Price incentives and unregulated deforestation: Evidence from Indonesian palm oil mills *

Valentin Guye,[†] Sebastian Kraus[‡]

January 31, 2023

Latest version

Abstract

We create a novel, spatially explicit microeconomic panel of Indonesian palm oil mills, to provide the first estimates of deforestation price elasticities based on observations of the actual prices paid at mill gates. To identify price elasticity, we spatially model how deforestation in upstream plantations is exposed to downstream, conditionally exogenous, shocks on mill-gate prices. We provide the first evidence that deforestation for smallholder plantations, as well as illegal deforestation, are price elastic. This implies that a price instrument can disincentivize deforestation where it is most difficult to monitor, and contain leakages from conservation regulations.

*JEL-codes: Q01, Q02, Q15, Q57, Q23, Q24, Q21

For their helpful comments, we thank Ludovic Bequet, Ondine Berland, Raja Chakir, Sabine Fuss, Jérémie Gignoux, Robert Heilmayr, Nicolas Koch, François Libois, Hugo Valin, an anonymous reviewer for the FAERE working paper series, and seminar participants at the 2021 International Development Economics Conference, the 26^{th} EAERE Annual Conference, the 2021 AERE Summer Conference, Restore+ Output sessions, and internal seminars at UMR Economie Publique and at MCC. We are grateful to Jason Benedict and Robert Heilmayr for their support with the Universal Mill List. We thank Claudia Guenther and Hanif Kusuma Wardani for their invaluable research assistance, and Andrew Hirst for his professional proofreading. The authors declare no conflict of interest. The parties this research received funding from are not accountable for its content. The authors are solely responsible for any error or omission.

†Valentin Guye. Contact author - valentin{dot}guye{at}inrae{dot}fr - Université Paris-Saclay, INRAE, AgroParisTech, Paris-Saclay Applied Economics, 91120, Palaiseau, France), Mercator Research Institute on Global Commons and Climate Change (MCC); Humboldt Universität zu Berlin.
Valentin Guye acknowledges funding from the French Ministry for Higher Education and Research, from the German Agency for Academic Exchange Services (DAAD) and from "l'Agence National de la Recherche" within the Programme CLAND ANR-16-CONV-0003.
‡Sebastian Kraus. Currently Bundesfinanzministerium. At time of writing Technische Universität Berlin; Berlin and Mercator Research Institute on Global Commons and Climate Change (MCC) EUREF Campus 19, Torgauer Straße 12-15, 10829 Berlin, GERMANY. Sebastian Kraus acknowledges funding by the RESTORE+ project (http://www.restoreplus.org/), part of the International Climate Initiative, supported by the Federal Ministry for the Environment, Nature Conservation, and Nuclear Safety (BMU) on the basis of a decision adopted by the German Bundestag.

1 Introduction

Tropical deforestation for oil palm plantations is a major source of global biodiversity loss and climate change. It accounts for 5% of the global greenhouse gas (GHG) emissions since 1986 (Hsiao 2022). Price incentives from ever-growing agricultural markets are one of the major drivers of tropical deforestation (Busch and Ferretti-Gallon 2017; Leblois et al. 2017), which hinders the effectiveness of conservation policies, as tropical deforestation is difficult to monitor. Regulators need local insights into the effect of price incentives on deforestation to design effective environmental policies or to predict deforestation leakages from their jurisdictions to the tropics through commodity markets (Hertel 2018). Yet, even in Indonesia, which supplies over half of the global market for palm oil, and from where nearly half of the world's GHG emissions from land use change and forestry have come since 2000 (WRI 2015), it is unclear how the various actors upstream of the palm oil supply chain react to the price incentives that actually pass through to them.

Do actual price incentives influence deforestation, and which segments of the oil palm sector are more or less responsive? We estimate price elasticities of deforestation across the Indonesian oil palm sector. We build the first spatially explicit dataset of prices paid at palm oil mill gates from 1998 to 2015. On average, we find that a 1% increase in crude palm oil price signals increases the conversion of primary forest to oil palms by 1.5%. Looking at different segments of the oil palm sector, we find that both industrial and smallholder plantations are price elastic, and that illegal deforestation only drives the effect. This constitutes evidence that the segments the most difficult to monitor can be incentivized away from deforestation. Our finding that legal deforestation is not price elastic indicates the existence of leakages from legal to illegal deforestation in the presence of economic opportunities and weak law enforcement. Furthermore, our empirical price elasticity estimates can be helpful for predicting how deforestation reacts to biofuel subsidies or trade policies against imported deforestation.[1] Providing a commodity- and country-specific estimate, breaking it down to different economic agents, deriving it from data on actual observed prices, and tackling concerns of endogeneity can help to perform more accurate predictions (Wicke et al. 2012).

Data on the Indonesian palm oil supply chain available to researchers has been limited. Previously, only the location and total capacity of most mills was known, together with the establishment date for a subset of them. We geo-localize palm oil mills based on the full Indonesian manufacturing census (IBS).[2] For approximately half of all known Indonesian mills, we observe, in particular, annual input (palm fruits) and output (crude palm oil), mill-gate prices, public, private and foreign ownership shares, as well as crude palm oil export shares. We use data based on satellite imagery to measure deforestation around mills at a high resolution. We detect deforestation as 30m-pixel events of primary forest loss, conditional on eventual oil palm plantation development. Industrial and smallholder plantations typically differ in scale and landscape. We define illegal deforestation as occurring outside a known concession and inside a state forest zone. Our sample for estimation is a 2002-2014 annual panel of $3 \times 3\,\mathrm{km}$

[1] Indonesia exports 75% of its palm oil production, representing 13% of all its exports (Pacheco et al. 2017).

[2] In the economics literature, this dataset has also been referred to as *Statistik Industri*; see, for instance, Amiti and Konings (2007).

plantation sites in the Sumatra and Kalimantan islands, where most deforestation due to oil palm plantations occurred during the period.

Our estimation strategy builds on the fact that palm fruits deteriorate quickly from harvest to processing. This means that each plantation can only reach a limited number of surrounding mills in time. For every plantation site that can reach at least one mill, we average the crude palm oil prices of reachable mills. By assigning higher weights to closer mills we model the relative influences of reachable mills on plantations in a way that is consistent with the palm oil sector's heterogeneity: the weights represent either the odds of being integrated plantation-mill systems, or differential transport costs from the independent plantations to the mills. Finally, we average these annual price signals over the four most recent years. This captures the information that we assume relevant and available to prospective plantations forming expectations of the profitability of the yield-lagging and perennial oil palm crop. Hence, we call our estimate a medium-run price elasticity.

Our identifying variation is the interaction of the plantation-mill spatial distribution, and mill-gate crude palm oil prices. In the latter, we isolate price shocks that are arguably exogenous to the local variations in distances to reachable mills and in deforestation. This identification strategy relates to a shift-share reduced-form design where the conditional exogeneity assumption is on the shifts (Borusyak et al. 2020). To isolate such idiosyncratic price variation, we difference out the annual district-market price, and we compare resulting price departures over time within plantations that can sell to the same mills (by the means of fixed effects). We argue that such spatially-specific price departures are exogenous to changes in local conditions over time for three reasons. First, because crude palm oil is a standardized good (Byerlee et al. 2016), so there is a single quality market. Second, this market is competitive for mills (Pirard et al. 2020), such that they cannot manipulate quantities to sell at a higher price. Third, this market is highly profitable (Byerlee et al. 2016), and therefore it is not at an equilibrium where mills managing to reduce their marginal costs sell at a lower price. Thus, we interpret our exploited variation as being driven by shocks in off-take agreements with, or transport conditions to, mill-specific buyers.

Our main estimate is robust in a range of alternative estimations, including mill catchment radius assumptions, price signal dynamics, control sets, fixed-effects, and clustering levels for statistical inference.

Our main results are threefold. First, we estimate medium-run crude palm oil price elasticity across the overall Indonesian oil palm sector at 1.5. Second, we find that deforestation in both industrial and smallholder plantations is price elastic. Third, we document that illegal deforestation is price elastic, whereas legal deforestation is not. Together, these results have three main implications.

First, segments of the oil palm sector that are more difficult to regulate - illegal industrial or smallholder plantations - can be incentivized away from deforestation. This is an important implication, because the existing conservation schemes - namely the Moratorium on new concessions (Groom et al. 2022) and the Roundtable on Sustainable Palm Oil (Heilmayr et al. 2020) - do not reach these segments. Yet, they are increasingly prevalent: smallholder relative expansion is expected to grow across the country (Schoneveld et al. 2019), and new oil palm frontiers

- in the island of Papua, in particular - seem to largely involve illegal deforestation.[3] To reach such informal segments, recent fiscal conservation policy proposals have devised a taxation on defaults whereby a commodity tax is uniformly levied at choke points (like palm oil mills), but can be refunded against proof of sustainable production (Heine et al. 2020). Our results indicate that such fiscal schemes can work in the Indonesian context.
Second, the results indicate that a price instrument can help conservation regulations to be more effective in similar contexts of weak monitoring. The licensing process that embeds conservation regulations is long and the cost of circumventing it is low. Thus, more stringent regulations are bypassed, unless economic opportunities for illegal deforestation are contained. We show that a price instrument can address such leakage. This is critical, because palm oil prices reached a historical peak in March 2021 and demand for palm oil is expected to keep growing.[4]
Thirdly, the 1.5 price elasticity of deforestation that we estimate for the whole sector implies that a 19% tax on crude palm oil can curb deforestation 29% below the 2002-2014 average (proportional to Indonesia's targeted reduction in GHG emissions under the Paris Agreement). We quantify that, for the whole country, this represents 38.5kha of avoided conversion of primary forest to oil palm plantations annually, and we discuss why this is probably a lower bound. Under a result-based payment scheme, like the United Nations program REDD+ (Reducing Emissions from Deforestation and Degradation), this corresponds to US$123M of yearly revenues.

We provide four additional results, that deepen our understanding of the determinants of illegal deforestation, smallholder deforestation, and of the formation of expectations.
The first shows that, unsurprisingly, price incentives do drive immediate conversion (within 4 years) of primary forest into industrial plantations, but not transitional deforestation, in which plantation development occurs 4 years or more after forest clearing. This is driven by illegal deforestation, indicating that medium-run economic opportunities incentivize to bypass licensing processes. A significant part of those rapid clearings do not lead to immediate plantation development though, probably because of conflicts with local communities or legal proceedings.
Second, we find that unstable prices tend to deter deforestation, especially for industrial plantations. A policy that stabilizes prices would thus risk to increase the pressure from plantation expansion on forests. This meets conclusions from (Lundberg and Abman 2022), in a different context.
Third, we disentangle the effects of palm fruit and crude palm oil prices (but caution is required here as our identification arguments do not apply to fruit prices). We document that deforestation in industrial plantations seem to be actually mainly driven by palm fruit prices, which vary consequentially to crude palm oil prices. The output of vertically integrated plantations, (i.e., plantation-mill systems) is crude palm oil, while the output of independent plantations is the palm fruit. Thus, we highlight the role of independent plantations in price-driven deforestation. More importantly, we find that deforestation in smallholder plantations decreases with palm fruit prices and increases with crude palm oil prices. This suggests that mill owners - usually

[3] `https://news.mongabay.com/2018/11/the-secret-deal-to-destroy-paradise/`

[4] Global demand grew by 7% annually between 1980 and 2013 (Cramb and McCarthy 2016), and is likely to keep doing so as the Government of Indonesia (GoI) increases its biodiesel blend mandates.

companies - wishing to benefit from higher output/input price ratios, are the ones setting up the timing and location of smallholder encroachment on forests.
Fourth, we disentangle the effects of short- and medium-run variation in crude palm oil prices. We find that short-run (annual) price changes alone do not affect deforestation, but strengthen the medium-run price elasticity. This indicates that oil palm decision-makers, to form distant expectations, look at short-run price signals only to confirm medium-run trends. Looking at elasticities to annual prices, we find that only the two most recent ones conditionally influence the long-run investment (with sunk economic and environmental costs) that is deforestation. We emphasize that providing medium-run elasticities is relevant to inform policies, such as taxes or tariffs, that are typically enforced for more than one year, but not necessarily expected to last more than a political mandate (Berry 2011).[5]

The remainder of this paper is organized as follows: Section 2 relates our work to the existing literature. Section 3 defines the main concepts used in the paper, and introduces our data. In Section 4, we present our empirical framework in five parts: the plantation-mill relationship model, and the estimation, identification, inference, and sampling strategies. In section 5, we present and discuss our main estimates, a mechanism analysis, and finally scaled-up counterfactuals. Section 6 concludes. References and appendix follow in this order.

2 Related literature

This paper principally contributes to the literature shedding light on the economic incentives of land use change.[6] Our results also relate to the literature on the relationships between conservation regulation and market incentives (Harding et al. 2021). The role of prices in oil palm-related deforestation is a case of particular interest, as indicated by recent efforts to relate time series of deforestation and palm oil prices in the Global Forest Review (Goldman et al. 2020) and in Gaveau et al. (2021). Yet, data availability has constrained the identification of causal relationships, as well as heterogeneity and mechanism analyses. Thanks to the new spatially explicit microeconomic data we produce, and to recent remote-sensing data sets, we are able to advance the literature in these directions.

We provide the first price elasticity estimates specific to smallholders and illegal oil palm deforestation.[7] These results relate our work to the field of development economics. We are also the first to explore how short-run prices and palm fruit prices (FFB) affect deforestation and how they interact with medium-run crude palm oil prices in doing so. Methodologically, we are the first to estimate country-level elasticities with actual price observations in the oil palm context. Other studies, using imputed price measures, provide estimates that can be interpreted similarly to some of ours. Yet, in these studies, the price elasticity of deforestation

[5] Yet, annual elasticities also provided in this paper can be useful to dynamic models and/or simulations of punctual market shocks.

[6] This is a large literature and we point in particular to Busch and Ferretti-Gallon (2017) for a review; Souza Rodrigues (2019) in the Amazon context; and Leblois et al. (2017) for a cross-country analysis.

[7] The closest literature on smallholders, based on survey data, estimates a positive correlation between crude palm oil and local land prices (Krishna et al. 2017) and opportunity costs of conservation (Cacho et al. 2014).

is not the main parameter, and thus their authors may have naturally focused less on identification concerns about it. Moreover, the estimates we provide are specific to primary forest conversion to oil palms, and hence exclude deforestation in broader senses (such as not imputable to oil palm plantations, or in already degraded forest), which is not always the case in previous comparable research.[8] In the following, we explain how observing actual local prices allows improved identification of the price elasticity compared to the existing literature. In Appendix F, we attempt to compare our results with estimates from other studies.

Wheeler et al. (2013) were the first to establish a positive correlation between time series of palm oil futures prices and forest loss alerts at a monthly rate. Subsequent studies have advanced the causal price effect identification by adding spatial variation. They proxied local farm prices by interacting international prices with agro-climatic measures of local suitability for palm plantations (Busch et al. 2015; Cisneros et al. 2021; Hsiao 2022). First, the suitability-price interaction proxy is subject to dynamic reverse causality bias. Indeed, past deforestation can influence current deforestation. It is also possible that more deforestation systematically occurs in more suitable places for oil palms and in years before the international price declines as a result of the increased production in Indonesia, the largest supplier globally. Our identifying variation exploits idiosyncratic shocks in CPO prices at the mill level and mills are price takers on the CPO market. Thus, our estimated price elasticity should suffer less from reverse causality bias. Second, the suitability for oil palm used in previous studies can correlate with suitability for other crops while palm oil prices can correlate with prices of other commodities, making it unclear whether the existing estimates measure the palm oil price elasticity specifically.[9] Another concern with prior approaches is that the suitability-price interaction proxy can be subject to important measurement error, including systematic error. Suitability variables and international prices may be systematically more precise measures of respectively the true exposure and the true treatment for some observations, biasing estimates if, for instance, information on suitability or price pass-throughs are correlated with deforestation. In contrast, our analysis relies on the actual prices paid at mill gates. We model how they are perceived across potential plantation sites, consistently with heterogeneity in plantation-mill integration and transport costs. Last but not least, this observational level is relevant because mills are pivotal in the palm oil value chain: they are the most influential actors over plantations, while they can still be influenced and monitored by downstream corporate or public actors. (Purnomo et al. 2018). This pivotal situation is critical for policy interventions, and it can also be leveraged by researchers as a quasi-experimental design to isolate downstream drivers in upstream relationships.

[8]See Hansen et al. (2014) for a discussion on the use of the Global Forest Change data to study deforestation.

[9]For instance, (Cisneros et al. 2021) find that a placebo interaction of the palm oil price with suitability for other crops still has an effect on deforestation.

3 Contextual background and data

This section presents the concepts of mills, plantations, deforestation and price signals used in this paper, together with the data we use to observe them empirically. Then, we explain our sampling strategy and show descriptive statistics.

3.1 A new, spatially explicit, microeconomic panel of palm oil mills

Palm oil mills are factories that process fresh fruit bunches (FFB) from palm trees into crude palm oil (CPO). In the paper, *mill-gate prices* refers to mill-level mean unitary values of either FFB or CPO. We semi-manually merge two existing data sets - the Indonesian manufacturing census (IBS) and the Universal Mill List (UML) - to produce an original, spatially explicit, microeconomic data set of palm oil mills in Indonesia from 1998 to 2015. In Appendix D.1 we describe these data sets in more detail and explain how we merged them. Input-output variables, as well as village identifiers, are usually not provided to researchers with IBS. They were essential in building the spatially explicit price data used in this paper. The final spatially explicit mill sample comprises 587 palm oil mills. 466 of them are matched with a mill referenced in the UML and hence have exact coordinates, while 121 are not matched with the UML but are approximately geo-localized at their village centroids. Table A.1 shows descriptive statistics of Indonesian palm oil mills, along with evidence that the subset of these mills used in the present analysis is not significantly different from the overall sample of palm oil mills in the manufacturing census.

3.2 Oil palm plantations

Throughout this paper, we use the term *plantations* to designate micro-economic agents deciding where and when to clear forest for the purpose of planting oil palms.[10] Empirically, we do not observe the actual boundaries between plantations. Thus, we approximate the theoretical individual plantations with square land parcels of an equal size. Each year, deforestation in each of these grid cells is assumed to result from decisions taken by an homogeneous, profit-maximizing plantation agent. We choose the typical size of grid cells to be $3 \times 3\,\mathrm{km}$ (900ha). This is the outcome of a trade-off: it is small enough to grasp very local variations in deforestation and in influence from surrounding mills. Yet, it is large enough to keep computation times reasonable.[11] We use the maps from Austin et al. (2017) to study industrial plantations. We pool small and mid-sized plantation maps from Petersen et al. (2016) to study smallholder plantations.[12] See Appendix D.2 for a description of these data sets.

[10]We purposely do not refer to 'landholders', 'landowners', 'growers' or 'farmers', in order to abstract as much as possible from notions of ownership, legality, or management. This seeks generality over the diversity of actors that may be involved in the decision process towards development of a plantation. Moreover, it is important to note that, in this study, we refer to plantations as agents prospecting to plant oil palms and not as a realized land use.

[11]This is also the size of grid cells in Busch et al. (2015).

[12]Where these maps overlap with the industrial plantation map, we characterize plantations as industrial, as remote sensing for this landscape is less error-prone.

