Title: Spatial Heterogeneity in Smallholder Oil Palm Production in Indonesia: Implications for Intervention Strategies

Authors: Kibrom T. Sibhatu[1,*], Linda Steinhübel[1], Hermanto Siregar [2], Matin Qaim[1], Meike Wollni[1]

Date: 01-07.2021

Abstract

Oil palm cultivation is a primary income source for millions of rural farm and non-farm households in the tropics but management systems of this tropical crop often vary in space. Understanding this spatial variation and driving factors is crucial in order to design effective and geographically targeted, and optimized interventions that support local farm productivity and sustainability. However, this has been hampered partly due to a shortage of data and methods to examine spatial heterogeneity in smallholder-dominated farming systems systematically. Here, this issue is addressed using primary household data and a structured additive regression model including nonlinear spatial effects—so-called *geosplines*—to analyze micro-level spatial variation in smallholder oil palm yield, input use, and output prices in Jambi Province, Indonesia. We add several standard covariates in our estimation to help investigate the causes of the spatial variation. We identify distinct spatial variation in oil palm production activities within the different parts of the farm households' settlements. Our results show that farm characteristics indicating stability (e.g., land titles) and specialization in oil palm production are associated with significantly higher oil palm yields, input use, and output prices. Further, proximity to a market center significantly increases input use and realized output prices. Finally, the estimated geosplines reveal that standard covariates explain only 50-60 percent of the spatial heterogeneity in our dependent variables. Controls for unexplained variation at smaller scales (e.g., village) can help, yet significant spatial patterns remain for input use and output prices. To explain these remaining patterns, a purely quantitative approach might not be sufficient. Thus, a combination of quantitative and qualitative information might be needed to target and optimize agricultural productivity and sustainability interventions geographically.

Key Words: *Spatial heterogeneity, agricultural intervention, oil palm, sustainability, smallholder farm households, Indonesia*

[1] Department of Agricultural Economics and Rural Development, University of Goettingen, 37073 Goettingen, Germany

[2] Faculty of Economics and Management, Bogor Agricultural University (IPB), Bogor, 16680, Indonesia

* Corresponding author: Department of Agricultural Economics and Rural Development, University of Goettingen, 37073 Goettingen, Germany. Email: ksibhat@uni-goettingen.de

1. Introduction

Smallholder farm households in developing countries represent the largest share of the world's most impoverished population (De La O Campos et al., 2018). Because those households rely on farming for their food and employment, improving their farming systems' productivity and sustainability remains a crucial development goal (UN, 2019). Among other things, progress towards achieving this goal requires understanding how agricultural production activities (farming systems) vary in local geographic space as well as identifying the different factors contributing to this spatial variation (Marenya and Barrett, 2007). Understanding spatial heterogeneity of household farming systems is vital because it helps target farmers that are most likely to benefit from appropriate agricultural (and welfare) policy intervention programs. However, this has been partly hampered due to a shortage of data and methods to systematically examine the spatial variation and its determinants in smallholder management systems.

The topic of spatial heterogeneity in smallholder farm management systems is increasingly gaining attention from researchers recently, with the ultimate aim of supporting agricultural and sustainability interventions. So far, it has been shown that geographic proximity to commercial centers supports the diffusion and adoption of environmentally friendlier and modern farm innovations (Holloway and Lapar, 2007; Knowler and Bradshaw, 2007; Wollni and Andersson, 2014; Tessema et al., 2016; Ebata et al., 2017; Vandercasteelen et al., 2018). Thus, the quality of infrastructure and environmental conditions significantly affect smallholders' access to markets and technology and their possibility to benefit from improved value chains and retail formats.

Interestingly, the studies available investigate spatial heterogeneity only in management systems of annual crops – spatial heterogeneity in perennial crops is yet to be examined. While perennial and plantation crops equally contribute to global smallholder income, their management system inherently differs from annual crops. For instance, they require costly planting investments, several years of growing, and year-round caring to bear yield (Corley and Tinker, 2016). One of these perennial crops largely factoring into smallholder incomes is oil palm. Millions of farm households in the tropics continue to adopt oil palm cultivation; about 50 percent of the worldwide oil palm land is estimated to be managed by smallholders (Qaim et al., 2020). As oil palm yields throughout the year with a low seasonal variation, it allows for a steady income stream independent of seasons (Edwards, 2019). Hence, oil palm expansion is often linked to reducing rural poverty and malnutrition (Sibhatu, 2019; Qaim et al., 2020). Nonetheless, it is also associated with adverse environmental effects and increased social conflicts, particularly in Southeast Asia, where over 85 percent of global palm oil is currently produced (Qaim et al., 2020; Cisneros et al., 2021). Consequently, more and more studies investigate determinants and trade-offs of smallholder oil palm production (Drescher et al., 2016; Euler et al., 2016b; Euler et al., 2017; Qaim et al., 2020). However, to the best of our knowledge, none of these studies explicitly examine spatial variation and its drivers in such perennial crop management systems. Here, we aim to fill this knowledge gap in the literature.

In particular, our study contributes to the existing literature in two ways. First, we investigate the spatial heterogeneity of oil palm production activities (indicated by yield, input use, and price of output) based on primary data of 793 smallholder households located in Jambi Province, Indonesia, a current hotspot of smallholder oil palm production (Clough et al., 2016; Drescher et al., 2016; Romero et al., 2019). Specifically, we evaluate whether the production activity indicators tend to cluster non-linearly and determine the physical location and scale of clustering. Our study area covers about 20,000 square kilometers and, thus, presents multiple environmental (e.g., altitude, rivers, (natural) forests) and infrastructural factors (e.g., cities, mills) to influence oil palm management and create spatial variation. Detecting the areas with higher or lower yield levels, input use, and output prices is crucial in geographically targeting agricultural and sustainability interventions and optimal resource utilization.

Second, our empirical approach allows us to control for unexplained, structured spatial heterogeneity explicitly. Studies examining spatial determinants of (annual) agricultural management systems often pick spatial proxies *a priori*. A frequent example is distance or travel times to the next market center, assuming that costs for transportation affect input and output costs (Damania et al., 2017; Vandercasteelen et al., 2018). However, recent studies also show that such one-dimensional proxies cannot capture more complex and nonlinear spatial patterns (Steinhübel et al., 2020; Steinhübel and von Cramon-Taubadel, 2021). Furthermore, studies usually do not check for any remaining spatial heterogeneity that is not explained by included covariates. Not considering this could lead to biased estimates and the interpretation of results as coefficients might not represent all factors determining clusters in agricultural management systems. By adding a so-called geospline to our predictor, we hypothesize and propose a straightforward approach to (a) investigate how much of the spatial variability in the dependent variables are picked up by standard covariates and (b) visualize remaining spatial patterns.

Our analysis detects significant spatial variation in oil palm production activities within the different parts of the farm households' settlements. Our results also show that farm characteristics associated with long-term stability (e.g., land titles, plantation age) and specialization are likely to improve smallholder oil palm systems (higher yields, input use, and output prices). Moreover, we find that—similar to annual cropping systems—access to market centers is associated with higher input use and realized output prices in smallholder oil palm systems. However, our results also suggest that only about 50 percent of the spatial heterogeneity in oil palm yields, input use, and output prices is explained by standard covariates included in our predictor. Only for oil palm yields, a control for unexplained factors at the village scale can pick up most of the remaining spatial variation. For input use and output prices, however, significant spatial clusters remain.

The rest of this article is structured as follows. The following section provides a brief overview of oil palm production in Jambi province. Data and the empirical strategy are presented in section 3. Section 4 reports and discusses the finding before concluding the paper in section 5.

2. Background on Jambi province and the cultivation of oil palm

A native to Central and West Africa, oil palm (*Elaeis guineensis*) was first introduced to then Dutch administered Sumatra island in the 18th century as an ornamental crop (Corley and Tinker, 2016). While oil palm has been produced commercially in Indonesia since the early 20th century (Qaim et al., 2020), it dramatically expanded recently through transmigration programs and allowing plantation companies to control up to 20,000 ha (Dharmawan et al., 2020). Consequently, oil palm cultivation areas increased massively from 1.1. Million ha in 1990 to 12.3 million ha in 2017 (Badan Pusat Statistik, 2017). Indonesia is now the number one producer and exporter of palm oil worldwide, economically contributing up to 10% of the country's GDP (Qaim et al., 2020).

Jambi province is located in the southeast of Sumatra Island, Indonesia, and was originally covered by tropical rainforests. Grass et al. (2020) report that by 2013 only 34.5 percent of the province was still covered by natural forests. The Sumatra island is known for its industrial plantations, including rubber and industrial woods (Beckert et al., 2014). In the study province, oil palm is mainly cultivated in lowlands and river basins, where the area is characterized by a humid tropical climate and frequent flooding during the rainy season (Merten et al., 2021).

Smallholders are equally active in oil palm production on Sumatra island, accounting for about 61 percent of the island's area cultivated with oil palm (Badan Pusat Statistik (BPS), 2019). Jambi Province was one of the target regions of the so-called *transmigration program* until the late nineties. This program entailed the government-supported migration of farmers from other parts of Indonesia to Jambi province and simultaneously encouraging them to enter company contracts to promote oil palm cultivation (Drescher et al., 2016). However, since then, thousands of farmers in Jambi (and other areas of Indonesia) have also established oil palm cultivation businesses independently without any government support or company contracts. Now, there are more independent farmers than supported smallholders (Schwarze et al., 2015), and oil palm cultivation is, along with rubber, the dominant agricultural practice in the province (Sibhatu et al., 2015; Kubitza et al., 2018a; Sibhatu and Qaim, 2018). Recent surveys show that about 213 thousand farm households in Jambi are involved in oil palm production (Badan Pusat Statistik (BPS), 2018). Unlike contracted farmers, who are tied to companies in long-term arrangements, independent oil palm farmers manage their plantations and supply oil palm fruit bunches to processing mills independently without government subsidies and company interventions (Schoneveld et al., 2019).

Since conversion to oil palm production often promises a higher and stable income, the number of smallholders starting to cultivate oil palm or expanding their production is still increasing (Cahyadi and Waibel, 2016; Santika et al., 2019; Sibhatu, 2019). Consequently, the demand for plantation land is constantly increasing, putting pressure on natural tropical forests because only a few areas remain that could be converted to oil palm plantation, particularly by smallholder farm households (Clough et al., 2016; Kubitza et al., 2018b). An alternative to expanding oil palm cultivation would be to increase the productivity of existing oil palm plantations. And some studies suggest that there is room for such improvement (Euler et al., 2016a), which might slow down the conversion of natural habitats.

