Market Structure, Vertical Integration and Farmers' Welfare in the Costa Rica Coffee Industry*

Fabrizio Leone
Bank of Italy and CESifo

Rocco Macchiavello
London School of Economics and CEPR

Josepa Miquel-Florensa
Toulouse School of Economics

Nicola Pavanini
Tilburg University and CEPR

Click here for the most recent version

March 2025

Abstract

Widespread market imperfections in agricultural value chains raise the possibility that state intervention may enhance efficiency and farmers' welfare. We develop a structural model of agricultural chains to evaluate trade-offs of common regulations, estimating it using rich transaction data from Costa Rica's coffee sector. Farmers supply differentiated mills that decide from which rural areas to source and bilaterally bargain prices with downstream exporters. Through counterfactuals, we assess two pro-competitive regulations. Revenue-sharing increases the average farmer's welfare but makes many worse off. Similarly, banning vertical integration raises farm-gate prices but harms most farmers by lowering valuable services provided to them.

Keywords: Agricultural Chains, Market Structure, Farmers' Welfare

JEL Classification: O12, Q13, L22.

*We thank conference and seminar participants at Bocconi LEAP, BSE Summer Forum, CEPR-JIE Conference on Applied Industrial Organization, Cowles Foundation at Yale, EIEF-UNIBO-IGIER Bocconi Workshop on Industrial Organization, EEA, EUI Workshop on Firms, Labour Market, and Development, IIOC, ITAM Business School, Leuven Summer Event, RIDGE Industrial Organization Workshop, Tilburg Structural Econometrics Group. We benefited from the feedback of Jose Ignacio Cuesta (discussant), Monica Morlacco (discussant), Milena Almagro, Tomás Domínguez-Iino, and Michael Rubens. Ariana Ocampo Cruz, Juan Quincan, Oliver Seager, and Valerie Umaña provided exceptional research assistance. The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not represent the views of the Bank of Italy or their affiliated organizations. Rocco Macchiavello gratefully acknowledges financial support from the European Research Council under the Grant Agreement No. 818767. Nicola Pavanini gratefully acknowledges financial support from the European Research Council under the Grant Agreement No. 101042069. Josepa Miquel-Florensa acknowledges funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). E-mails: fabrizio.leone@esterni.bancaditalia.it; R.Macchiavello@lse.ac.uk; pepita.miquel@tse-fr.eu; and n.pavanini@tilburguniversity.edu.

1 Introduction

Approximately 75% of the world's poor live in rural areas and earn their livelihood in agriculture (World Bank, 2007, 2017). Linking these smallholder farmers to markets has the potential to lift millions out of poverty and kick-start structural transformation in many contexts. Agricultural value chains, however, are plagued with multiple market imperfections,[1] raising the possibility that state intervention may increase both efficiency and farmers' welfare. Indeed, government interventions to support farmers are commonplace. Besides direct support through, e.g., subsidized extension services, inputs, or loans, many regulations aim at altering market structures and/or rent-distribution along the chain to support farmers. Minimum prices, forced divestitures and/or consolidations, and zoning regulations are but a few common examples.

What are the impacts of these policies? Two challenges make it difficult to answer this question. First, the interaction of multiple market imperfections along the chain implies that these policies often have *a priori* ambiguous effects.[2] Second, data are often scarce due to the informality of many agricultural chains. This paper deploys tools from the empirical industrial organization literature to develop a structural model of agricultural chains to evaluate these trade-offs. We estimate the model in the Costa Rica coffee chain, a context of intrinsic interest due to its regulation and that also provides uniquely rich data to transparently identify the model primitives. Through counterfactuals, we quantify how common policy interventions affect market structure and welfare along the chain.

Our analysis highlights the nuanced, and potentially counter-productive, effects of commonly observed pro-competitive regulations on farmers' welfare. We proceed in three steps. First, we collect original data on the universe of transactions between farmers (who harvest coffee) and mills (who process it), and between mills and exporters (who sell it to international markets). We then build an empirical model that captures the key features of this—and others—agricultural chains: a multi-layered vertical structure in which (i) farmers supply oligopsonistic mills differentiated along both price and non-price (e.g., loans, training programs) dimensions; (ii) mills choose the local markets they source coffee from; and (iii) mills negotiate prices with exporters varying in demand and bargaining power. Finally, we use the model to assess the effects of two commonly observed regulations: (a) a rent-sharing rule operating in Costa Rica coffee (and in other countries); (b) a ban on vertical integration

[1] For example, farmers' lack access to well-functioning input (Duflo et al., 2008, 2011; Ashraf et al., 2009; Bold et al., 2017), credit (Karlan et al., 2014; Burke et al., 2019), insurance (Cai et al., 2015; Casaburi and Willis, 2018), saving (Casaburi and Macchiavello, 2019) and output (Bergquist and Dinerstein, 2020; Macchiavello and Morjaria, 2020; Casaburi and Reed, 2022) markets.

