

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

Valentin Guye[†], Sebastian Kraus[‡]

May 12, 2023

[Latest version](#)

Abstract

Global demand shifts and supply chain interventions have the potential to reduce palm oil's environmental footprint, especially in otherwise unregulated (informal or illegal) plantations. This ultimately depends on deforestation reacting to prices in upstream, complex plantation-mill systems. We produce the first microeconomic panel of geolocalized palm oil mills, and we model their influence on palm plantations across Indonesia where the issue is most critical. We leverage our data granularity and the nature of the value chain to isolate downstream mill-gate price shocks that are exogenous to deforestation upstream. We find that deforestation for smallholder (informal) plantations and for illegal industrial plantations is elastic to the mill-gate price of crude palm oil. However, smallholder deforestation decelerates as palm fruit prices increase. These results inform the design of fair and effective conservation policy.

*JEL-codes: Q23, Q57, Q58, Q56, Q15

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 26th 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, AgroParis-Tech, 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

In this century's first decade, the conversion of Indonesian forests to oil palm plantations has released 1% to 4% of anthropogenic greenhouse gas emissions (Busch et al. 2015). Consequences in other environmental and socio-economic dimensions are substantial too (Qaim et al. 2020). Oil palm plantation expansion and associated deforestation rates have declined since 2013 (Gaveau et al. 2022). This is commonly attributed to lower palm oil prices, as price incentives are identified as a major driver of tropical deforestation in general (Busch and Ferretti-Gallon 2017; Berman et al. 2023), and effectiveness of the existing conservatory initiatives is limited by weak institutions in this context (Heilmayr et al. 2020; Drost et al. 2021; Groom et al. 2022). Hence, long-run growth in palm oil demand and downturns in the supply of substitutes (fuels and vegetable oils) may raise price signals, triggering a new crisis and sustaining deforestation unaddressed by regulations. Although such statement is widespread, and market-based solutions have been proposed (Heine et al. 2020), there is still little evidence that, upstream the palm oil supply chains, prices actually incentivize unregulated deforestation.

In this paper, we investigate whether and how price signals cause deforestation in the complex landscapes upstream the Indonesian palm oil supply chains. We primarily seek to identify whether price signals have a causal effect on oil palm deforestation, overall and for specific kinds of unregulated plantations — namely smallholder plantations and illegal industrial plantations. We question how prices incentivize actors with differential agency and interest into deforesting for industrial and smallholder plantation expansion. We also explore heterogeneities in price responsiveness and gauge response magnitudes.

We advance beyond previous studies that estimated a relationship between international palm oil prices and deforestation for undistinguished oil palm plantations at national scale (Wheeler et al. 2013; Busch et al. 2015; Cisneros et al. 2021; Gaveau et al. 2022; Busch et al. 2022; Hsiao 2022). Wheeler et al. (2013) first established a positive correlation between time series of palm oil futures prices and forest loss alerts at a monthly rate. Gaveau et al. (2022) found a similar relationship at annual rates, although less precise in the distinct case of smallholder plantations. Subsequent studies advanced identification of the causal price effect by adding spatial variation. They proxied local incentives by interacting international prices with agro-climatic measures of local suitability for palm plantations (Busch et al. 2015; Cisneros et al. 2021; Hsiao 2022). Although powerful, this approach is subject to three causal inference challenges

(partially addressed in these studies): reverse causality due to deforestation in most suitable areas affecting the international price; correlation patterns with other crops in terms of suitability and prices; and systematic measurement errors if deforestation rates correlate with information on suitability or international price pass-throughs. Moreover, these studies did not focus primarily on the link between palm oil price and deforestation and thus, they did not provide in depth analyses of this relationship. Specifically, they did not study unregulated deforestation distinctly, which requires adequate conceptual and empirical approaches.

We produce new country-wide data that allows us to advance on the internal validity front and to study in new depth the influence of price incentives on deforestation, notably unregulated, in the upstream stages of the supply chain. Data on the Indonesian palm oil supply chain available to researchers is limited. Previously, only the location and total capacity of most palm oil mills was known, together with the establishment date for a subset of them. For about half of all known Indonesian palm oil mills, we extend the geo-localized information to the full set of attributes in the Indonesian manufacturing census (IBS).¹ We notably observe annual input (palm fruits) and output (crude palm oil), mill-gate prices, public, private and foreign ownership shares, as well as export shares. We use data based on satellite imagery to measure deforestation as 30m-pixel events of primary forest loss eventually converted to oil palm plantations. These secondary data distinguish industrial from smallholder plantations based on scale and shape features that match economically and politically relevant differences. We further detect illegal deforestation by associating knowledge of the legal requirements with the best available data on palm concessions and on land use zoning. To measure price signals, we leverage the agronomic constraint on the distance between palm fruit harvest, in plantations, and processing into crude palm oil, in mills. For every plantation site that can reach at least one mill, we average the crude palm oil prices received at reachable mills. By assigning higher weights to closer mills, we model the relative influences of reachable mills in a way that is consistent with unobserved vertical integration and transport cost structures. By averaging annual price signals over the four last years, we capture the information that we assume relevant and available to prospective plantation agents forming expectations of the profitability of the perennial oil palm crop.

Our identification strategy relates to a reduced-form shift-share instrument where exoge-

¹In the economics literature, this dataset has also been referred to as *Statistik Industri*; see, for instance, Amiti and Konings (2007).

nous variation is in the shifts (the mill-gate prices), as proposed by Borusyak et al. (2022). We exploit that mill-gate prices have an idiosyncratic component pertaining to the deals mills close with downstream buyers (Pirard et al. 2020; Wiggs et al. 2020). We assume that the terms of such deals are orthogonal to mills' own or upstream characteristics, given the concentration of buyers (global traders and refineries) and the profitability of the milling sector (Byerlee et al. 2016; Pirard et al. 2020). The granularity in our primary data allows us to isolate this exogenous variation in price signals from local dynamics up to levels beyond which crude palm oil prices are unlikely endogenous (i.e., by the means of contextually highly resolved fixed effects). Notably, we allow for endogenous price formation at the levels of evolving plantations' purchasing sheds and districts' yearly political economy.

Beyond internal validity improvements, we pinpoint a price elasticity estimate to the narrower, politically promising, plantation-mill segment of the supply chain. Furthermore, with external validity stemming from country-wide, 15-year period representative data, we discover how deforestation responds to price incentives in the upstream end of the Indonesian palm oil supply chain. First, our primary mill-gate price signal measurements allow us to reveal the price responsiveness of specific but important forms of deforestation, namely legal and illegal deforestation for either industrial or smallholder plantations. Second, we disentangle the influences of actors with different agency and interest on deforestation for smallholder plantations, in reaction to a change in prices. Finally, we analyze the dynamics of price signals and forest conversion revealed in our data to document the decision rules of deforesting agents.

Our main results are threefold. First, we find that a 1% increase in crude palm oil price signals increases the average conversion of primary forest to oil palms by 1.5%. This estimate is robust in a range of alternative settings, including mill catchment radius assumptions, price signal dynamics, control sets, fixed effect and standard error clustering levels. Second, looking at unregulated deforestation, we find that deforestation for smallholder plantations, and for illegal industrial plantations, are price responsive and drive the overall effect. The effect of price signals on legal deforestation is consistently smaller and undistinguished from zero. Thirdly, while deforestation for smallholder plantations increases with crude palm oil price signals, it decreases with palm fruit price signals. We keep finding such negative elasticity after controlling for an eviction effect from industrial plantations.

Together, these results have several implications. First, deforestation left unregulated will follow prices. Our results suggest that last decade's price contraction helped deforestation to

slow down, and that the recent surge in prices could overtake conservation efforts that rely on strong institutions. We estimate that a positive standard deviation in price signals leads 10 thousand hectares of primary forest to be converted the next year. Our results can feed into more dedicated efforts to model deforestation impacts of e.g. trade or biofuel policies. Notably, our 1.5 crude palm oil price elasticity estimate is one order of magnitude higher than that used by Busch et al. (2022) to conclude that export bans have little effectiveness in curbing deforestation.

With consumer-side restrictions facing coordination (Hsiao 2022) and traceability challenges (Lyons-White and Knight 2018), price interventions at upstream choke points, like mills, are promising policy propositions — especially to steer hardly regulated deforestation (Heine et al. 2020). Our finding that deforestation for smallholder and for illegal plantations respond to prices implies that such interventions could work. This matters because the relative footprint of smallholder expansion is predicted to grow (Schoneveld et al. 2019) and illegal deforestation is documented to prevail in active frontiers.² Our back-of-the-envelope estimation indicates that a 19% tax levied uniformly on palm oil mills could curb deforestation 29% under average and earn Indonesia about USD 120 million a year for avoided emissions (at a USD 5/tCO₂ price). Refunding against proof of sustainable production would further improve effectiveness, at low monitoring costs (Heine et al. 2020).

Finally, our results indicate that protecting smallholder revenues further increases the effectiveness of a conservatory tax on crude palm oil, as smallholders otherwise compensate depressed prices by expanding plantations into the forest.