3. Methodology

3.1. Household survey

This cross-sectional study uses survey data of 793 smallholder oil palm farmers in Jambi Province, Sumatra, Indonesia (Figure 1). The data were collected as part of the Collaborative Research Centre 990: Ecological and Socioeconomic Functions of Tropical Lowland Rainforest Transformation Systems (EFForTS-CRC990). The EFForTS-CRC990 is an interdisciplinary and collaborative research program that investigates ecological and socio-economic effects of a significant transformation from forests towards a cash crop-dominated landscape of rubber and oil palm (Drescher et al., 2016). The farm households were selected in a multistage-cluster sampling approach (Romero et al., 2019). At the first stage, five oil palm growing districts, namely, Muaro Jambi, Batanghari, Sarolangun, Tebo, and Bungo, were purposely selected (Figure 1). These regencies cover most of the tropical lowland areas in Jambi affected by the ongoing oil palm expansion, and smallholder farm households are the dominant producers (Romero et al., 2019).[1]

At the second stage, a list of oil palm growing villages (n = 90) was compiled using the Village Potential Statistics (PODES) census data of the Indonesian Central Bureau of Statistics (Romero et al., 2019). Out of the 90 villages, 27 were randomly selected. Because the selected villages were all transmigrant households—i.e., households migrated to Jambi from other islands in Indonesia through government transmigration programs—nine autochthonous villages were further purposely included. This brought the total number of villages to 36. At the final stage, farm households were randomly selected from each selected village; the number of households sampled per village varied proportionally to village size. Finally, after excluding a few observations with missing or unclear information, 793 household observations are eligible for our analyses. The full dataset is georeferenced, including village and household coordinates (latitude and longitude), which is crucial for spatial analyses. Furthermore, a wide range of information on socioeconomic, demographic, and

[1] Figure 1 shows that some processing mills are located outside of the area covered in our study, suggesting that some oil production areas are not included in our study area. This is not an oversight during the sample selection, but those areas represent areas where plantation companies are dominant and where smallholder farmers are contracted and supply to plantation companies' processing mills. Moreover, the management activities of those contracted farmers are managed by the plantation companies, not by the farmers themselves. Since we focus our study on independent oil palm producers, we neglect these particular areas of Jambi province.

farm characteristics were captured for each household using a structured questionnaire and face-to-face interviews with household heads. Carefully trained data collectors did the household interviews in the local language Bahasa Indonesia. The survey was conducted between October and December 2015.

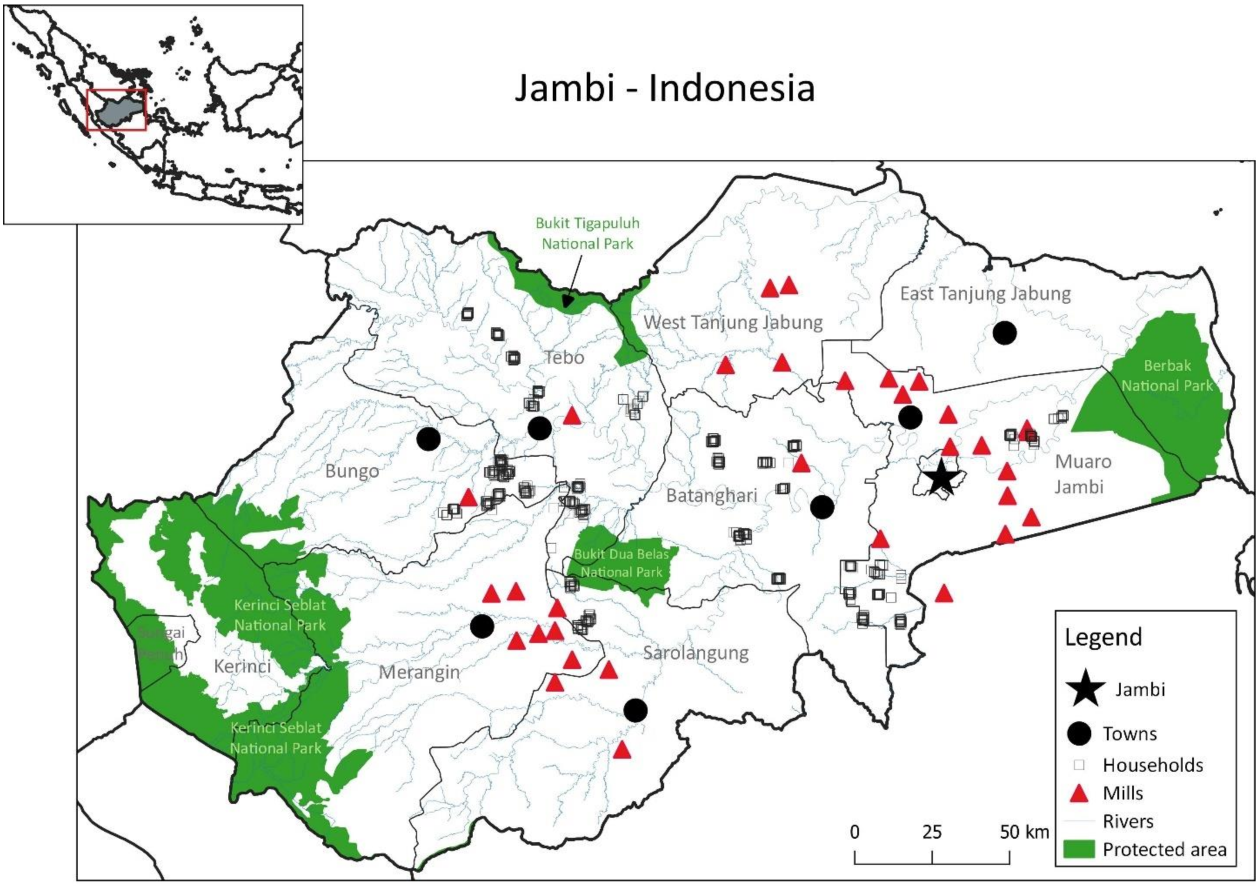

Figure 1. Jambi provivce, Sumatra, Indonesia.

3.2. Outcome variables

This article aims to characterize the spatial heterogeneity of oil palm production activities in smallholder households. To achieve this, identifying outcome variables that reflect farmers' management decisions and intensities is paramount. In the literature, variables commonly used to measure differences in farm management intensity are yield, input use (amount of money invested in inputs), and output prices (Minten et al., 2013; Asfaw et al., 2016; Damania et al., 2017; Ebata et al., 2017; Vandercasteelen et al., 2018). Therefore, based on the survey data, we calculate the fruit bunch harvest in kilograms per hectare and use average prices per kilogram received by farmers in the last 12 months. We also calculate the amount of money invested in inputs (hereafter, we call it input use) for each household's oil palm plantation, equal to the total input cost per ha in the past 12 months expressed in Indonesian Rupiah (IDR). Input use includes the amount of labor, fertilizer, herbicides, and pesticide applied and their respective prices in the past 12 months. As was done in previous studies, we do not disaggregate input use by type (Steinhübel and von Cramon-Taubadel, 2021). This is because many respondents hardly apply many types of inputs, which is a typical characteristic among oil palm farm households (Euler et al., 2016a).

3.3. Covariate selection

Investigating the drivers of spatial patterns in oil palm production systems in Jambi is this study's goal. Table 1 presents all covariates included in our analysis grouped into either socioeconomic or geographic categories. The latter category is particularly important to generate a nuanced insight into spatial patterns of agricultural production systems because they all present exogenous spatially-clustered factors that can have important implications for agricultural productivity and management decisions.

The socioeconomic variables comprise standard household controls such as age and education of the household head, household size, and specific to the Jambi context, whether the household lives in an autochthonous or a transmigrant village (measured at the individual household level). Furthermore, we include acreage variables, the share of plots with systematic land titles, input use, oil palm yield[2], other crops managed besides oil palm, and awareness of the RSPO (Roundtable on Sustainable Palm Oil), the globally most recognized certification scheme for oil palm (Kunz et al., 2019) to capture differences among smallholder households. Finally, we note that all these variables are either inherent to or under the control of the household, and for any spatial clusters in these variables, we cannot assume exogeneity.

In comparison, the geographic variables refer to variables that describe the environment in which smallholders navigate their farming activities and patterns beyond the (spatial) scale of the household. Assuming that these larger-scale factors are important drivers of the geographic variability in oil palm production, we define four subgroups based on spatial proxies commonly used in the literature. These are infrastructure, geomorphology, administrative units, and exact household location in space. By pulling together those different spatial factors

[2] Input use and oil palm yield are outcome variables as well. Thus, input use is only considered as explanatory variable with the yield outcome variable. Oil palm yield is considered as explanatory variable with the price outcome variable.

in one analysis, we aim to present a framework that allows for a systematic analysis of geographic patterns in agricultural production systems.

Table 1. Types of explanatory variables included in the selection algorithm

Type	Indicator	Remark
Socio-economic variables		
1. Household characteristics	Age of household head	Years
	Education	Years
	Household size	Count
	Type of village a household belongs	Dummy: autochthonous =1; 0 = transmigrant
2. Production characteristics	Total land owned	Ha, log
	Input cost/ha	'000 Indonesian Rupiah (IDR)
	Oil palm yield	Kg/ha a year
	Age of plantation	Years
	Share of plots with systematic land title	Percentage (0-100)
	Non-oil palm production	Dummy
	Awareness of RSPO certification	Dummy
Geographic variables		
3. Infrastructure characteristics		
	Distance to the nearest city	Km
	Distance to Jambi city	Km
	Distance to the nearest mill	Km
4. Geomorphological characteristics	Altitude	Meters above sea level (masl)
	Access to a river	Dummy: Yes =1; No=0
	Household located at the edge of a forest	Dummy: Yes =1; No=0
5. Administrative units	Regency	Categorical variable
	Village	Categorical variable
6. Household location	Household location	Longitude, latitude

The infrastructure characteristics contain proxies for households' access to markets, such as the distance to the closest town, the distance to the province capital Jambi, and the distance to the closest palm oil mill. Access to urban centers and markets will likely lead to more intensified production systems (Damania et al., 2017; Vandercasteelen et al., 2018; Steinhübel and von Cramon-Taubadel, 2021). With proximity to markets, transportation costs normally decrease and, thus, households pay lower net input prices and receive higher net output prices. As a consequence, farmers might be more likely to modernize and intensify their production systems.

Furthermore, previous studies have shown that biophysical and geomorphological factors significantly affect oil palm management intensity (Kubitza et al., 2018b; Romero et al., 2019; Merten et al., 2021). To capture this aspect, we include the distance to the edge of a forest, access to rivers, and altitude above sea level as geomorphological variables in our analysis. Access to a river might affect the possibility of irrigation. While farms located closer to forest edges are often more likely to lack formal land titles (Kubitza et al., 2018b), being near forest might also imply freshly converted fertile land, which positively affects farmers' intensity production management. Higher altitudes can negatively affect oil palm productivity and indicate less accessible areas (Krishna et al., 2017a; Sibhatu, 2019).