[2] For example, minimum farm-gate prices might drive intermediaries out of the market and hinder farmers' ability to sell; forced divestitures might curb buyers' power but also lower operational efficiency, and so on.

between mills and exporters (a policy also commonly observed). While farmers on average benefit from the rent-sharing rule, further tightening mills' share of rents would lower welfare for many farmers, once changes in market structure are accounted for. Banning vertical integration between mills and exporters also harms farmers, as the positive effect on farm-gate prices is more than offset by the loss of valuable services provided by integrated chains.

We begin providing background information on the Costa Rica coffee chain and the data. Coffee cherries are harvested by smallholder farmers and delivered to mills. Then, mills process the coffee and sell it to downstream exporters either through forward contracts agreed towards the beginning of the harvest season, or through spot contracts, agreed towards the end of harvest. Finally, exporters consolidate the coffee before selling it to international markets. This paper focuses on standard—as opposed to differentiated—coffee, which accounts for about 70% of total production in Costa Rica. Standard-grade coffee is a relatively homogeneous commodity, where differences in prices paid to farmers mostly reflect competitive forces and variations in mill technology—which are the focus of this paper—rather than differences in the coffee itself.

The Costa Rica coffee chain is regulated by the local coffee board (*Instituto del Café de Costa Rica*, or ICAFE). The regulation requires all coffee transactions along the chain to be registered as contracts with ICAFE. Therefore, it generates unique data on coffee sales from farmers to mills, from mills to downstream exporters, and beyond. For each transaction, we observe prices and quantities separately for ten harvest seasons, from 2010 to 2020. We also observe other farmers', mills' and exporters' characteristics, such as size, costs, and organizational forms. We distinguish between mills owned by farmers (cooperatives), by downstream exporters (vertically integrated), or operating as independent private entities. We complement these administrative data with our own surveys of mills to gather evidence about services that they offer to farmers, such as advance payments, loans, and training programs.

We documents three facts. First, in any given season, farmers sell their coffee to one mill only within the district where they operate. Second, vertically integrated mills source from more local markets than cooperatives and non-integrated mills. Finally, integrated mills and exporters trade a higher share of their coffee through forward contracts. Such contracts provide demand and supply assurance, but carry default risk when world prices fluctuate (Blouin and Macchiavello, 2019)—a friction mitigated by vertical integration (Macchiavello and Miquel-Florensa, 2018).

These facts inform the three building blocks of the structural model: farmers' choice to which mill they sell their coffee, bilateral bargaining between mills and exporters over prices, and mills' sourcing strategies across local coffee markets. Although this structure applies to other contexts, we explicitly model features specific to coffee in Costa Rica to enhance the transparency of identification and estimation of the model's structural parameters.

We model farmers' supply as a discrete choice among a set of differentiated mills, "inverting" the standard framework used to analyze the demand for differentiated products (Berry, 1994). This approach accounts for the fact that due to transport costs and market imperfections, prices offered by mills are not the only driver of farmers' decisions. Conditional on prices, farmers may prefer certain mills, such as those that provide credit, inputs, and other services, which our model captures through mill-market specific supply shifters.[3]

We model trade between mills and exporters as a bilateral-oligopoly bargaining process. To pin down bilateral prices, we rely on the "Nash-in-Nash" solution concept (Horn and Wolinsky, 1988; Collard-Wexler et al., 2019), which rationalizes a negotiated price between a pair of upstream and downstream firms as the Nash bargaining solution, assuming all other pairs have reached an agreement. The fact that mills and exporters can trade coffee in either the spot or forward market provides a natural way to define outside options if bargaining collapses. Unlike much of the existing literature (Crawford and Yurukoglu, 2012; Crawford et al., 2018), we observe mills' and exporters' margins, not just who trades with whom. Margins provide direct information about the relative bargaining power of mills and exporters, allowing for a transparent estimation.

We model mills' sourcing decisions as a two-stage game, as in Eizenberg (2014). In the first stage, mills choose which local districts to source from, taking into account both expected revenues and fixed sourcing costs. In the second stage, they receive coffee from farmers and sell it to downstream exporters, as described above. We use a revealed preference approach to estimate bounds on the fixed costs mills incur to source in each district.

The model estimates deliver several key insights, two of which are highlighted here. First, non-price factors appear to be key drivers of farmers' supply decisions and welfare. Estimated mill-district fixed effects, that capture non-price factors differentiating mills that source in the same district, shape farmers' supply decisions to a significant extent: a one-standard-deviation increase in the mill-district fixed effect is equivalent to a 2.5 standard deviations higher price. Mill fixed-effects accounts for approximately 60% of the estimated variation in mill-district fixed effects, suggesting that mill-level characteristics drive farmers' choices. Our original survey reveals that estimated supply shifters positively correlate with the share of farmers receiving advance payments, loans, and agricultural inputs and with the mills'

[3] Even though specific features of our context assuage endogeneity concerns, estimating farmers' supply still requires an instrument for the price paid by the mill in a given season. Because of the regulation formula, the price paid to farmers depends on the mill yield factor—the volume of parchment coffee produced from one kilo of coffee cherries. We instrument for prices using the *rendimiento tecnico*, the minimum, mill-specific, yield factor allowed by ICAFE in a given season.

participation in certification programs. Second, vertical integration plays an important, but ambiguous role, in farmers' supply decisions and welfare. On the one hand, vertical integration depresses prices paid to farmers: integrated mills negotiate lower prices with "their" exporters; integrated exporters also pay lower prices when sourcing from other mills. On the other hand, integrated mills provide more services and are more appealing to farmers, as revealed by their large market shares conditional on prices.