The results and their implications relate this paper to the literature on the economic drivers of land use change³ and smallholder agriculture in particular (Krishna et al. 2017; Dalheimer et al. 2022); and to the literature on the interdependent roles of regulatory and incentive-based tropical supply chain governance (Godar et al. 2014; Busch et al. 2015; Harding et al. 2021; Lambin and Furumo 2023).

Next, we provide background and a conceptual framework in Section 2; data and measurements in Section 3; estimation and identification strategies in Section 4; results in Section 5; and concluding discussions, including of the policy implications, in Section 6.

²Notably in the island of Papua: <https://news.mongabay.com/2018/11/the-secret-deal-to-destroy-paradise/>

³We point in particular to Busch and Ferretti-Gallon (2017) for a review; Leblois et al. (2017) and Berman et al. (2023) for global analyses; Souza Rodrigues (2019) in the Amazon context.

2 Background and conceptual framework

2.1 Background: unregulated upstream plantations and pivotal mills in the palm oil supply chain

The Indonesian palm oil supply chain. The global demand for vegetable oil has led to an increase in the cultivation of oil palm, the most productive oil plant (Corley and Tinker 2015). A significant share of the global area suitable for oil palm cultivation is in Indonesia, where a complex supply chain has developed to currently account for more than half of the global production. Across the country, more than 15M hectares of oil palm tree plantations provide fresh fruit bunches (FFB) that are processed locally by a thousand mills into crude palm oil (CPO) (Pirard et al. 2020). The majority of the planted area and production comes from large, grid-shaped landscapes, ranging from a hundred hectares to hundreds of thousands of hectares, called industrial plantations (Gaveau et al. 2016; Austin et al. 2017). Foreign and domestic companies, as well as local governments invest in these plantations and operate their development. Some industrial plantations are vertically integrated with mills, but this integration seems limited (Pirard et al. 2020).⁴ Around 400 refineries and exporters take custody of mills' CPO, in one of the 61 Indonesian ports (Pirard et al. 2020). Crude palm oil is a standardized, fungible commodity that, unlike palm fruits, can be transported across long distance without degrading (Byerlee et al. 2016). Part of the milling and refining capacity is vertically integrated, but this is not the dominant model (Pirard et al. 2020). Refined palm oil is sold to a myriad of manufacturers for further processing into final products including food, biodiesel and cosmetics. The palm oil supply chain is hence hourglass shaped and characterized by limited vertical integration instead of a transparent ownership structure (Glenday et al. 2015; Pirard et al. 2020). This hampers the traceability of physical and financial flows with upstream plantations, thus preventing supply chain governance from downstream interventions (Lyons-White and Knight 2018; Pirard et al. 2020; zu Ermgassen et al. 2022).

The unregulated oil palm plantations. Governing oil palm plantations through upstream, place-based interventions is challenging as well, in particular because of the prevalence of unregulated plantations in Indonesia (Heilmayr et al. 2020; Drost et al. 2021; Groom et al. 2022).

⁴This is also suggested, in particular until 2013, by Agriculture Regulation No. 98/2013 which stipulates that plantations larger than 1000 hectares must integrate with mill operations, and that mills must source at least 20% FFB from their own plantations (Glenday et al. 2015).

We distinguish two non-exclusive types of unregulated plantations: informal and illegal plantations.

Informal plantations refer to palm trees which planting and cultivation are not or weakly identified by the institutions in place. Typically, this echoes with smallholder plantations. This term has no single definition, but is often used in contrast with industrial plantations of larger scale and formally linked to a company (although not necessarily legal).⁵ Yet, not all smallholders are informal: some are developed with mandatory and formal support from a plantation company, as part of a so-called “plasma” scheme (Paoli et al. 2013; Byerlee et al. 2016). Smallholder plantations developed outside such schemes have driven the smallholder expansion during our study period (Byerlee et al. 2016; Euler et al. 2016), until accounting for more than a third of the total oil palm acreage (Gaveau et al. 2022) and a tenth of oil palm deforestation (Lee et al. 2014). They are characterized by their non-industrial size and shape, and by not being tied to a single mill to which they sell exclusively. Rather, these smallholders sell their fruits to different mills through intermediaries (called ‘middle-men’), and are thus called independent smallholders (Cramb and McCarthy 2016; Baudoin et al. 2017). However, they are not independent regarding the extent and timing of their expansion, which also depends on locally more powerful actors like companies operating mills and industrial plantations — partially in the context of land bargaining associated with transmigration schemes (McCarthy et al. 2012b,a; Potter 2012; Li 2015; Euler et al. 2016; Jelsma et al. 2017). In this paper, we use the term ‘smallholder plantations’ to refer to small-scale plantations that tend to be informal, selling on a spot-market, and having incomplete power over expansion decisions.

Illegal plantations are typically developed in forested areas that are not legally designated for oil palm cultivation, and/or without all the required authorizations (see Paoli et al. (2013) for a comprehensive overview of those). According to the legal land designation criteria, both industrial and smallholder plantations can be illegal. However, only industrial plantation development formally requires several permits that establish so-called oil palm concessions. The complexity of land institutions implies that illegality can take many forms. Yet, it is generally characterized as a problem of weak central law enforcement capacity, which is commonly at-

⁵Plantations 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). This size threshold provides one of the definitions for smallholders, although plantations both below and above this threshold rarely hold the corresponding permit (Craw 2019).

tributed to the decentralization process that started in 2001.⁶

The pivotal situation of palm oil mills. Governing the expansion of oil palm plantations on forest is important, because it has significant adverse consequences on biodiversity, climate change and forest-reliant livelihoods in particular (Petrenko et al. 2016). However, the hourglass-shaped supply chain and unregulated expansion hamper the effectiveness of upstream and downstream interventions. In this perspective, mills have a pivotal position in the palm oil supply chain. They are sufficiently upstream to have the potential to trace their sourcing to the plantations; and they may have sufficient local power to influence expansion decisions directly. Moreover, it has been shown that FFB prices locally result from mills' discretionary decisions based on their monopsonic market power, the quality of FFB purchased, and their respective CPO sales (Maryadi et al. 2004; Masliani et al. 2014).⁷ On the other hand, mills are sufficiently downstream to be governable, in particular through a price instrument (Purnomo et al. 2018). Hence, palm oil mills represent a strategic choke point in the supply chain, from where it is possible to steer the expansion of plantations, and especially of unregulated ones (Heine et al. 2020).

However, the effect on deforestation from any such intervention — and more generally from any downstream shock or demand shift — depends on whether forest conversion upstream the supply chain reacts to crude palm oil price changes, on mills' downstream market. In light of the context exposed above, we now explicit how we conceptualize this reaction.

2.2 Conceptual framework

A standard conceptual framework posits that a deforestation event results from the decision of an optimizing micro-economic agent. The typical decision rule is the comparison of the expected discounted present utilities (or profits) from alternative inter-temporal scenarios, defined by the kind, the timing and the amount of deforestation.⁸ To form such expectations, every agent grounds on privately observed informational elements (Stavins 1999). Here, we

⁶Mechanisms include corruption being fostered by competition for resource rents between local authorities (Burgess et al. 2012; McCarthy et al. 2012b), and law bypassing being facilitated by ambiguous and overlapping authority between institutions (Setiawan et al. 2016).

⁷Although FFB prices are supposed to be collectively determined by provincial governments, firms and farmers.

⁸The counterfactual scenario includes both conservation and deforestation to other land uses. Conservation includes both expansion outside forest or no expansion (i.e., intensification or not entering the oil palm fruit market). We do not distinguish between these alternative scenarios in our analysis.

are interested in those of such elements that are prices, and we call them price signals. Price signals can inform about future prices and profits, and can also reflect accrued revenues available for new investments. Hence, the price elasticity of deforestation essentially refers to the observation of a price signal and the act of deforesting accordingly, by what we call a plantation agent.

In the present context, detailed above, plantation agents are not homogeneous in their decision rules. Plantation agents owning large plantations and/or milling capacity are sufficiently powerful in the local political economy to influence the timing and location of small-scale plantations they do not own. The reverse is not true: large-scale expansion is not influenced by small-capital plantation agents. In addition, plantation agents differ in their own milling capacities. Therefore, some are more attentive to price signals for crude palm oil, while others consider primarily the price signals of fresh fruit bunches. Hence, we conceptualize three types of plantations agents to reflect this essentially heterogeneous context. The first type of plantation agents, *A*, has only agency on its own expansion and expect to sell only FFB. This typically corresponds to smallholders. The second type, *B*, corresponds to industrial plantations without milling capacity: they have agency on the expansion of small plantations they do not own, and their expected revenues depend on FFB sales. The third type, *C*, corresponds to industrial plantations with milling capacity: they have agency on the expansion of small plantations they do not own, and both FFB and CPO prices are relevant to their decision rule (as opportunity cost and output price respectively).