Industrial plantations. Industrial plantations are large, grid-shaped landscapes, ranging from a hundred hectares to hundreds of thousands of hectares (Gaveau et al. 2016; Austin et al. 2017). They represent the majority of the planted area and production in Indonesia. They are developed by companies or public governments. Some industrial plantations are integrated with mills and sometimes also further downstream with refineries and exporters, but, in the light of the best knowledge of the field, this integration seems limited (Pirard et al. 2020). Hence, industrial plantations are heterogeneous in how they sell their fruits, from internal transactions to partial off-take agreements, to selling on the local spot market.

Smallholder plantations. The term 'smallholder' lacks a common definition, but is often used in contrast with some or all of the characteristics of industrial plantations. Here, we refer to smallholder plantations in accordance with the measurement in our secondary data: patches of palm trees smaller than 100ha, comprising at least 50% of a mosaic landscape wider than 100ha (Petersen et al. 2016). For two reasons, this definition implies that the smallholder plantations we observe tend to be informal, in the sense that their development is framed by weaker to no institutions. First, because it excludes *plasma* plantations which are jointly developed with industrial plantations wider than 250ha as part of a smallholder scheme, and therefore not distinguishable from them (Paoli et al. 2013; Byerlee et al. 2016; Gaveau et al. 2022). Rather, our smallholder definition captures independent smallholder plantations, developed outside such schemes and which fruits do not have to be sold to one mill exclusively (Cramb and McCarthy 2016; Baudoin et al. 2017).[13] These independent smallholders have driven smallholder expansion during our study period (Byerlee et al. 2016; Euler et al. 2016). Second, because plantations smaller than 25ha only require a Plantation Registration Certificate (STD-B), not a Plantation Business License (IUP-B), and are thus exempt from most legal processes (Paoli et al. 2013; Jelsma et al. 2017).

3.3 Deforestation

We use the term *deforestation* to refer to land use change from primary forest to an oil palm plantation. To compute annual maps of deforestation, we overlay a map of the extent of primary forest in 2000 (Margono et al. 2014), annual maps of forest loss (Hansen et al. 2013), and the maps of oil palm plantations mentioned above. These data sets and how we combine them are described in more detail in Appendix D.2.[14]

Conceptually, our outcome of interest is the decision to clear primary forest land with the aim of growing oil palms. Empirically, we count 30m pixel-year events of primary forest loss that occur within the industrial and smallholder plantation extents in 2015 and 2014 respectively.

[13] Independent smallholders are heterogeneous in ownership and management (households, cooperatives, or companies), and in their integration in the supply chain (through middle-men or not). See Euler et al. (2016) and Jelsma et al. (2017) for more insights into independent smallholders specifically.

[14] We note here that our main approach does not count forest degradation as deforestation, because the tree loss pixel-event is counted only once, the year a near-zero canopy closure is observed (Hansen et al. 2013). To document how degradation reacts to prices, we alternatively use a secondary forest measure, described in Section D.2.

For this to reflect the conceptual outcome, we need to assume that on average, where such land use change is observed, developing a plantation was the prime motivation to clear forest land, irrespective of the nature of the intermediary land use or of its duration, up to 12 years.[15] This assumption is not critical for two reasons. First, where it can be grown, oil palm is the most lucrative land use (Byerlee et al. 2016). Second, delays (voluntary or not) between forest clearing and plantation development are uncommon (Gaveau et al. 2022).

Immediate and transitional conversion. To better understand the conversion process and how it reacts to prices, we make a further distinction in our analysis, between immediate and transitional conversion. For industrial plantations only, we can observe the time lapse between the forest loss event and the year when a plantation is observed for the first time in the half-decadal data from Austin et al. (2017). Conversion is deemed immediate if the time lapse is between 0 and 4 years. It is deemed transitional if the time lapse is between 5 and 12 years. We observe that immediate conversion represents two thirds of the deforestation we measure in 2002-2010 (when we can observe transitional conversion), while it occurs slightly less often. This indicates that transitional conversion occurs at the margin of immediate conversion.

Illegal deforestation. Observing illegality is generally challenging, and this is all the truer in the outer islands of Indonesia where the line between legality and illegality is blurred by weak institutions, and data is scarce and from diverse sources. To overcome these challenges, we focus our efforts on mitigating commission errors in observing illegal deforestation and we assume in what follows that omission errors are conditionally independent from the price elasticity.[16]

The best available maps of oil palm concessions and Indonesian legal land designation are known to be incomplete (Greenpeace 2011; *Indonesia legal classification* 2023). To mitigate commission errors due to this incompleteness, we impose a combination of conditions: we deem deforestation as illegal if it occurs outside a known concession *and* inside a forest zone designation where oil palm cultivation is forbidden.[17]

A further issue is that both maps are composite snapshots — they do not specify the date the concessions were issued, nor changes in land designation. Different settings could thus lead to commission error in observing illegal deforestation. Where the land designation snap-

[15] Both plantation data sets recognize areas with signs of future cultivation as plantations. Hence, we can observe deforestation up to 2014, the latest common year for both industrial and smallholder plantations. Since the first year in our sample for analysis is 2002, the maximum time laps is 12 years.

[16] Indeed, we do not seek to gauge total illegal deforestation, but rather to estimate the price elasticity for a representative sample of illegal deforestation.

[17] Any of Permanent Production Forest (*Hutan Produksi Tetap*), Limited Production Forest (*Hutan Produksi Terbatas*), Conservation Forest (*Hutan Konservasi*) or Protection Forest (*Hutan Lindung*). Concretely, those umbrella groups comprise the following classes in our data: HL, HP, HPT, HK, KSA, KPA, KSAL, CA, SM, TN, TWA, Tahura, KSAL, KPAL CAL, SML, TNL, TWAL, TB and Hutan Cadagan. See (MoF 2019) for definitions of specific acronyms. The land designation map and the concession map can be downloaded from the Global Forest Watch website (*Indonesia legal classification* 2023; Greenpeace 2011).

shot precedes a reclassification into convertible forest, or the concession snapshot precedes a concession issuance, we could incorrectly deem deforestation occurring afterwards as illegal. Such commission error in illegal deforestation due to the snapshots being outdated is probably limited though. The land designation and the concession maps are based on 2010 official data and the Moratorium on new concessions in primary forest came into force in 2011. New concessions have been issued by local governments despite the moratorium, but this may be considered part of questionable legal processes, especially with respect to the central government (Enrici and Hubacek 2016). Moreover, alternatively using a map of concessions in 2020[18] to identify illegal deforestation shows that our results are robust to concession issuance post-2010 (Table A.4). The opposite problem of too recent snapshots causing commission error in illegal deforestation is even more unlikely, as it would emerge from cases of land reclassification into protected forest, or concession revocation, after deforestation occurred.[19] Finally, note that too recent snapshots risk to yield higher commission errors in legal deforestation. For this reason, we do not use the 2020 concession map in our main analysis, and even with the 2010 concession map, our results in the class of legal deforestation should be taken more cautiously.

3.4 Price signals: a model of plantation-mill relationships

We assume that deforestation results from decisions taken by plantations. The typical decision rule is the comparison of the expected discounted present utilities (or profits) from alternative inter-temporal scenarios, defined by the timing and the amount of deforestation.[20] To form such expectations, we assume that plantations ground on privately observed informational elements (Stavins 1999). The one such element we are interested in here is the price signal. For every plantation, every year, there is a true, privately observed set of prices that plantations ground on to form their expectations. We approximate this true price signal with a mix of the prices at the gates of the mills the plantations can reach in time, before the fruits spoil. What constitutes this mix has implications for how the different segments of the oil palm sector contribute to our estimation. The next four paragraphs explain how.

The set of reachable palm oil mills. Oil palm trees produce fruits that can be harvested around 10 times a year, for around twenty years. Once harvested, these fruits damage quickly because they rot fast and bruise easily during transport. The fruits are brought by trucks and/or by river boats to factories, called mills, that process them into crude palm oil. The quantity of oil derived from a tonne of fruit increases with the quality of the fruits and thus decreases with

[18](*Greenpeace Kepo Hutan Public Downloads - Google Drive* 2023)

[19]Revocation is intended in cases of chronic non-compliance with the plantation development laws (Paoli et al. 2013). Moreover, there is no anecdotal support for significant oil palm concession revocation in the period of interest (Morel et al. 2016; Mongabay 2022).

[20]The counterfactual scenario includes both conservation and deforestation to other land uses. Conservation includes both expansion of agriculture outside forests, or no expansion (i.e., intensification or not entering the market as a new plantation). We do not distinguish between these alternative scenarios in our analysis.

the distance from the trees to the mill (Byerlee et al. 2016). This constraint leads to spatial proximity between mills and plantations.

For each plantation, we determine a set of reachable mills for each year. Mills are considered reachable if they are within a circular area around the plantation, determined by a catchment radius parameter. We assume that freshly harvested palm fruits can potentially be brought to any mills within this area without deteriorating too much. Mills beyond the catchment radius are not reachable and thus are assumed to have no influence on the plantation's decision to deforest.

In this study, our preferred catchment radius is $30\,\mathrm{km}$ in Sumatra and $50\,\mathrm{km}$ in Kalimantan. Choosing the value for this parameter results from a trade-off. On the one hand, a too short catchment radius implies observing too few of the plantations experiencing deforestation and biasing our observations towards areas near palm oil mills. On the other hand, a large catchment radius implies spuriously relating plantations to more mills that, despite being reachable, are actually unrelated. This would, in turn, make our price elasticity estimate less precise. This trade-off justifies that we assume a different catchment radius for Sumatra and Kalimantan. First, in Sumatra, typically most deforestation occurs within $30\,\mathrm{km}$ of mills, while in Kalimantan a significant share occurs farther away (see Table 1). Second, the higher mill concentration in Sumatra reduces the likelihood that a plantation will be influenced by prices from mills located farther than $30\,\mathrm{km}$ away.[21] In Appendix C we discuss estimates under alternative catchment assumptions.

Table 1: Deforestation accumulated over 2002-2014, in kha.

	Sample	30km from sample mill	50km from sample mill	Total
Sumatra	221.72	564.55	702.02	801.40
Kalimantan	150.32	321.92	565.81	1015.62
Both	372.05	886.47	1267.83	1817.02

NOTE. This table shows measures of accumulated deforestation from 2002 to 2014 in different groups of Indonesian plantation sites. Deforestation is counted as primary forest loss eventually (by 2015) replaced with oil palm plantations (either industrial or smallholders). The sample of plantation sites is the one we actually use in estimations. Sample mills are the 587 palm oil processing plants from the Indonesian manufacturing census that we have geo-localized.

Mill influence intensities. A plantation can reach several mills, but not all mills are equally influential. We do not directly observe how prices paid at every reachable mill enter the price

[21]The existing literature helps us get a sense of magnitudes for catchment radii of palm oil mills. According to Harris et al. (2013), only 15.3% of oil palm plantations are farther than $30\,\mathrm{km}$ from a mill. This study is based on Gunarso et al. (2013) for plantation data and Global Forest Watch for palm oil mill data, for Indonesia, Malaysia, and Papua New Guinea. 44.5% of oil palm plantations are within $10\,\mathrm{km}$ of a mill, and 8.1% are farther than $50\,\mathrm{km}$. The Center for International Forestry Research (CIFOR), in its online atlas (`https://atlas.cifor.org/borneo/#en`) applies a $10\,\mathrm{km}$ buffer around mills. In Peninsular Malaysia, a region comparable to Sumatra, Shevade and Loboda (2019) reports almost no deforestation due to oil palms beyond $40\,\mathrm{km}$ to a mill.

signal that is observed by each plantation. Therefore, we model these intensities using straight-line distances between each plantation and its annual set of reachable mills. More precisely, we model the price signal as the standardized invert-distance weighted average of prices at reachable mills.[22] Hence, every plantation i can reach M_{it} mills at time t and each of these mills, m, is at a distance d_{im} and has a mill-gate price P_{mt}. The annual (short-run) price signal perceived at plantation i in year t is:

$$S_{it} = \sum_{m}^{M_{it}} \frac{d_{im}^{-1}}{\sum_{n}^{M_{it}} d_{in}^{-1}} * P_{mt} \tag{1}$$

As depicted in Section 3.2, both industrial and smallholder plantations may be vertically integrated to different extents, from full integration with one mill to having partial off-take agreements, to selling on the spot market only (full independence). We do not observe the degree of integration of each plantation. Yet, the standardized inverse-distance weights enable modeling of the relative influence from reachable mills in a way that is consistent over degrees of integration. To see this, consider two main types of plantations: those selling exclusively to (and hence getting a price signal from) one mill, and those selling at least some of their outputs on the local spot market, i.e., to any reachable mill (and getting a composite price signal). Plantations in the former category are typically close to the mill they sell to. Thus, the standardized inverse-distance weights approximate the odds to be integrated with each reachable mill. For plantations in the second category, the standardized inverse-distance weights approximate the expected transport costs to every reachable mill (including fuel costs and fruit quality decline). Prices at mills relatively farther away are less influential, because reaching them from the plantation site is more costly.[23]

Prices of palm fruits and prices of crude palm oil. We know the annual average prices offered at mills' gates for fresh fruit bunches (FFB, the output of independent plantations), as well as those received for crude palm oil (CPO, the output of mills, i.e. of integrated plantation-mill systems of and independent mills). The FFB prices reported by mills in our data may reflect only what they paid to independent plantations, i.e. integrated plantation-mill systems may not report internal FFB transactions. Thus, focusing on FFB prices would risk to exclude integrated plantations from the analysis. Focusing on CPO prices, however, does not exclude independent plantations, because mills are potential price makers on the FFB market. FFB prices are supposed to be collectively determined by provincial governments, firms and farmers. However, it has been shown that they actually result from each mill's discretionary decisions based on its monopsonic market power, the quality of FFB purchased, and each mill's CPO sales (Maryadi et al. 2004; Masliani et al. 2014). Thus, the CPO prices can pass-through to the FFB prices paid to

[22]This is explained here for the price signal, our explanatory variable of interest, but the same method is applied to all mill-level covariates.

[23]This price signal measure implies a higher variance in price signals for plantations that can reach fewer mills (assuming price correlation across reachable mills). From a causal inference perspective, this is not worrisome as we control for the number of reachable mills. From a statistical inference perspective, this is handled by clustering standard errors at the level of the set of reachable mills.

independent plantations.[24] Therefore, in our main analyses we focus on CPO prices to capture the effect in both independent and integrated plantations. Moreover, potential price instruments are more conceivable at the level of the CPO market, and thus CPO price elasticities are more relevant to policy implications. In Section **??**, we disentangle the roles of FFB and CPO prices with respect to each other, notably to gain insights into independent plantations specifically.

Medium-run price signal: assumptions on the formation of expectations. Because oil palm is a perennial crop with a roughly 20-year lifetime, prospective plantations have to form price expectations far ahead into the future. How we model these expectations grounds on three distinct assumptions. First, we assume that plantations attempt to form expectations on the local price as of four years ahead, when the first palm fruits can be sold if clearing and planting occurs contemporaneously (which is not always the case, see Section 3.3).We assume they expect more distant prices to be equal (on a discounted average) to the one in four years. Second, we assume that every year, the best information available to the plantations are the price observations in the four most recent years. This assumption results from a trade-off: on the one hand, it is not credible to assume that plantations look only at the last annual price observation to anticipate the price in four years (in other words, that they form naive expectations). Indeed, an ARIMA analysis over a random sample of plantations' price signal (stationarized) time series shows that crude palm oil price signals are partially auto-correlated beyond the fourth lag of themselves. On the other hand, partial auto-correlation subsists even beyond the eighth lag, suggesting that more price observations than the four most recent ones would be relevant in forecasting the price four years ahead. However, it is not credible either to assume that plantations have access to local price information (mill-gate prices at the different reachable mills) too far in the past.[25] In Section C, we check the robustness of the four-year assumption to shorter and longer lengths. Third, we assume that each annual observation is equally informative. In Table A.5, we show that relaxing this assumption, i.e. jointly estimating elasticities to annual price signals S_{it} and adding them up, yields very similar results. Given this robustness we use the less flexible estimation in the main analyses for the sake of simplicity (when exploring interaction effects in particular). Further, in Section **??**, we show the respective roles and the interaction between the most and least recent price observations. Grounding on these assumptions, in our main approach we model the price expectations with a four-year, equally-weighted moving average, i.e., the average of the current prices and the three past-year

[24] This pass-through cannot readily be estimated in our data, as finding a quasi-experimental setting at the mill-level is not trivial. Hence, using CPO prices as the signal treatment for independent plantations is a reduced form with respect to the mill-level pass-through. From a causal identification perspective, the FFB prices are an intermediate outcome, as we argue that they do not affect CPO prices that are not determined locally, see Section 4.2.

[25] In addition, due to price data unavailability before 1998, assuming that plantations' price expectations ground on longer lags would shorten our study period; and assuming differential lag lengths, depending on data availability, would complicate the interpretation of the results.

prices. We call this the medium-run price signal.

$$Price_{it} = \frac{1}{4}\sum_{l=3}^{0} S_{it-l} \tag{2}$$

Using such medium-run variation allows us to identify price elasticities that are relevant to policy instruments that are typically enforced for more than one year, but not necessarily expected to last more than a political mandate.[26]

3.5 Sampling

Our sample is an annual unbalanced panel of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014.[27] Sumatra and Kalimantan are the two main Indonesian regions where oil palm expansion occurred during our study period (Austin et al. 2017). We further restrict the sample in several dimensions.

First, we include only observations of grid cells from years when at least one geo-localized mill from the manufacturing census is reachable.

Second, we remove observations as soon as they are included in an RSPO certified concession.[28] Indeed, we expect the effect of prices on deforestation to be systematically different there from other areas .[29]

Third, we remove annual records as soon as one of the variables in Equation 3 (the price signal variable mainly) has a missing value. In our case, this has a particular influence on the final sample, because the likelihood that a price signal value is missing decreases with the number of reachable mills. Thus, removing observations with missing values implies that we tend to sample fewer grid cells in remote areas. Another particular implication of removing observations with missing values in our case is that we do not sample records of grid cells in the first 4 years after the first reachable mill is established (as our main, medium-run, price signal measure runs over 4 years). Table 2 shows that this does make a significant difference for the distribution of deforestation, the average price signal, and for the number of reachable mills, in particular. This is not surprising, since inclusion in the sample is a function of the number of reachable mills. We argue that this necessary sampling step does not risk introducing a selec-

[26] We further note that a dynamic structural estimation, as in (Scott 2014; Araujo et al. 2021; Hsiao 2022), or using cross-sectional variation as in (Souza Rodrigues 2019), would constrain the heterogeneity and mechanism analyses that are parts of this paper's contribution. Finally, we note that the existing literature on palm oil price forecasting, to the best of our knowledge, has focused on identifying best models and lags, using monthly and macro price series, and only for short-run forecasting. Thus it is not informative for the present case.

[27] Precisely, 27.8x27.6m pixels aggregate to 3002.4x3008.4m grid cells. We do not include observations from other Indonesian islands, where data is too scarce. In Papua, we have very few observations and in other islands data on oil palm plantation extents are lacking. Although we have data on year 2015 for industrial plantations, we do not include these observations in order to observe them in the same time period as smallholder plantations. We start observing 4-year average price signals in 2002.

[28] Using data from Carlson et al. (2018)

[29] This applies to very few grid cells of our sample because few certifications were issued in the first years of the RSPO, from 2009 to 2014.

tion bias, as we control in our regressions precisely for the criterion behind it: the number of reachable mills.