Finally, we use the model to evaluate the effects of common pro-competitive policy interventions aimed at protecting farmers. We consider changes in the rent-sharing rule in Costa Rica and a ban on vertical integration.[4] Increasing the share of mills' value added to be paid to farmers, mechanically increases payments to farmers. However, this direct effect causes mills to source from fewer local markets, resulting in lower welfare for many farmers. Banning vertical integration, instead, increases competition at the farm gate but, while beneficial for non-integrated mills and exporters, ultimately reduces farmers' welfare as the positive price effect is more than offset by the loss in valuable services provided by integrated mills.

Related Literature. We aim to contribute a tractable, yet realistic, empirical framework to evaluate common policies in agricultural value chains through counterfactual analysis. Two recent papers in the context of agricultural value chains are closest to ours. In a fascinating study, Rubens (2023) examines the consequences of oligopsony power in the Chinese tobacco industry, estimating markdown through a cost-side approach. He shows that a regulatory reform eliminating small, inefficient producers to increase productivity resulted in higher markdowns on tobacco leaves. While the cost-side approach imposes less assumptions than ours, it precludes detailed counterfactual analysis—our main goal. We estimate an equilibrium model of demand and supply along the chain to quantify the welfare implications of vertical, rather than horizontal, integration for farmers, mills, and exporters. Dominguez-Iino (2024) builds an empirical model of the South American agricultural sector to study how environmental policies are transmitted along the supply chain when regulation at the externality's source is infeasible. He finds that downstream agribusiness taxes reduce upstream emissions but poor targeting—exacerbated by agribusinesses' monopsony power—hinders their effectiveness. In his framework, farmers consider output and factor prices, and sell to homogeneous intermediaries. We extend our framework to capture salient features of agricultural value chains in developing countries, such as farmers' valuation for

[4] Guatemala, Nicaragua, El Salvador and Burundi have at some point adopted in their coffee sector rent-sharing rules similar to the one in Costa Rica. The Kenya and Rwanda tea sectors are currently also partially regulated along similar lines. Coffee in Kenya and Ethiopia, cocoa in Ghana and Ivory Coast, cotton in Tanzania, rice in Vietnam and Thailand, palm oil in Indonesia, and rubber in Malaysia and Thailand, are examples of contexts in which exporters cannot vertically integrate to purchase directly from farmers.

mills' provision of services, including loans, advances, and training (as in Macchiavello and Morjaria, 2020), as well as the oligopsonistic behaviour of intermediaries that strategically enter rural markets upstream and negotiate prices downstream. While both Rubens (2023) and Dominguez-Iino (2024) assume that intermediaries are price-takers downstream, we relax this assumption by introducing a bilateral bargaining model between mills and exporters.[5]

Our paper adapts tools developed in the industrial organization literature to a new context, agricultural value chains (see Macchiavello et al., 2022, for a review). We estimate an "inverted" demand framework for differentiated products (Berry, 1994; Berry et al., 1995; Nevo, 2000) to model farmers' supply.[6] We rely on the literature on product variety (Lee and Pakes, 2009; Eizenberg, 2014; Pakes et al., 2015; Wollmann, 2018; Fan and Yang, 2020, 2022; Montag, 2023) to model mills' sourcing decision across geographic markets, estimating bounds to their fixed costs of entry.

Within the structural industrial organization literature, existing studies (Crawford and Yurukoglu, 2012; Crawford et al., 2018; Collard-Wexler et al., 2019; Lee et al., 2021; Alviarez et al., 2023) focus on aspects of vertical integration that are pro-competitive (no double marginalization) as well as anti-competitive (foreclosure, rising rivals' costs). Albeit in a reduced-form, our model captures contracting problems that are particularly salient in agricultural chains in developing countries. First, integrated mills' advantage in providing farmers with interlinked transactions typical of these chains (see, e.g. Macchiavello, 2022; Boudreau et al., 2023; Casaburi and Willis, 2024). Second, integrated chains' advantage in curbing the risk of strategic default (Blouin and Macchiavello, 2019; Brugues, 2024) on forward contracts. In addition, leveraging unique data on prices along the chain and mills' costs, we are able to recover the full empirical distribution of bargaining parameters at the mill-exporter-season level with no need for supply-side estimation. Overall, our model allows us to quantify differences across mills' organizational structures in processing costs, fixed entry costs, bargaining power, and farmers' preferences.

The paper is structured as follows: Section 2 describes the Costa Rica coffee industry and presents the data. Section 3 contains the empirical facts. Section 4 introduces the empirical model. Section 5 describes identification and estimation. Section 6 presents our counterfactual excercises. Finally, Section 7 concludes.