Our parameter of interest is α_{CPO}^j , the CPO price elasticity of deforestation for the plantations of agent type *j*. We focus on CPO as its market is most exposed to national and international policy and demand shocks. Given the supply chain structure, we assume that mills can have market power in their local FFB markets, but not in the globalized CPO market downstream, such that local price shocks are passed-through upwards only. Hence, we decompose α_{CPO}^j into theoretical mechanisms as follows.

$$\alpha_{CPO}^j = (\iota_{FFB}^A + \phi_{FFB}^B) \times \epsilon_{CPO}^{FFB} \times \mathbb{1}[j = A] + \phi_{CPO}^C \times \mathbb{1}[j = A] + \iota_{FFB}^B \times \epsilon_{CPO}^{FFB} \times \mathbb{1}[j = B] + \iota_{FFB}^C \times \epsilon_{CPO}^{FFB} \times \mathbb{1}[j = C] + \iota_{CPO}^C \times \mathbb{1}[j = C] \quad (1)$$

ε_{CPO}^{FFB} is the elasticity of FFB price to CPO price in the local FFB market (i.e., common to all agents in this market) — or price pass-through. ι parameters represent plantation agents' own elasticities, i.e., price-driven decisions to deforest for their own plantations. These coefficients may be positive or negative, depending whether price signals encourage investments at the extensive or intensive margins respectively (which depends on the relative costs of land, capital and labour).⁹ ϕ parameters represent type *B* or *C* plantation agent's decisions to deforest for small-scale plantations they do not own. We hypothesize that ϕ_{FFB}^B is negative, as large-scale plantations compete for land with small-scale plantations, and tend to evict them when price signals are high or release land when they are low. ϕ_{CPO}^C is the influence of type *C* plantation agents over deforestation for small-scale plantations, in reaction to a CPO price change. It may be driven by an eviction mechanism similar to ϕ_{FFB}^B , but it may also be driven by the need for a larger supply base when signals of future profit from CPO sales increase.¹⁰ Thus, we do not hypothesize the sign of ϕ_{CPO}^C .

3 Data, measurements and sampling

3.1 Deforestation for oil palm plantations

Conceptually, our outcome of interest is the clearing of forest land for the purpose of growing oil palms, resulting from the decisions of plantation agents. For brevity, we call this quantity *deforestation*.

To measure deforestation annually, 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 a map of eventual oil palm plantations (Petersen et al. 2016; Austin et al. 2017). These remote sensing products and how we combine them are described in more detail in Appendix D.2.¹¹

Here, we explain how we measure the extent of conceptually and politically relevant sub-types of deforestation in the available data.

⁹ ι_{FFB}^C represents the opportunity cost of milling, instead of selling, palm fruits.

¹⁰Moreover, agents with milling capacity have interest in maximizing fruit production from small-scale plantations in their supply shed as a way to depress local FFB prices and access cheap adjustment supply.

¹¹Alternatively, we use a measure of secondary forest deforestation, described in Section D.2. Besides, we note 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).

Plantation sites. We do not observe the true boundaries of the places affected by the decisions of individual plantation agents. To approximate these boundaries empirically, we arbitrarily delineate *plantations sites* according to a trade-off between efficiency loss and precision inflation.¹² We choose our typical plantation sites to be 3×3 km (900ha) grid cells, as in (Busch et al. 2015).

Industrial and smallholder plantations. Within plantation sites, we can observe where oil palms are effectively grown. Moreover, we distinguish whether the size and shape of palm tree patches correspond to industrial or to smallholder plantations, using remote sensing products from Austin et al. (2017) and Petersen et al. (2016) respectively.¹³ Hence, we measure industrial plantations as large rectangular grid landscapes and smallholder plantations as small- and mid-sized plantation mosaic landscapes (palm tree patches smaller than 100ha, comprising at least 50% of a mosaic landscape wider than 100ha (Petersen et al. 2016)). According to the literature referred to in Section 2.1, this latter measurement captures smallholder plantations broadly characterized by informality, spot-market sales and incomplete power over expansion decisions¹⁴ — i.e., plantation agent type *A* in our conceptual framework. On the other hand, the industrial plantation measurement does not allow to capture plantation agent types *B* and *C* distinctly.

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 deforestation price elasticity.¹⁵

The best available maps of oil palm concessions and Indonesian legal land designation are

¹²This is a standard problem of spatial analysis (Atkinson et al. 2022). In addition, here, we wish to keep computation times reasonable. We further mitigate precision inflation with our inferential clustering strategy explained in Section 4.1.

¹³Where these maps overlap, we characterize plantations as industrial, since remote sensing for this landscape is less error-prone.

¹⁴In particular, it excludes formalized “plasma” plantations, jointly developed near and alike industrial plantations wider than 250ha as part of a smallholder scheme (Paoli et al. 2013; Byerlee et al. 2016; Gaveau et al. 2022).

¹⁵Indeed, we do not seek to gauge total illegal deforestation, but rather to estimate the price elasticity for a representative sample of illegal deforestation.

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.¹⁶

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 snapshot 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 2020¹⁷ to identify illegal deforestation shows that our results are robust to concession issuance post-2010 (Table A.1). 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.¹⁸ 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.

¹⁶Any 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 Cadangan. 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).

¹⁷(*Greenpeace Kepo Hutan Public Downloads - Google Drive* 2023)

¹⁸Revocation 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).

Immediate and transitional conversion. We measure deforestation as annual primary forest loss events occurring within the plantation extent and prior to the first year palms are detected (2015 and 2014 for industrial and smallholder plantations respectively). For this to reflect our 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.¹⁹ 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).

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 oil palms are 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).

3.2 Price signals: a model of plantation-mill relationships

Conceptually, our treatments of interest are the fresh fruit bunches (FFB) and crude palm oil (CPO) price signals perceived by plantation agents. Here, we explain how we measure these, with primary and secondary data, based on a model of the influence of mills on plantation sites.

A primary microeconomic panel of geolocalized palm oil mills. We semi-manually merge two existing data sets - the Indonesian manufacturing census (IBS) and the Universal Mill List (UML) - to produce an original microeconomic data set of geo-referenced palm oil mills in Indonesia from 1998 to 2015. In Appendix D.1 we describe the IBS and the UML 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

¹⁹Both plantation data sets recognize areas with signs of future cultivation as oil palm 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.

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.2 shows primary 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 Indonesian manufacturing census. With this primary data set, we know the annually averaged prices offered at mills' gates for fresh fruit bunches as well as those faced by mills for crude palm oil.²⁰ In the following paragraphs, we describe how we measure price signals for either commodity.

The set of reachable mills. The quantity of oil derived from a tonne of palm fruit increases with the quality of the fruits and thus decreases with the duration between harvest and milling (Byerlee et al. 2016). This constraint leads to spatial proximity between plantation sites and mills.

For every plantation site, we determine a set of reachable mills for every year. Mills are considered reachable if they are within an euclidean distance from the plantation site, which we call catchment radius parameter.²¹ We assume that freshly harvested palm fruits can potentially be brought to any mill within the concentric area without deteriorating too much. Mills beyond the catchment radius are not reachable and thus we assume their prices have no influence on the decisions to deforest in the plantation site.

In this study, our preferred catchment radius is 30 km in Sumatra and 50 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 plantation sites 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 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

²⁰Precisely, we measure annual prices as mean unitary values.

²¹In Appendix C we compare results under the euclidean model with results from a travel time model.

mills located farther than 30 km away.²² 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 replaced with oil palm plantations (either industrial, in 2015, or smallholder plantations, in 2014). The first column on the left describes the sample of plantation sites that have non-missing price signal information and that we hence use in main estimations. Sample mills are the 587 palm oil processing plants from the Indonesian manufacturing census that we have geo-localized.

Differential mill influence. From a plantation site several mills can be reached, but not all mills are equally influential. Plantation agents may expect different sales plans, from selling to one mill exclusively, to partial off-take agreements with some of the reachable mills, to selling on the local spot market only. We do not directly observe how prices paid at all reachable mills enter the price signals that are observed by agents from all plantation sites. Therefore, we model the differential influence of the prices at reachable mills using plantation-mill straight-line distances. More precisely, we model the price signal as the standardized invert-distance weighted average of prices at annually reachable mills.²³ Hence, every plantation site 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 site i in year t is:

$$S_{it} = \sum_m \frac{d_{im}^{-1}}{\sum_n M_{in} d_{in}^{-1}} * P_{mt} \quad (2)$$

²²The 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 palms are farther than 30 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 palms are within 10 km of a mill, and 8.1% are farther than 50 km. The Center for International Forestry Research (CIFOR), in its online atlas (<https://atlas.cifor.org/borneo/#en>) applies a 10 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 km to a mill.

²³This is explained here for the price signal, our explanatory variable of interest, but the same method is applied to any mill-level covariate.

This models the relative influence of reachable mills in a way that is consistent over the unobserved heterogeneity in the expected sales plans of plantation agents. Indeed, the standardized inverse-distance weights approximate the expectable, relative transport costs within a set of reachable mills (including constant fuel costs and fruit quality decline).²⁴

Medium-run expectations and price signals. Oil palm trees become productive three years after planting and can yield fruits for two decades onward. Therefore, prospective plantation agents have to form price expectations far ahead into the future. How we model these expectations grounds on three distinct assumptions.