Variables controlling for administrative units are generally introduced to control for effects of unobserved factors on larger spatial scales (e.g., village and district) or to capture administrative conditions or local policies influencing household decisions (Kubitza et al., 2018b; Krishna and Kubitza, 2021). In our case, we include controls for the village and regency levels.

Finally, we also consider the household coordinates (longitude, latitude) as an explanatory variable. Based on this information, we can estimate so-called *geosplines* or *effect surfaces* representing the effect of precise household location in two-dimensional space on the three outcome variables (see section 3.4 for details on the estimation). These effect surfaces also allow to capture and visualize spatial clusters not explained by the other variables presented in Table 1, and they control for potential social interaction effects because interpolation is based on averaging over neighboring households.

3.4. Empirical Strategy

To determine whether oil palm yield, input use, and output prices tend to cluster and delineate the location of the clusters and to understand how demographic, socioeconomic, and geomorphological covariates affect the distribution of oil palm yield, input use, and output prices in Jambi, we apply a Structured Additive Regression (STAR) framework. This approach allows the estimation of different effect types (e.g., linear, nonlinear, and random) in the same model predictor.

The STAR framework could be exceptionally helpful when analyzing spatial patterns since so-called *geosplines* could be added. Geosplines, one of the vital approaches to spatial analyses, uses the physical location of households (GPS coordinates of a household residence) within a given geographic area to determine nonlinear spatial clusters and map the location of spatial clusters. Geosplines are a straightforward tool to estimate and visualize spatial effects (Sharma et al., 2011; Steinhübel et al., 2020; Steinhübel and von Cramon-Taubadel, 2021). The general idea of geosplines is that the GPS coordinates of a household residence can be used as a bivariate explanatory variable. Thus, longitude and latitude are each treated as a continuous variable and define a household's position relative to the other households in the dataset. Utilizing the GPS coordinates also helps determine whether yield, input use, and output prices tend to cluster and delineate the location of the clusters. This approach can help agricultural and sustainability intervention programs target and determine the type and scale of planned interventions by determining and visualizing the physical location and the low and high levels of the indicators of farmers' management activities.

To estimate such nonlinear effects, we use penalized splines (P-splines) (Fahrmeir et al., 2013). For that purpose, the value space of the continuous variable is split into $k-1$ intervals, and for each interval, a polynomial of degree $l \geq 0$ is fitted. To produce one continuous effect function, i.e., the different polynomials must add up to one function, the function must be $(l-1)$ differentiable. A penalty term ensures that the function strikes the right balance between flexibility and smoothness. The smaller the intervals and the more polynomials are estimated, the more flexible the function will be and at some point, might be hard to interpret. Thus, smoothing based on the differences between estimated coefficients of neighboring observations ensures that the general patterns become visible. In statistics, splines have become a common tool and user-friendly packages in most statistical software programs facilitate the implementation (see our code in the supplementary material for an application in R). For a detailed introduction to the methodology see, for example (Fahrmeir et al., 2013; Umlauf et al., 2015). For spatial analysis, the interesting extension is the estimation of geosplines, i.e., two-dimensional P-splines based on longitude and latitude. The result is a smooth surface (instead of a smoothed line for a one-dimensional variable) that can be mapped and, thus, visualizes spatial patterns flexibly.

In the subsequent analysis, we employ two model specifications to investigate spatial differences in how smallholders in Jambi produce oil palm. To ease interpretation and because the variables are skewed and/or quite spread out, we log-transform all three of them. For the input use, we use a $log(y+1)$-transformation to adjust for zero observations.

In the first model specification, we only estimate separate geosplines for each of the three outcome variables $a = (yield, input\ use, output\ price)$:

$$y_{a,h} = exp(f(latitude_h, longitude_h) + \varepsilon_h) \tag{1}$$

where y refers to the three log-transformed outcome variables a, the function $f(latitude_h, longitude_h)$ is the geospline based on the GPS coordinates of household h, and ε a random error term. By only including $f(latitude_h, longitude_h)$ in Equation (1), the resulting estimated effect surface shows the full structured spatial variation (i.e., spatial variation excluding the white noise) of the three outcome variables.

Using the results of Equation (1) as the base, in the second model specification, we add the variables presented in types 1 to 5 in Table 1:

$$y_{a,h} = exp(X_h\beta + v_h + f(latitude_h, longitude_h) + \varepsilon_h) \quad (2)$$

Except for the village dummy, all variables are included as standard fixed effects $X_h\beta$. Village random effects v_h control for unobserved variables at the village level, i.e., the random effects allow for a village-specific deviation from the overall sample intercept.

Finally, by comparing estimation results of equations (1) and (2), we can deduce if and to which extend the control variables in Equation (2) explain the spatial variation in the outcome variables (Equation (1)). We also expect that the geosplines in Equation (2) would reveal any remaining spatial clusters after controlling for all control and geographic variables presented in Table 1 (types 3 to 6 in Table 1).

4. Results and Discussion

4.1. Descriptive statistics

Our survey data reveals that, on average, farmers own about 5.7 ha of land. Out of this, only 0.7 ha is used to grow other crops next to oil palm, implying that the sample households are highly specialized in oil palm production. Furthermore, approximately 70 percent of the oil palm plots that the sample household cultivates are assigned by systematic (formal) land titles, the most legally recognized title for farmlands in Indonesia. This is a relatively high proportion for Indonesian smallholder farms. Note, however, that we only focus on oil palm farmers, of whom many received formal land titles through the government's transmigration program (Gatto et al., 2017). For other crops, the share is much lower (Krishna et al., 2017b; Kubitza et al., 2018b).

Less than seven percent of the households report being aware of the RSPO certification program in our sample. Thus, we can assume that the penetration of certification standards among independent oil palm farm households in Jambi is still low. Plantations are on average 15 years old and produce 19 tons of fruit bunches per year. This average yield seems slightly high for smallholders, particularly compared with other areas in Indonesia and other oil palm-producing regions in Asia, Africa, and Latin America (Qaim et al., 2020). However, in Jambi, oil palm is still predominantly cultivated by transmigrant households, who received training and technical and financial support for plantation establishment from companies and/or government (Euler et al., 2016b; Euler et al., 2017). Moreover, the province's market and rural infrastructure (compared with Sulawesi, Kalimantan, or Papua) are well developed, and Jambi farmers also have extensive experience in handling commercial plantation crops such as rubber and industrial timber (Sibhatu, 2020). The average price received per kilogram of fruit bunch in 2015 is about 850 IDR, and an average farmer invests about 9.9 million IDR per ha a year on agricultural inputs. This relatively high spending on inputs might also be a reason for the somewhat higher yields in Jambi.

The last column in Table 2 presents the test statistics to analyze the statistical significance of differences in the farm, geomorphology, and household characteristics across regencies. The statistics suggest that almost all variables in our analysis show statistically significant variation across Jambi regencies. The only two variables not yielding a significant test statistic are household size and the gender of the household head.

As for the dependent variables, the highest average yield per ha is reported in Bungo with almost 22 kg/ha, while farmers in Betanghari report only 16.34 kg/ha on average. Interestingly, however, farmers in Betanghari receive the highest average output price of 1,340 IDR per kg. Conversely, the lowest output prices are reported for Muaro Jambi, where households, nonetheless, spend about two times more on inputs than the overall mean of our sample (~ 12,940,000 IDR per ha / 908 United States Dollar). Also, other production characteristics vary across regencies. For example, Batanghari and Tebo's farms are the largest, while the smallest ones are Sarolangung and Muaro Jambi. The awareness of the RSPO certification schemes appears to somewhat coincide with the differences in output prices. One-quarter of the respondents in Batanghari reported being aware of RSPO certification, but none is reported from Sarolangun.

Furthermore, the oil palm farmers in our sample are, on average, 50 years old and received formal schooling for about 7.5 years. Nearly one-fifth of the households have access to rivers, but shares vary significantly among regencies. In Bungo, for example, almost 60 percent of farmers reported river access, whereas none had direct access to a river in Sarolangung. The same holds for the location close to forest edges, where zero percent of households are located in the proximity of natural forest in Batanghari, Bungo, and Muaro compared with 44 percent in Tebo. Household dwellings are located around 50 meters above sea level on average, with a variation of about 45 meters across all regencies.

4.2. Structured spatial heterogeneity of oil palm yield, inputs, and output prices

In order to find out whether there is a distinct spatial variation, we map the effects of household location (geosplines) separately on the three outcome variables (yield, input use, and output prices) without any covariates, as defined in Equation (1). The results are depicted in Figure 2. As one might expect, all three variables show spatially distinct clusters of above-mean (red areas) and below-mean (blue area) observations. Areas colored in red indicate hotspots of higher yields per hectare (Figure 2a), higher input use (Figure 2b), and higher output prices (Figure 2c), while those in blue are clusters of lower yields, input use, and output prices. The scales in Figure 2 represent coefficient estimates. Because we log-transformed the outcome variables, we have to transform the coefficients to their linear equivalent to derive effects in percentage changes, $((exp(\beta) - 1) \times 100)$. For example, an absolute coefficient magnitude of 0.5 implies a 64.87 percent increase above-average yields, input use, or output prices, *ceteris paribus*. Considering the scales of Figure 2, the spatial variation of oil palm yields and input costs appear particularly large but also output prices show spatial clusters up to 28 percent above and below the mean of 1,290 IDR per kilogram of oil palm fruit bunch (Figure 2c).