[5] Other papers develop structural approaches to evaluate agricultural policies through counterfactuals, but with a rather different focus from ours. For example, Garg and Saxena (2022) and Bergquist et al. (2022) focus on input subsidies to farmers in India and Uganda, respectively.

[6] An alternative approach to measuring monopsonony power involves estimating the processors' production function (Avignon and Guigue, 2022; Rubens, 2023). See De Loecker and Scott (2022) for a discussion about differences and complementarities between these two approaches in the context of markup estimation.

2 Background: Context and Data Sources

2.1 The Costa Rica Coffee Chain

Harvest and Processing. Coffee cherries are harvested from coffee trees. The harvest season in Costa Rica typically lasts from November to March, with the exact timing varying slightly across regions, depending on altitude and weather patterns during the flowering season. After harvesting, the pulp of the coffee cherry is removed, leaving the bean, which is then dried to produce parchment coffee. There are two processing methods for obtaining parchment coffee: the dry method and the wet method. In the dry method, farmers clean cherries before drying them. In the wet method, instead, cherries are taken to a mill—known as coffee washing stations or *beneficios* in Costa Rica—where de-pulping, washing, drying, and sorting take place. The wet method requires specific equipment but results in more consistent quality, higher prices in export markets, and higher domestic value addition compared to the dry method.[7]

Unlike in other coffee-producing countries, nearly all coffee in Costa Rica is processed through the wet method. Coffee cherries must be processed within hours of harvest. Mills therefore process cherries harvested on farms located relatively nearby. A distinction can be made between regular mills and micro-mills. Regular mills are large-scale facilities that process coffee cherries from multiple farmers within a region. These mills handle substantial volumes of coffee and focus on efficiency and consistency across large batches. In contrast, micro-mills operate with lower fixed costs and capacity, allowing individual farmers or small groups of farmers to process their cherries.

Costa Rica coffee is classified as either standard (*convencional*) or differentiated (*diferenciado*).[8] *Diferenciado* refers to coffee that is differentiated along one or more dimensions: e.g., high-quality beans grown at higher altitudes, beans processed with less common honey or natural methods, specialty-grade coffee, organic coffee, fair trade, micro-lots with unique flavor profiles, and so on. *Convencional* coffee, instead, refers to commercial-grade coffee, produced in higher volumes with a focus on yield and efficiency rather than distinct cup quality. It is usually sold in bulk to larger importers and is blended for consistency rather than uniqueness. This paper focuses exclusively on standard coffee, which constitutes approximately 70% of total coffee production in Costa Rica each year. The undifferentiated nature of standard coffee implies that prices paid to farmers primarily reflect competition and technology in local markets rather than any inherent differences in the coffee itself.

[7] See Technoserve (2020) for a comprehensive guide to the wet method.

[8] Costa Rica exclusively cultivates Arabica coffee (cultivation of Robusta coffee is banned), mostly of the Caturra and its descendant Catuai varieties.

Marketing and Sales. After harvest, farmers produce coffee cherries and sell them to mills. Mills may or may not provide services to farmers, such as advance payments, loans, and training programs. Mills process the cherries and then sell the resulting parchment coffee to both national and international exporters. Exporters consolidate and prepare the coffee before selling it to foreign buyers or domestic roasters. Each year, approximately 90% of Costa Rica's total coffee production is exported. At the end of the harvest season, mills pay farmers a final liquidation price, determined by the regulation described below.

Coffee cherries must be delivered to mills within hours of harvest. Farmers deposit cherries either directly at a mill's premises nearby their farms, or at buying points operated by mills, with mills then arranging transportation to the mill's premises for processing. Farmers thus deliver to mills operating in their proximity. Otherwise, there are no regulations, e.g., zoning rules, which restrict where farmers (mills) can sell to (buy from).

Mills differ in their organizational forms. We distinguish between mills owned by backward integrated exporters (i.e., vertically integrated mills), private non-integrated mills, and cooperatives. Thanks to more secure marketing downstream, backward integration enables mills to source large quantities of coffee each season, making them some of the largest organizations in Costa Rica. Cooperatives are mills collectively owned by local farmers. While most cooperatives have bylaws that require members to sell cherries exclusively to the cooperative, these bylaws are often not enforced in practice. Still, cooperatives rarely purchase cherries from non-members and typically do not operate buying points outside their designated catchment area. While prices are an important factor in farmers' decision of which mill to sell to, mills can also compete through non-price factors, such as advance payments, provision of pre-harvest loans, and training programs.

Mills and exporters trade coffee either through spot or forward contracts. Spot transactions are agreed upon towards, or even after, the end of the harvest season. Spot contracts entail the prompt delivery of already processed parchment coffee and are signed over a relatively short period when all mills and exporters are actively looking to sell and buy coffee. In contrast, forward contracts entail promises to deliver parchment coffee that has not yet been produced. They are signed over a longer period, often well before the beginning of the harvest season, to allow mills and exporters to manage their specific needs for demand and supply assurance. Forward contracts thus greatly differ in terms of when they are agreed, when coffee is to be delivered, and how much they leave parties exposed to price risk and potential strategic defaults (Blouin and Macchiavello, 2019). While spot contracts are relatively homogeneous and agreed upon in an active market, forward contracts are bilaterally negotiated between mills and exporters with pre-existing relationships (Macchiavello and Miquel-Florensa, 2018).