First, we assume that plantation agents 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 3.1). 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 plantation agents 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 plantation agents 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 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 plantation agents 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.

²⁴This grounds on a Von Thünenian rationale, according to which prices at mills relatively farther away are less influential, because reaching these sale points is more costly. Note that 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. See below for both points.

²⁵In addition, due to price data unavailability before 1998, assuming that 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.

Third, we assume that each annual observation is equally informative. In Table A.3, 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 5, 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 prices. We call this the medium-run price signal.

$$Price_{it} = \frac{1}{4} \sum_{l=3}^0 S_{it-l} \quad (3)$$

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

3.3 Sampling

Our sample is an annual unbalanced panel of 3×3 km 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. We note that it is common that mills need a minimal fruit supply basis to operate. This implies that at usual mill capacity and plantation yield, a minimum plantation size of ca. 3000 hectares is developed alongside any new mill opening (Paoli et al. 2013). Because of the lag between planting and harvesting, deforestation

²⁶We 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.

²⁷Precisely, 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.

for these integrated plantations occurs before the mill starts operating. Such deforestation does probably not occur because of price signals from reachable mills already operating. In a robustness exercise, in Table A.4 we show results following a sampling that excludes plantation sites in the direct vicinity (3000 hectares) of a mill before it starts operating.

Second, we remove plantation sites within RSPO certified concessions.²⁸ Indeed, we expect both price and deforestation dynamics to differ systematically in these plantation sites, including before certification procurement.

Third, we remove annual records as soon as one of the variables in Equation 4 (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 selection 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.

3.4 Descriptive statistics

Table 2 provides descriptive statistics for the final sample used in the main estimation (in Table A.2, we provide mill-level descriptive statistics). Deforestation is a count of 30m pixel-level events of primary forest loss eventually replaced by oil palms. In Table 2, it is converted to

²⁸RSPO stands for the Roundtable on Sustainable Palm Oil, the main voluntary certification scheme for palm oil. Few certifications were issued in the first years of the RSPO, from 2009 to 2014. Thus, this sampling exclusion applies to very few observations during our study period. We observe certified areas using data from Carlson et al. (2018).

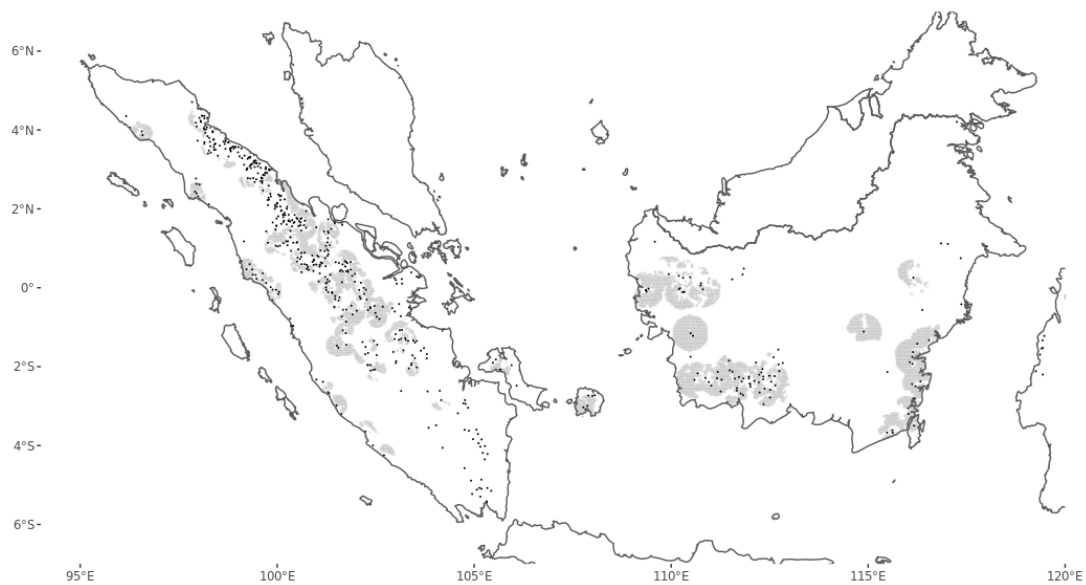


Figure 1: Study samples of palm oil mills and plantations

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 (USD/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 (3×3 km 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.

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.5 and A.6, 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

The conceptual framework outlined above can be summarized in a reduced-form relationship between deforestation on the left-hand side, and the crude palm oil (CPO) price signal and a structural error term on the right-hand side.²⁹ We model this reduced-form as follows:

$$Deforestation_{idt}^k = \exp(\alpha^k \ln(Price_{idt}) + \beta^k X_{idt} + \lambda_{idt}^k + \gamma_{dt}^k + e_{idt}^k) \quad (4)$$

$Deforestation_{idt}^k$ is deforestation for $k = \text{industrial, smallholder or all plantations}$, in plantation site i , in district d , from 2002 to 2014 ($t = 1, \dots, 13$). $Price_{idt}$ is a measure of the price signal (constructed with equations 2 and 3) observed from plantation site i in district d in year t . In our main specification, $Price_{idt} = Price_{idt}^{CPO}$, and thus $\alpha^k = \alpha_{CPO}^k$ is the crude palm oil price elasticity of deforestation for smallholder ($k = A$), industrial ($k = \{B, C\}$), or all ($k = \{A, B, C\}$) plantations. α_{CPO}^k is a reduced-form parameter with respect to the more structural elasticities in the right hand

²⁹Which comprises all the other determinants of deforestation — probably including perceived information on investment costs (e.g., of land acquisition and conversion), discount rates, operating costs (e.g., of labor, energy and fertilizers), institutional costs (either fixed or marginal, positive or negative, formal or not), opportunity costs, and attainable yields.

side of Equation 1. In alternative specifications, we add the price signal of fresh fruit bunches (FFB) in $Price_{idt}$ to document some of these mechanisms.

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. We decompose the other determinants of deforestation into heterogeneity sources that can be either specific to a local fresh fruit bunch market (local market fixed effects, λ^k_{idt} , indexed by t as the local market may evolve, although not necessarily every year); annual shocks common to a whole district (district-year fixed effects, γ^k_{dt}); or error terms, e^k_{idt} . We elaborate on observed and unobserved heterogeneity in the next subsection.

$Deforestation^k_{idt}$ is a count of non-negative integers (deforestation pixel-events) 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 4 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 may be spurious because of omitted variable bias, reverse causality or 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. Prices and deforestation may also have an underlying driver, the omission of which would bias a 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 and Loboda 2019) and to roads (Hughes 2018); the decentralization of authority on land (The Gecko Project³¹ 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.

Identifying variation in a shift-share reduced-form design. Absent a sharp experimental framework, we reflect on the observational variation at hand to make explicit under what assumptions we can mitigate these threats to the identification of the causal parameter α in the data. 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 crude palm oil (CPO) prices. Hence, the price

³⁰Such 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.

³¹<https://thegeckoproject.org/>

signal is a shift-share instrument and Equation 4 an instrumental reduced-form. Thus, for $\hat{\alpha}^k$ to identify the price elasticity, it is sufficient that the mill-gate CPO prices (the shifts) be conditionally exogenous to plantation-mill distance weights (the shares) and to deforestation in plantation sites (Borusyak et al. 2022). This condition can hold because mill-gate CPO prices are partially driven by two kinds of idiosyncratic deals with downstream buyers. First, mills make off-take agreements with traders and refineries, typically for periods of three to nine months (Wiggs et al. 2020). Given the concentration of the trading and refining level of the supply chain, the pricing conditions in such agreements are most likely unaffected by mills' own or upstream characteristics. Second, traders and refineries make opportunistic and speculative purchases to mills, to take advantage of downstream market conditions in the short-run (Pirard et al. 2020; Wiggs et al. 2020). The pricing conditions in such deals are essentially exogenous to deforestation. The corporate ties between mills, traders and refineries are still mostly hidden from the public and the terms of their transactions are even more so. This makes it impossible to effectively track these sources of idiosyncratic variation in mill-gate prices. Instead, we endeavor to isolate this identifying variation from potential confounders.

Omitted-variable bias. First, we include district-year fixed effects in the main specification (equation 4), effectively removing variation common to observations in the same year and district. The remaining variation in price signals can be interpreted as departures from the district average. 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. These price departures are independent from macro-drivers of prices on markets larger than a district, and that could spuriously correlate with macro-level drivers of deforestation.³² Because departures are annual, they are further independent from district-specific dynamics (including slower-than-annual 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

³²For 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).

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 in the same local FFB market (i.e. that can reach and sell to the same set of mills). This local-market fixed effect isolates our identifying variation from systematic differences in local determinants of price departures that would endogenously correlate with deforestation. Such differences at this level include in particular: i) agro-climatic conditions (typically constant in time and common to larger areas than a plantation site); 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 Given the high level of standardization and homogeneity of the crude palm oil commodity (Byerlee et al. 2016), we can confidently assume that the local market fixed-effect absorbs differences in FFB quality that are sufficient to explain differences in CPO quality (and price, in fine).