Table 2. Descriptive statistics disaggregated by regency

Variable	All Households		Batanghari		Bungo		Muaro Jambi		Sarolangung		Tebo		Test statistic[a, b]
	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev	
Harvest /ha ('000 kg)	19.46	7.71	16.34	9.80	21.71	6.46	20.25	8.63	19.11	7.71	19.54	10.07	13.22***
Input cost / ha ('000 IDR)	4344.17	78765.22	1185.66	2928.75	2769.44	9506.04	12940.85	157382.20	422.35	2382.27	1068.06	4614.42	3.89***
Output price /kg ('000 IDR)	1.29	0.25	1.34	0.27	1.31	0.25	1.23	0.20	1.28	0.27	1.27	0.26	57.93***
Control variables													
Age of hh head (years)	49.68	10.26	46.99	9.29	51.94	10.17	50.57	10.90	48.94	9.80	49.62	10.19	5.56***
Education hh head (years)	7.49	3.62	7.85	3.46	7.17	3.69	7.19	3.61	6.94	2.82	7.92	3.90	2.16*
Household size (count)	3.97	1.50	4.06	1.56	3.77	1.51	4.06	1.47	4.17	1.49	3.89	1.47	1.14
Village type (auto. / transmig.)	0.24		0.38		0.00		0.23		0.00		0.41		35.72***
Gender hh head (Male/female)	0.98		0.98		0.99		0.96		0.98		0.99		1.80
Total land owned (ha)	5.72	6.88	6.91	9.24	4.93	4.03	4.81	7.67	4.70	4.12	6.56	5.81	3.84***
Oil palm production area (ha)	4.98	6.44	5.86	8.58	4.37	3.67	4.63	7.69	4.34	4.06	5.24	5.00	6.54***
Age of plantation (years)	15.03	6.39	15.04	5.99	16.74	5.36	15.74	7.17	16.46	5.72	12.36	6.10	13.52***
Share of plots with systematic land title	0.69	0.42	0.64	0.43	0.86	0.30	0.63	0.44	0.80	0.36	0.61	0.46	11.44***
Aware of RSPO certification (dummy)	0.07		0.26		0.01		0.01		0.00		0.03		40.05***
Land areas allocated to non-oil palm crops (ha)	0.74	1.88	1.05	2.36	0.55	1.32	0.19	0.59	0.36	0.61	1.31	2.59	11.62***
Distance to nearest city (Km)	27.33	8.19	27.35	8.22	22.00	4.39	31.67	6.62	32.08	0.94	25.64	10.01	47.47***
Distance to Jambi city (Km)	95.20	44.22	62.42	9.08	142.80	9.43	40.81	7.86	124.39	1.53	131.62	17.58	2853.32***
Distance to nearest mill (Km)	16.30	9.31	14.51	2.79	9.82	5.09	13.65	7.06	12.09	5.64	26.36	10.71	146.96***
Access to a river (dummy)	0.22		0.24		0.56		0.11		0.00		0.11		47.96***
Household located at the edge of forest (dummy)	0.13		0.00		0.00		0.00		0.33		0.44		94.46***
Altitude	51.08	24.75	44.62	17.49	75.32	18.27	29.98	20.71	55.77	11.31	56.95	21.47	135.24***
No. Observations	794		177		160		198		66		193		

Notes: **a** Test for equality of 5 group means, assuming heterogeneity (Wald Chi-squared statistics are reported). **b** Non-normally distributed continuous variables are log-transformed for the test. Asterisks represent p-values, *** <0.01, ** <0.05 and *<0.1. auto. / transmig – autochthonous / transmigrant. hh – household head.

In addition, Figure 2 suggests that the spatial clusters identified by the heatmap do not necessarily match regency boundaries, and we have to assume that additional factors influence the spatial heterogeneity in our dependent variables. Note, for example, the patterns around the Bukit 12 National Park in the south of our study area. Especially towards the east of this conservation area for tropical rainforest in South Batanghari, Figures 2a and b reveal much lower yield levels and input use. In contrast, in northern Batanghari, we observe some of the highest yields in our sample. Thus, based on our estimated geosplines, we can identify clear in-regency spatial patterns that do not become evident in the descriptive analysis presented in Table. This means that we have to assume that—if a set of different factors drives spatial heterogeneity in yields, input costs, and output prices (e.g., forests, administrative boundaries)—one spatial scale is not enough to account for the total spatial variation in the dependent variables. Instead, only a spatially explicit approach, such as presented in Figure 2 based on the smallest scale of observation (i.e., households), will allow for a complete picture.

If we assume that the three dependent variables represent the oil palm production systems in Jambi, the comparison of the three maps in Figure 2 also reveals some overall spatial patterns. For example, we observe above-average levels for all three dependent variables in the west of our study area. That means smallholders spend relatively much on inputs and produce high yields, for which they receive above-average prices. In the center of our research area, we find a large cluster of above-average output prices (Figure 2c), but a sharp north-south divide in above and below-average yields and input spend. In contrast, farmers in the east spend average amounts on inputs, and also harvest quantities are close to the sample mean. Nonetheless, they receive below-average prices for their produce. This is particularly interesting because the province capital, Jambi, is located in this area. Normally, the literature suggests that transaction costs decrease with proximity to urban centers and markets, and farmers receive higher net output prices (Damania et al., 2017; Vandercasteelen et al., 2018). However, recent studies also highlight those smaller towns might be more relevant for agricultural production than large urban centers (Steinhübel and von Cramon-Taubadel, 2021).

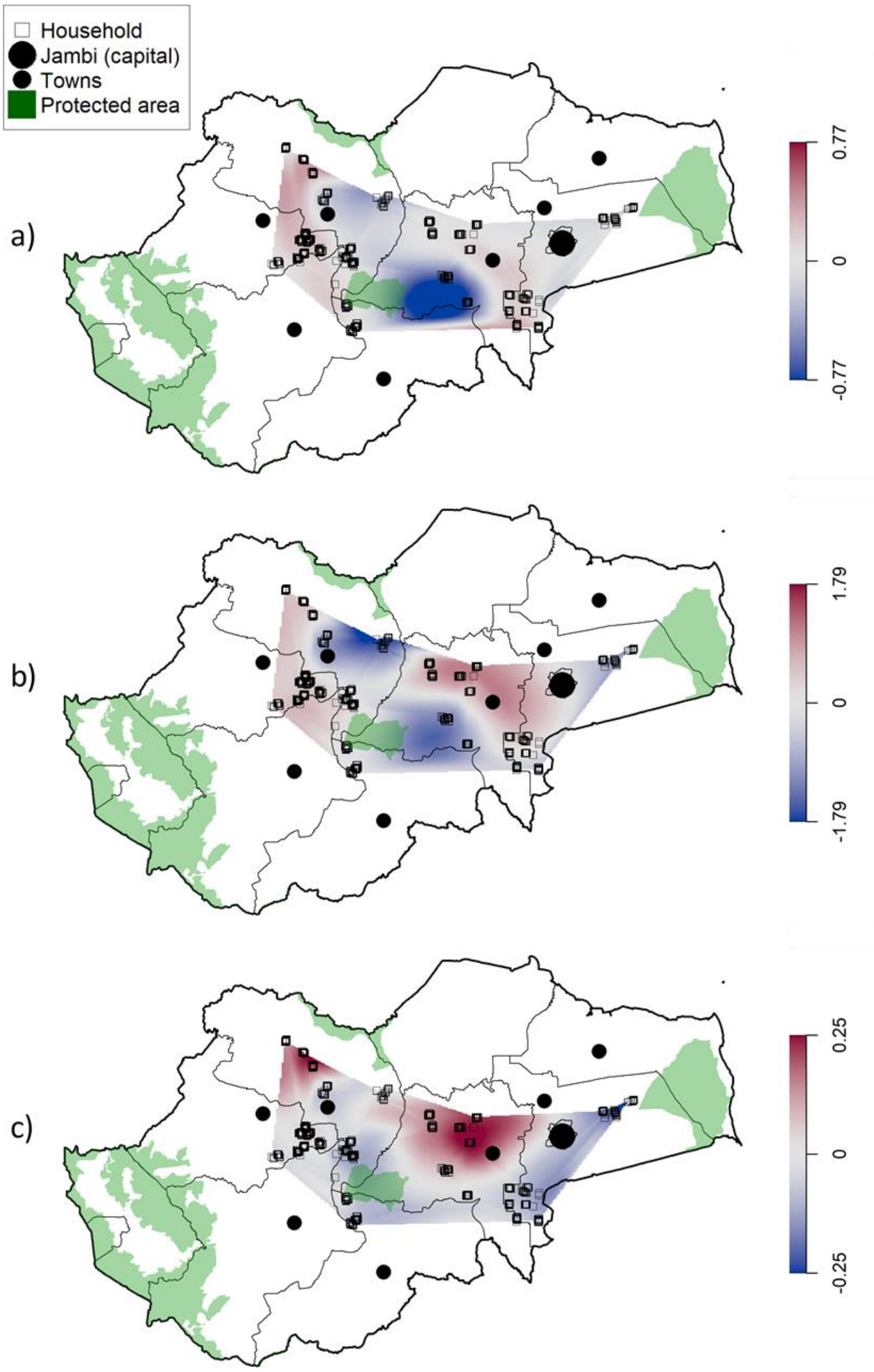

Figure 2: Estimated geosplines based on Equation (1), a) oil palm yields, b) input use, c) output price.
Notes: Areas colored in red indicate above-mean yields per hectare (a), higher input use (b), and higher output prices (c), while those in darker blue, clusters of lower yield, input use, and output prices. The scales represent coefficient estimates. Outcome variables are log-transformed. Regency names (depicted in Figure 1) are left out for clarity.

4.3. Drivers of spatial heterogeneity of oil palm yield, inputs, and prices

Our analysis in section 4.2 shows that oil palm production in Jambi varies in space. Now, the question is whether farm, environmental, and socioeconomic factors can explain the observed spatial clusters in oil palm yield, input use, and output prices. Hence, in Table 3, we present estimation results for the model in Equation (2). The estimated effect surfaces (geosplines) as specified in Equation (2) are shown in Figure 3.

Column (1) in Table 3 shows the factors that are associated with oil palm yield. The intensity of input use, age of plantations, and systematic land title all have positive and statistically significant coefficients. With every additional 1,000 IDR a farmer spends on inputs per hectare, yields per hectare increase by 4.3 percent ($(exp(0.042) - 1) \times 100$). Similarly, with every additional year, a plantation produces 3.2 percent ($(exp(0.042) - 1) \times 100$) more output. However, this effect is not linear. In Figure A.1, we present a nonlinear estimate of the effect of plantation age (keeping the rest of the model specification the same), and it suggests that the positive age effect only holds until ten years of age. After 25 years there is even a slight decline. Smallholders with systematic land titles reported 10.6 percent ($(exp(0.101) - 1) \times 100$) higher oil palm yields *ceteris paribus*. Particularly in perennial (i.e., long-term) crop systems, a land title provides security to smallholders that investments and intensification pay off in the future. Thus, the positive effect is plausible. Also, a land title allows farmers to use their land as collateral to access rural financial markets or to diversify their off-farm livelihood systems (Krishna et al., 2017b). Only slightly above conventional significance levels ($p - value = 0.139$), also a location at the edge of forests seems to be associated with high oil palm yields. Note that this dummy refers to any kind of forest and is not exclusively tied to conservation areas. Farms at forest margins could have higher availability of organic matter and other nutrients in their soil system from the forest covers in the recent past, thus supporting higher yields at least in the first planting. In addition, plantations close to natural forests might be younger because they result from more recent land conversion. Kubitza et al. (2018b), for example, show that, in Jambi, proximity to a forest is positively associated with bigger farm size, possibly acquired by deforestation but often without official land titles. Therefore, short-term rewards from higher yields from deforestation might be a smallholder strategy to compensate for missing formal land titles. Finally, a location at higher altitudes significantly decreases oil palm yields. This is consistent with previous studies that show that environmental conditions for oil palm production are better in lower altitudes and, thus, also discourage oil palm adoption in higher altitudes (Corley and Tinker, 2016; Krishna et al., 2017a; Sibhatu, 2019).