The Sistema de Liquidación Final. The coffee sector in Costa Rica is regulated through the *Instituto del Café de Costa Rica* (ICAFE), a non-government institution representing the interests of farmers, mills, and exporters. The main goal of ICAFE is "to achieve an equitable system of relationships between farmers, mills, and exporters of coffee that guarantees a rational and secure participation of each stage in the coffee business".[9] ICAFE's main tool in pursuing its goal is the *Sistema de Liquidación Final*, illustrated in Online Figure B.1.

The *Sistema de Liquidación Final* is a revenue-sharing rule that regulates the price that mills must pay to farmers at the end of the harvest season. This final liquidation price is directly tied to the mill's margin and is computed as follows. Mills are allowed to retain as profits 9% of the difference between the revenues generated by sales to downstream buyers and their processing costs, both of which must be approved by the regulator. The remaining amount must be paid to farmers as the liquidation price. These prices are published in newspapers, and mills are required to make the corresponding payments within eight days, deducting any advance payments made during the harvest season. Mills must pay the same liquidation price to all farmers they source coffee from during a given harvest season. While, in theory, this could affect farmers' and mills' quality choices, this is not a concern for standard coffee—the focus of this paper—which should be understood as a relatively homogeneous good.

To calculate the final liquidation price and implement the *Sistema de Liquidación Final*, ICAFE must be able to calculate mills' sales, costs, and yields. On the sales side, all contracts between farmers and mills, as well as between mills and exporters, must be registered with ICAFE.[10] ICAFE also enforces contracts throughout the chain.[11] Additionally, mills must report their processing costs to ICAFE, which approves them following a detailed technical assessment of the mill's technology and prevailing input prices. Finally, ICAFE computes a minimum yield—the quantity of parchment coffee produced per kilo of cherries—for each mill. Based on coffee production from previous harvests, ICAFE identifies a set of representative districts. During the harvest season, ICAFE technicians visit the buying stations (*recibidores*) of active mills in these districts and sample the coffee beans they receive from farmers. By analyzing the beans' physical characteristics, such as size and humidity, ICAFE estimates a mill-specific yield (*rendimiento técnico*), which indicates how many kilograms of coffee can be produced from one kilogram of input. This estimated yield is then used to cal-

[9] For more details, see: www.icafe.go.cr.

[10] Prices between mills and exporters cannot fall below an undisclosed minimum calculated by ICAFE based on market conditions. This ensures a meaningful interpretation of prices within vertically integrated organizations.

[11] For example, once a forward contract is registered, the buyer cannot refuse delivery at the agreed-upon price. Mills may only default on forward contracts under exceptional circumstances that prevent them from processing.

culate the minimum price. Mills located in areas with higher estimated yields are required to pay a higher minimum price. At the same time, mills calculate their own yield (*rendimiento reportado*) and submit a price proposal to ICAFE. The price that each mill pays is either the self-reported price or the minimum price set by the regulator, whichever is higher. We exploit this feature of the regulation to create instruments to identify farmers' supply.

The revenue-sharing rule in the *Sistema de Liquidación Final* ensures that farmers receive competitive prices despite selling in relatively concentrated local markets where mills potentially have substantial market power. Furthermore, farmers are paid a final price tied to market conditions prevailing throughout the whole harvest season, rather than at the specific time they harvest. Given sudden swings in international coffee prices, this provides within-season price insurance. Finally, mills pay the final liquidation to farmers *after* having been paid by downstream buyers. This reduces mills' working capital requirements, a significant challenge in the industry (Blouin and Macchiavello, 2019).

2.2 Data Sources and Descriptive Statistics

Data Sources. From ICAFE, we obtain transaction-level data for all sales between farmers and mills and between mills and exporters over a sample period for the harvest seasons from 2010/2011 to 2019/2020. The type of coffee, *convenicional* or *diferenciado*, is observed in each data source and, as mentioned above, we focus on *convenicional* coffee.

Farmer-Mill Transactions. We observe the quantity of coffee each farmer sells to each mill—including micro-mills possibly owned by the farmer—in a given district and season. On the farmers' side, we aggregate this information to construct a measure of farmer size. We also know the farmer's gender. Concerning mills, in each season, we can compute the quantity of cherries purchased in every district they source from. We also observe the final liquidation price paid by mills to farmers. Finally, we observe audited mills' cost of processing coffee as well as their yield factor (*rendimiento*)—the kilograms of parchment of coffee obtained per kilogram of purchased cherries. We convert all the nominal quantities (prices and costs) in USD per kilogram and express trade volumes in kilograms. Finally, we observe mills' organizational form (private vertically integrated, private not integrated, or cooperative). No mill ever changes organizational form during our sample.