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. 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 (see Kraus et al. (2022) for a detailed discussion of the drivers thereof).³⁵

Overall, this specification effectively imposes comparisons in deforestation departures from the district average, between plantation sites in the same local FFB market, that only differ in

³³Indeed, 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.

³⁴On 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. 2022). 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 a fixed effect at the plantation site level.

³⁵Moreover, adding a control for past deforestation in the last 4 years to mitigate the endogeneity of mill density and deforestation does not change the results (see Appendix C).

their relative distances to mills facing different price departures (from the district average).

Reverse causality. Even within a local market, shocks (i.e., annual departures from district average) in FFB volumes or prices may affect CPO prices. Such shocks may notably arise if plantation agents of types *B* (independent industrial plantations) or *C* (integrated industrial plantations) locally contain expansion (and deforestation) in order to realize higher margins (Maryadi et al. 2004; Masliani et al. 2014)). This would open a confounding reverse causal path from deforestation to CPO prices. However, mills sell CPO on a large global market where they are price-taker.³⁶ Thus, we are confident in assuming that mills cannot raise the CPO price they offer to accommodate a local input price surge, and that they cannot obtain better CPO prices by shrinking supplied (and thus sourced) volumes.³⁷ In Section C we check our results' robustness to this assumption. We find similar results when controlling for temporally, and for temporally and spatially, lagged deforestation (i.e., when blocking the causal channel through which local FFB supply shocks associated with deforestation could reversely impact CPO prices).

Finally, one might be concerned about local shocks in milling marginal costs being endogenous to both deforestation and CPO price departures (through economies of scale for instance). We dispel this concern by observing that the crude palm oil market is not at equilibrium.³⁸ This implies indeed that mills can sell all their production at the prices they face, and thus do not transform productivity gains into more competitive prices.

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: 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.2 shows that there is no systematic

³⁶The average mill in our data (Table A.2) produces roughly one thousandth of global supply in the period.

³⁷Moreover, we note that supply chain initiatives like the RSPO or NDPE (No Deforestation No Peat No Exploitation) were implemented during our study period. Thus, lower deforestation was not yet associated with potential price premiums for sustainable palm oil.

³⁸See Box 2.1. in Byerlee et al. (2016) for an analysis of mill profitability.

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.

Identification of elasticity to fresh fruit bunches price signals. In secondary analyses, we add the price signal of fresh fruit bunches (FFB) to the main specification to infer additional insights from comparisons with the main crude palm oil (CPO) price elasticity. This aims at disentangling some of the mechanisms featured in Equation 1, in the conceptual framework presented in Section 2.2. This requires to assume a non-negative pass-through from CPO to FFB mill-gate prices, which is reasonable according to our understanding of the palm oil value chain, detailed in Section 2.1.³⁹

Yet, the identification strategy presented so far for the CPO price signal treatment is not readily portable to FFB price signals. Here, we discuss the additional sources of endogeneity in specifications of Equation 4 where $Price_{idt}$ includes the price signal of FFB.

First, the price pass-through allows to consider mill-gate FFB price departures as an intermediate outcome of CPO's. Therefore, in the specification with $Price_{idt} = [Price_{idt}^{FFB} Price_{idt}^{CPO}]$ we do not seek to interpret α_{CPO}^k the coefficient on the CPO price signal conditional on the FFB price signal. Indeed, conditioning on the FFB price signal is very likely to introduce a collider bias in α_{CPO}^k , that we do not deem reasonable to attempt interpreting. However, we reflect on the possible interpretation of the estimated coefficient on the FFB price signal itself.

We consider two main sources of endogeneity, conditional on the fixed-effects and controls discussed above and conditional on the CPO price signal — an obvious potential confounder in this case.⁴⁰ First, local shocks in FFB quality may drive both deforestation and FFB price departures. The bias from such omitted variable would most likely be positive. Thus, our FFB price elasticity estimate is to be interpreted as an upper bound. Second, independent industrial plantations' may have local oligopolistic power and hence contain expansion (including into forest) to maintain high FFB prices. This can introduce a negative reverse causality bias. Thus, our FFB price elasticity estimate is to be interpreted as a lower bound for industrial plantations.

³⁹This pass-through cannot readily be estimated in our data, not the least because finding a quasi-experimental setting at the mill-level, and modelling local mill monopsones is not trivial.

⁴⁰Because of the price pass-through and of the potential direct influence of mills on deforestation for smallholder plantations, θ_{CPO}^C in Equation 1.

5 Results

Table 3 shows our estimates of the crude palm oil price elasticity of deforestation across 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 reacts positively to crude palm oil price signals. Going forward, we show results for unregulated deforestation specifically; then we show results on the mechanisms of smallholder expansion elasticity; finally we show heterogeneities and magnitudes — discussing the external validity of our results. We discuss policy implications in the conclusion. In Appendix C, we conduct a robustness analysis summarized in the specification chart presented in Figure B.1. We compare our estimates to existing ones in Appendix F.

Table 3: **CPO price elasticity 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 crude palm oil (CPO) 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 CPO 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.

5.1 The price elasticities of unregulated deforestation

Industrial and smallholder plantations. We find that a 1% increase in crude palm oil price signals increases deforestation for industrial plantations by 1.8% and deforestation for smallholder plantations by 1.4%.

Assuming a non-negative pass-through from crude palm oil (CPO) prices to fresh fruit bunch (FFB) prices, the positive CPO price elasticity of deforestation for smallholder plantations indicates that smallholders are price elastic and/or that more influential plantation agents (with or without milling capacity) steer deforestation for smallholder plantations in the same direction as CPO price signals (as detailed in the first line of Equation 1). In Section 5.2, we disentangle these potential mechanisms where possible.

The positive CPO price elasticity of deforestation for industrial plantations indicates that corporate actors of the oil palm sector do engage in large-scale deforestation as CPO prices increase. In particular, this suggests that medium-run price signals (over 4 years here) influence

deforestation for palm oil production in the long-run. In Section 5.3, we provide more insights into the dynamics of the price elasticity.

The price elasticities of deforestation for industrial and for smallholder plantations are not significantly different. Yet, looking specifically at illegal deforestation, the point estimate for industrial plantations is more than twice as large as for smallholder plantations and this difference is significant at the 95% confidence level (Table A.7).

Legal and illegal deforestation. We further break down the estimation according to the legal status of deforestation.⁴¹ We find close to zero effects of price signals on legal deforestation for either industrial or smallholder plantations. On the other hand, illegal deforestation appears to be price elastic for both plantation types. Overall, the price elasticity of illegal deforestation is 3. Illegal deforestation for industrial and for smallholder plantations increases by 5.3% and 1.8% respectively, when crude palm oil price signals increase by 1%. 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 large number of observations and clusters (sets of reachable mills), we believe that it is rather truly inelastic to prices.

Table A.7 documents that the differences in the price elasticities of legal and illegal deforestation are statistically significant (except for smallholders, for which this distinction is less essential). The estimated price elasticity of illegal deforestation is also of larger magnitude than that of all deforestation (legal, illegal and unknown combined), especially for industrial plantations. Altogether, these findings about legal and illegal deforestation indicate that, across plantation types, positive price elasticity is driven by illegal deforestation. This suggests that economic opportunities encourage plantations to circumvent land use regulations. This is most likely the result of two factors. First, the processes to obtain the legal green lights to clear the forest and plant palm trees takes several years (to count on top of the years between planting and harvesting), a deforestation permit and a plantation license are lengthy.⁴² Plantations probably rely on more stable signals than those that we capture in this study to formulate

⁴¹For smallholder plantations, we show results disaggregated by legal status although this distinction is not essential — see Section 3.

⁴²See Paoli et al. (2013) for more detail on the licensing process, involving, for instance, measuring environmental suitability and community consultation.

long-term expectations about the profitability of engaging today in legal deforestation. Second, formal procedures for the conversion of forest to oil palm are relatively easy to bypass (Setiawan et al. 2016).

5.2 Mechanisms of smallholder expansion elasticity

Here, we add the price signal of fresh fruit bunches (FFB) to the main specification and we infer additional insights from comparisons with the main crude palm oil (CPO) price elasticity presented above, under the assumption that the pass-through from CPO to FFB prices is non-negative. This aims at disentangling some of the mechanisms featured in Equation 1, in the conceptual framework presented in Section 2.2.

Table 4 shows our estimates of the coefficient on FFB price signals for smallholder plantations.⁴³ Since smallholder plantations are unlikely to have even local oligopolistic power, we can interpret the price elasticity estimate as an upper bound. Table 4 shows that this estimate is negative, in particular after controlling for the confounding variation in CPO prices. Hence, we infer that the effect of FFB price signals on deforestation for smallholder plantations is negative. In our conceptual framework, this means that in Equation 1, $\iota_{FFB}^A + \phi_{FFB}^B < 0$. We infer two conclusions from this negative upper bound result.