Finally, our quasi-Poisson estimation procedure (see Section 4.1) removes observations from clusters in the fixed-effect dimensions that have a constantly null outcome (i.e. no deforestation during our study period). This naturally mitigates zero-inflation. Grey shapes in Figure 1 represent the area covered by our estimation sample.

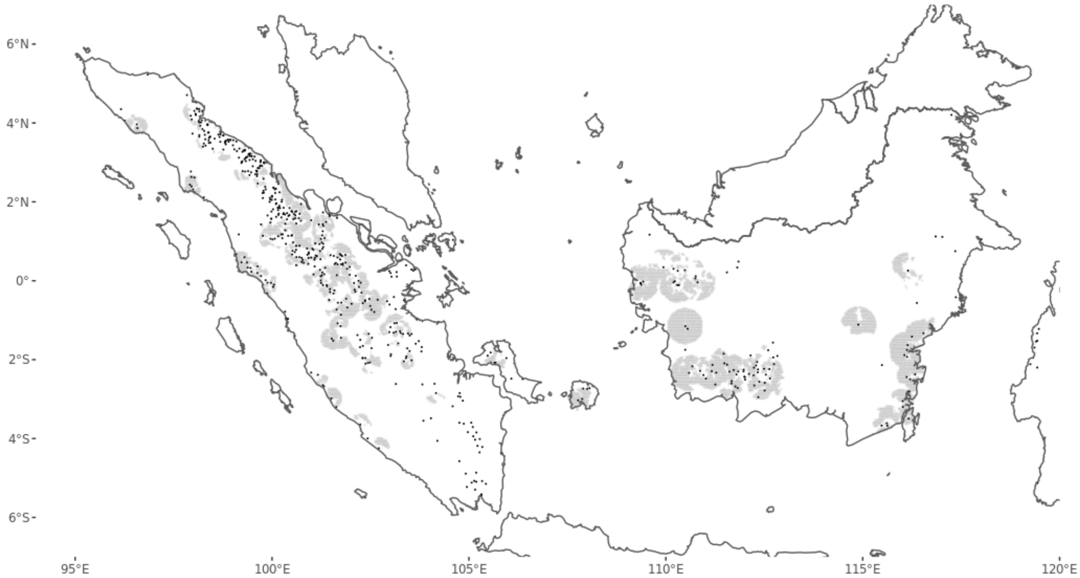

Figure 1: Study samples of palm oil mills and plantations

NOTE. This figure maps the samples of palm oil mills (dark dots) and plantations (light grey area) used in this study. The geographical area includes the Indonesian regions of Sumatra (left) and Kalimantan (right).

Table 2: Estimation sample - descriptive statistics

	Without missing values			**With missing values**			**t test**	**KS test**
	# grid cells = 12687 # grid cell-year = 71926			# grid cells = 22570 # grid cell-year = 215667				
	mean	std.dev.	median [min; max]	mean	std.dev.	median [min; max]	p-value	p-value
Deforestation (ha)	5.17	29.6	0 [0; 847.5]	5.18	30.95	0 [0; 903.1]	0.984	0.000
Price signal ($/tCPO)	672.9	92.46	671.7 [349.8; 926.4]	673.9	92.42	673.2 [349.8; 926.4]	0.031	0.276
# reachable mills	7.79	5.12	7 [1; 37]	5.72	4.3	5 [1; 37]	0.000	0.000

NOTE. This table shows descriptive statistics of the variables used in our main regression, for the sample of plantation sites (3x3km grid cells) actually used in estimations (without missing values), and the same sample but without removing observations with missing values. # means "number of". The two right-most columns show p-values of Welch two-sided t-tests, where the null hypothesis is that the true difference in means between the two groups is null, and the groups' variances are not assumed to be equal; and p-values of Kolmogorov-Smirnov tests where the null hypothesis is that the variables in the two groups are drawn from the same continuous distribution.

3.6 Descriptive statistics

Table 2 provides descriptive statistics for the final sample used in the main estimation (in Table A.1, we provide mill-level descriptive statistics). Deforestation is a count of pixel-level events of primary forest loss eventually (by 2015) replaced by oil palms. In Table 2, it is converted to hectares for more readability. The average deforestation annually observed is 5ha and the maximum is 847ha or almost the whole 900ha grid cell area. Price signal is the plantation-level inverse-distance-to-mill weighted average of CPO prices at reachable mills, averaged over the 4 previous years, as detailed above. In our estimation sample, it averages to 673 2010-constant USD per tonne CPO. The number of reachable mills is the annual count of known palm oil mills (as from the UML) within a 30km (50km in Kalimantan) catchment radius from a plantation. It ranges from 1 to 37, and half of the observations can reach more than 7 mills.

In Tables A.2 and A.3, we break down these descriptive statistics across the sub-categories of industrial, smallholder, legal, and illegal deforestation. We note two particular patterns. First, in industrial plantations, legal deforestation is, on average, twice larger than illegal deforestation, while in smallholder plantations this is the opposite. Second, illegal deforestation in both industrial and smallholder plantations is exposed to slightly higher price signals.

4 Estimation and identification

4.1 Estimation strategy

Here, we present the assumptions associated with our reduced form, and the framework to estimate it.

In addition to price signals, plantations use information related to investment costs (e.g., of land acquisition and conversion), operating costs (e.g., of labor, energy and fertilizers), institutional costs (either fixed or marginal, positive or negative, formal or not) and attainable yields. Plantations also take into account the expected relative costs and benefits of the alternative land uses. Hardly observable parameters, such as the plantations' own discount rates and abilities to form expectations, are also at play in the decision rule. We do not attempt to formally model how complete information sets determine deforestation decisions. Conceptually, this may all be summarized in a reduced-form relationship that is not necessarily linear, between deforestation on the left-hand side, and the true price signal perceived by the representative plantation on the right-hand side. This reduced form has a structural error term that includes all the information elements mentioned above. We approximate it as follows:

$$Deforestation_{idt} = exp(\alpha ln(Price_{idt}) + \beta X_{idt} + \lambda_{id} + \gamma_{dt} + e_{idt}) \quad (3)$$

From 2002 to 2014 ($t = 1, ..., 13$), we observe $Deforestation_{idt}$, which is the sum of pixel-level deforestation events in plantation site i in district d in year t (see Section 3.3). $Price_{idt}$ is a measure of the price signal (constructed with equations 1 and 2) observed by plantation i in district d in year t. α is the price elasticity of deforestation. X_{idt} is a vector of other observed determinants of deforestation that vary both locally and annually. In our main specification, X_{idt} comprises the number of known reachable mills. The unobserved determinants of deforestation can be decomposed as a sum of heterogeneity sources that can be either fixed attributes of plantations (local fixed effects, λid), or annual shocks common to a whole district (district-year fixed effects, γ_{dt}), or error idiosyncratic terms, e_{idt}. We elaborate on observed and unobserved heterogeneity in the next subsection.

$Deforestation_{idt}$ is a count of non-negative integers that may be null for a significant proportion of observations. Effectively, Table 2 shows that it is positively skewed, with a substantial amount of zero values. Therefore, we estimate Equation 3 as an exponential mean model, by Poisson Quasi maximum likelihood (Wooldridge 1999) (see Appendix E.1 for more details). The quasi-Poisson distribution imposes weaker assumptions on the data, as it only requires the mean (and not the variance) to be correctly specified. It is also more robust to distributional assumptions than negative binomial models for count data (Wooldridge 2002), and more appropriate than the inverse hyperbolic sine transformation given the small values of the measured land use change responses for many observations (Bellemare and Wichman 2020). Finally, unlike for zero-inflated Poisson models, the available algorithms to fit quasi-Poisson models accommodate fixed effects well.

Inference. We do not assume that annual records of price signals are independent and identically distributed. Rather, we allow arbitrary correlations within clusters of observations. Abadie et al. (2022) explain that clusters should be set at the level the treatment is randomly assigned. In our case, as we do not use experimental data, identifying the proper clustering level is not straightforward. As explained in more detail in Section 4.2, our treatment assignment mechanism is the interaction between distances to reachable mills and conditionally independent mill-gate price shocks. Its randomness grounds on the simultaneous variation in both dimensions. When, across some observations, there is no variation in one of these dimensions, such observations should be counted as one cluster and not as random draws with respect to each other. Consider plantation sites around a single mill, over several periods of time. Across these observations, the price signal varies in only one dimension (the mill-gate price shock, over time). This is also the case of repeated observations of a plantation site over a time period when the same set of mills is reachable. To count such observations as a single random draw, we cluster standard errors at the level of the set of reachable mills.[30] Note that this is a conservative choice, because for the many plantations that have the same set of *several* reachable mills, the treatment assignment is random (their relative distances to these mills differ in a way that is conditionally independent to the price shocks at these mills.)

4.2 Identification strategy

The causal interpretation of the observed correlation between prices and deforestation - identification of the price elasticity - is threatened by reverse causality, omitted variable bias and measurement error. Reverse causality can arise, for instance, if deforestation increases the palm oil supply, or expectations about it, pushing prices downwards. This is even more likely in the presence of spatial auto-correlation in deforestation. A non-included variable could also drive both prices and deforestation, biasing the causal interpretation of the observed correlation. In particular, this could be one of the already identified drivers of deforestation: agro-climatic suitability (Byerlee et al. 2016); the proximity to existing plantations (Gunarso et al. 2013; Shevade

[30] Such that a plantation that can reach mills A and B is not in the same cluster as a plantation that can reach mills A, B and C. We do not use Conley standard errors that do not account for auto-correlation over time.

and Loboda 2019) and to roads (Hughes 2018); the decentralization of authority on land (The Gecko Project[31] and Burgess et al. (2012)); and local political cycles opening up land and creating new infrastructure (Cisneros et al. 2021). Finally, measurement error, random or systematic, may also lead to spurious causal conclusions. It is possible, for instance, that the international price is a more precise measure of the true price incentive for plantations integrated in large companies, which also demonstrate systematically different deforestation patterns.

Absent a sharp experimental framework, we reflect on the observational variation at hand to make explicit under what assumptions we can avoid those threats to causal identification. The variation in price signal, our treatment variable, arises from the interaction of two variation sources. The first one is the spatial distribution of mills and plantations - i.e., the differences between plantation sites in their relative distances to reachable mills. The second source of variation is the differential mill-gate CPO prices. This interaction relates to a shift-share reduced-form design. Here, the shift-share "instrument" is the price signal and its reduced-form coefficient in Equation 3 is thus directly interpretable as the price elasticity. For this coefficient to be identified in our data it is sufficient to assume the conditional exogeneity of the mill-gate CPO prices (the shifts) with respect to plantation-mill distance weights (the shares) and with deforestation in plantation sites (Borusyak et al. 2020). In what follows, we clarify on what conditions and arguments this identification assumption relies.

First, we include district-year fixed effects in the main specification (equation 3), effectively removing variation common to observations in the same year and district. The remaining variation in price signals can be interpreted as yearly departures from the district market price. We quantify the within district-year standard deviation of mill-gate CPO prices as USD 138/tonne. This confirms that mills, even in the same district, idiosyncratically depart from a unique market price, and so do price signals across plantation sites in a given year and district. These price shocks are independent from macro-determinants of prices on markets larger than a district, and that could spuriously correlate with macro-level drivers of deforestation.[32] Because these shocks are departures down to annual *district* market prices, they are further independent from district-specific dynamics. This is crucial in our analysis, because districts are powerful jurisdictions in the administration of land in Indonesia and the control over land can unlock substantial revenues from natural resources. Therefore, political cycles at the district level can explain much of deforestation and of prices through general equilibrium effects on the district markets for, in particular, land, labor and energy.[33] In Figure B.1, we show estimates under alternative time fixed effects and thus with different market price departure interpretations.

We further impose that our main regression compares price signal shocks across plantation sites that sell to the same mills (i.e. that share the same set of reachable mills). This second fixed effect isolates our identifying variation from systematic differences in local determinants of price departures that would endogenously correlate with deforestation. Such differences at

[31] `https://thegeckoproject.org/`

[32] For instance, large-scale meteorological events like El Niño, affect both annual agronomic palm conditions and international prices of palm oil and substitutable crops, in turn affecting the Indonesian market prices (Rahman et al. 2013; Sanders et al. 2014; Santeramo and Searle 2019).

[33] Indeed, district splits (Burgess et al. 2012) and competition for election as district head (Cisneros et al. 2021) have been shown to be determinants of deforestation.

this level include in particular: i) agro-climatic conditions (typically common to large areas and constant in time); ii) average distances to downstream buyers (capturing the time-constant part of transport costs, as well as the intensity of monitoring by law or civil society, hence proxying institutional costs). iii) the number of reachable mills (continuously capturing differences between frontier and mature markets, including local infrastructures and hence a time-varying part of transport costs).[34]

As the sets of reachable mills are made only of the mills we have geo-localized, we complement this fixed effect with a control on the number of all known reachable mills.[35]

The two fixed effect dimensions (the district-year and the set of reachable mills) ensure that macro to local, time varying, notoriously endogenous, heterogeneity be removed from our shift-share identifying variation. Thanks to the granularity in our data, enough variation remains available for identification. Yet, it may be argued that some of this hyper-local heterogeneity in CPO price signals may still be potentially endogenous to deforestation. For instance, changing local palm fruit supply or milling marginal costs could affect mill-gate price departures and bias our estimate through reverse causality or omitted variable bias. We argue that these remaining threats to identification can be ruled out, given the nature of the market into which Indonesian mills sold CPO during our study period.

First, the quality of CPO cannot explain these differences, since CPO is a highly standardized good (Byerlee et al. 2016). Put differently, there is a single quality market for CPO.

Second, on this market, palm oil mills are price-takers, given the number of palm oil mills in Indonesia (more than a thousand) and the relatively low number of downstream actors (less than a hundred refineries and exporters). Pirard et al. (2020) show this pyramidal shape of the palm oil value chain. Therefore, mills cannot manipulate the quantities they sold to receive higher prices for CPO.

Third, the Indonesian CPO market is highly profitable and demand-driven, i.e. it is not at an equilibrium where the marginal cost equals the price for all market participants. Therefore, mills achieving lower marginal costs do not sell at more competitive prices. See Byerlee et al. (2016) on this, and in particular their Box 2.1. for a mill-level analysis of profitability.

Finally, the differential prices of CPO at mills' gates can be explained by shocks in the costs of transport to particular refiners or exporters, or by shocks in particular markets mills sell to, including in off-take agreements with specific buyers. As shown in Pirard et al. (2020), there is little integration between mills and refineries, and the hourglass-shaped industry is centered

[34]On a more technical note, this fixed effect is at the level where the distance weights (the shares) are constant in time, which is sound in a shift-share design where the shares vary in time (Borusyak et al. 2020). In other words, it is perfectly collinear to the shares, thus controlling for this part of the shift-share interaction (the district-year fixed effect absorbs the other part). This would not be the case of plantation-level fixed effects. Yet, in Figure B.1 we show that the main estimate is barely changed with plantation and district-year fixed effects.

[35]This mill-density control may be a collider (so-called "bad control"), if local market development and deforestation have a common cause (like past deforestation) and if local market development *results* from local prices. We argue that variations in medium-run price signal shocks are unlikely to influence the timing and location of a multi-million dollar investment in a new mill (Hsiao 2022). Moreover, adding a control for past deforestation in the last 4 years to attenuate the endogeneity of mill density and deforestation does not change the results (see Appendix C).

around refining and exporting groups. Monopsonic power on the CPO market does surely not affect all sellers (mills) homogeneously in the cross-section and in time. The ties between operating companies are mostly hidden from the public as of now, thus preventing us from modelling this source of variation in mill-gate prices. However, it is sufficient here that these downstream shocks in price signal departures be exogenous to hyper-local changes in deforestation (within plantations reaching the same mills at the broadest).

Hence, our reduced-form shift-share, fixed-effects estimator identifies the price elasticity under the assumptions that the four points above are valid. In the remainder of this section, we briefly discuss the robustness of our identification to relaxing some of these assumptions, and we add a note on measurement error.

Reverse causality through mills' market power or marginal costs. In our setting, reverse causality could arise and bias our estimates, if deforestation affected prices. It is conceivable, indeed, that deforestation leads to increased production and hence affects prices. As explained above, our price signal variable is a 4-year average. In addition, there is a 4-year time lapse between planting and first harvests and oil palm trees are not always planted immediately after forest clearing. Thus, there is a long time lag between the moment we measure the price signal and the moment current deforestation would affect prices. This is a first argument against the presence of reverse causality. However, as demonstrated in Bellemare et al. (2017), such an argument relies on the assumption that there are no dynamics in the confounders. In our case, deforestation may, indeed, be correlated over time. Past deforestation may affect current deforestation, while also affecting past and current mill-gate prices. If mills actually have some market power, their increased production can allow them to sell at higher prices. If, on the other hand, the CPO market is competitive and at equilibrium, and there are economies of scale in the production of CPO, increased production can lead mills to sell at lower prices. To check the robustness of our results to relaxing any of these assumptions, we alternatively control for 4-year lagged deforestation, and for 4-year lagged deforestation in the 8 neighboring plantation sites. Both specifications yield a similar estimate to the main one. In Appendix C, we discuss these robustness checks, in more detail.
Finally, we note that the fixed effect on the set of reachable mills and the control for the number of all known mills introduced above capture a large share of the variation in independent plantations' bargaining power (higher in high mill-density areas (Maryadi et al. 2004; Masliani et al. 2014)). Thus, were the assumption on marginal costs to be relaxed, our estimates are rather protected from a bias due to a pass-through from lower fruit prices (lower marginal costs) to lower CPO prices.[36]

Measurement error. We believe that our data and estimation strategy enable us to get the most accurate measure of the true price incentives privately observed by oil palm plantations in Indonesia to date. However, some measurement error remains. Here are its main sources:

[36]We do not attempt to control for FFB inputs or price signals in our main specification, because conditioning on these variations risks to introduce endogeneity as they are at least partly caused by CPO prices, i.e. they could be "bad controls".

First, we observe the annual mean unitary values and not the prices that mills publicly disclose (at a higher frequency than annually). Second, we can only model the price signal that reaches individual plantations (cf. Section 4.1). Third, our sample of geo-localized IBS mills does not cover the whole population. Therefore, in areas with mills both from and not from our sample, our measure of the price signal is incomplete. We do not suspect any of these to be prone to systematic measurement error. In particular, Table A.1 shows that there is no systematic difference between the IBS mills we have geo-localized and the others. Finally, any systematic measurement error in price signal between legal and illegal, or industrial and smallholder plantations, or other time-invariant distinction is absorbed by unit fixed effects.

5 Results

In this section, we first present and discuss our main results: estimates of price elasticities of deforestation in different segments of the Indonesian palm oil sector. Then, we estimate the effects of fruit prices and short-run prices, to provide insights into vertical integration and the formation of expectations in the oil palm sector. Finally, we discuss the external validity of our results and calculate back-of-the-envelope scaled-up counterfactual effects.

