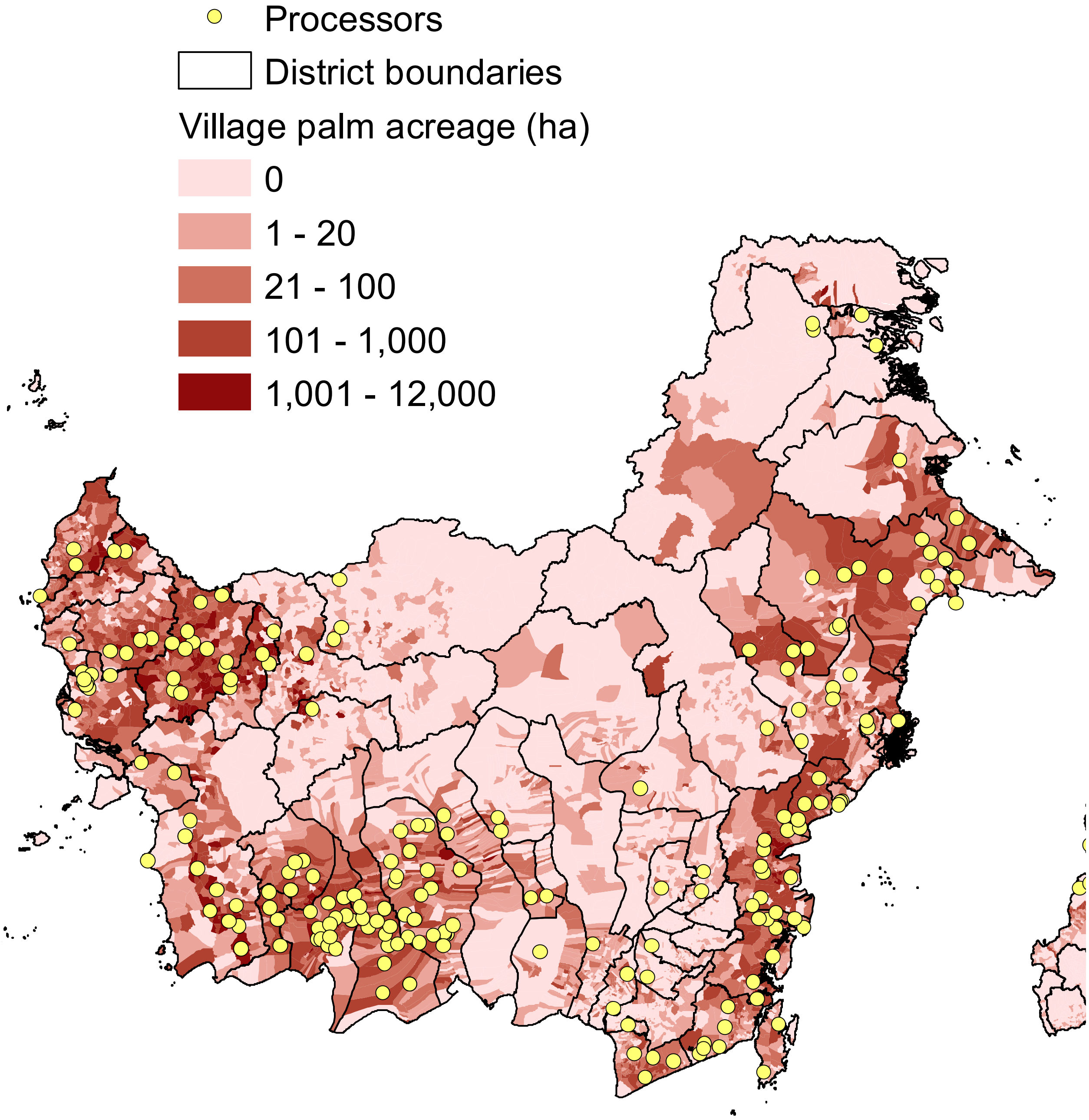


Figure A10: Processors and Village Oil Palm Acreage—Eastern Indonesia