First, together with the former result that CPO price signals accelerate deforestation for smallholder plantations (Table 3: $\hat{\alpha}_{CPO}^A = 1.37$), the negative effect of FFB price signals allows to infer that when CPO prices increase, mills react by accelerating deforestation for smallholder plantations ($\phi_{CPO}^C > 0$). This actually more than compensates the reduction in deforestation for smallholder plantations caused by the passed-through increase in FFB price signals. This over-compensation indicates that the price pass-through (assumed non-negative) is weak relative to mills' direct influence.

Second, either or both of smallholders' own deforestation elasticity to FFB prices (ι_{FFB}^A) and of the eviction effect from independent industrial plantations (ϕ_{FFB}^B) are negative. To document which of these mechanisms primarily drives the negative effect, we additionally control for the immediate conversion of forest to industrial plantations, as a proxy for independent industrial plantation expansion. In addition, we also control for the interaction of this proxy with the FFB price signal. The two right-most columns of Table 4 show that the effect of FFB price signals

⁴³We do not estimate a similar coefficient for industrial plantations, because, as explained in Section 4.2, potential biases can go in either direction in this case.

on deforestation for smallholder plantations remains negative after the eviction effect (ϕ_{FFB}^B) is absorbed away. Thus, we conclude that smallholder agents have a negative price elasticity in terms of their own deforestation decisions and observed prices ($i_{FFB}^A < 0$).

Table 4: **FFB price elasticity of deforestation for smallholder plantations**

FFB price signal				
Estimate	-2.02	-1.92	-1.94	-1.97
95% CI	[-3.61; -0.43]	[-3.48; -0.36]	[-3.51; -0.37]	[-3.54; -0.39]
CPO price signal				
Estimate		1.71	1.71	1.71
95% CI		[-0.14; 3.57]	[-0.14; 3.57]	[-0.15; 3.56]
Additional controls				
Industrial expansion proxy				Yes
... interacted with FFB price			Yes	Yes
Observations	18207	15880	15880	15880
Clusters	501	484	484	484

NOTE. This table shows our estimates of the palm fruit (FFB) price signal partial effects on deforestation for smallholder plantations. They are to be interpreted as points of percentage change in average deforestation associated with a 1% increase in price signal. We show partial effect estimates for the crude palm oil (CPO) price signal for informative purpose only, as they only have a control role here. The price signal is measured as the 4-year average of annual inverse-distance weighted averages of either FFB or CPO prices at the gates of reachable mills. Deforestation is measured as primary forest loss eventually replaced with smallholder oil palm plantations. 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 Further dynamics and heterogeneity

Immediate and transitional conversion to oil palm. 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 5 shows our estimates of the price elasticity of immediate and transitional deforestation for industrial plantations, again distinguishing legal, illegal, and overall (legal, illegal and unknown legal status) conversion. We show results for immediate conversion over the whole period and over 2002-2010, the sub-period when transitional conversion can be observed. Over the whole period, the price elasticities of immediate conversion 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.

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.

Table 5: **CPO price elasticity of deforestation for industrial plantations, by conversion dynamics**

	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.

Expectation formation: short- and medium-run price signals. Here, we estimate the partial effects of the short-run price signal, an alternative medium-run price signal, and their interaction. The short-run price signal is S_{it} from Equation 2, where t is the year deforestation occurs. The alternative medium-run price signal runs over the preceding 3 years and does not include the contemporaneous price signal S_{it} .

Table 6 displays the partial effects of short-run price signals unconditional and conditional on the alternative medium-run price signals. All models are based on the same specifications as the main one, from Equation 4.

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 smallholder plantations). 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 (or vice-versa). In Table A.3, we show elasticities to annual price signals, conditional on the other annual price signal lags. This reveals that deforestation is 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 CPO price elasticity of deforestation

	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 crude palm oil (CPO) 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 CPO 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.

Price variability. We here report on deforestation responsiveness to crude palm oil price variability, rather than level (but conditional on level). We find that the effect of a standard deviation in the medium-run price signal is generally not distinguishable from zero. (Table A.8). However, this hides notable patterns: price variability deters legal deforestation (for both industrial and smallholder plantations) and deforestation for industrial plantations in general. This is in line with our interpretation of the results on legal and illegal deforestation discussed above. Lundberg and Abman (2022) demonstrate that this may reflect a positive option value of unprotected forests (they find similar results in a non-subsistence smallholder context).

Spatial heterogeneity. We estimate the price elasticities of deforestation for Sumatra and Kalimantan separately. Table A.9 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 our main model on a measure of deforestation in secondary forest only. Table A.10 shows that deforestation in such forests is generally not price elastic. We see two non-exclusive potential explanations for this absence of effect.⁴⁴ 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.

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

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.11). Finally, we note that our estimates mainly capture effects on deforestation at the intensive margin, i.e., occurring after at least one mill opened.

⁴⁴Given the large number of observations and clusters, we do not attribute it to a lack of statistical power.

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 crude palm oil (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 $tCO_2 ha^{-1}$ emissions due to deforestation (Guillaume et al. 2018).⁴⁵ CO₂ revenues are based on the USD 5/tCO₂ agreed price Norway paid to Indonesia for its recently avoided deforestation.⁴⁶

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.⁴⁷ In the presence of a result-based payment scheme, this represents a yearly oppor-

⁴⁵We 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.

⁴⁶<https://www.regjeringen.no/en/aktuelt/noreg-betaler-530-millionar-for-redusert-avskoging-i-indonesia/id2722135/>

⁴⁷We compute standard deviations in our price signal regressor variable, in the estimating sample, after removing variations in fixed-effect dimensions (Mummolo and Peterson 2018).

tunity cost of USD 32.3 million. To curb annual deforestation 29% below the 2002-2014 average with price incentives alone, price signals for individual plantations should be lowered by 19%.⁴⁸ This would save 38.5kha of primary forest annually, corresponding to revenues from a potential result-based payment scheme of USD 123 million.

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 (million USD)	-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 CO₂ revenues correspond to result-based payments paid at a price of USD 5 per tCO₂ avoided, assuming average emissions of 174tC per hectare deforested.

⁴⁸Aligning 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 Discussion and conclusion

To conclude, we summarize our main findings, we state the main methodological limitations and we propose further research avenues. Finally, we discuss the policy relevance of our results.

Summary. 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. Unregulated deforestation (for smallholder plantations and for illegal industrial plantations) reacts positively to crude palm oil (CPO) prices, while legal deforestation is inelastic to prices. On the other hand, fresh fruit bunch (FFB) price signals have a negative effect on deforestation for smallholder plantations.

Limitations. Our country-wide deforestation-price elasticity estimates advance the existing literature in terms of internal validity and of analyses of heterogeneity and mechanisms. Yet, they necessarily rely on observational data, which prevents us from ascertaining that our estimates exactly identify the causal price elasticity parameters of interest.

In terms of heterogeneity, investigating the empirics of informal or illegal economic behaviours, while being important in this context, is essentially prone to measurement error. Our measurement of plantations, in particular smallholder ones, does not allow us to look further into the local diversity of politically relevant situations. In Section 3.1 we explain our efforts to minimize the error in characterizing illegal deforestation. Yet, our results that legal deforestation is not price elastic and that the price elasticity differs between legal and illegal deforestation may be misleading, if those deforestation events deemed legal actually reflect a specific form of illegality that could not be captured by our measurement.

In terms of mechanisms, we are not able to disentangle the price elasticities of deforestation for industrial plantations with and without milling capacity. Moreover, we are not able to disentangle whether the negative effect of FFB price signals on deforestation for smallholder plantations is due to both or either an eviction effect from industrial plantations or smallholders' own elasticity being negative.

Regarding external validity, 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 (Hsiao 2022; Kraus et al. 2022). Yet, the large structural gap between installed milling capacity and actual CPO production (Pirard et al. 2020) indicates a significant risk of deforestation at the intensive margin (where milling capacity is already installed), for which our estimates are valid. 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.

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.

6.1 Policy Relevance

There are three main implications from our results that are relevant to sustainable and fair policy in the Indonesian palm oil context.

Deforestation left unregulated will follow prices. First, extrapolating our general result to the last decade suggests that the downward trend in palm oil prices is at least partly responsible for the decrease in deforestation rates. As a corollary, a surge in price signals in the future could overtake supply-side conservation efforts. This is especially likely as unregulated deforestation — for illegal and for smallholder plantations — remains weakly addressed by existing institutionally-reliant interventions (Heilmayr et al. 2020; Drost et al. 2021; Groom et al. 2022).⁴⁹ Our results indicate that this is all the more likely that industrial plantations just by-pass regu-

⁴⁹Adding incentives to the conservation policy toolkit has been advocated in the Amazonian context for similar reasons (Godar et al. 2014).

lations to seize economic opportunities signaled in the medium-run and otherwise unseizable through legal expansion.⁵⁰

Feasible price interventions can steer unregulated deforestation. Since maintaining low prices through international demand requires challenging coordination (Hsiao 2022), upstream price intervention is a promising policy option to effectively govern deforestation. Our results confirm the potential effectiveness of such intervention, as they show that deforestation — for unregulated plantations in particular — decelerates in reaction to lower mill-gate crude palm oil prices. Our back-of-the-envelope estimation indicates that a 19% tax levied uniformly on palm oil mills could curb deforestation 29% under the critically high average of 2002-2014.⁵¹ This would credit Indonesia a minimum of USD 120 million a year for avoided emissions. In addition, we find that a uniform mill-tax would only curb illegal (and not legal) deforestation, thus possibly fostering structural institutional change in Indonesian palm oil landscapes.