5.1 Price elasticities of deforestation in Indonesian oil palm plantations

Table 3 shows our estimates of the price elasticity of deforestation for different kinds of oil palm plantations in Indonesia. The right-most column in Table 3 features our overall 1.5 price elasticity of deforestation, showing that, pooling together all kinds of plantations, deforestation due to oil palms does react positively to price signals. In Appendix C, we conduct a robustness analysis on that estimate, summarized in the specification chart presented in Figure B.1. The next paragraphs, and Tables 3 and A.6, document which subgroups of the Indonesian plantation sector contribute to making this estimate lower and/or less precise, and which do not. Indeed, the magnitudes and precision of price elasticity estimates are heterogeneous over the different segments of deforestation.[37]

Table 3: Price elasticities of deforestation across Indonesian oil palm plantations

	Industrial plantations			Smallholder plantations			All		
	Legal	Illegal	All	Legal	Illegal	All	Legal	Illegal	All
Estimate	0.53	5.3	1.79	1.32	1.79	1.37	0.33	3.09	1.5
95% CI	[-1.13; 2.19]	[2.19; 8.41]	[0.27; 3.31]	[-1.53; 4.17]	[0.62; 2.96]	[0.18; 2.57]	[-1.07; 1.73]	[1.43; 4.75]	[0.37; 2.62]
Observations	24131	17091	65368	5885	5704	20721	26079	20695	71926
Clusters	629	451	1143	203	276	529	738	640	1441

NOTE. This table shows our main estimates of the price elasticity of deforestation. They are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signals. The price signal is measured as the 4-year average of annual inverse-distance weighted averages of crude palm oil prices at the gates of reachable mills. Deforestation is measured as primary forest loss eventually replaced with oil palm plantations. We differentiate industrial from smallholder plantations based on scale and landscape criteria (Austin et al. 2017; Petersen et al. 2016). We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. Sample observations are annual records of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014. They all have a positive extent of remaining primary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.

[37] In Table A.9, we also present effects of interactions between the price signal and local market development. It appears that the price elasticity of deforestation does not substantially depend on this covariate.

Industrial and smallholder plantations. Breaking down the estimation into plantation types, we find that to a 1% increase in price signals, industrial and smallholder plantations react with a 1.8% and a 1.4% increase in average deforestation, respectively.[38]

This positive price elasticity of deforestation in industrial plantations indicates that corporate actors of the oil palm sector engage in large-scale deforestation where and when prices are higher than usual. This suggests that medium-run price signals (over 4 years here) do influence large long-term investments, typically over more than a decade. In the next subsection, we disentangle annual price variations to provide more insights into the dynamics of price signals.

The positive price elasticity of deforestation we estimate in smallholder plantations indicates that smaller plantations, organized in mosaic landscapes with other land uses, encroach on forests when prices are higher than usual. This responsiveness to crude palm oil prices suggests that it is actually mill owners - most usually companies - that decide upon the timing and location of smallholder plantation expansion. In the next subsection, we differentiate the effect of palm fruit prices to provide more insights into this direction.

Deforestation in industrial plantations seems more price-elastic than in smallholder plantations. The difference is especially pronounced in the case of illegal deforestation, where the point estimate for industrial plantations is more than twice as large as for smallholders, and this difference is significant at the 95% confidence level (see Table A.10). Hence, industrial plantations seem more reactive than smallholders in illegally encroaching on forests when prices increase.

Legal and illegal deforestation. We further break down the estimation according to the legal status of deforestation.[39] We find close to zero effects of price signals on legal deforestation, irrespective of the plantation type. On the other hand, illegal deforestation appears to be price elastic in every plantation type. Overall, the price elasticity of illegal deforestation is 3. Industrial and smallholder plantations react to a 1% increase in price signals by illegally deforesting 5.3% and 1.8% more respectively.

The positive price elasticity we estimate for illegal deforestation indicates that economic opportunities encourage plantations to circumvent land use regulations. On the other hand, we estimate that legal deforestation is not price elastic. This may come from a lack of statistical power to detect a true positive price elasticity, or to legal deforestation being truly inelastic to prices. Given the magnitude of the estimate (0.3 across plantation types) and the number of observations and clusters (sets of reachable mills), we believe that it is rather truly inelastic to prices. This is most likely the consequence of the long processes necessary to acquire a plantation license (involving, for example, measuring environmental suitability and community consultation; see Paoli et al. (2013) for more detail on the licensing process). If obtaining the legal green lights to clear the forest and plant palm trees takes several years, (in addition to the

[38] As detailed in Appendix D, the distinction between industrial and smallholder plantations is based on the landscape and size differences between plantations mapped by Austin et al. (2017) and the mid and small-sized plantations mapped by Petersen et al. (2016).

[39] Cf. Section 3 for definitions.

lag between planting and harvesting), it is not surprising that medium-run price signals do not influence legal deforestation. Plantations probably rely on more stable signals than those that we capture in this study to formulate long-term expectations about the profitability of engaging today in legal deforestation.

Table A.10 documents that the differences in the price elasticities of legal and illegal deforestation are statistically significant (except for smallholders). The estimated price elasticity of illegal deforestation is of larger magnitude than for all deforestation (legal, illegal and unknown combined). This is true for any plantation type (industrial, smallholder, or both). This is especially pronounced for industrial plantations, where the price elasticity point estimate is more than twice as large for illegal deforestation. Altogether, these findings about legal and illegal deforestation indicate that, across plantation types, positive price elasticity is driven by illegal deforestation.

Immediate and transitional deforestation. As explained in more detail in Section 3, we observe both the moments of forest loss and of planting and, for industrial plantations only, we can calculate time lags between the two. We consider deforestation to be transitional if between 5 and 12 years elapse between forest loss and plantation development. Table 4 shows our estimates of the price elasticity of immediate and transitional deforestation, again distinguishing legal, illegal, and overall (legal, illegal and unknown legal status) deforestation. We show results for immediate conversion over the whole period and over 2002-2010, the subperiod when transitional conversion can be observed. Over the whole period, the price elasticities of immediate deforestation are not substantially different from the main results. Over the 2002-2010 period, looking at immediate conversion only, we estimate economically and statistically significant price elasticities of -3.8 and 14.4 for legal and illegal deforestation respectively. In the same period, we find a price elasticity of 6.7 for transitional, illegal conversion. Irrespective of the legal status, in 2002-2010, the price elasticities of immediate and transitional conversion are imprecisely measured.

Table 4: Price elasticities of immediate and transitional deforestation in Indonesian industrial plantations

	2002 - 2014			2002 - 2010					
	Immediate conversion			Immediate conversion			Transitional conversion		
	Legal	Illegal	All	Legal	Illegal	All	Legal	Illegal	All
Estimate	1.32	6.34	2.41	-3.81	14.37	-1.6	-0.77	6.7	1.26
95% CI	[-0.67; 3.31]	[2.14; 10.54]	[0.53; 4.29]	[-6.61; -1.01]	[7.47; 21.28]	[-4.91; 1.7]	[-3.55; 2]	[2.86; 10.55]	[-1; 3.52]
Observations	21648	13885	59664	10954	5608	29749	13315	8528	36347
Clusters	583	411	1052	421	223	706	451	300	817

NOTE. This table shows our estimates of the price elasticity of deforestation in industrial oil palm plantations. They are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signals. The price signal is measured as the 4-year average of annual inverse-distance weighted averages of crude palm oil prices at the gates of reachable mills. Deforestation is measured as primary forest loss eventually replaced with oil palm plantations. We differentiate immediate from transitional conversion based on the time lapse between forest loss and plantation development. The cut-off point is 4 years. We can observe transitional conversion only up to 2010. We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. Sample observations are annual records of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014. They all have a positive extent of remaining primary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.

These results provide further understanding of the conversion process and how it reacts to price signals. Looking at immediate conversion only during the shorter and earlier 2002-2010 period, the substantially negative and positive price elasticities of legal and illegal deforestation indicate that if prices are high, plantations do not engage in the lengthy licensing process (thus reducing legal deforestation), but rather rush to benefit from medium-run economic opportunities (thus increasing illegal deforestation). However, some of this hurried illegal forest clearing does apparently not lead to immediate conversion, as indicated by the positive elasticity of illegal transitional conversion. Such delays (5 years or more) may be due to conflicts with local communities over land rights, or to legal proceedings.

Price variability. We here report on the effect of price variability, rather than level (but conditional on price level, as in Lundberg and Abman (2022)). We find that the effect of a standard deviation in annual price signal over the last 4 years on deforestation is generally not distinguishable from zero. (Table A.8). However, this hides notable patterns: price variability deters legal deforestation (in both industrial and smallholder plantations) and deforestation for industrial plantations in general.

Spatial heterogeneity. We estimate the price elasticities of deforestation for Sumatra and Kalimantan separately. Table A.6 shows that, in Kalimantan, there are fewer clusters (sets of reachable mills) and observations than in Sumatra, (especially for smallholders) and thus estimates are less precise. It is also possible that, in Kalimantan, we managed to geo-localize a lower share of the universe of palm oil mills, and thus suffer from more noise in the price signal variable, yielding downward biased estimates. Yet, it is also possible that, during our study period, deforestation in Kalimantan was driven by different dynamics than in Sumatra, and that prices were, indeed, less influential (with a relatively larger role played by political economy factors, for instance).

Deforestation in secondary forest. Finally, we estimate the main model on a measure of deforestation in secondary forest only. Table A.7 shows that deforestation in such forests is generally not price elastic. We see two non-exclusive potential explanations for this absence of effect.[40] First, secondary forest is, by definition of primary forest, more scattered (see Appendix D.2). Deforestation in secondary forest plots may be decided marginally, at too small a scale for price signals to be significantly influential. Second, deforestation in secondary forest can include rotations in existing tree plantations (non-industrial oil palms or others), that are not related to the price signals in our model.

5.2 Vertical integration and expectation formation

Here, we investigate how the medium-run crude palm oil price signal affects deforestation. We disentangle the price elasticity of deforestation in two dimensions: vertical integration and the time length of price signals. To do so, we use, in turn, two new variables: the medium-run palm fruit price signal and the short-run crude palm oil price signal. Each of them is arguably a

[40] Given the large number of observations and clusters, we do not attribute it to a lack of statistical power.

post-treatment variable in the sense that they do not affect, but are affected by, the treatment (the medium-run crude palm oil price signal). For the palm fruit price signal, this hinges on the assumption that mills have market power on their input (palm fruit) market, but not on their output (crude palm oil) market. The post-treatment status of the short-run price signal relies on the temporal causal argument that past prices affect current prices, but the reverse is false. Therefore, each of them can have an indirect effect on deforestation, whereby the medium-run crude palm oil price signal affects the post-treatment variable, which then affects deforestation. These post-treatment variables can also have a moderation effect on deforestation, whereby they affect the treatment effect. We estimate the partial effects of the post-treatment variables, unconditional and conditional on the treatment.[41] When conditional on the treatment, the partial effects of the post-treatment and of the treatment variables exclude the indirect effect. In any case, the partial effects include the moderated and the unmoderated effects (see Appendix E.2 for more detail). We report the moderation effects as partial effects of terms of interactions between the post-treatment variable and the treatment.

Vertical integration: palm fruit and crude palm oil price signals. In this study, our main measure of price signals uses crude palm oil prices (see Section 4.1). Palm tree fruits, commonly called fresh fruit bunches (FFB), are sold by independent plantations to mills. The effect of palm fruit price signals on deforestation may thus document the price elasticity of less vertically integrated plantations. Table 5 shows our estimates of palm fruit and crude palm oil price elasticities, along with the partial effects of their interactions on deforestation. Table 5 also displays the partial effects of palm fruit price signals unconditional on the effect of crude palm oil prices. All models are based on the same specifications as the main one, from Equation 3. Estimates in this exercise should be taken with caution, because arguments critical to our identification strategy apply to the crude palm oil market but not to the palm fruit market. Hence, estimates from Table 5 should more cautiously be seen as descriptive rather than causal.
Palm fruit price signals seem to influence deforestation, but in opposite directions in industrial and smallholder plantations. In industrial plantations, a palm fruit price increase of 1% causes an increase in average deforestation of 1.6%. On the other hand, in smallholder plantations, it causes a decrease in average deforestation of 2%. Over all plantation types, these effects balance to a positive price elasticity (1.1). This pattern is similar whether conditional or not on crude palm oil prices. For any plantation type, the effect of crude palm oil prices vanishes once the effect of palm fruit price signals on deforestation is taken into account. The interaction partial

[41]The causal interpretation of all these partial effects relies on the same identification strategy as presented in Section 4.2: reachable-mills and district-year fixed effects plus a control on the number of all known reachable mills. In particular, it relies on the assumption that fixed-effects and the control rule out post-treatment confounders that would affect both the post-treatment variable and deforestation. This assumption may be stronger in the case of palm fruit price signals than short-run price signals. Under these assumptions, the conditional partial effects of the post-treatment and treatment variables can be interpreted as net of the indirect effect. The partial effect of the treatment variables, unconditional on the post-treatment variables presented in the previous section, can be interpreted as total effects. Thus, the difference with conditional partial effects presented here documents indirect effects.

effect on deforestation is positive. This means that the effect of crude palm oil price signals on deforestation increases with an increase in palm fruit price signals (and vice-versa).

In industrial plantations, the bulk of the effect of crude palm oil price signals on deforestation (as estimated in our main analysis, see Table 3) is actually attributable to the mechanism of local crude palm oil prices influencing local palm fruit prices which, in turn, affect deforestation decisions. This suggests that deforestation in industrial plantations occurs mainly in independent plantations - presumably as a result of the low vertical integration, even in the downstream part of the sector (Pirard et al. 2020). The positive interaction effect indicates that palm fruit price elasticity is even larger where crude palm oil price signals are high. This suggests that higher crude palm oil prices reinforce expectations about high palm fruit prices and, hence, motivate deforestation.

In smallholder plantations, our results indicate that deforestation increases in times and places of low palm fruit prices but high crude palm oil prices. This suggests that it is the companies owning the mills, wishing to benefit from higher output/input price ratios, that decide upon the timing and location of smallholder plantations. This negative elasticity indicates that the expansion of independent smallholders (that was predominant relative to that of supported smallholders since the 2000s) onto forests is actually driven by mill-level decisions.

Table 5: Palm fruit and crude palm oil price elasticities of deforestation across Indonesian oil palm plantations

	Industrial plantations		Smallholder plantations		All plantations	
FFB price signal						
Estimate	1.65	2.55	-2.02	-1.56	1.15	1.79
95% CI	[0.31; 2.99]	[0.71; 4.4]	[-3.61; -0.43]	[-3.01; -0.12]	[0.07; 2.24]	[0.49; 3.1]
CPO price signal						
Estimate		0.5		1.4		0.68
95% CI		[-1.98; 2.99]		[-0.43; 3.23]		[-0.97; 2.33]
Interaction						
Estimate		0.12		0.03		0.07
95% CI		[0.03; 0.21]		[-0.05; 0.1]		[0.01; 0.13]
Observations	60077	44379	18207	15880	66191	49829
Clusters	1040	993	501	484	1334	1281

NOTE. This table shows our estimates of the palm fruit and crude palm oil price elasticity of deforestation. They are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signals. The price signal is measured as the 4-year average of annual inverse-distance weighted averages of either palm fruit or crude palm oil prices at the gates of reachable mills. The last block of rows shows estimates of the partial effects of the interaction of both, evaluated at the sample mean. Deforestation is measured as primary forest loss eventually replaced with oil palm plantations. We differentiate industrial from smallholder plantations based on scale and landscape criteria (Austin et al. 2017; Petersen et al. 2016). We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. Sample observations are annual records of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014. They all have a positive extent of remaining primary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.

Expectation formation: short- and medium-run price signals. The short-run price signal is the inverse-distance weighted average of prices at the gate of reachable mills the year deforestation occurs. In this part, we consider an alternative medium-run price signal that runs over the preceding 3 years, but does not include the contemporaneous price signal. Thus, the average of these short- and medium-run price signals corresponds to our main measure of price signal (which we also called medium-run). (see Equation 2). Table 6 shows the estimates of short- and medium-run price elasticities, along with the partial effects of their interactions on deforestation. Table 6 also displays the partial effects of short-run price signals alone - i.e., not conditional on the effect of medium-run price signals. All models are based on the same specifications as the main one, from Equation 3.

Short-run price signals alone do not explain deforestation. However, once medium-run price signals are included in the model, the partial effects in the short-run increase substantially (except for smallholders). The interaction partial effect on deforestation is positive. This means that the effect of medium-run price signals on deforestation increases with an increase in short-run price signals (and vice-versa).

In Table A.5, we show elasticities to annual price signals, conditional on the other annual price signal lags. This reveals that plantations are most clearly sensitive to the price signals in the two most recent years.

These results may reflect the fact that more recent developments in prices weigh more on expectations and hence on deforestation decisions than older prices. Moreover, it seems that short-run prices influence deforestation only when longer variations are also accounted for. This is at least partly due to the positive moderating effect of short-run price signals on medium-run ones. Together, these results suggest that plantations, to form distant expectations on the profitability of their perennial and yield-lagging crop, look at short-run price signals only to confirm medium-run dynamics. It is also notable that this pattern comes from industrial plantations, and that it is reversed in smallholder plantations. Indeed, among smallholders, the total price signal effect seems to be driven by medium-run variations. One hypothetical explanation for this difference is that, in times of short-run price spikes, companies prioritize deforestation for industrial plantations, and then allocate forest land to smallholder plantation development.

Table 6: Short-run and medium-run price elasticities of deforestation across Indonesian oil palm plantations

	Industrial plantations		Smallholder plantations		All plantations	
Short-run price signal						
Estimate	0.35	1.072	0.15	0.345	0.31	0.773
95% CI	[-0.05; 0.74]	[0.571; 1.573]	[-0.28; 0.59]	[-0.224; 0.914]	[-0.01; 0.64]	[0.379; 1.168]
Medium-run price signal						
Estimate		0.751		1.05		0.77
95% CI		[-0.469; 1.971]		[0.16; 1.94]		[-0.131; 1.671]
Interaction						
Estimate		0.029		0.035		0.027
95% CI		[0.002; 0.056]		[-0.002; 0.072]		[0.008; 0.047]
Observations	140405	65368	44393	20721	152099	71926
Clusters	1430	1143	660	529	1779	1441

NOTE. This table shows our estimates of the short- and medium-run price elasticity of deforestation. They are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signals. The short-run price signal is measured as the inverse-distance weighted average of crude palm oil prices at the gates of reachable mills. The medium-run price signal is the 4-year average of short-run price signals. The last block of rows shows estimates of the partial effects of the interaction of both, evaluated at the sample mean. Deforestation is measured as primary forest loss eventually replaced with oil palm plantations. We differentiate industrial from smallholder plantations based on scale and landscape criteria (Austin et al. 2017; Petersen et al. 2016). We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. Sample observations are annual records of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014. They all have a positive extent of remaining primary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.

5.3 Back-of-the-envelope scaled-up counterfactuals

In this subsection, we attempt to give a sense of the magnitudes that are implied by our estimated 1.5 micro-level price elasticity of deforestation. First, we discuss the external validity of our results. In light of this, we then describe how we scale up average partial effects. Finally, we present and discuss scaled effects of counterfactual price changes.