In agricultural chains, downstream mills and intermediaries often provide services to farmers, e.g., agricultural inputs, credit, agronomic training, and other extension services (Macchiavello et al., 2022). Holding constant prices and distance, the provision of such services potentially influences farmers' supply choices across mills. The provision of these services is rarely recorded in administrative data, and Costa Rica is no exception. We thus complement the transaction-level data with a survey jointly conducted with ICAFE in 2024. We selected a representative sample of 70 mills, stratified by geographical area and organizational form, and obtained a 58% response rate. The survey covered mills' processes and activities during the 2022/2023 harvest season. A designated module covered mills' provision of advance payments, agricultural inputs, training and extension services, loans to farmers, and participation in sustainability programs. We use survey responses to interpret the estimated structural parameters of farmers' supply preferences across differentiated mills.

Mill-Exporter Transactions. We observe the bilateral prices and quantities exchanged by each mill and exporter in a given season. We distinguish between forward and spot contracts. We classify as "forward trade" all contracts that specify at least a 30-day interval between signature and delivery. We classify as "spot trade" all the others. We express prices in USD and quantities in kilograms. Moreover, we observe the price that each exporter receives on international markets in every season. We also observe if exporters are integrated or not with upstream mills. No exporter ever changes organizational form in our sample.

2.3 Descriptive Statistics

Farmers and Mills. Our sample includes 105 districts where mills compete to source coffee from farmers.[12] In the average district-season, about 460 farmers (28% of which are female) and 6 mills are active.

Online Table A.1 reports summary statistics for mills by organizational form. In the average season, there are 165 active mills, 9 of which are vertically integrated. Integrated mills collectively source between 35% and 55% of all coffee produced, depending on the season. Private integrated mills source coffee from more farmers and more districts (the average mill sources coffee from 3 districts). Private integrated mills have higher market shares in the average district they source from, have lower processing costs, and pay lower prices to farmers. Our survey data offer insights into the differences in mills' service provision to farmers. As shown in Online Table A.2, the average mill pays approximately 20% of its farmers in advance, provides agricultural inputs to 25%, and offers loans to 19%. Around 60% of mills participate in at least one voluntary sustainability program.[13]

Mills and Exporters. In an average season, there are 71 active exporters, each mill sells to 2 exporters, and each exporter sources from 5 mills. Among the exporters, 7 are vertically integrated (i.e., own at least one mill). Online Table A.3 presents summary statistics from the

[12] There are 174 administrative districts in Costa Rica where coffee is grown. To define the geographical boundaries of the areas where mills compete, we aggregate small districts into larger ones based on geographical proximity. After the aggregation, we end up with the 105 districts shown in Online Figure B.2.

[13] We exclude programs (such as Fair Trade) available only to cooperatives or farmer associations.

perspective of mills, while Online Table A.4 provides them from the exporters' perspective. Integrated mills typically trade with fewer exporters (typically just their parent company) and therefore maintain longer relationships, trade larger coffee volumes, and transact at lower prices than cooperatives and private non-integrated mills. In contrast, integrated exporters source from both their own mills and from other mills. Even when trading with outside mills, they maintain longer relationships and at trade lower prices than non-integrated exporters.

3 Motivating Facts

Figure 1 documents three empirical facts that motivate our modeling assumptions in the next section.

Fact 1. *In any given season, each farmer sells to one mill only among the few that source from the farmer's district in that season.*

The top-left panel of Figure 1 shows the share of farmers that sell to a given number of mills within a district-season. About 95% of farmers sell to a single mill.[14] The top-right panel displays the distribution of the number of mills sourcing from a given district. Around 6 mills (standard deviation 3.5) source cherries in the average district in the typical season.

Two main reasons explain this fact. First, coffee cherries are perishable and must be processed by mills within hours of harvest. This limits the distance that farmers can travel to deliver their harvest. Second, to source cherries from multiple districts, mills incur costs for setting up and operating collection centers, establishing buying points, and covering transportation expenses. This explains why only a few mills source in each district. As a result, farmers have a limited number of mills to choose from. From a modeling perspective, farmers' supply is well-represented by a discrete choice over a finite set of differentiated mills sourcing in their district.

Fact 2. *Private integrated mills enter more districts and source more coffee than cooperatives and private non-integrated mills.*

The middle-left panel of Figure 1 shows the number of districts from which mills with different organizational forms source coffee. The middle-right panel shows the log of the total quantity of coffee they source across all seasons. The average private integrated mill sources coffee from 20 districts, whereas private non-integrated mills and cooperatives enter

[14] Few farmers operate multiple farms across districts. Indeed, in any given season, 85% of farmers sell in a single district. Since cherries must be processed within hours of harvest, proximity is key and we consider the farmer-district as the relevant unit of analysis.

less than 10 districts. Additionally, private integrated mills source about twice as much coffee as the other two types of organizations.

Private integrated mills thus source larger volumes than other mills both because they source larger volumes in a given district, but also because, in any given season, they source from more districts. To capture this, the model's structural parameters on the supply side (farmers' price sensitivity and mills' appeal to farmers in a given district) and on the demand side (mills' entry costs across districts) are allowed to vary with mills' organizational form.