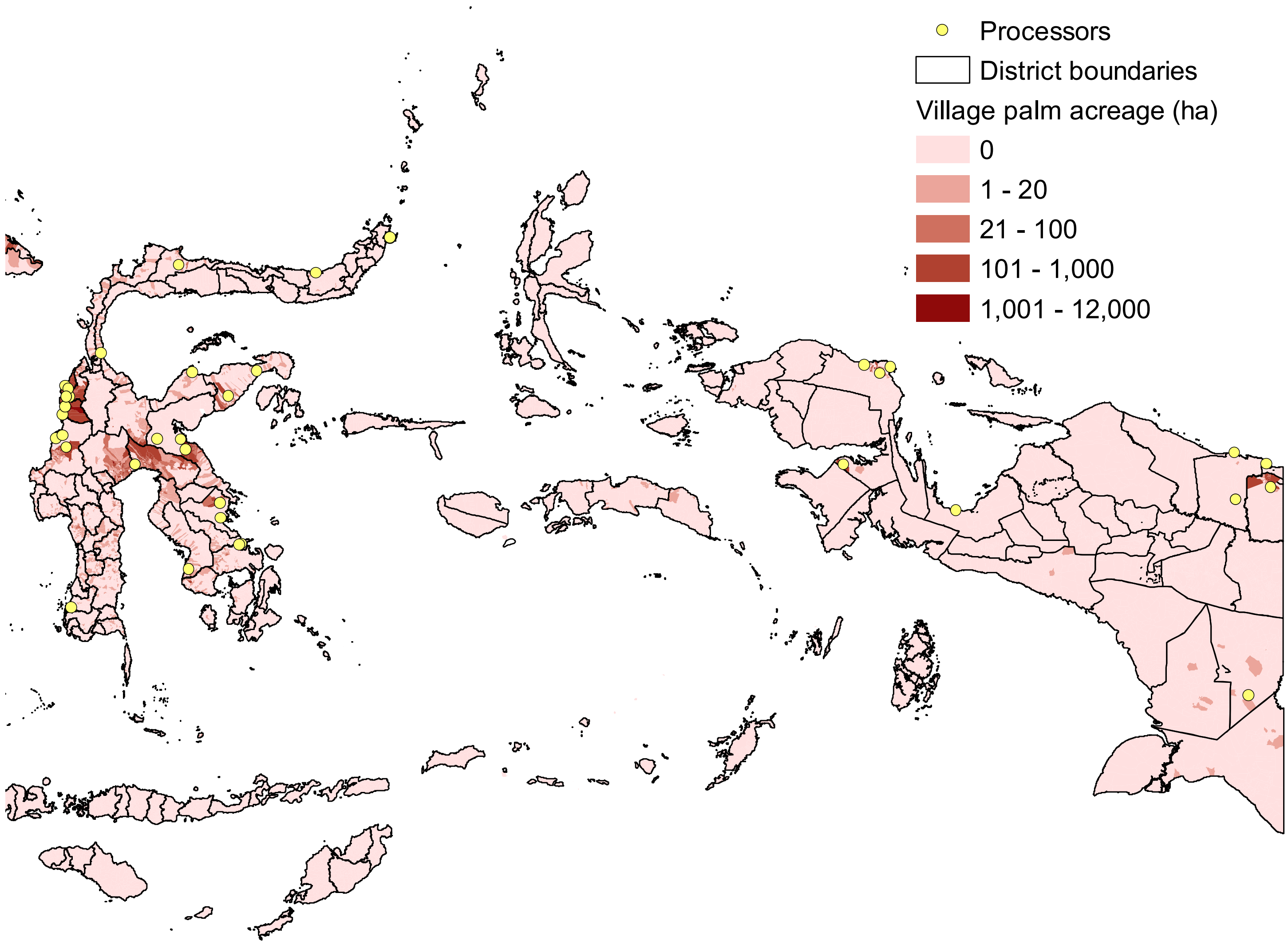
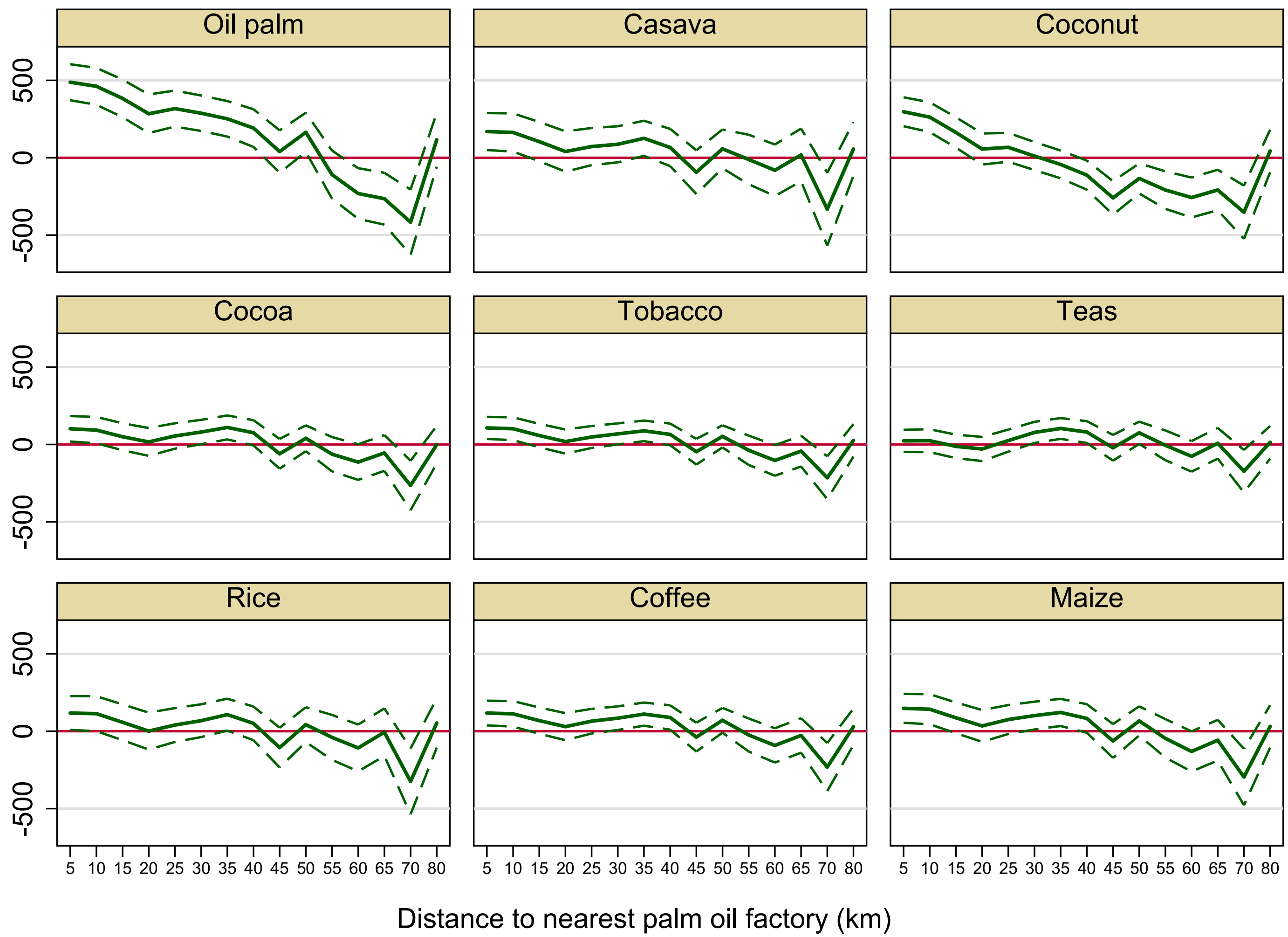
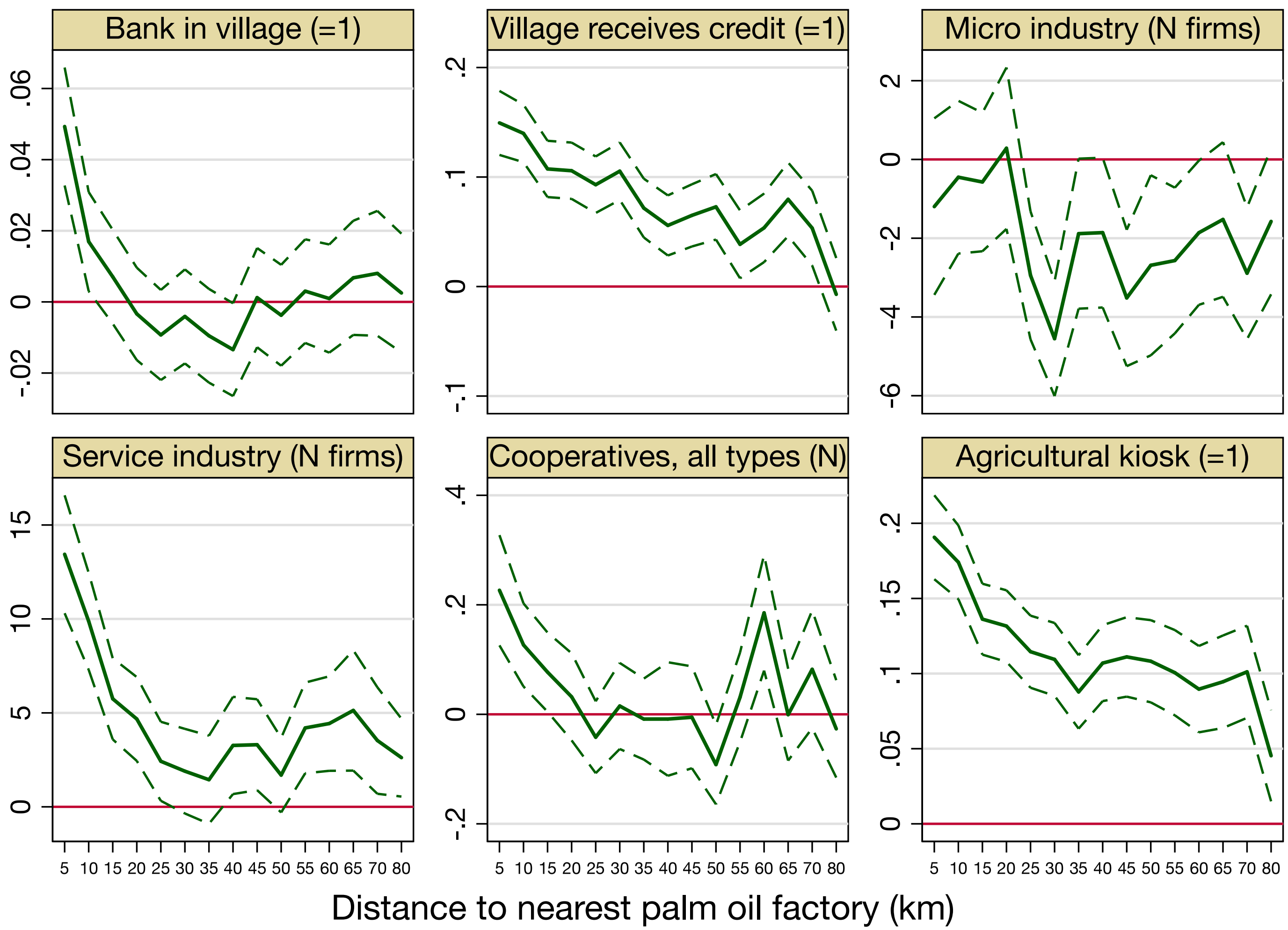