External validity. Given the specific organisation of the palm oil sector in Indonesia, our results cannot automatically be extrapolated to other crops or countries. Even within Indonesia, given the differences between Sumatra and Kalimantan observed in this study, one should be cautious in extrapolating our results to specific regions like the new deforestation frontier in Papua. However, as the regions in our analysis include most existing Indonesian oil palm plantations and deforestation, we are confident in claiming external validity with respect to the country as a whole. Extending our conclusions in time should also be done with caution, since our study does not cover recent developments in oil palm-related policies, such as the biofuel mandates, (Kharina et al. 2016) or the *No Deforestation, No Peat, No Exploitation* commitments from the private sector (Pirard et al. 2015). We believe that, although our sample is restricted to plantations within 30km (50km in Kalimantan) from at least one mill (to avoid introducing too much noise into our sample), the results can be extrapolated to plantations located even further away. This is supported by our finding that price elasticity is not contingent on our measure of remoteness - the number of reachable mills (Table A.9). Finally, we note that our estimates mainly capture effects on deforestation at the intensive margin, i.e., occurring after at least one mill opened.[42]

Scaling factor. To scale up our estimated average price effects to the whole country of Indonesia, we count the number of individual plantation sites (grid cells) where deforestation is possible in Sumatra and Kalimantan. Hence, we first count grid cells that are within 82km of at least one known (as from the UML) palm oil mill. This follows Heilmayr et al. (2020), who analyzed from RSPO audit reports that 99% of mills' supply bases were within this straight line distance. Because, in this area, many plantation sites are actually unlikely to experience deforestation (either because there is no forest or because of unsuitability to oil palms), we excluded those that never experienced any deforestation from 2002 to 2014 (which is probably conservative regarding the total extent of primary forest where oil palm can be grown in the country). Note that, for the sake of simplicity, we count in the scaling area the plantation sites where deforestation occurred before the first mill opened in the catchment radius - i.e., at the extensive margin. Finally, we aggregate our results over 11396 3x3km plantation sites in Sumatra and Kalimantan. We assume that this population of plantation sites has the same average deforestation as in our sample. Under this assumption, we multiply by the scaling factor to estimate a baseline total deforestation of 132835ha.

Counterfactual effects. Table 7 shows the aggregated annual effects of different counterfactual CPO price changes on deforestation in Indonesia. For different price changes, we quantify the relative change in average deforestation, the scaled effect on deforestation, and the corresponding potential revenue from a CO2 payment. The effect is scaled based on the aggregation factor presented above. We estimate corresponding carbon pricing revenues from a potential result-based payment for reducing emissions from deforestation. We apply an average of $638\ tCO2\ ha^{-1}$ emissions due to deforestation (Guillaume et al. 2018).[43] CO2 revenues are based on the $5/tCO2 agreed price Norway paid to Indonesia for its recently avoided deforestation .[44]
Hence, given a 1.5 price elasticity of deforestation, we estimate that average variations (+5%) in CPO price signals incentivize Indonesian oil palm plantations to clear 10.1kha of primary forest annually.[45] In the presence of a result-based payment scheme, this represents a yearly opportunity cost of M$32.3. To curb annual deforestation 29% below the 2002-2014 average with

[42]Mills need a minimal fruit supply basis to operate. At usual mill capacity and plantation yield, this implies a minimum plantation size of ca. 3000 hectares to be developed alongside any new mill opening (Paoli et al. 2013). Because of the lag between planting and harvesting, deforestation occurs before the mill starts operating. On this margin, deforestation occurs far from already operating mills, and thus local price signals do not exist.

[43]We apply the 44/12 C to CO2 conversion factor to their $174\ Mg\ C\ ha^{-1}$ lost in conversion of Sumatra rainforests into oil palm monocultures.

[44]`https://www.regjeringen.no/en/aktuelt/noreg-betaler-530-millionar-for-redusert-avskoging-i-indonesia/id2722135/`

[45]We compute standard deviations in our price signal regressor variable, in the estimating sample, after removing variations in fixed-effect dimensions (Mummolo and Peterson 2018).

price incentives alone, price signals for individual plantations should be lowered by 19%.[46] This would save 38.5kha of primary forest annually, corresponding to revenues from a potential result-based payment scheme of M$123.

Table 7: Counterfactual annual effects of different CPO price changes on deforestation in Indonesia

	+1 std. dev.	-1%	-19%
Relative change (%)	7.63	-1.49	-29
Total change (ha)	10137	-1978	-38522
Potential CO_2 revenues (M$)	-32.3	6.3	122.9

NOTE. This table shows scaled-up effects of counterfactual changes in crude palm oil (CPO) price signals. To compute total change effects, we apply relative changes to average deforestation from our main econometric model, with a scaling factor of 11396, equal to the number of 3x3km grid cells in Sumatra and Kalimantan within 82km to any known palm oil mill where deforestation occurred at least once between 2002 and 2014. Potential CO_2 revenues correspond to result-based payments paid at a price of $5 per tCO_2 avoided, assuming average emissions of 174tC per hectare deforested.

[46] Aligning annual deforestation reduction to Indonesian Paris Agreement targets, i.e., 29% GHG emission (including LUC) below business as usual by 2030 (GoI 2016). Note that this is a rather arbitrary target though, since a 29% reduction in primary forest loss does not necessarily yield a 29% reduction in GHG emissions.

6 Conclusion

In this study, we estimate different price elasticities of primary forest conversion to oil palm plantations in Indonesia. We find that medium-run crude palm oil price signals have an overall positive effect on deforestation in the Indonesian oil palm sector. The price elasticity is 1.5. Industrial, smallholder and illegal plantations are responsive to prices. On the other hand, price signals have no effects on legal deforestation. To conclude, we discuss some limitations the reader should be aware of, we present the policy relevance of our results, and propose further research avenues.

Study limitations. Our nation-wide estimates of the price elasticities of smallholders and illegal deforestation are, to the best of our knowledge, the first in the literature on oil palms. Yet, they necessarily rely on observational data that are still scarce and incomplete. This prevents us from ruling out some confounding threats. Notably, the concession data we use to identify legal and illegal deforestation are known not to be exhaustive (see Section D.2). The land zoning data are time-invariant and thus do not inform us about land releases. For these two reasons, we may identify too much illegal deforestation. This imprecision may bias our results if it is correlated locally with price signals and deforestation. For instance, a district jurisdiction could release forest estate land to oil palm production in some areas, impacting local palm oil prices there, as well as deforestation. This systematic measurement error would bias the overall estimate.

We also highlight that the external validity of our study may be limited by the exclusion of the extensive margin in our analysis, i.e., deforestation occurring where no mill is already operating. We would expect that such deforestation is less price elastic, because it depends more on other elements that determine the mill establishment, like capital availability, or the regional political economy and infrastructure.

Finally, we note that our definition of deforestation excludes the displacement of other land uses onto forests due to oil palm expansion influenced by prices.

Policy Relevance. Oil palm is a highly profitable crop in Indonesia, with large, suitable but forested, areas still undeveloped (Pirker et al. 2016). Moreover, installed processing capacities are far from saturated (Pirard et al. 2020). Thus, the ever growing demand and associated economic incentives pose the risk of a continued threat to the country's primary forest. The existing conservation schemes have limited effectiveness due to the prevalence of smallholders and illegal plantations. This study shows that these unregulated segments of the oil palm sector can be incentivized away from deforestation with a price instrument. We find that such an instrument would be most effective on illegal deforestation for industrial plantations. In addition, several parts of our results suggest that smallholder encroachment on forests is decided by mills. In this case, a market-based conservation scheme applied at the mill level, on CPO prices, would reach deforestation for smallholder plantations. However, complementary schemes would be necessary to protect smallholders from being one of mills' variables of adjustment to the price intervention.

Furthermore, our finding that legal deforestation is inelastic to prices suggests that legal deforestation does not react to medium-run market signals because of long licensing processes. On the other hand, we estimate a substantial price elasticity of illegal deforestation. This indicates the existence of strong incentives to circumvent land use regulations in order to seize economic opportunities for palm expansion. These two phenomena probably interact. More stringent conservation regulations may make the licensing process even longer and, in the absence of strong monitoring, encourage illegal deforestation in the presence of high price incentives. However, this leakage effect can be contained if price incentives are controlled. Hence, our results suggest that, in the context of weak monitoring, a market-based instrument may help regulatory instruments be more effective. We also find that price volatility deters legal deforestation, displacing deforestation for smallholders outside of concessions. Therefore, a market instrument that stabilizes prices can also increase the effectiveness of regulations.

A sector-wide tax on CPO, levied at palm oil mills and refunded against proof of sustainable production could be less dependent on local monitoring and hence reduce the risk that weak institutions hinder effective forest conservation intervention (Heine et al. 2020).[47] This would still require to implement a monitoring capacity that allows the tax authority to verify the claims from the mills. Yet, this could be more easily achieved when the cost of not implementing it is reversely borne by the mills. A tax on palm oil, although refunded to deforestation-free plantations, may cause companies to issue less land and employment, in particular, to rural communities. Therefore, to be fair, part of the revenues from such a conservation tax, and possibly external revenues from schemes like REDD+, should compensate the communities most reliant on palm oil for their development.

We estimate that a 19% tax on CPO could reduce average deforestation rate by 29%, which is approximately in line with the Indonesian Nationally Determined Contributions (NDC) to the Paris Agreement. There are several reasons why the actual effect of such a tax could depart from what we estimate though. First, refunding the producers that can testify sustainable palm oil production can strengthen the incentives provided by the tax beyond the price elasticity we estimate here based on price variations unrelated with sustainability commitments. Incentives could be further strengthened and targeted to specific objectives if all tax revenues were redistributed to sustainable producers. Second, general-equilibrium effects can modulate the effect of the tax in either direction. For instance, if the elasticity of demand for deforestation-based palm oil is low, producers could transfer the burden of the tax to consumers and thus be less affected on equilibrium. A tax on CPO only may also achieve lower emission reductions than estimated if forest is left vulnerable to uncontrolled production of another commodity. Third, a tax reducing leaked deforestation could encourage the development of new regulation that could further reduce deforestation more effectively.

Finally, our results seem to suggest that the price incentives provided by the Roundtable on Sustainable Palm Oil (RSPO) are insufficient to reach zero-deforestation palm oil. Indeed, the price premium offered by the RSPO is around 2% according to Levin (2012), and 7% according to Preusser (2015).

[47] Currently, an export tax applies to crude palm oil, and its revenues are meant to support rural development. This embeds no conservation incentives though.

Further research. We do not attempt in this paper to properly simulate policy effects on deforestation through prices. We do not model a separation between deforestation-free and deforestation-based markets (and prices) that is caused by a label or by downstream due diligence on sustainability. Hence, our study does not provide strong insights into the incentivizing scheme of the Roundtable on Sustainable Palm Oil (RSPO). We leave such efforts to future research.

We note that our new, spatially explicit, microeconomic panel dataset of palm oil mills could be useful to study the economic causes of other important phenomena in Indonesia, like land conflicts or intentional forest and peat fires. These data can also help further the understanding of the economics of palm oil mills, whose operations have remained a black box so far.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge (2022). "When Should You Adjust Standard Errors for Clustering?*". In: *The Quarterly Journal of Economics*, qjac038.

Ai, Chunrong and Edward C. Norton (2003). "Interaction Terms in Logit and Probit Models". In: *Economics Letters* 80.1, pp. 123–129.

Amiti, Mary and Jozef Konings (2007). "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia". In: *The American Economic Review* 97.5, pp. 1611–1638.

Araujo, Rafael, Francisco J M Costa, and Marcelo Sant'Anna (2021). *Efficient Forestation in the Brazilian Amazon: Evidence from a Dynamic Model*. Preprint. SocArXiv.

Austin, Kemen, A. Mosnier, J. Pirker, I. McCallum, S. Fritz, and P.S. Kasibhatla (2017). "Shifting Patterns of Oil Palm Driven Deforestation in Indonesia and Implications for Zero-Deforestation Commitments". In: *Land Use Policy* 69, pp. 41–48.

Baudoin, Alice, P.M. Bosc, C Bessou, and P Levang (2017). *Review of the Diversity of Palm Oil Production Systems in Indonesia: Case Study of Two Provinces: Riau and Jambi*. Center for International Forestry Research (CIFOR).

Bellavia, Andrea, Matteo Bottai, Andrea Discacciati, and Nicola Orsini (2015). "Adjusted Survival Curves with Multivariable Laplace Regression:" in: *Epidemiology* 26.2, e17–e18.

Bellemare, Marc F., Takaaki Masaki, and Thomas B. Pepinsky (2017). "Lagged Explanatory Variables and the Estimation of Causal Effect". In: *The Journal of Politics* 79.3, pp. 949–963.

Bellemare, Marc F. and Casey J. Wichman (2020). "Elasticities and the Inverse Hyperbolic Sine Transformation". In: *Oxford Bulletin of Economics and Statistics* 82.1, pp. 50–61.

Benedict, Jason Jon, Kimberly M. Carlson, Ramada Febrian, and Robert Heilmayr (2023). *Characteristics of Indonesian palm oil mills*. Harvard Dataverse.

Bergé, Laurent R (2018). "Efficient Estimation of Maximum Likelihood Models with Multiple Fixed-Effects: The R Package FENmlm". In: p. 39.

Berry, Steven T (2011). *Biofuels Policy and the Empirical Inputs to GTAP Models*, p. 29.

Borusyak, Kirill, PrincetonPeter Hull, and U ChicagoXavier Jaravel (2020). "Quasi-Experimental Shift-Share Research Designs". en. In: p. 26.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber (2012). "The Political Economy of Deforestation in the Tropics*". In: *The Quarterly Journal of Economics* 127.4, pp. 1707–1754.

Busch, J., R. N. Lubowski, F. Godoy, M. Steininger, A. A. Yusuf, Kemen Austin, J. Hewson, D. Juhn, M. Farid, and F. Boltz (2012). "Structuring Economic Incentives to Reduce Emissions from Deforestation within Indonesia". In: *Proceedings of the National Academy of Sciences* 109.4, pp. 1062–1067.

Busch, Jonah and Kalifi Ferretti-Gallon (2017). "What Drives Deforestation and What Stops It? A Meta-Analysis". In: *Review of Environmental Economics and Policy* 11.1, pp. 3–23.

Busch, Jonah, Kalifi Ferretti-Gallon, Jens Engelmann, Max Wright, Kemen Austin, Fred Stolle, Svetlana Turubanova, Peter V. Potapov, Belinda Margono, Matthew C. Hansen, and Alessandro Baccini (2015). "Reductions in Emissions from Deforestation from Indonesia's Moratorium on New Oil Palm, Timber, and Logging Concessions". In: *Proceedings of the National Academy of Sciences* 112.5, pp. 1328–1333.

Byerlee, Derek, P. Falcon Walter, and L. Naylor Rosamond (2016). *The Tropical Oil Crop Revolution, Food, Feed, Fuel, and Forests*. New York: Oxford University Press.

Cacho, Oscar J., Sarah Milne, Ricardo Gonzalez, and Luca Tacconi (2014). “Benefits and Costs of Deforestation by Smallholders: Implications for Forest Conservation and Climate Policy”. In: *Ecological Economics* 107, pp. 321–332.

Carlson, Kimberly M., Robert Heilmayr, Holly K. Gibbs, Praveen Noojipady, David N. Burns, Douglas C. Morton, Nathalie F. Walker, Gary D. Paoli, and Claire Kremen (2018). “Effect of Oil Palm Sustainability Certification on Deforestation and Fire in Indonesia”. In: *Proceedings of the National Academy of Sciences* 115.1, pp. 121–126.

Cisneros, Elías, Krisztina Kis-Katos, and Nunung Nuryartono (2021). “Palm Oil and the Politics of Deforestation in Indonesia”. In: *Journal of Environmental Economics and Management* 108, p. 102453.

Cramb, Rob and John F. McCarthy (2016). “Chapter 2 Characterizing Oil Palm Production in Indonesia and Malaysia”. In: *The Oil Palm Complex: Smallholders, Agribusiness and the State in Indonesia and Malaysia*. NUS Press.

Enrici, Ashley and Klaus Hubacek (2016). “Business as Usual in Indonesia: Governance Factors Effecting the Acceleration of the Deforestation Rate after the Introduction of REDD+”. In: *Energy, Ecology and Environment* 1.4, pp. 183–196.

Euler, Michael, Stefan Schwarze, Hermanto Siregar, and Matin Qaim (2016). “Oil Palm Expansion among Smallholder Farmers in Sumatra, Indonesia”. In: *Journal of Agricultural Economics* 67.3, pp. 658–676.

Gaveau, David, Bruno Locatelli, Mohammad Salim, Husnayaen Husnayaen, Timer Manurung, Adrià Descals, Arild Angelsen, Erik Meijaard, and Douglas Sheil (2021). “Slowing Deforestation in Indonesia Follows Declining Oil Palm Expansion and Lower Oil Prices”. In.

Gaveau, David, Douglas Sheil, Husnayaen, Mohammad A. Salim, Sanjiwana Arjasakusuma, Marc Ancrenaz, Pablo Pacheco, and Erik Meijaard (2016). “Rapid Conversions and Avoided Deforestation: Examining Four Decades of Industrial Plantation Expansion in Borneo”. In: *Scientific Reports* 6.1.

Gaveau, David L. A., Bruno Locatelli, Mohammad A. Salim, Husnayaen, Timer Manurung, Adrià Descals, Arild Angelsen, Erik Meijaard, and Douglas Sheil (2022). “Slowing Deforestation in Indonesia Follows Declining Oil Palm Expansion and Lower Oil Prices”. In: *PLOS ONE* 17.3. Ed. by RunGuo Zang, e0266178.

GoI (2016). “First Nationally Determined Contribution Republic of Indonesia”. In.

Goldman, Elizabeth Dow, Mikaela Weisse, Nancy Harris, and Martina Schneider (2020). “Estimating the Role of Seven Commodities in Agriculture-Linked Deforestation: Oil Palm, Soy, Cattle, Wood Fiber, Cocoa, Coffee, and Rubber”. In.

Greene, W. H. (2012). *Econometric Analysis*. 7th ed. Prentice Hall: Upper Saddle River, NJ.

Greenpeace (2011). *Indonesia Ministry of Forestry, Greenpeace, and WRI. “Indonesia Oil Palm Concessions.” Accessed through www.globalforestwatch.org in October 2020.*

Greenpeace Kepo Hutan Public Downloads - Google Drive (2023). URL: `https://drive.google.com/drive/folders/1-pA7sqXwm6MYiVqzydnUyDIxJST-JCzX` (visited on 01/20/2023).

Groom, Ben, Charles Palmer, and Lorenzo Sileci (2022). “Carbon Emissions Reductions from Indonesia’s Moratorium on Forest Concessions Are Cost-Effective yet Contribute Little to Paris Pledges”. In: *Proceedings of the National Academy of Sciences* 119.5, e2102613119.

Guillaume, Thomas, Martyna M. Kotowska, Dietrich Hertel, Alexander Knohl, Valentyna Krashevska, Kukuh Murtilaksono, Stefan Scheu, and Yakov Kuzyakov (2018). “Carbon Costs and Benefits of Indonesian Rainforest Conversion to Plantations”. In: *Nature Communications* 9.1, p. 2388.

Gunarso, Petrus, Manjela Eko Hartoyo, Fahmuddin Agus, and Timothy J Killeen (2013). “Oil Palm and Land Use Change in Indonesia, Malaysia and Papua New Guinea”. In: p. 36.

Hansen, M., P. Potapov, B. Margono, S. Stehman, S. Turubanova, and A. Tyukavina (2014). “Response to Comment on "High-Resolution Global Maps of 21st-Century Forest Cover Change"”. In: *Science* 344.6187, pp. 981–981.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend (2013). “High-Resolution Global Maps of 21st-Century Forest Cover Change”. In: *Science* 342.6160, pp. 850–853.