Fact 3. *Vertically integrated chains trade a higher share of coffee through forward contracts than non-integrated ones.*

The bottom-left panel of Figure 1 shows the share of forward trade in mills' total coffee sales to exporters. Forward contracts provide supply assurance for exporters and demand assurance for mills, reducing inventory risks, but are potentially subject to the risk of strategic default when market conditions change (Blouin and Macchiavello, 2019). Private integrated mills trade a higher share of coffee through forward contracts (about 65%) than cooperatives (47%) and other private mills (34%). The bottom-right panel of the figure shows that mills transact with only a limited number of exporters on the forward market. Private integrated mills typically only sell coffee to the downstream exporter that owns them. In the average season, private non-integrated mills and cooperatives sell to 1.5 and 2.7 exporters, respectively.

Motivated by this evidence, we model the forward market as a bilateral oligopolistic market in which mills and exporters with heterogeneous demand, costs, and negotiating power bargain over prices. The model captures heterogeneity in volumes traded forward across organizational forms while abstracting from strategic defaults on forward contracts.[15]

4 Model

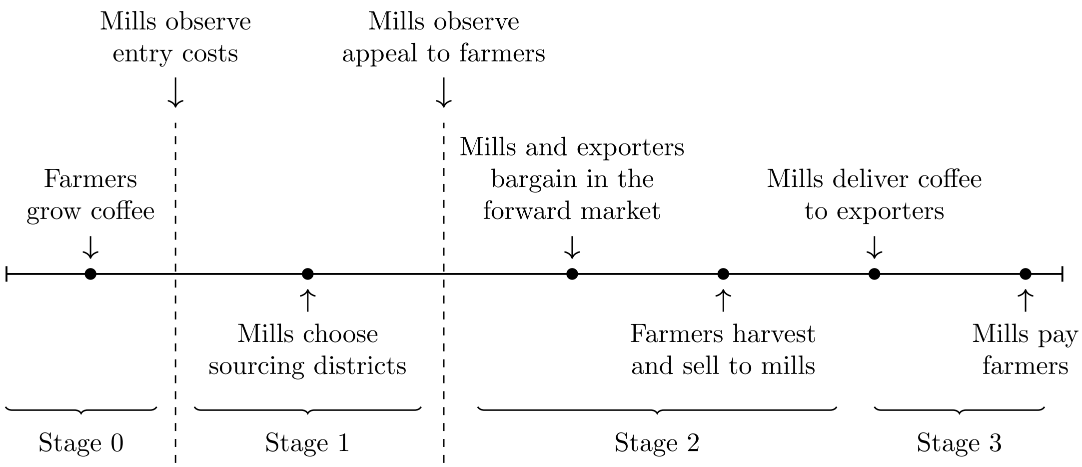
This section introduces an empirical model of the coffee value chain. While the structural framework reflects the stylized facts described above, as well as the institutional and regulatory features of the Costa Rica coffee chain, it could be generalized to other settings.

[15] Macchiavello and Miquel-Florensa (2018) show that volumes traded forward are calibrated to the organizational form's ability to avoid strategic defaults, which are thus rare in the data. Vertical integration entirely suppresses a mill's ability to strategically default, and thus integrated mills supply larger volumes of coffee through forward sales to their exporters. Long-term relationships between mills and exporters mitigate strategic default, but to a lesser extent, and therefore non-integrated vertical chains are constrained in the volumes of coffee they can trade forward.

4.1 Timeline

The model closely follows the unfolding of the coffee harvest season and consists of three parts. Proceeding backward, first, we model farmers' supply as a discrete choice over mills that source in their district. Second, we model the bargaining between mills and exporters. Last, we model mills' entry decisions across districts. The Figure illustrates the timeline of the events within each harvest season in the model.

Timeline of Events



Note: The Figure illustrates the timeline of events in the model. The timeline closely follows the unfolding of the coffee season. In stage 0, farmers make planting decisions. Those determine the quantity of coffee produced, which is exogenous and outside the model. In stage 1, mills form rational expectations over the subsequent unfolding of the coffee season and decide whether to pay the fixed costs to source from a given district (subsection 4.4). In stage 2, the harvest season unfolds. Farmer-specific taste shocks over mills sourcing in their district are realized. Mills negotiate forward contracts with exporters (subsection 4.3) and farmers form rational expectations over prices and decide the mill to which they supply coffee among those sourcing in their district (subsection 4.2). In stage 3, farmers harvest coffee and deliver it to mills, mills process coffee and deliver it to exporters, and, finally, all payments are made. These events happen mechanically and are thus not modeled as separate stages.

In stage 0, farmers grow coffee in their districts. Coffee trees require about three years to produce after planting. In a given season, the volume of coffee harvested is determined by past planting decisions and weather conditions during the flowering season, which are both observable to all market participants. For simplicity, we do not model farmers' investment choices and take the volume of coffee harvested as exogenous.

In stage 1, mills observe the cost of entering each district and form expectations about their appeal to farmers. They then compare expected revenues against entry costs and decide which districts to source coffee from. Following Eizenberg (2014), we assume that mills observe entry costs before making sourcing decisions but learn their appeal to farmers only upon entry, i.e., between stage 1 and stage 2.