In column (2) in table 3, we report factors associated with input use. Again, we find significant positive effects of formal land titles, the number of plots, and the distance to Jambi. As mentioned earlier, stronger land property rights combined with bigger farm sizes could encourage farmers to intensify input application and increase productivity. This matches the findings by (Kubitza et al., 2018b), which show that farmers who own formal land titles deforest less but increase their input intensity in Jambi. The cultivation of other crops (i.e., a higher agricultural production diversity) is associated with lower use of inputs of 51 percent ($(exp(-0.717) - 1) \times 100$). This is plausible since we can assume these farms to be less specialized in oil palm production. Interestingly, the distance to the capital Jambi decreases input use, whereas input use seems to increase with proximity to other (nearest) towns, though the p-value is slightly above conventional levels ($p - value = 0.166$). This matches results of recent studies that suggest that smaller towns might be more relevant for agricultural intensification than large urban centers (Steinhübel and von Cramon-Taubadel, 2021). The latter often offer more non-farm employment opportunities and, thus, increases opportunity costs for agricultural intensification (e.g., farm labor becomes more expensive).

When it comes to output prices, it appears that established, specialized, and large farms have the best chance of receiving above-average prices for their oil palm produce. That is, farm size, yields, plantation age, and formal land titles have a statistically significant association with higher output prices, whereas the management of additional crops is related to lower prices. Literature has shown that large farms tend to be more specialized and have a higher potential for sustainably commercialized production (Meemken, 2021). Hence, it is plausible that these smallholders might be better informed about prices and choose market places. Furthermore, we find that also the location at a forest edge is associated with 19 percent ($(exp(0.101) - 1) \times 100$) higher realized output prices. Again, this might point towards some short-term gains from deforestation without formal land titles. In addition, similar to input spending, our results suggest that remote smallholder farms receive lower prices for their oil palm yields. With every additional kilometer away from an urban center, prices decrease by 0.4 percent ($(exp(-0.004) - 1) \times 100$). These results match findings in the literature that show that market access is an important factor in agricultural prices (Levi et al., 2020). Note that this effect is only statistically significant for the closest town but not for the distance to the province capital Jambi.

Table 3. Estimation results, Equation (2)

	(1) Yield /ha (log, kg)	(2) Input use / ha (log, '000 IDR)	(3) Output price /kg (log, '000 IDR)
Socio-economic variables			
Age of hh head (years)	0.001 (0.002)	-0.001 (0.008)	0.0003 (0.001)
Education hh head (years)	0.005 (0.005)	0.033 (0.022)	0.002 (0.001)
Household size (count)	0.010 (0.012)	-0.067 (0.049)	0.001 (0.003)
Village type (dummy, trad. /transm.)	-0.036 (0.109)	-0.219 (0.341)	-0.017 (0.024)
Total land owned (log, ha)	0.008 (0.036)	0.036 (0.146)	0.038*** (0.009)
Input cost / ha (log, '000 IDR)	0.042*** (0.009)		0.005** (0.002)
Harvest /ha (log, kg)	NA		0.043*** (0.009)
Number of plots (count)	0.030 (0.022)	0.162* (0.087)	0.008 (0.005)
Age of plantation (years)	0.033*** (0.004)	0.018 (0.015)	0.001 (0.001)
Share systematic land title	0.109** (0.054)	0.524** (0.217)	0.034** (0.013)
Cultivate other crops (yes=1; no=0)	-0.014 (0.049)	-0.704*** (0.197)	-0.033*** (0.012)
Aware of RSPO certification (dummy)	-0.043 (0.096)	0.484 (0.387)	0.016 (0.024)
Spatial variables			
Distance to nearest city (Km)	-0.001 (0.006)	-0.031 (0.023)	-0.004** (0.002)
Distance to Jambi city (Km)	0.001 (0.003)	0.024* (0.014)	0.001 (0.001)
Distance to nearest mill (Km)	-0.005 (0.005)	0.006 (0.023)	-0.001 (0.002)
Altitude (meter above sea level)	-0.003*** (0.001)	-0.002 (0.005)	-0.0003 (0.0003)
Access to a river (dummy)	0.025 (0.107)	0.264 (0.286)	-0.010 (0.017)
Hh located at edge of forest (dummy)	0.201 (0.136)	-0.034 (0.492)	0.173*** (0.040)
Regency dummy (ref. Batanghari)	✓	✓	✓
Bungo	0.260 (0.264)	-1.530 (1.136)	-0.160* (0.086)
Muaro Jambi	0.252* (0.143)	0.954 (0.778)	-0.195*** (0.072)
Sarolangun	0.064 (0.246)	-1.228 (1.315)	-0.296** (0.129)
Tebo	0.197 (0.217)	-1.744* (0.940)	-0.254*** (0.080)
Village random effect	✓	✓	✓
Intercept	8.678*** (0.316)	4.860*** (1.346)	-0.507*** (0.150)
No. Obser.	793	793	793
AIC	-310.237	1908.9	-2526.2

Notes: Coefficients with standard errors in parentheses are reported. Asterisks represent p-values, *** <0.01, ** <0.05 and *<0.1. NA – not applicable; all estimations with geosplines. Hh/hh – household head.

When comparing all three columns in Table 3, some overall patterns emerge for the socioeconomic variables. Interestingly, official land titles, farm size, and specialization in oil palm production seem to be the important factors determining high-yield and economically successful oil palm farms[3]. This is particularly relevant because we investigate a perennial agricultural management system. This type of plantation system requires massive initial investments while yields only increase with time. Thus, secure land property rights are fundamental in terms of collateral when borrowing money from a formal institution, which is often a challenge for smallholder farmers (Fenske, 2011; Lawry et al., 2017). This is also interesting when considering that a location at a forest margin also improves yields and output prices. Since previous research suggests that this could be related to harvesting short-term gains from deforestation without official land titles (McCarthy, 2010; Krishna et al., 2017c; Kubitza et al., 2018b), easier access to strong land property rights might be a good policy strategy to improve farm investments and productivity while also preventing deforestation. Furthermore, our results highlight the importance of urban centers, particularly when it comes to input use and oil palm prices. However, in the Jambi context, regional towns appear to be more relevant for agricultural development than the province capital Jambi. Our results, thus, support claims in previous studies that differentiation of different types of towns is necessary, and smaller towns are often more important for the agricultural sector than large urban centers (Steinhübel and von Cramon-Taubadel, 2021).

Finally, the remaining question is whether the coefficients we discussed so far (Table 3) capture all the spatial variation presented in Figure 2. By comparing Figure 2 and in Figure 3, we observe a reduction in the magnitude of estimated coefficients (scale on the right of each map). While it appears that almost all of the spatial variation in yields (scale extremes: 0.77 → 0.01; Figure 2a and Figure 3a) is controlled for, we observe a reduction of only 40 to 60 percent for input use and output prices (scale extremes: 1.79 → 0.75 and 0.25 → 0.15, respectively; Figure 2b/c and Figure 3b/c). In Figure 3b., red/blue areas still show areas where input spending is up to 144 percent above or below the sample mean. For output prices, the variation ranges from 16 percent above and below the sample mean. This means that—at least for input use and output prices—the spatial clusters in Figure 2 are not entirely explained by our control variables, and some structured spatial heterogeneity remains unexplained.

---

[3] For input use and output prices we also observe positive effects of education. However, signicance levels are slightly above conventional thresholds ($p - value = 0.128$ and $p - value = 0.125$, respectively).

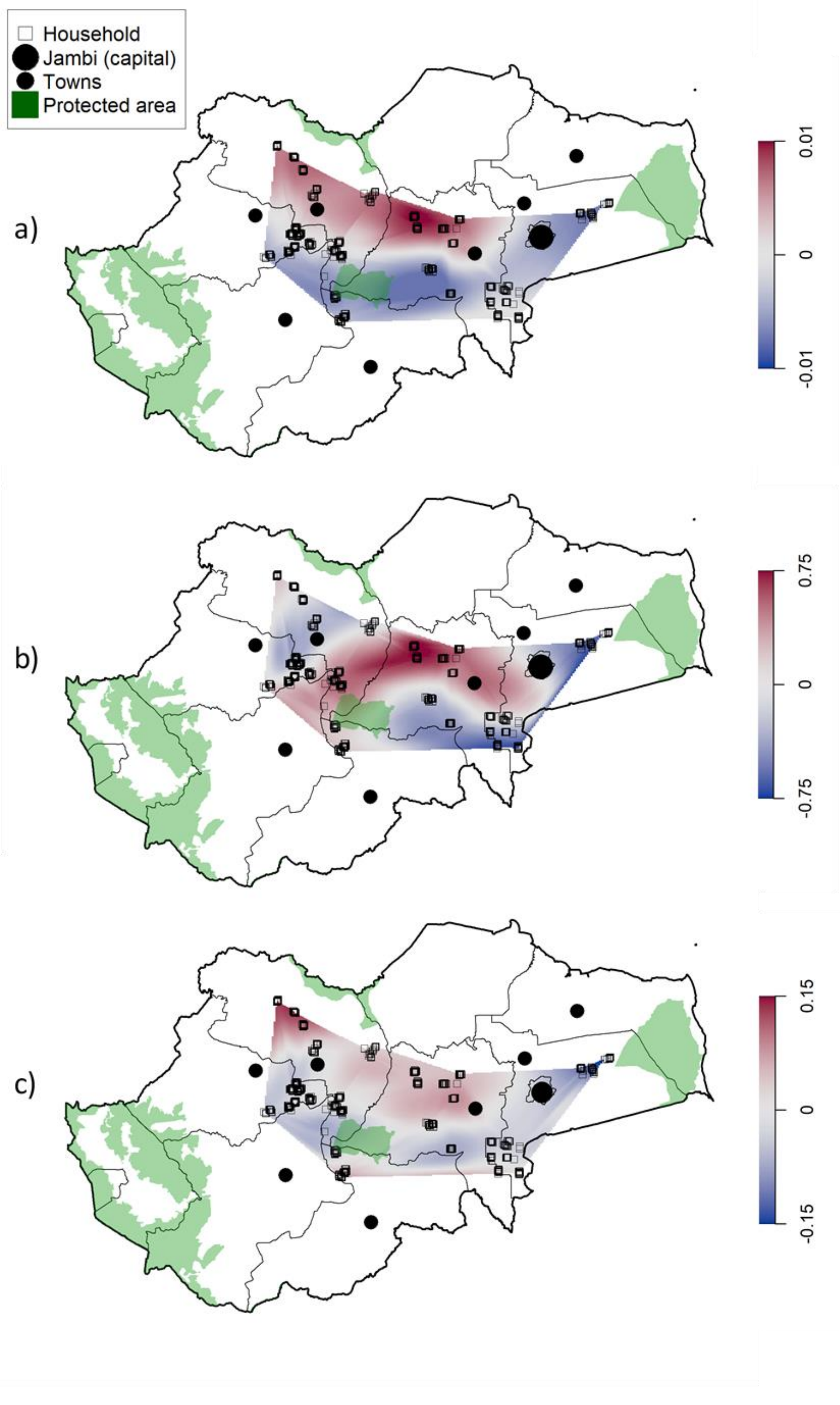

Figure 3. Estimated geosplines based on Equation (2), a) oil palm yields, b) input use, c) output price.
Notes: Areas colored in red indicate above-mean yields per hectare (a), higher input use (b), and higher output prices (c), while those in darker blue, clusters of lower yield, input use, and output prices. The scales represent coefficient estimates. Outcome variables are log-transformed. Regency names (depicted in Figure 1) are left out for clarity.