Harahap, Fumi, Sylvain Leduc, Sennai Mesfun, Dilip Khatiwada, Florian Kraxner, and Semida Silveira (2019). “Opportunities to Optimize the Palm Oil Supply Chain in Sumatra, Indonesia”. In: *Energies* 12.3, p. 420.

Harding, Torfinn, Julika Herzberg, and Karlygash Kuralbayeva (2021). “Commodity Prices and Robust Environmental Regulation: Evidence from Deforestation in Brazil”. en. In: *Journal of Environmental Economics and Management* 108, p. 102452.

Harris, Nancy L, Kevin Brown, Michael Netzer, and Petrus Gunarso (2013). “Projections of Oil Palm Expansion in Indonesia, Malaysia and Papua New Guinea from 2010 to 2050”. In: p. 28.

Heilmayr, Robert, Kimberly M Carlson, and Jason Jon Benedict (2020). “Deforestation Spillovers from Oil Palm Sustainability Certification”. In: *Environmental Research Letters* 15.7, p. 075002.

Heine, Dirk, Erin Hayde, and Michael Faure (2020). “Letting Commodity Tax Rates Vary With the Sustainability of Production”. In: p. 47.

Hertel, Thomas W (2018). “Economic Perspectives on Land Use Change and Leakage”. In: *Environmental Research Letters* 13.7, p. 075012.

Hsiao, Allan (2022). “Coordination and Commitment in International Climate Action: Evidence from Palm Oil”. In.

Hughes, Alice C. (2018). “Have Indo-Malaysian Forests Reached the End of the Road?” In: *Biological Conservation* 223, pp. 129–137.

Indonesia legal classification (2023). URL: `https://data.globalforestwatch.org/datasets/gfw::indonesia-legal-classification/about` (visited on 01/19/2023).

Jelsma, Idsert, G.C. Schoneveld, Annelies Zoomers, and A.C.M. van Westen (2017). “Unpacking Indonesia’s Independent Oil Palm Smallholders: An Actor-Disaggregated Approach to Identifying Environmental and Social Performance Challenges”. In: *Land Use Policy* 69, pp. 281–297.

Kharina, Anastasia, Chris Malins, and Stephanie Searle (2016). *Biofuels Policy in Indonesia: Overview and Status Report*, p. 20.

Khatiwada, Dilip, Carl Palmén, and Semida Silveira (2018). "Evaluating the Palm Oil Demand in Indonesia: Production Trends, Yields, and Emerging Issues". In: *Biofuels*, pp. 1–13.

Krishna, Vijesh V., Christoph Kubitza, Unai Pascual, and Matin Qaim (2017). "Land Markets, Property Rights, and Deforestation: Insights from Indonesia". In: *World Development* 99, pp. 335–349.

Leblois, Antoine, Olivier Damette, and Julien Wolfersberger (2017). "What Has Driven Deforestation in Developing Countries Since the 2000s? Evidence from New Remote-Sensing Data". In: *World Development* 92, pp. 82–102.

LeSage, James P. (2014). "What Regional Scientists Need to Know About Spatial Econometrics". In: *SSRN Electronic Journal*.

Levin, J (2012). *Sustainability and Profitability in the Palm Oil Sector*. WWF, FMO, CDC.

Lundberg, Clark and Ryan Abman (2022). "Maize Price Volatility and Deforestation". In: *American Journal of Agricultural Economics* 104.2, pp. 693–716.

Margono, Belinda Arunarwati, Peter V. Potapov, Svetlana Turubanova, Fred Stolle, and Matthew C. Hansen (2014). "Primary Forest Cover Loss in Indonesia over 2000–2012". In: *Nature Climate Change* 4.8, pp. 730–735.

Maryadi, Yusuf, A. K., and A. Mulyana (2004). "Pricing of Palm Oil Fresch Fruit Bunches for Smallholders in South Sumatra".

Masliani, M. Muslich Mustadjab, Syafrial, and Ratya Anindita (2014). "Price Determination of Palm Oil Fresh Fruit Bunches on Imperfect Competition Market in Central Kalimantan Province, Indonesia". In: *Journal of Economics and Sustainable Development* 5.1, pp. 134-139–139.

MoF (2008). *Reducing Emissions from Deforestation and Forest Degradation in Indonesia. Indonesian Forest Climate Alliance Consolidation Report*. Ministry of Forestry of the Republic of Indonesia, p. 185.

— (2019). *Kawasan Hutan 2019 - Kementerian Lingkungan Hidup Dan Kehutanan Republik Indonesia*. Ministry of Forestry of the Republic of Indonesia.

Mongabay (2022). *Palm Oil Firm Hit by Mass Permit Revocation Still Clearing Forest in Indonesia*. Mongabay Environmental News. URL: `https://news.mongabay.com/2022/02/palm-oil-firm-hit-by-mass-permit-revocation-still-clearing-forest-in-indonesia/` (visited on 01/17/2023).

Morel, Alexandra, Rachel Friedman, Daniel J Tulloch, and Ben Caldecott (2016). "A Case Study of Indonesia". In.

Mummolo, Jonathan and Erik Peterson (2018). "Improving the Interpretation of Fixed Effects Regression Results". In: *Political Science Research and Methods* 6.4, pp. 829–835.

Pacheco, P, S Gnych, A Dermawan, and B Okarda (2017). *The Palm Oil Global Value Chain: Implications for Economic Growth and Socialand Environmental Sustainability*. Center for International Forestry Research (CIFOR).

Paoli, Gary D, Piers Gillespie, Philip L Wells, Lex Hovani, Aisyah Sileuw, Neil Franklin, and James Schweithelm (2013). *Governance, Decision Making & Implications for Sustainable Development*, p. 70.

Petersen, Rachael, Dmitry Aksenov, Elena Esipova, Elizabeth Goldman, Nancy Harris, Irina Kurakina, Tatiana Loboda, Alexander Manisha, Sarah Sargent, and Varada Shevade (2016).

"Mapping tree plantations with multispectral imagery: preliminary results from seven tropical countries". In: p. 18.

Pirard, Romain, S Gnych, P Pacheco, and S Lawry (2015). *Zero-Deforestation Commitments in Indonesia: Governance Challenges*. Center for International Forestry Research (CIFOR).

Pirard, Romain, Nils Schulz, Jason Benedict, Robert Heilmayr, Ramada Febrian, Ben Ayre, and Helen Bellfield (2020). *Corporate Ownership and Dominance of Indonesia's Palm Oil Supply Chains*, p. 7.

Pirker, Johannes, Aline Mosnier, Florian Kraxner, Petr Havlík, and Michael Obersteiner (2016). "What Are the Limits to Oil Palm Expansion?" In: *Global Environmental Change* 40, pp. 73–81.

Potapov, Peter, Aleksey Yaroshenko, Svetlana Turubanova, Maxim Dubinin, Lars Laestadius, Christoph Thies, Dmitry Aksenov, Aleksey Egorov, Yelena Yesipova, Igor Glushkov, Mikhail Karpachevskiy, Anna Kostikova, Alexander Manisha, Ekaterina Tsybikova, and Ilona Zhuravleva (2008). "Mapping the World's Intact Forest Landscapes by Remote Sensing". In: *Ecology and Society* 13.2.

Preusser, S (2015). *The Correlation between Economic and Financial Viability with Sustainability for Palm Oil Plantations*. RSPO online.

Purnomo, Herry, Beni Okarda, Ade Ayu Dewayani, Made Ali, Ramadhani Achdiawan, Hariadi Kartodihardjo, Pablo Pacheco, and Kartika S. Juniwaty (2018). "Reducing Forest and Land Fires through Good Palm Oil Value Chain Governance". In: *Forest Policy and Economics* 91, pp. 94–106.

Rahman, Ayat K Ab, Ramli Abdullah, N Balu, and Mohd Shariff (2013). "The Impact of La Niña and El Niño Events on Crude Palm Oil Prices: An Econometric Analysis". In: *Oil Palm Industry Economic Journal* 13, p. 14.

Rifin, Amzul (2014). "The Effect of Progressive Export Tax on Indonesian Palm Oil Industry". In: *Oil Palm Industry Economic Journal* 14, p. 8.

Robalino, Juan A. and Alexander Pfaff (2012). "Contagious Development: Neighbor Interactions in Deforestation". In: *Journal of Development Economics* 97.2, pp. 427–436.

Sanders, D. J., J. V. Balagtas, and G. Gruere (2014). "Revisiting the Palm Oil Boom in South-East Asia: Fuel versus Food Demand Drivers". In: *Applied Economics* 46.2, pp. 127–138.

Santeramo, Fabio Gaetano and Stephanie Searle (2019). "Linking Soy Oil Demand from the US Renewable Fuel Standard to Palm Oil Expansion through an Analysis on Vegetable Oil Price Elasticities". In: *Energy Policy* 127, pp. 19–23.

Schoneveld, G C, D Ekowati, A Andrianto, and S van der Haar (2019). "Modeling Peat- and Forestland Conversion by Oil Palm Smallholders in Indonesian Borneo". In: *Environmental Research Letters* 14.1, p. 014006.

Scott, Paul T (2014). *Dynamic Discrete Choice Estimation of Agricultural Land Use*. en. 526, p. 56.

Shevade, Varada S. and Tatiana V. Loboda (2019). "Oil Palm Plantations in Peninsular Malaysia: Determinants and Constraints on Expansion". In: *PLOS ONE* 14.2. Ed. by Gopalasamy Reuben Clements, e0210628.

Souza Rodrigues, Eduardo. (2019). "Deforestation in the Amazon: A Unified Framework for Estimation and Policy Analysis". In: *The Review of Economic Studies Limited* 86.6, pp. 2713–2744.

Stavins, Robert N (1999). “The Costs of Carbon Sequestration: A Revealed-Preference Approach”. In: *American Economic Review* 89.4, pp. 994–1009.

UML (2018). *World Resources Institute, Rainforest Alliance, Proforest, and Daemeter. "Universal Mill List". Accessed through www.globalforestwatch.org on 01/2020.*

Wheeler, David, Dan Hammer, Robin Kraft, Susmita Dasgupta, and Brian Blankespoor (2013). “Economic Dynamics and Forest Clearing: A Spatial Econometric Analysis for Indonesia”. In: *Ecological Economics*, p. 12.

Wicke, Birka, Pita Verweij, Hans van Meijl, Detlef P van Vuuren, and Andre PC Faaij (2012). “Indirect Land Use Change: Review of Existing Models and Strategies for Mitigation”. In: *Biofuels* 3.1, pp. 87–100.

Wooldridge, Jeffrey (2002). “Econometric Analysis of Cross Section and Panel Data”. In: p. 741.

Wooldridge, Jeffrey M. (1999). “Distribution-Free Estimation of Some Nonlinear Panel Data Models”. In: *Journal of Econometrics* 90.1, pp. 77–97.

WRI (2015). *CAIT Country Greenhouse Gas Emissions: Sources and Methods*. World Resource Institute / World Resource Institute.

Appendix

A Tables

Table A.1: Descriptive statistics of palm oil mills in the Indonesian manufacturing census

	Geo-localized IBS palm oil mills n = 587 mills			**All IBS palm oil mills n = 930 mills**			**t-test**	**KS test**
	mean	std.dev.	median [min; max]	mean	std.dev.	median [min; max]	p-value	p-value
First year in IBS	1999	8.19	2001 [1975; 2015]	2000	8.78	2002 [1975; 2015]	0.000	0.000
FFB farm gate price (USD/tonne)	124.7	35.69	127.4 [16.84; 241.5]	123.3	35.73	125.8 [16.84; 242.2]	0.108	0.274
FFB input (tonne)	149047	115114	133193 [0; 1035319]	148035	114416	132552 [0; 1035319]	0.692	1.000
CPO farm gate price (USD/tonne	684.9	172.5	706.8 [170.1; 1191]	679.8	173.4	700.8 [170.1; 1191]	0.192	0.287
CPO output (tonne)	36082	24384	32902 [0.64; 179142]	35795	24363	32389 [0.64; 179142]	0.587	0.999
PKO farm gate price (USD/tonne)	399.9	140	389.4 [12.53; 827]	398.4	139.8	386 [12.53; 832.9]	0.676	1.000
PKO output (tonne)	8441	8918	6917 [0.11; 96775]	8368	8861	6846 [0.11; 96775]	0.724	1.000
CPO export share (%)	16.85	33.37	0 [0; 100]	15.75	32.55	0 [0; 100]	0.072	0.375
Central government ownership (%)	15.39	35.48	0 [0; 100]	14.64	34.76	0 [0; 100]	0.227	0.961
Local government ownership (%)	2.25	14.65	0 [0; 100]	2.1	14.17	0 [0; 100]	0.562	1.000
National private ownership (%)	65.75	46.02	100 [0; 100]	66.76	45.7	100 [0; 100]	0.214	0.831
Foreign ownership (%)	16.62	34.89	0 [0; 100]	16.51	34.88	0 [0; 100]	0.862	1.000

NOTE. This table reports summary statistics for a set of variables from the Indonesian manufacturing census (IBS), at the palm oil mill level, annually in 1998-2015. The sample of geo-localized IBS palm oil mills is a sub-sample of all IBS palm oil mills. IBS palm oil mills are identified here as IBS plants that report crude palm oil (CPO) or palm kernel oil (PKO) outputs, or fresh fruit bunches (FFB) inputs at least one year, and are not in Java or Bali islands. Farm gate prices are measured with mean unitary values (the ratios of value on quantity). USD is 2010-constant. We report p-values of Welch two-sided t-tests where the null hypothesis is that the true difference in means between the two groups is null, and the groups' variances are not assumed to be equal; and p-values of Kolmogorov-Smirnov tests where the null hypothesis is that the variables in the two groups are drawn from the same continuous distribution.

Back to Section 3.1

Table A.2: Estimation sample for industrial plantations - descriptive statistics

	Legal			Illegal			All		
	# grid cells = 3983 # grid cell-year = 24131			# grid cells = 3189 # grid cell-year = 17091			# grid cells = 11782 # grid cell-year = 65368		
	mean	std.dev.	median [min; max]	mean	std.dev.	median [min; max]	mean	std.dev.	median [min; max]
Deforestation (ha)	6.64	34.21	0 [0; 847.5]	3.14	24.32	0 [0; 763.2]	4.37	28.54	0 [0; 847.5]
Price signal ($/tCPO)	664	88.46	659.8 [394.9; 926.4]	665.6	98.64	663.8 [349.8; 921.4]	668.3	92.84	665.2 [349.8; 926.4]
Public ownership (%)	12.51	23.43	0 [0; 100]	19.12	29.8	0 [0; 100]	14.8	25.75	0 [0; 100]
Domestic private ownership (%)	69.3	31.99	78.93 [0; 100]	65.55	33.25	71.32 [0; 100]	68.06	32.17	75.64 [0; 100]
Foreign ownership (%)	18.19	27.55	0 [0; 100]	15.32	26.81	0 [0; 100]	17.13	26.78	0 [0; 100]
# reachable mills	8.82	6.08	7 [1; 37]	6.74	4.39	6 [1; 34]	7.67	5.17	7 [1; 37]

NOTE. This table shows descriptive statistics of the variables used in our main regression, for the samples of industrial and smallholder plantations together. We break it down to legal, illegal, and both or unknown ("All") categories. # means "number of". Price signal and ownership variables at the plantation level are inverse-distance weighted averages of these variables at reachable mills.

Table A.3: Estimation sample for smallholder plantations - descriptive statistics

	Legal			Illegal			All		
	# grid cells = 746 # grid cell-year = 5885			# grid cells = 1056 # grid cell-year = 5704			# grid cells = 3211 # grid cell-year = 20721		
	mean	std.dev.	median [min; max]	mean	std.dev.	median [min; max]	mean	std.dev.	median [min; max]
Deforestation (ha)	3.73	16.93	0 [0; 438.7]	9.8	34.55	0 [0; 653]	4.15	21.13	0 [0; 653]
Price signal ($/tCPO)	704.5	85.04	719.8 [349.8; 898.5]	724.9	79.53	740.8 [349.8; 898.3]	705.7	86.61	722.3 [349.8; 905.4]
Public ownership (%)	11.09	20.05	0 [0; 100]	14.36	22.6	0 [0; 100]	16.65	26.23	0 [0; 100]
Domestic private ownership (%)	77.72	24.33	83.59 [0; 100]	74.91	25.8	77.74 [0; 100]	71.17	28.93	76.23 [0; 100]
Foreign ownership (%)	11.19	16.48	0 [0; 96.46]	10.73	15.92	0 [0; 97.43]	12.18	18.71	0 [0; 100]
# reachable mills	8.87	3.91	8 [1; 27]	7.9	4.37	7 [1; 22]	7.96	4.09	7 [1; 27]

NOTE. This table shows descriptive statistics of the variables used in our main regression, for the samples of industrial and smallholder plantations together. We break it down to legal, illegal, and both or unknown ("All") categories. # means "number of". Price signal and ownership variables at the plantation level are inverse-distance weighted averages of these variables at reachable mills.

Table A.4: Price elasticities of illegal deforestation according to 2020 concession map

	Industrial plantations	Smallholder plantations	All
Estimate	4.85	1.96	2.89
95% CI	[1.86; 7.84]	[0.83; 3.09]	[1.29; 4.5]
Observations	18811	5843	22371
Clusters	467	279	656

NOTE. This table shows our estimates of the price elasticity of illegal deforestation, where illegal deforestation is identified as occurring within a protected forest zone and outside a known concession in 2020 (*Greenpeace Kepo Hutan Public Downloads - Google Drive* 2023). They are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signals. The price signal is measured as the 4-year average of annual inverse-distance weighted averages of crude palm oil prices at the gates of reachable mills. Deforestation is measured as primary forest loss eventually replaced with oil palm plantations. We differentiate industrial from smallholder plantations based on scale and landscape criteria (Austin et al. 2017; Petersen et al. 2016). We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. Sample observations are annual records of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014. They all have a positive extent of remaining primary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.

Back to Section 3.3

Table A.5: Cumulative and annual price elasticities of deforestation across Indonesian oil palm sectors

	Industrial plantations			Smallholder plantations			All		
	Legal	Illegal	All	Legal	Illegal	All	Legal	Illegal	All
Cumulative elasticity									
Estimate	0.83	4.16	1.94	2.26	1.53	1.47	0.67	2.64	**1.51**
95% CI	[-0.72; 2.37]	[1.44; 6.87]	[0.61; 3.27]	[-0.51; 5.03]	[0.27; 2.79]	[0.39; 2.54]	[-0.63; 1.97]	[1.1; 4.18]	**[0.5; 2.53]**
Elasticities to prices in:									
t									
Estimate	0.43	1.57	1.03	0.7	0.14	0.35	0.42	0.88	**0.73**
95% CI	[-0.31; 1.16]	[0.55; 2.59]	[0.54; 1.51]	[-0.57; 1.97]	[-0.56; 0.85]	[-0.23; 0.93]	[-0.2; 1.05]	[0.27; 1.5]	**[0.34; 1.13]**
t-1									
Estimate	0.25	1.66	0.75	-0.56	0.68	0.21	0.13	1.03	**0.62**
95% CI	[-0.22; 0.72]	[0.63; 2.7]	[0.29; 1.21]	[-1.45; 0.33]	[0.11; 1.25]	[-0.34; 0.76]	[-0.27; 0.53]	[0.41; 1.64]	**[0.27; 0.98]**
t-2									
Estimate	0.15	0.18	0.31	0.87	0.11	0.38	0.05	0.26	**0.29**
95% CI	[-0.4; 0.7]	[-0.83; 1.2]	[-0.2; 0.82]	[0.04; 1.7]	[-0.47; 0.69]	[-0.06; 0.81]	[-0.42; 0.53]	[-0.3; 0.82]	**[-0.11; 0.68]**
t-3									
Estimate	0	0.74	-0.14	1.25	0.6	0.53	0.06	0.47	**-0.13**
95% CI	[-0.62; 0.62]	[-0.49; 1.97]	[-0.64; 0.35]	[0.44; 2.05]	[-0.15; 1.34]	[0.1; 0.96]	[-0.45; 0.58]	[-0.22; 1.16]	**[-0.51; 0.24]**
Observations	24131	17091	65368	5885	5704	20721	26079	20695	**71926**
Clusters	629	451	1143	203	276	529	738	640	**1441**

NOTE. *For illegal smallholder plantations, the GLM algorithm did not converge, even with high number of iterations. Estimates are presented for informative purpose but should be taken with caution.* This table shows the elasticities of deforestation to the contemporaneous price signal and to the price signals in the three past years. The four elasticities are estimated jointly. The cumulative elasticity is the sum of these four elasticities. Price elasticity estimates are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signals. The annual price signal is measured as the inverse-distance weighted average of crude palm oil prices at the gates of reachable mills. Deforestation is measured as primary forest loss eventually replaced with oil palm plantations. We differentiate industrial from smallholder plantations based on scale and landscape criteria (Austin et al. 2017; Petersen et al. 2016). We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. They all have a positive extent of remaining primary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.