Figure A11: Selection on FAO-GAEZ Agro-Climatic Suitability



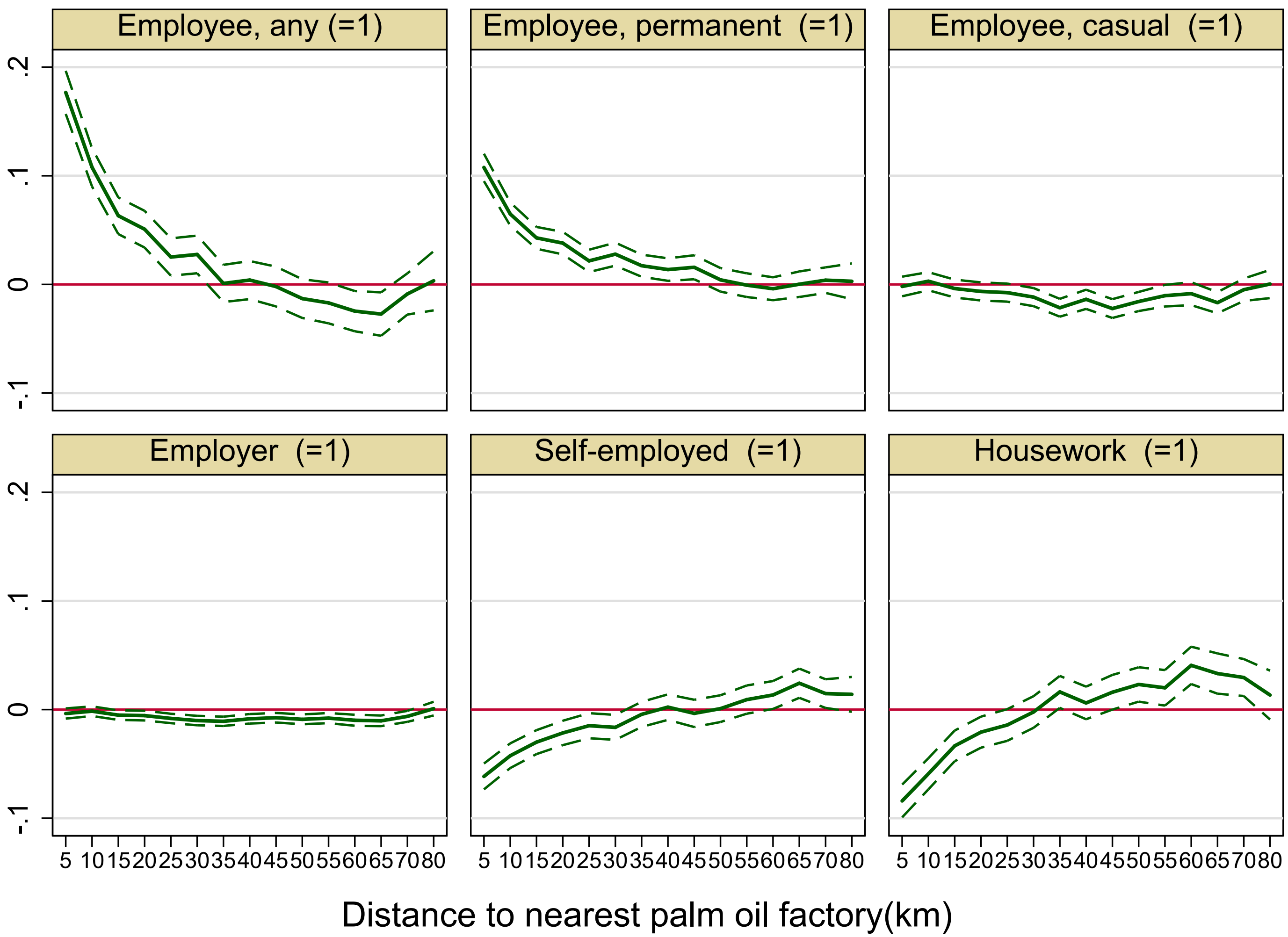
Notes: The figures plot the coefficients from estimating Equation 1 at 5km bins using as dependent variables the agro-climatically attainable yields, for each crop, of the overlapping or nearest gridcell for each village. Confidence intervals are for 95%.