Moreover, the key to explaining the much higher reduction of spatial variation in oil palm yields is the village random effects. If we estimate Equation (2) without the village controls, the fixed effect estimates are robust

(Table A.1). However, while Figures 3b-c and Figures A.2b-c show mainly the same patterns, the estimated geosplines for yields reveal a significant difference (Figure 3a and Figure A.2a). Without the control on the village scale, the reduction of spatial variation is similar in all three dependent variables, i.e., also for yields only about 60 percent (scale extremes: 0.77 → 0.26; Figure 2a and Figure 3a) of the spatial variation is explained by the variables in Table 3/Table A.1. This allows for two conclusions. First, small-scale (i.e., village) factors are crucial in explaining variation in oil palm productivity, but some of these factors seem not to be captured by standardized socioeconomic or geographic variables (Table 2). This could be, for example, village-specific networks or extension services and local environmental conditions that affect oil palm yields. Second, from a technical/modeling point of view, controls for unobserved variability on small spatial scales seem to be sufficient to capture most of the remaining structured spatial heterogeneity in oil palm yields.

The picture is different for input use and output prices. Especially, the spatial variation in farmers' decisions concerning how much they invest in their oil palm production systems is still substantial. Thus, neither the village nor the regency scale seems appropriate to capture high or low input spending clusters. Instead, based on Figure 3b, it appears that a larger region, arching through the research area from the southern areas in the east and west to the north in the center, shows relatively high levels of input use.

However, are there any other patterns not captured by the control variable that could explain such a cluster? There is always the possibility that large-scale environmental patterns (e.g., groundwater, soils) or infrastructure (e.g., roads, rivers), which are difficult to capture in standardized socioeconomic surveys, lead to remaining spatial variation. This also holds for institutions or services that are only available in spatially bound areas. We, for example, suspect that oil palm certification schemes might play a role in explaining the remaining clusters in input use. Unfortunately, we only have information on whether farmers are aware of the RSPO scheme but not if they participate. The dummy in Table 3 (column 2) is not statistically significant at conventional levels ($p - value = 0.221$) but shows a large positive effect size and might be a signal that certifications schemes contribute to an above-average input spending behavior in this area. Furthermore, a significantly larger share of farmers in Batanghari (regency in the center of the research area) is aware of the RSPO scheme (Table 2), and one of the few RSPO-certified processing mills is located in the north of Batanghari. Based on a village questionnaire, we also know that three villages of our sample villages actively participate in the RSPO scheme. About 82 percent (44 out of 54 observations) of the farmers aware of the

RSPO certification scheme in our sample are from these three villages. Nonetheless, this is only anecdotal evidence, and additional analysis is necessary to examine whether production systems differ in the area due to certification schemes.

In any way, our results show that in two out of three dependent variables, standard explanatory variables and controls on different administrative scales are not enough to capture all of the structured spatial heterogeneity in oil palm systems in Jambi. Assuming that this is the case in other study areas, we have to think about what this means in general to analyze agricultural systems based on socioeconomic survey data.

First, we have to expect biased estimates from a methodological perspective if we do not include spatially explicit controls such as geosplines in our analysis or at least test for remaining structured spatial patterns. In our study area, such a bias does not seem to be a big issue because the results of fixed effects presented in Table 3 match the finding of previous studies in Jambi (Mehraban et al., 2021). Nonetheless, it is difficult to assume the robustness of coefficient estimates *a priori*.

Second, remaining unexplained spatial variation should also be considered in efforts to support agricultural development and environmental conservation. While focusing on the significant effects of proxies such as formal land titles in our case is important to inform policymakers, visualizing remaining spatial clusters of low yields, input use, and output prices (Figure 3) allows us to target disadvantaged regions specifically. It might also be necessary to engage qualitative approaches or work with grass-root organizations to understand the reasons for the lower levels of agricultural developments.

5. Conclusion

In this study, we have investigated the spatial heterogeneity in perennial agricultural systems—a research question, so far, only attempted for annual cropping systems. To do so, we have applied a Structured Additive Regression framework in order to detect a distinct spatial variation and explore the association between socioeconomic/geographic covariates and oil palm yield, input use, and output price. We have used a primary data set of 793 smallholder households in Jambi province, Indonesia. Moreover, the estimation of so-called *geosplines* has allowed us to control for unexplained spatial patterns in smallholder oil palm systems and to visualize remaining spatial heterogeneity after the inclusion of all standard covariates.

Our findings can be summarized as follows. First, farm characteristics that indicate long-term stability and a specialization in oil palm production are associated with significantly higher oil palm yields, input use, and realized output prices. In this context, we have found the strongest effects (in statistical significance and magnitude) for the possession of formal land titles and plantation age or the size of the farm play a role. Furthermore, our results suggest that, in the absence of a formal land title, smallholders might be more likely to cut natural forests to gain short-term benefits from higher oil palm yields due to improved soil conditions.

Second, we have discovered that proximity to market centers is associated with significantly higher input use and realized output prices in oil palm (perennial) management systems, consistent with the findings from the studies conducting spatial analyses of annual cropping systems. Note that, in this context, smaller towns seem to be more important than proximity to the province capital Jambi. This means that for cropping systems with a longer planning horizon, market access seems to be crucial for smallholders to intensify/commercialize their production.

Third, the estimated geosplines in our empirical analysis have shown that standard covariates explain only 40-60 percent of the original spatial heterogeneity in oil palm yields, input use, and output prices. This includes standard spatial proxies such as distance to the closest town or mill as well as geographic variables such as altitude or the location at the edge of a forest. Further, we have shown that control for unexplained variation at the village scale (i.e., village random effect) can correct most of the remaining spatial variation for oil palm yield. However, significant spatial patterns in input use and output prices remain, even after controlling for different administrative scales (village, regency) and standard explanatory variables. We conclude that, even if we could not explain the origins of these remaining spatial variations, we are able to detect them using the geosplines, preventing estimation bias. Thus, visualization of spatial clusters might be helpful for more targeted policy initiatives to improve oil palm management systems.

Finally, we suggest that future quantitative studies might combine qualitative data to fully understand the remaining spatial variation and validate and strengthen our findings both in the context of annual or perennial crops and in the context of Indonesia or other tropical regions. We also suggest that intervention programs collaborate with grass-root organizations to identify disadvantaged farmers and account for the observed spatial variation. Also, all results reported in this study are based on a cross-sectional data set. Hence, everything we present may not refer to causal relationships, which should be considered when interpreting our findings. Especially when considering the longer planning horizon of perennial cropping systems, a panel analysis including explicit spatial predictors such as geosplines could be useful to obtain a more nuanced understanding of the drivers of spatial clusters.

References

Asfaw, S., Di Battista, F., Lipper, L., 2016. Agricultural technology adoption under climate change in the Sahel: Micro-evidence from Niger. J Afr Econ 25 (5), 637–669. 10.1093/jae/ejw005.

Badan Pusat Statistik, 2017. Statistik Kelapa Sawit Indonesia 2017. Badan Pusat Statistik. https://www.bps.go.id/publication/2018/11/13/b73ff9a5dc9f8d694d74635f/statistik-kelapa-sawitindonesia-2017.html. Accessed 2 June 2020.

Badan Pusat Statistik (BPS), 2018. Provinsi Jambi Dalam Angka. BPS Provinsi Jambi, Jambi, Indonesia.

Badan Pusat Statistik (BPS), 2019. Indonesian Oil Palm Statistics 2018. Badan Pusat Statistik, Jakarta.

Beckert, B., Dittrich, C., Adiwibowo, S., 2014. Contested land: An analysis of multi-layered conflicts in Jambi Province, Sumatra, Indonesia. Austrian Journal of South-East Asian Studies 7 (1), 75–92. 10.14764/10.ASEAS-2014.1-6.

Cahyadi, E.R., Waibel, H., 2016. Contract farming and vulnerability to poverty among oil palm smallholders in Indonesia. Journal of Development Studies 52 (5), 681–695. 10.1080/00220388.2015.1098627.

Cisneros, E., Kis-Katos, K., Nuryartono, N., 2021. Palm oil and the politics of deforestation in Indonesia. Journal of Environmental Economics and Management 108, 102453. 10.1016/j.jeem.2021.102453.

Clough, Y., Krishna, V.V., Corre, M.D., Darras, K., Denmead, L.H., Meijide, A., Moser, S., Musshoff, O., Steinebach, S., Veldkamp, E., Allen, K., Barnes, A.D., Breidenbach, N., Brose, U., Buchori, D., Daniel, R., Finkeldey, R., Harahap, I., Hertel, D., Holtkamp, A.M., Hörandl, E., Irawan, B., Jaya, I.N.S., Jochum, M., Klarner, B., Knohl, A., Kotowska, M.M., Krashevska, V., Kreft, H., Kurniawan, S., Leuschner, C., Maraun, M., Melati, D.N., Opfermann, N., Pérez-Cruzado, C., Prabowo, W.E., Rembold, K., Rizali, A., Rubiana, R., Schneider, D., Tjitrosoedirdjo, S.S., Tjoa, A., Tscharntke, T., Scheu, S., 2016. Land-use choices follow profitability at the expense of ecological functions in Indonesian smallholder landscapes. Nature Communications 7, 13137. 10.1038/ncomms13137.

Corley, R., Tinker, P.B., 2016. The Oil Palm, 5th ed. Wiley Blackwell, Chichester, West Sussex.

Damania, R., Berg, C., Russ, J., Federico Barra, A., Nash, J., Ali, R., 2017. Agricultural technology choice and transport. American Journal of Agricultural Economics 99 (1), 265–284. 10.1093/ajae/aav073.

De La O Campos, A.P., Villani, C., Davis, B., Takagi, M., 2018. Ending extreme poverty in rural areas – Sustaining livelihoods to leave no one behind. FAO, Rome.

Dharmawan, A.H., Mardiyaningsih, D.I., Komarudin, H., Ghazoul, J., Pacheco, P., Rahmadian, F., 2020. Dynamics of rural economy: A socioeconomic understanding of oil palm expansion and landscape changes in East Kalimantan, Indonesia. Land 9 (7), 213. 10.3390/land9070213.