Back to Section 5.2

Table A.6: Price elasticities of deforestation across the oil palm sector, by island

	Industrial plantations			Smallholder plantations			All		
	Legal	Illegal	All	Legal	Illegal	All	Legal	Illegal	All
Sumatra									
Estimate	1.89	5.78	2.86	1.34	1.79	1.39	0.76	3.11	1.77
95% CI	[-0.22; 3.99]	[2.15; 9.41]	[0.87; 4.86]	[-1.52; 4.2]	[0.62; 2.97]	[0.19; 2.58]	[-1.08; 2.61]	[1.36; 4.85]	[0.47; 3.07]
Observations	7218	7908	26873	4347	5677	16050	8615	11435	32443
Clusters	277	306	680	190	274	512	378	494	972
Kalimantan									
Estimate	0.05	2.98	0.92	-16.25	2542.81	-6.16	0.04	2.95	0.97
95% CI	[-2.12; 2.21]	[-1.89; 7.86]	[-1.23; 3.07]	[-55.31; 22.81]	[869.67; 4215.96]	[-34.53; 22.22]	[-2.12; 2.19]	[-1.92; 7.83]	[-1.15; 3.08]
Observations	16905	9176	38087	1530	27	4427	17456	9197	38941
Clusters	352	145	465	13	2	17	360	146	472

NOTE. *For illegal industrial plantations, and illegal grouped ("All") plantations in Kalimantan the GLM algorithm did not converge, even with high number of iterations. Estimates are presented for informative purpose but should be taken with caution.*
This table shows our estimates of the price elasticity of deforestation by island. They are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signals. The price signal is measured as the 4-year average of annual inverse-distance weighted averages of crude palm oil prices at the gates of reachable mills. Deforestation is measured as primary forest loss eventually replaced with oil palm plantations. We differentiate industrial from smallholder plantations based on scale and landscape criteria (Austin et al. 2017; Petersen et al. 2016). We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. Sample observations are annual records of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014. They all have a positive extent of remaining primary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.

Back to Section 5.1

Table A.7: Price elasticities of deforestation in secondary forest, across the oil palm sector

	Industrial plantations			Smallholder plantations			All		
	Legal	Illegal	All	Legal	Illegal	All	Legal	Illegal	All
Estimate	0.49	2.24	0.28	-0.95	-1.38	-0.61	0.34	0.99	0.15
95% CI	[-1.11; 2.1]	[-1.41; 5.89]	[-0.88; 1.43]	[-1.99; 0.09]	[-2.83; 0.07]	[-1.28; 0.06]	[-0.96; 1.64]	[-1.16; 3.14]	[-0.72; 1.01]
Observations	41873	21208	125797	17726	9178	72517	44975	26580	139870
Clusters	1103	648	2508	467	433	1669	1292	918	3212

NOTE. This table shows our estimates of the price elasticity of deforestation in secondary forest. They are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signals. The price signal is measured as the 4-year average of annual inverse-distance weighted averages of crude palm oil prices at the gates of reachable mills. Deforestation is measured as forest loss outside the 2000 primary forest extent, eventually replaced with oil palm plantations. We differentiate industrial from smallholder plantations based on scale and landscape criteria (Austin et al. 2017; Petersen et al. 2016). We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. Sample observations are annual records of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014. They all have a positive extent of remaining secondary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.

Back to Section 5.1

Table A.8: Effects of price variability on deforestation across the Indonesian oil palm sector

	Industrial plantations			Smallholder plantations			All		
	Legal	Illegal	All	Legal	Illegal	All	Legal	Illegal	All
Estimate	-0.2	-0.08	-0.25	-0.25	0.29	0.1	-0.2	0.06	-0.11
95% CI	[-0.4; 0]	[-0.57; 0.41]	[-0.45; -0.04]	[-0.51; 0]	[0; 0.57]	[-0.07; 0.28]	[-0.37; -0.04]	[-0.19; 0.31]	[-0.26; 0.04]
Observations	24131	17091	65368	5885	5704	20721	26079	20695	71926
Clusters	629	451	1143	203	276	529	738	640	1441

NOTE. * For illegal smallholder plantations the GLM algorithm did not converge, even with high number of iterations. Estimates are presented for informative purpose but should be taken with caution.*
This table shows points of percentage change in average deforestation associated with a standard deviation in price signals across the 4 past years, with annual price signals measured as the inverse-distance weighted averages of crude palm oil prices at the gates of reachable mills. Deforestation is measured as forest loss outside the 2000 primary forest extent, eventually replaced with oil palm plantations. We differentiate industrial from smallholder plantations based on scale and landscape criteria (Austin et al. 2017; Petersen et al. 2016). We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. Sample observations are annual records of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014. They all have a positive extent of remaining primary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.
Back to Section 5.1

Table A.9: Price elasticity heterogeneity across local market development

	Industrial plantations			Smallholder plantations			All		
	Legal	Illegal	All	Legal	Illegal	All	Legal	Illegal	All
Price signal									
Estimate	0.2367	5.4404	1.6537	1.6296	2.0657	1.668	0.2683	3.3461	1.4736
95% CI	[-1.4397; 1.913]	[1.8292; 9.0516]	[0.0366; 3.2707]	[-1.2571; 4.5163]	[0.7332; 3.3983]	[0.3975; 2.9385]	[-1.1432; 1.6798]	[1.5366; 5.1556]	[0.3149; 2.6323]
Interaction with									
# reachable mills									
Estimate	0.004	0.008	8e-04	5e-04	-0.0036	-0.002	0.0032	2e-04	0
95% CI	[3e-04; 0.0077]	[-9e-04; 0.0169]	[-0.0021; 0.0036]	[-0.0058; 0.0069]	[-0.0103; 0.003]	[-0.0054; 0.0013]	[5e-04; 0.006]	[-0.0044; 0.0049]	[-0.0021; 0.0021]
Observations	24131	17091	65368	5885	5704	20721	26079	20695	71926
Clusters	629	451	1143	203	276	529	738	640	1441

NOTE. * For illegal smallholder plantations the GLM algorithm did not converge, even with high number of iterations. Estimates are presented for informative purpose but should be taken with caution.*
This table shows our estimates of the price elasticity of deforestation, along with estimated partial effects of interaction variables. Price elasticity estimates are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signals. The price signal is measured as the 4-year average of annual inverse-distance weighted averages of crude palm oil prices at the gates of reachable mills. Interaction terms are the product of the price signal and interacting variables, or covariates. The interacting variable is the annual count of all known reachable mills. The partial effects of interaction terms are second-order cross derivatives evaluated at the sample mean. Deforestation is measured as primary forest loss eventually replaced with oil palm plantations. We differentiate industrial from smallholder plantations based on scale and landscape criteria (Austin et al. 2017; Petersen et al. 2016). We identify illegal deforestation as occurring outside a known oil palm concession and inside a permanent forest zone designation. There are places where not enough information is available to designate the legal status. All estimates are derived from a generalized linear model of the quasi-Poisson family. All regressions include (geo-localized IBS) reachable-mills and district-year fixed effects, as well as the annual count of all known reachable mills as covariate. Sample observations are annual records of 3x3km grid cells in Sumatra and Kalimantan from 2002 to 2014. They all have a positive extent of remaining primary forest, and are within a 50km (30km in Sumatra) radius from at least one of our sample mills. 95% confidence intervals (CI) are based on standard errors computed with the delta method and clustered at the set of reachable mills.

Table A.10: p-values from equality tests of price elasticities

	All plantations			Industrial plantations	Smallholder plantations
Ho	Legal	Illegal	All		
industrial = smallholders	0.6773	0.0427	0.5561		
legal = illegal			0.0121	0.0113	0.6925

NOTE. This table shows p-values of two-sided t-tests, where the null hypothesis is that the true difference in price elasticities of deforestation between two groups is null.
Back to Section 5.1

B Figures

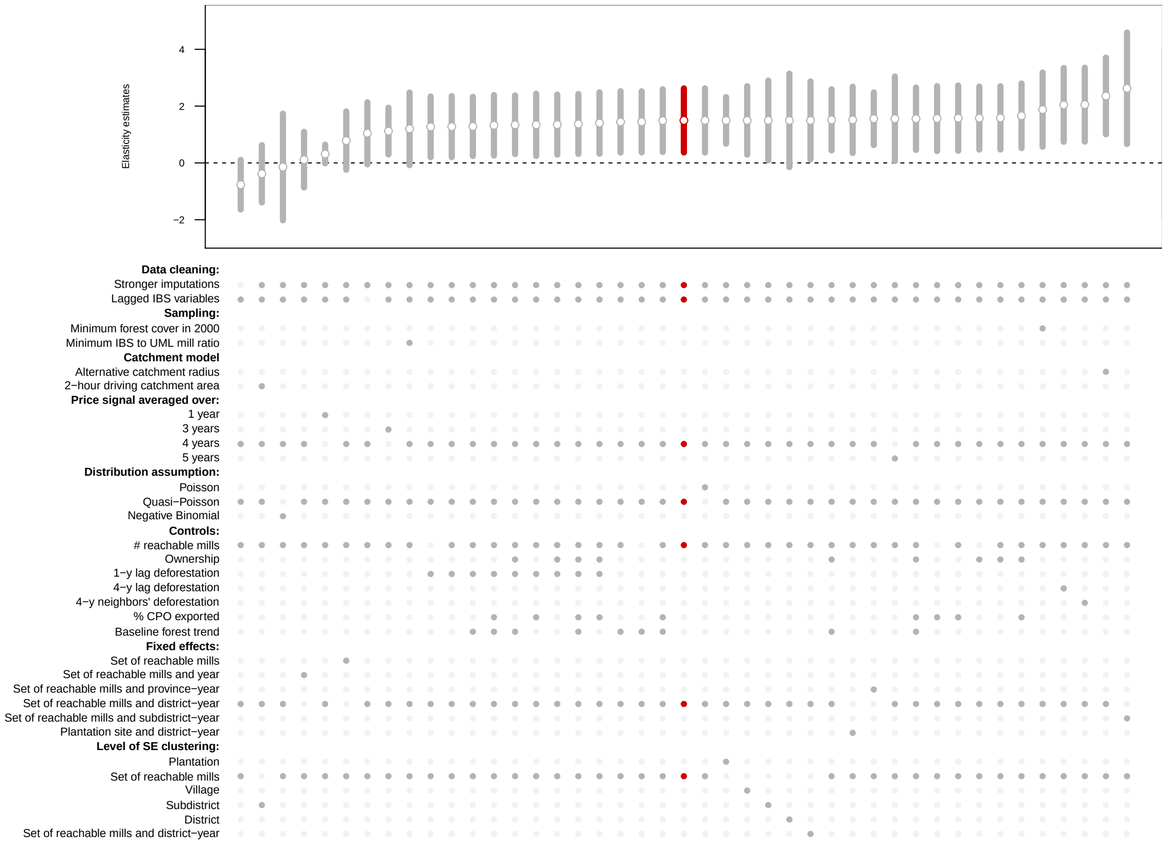

Figure B.1: Estimates of the Indonesian price elasticity of deforestation under different specifications

NOTE. This figure shows point estimates (white dots in upper panel) of the overall Indonesian price elasticity of deforestation estimated in this paper.Grey bars in the upper panel represent 95% confidence intervals. Darker marks in the lower panel mean that the corresponding vertical estimate is derived from a model that has the corresponding horizontal feature. The main specification is highlighted.
The minimum forest cover in 2000 is 50%. The IBS to UML mill ratio designates the number of mills from our sample relative to the total number of known reachable mills. It is also set to 50% (included). Alternative catchment radius is 50km in Sumatra and 30km in Kalimantan.
Back to Section 5.1

C Robustness analysis

Here, we document a battery of alternative estimation and identification strategies. We explain why these different specifications are relevant and we justify why we do not keep them in our main analysis. Figure B.1 shows how they compare with the overall price elasticity in Indonesian plantations estimated with the main specification (Equation 3) and sample described above. We mention only single departures from the main specification. We do not discuss combinations of alternative specifications.

IBS data cleaning. We check two departures from our main analysis in terms of preparation of IBS variables.

The first departure is the imputation described in Appendix D to clean price variables. In our main analysis, we use a stronger imputation, in order to reduce statistical noise due to duplicates. The softer cleaning choice appears to not only cause more statistical noise in the regressors (that would "just" yield an attenuation bias). It also maintain many mill-level price observations that have systematic measurement error. All these mill observations (829), that are removed in building the main dataset, are duplicates in terms of either CPO quantity or CPO value, (but not both). In most cases (816), the observation of quantity is duplicated from another year, while the value of the output is (presumably) correctly reported to the Indonesian office of statistics (BPS). We do not speculate here about the different ways how this measurement error can be systematically associated with deforestation. Rather, we argue that such source of variation, because it is not well understood, should not enter our main analysis.

The second data preparation we check is lag-adjusting price signals. Recall that, in our main analysis, we lag IBS variables to correct for a suspected measurement lag between them and remote sensing variables. Not lagging IBS variables yields a slightly lower and less precise estimate. This does not disprove our belief that prices recorded in IBS in a given year have little effect on the deforestation recorded that same year.

Sampling. We report the price elasticity estimates for two additional sampling conditions. In our main analysis, no such conditions are applied. Both conditions yield very similar estimates to the main one.

Under the first additional condition, we include in the sample only plantations where more than 50% of the area was covered with primary forest in 2000. This condition is relevant because it makes the sample more homogeneous in terms of initial land use. It is not included in our main analysis because it also limits the external validity of our results.

Under the second condition, we include in the sample only the plantations for which the set of known reachable mills comprises at least 50% of IBS geo-localized mills. This excludes plantations for which the measurement error is too high due to our geo-localized IBS mill data set not being exhaustive. In our main analysis, we do not apply this condition for the sake of generality and simplicity.

Catchment modelling. How we model the true relationships between mills and plantations is a critical point in our analysis. Therefore, we explore three alternatives to the model used in our main estimation strategy - catchment radii of 30km in Sumatra and 50km in Kalimantan.

The first alternative consists in the assumption that plantations are only influenced by prices at the nearest mill. This is the simplest model possible. Not surprisingly, it is very imprecise. This estimate's confidence interval is so large that we do not feature it in Figure B.1 for the sake of readability.

The second alternative is a different catchment radius in each island: 50km in Sumatra and 30km in Kalimantan. In Section 4.1, we discuss the size of the catchment radius and the reason why it should be lower in Sumatra than in Kalimantan. The alternative catchment radii yield a higher estimate.[48]

Finally, we model the catchment area of each mill not as a circle defined by a radius, but as the set of plantations that can reach the mill within two hours of driving (see Harahap et al. (2019) for a discussion on the driving time.[49]). This modelling is highly relevant because often, mills, although close to plantations in straight line distance, may actually not be reachable in time by trucks following weaving roads (and the opposite is also true). However, this modelling is not done in our main, preferred analysis because it may introduce endogeneity. Indeed, plantations likely expand (and hence deforest more) in parts of districts where the road infrastructure is better, while in the same area, prices are probably affected by the better access to markets enabled by better roads. This bias should be attenuated in our main analysis as we arbitrarily draw a line beyond which plantations are not connected to a mill although the road infrastructure would actually make the mill's prices influence deforestation. The estimate under this catchment area model is negative and imprecise, which may result from significant discrepancies between the road network available in our data (OpenStreetMap) and the actual roads used by palm oil producers.

Price signal time average. As explained in Section 4.1, our main measure of price signal is a 4-year average of annual price signals. We present here price elasticity estimates with different time average lengths.

Unsurprisingly, the short-run, annual price signal measure alone yields a non-significant estimate. Indeed, we expect the development of perennial crops to have little responsiveness to annual variations. This is confirmed by the narrow confidence interval.

The price elasticity point estimate increases with the average length of the price signal time, while precision decreases. With a 5-year average, too much noise enters the price signal measure and the price elasticity becomes less precise.

[48] We also get an estimate under a 10km catchment radius assumption, but here again we do not present it in Figure B.1 as the confidence interval is so wide that it complicates the reading of the whole figure.

[49] Harahap et al. (2019) use a four-hour constraint, grounding on `https://goldenagri.com.sg/plantation-mill-24-hours/` Here we present a twice shorter constraint because the estimation with the four-hour constraint yields too large a confidence interval to be displayed next to the other estimates.

Distributional assumptions. Our preferred distributional assumption is a quasi-Poisson distribution (which allows the variance to be different from the mean). A Poisson distribution assumption yields the same point estimate and very similar standard errors. This suggests that our data are not subject to over- or under-dispersion. The negative binomial distribution assumption is another option for count data. In our case, it yields a slightly higher but less precise estimate.