Figure A12: Impacts on Village Credit and Organizations



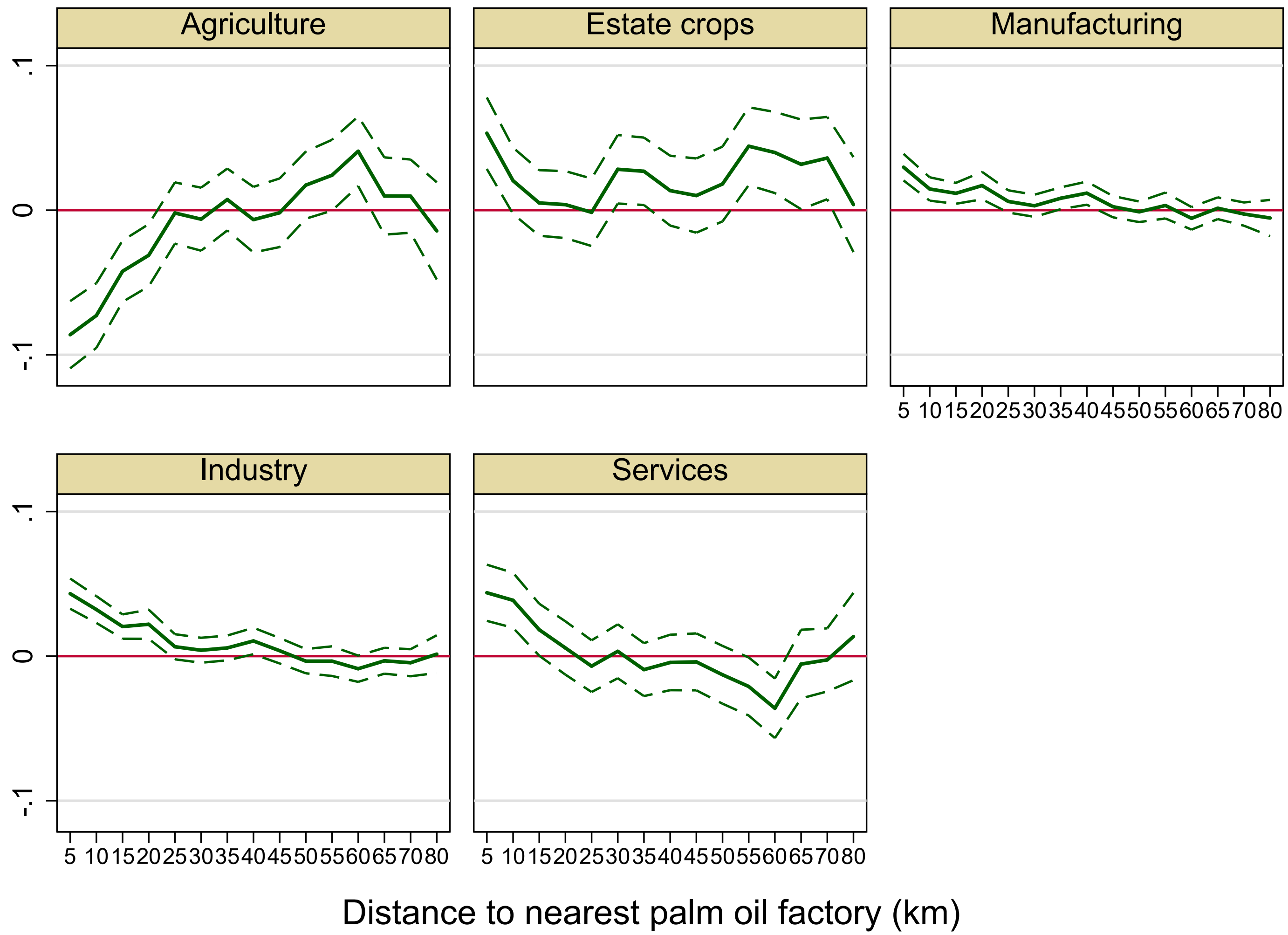
Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using as a dependent variable different outcomes reported in PODES 2014. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java. Confidence intervals are for 95%.