Drescher, J., Rembold, K., Allen, K., Beckschäfer, P., Buchori, D., Clough, Y., Faust, H., Fauzi, A.M., Gunawan, D., Hertel, D., Irawan, B., Jaya, I.N.S., Klarner, B., Kleinn, C., Knohl, A., Kotowska, M.M., Krashevska, V., Krishna, V., Leuschner, C., Lorenz, W., Meijide, A., Melati, D., Nomura, M., Pérez-Cruzado, C., Qaim, M., Siregar, I.Z., Steinebach, S., Tjoa, A., Tscharntke, T., Wick, B., Wiegand, K., Kreft, H., Scheu, S., 2016. Ecological and socio-economic functions across tropical land use systems after rainforest conversion. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 371 (1694). 10.1098/rstb.2015.0275.

Ebata, A., Velasco Pacheco, P.A., Cramon-Taubadel, S. von, 2017. The influence of proximity to market on bean producer prices in Nicaragua. Agricultural Economics 48 (4), 459–467. 10.1111/agec.12347.

Edwards, R.B., 2019. Export agriculture and rural poverty: Evidence from Indonesian palm oil, Brisbanem, QLD.

Euler, M., Hoffmann, M.P., Fathoni, Z., Schwarze, S., 2016a. Exploring yield gaps in smallholder oil palm production systems in Eastern Sumatra, Indonesia. Agricultural Systems 146, 111–119. 10.1016/j.agsy.2016.04.007.

Euler, M., Krishna, V., Schwarze, S., Siregar, H., Qaim, M., 2017. Oil palm adoption, household welfare, and nutrition among smallholder farmers in Indonesia. World Development 93, 219–235. 10.1016/j.worlddev.2016.12.019.

Euler, M., Schwarze, S., Siregar, H., Qaim, M., 2016b. Oil palm expansion among smallholder farmers in Sumatra, Indonesia. J Agric Econ 67 (3), 658–676. 10.1111/1477-9552.12163.

Fahrmeir, L., Kneib, T., Lang, S., Marx, B., 2013. Regression: Models, Methods and Applications. Springer, Dordrecht.

Fenske, J., 2011. Land tenure and investment incentives: Evidence from West Africa. Journal of Development Economics 95 (2), 137–156. 10.1016/j.jdeveco.2010.05.001.

Gatto, M., Wollni, M., Asnawi, R., Qaim, M., 2017. Oil palm boom, contract farming, and rural economic development: Village-level evidence from Indonesia. World Development 95, 127–140. 10.1016/j.worlddev.2017.02.013.

Grass, I., Kubitza, C., Krishna, V.V., Corre, M.D., Mußhoff, O., Pütz, P., Drescher, J., Rembold, K., Ariyanti, E.S., Barnes, A.D., Brinkmann, N., Brose, U., Brümmer, B., Buchori, D., Daniel, R., Darras, K.F.A., Faust, H., Fehrmann, L., Hein, J., Hennings, N., Hidayat, P., Hölscher, D., Jochum, M., Knohl, A., Kotowska, M.M., Krashevska, V., Kreft, H., Leuschner, C., Lobite, N.J.S., Panjaitan, R., Polle, A., Potapov, A.M., Purnama, E., Qaim, M., Röll, A., Scheu, S., Schneider, D., Tjoa, A., Tscharntke, T., Veldkamp, E., Wollni, M., 2020. Trade-offs between multifunctionality and profit in tropical smallholder landscapes. Nature Communications 11 (1), 107. 10.1038/s41467-020-15013-5.

Holloway, G., Lapar, M.L.A., 2007. How big is your neighbourhood? Spatial implications of market participation among Filipino smallholders. Journal of Agricultural Economics 58 (1), 37–60. 10.1111/j.1477-9552.2007.00077.x.

Knowler, D., Bradshaw, B., 2007. Farmers’ adoption of conservation agriculture: A review and synthesis of recent research. Food Policy 32 (1), 25–48. 10.1016/j.foodpol.2006.01.003.

Krishna, V.V., Euler, M., Siregar, H., Qaim, M., 2017a. Differential livelihood impacts of oil palm expansion in Indonesia. Agricultural Economics 48 (5), 639–653. 10.1111/agec.12363.

Krishna, V.V., Kubitza, C., 2021. Impact of oil palm expansion on the provision of private and community goods in rural Indonesia. Ecological Economics 179, 106829. 10.1016/j.ecolecon.2020.106829.

Krishna, V.V., Kubitza, C., Pascual, U., Qaim, M., 2017b. Land markets, property rights, and deforestation: Insights from Indonesia. World Development 99 (2017), 335–349.

Krishna, V.V., Kubitza, C., Pascual, U., Qaim, M., 2017c. Land markets, property rights, and deforestation: Insights from Indonesia. World Development 99, 335–349. 10.1016/j.worlddev.2017.05.018.

Kubitza, C., Krishna, V.V., Alamsyah, Z., Qaim, M., 2018a. The economics behind an ecological crisis: Livelihood effects of oil palm expansion in Sumatra, Indonesia. Human Ecology 46 (1), 107–116. 10.1007/s10745-017-9965-7.

Kubitza, C., Krishna, V.V., Urban, K., Alamsyah, Z., Qaim, M., 2018b. Land property rights, agricultural intensification, and deforestation in Indonesia. Ecological Economics 147, 312–321. 10.1016/j.ecolecon.2018.01.021.

Kunz, Y., Otten, F., Mardiana, R., Martens, K., Roedel, I., Faust, H., 2019. Smallholder telecoupling and climate governance in Jambi Province, Indonesia. Social Sciences 8 (4), 115. 10.3390/socsci8040115.

Lawry, S., Samii, C., Hall, R., Leopold, A., Hornby, D., Mtero, F., 2017. The impact of land property rights interventions on investment and agricultural productivity in developing countries: A systematic review. Journal of Development Effectiveness 9 (1), 61–81. 10.1080/19439342.2016.1160947.

Levi, R., Rajan, M., Singhvi, S., Zheng, Y., 2020. The impact of unifying agricultural wholesale markets on prices and farmers' profitability. Proceedings of the National Academy of Sciences of the United States of America 117 (5), 2366–2371. 10.1073/pnas.1906854117.

Marenya, P.P., Barrett, C.B., 2007. Household-level determinants of adoption of improved natural resources management practices among smallholder farmers in western Kenya. Food Policy 32 (4), 515–536. 10.1016/j.foodpol.2006.10.002.

McCarthy, J., 2010. Processes of inclusion and adverse incorporation: oil palm and agrarian change in Sumatra, Indonesia. Journal of Peasant Studies 37 (4), 821–850. 10.1080/03066150.2010.512460.

Meemken, E.-M., 2021. Large farms, large benefits? Sustainability certification among family farms and agro-industrial producers in Peru. World Development 145, 105520. 10.1016/j.worlddev.2021.105520.

Mehraban, N., Kubitza, C., Alamsyah, Z., Qaim, M., 2021. Oil palm cultivation, household welfare, and exposure to economic risk in the Indonesian small farm sector. J Agric Econ. 10.1111/1477-9552.12433.

Merten, J., Nielsen, J.Ø., Rosyani, Soetarto, E., Faust, H., 2021. From rising water to floods: Disentangling the production of flooding as a hazard in Sumatra, Indonesia. Geoforum 118, 56–65. 10.1016/j.geoforum.2020.11.005.

Minten, B., Koru, B., Stifel, D., 2013. The last mile(s) in modern input distribution: Pricing, profitability, and adoption. Agricultural Economics 44 (6), 629–646. 10.1111/agec.12078.

Qaim, M., Sibhatu, K.T., Siregar, H., Grass, I., 2020. Environmental, economic, and social consequences of the oil palm boom. Annual Review of Resource Economics 12 (6), 1–24. 10.1146/annurev-resource-110119-024922.

Romero, M., Wollni, M., Rudolf, K., Asnawi, R., Irawan, B., 2019. Promoting biodiversity enrichment in smallholder oil palm monocultures – Experimental evidence from Indonesia. World Development 124, 104638. 10.1016/j.worlddev.2019.104638.

Santika, T., Wilson, K.A., Budiharta, S., Law, E.A., Poh, T.M., Ancrenaz, M., Struebig, M.J., Meijaard, E., 2019. Does oil palm agriculture help alleviate poverty? A multidimensional counterfactual assessment of oil palm development in Indonesia. World Development 120, 105–117. 10.1016/j.worlddev.2019.04.012.

Schwarze, S, Euler, M, Gatto, Hein, J, Hettig, E, Holtkamp, M, A., Izhar, L, Kunz, Y, Lay, Merten, Moser, Mußhoff, O, Otten, F, Qaim, Soetarto, Steinebach, Trapp, K, Vorlaufer, Faust, H, 2015. Rubber vs. oil palm: An analysis of factors influencing smallholders' crop choice in Jambi, Indonesia. University of Goettinge. EFForTS Discussion Paper Series No. 11, Goettingen.

Sharma, A., Bailey, A., Fraser, I., 2011. Technology adoption and pest control strategies among UK cereal farmers: Evidence from parametric and nonparametric count data models. Journal of Agricultural Economics 62 (1), 73–92. 10.1111/J.1477-9552.2010.00272.X.

Sibhatu, K.T., 2019. Oil palm boom and farm household diets in the tropics. Frontiers in Sustainable Food Systems 3, S133. 10.3389/fsufs.2019.00075.

Sibhatu, K.T., 2020. Response: Commentary: Oil palm boom and farm household diets in the tropics. Frontiers in Sustainable Food Systems 4. 10.3389/fsufs.2020.00107.

Sibhatu, K.T., Krishna, V.V., Qaim, M., 2015. Production diversity and dietary diversity in smallholder farm households. Proceedings of the National Academy of Sciences of the United States of America 112 (34), 10657–10662. 10.1073/pnas.1510982112.

Sibhatu, K.T., Qaim, M., 2018. Farm production diversity and dietary quality: linkages and measurement issues. Food Security 10 (1), 47–59. 10.1007/s12571-017-0762-3.

Steinhübel, L., von Cramon-Taubadel, S., 2021. Somewhere in between Towns, Markets and Jobs – Agricultural Intensification in the Rural–Urban Interface. Journal of Development Studies 57 (4), 669–694. 10.1080/00220388.2020.1806244.

Steinhübel, L., Wegmann, J., Mußhoff, O., 2020. Digging deep and running dry – the adoption of borewell technology in the face of climate change and urbanization. Agricultural Economics 51 (5), 685-706. 10.1111/agec.12586.

Tessema, Y.M., Asafu-Adjaye, J., Kassie, M., Mallawaarachchi, T., 2016. Do neighbours matter in technology adoption? The case of conservation tillage in northwest Ethiopia.