Control variables. We explore specifications with all combinations of control variables.[50] These include the number of reachable mills control variables in our main specification, and four additional control variables.
The first of the additional controls is the 1-year lagged outcome variable, i.e., deforestation. Deforestation has been often shown to be an auto-regressive process, and indeed we find that, in our data, lagged deforestation is positively correlated with current deforestation (results available upon request). Furthermore, we expect that prices from the 4 past years that we average in our price signal measure also influenced past deforestation. Indeed, in our data, we find that a price signal measured as an average of prices over 3 years does influence deforestation (cf. the above paragraph on different time average lengths). However, we do not believe that 1-year lagged deforestation can impact price signals (because of the time lag between planting and harvesting). Therefore, we suspect 1-year lagged deforestation to be an intermediate factor. We find that neither the magnitude nor the precision of our estimate varies with the inclusion of 1-year lagged deforestation. Thus, we conclude that the effect we measure is not inflated by the spurious accumulation of intermediate effects by which past prices would cause past deforestation that would then cause present deforestation.
We extend this robustness check to controlling for the 4-year lagged deforestation. Again, this control is motivated by the auto-regressive nature of deforestation. But here, we control for the risk that past deforestation affected prices through reverse causality. Indeed, it is possible that past-enough deforestation (four years ago) leads to increased production of fresh fruit bunches, that affects the marginal costs of surrounding mills (e.g. through increased market power of plantations, or economies of scale) and hence price signals. We find that adding this control to the main specification yields a price elasticity point estimate of 2.3. This contrasts slightly with our main estimate of 1.6. This difference seems to be due to the time period over which the regression with this control variable is estimated. The long lag (4 years) in deforestation restricts the estimating sample to the time period 2005-2014 (as deforestation is observed only as of 2001). Estimating our main model over this same period yields a point estimate of 2.3 (confidence interval [1.1; 3.6]). Hence, this robustness check makes us more confident that our results are not confounded by reverse causality.
The second additional control variable is the (inverse-distance weighted) average share of crude palm oil (CPO) exported by reachable mills. This proxies plantation exposure to the Indonesian export tax (Rifin 2014) and to international supply chains and hence might control for additional potentially confounding systematic differences between plantations. Adding it to the main con-

[50] Except the case without any control, and the 4-year lagged outcome variables that constrain estimations over a significantly different time period.

trol set yields a similar estimate.
The third additional control variables are the (inverse-distance weighted) average ownership shares of reachable mills: the share of domestic private capital and the share of foreign capital (we exclude the share of public capital to avoid perfect collinearity). We might be concerned that, for instance, local government mills have different deforestation motivations than foreign mills, while also having different marketing conditions. However, ownership changes may react to price shocks, while also being endogenous to local conditions. This makes ownership shares potential colliders, or "bad controls", that we prefer to exclude from our main analysis. Including them in the regression yields a similar estimate.
The fourth additional control variable is the baseline forest trend. This is built as an interaction between the primary forest cover in 2000 and the year. It captures differential trends between plantations with different initial land uses. These trends likely explain deforestation. If they are also correlated with price signals, they can bias our estimate. However, adding them to the main control set yields a similar estimate.
Finally, we present estimates with a last control variable: the 4-year lagged deforestation in neighboring sites. This is measured as the average of deforestation in the 8 neighboring plantation sites (grid cells) four years ago. As such, this variable captures the potential bias that could arise from global spatial spillovers (LeSage 2014). These spillovers occur when deforestation in surrounding areas affects local deforestation. They are likely to occur (Robalino and Pfaff 2012; Shevade and Loboda 2019), and in particular it is possible that surrounding deforestation in the past, (i.e., temporally and spatially lagged) affects current local deforestation. Such spillovers can bias our estimates if past surrounding deforestation also affects current local price signals (which are 4-year averages). This would occur if, around a plantation site in a given year t-4, deforestation was important enough so that four years later, when palm trees bear their first fruits, local prices in year t are impacted.
Over all plantation and deforestation types, controlling for the neighbors' past deforestation, the price elasticity point estimate is 2.3 (confidence interval [0.8; 3.8]). This contrasts slightly with our main estimate of 1.6. This difference seems to be due to the time period over which the regression with this control variable is estimated. The long lag (4 years) in deforestation restricts the estimating sample to the time period 2005-2014 (as deforestation is observed only as of 2001). Estimating our main model over this same period yields a point estimate of 2.3 (confidence interval [1.1; 3.6]). We see at least two explanations for the absence of bias from the neighbors' past deforestation. First, the important and heterogeneous time lapse between deforestation (observed in t-4) and palm tree planting mitigates the effect of deforestation - even aggregated over 8 plantation sites - on prices. Second, the limited market power of mills on the crude palm oil market makes it less likely that deforestation - even aggregated over 8 plantation sites - affects prices.
Given the substantial change in the sampling time period implied by the addition of this control variable and its negligible incidence, we do not investigate it in combination with the other robustness control variables presented above.

Fixed effects. Our main analysis uses a combination of reachable mills and district-year fixed effects, as we believe that most price endogeneity arises at the district level. Different fixed effects absorb variations at different levels. The reachable mills fixed effect alone removes little time heterogeneity, thus allowing aggregate shocks to confound the estimate, leading to a lower estimate. Adding a year fixed effect additionally controls for country-wide annual shocks that would apparently introduce a positive bias. Adding, rather, a local-year fixed effect, i.e., ruling out common confounding shocks at the level of province, district, subdistrict or village, yields positive estimates. These are precise in the case of province-year and district-year fixed effects, larger but less precise in the case of subdistrict-year fixed effects, and very imprecise in the case of village-year fixed-effects (which we do not display in Figure B.1 in order to better read it). This shows that most of the effect of price signals on deforestation is at play above the village-year level. Finally, holding the price departure with respect to the district market level, we change the other fixed effect from the set of reachable mills to the plantation level. This bans inter-plantation comparisons (contemporaneous or not) from identification. However, it introduces a new type of identifying comparisons: within the same plantation, but across years when it can reach a different set of mills. The resulting estimate is very similar to the main one.

Clustering. We show in Figure B.1 how allowing correlations in standard errors within different observation clusters affects confidence intervals. Price elasticity estimates are statistically different from zero with more clusters than in our main analysis - i.e., with plantation and village clusters. They also remain significant with larger and hence fewer clusters; namely, with district clusters and two-way plantation and district-year clusters.

D Data

In this section, we present the data we use to measure the components of Equation (3). The first subsection documents our original micro-economic dataset of geo-localized palm oil mills. The methodology to measure price signals and transform the mill data into the final sample of plantations is not described here but in Section 3.4. The second subsection presents the land use data, along with the methodology to measure deforestation.

D.1 Micro-economic data: an original merge of the Indonesian manufacturing census and the Universal Mill List

We semi-manually matched two existing data sets to produce an original, spatially explicit, microeconomic data set of palm oil mills in Indonesia from 1998 to 2015.

Indonesian manufacturing census (IBS). The Indonesian manufacturing census (IBS) is issued by the Indonesian office of statistics (BPS).[51] It reports annual establishment-level data for all manufacturing facilities employing at least 20 employees.[52] We identified palm oil mills with 9-digit commodity codes from 1998 to 2015. We use KKI codes 151410102 or 151410103 for crude palm oil and crude palm kernel oil respectively, and 011340101 or 011340501 for fresh fruit bunches. The variables available in the manufacturing census and used in our analysis are geographic variables;[53] mill-level input and output quantities and values at the 9-digit commodity level; mill-level ownership shares across four categories (national public, regional public, domestic private and foreign private); and product-level export shares.

Cleaning IBS data We use two main routines to clean input and output quantity and value variables: we remove duplicates, and we remove outliers. For each routine, we construct two cleaned variables: one with the stronger imputations (suffixed "imp1"), and one with the weaker imputations (suffixed "imp2"). The one with the stronger imputations described a more modified sample, in an attempt to reduce statistical noise (the term "removed" means "is given a missing value" throughout the paragraph). For duplicates within a firm identifier, imp1-variables observations are removed if either quantity or value is duplicated. For imp2-variables, observations are removed only if both quantity and value are duplicated. For duplicates within a year, imp1- and imp2-variables observations are removed only if both quantity and value are duplicated.

We define statistical outliers as observations that, within a year, are higher than $p75 + 1.5iqr$ where $p75$ is the 75^{th} percentile value and iqr is the interquartile range. We define outliers as observations of quantity variables that are statistical outliers and fail one of three tests. The first test asks whether the observation's input-output ratio is also a statistical outlier. The

[51] The data has also been referred to as *Statistik Industri* in the literature

[52] The average mill in IBS has 137 employees, and 75% of the mills have more than 87 employees. Thus, we are not worried that the 20-employee threshold is a threat in terms of selection bias.

[53] The data we obtained from BPS provided the district (*kabupaten*) information over the 1998-2015 period. However, the sub-district (kecamatan) and the village (desa) information were provided over 1998-2010 only.

second test asks whether the observation's crude palm oil-palm kernel oil ratio is a statistical outlier. The third test asks whether an observation's variation rate with respect to the previous period is an outlier. This procedure allows us to use all available information to deem an observation an outlier. For value variables, this is not possible and we deem an observation an outlier as long as it is a statistical outlier within a year. We express all monetary values used in the analysis in 2010 USD. We then compute price variables as mean unitary values: the ratios of quantities and values. We finally remove observations whose price variables are either upper or lower statistical outliers. Removing price upper outliers removes observations whose quantity is mismeasured (too low) relative to value, or whose value is mismeasured (too high though not outlier) relative to a true small quantity. Removing price lower outliers removes observations whose value is mismeasured (too low) relative to quantity, or whose quantity is mismeasured (too high though not outlier) relative to a true small value.

In addition, we lag all variables from the Indonesian manufacturing census, including prices, by one year. This merely aims at correcting a measurement lag. We do this because remotely sensed annual deforestation does not necessarily represent the actual state at the end of the year, while IBS variables should, a priori, reflect census respondents' observations for the whole year. Because this does not have conceptual implications for our empirical strategy, we do not annotate these lags or refer to them further.

Finally, with these cleaned variables, we identified 930 plants as palm oil mills, based on the criteria that they sourced FFB at least once or sold CPO or PKO at least once, and that they are not located in Java or in Bali.

Universal Mill List (UML). In the latest version we use, the Universal Mill List features 1140 Indonesian palm oil mills, with their names and coordinates (UML 2018). We merge the UML with a newer data set of palm oil mills (Benedict et al. 2023), containing information on parent companies and establishment dates, but we further refer to the whole data set as the UML.

Matching the manufacturing census and the UML. We matched the palm oil mills from these two data sets to make the manufacturing census economic data spatially explicit. The matching strategy leverages a third document: the manufacturing directories. This is a list of manufacturing establishments, with their names, 5-digit industry codes, main commodity names, addresses (often incomplete), and number of workers. Although they are edited annually, we could find them only for years 2003, 2006, 2009-2015. Since the number of workers in the directories is sourced from the manufacturing census (although with many lags, leads, and inconsistencies between the two), we used this variable together with district (and village when available) information to match mills from the manufacturing census with manufacturing directories' names. These names were then used to match the manufacturing census mills with UML coordinates. All conflicts were resolved after a case-by-case investigation. Finally, we match 466 mills from the manufacturing census with a UML palm oil mill (and four more which never reported CPO or PKO output, nor FFB input, or are located in Java)

There are 464 palm oil mills from the manufacturing census that could not be matched with the UML by the method explained above. Out of these, we approximate the geo-localization of the

121 additional mills for which village information is reported in the manufacturing census. To do so, we use the centroids of the polygons of the most recent valid village identifier. Because, in Indonesia, since 2000, there is a trend to village splits rather than to village mergers, the most recent information also tends to be the most spatially accurate.[54]

D.2 Land use change from forest to oil palm plantations

In this section, we explain how we construct our measures of land use change from forest to oil palm plantation (referred to as 'deforestation' here).[55]

Forest loss. We use maps from the Global Forest Change (GFC) dataset (Hansen et al. 2013). They cover the whole of Indonesia with a resolution of 1 arc-second per pixel (i.e. 27.8 x 27.6 meter pixels with our projection) annually from 2001 to 2018. A forest loss event is defined at the pixel level, as the year when complete removal of tree canopy cover (with a minimum height of 5m) is observed where such cover was still present in 2000. A minimum canopy cover threshold defines what is counted as forest in 2000 at the pixel level. However, the GFC dataset does not enable us to distinguish between 2000 tree canopy cover (and hence loss) in primary forest, secondary forest, or tree plantations.

Primary forest extent in 2000. The map we use to measure primary forest extent in 2000 comes from Margono et al. (2014). It covers the whole country, with the same resolution as the GFC data set. Primary forest in 2000 is a subset of the 2000 tree canopy cover from the GFC data set, with canopy cover of at least 30%. It is defined as "mature natural forest cover that has not been completely cleared in recent history and consisted of a contiguous block of 5ha or more" (Margono et al. 2014). Two primary forest types are distinguished: intact and degraded. The former, following Potapov et al. (2008), shows no sign of alteration by humans, while the second has been subjected to human disturbances, such as selective logging. They correspond to the Indonesian Ministry of Forestry's primary and secondary forest cover types (Margono et al. 2014). In this study, we regroup them.

Oil palm plantations. In this study, we use two different maps, from Austin et al. (2017) and Petersen et al. (2016). These maps have been produced by visual interpretation of Landsat imagery. They both recognize areas with signs of future cultivation as plantations. The former product, from Austin et al. (2017), includes only large-scale oil palm plantations and covers the

[54] Due to administrative village splits, plants do not necessarily report their correct village names or codes every year. This can be particularly misleading because codes for "parent" villages may be re-used in the next iteration but for different villages than their "child" villages. Therefore, we deemed that the village information a plant reported in a given year was valid if the corresponding "parent" village (in 2000) matched with the mode of all annual village information reported by the plant (also expressed in "parent" village).

[55] All rasters used in this study are aligned with the resolution of forest loss maps from Hansen et al. (2013) and all spatial data are projected with a Cylindrical Equal Area projection centered on Indonesia (longitude = 115, latitude = 0).

regions of Sumatra, Kalimantan, and Papua for the years 1995, 2000, 2005, 2010 and 2015, with a 250m pixel resolution. The latter product, from Petersen et al. (2016), includes and distinguishes between large plantations of more than 100ha, mid-size plantations and small-size plantations. It is a snapshot of the whole of Indonesia, computed with images from 2013 and 2014. Mid and small-size plantations are mosaic landscapes. Mid-size plantation mosaic landscapes are at least 100 hectares wide, have oil palm patches between 10 and 100 hectares, comprising at least 50% of the landscape. Small-size plantation mosaic landscapes have oil palm patches smaller than 10 hectares, again comprising at least 50% of the landscape.
In our main analysis, we use the maps from Austin et al. (2017) to study industrial plantations, and we pool small and mid-sized plantation maps from Petersen et al. (2016) to study smallholder plantations. Where these map sources overlap, we characterize plantations as industrial, as remote sensing for this landscape is less error-prone.

Measuring deforestation. We combine these data sets to compute annual maps of deforestation for oil palm plantations. Our main forest definition at the pixel level, hence determining our baseline forest extent in 2000, is any (i.e., intact or degraded) primary forest.[56] This corresponds to the official forest definition by the Government of Indonesia (MoF 2008; Austin et al. 2017) which justifies that this is retained in our main analysis. In Table A.7 we present alternative results from secondary forest. We define secondary forest in 2000, at the pixel level, as tree canopy cover of at least 30 percent, outside primary forest, and notably outside 2000 industrial oil palm plantations (as observed by Austin et al. (2017)).[57]
Then, annual (primary or secondary) forest loss pixel events observed within the 2000 baseline forest extent are deemed deforestation events if they later fall within an oil palm plantation. This means that we count a deforestation pixel-event the year the forest is cleared, and not the year the palm trees are planted or when they become productive.

[56]We routinely exclude pixels categorized as industrial plantations in 2000, although the primary forest map should already exclude them.

[57]This ensures that canopy closure removals within already existing plantations (i.e., palm replacements) are not counted as deforestation. This approach is the best we can do in the absence of other tree plantation maps for 2000, but it still has some pitfalls. For instance, if an area was covered with another plantation type (like timber) in 2000, cleared and converted to an oil palm plantation before 2015, it would be mistakenly counted as deforestation.

E Empirical framework

E.1 Estimation strategy

Functional form and estimation In this study, we estimate an exponential mean model by Poisson Quasi maximum likelihood. The Poisson distributional assumption has been made elsewhere in statistical studies of (Indonesian) deforestation (e.g., Burgess et al. (2012), Busch et al. (2012), and Busch et al. (2015)). Hence, we also seek comparability of our results with, in particular, Busch et al. (2015). The quasi-Poisson distribution imposes weaker assumptions on our data, as it only requires the mean (and not the variance) to be correctly specified. We use the standard log-link function. We estimate Equation 3 with the feglm algorithm from the R package *fixest*. This method estimates generalized linear models using weighted ordinary least squares (OLS) estimations with demeaning along fixed effect dimensions in the OLS steps and no presence of the incidental parameter problem (Bergé 2018).

E.2 Partial effects

In all regressions, the price signal variable is scaled to the natural logarithm. The partial effects of price signals on deforestation are computed as the relative difference between predicted deforestation at the sample means, with and without a 1% increase in the price signal, multiplied by 100 (hence, all estimates are scaled to percentage points). From Equation 3, this simplifies to $100(1.01^{\hat{\alpha}} - 1)$% and hence does not depend on sample means (Bellavia et al. 2015). This only slightly differs from the exponential of regression coefficients as it gauges the effect for a "full" 1% change in a right-hand-side variable and not for an infinitesimal change. We present results this way because it is more consistent with computation of effects for larger changes (e.g., one standard deviation) or when second-order terms are included on the right-hand side. We estimate the variance of the partial effect with the delta method (Greene 2012).

To investigate synergies, we use interaction terms: right-hand-side variables computed as the product of the treatment (price signal here) and an interacting variable which is also featured in the right-hand side. Because our model is not linear, the informative estimate is the partial effect of the interaction term, not its coefficient (Ai and Norton 2003). Hence, interaction estimates discussed in Section **??** and displayed in Tables 6, 5 and A.9 are second-order cross-derivatives of predicted deforestation, evaluated at the sample mean.

F Comparison with existing estimates

Here, we attempt to compare our findings with the closest estimates in the literature. Yet, we remark that none of the studies discussed here have provided a price elasticity of deforestation as their main estimate. Therefore, they may naturally have focused less on identification concerns about this parameter. The first (in time) study we can compare our estimates to, is Wheeler et al. (2013). They estimate a log-log regression of deforestation on a time series of palm oil futures prices and other economic variables. We can compare our estimated price elasticity to their model coefficient of 0.816. Using our spatial variation, we hence find a price elasticity twice as large as theirs. We shall note that this difference may also come from differences in the measure of deforestation between our two studies.

Comparing with Busch et al. (2015) requires more assumptions, because this study provides an estimate of the effect of agricultural revenue - and not price - on deforestation. They find that an additional $100 (in 2005 USD) is associated with a 1.02-1.18% increase in deforestation. Converting to 2010 USD, assuming an average yield of 3.5 ton CPO per hectare (Khatiwada et al. 2018) and an average price of $680 / ton CPO over the period (based on our own data), we convert their estimates into a 0.13-0.15 price elasticity.[58] This is lower but comparable to our estimated 1.8 price elasticity of deforestation in industrial plantations, which is the most similar setting to theirs. One should note that the agricultural revenue in Busch et al. (2015) is computed at each land parcel for the most potentially lucrative crop, which is oil palm 69% of the time.

In Cisneros et al. (2021) the effect of price exposure (calculated as the interaction of international prices and suitability for oil palm) on deforestation is expressed for one standard deviation. Thus, in order to compare our analyses to theirs, we compute our partial effects for one standard deviation in our data (remaining after fixed-effect variations are absorbed). In their study, a one standard deviation higher palm oil price exposure results in an 8% increase in deforestation. This is exactly equivalent to the effect of one standard deviation in our setting (corresponding to our main 1.5 price elasticity estimate). However, for the two studies to be more aligned, we compare our price elasticity in industrial plantations (10.2% increase in deforestation for a one-standard-deviation increase in price signals) to their estimated effect of price exposure on deforestation in new industrial oil palm plantations by 2015 (3% and imprecise). Hence, here too, our research setting seems to capture a larger effect of prices on deforestation in the Indonesian oil palm sector.

[58] We convert the additional $100 to a $100 * \$100/(0.518 * 3.5 * 680) \approx 8.110924$ percentage change in CPO prices (where 0.518 is approximately the deflator we use). We then scale the associated percentage change in deforestation - either 1.02 or 1.18% - by this relative price change.