Yet, the effectiveness and equity of a mill-tax on crude palm oil would vary with its design. In a generic context alike the present one, Heine et al. (2020) discuss extensively how such tax could be redistributed to incentivize conservation even further.⁵² In particular, they suggest targeting choke points in the value chain, much alike palm oil mills, where the state has sufficient capacity to levy a tax and that cannot be circumvented by upstream producers. They propose that the tax be refunded against proof of sustainable production. Reversing the bur-

⁵⁰Currently, an export tax applies to crude palm oil and its revenues are meant to support rural development. Our results suggest that the export tax has contributed to avoid illegal deforestation. Moreover, its stabilizing effect may have increased the demand for legal deforestation, according to our results. However, the effect of the export tax on crude palm oil price signals is not clear, for at least three reasons: the tax varies with international prices; the domestic demand for crude palm oil has increased; and the tax embeds no conservation incentives.

⁵¹General-equilibrium effects could modulate the effect of such a 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 palm oil only may also achieve lower emission reductions than we estimate if forest is left vulnerable to uncontrolled production of another commodity. Finally, our results suggest that a tax would disincentivize illegal deforestation more than legal deforestation, and that if it stabilized prices the tax could even encourage legal deforestation. The effect of the tax therefore depends on the regulatory framework and its capacity to handle the additional demand for legal deforestation permits.

⁵²Indeed, linking the actual tax level to conservation outcomes strengthens the incentives provided by a tax beyond the price elasticity we estimate here based only on price variations unrelated with any conservation commitment. For the same reason, our results cannot serve to evaluate the price incentives provided by the Roundtable on Sustainable Palm Oil (RSPO). Yet, it seems to us that given our elasticity estimates, the RSPO premiums — ranging from 2% (Levin 2012) to 7% (Preusser 2015) — are insufficient to promptly reach zero-deforestation palm oil.

den of the proof accommodates possibly weak monitoring capacity of the tax authority. This effectively enlists the entity at the choke point as a “voluntary private enforcer” of the state’s desired policy (Heine et al. 2020). Our results indicate that palm oil mills could serve as such entities, as they steer deforestation, in particular for illegal and for smallholder plantations, upon price changes. More generally, structural change — like smallholder formalization — in remote and unregulated segments of the palm oil supply chain may be governed through such feebate targeted at the locally potent palm oil mill (Heine et al. 2020).

Protecting smallholder revenues further reduces deforestation. Regarding equity, how revenues from the tax, and possibly from international compensation for avoided emissions, are redistributed to land and palm oil-reliant communities (including smallholders) is critical. Our main results indicate that a tax on crude palm oil would lead palm oil mills to restrict forest clearing for smallholder plantations. While this is a welcome impact from a gross environmental perspective, it is not from an environmental justice perspective, nor from an economic development one. Importantly, our results are consistent with the proposition that smallholders compensate a loss in palm fruit revenues by expanding further their plantations into the forest. Notably, such a reaction to a drop in fresh fruit prices induced by the tax would be more than compensated by the influence of palm oil mills. These results imply that both the effectiveness and equity of a tax on crude palm oil could be improved by additionally protecting fresh fruit prices or revenues for smallholders.

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*, qjaco38.
- 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.
- Atkinson, Peter M., A. Stein, and C. Jeganathan (2022). "Spatial Sampling, Data Models, Spatial Scale and Ontologies: Interpreting Spatial Statistics and Machine Learning Applied to Satellite Optical Remote Sensing". In: *Spatial Statistics*. Special Issue: The Impact of Spatial Statistics 50, p. 100646.
- 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.

- Berman, Nicolas, Mathieu Couttenier, Antoine Leblois, and Raphael Soubeyran (2023). "Crop Prices and Deforestation in the Tropics". In: *Journal of Environmental Economics and Management* 119, p. 102819.
- Borusyak, Kirill, Peter Hull, and Xavier Jaravel (2022). "Quasi-Experimental Shift-Share Research Designs". In: *The Review of Economic Studies* 89.1, pp. 181–213.
- 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, Oyut Amarjargal, Farzad Taheripour, Kemen G Austin, Rizki Nauli Siregar, Kellee Koenig, and Thomas W Hertel (2022). "Effects of Demand-Side Restrictions on High-Deforestation Palm Oil in Europe on Deforestation and Emissions in Indonesia". In: *Environmental Research Letters* 17.1, p. 014035.
- 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.
- 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.

- Corley, R.H.V. and P.B. Tinker (2015). "The Origin and Development of the Oil Palm Industry". In: *The Oil Palm*. Chichester, UK: John Wiley & Sons, Ltd, pp. 1–29.
- 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.
- Craw, M (2019). *Palm Oil Smallholders and Land-Use Change in Indonesia and Malaysia: Implications for the Draft EU Delegated Act of the Recast Renewable Energy Directive*.
- Dalheimer, Bernhard, Christoph Kubitzka, and Bernhard Brümmer (2022). "Technical Efficiency and Farmland Expansion: Evidence from Oil Palm Smallholders in Indonesia". In: *American Journal of Agricultural Economics* 104.4, pp. 1364–1387.
- Drost, Sarah, Barbara Kuepper, and Matt Piotrowski (2021). "Indonesian Moratoria: Loopholes, Lack of Sanctions Fail to Stop Palm Oil-Linked Deforestation". In: *Chain Reaction Research*.
- 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, 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, Adria 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.
- Glenday, Sky, Yusurum Jagau, Suharno Suharno, and Agnes Safford (2015). "Central Kalimantan's Oil Palm Value Chain: Opportunities for Productivity, Profitability and Sustainability Gains". In:
- Godar, Javier, Toby A. Gardner, E. Jorge Tizado, and Pablo Pacheco (2014). "Actor-Specific Contributions to the Deforestation Slowdown in the Brazilian Amazon". In: *Proceedings of the National Academy of Sciences* 111.43, pp. 15591–15596.

- Gol (2016). "First Nationally Determined Contribution Republic of Indonesia". 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 Krashevskaya, Kukuh Murti Laksono, 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. 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 Khatriwada, 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.

- Hsiao, Allan (2022). "Coordination and Commitment in International Climate Action: Evidence from Palm Oil".
- 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.
- Kraus, Sebastian, Robert Heilmayr, and Nicolas Koch (2022). "Spillovers to Manufacturing Plants from Multi-Million Dollar Plantations: Evidence from the Indonesian Palm Oil Boom". In: Krishna, Vijesh V., Christoph Kubitz, Unai Pascual, and Matin Qaim (2017). "Land Markets, Property Rights, and Deforestation: Insights from Indonesia". In: *World Development* 99, pp. 335–349.
- Lambin, Eric F. and Paul R. Furumo (2023). "Deforestation-Free Commodity Supply Chains: Myth or Reality?" In: *Annual Review of Environment and Resources* 48.1, null.
- 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.
- Lee, Janice Ser Huay, Sinan Abood, Jaboury Ghazoul, Baba Barus, Krystof Obidzinski, and Lian Pin Koh (2014). "Environmental Impacts of Large-Scale Oil Palm Enterprises Exceed That of Smallholdings in Indonesia: Forest Loss from Sumatra's Oil Palm Industry". In: *Conservation Letters* 7.1, pp. 25–33.
- 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.

- Li, Tania Murray (2015). *Social Impacts of Oil Palm in Indonesia: A Gendered Perspective from West Kalimantan*. Center for International Forestry Research (CIFOR).
- Lundberg, Clark and Ryan Abman (2022). "Maize Price Volatility and Deforestation". In: *American Journal of Agricultural Economics* 104.2, pp. 693–716.
- Lyons-White, Joss and Andrew T. Knight (2018). "Palm Oil Supply Chain Complexity Impedes Implementation of Corporate No-Deforestation Commitments". In: *Global Environmental Change* 50, pp. 303–313.
- 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 Fresh 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.
- McCarthy, John F., Piers Gillespie, and Zahari Zen (2012a). "Swimming Upstream: Local Indonesian Production Networks in "Globalized" Palm Oil Production". In: *World Development* 40.3, pp. 555–569.
- McCarthy, John F., Jacqueline A.C. Vel, and Suraya Afiff (2012b). "Trajectories of Land Acquisition and Enclosure: Development Schemes, Virtual Land Grabs, and Green Acquisitions in Indonesia's Outer Islands". In: *Journal of Peasant Studies* 39.2, pp. 521–549.
- 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".