Umlauf, N., Adler, D., Kneib, T., Lang, S., Zeileis, A., 2015. Structured Additive Regression Models: An R Interface to BayesX. J. Stat. Soft. 63 (21), 1–46. 10.18637/jss.v063.i21.

UN, 2019. Official homepage United Nations (UN): Sustainable Development Goals. UN. https://sustainabledevelopment.un.org/sdgs. Accessed 20 January 2021.

Vandercasteelen, J., Beyene, S.T., Minten, B., Swinnen, J., 2018. Cities and agricultural transformation in Africa: Evidence from Ethiopia. World Development 105, 383–399. 10.1016/j.worlddev.2017.10.032.

Wollni, M., Andersson, C., 2014. Spatial patterns of organic agriculture adoption: Evidence from Honduras. Ecological Economics 97, 120–128. 10.1016/j.ecolecon.2013.11.010.

Appendix

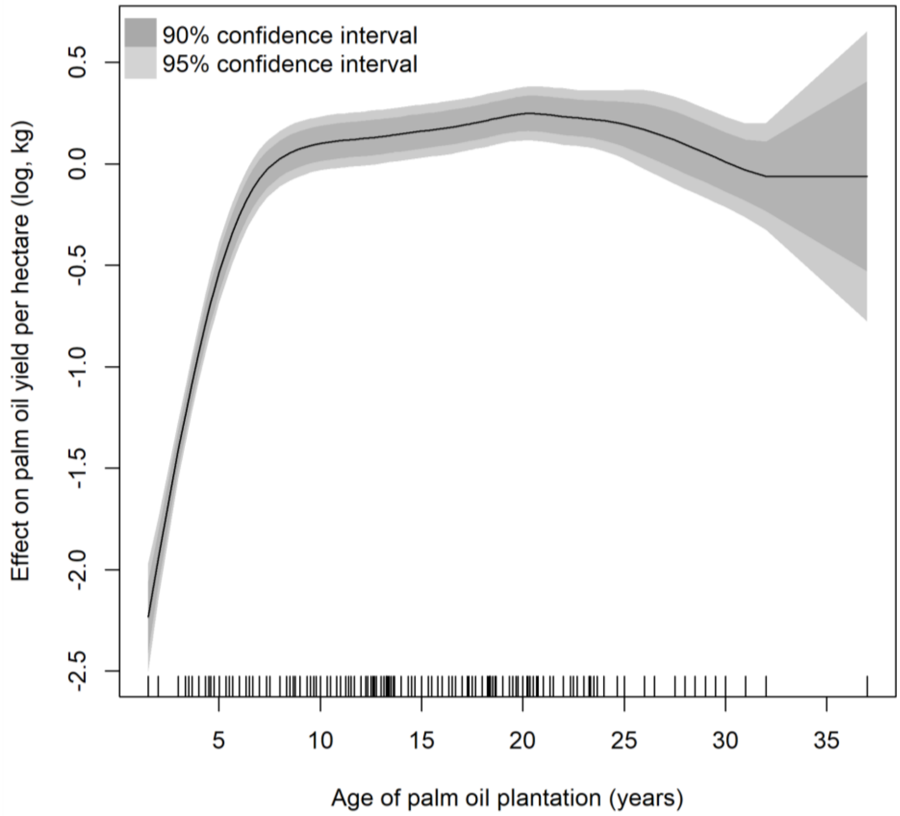

Figure A.1: Estimated one-dimensional splines from the model without geosplines. (*Note*: CI: Gray = 90 percent and dark gray = 95 percent)

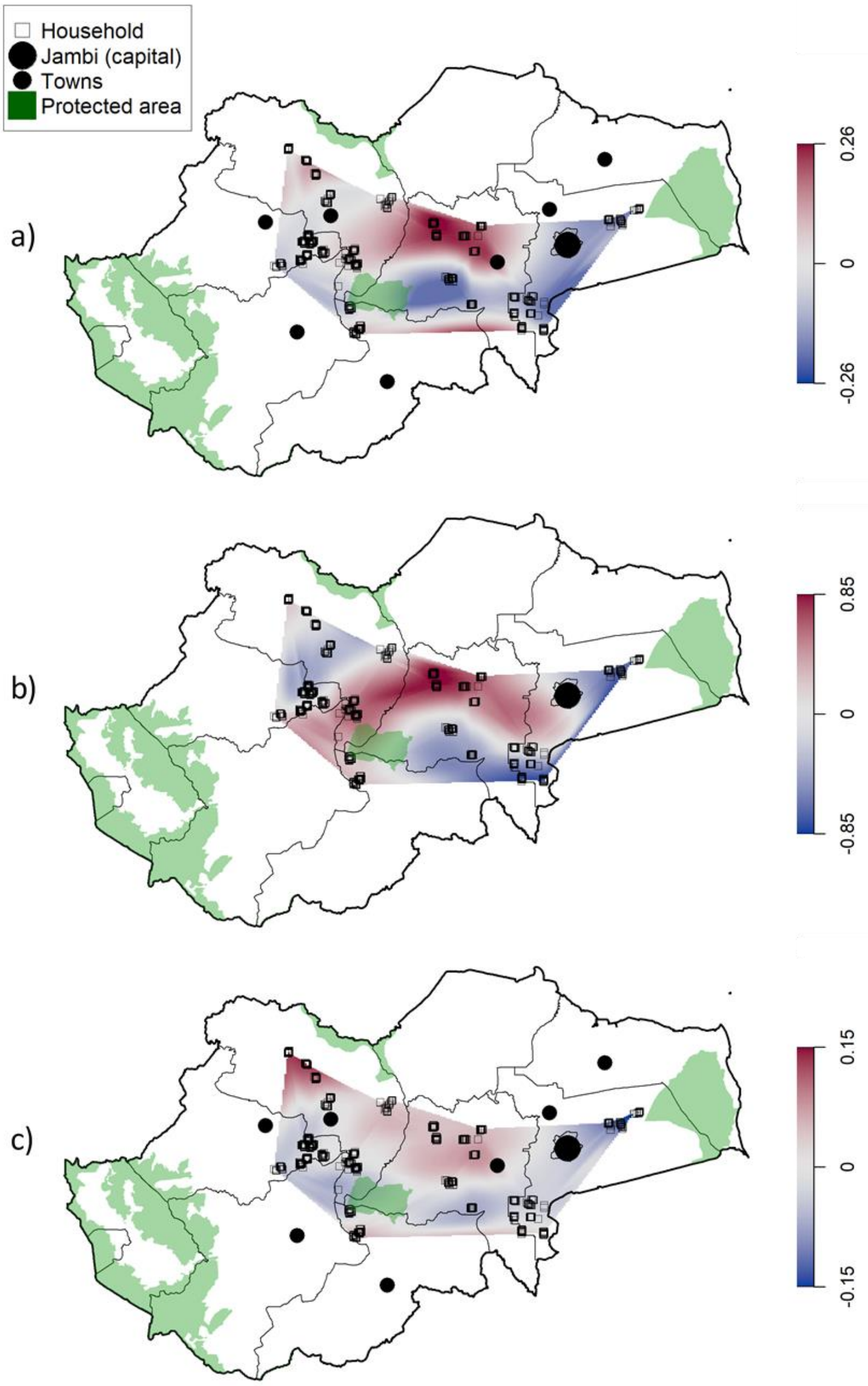

Figure A.2. Estimated geosplines based on Equation (2) excluding village random effects, a) oil palm yields, b) input use, c) output price.

Notes: Areas colored in red indicate above-mean yields per hectare (a), higher input use (b), and higher output prices (c), while those in darker blue, clusters of lower yield, input use, and output prices. The scales represent coefficient estimates. Outcome variables are log-transformed. Regency names (depicted in Figure 1) are left out for clarity.

Table A.1. Estimation results, Equation (2), excluding village random effects

	(1) Yield /ha (log, kg)	(2) Input use / ha (log, '000 IDR)	(3) Output price /kg (log, '000 IDR)
Socio-economic variables			
Age of hh head (years)	0.001 (0.002)	-0.001 (0.008)	0.0003 (0.001)
Education hh head (years)	0.007 (0.005)	0.032 (0.022)	0.002 (0.001)
Household size (count)	0.009 (0.012)	-0.066 (0.049)	0.0012 (0.003)
Village type (dummy, trad. /transm.)	0.057 (0.088)	-0.183 (0.327)	-0.016 (0.024)
Total land owned (log, ha)	0.007 (0.036)	0.037 (0.146)	0.038*** (0.009)
Input cost / ha (log, '000 IDR)	0.042*** (0.009)	NA	0.005** (0.002)
Harvest /ha (log, kg)	NA	NA	0.043*** (0.009)
Number of plots (count)	0.031 (0.022)	0.163* (0.087)	0.008 (0.005)
Age of plantation (years)	0.031*** (0.004)	0.017 (0.015)	0.001 (0.001)
Share systematic land title	0.102* (0.055)	0.508** (0.217)	0.035** (0.013)
Cultivate other crops (dummy)	-0.0002 (0.050)	-0.698*** (0.196)	-0.033*** (0.012)
Aware of RSPO certification (dummy)	-0.035 (0.098)	0.499 (0.387)	0.016 (0.024)
Geographic variables			
Distance to nearest city (Km)	-0.005 (0.006)	-0.036 (0.023)	-0.004** (0.002)
Distance to Jambi city (Km)	0.003 (0.004)	0.026* (0.014)	0.001 (0.001)
Distance to nearest mill (Km)	-0.002 (0.007)	0.010 (0.023)	-0.001 (0.002)
Altitude (meter above sea level)	-0.003** (0.001)	-0.002 (0.005)	-0.0003 (0.0003)
Access to a river (dummy)	0.017 (0.068)	0.303 (0.262)	-0.010 (0.017)
Hh located at edge of forest (dummy)	0.236* (0.135)	-0.0003 (0.481)	0.172*** (0.039)
Regency dummy (ref. Batanghari)	✓	✓	✓
Bungo	0.238 (0.309)	-1.586 (1.131)	-0.160* (0.085)
Muaro Jambi	0.537** (0.229)	1.148 (0.785)	-0.196*** (0.071)
Sarolangun	0.064 (0.402)	-1.272 (1.355)	-0.294** (0.127)
Tebo	0.054 (0.268)	-1.882** (0.944)	-0.253*** (0.079)
Intercept	8.490*** (0.400)	4.785*** (1.363)	-0.508*** (0.148)
No. Obser.	793	793	793
AIC	-288.41	1909.54	-2526.33

Notes: Coefficients with standard errors in parentheses are reported. Asterisks represent p-values, *** <0.01, ** <0.05 and *<0.1. NA – not applicable; all estimations with geosplines. Hh – household head.