Figure A13: Impacts on Employment Status, 2010 Population Census



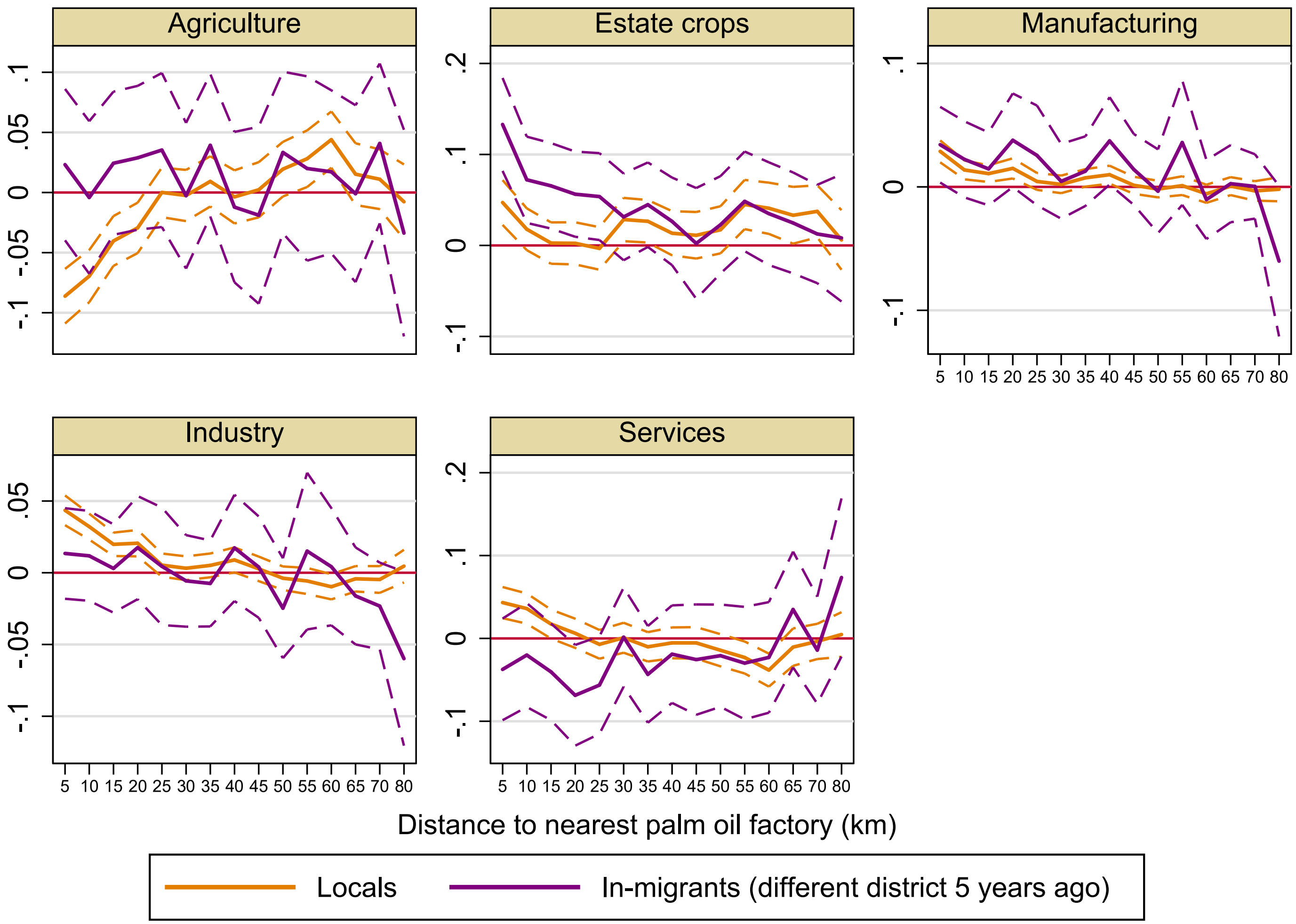
Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using as a dependent the employment status in the 2010 Population Census. Confidence intervals are for 95%. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

FIGURE A14: IMPACTS ON SECTOR OF EMPLOYMENT, 2010 POPULATION CENSUS



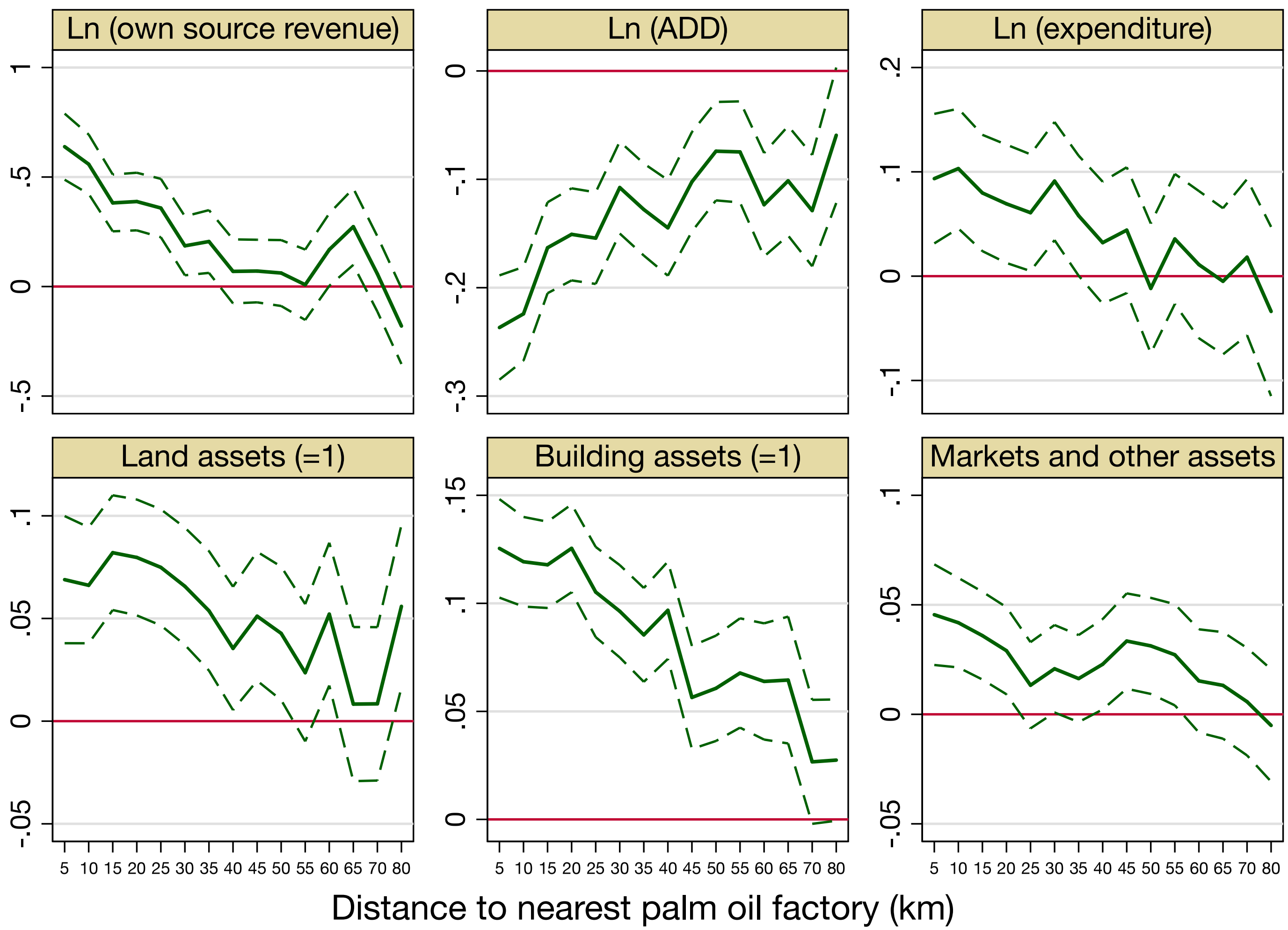
Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using as a dependent variable the main sector of employment in the 2010 Population Census. Confidence intervals are for 95%. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

Figure A15: Employment Sector Impacts By Migration Status



Notes: The figures plots the coefficients from estimating Equation 1 at 5km bins using as a dependent variable the main sector of employment in the 2010 Population Census. Confidence intervals are for 95%. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java.

Figure A16: Impacts on Village Finances and Assets



Notes: These figures plot the coefficients from estimating Equation 1 at 5km bins using as a dependent variable different fiscal outcomes reported in PODES 2014. The sample is all villages within 100 km of a palm oil factory, excluding those that are in major cities (i.e., kota) or on Java. Confidence intervals are for 95%.