- 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.
- 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.
- Petrenko, Chelsea, Julia Paltseva, and Stephanie Searle (2016). "Ecological Impacts of Palm Oil Expansion in Indonesia". In:
- 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.
- 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.
- Potter, Lesley (2012). "New Transmigration "Paradigm" in Indonesia: Examples from Kalimantan". In: *Asia Pacific Viewpoint* 53, pp. 272–287.
- 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.
- Qaim, Matin, Kibrom T. Sibhatu, Hermanto Siregar, and Ingo Grass (2020). "Environmental, Economic, and Social Consequences of the Oil Palm Boom". In: *Annual Review of Resource Economics* 12.1, pp. 321–344.

- 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 Forest-land 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.
- Setiawan, Eko N., Ahmad Maryudi, Ris H. Purwanto, and Gabriel Lele (2016). "Opposing Interests in the Legalization of Non-Procedural Forest Conversion to Oil Palm in Central Kalimantan, Indonesia". In: *Land Use Policy* 58, pp. 472–481.
- 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.

- Wiggs, Chris, B. Kuepper, M. Piotrowski, Aidenvironment Moulin, Okita Miraningrum, Aidenvironment Steinweg, and Gerard Rijk (2020). "Spot Market Purchases Allow Deforestation-Linked Palm Oil to Enter NDPE Supply Chains". In: *Chain Reaction Research*, p. 11.
- 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.
- zu Ermgassen, Erasmus K. H. J., Mairon G. Bastos Lima, Helen Bellfield, Adeline Dontenville, Toby Gardner, Javier Godar, Robert Heilmayr, Rosa Indenbaum, Tiago N. P. dos Reis, Vivian Ribeiro, Itohan-osa Abu, Zoltan Szantoi, and Patrick Meyfroidt (2022). "Addressing Indirect Sourcing in Zero Deforestation Commodity Supply Chains". In: *Science Advances* 8.17, eabn3132.

Appendix

A Tables

Table A.1: 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 3.1

Table A.2: 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 3.2

Table A.3: 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 5.3

Table A.4: CPO price elasticity of deforestation across Indonesian oil palm plantations, away from direct mill vicinity

	Industrial plantations			Smallholder plantations			All		
	Legal	Illegal	All	Legal	Illegal	All	Legal	Illegal	All
Estimate	0.65	5.34	1.91	1.14	1.81	1.28	0.49	3.11	1.59
95% CI	[-1.16; 2.46]	[2.24; 8.45]	[0.36; 3.46]	[-1.83; 4.11]	[0.63; 3]	[0.06; 2.49]	[-1; 1.98]	[1.44; 4.77]	[0.45; 2.72]
Observations	22921	16955	62995	5669	5612	20028	24799	20482	69359
Clusters	619	446	1130	201	273	521	726	633	1423

NOTE. This table shows estimates of the crude palm oil (CPO) price elasticity of deforestation, excluding deforestation within 3000 hectares around a mill and before it starts operating. 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 CPO 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.

Table A.5: 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 (USD/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.6: 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 (USD/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.7: p-values from equality tests of price elasticities

Ho	All plantations			Industrial plantations	Smallholder plantations
	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

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 5.3

Table A.9: 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 5.3

Table A.10: 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 5.3

Table A.11: 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.

B Figures

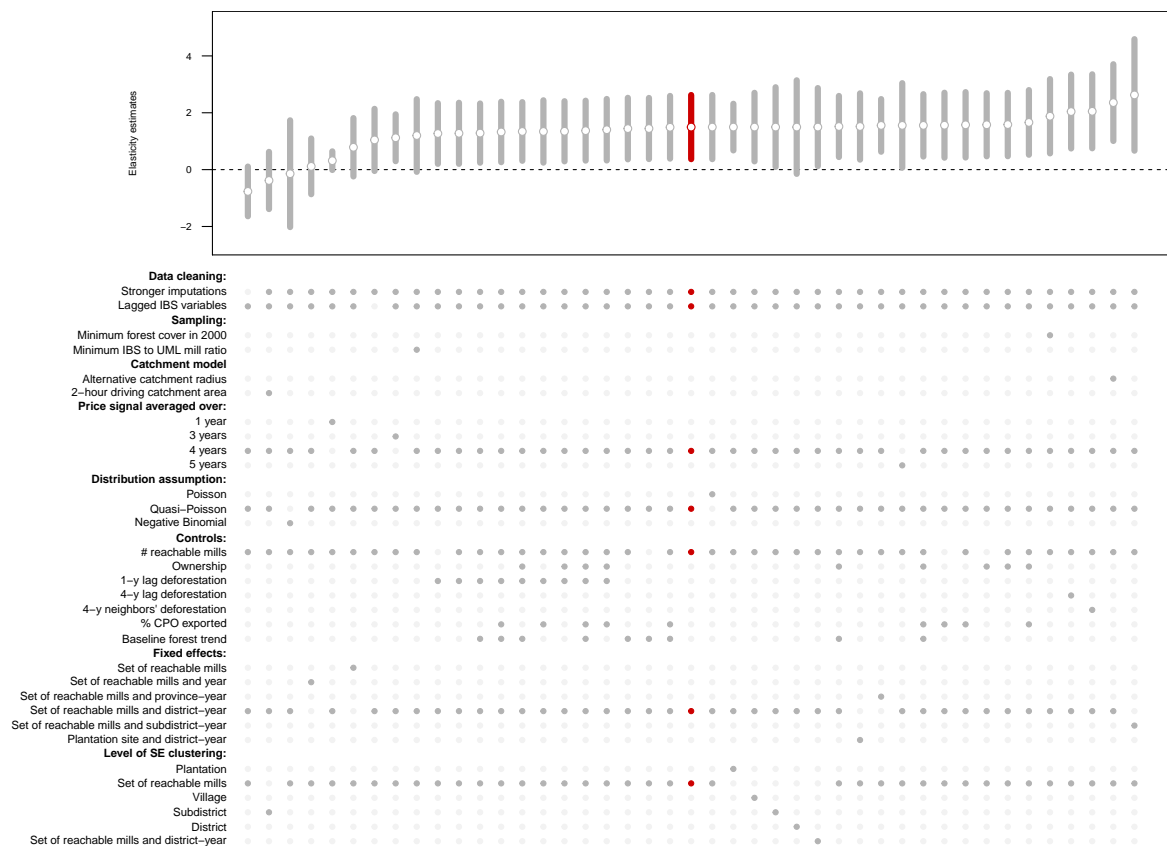


Figure B.1: Estimates of the Indonesian price elasticity of deforestation under different specifications

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

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 4) 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. In our main analysis, we lag IBS variables to correct for a suspected measurement lag between them and the remote sensing forest loss measurement in the outcome variable. Indeed, IBS variables are representative of a whole year, since they are census-based measurements. On the other hand, true forest loss events may not be detected instantly, in particular because of haze and seasonal clouds (Gaveau et al. 2022). Moreover, events could occur at the beginning of the year and be spuriously counted as a decision taken this year. Not taking this into account (i.e., not lagging IBS variables) yields a slightly lower and less precise estimate.

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.⁵³

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.⁵⁴). 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,

⁵³We 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.

⁵⁴Harahap 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.

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.

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.⁵⁵ These include the number of reachable mills control variables in our main specification, and four additional control variables.

The first type of additional control variables is lagged outcome. First, we control for the 1-year lagged deforestation. Deforestation has been often shown to be an auto-regressive pro-

⁵⁵Except the case without any control, and the 4-year lagged outcome variables that constrain estimations over a significantly different time period.

cess, 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. It is conceivable, indeed, that deforestation is associated with 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. 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.

Furthermore, we present estimates from a specification controlling for 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. 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 can only be observed as of 2001). Estimating our main model over this same period yields a point estimate of 2.3 (confidence interval [1.1; 3.6]). Given the substantial change in the sampling time period implied by the addition of these lagged-outcome control variables and its negligible incidence, we do not investigate it in combination with the other robustness control variables presented below.

The second type of additional control variables 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 control set yields a similar estimate.

The third type of 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 type of 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.

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 (4). 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 3.2. 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).⁵⁶ It reports annual establishment-level data for all manufacturing facilities employing at least 20 employees.⁵⁷ 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;⁵⁸ 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

⁵⁶The data has also been referred to as *Statistik Industri* in the literature

⁵⁷The 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.

⁵⁸The 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.

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 75th 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 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.⁵⁹

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).⁶⁰

⁵⁹Due 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).

⁶⁰All 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).

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 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 small-holder 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.⁶¹ 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.10 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)).⁶²

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.

⁶¹We routinely exclude pixels categorized as industrial plantations in 2000, although the primary forest map should already exclude them.

⁶²This 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 4 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 4, 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 5 and displayed in Tables 6, 4 and A.11 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 USD 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 (Khatriwada et al. 2018) and an average price of USD 680/tCPO over the period (based on our own data), we convert their estimates into a 0.13-0.15 price elasticity.⁶³ 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 im-

⁶³We convert the additional USD 100 to a $100 * USD100 / (0.518 * 3.5 * 680) \approx 8.110924$ 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.

precise). Hence, here too, our research setting seems to capture a larger effect of prices on deforestation in the Indonesian oil palm sector.