In stage 2, the harvest season unfolds. Farmer-specific taste shocks over mills sourcing in their district are realized. Before receiving coffee from farmers, mills negotiate with exporters a forward price for a quantity of coffee to be exchanged after the harvest. Farmers form rational expectations about the price of coffee paid by each mill, which—through the ICAFE regulation—depends on the prices and volumes mills negotiate with exporters and future spot market prices.

Finally, in stage 3, after harvest, mills receive coffee from farmers, process it, and deliver it to downstream exporters. Any coffee unsold by mills, or unpurchased by exporters, through forward contracts is traded on the spot market at exogenous prices determined by the international market. Mills pay farmers in compliance with the regulation. We assume that all negotiated contracts and positions are executed with no default, and thus do not explicitly model this stage.

Coffee cherries (mills' raw material) are perishable and mills thus do not carry inventories. In principle, parchment coffee (mills' output) is storable. However, its market value drastically depreciates as the following harvest season approaches, and therefore neither mills nor exporters carry inventories. We thus solve the model separately for each season.

4.2 Farmers' Choice of Mills

There are $f = 1, \ldots, F$ farmers located in $d = 1, \ldots, D$ districts. In each season $s = 1, \ldots, S$, each mill $m = 1, \ldots, M$ decides whether to enter district d. Farmer f sells all her harvest to one mill $m \in \mathcal{M}_{ds}$, with $\mathcal{M}_{ds}$ the (endogenous) set of mills sourcing in district d in season s. Each farmer sells to the mill that maximizes her indirect utility, given by:

$$U_{fmds} = \underbrace{\alpha_0 + \alpha_1 P_{ms} + \eta_{md} + \zeta_{ds} + \xi_{mds}}_{\delta_{mds}} + \varepsilon_{fmds}. \tag{1}$$

Farmers' choice of mills is driven by both price and non-price factors. P_{ms} is the price offered by mill m in season s. η_{md} is the average appeal of mill m in district d. This term captures—in a reduced form—non-price attributes that are relevant to farmers' decisions, such as the convenience of the mill's collection points in the district and interlinked transactions which are common in agricultural chains (Macchiavello et al., 2022). It explains why some mills systematically receive more coffee than others, conditional on prices. For example, mills offering more services (e.g., cash advances, loans, and inputs) may have a higher η_{md}. Mills' ability to provide, and farmers' valuation for, such services may vary over space and thus

η_{md} is mill-district specific. ζ_{ds} is a district-season fixed effect, and ξ_{mds} is a mill-district-season demand shifter unobserved to the econometrician. ε_{fmds} is an idiosyncratic shock that follows a type 1 extreme value distribution. α_0 is the average utility level, while α_1 is the price sensitivity of farmers.

Based on this utility function, we can express the market share of mill m in district d in season s and farmers' surplus in district d in season s as:

$$S_{mds} = \frac{\exp(\delta_{mds})}{\sum_{n \in \mathcal{M}_{ds}} \exp(\delta_{nds})} \quad \text{and} \quad \mathrm{FS}_{ds} = \frac{1}{\alpha_1} \log \left(\sum_{n \in \mathcal{M}_{ds}} \exp\left(\delta_{nds}\right) \right). \tag{2}$$

Market shares positively depend on mills' prices, mill-district and district-season fixed effects, and unobserved mill-district-season demand shifters. The same applies to farmers' surplus, which—all else equal—also positively depends on the number of mills active in a district. However, if higher competition between mills hinders mills' ability to provide services to farmers (as in Macchiavello and Morjaria, 2020) lower estimated η_{md} would captured that.

4.3 Mills and Exporters

Mills' Production Function and Payoff. Mills convert coffee cherries into parchment coffee. The production function of mill m in season s is characterized by two parameters: a yield factor z_{ms} and a constant marginal cost c_{ms}. The yield factor captures the kilos of parchment coffee produced from one kilo of coffee cherries and depends on mill-specific technology and management practices, as well as prevailing agronomic and weather conditions that affects cherries humidity and moisture. Mill also use variable inputs (mainly water, electricity, fuel, labor, and other materials). Conditional on the mill's overall installed capacity, the mill's technology is well approximated by constant marginal costs.

Every season, mills sell coffee to exporters $e = 1, \ldots, E$. Mills and exporters trade coffee either through forward contracts, where prices are negotiated bilaterally, or on the spot market, where prices reflect conditions prevailing in the international coffee markets. The profits of mill m in season s are:

$$\pi_{ms} = \mu \left(\sum_{e \in \mathcal{E}_{ms}} \left(p_{mes} - c_{ms}\right) q_{mes} + \left(p_{ms}^{spot} - c_{ms}\right) q_{ms}^{spot} \right) - \sum_{d \in \mathcal{D}_{ms}} F_{mds}. \tag{3}$$

p_{mes} denotes the forward price that mill m negotiates with exporter e in season s, and q_{mes} is the corresponding quantity of parchment traded. $q_{ms}^{spot} \geq 0$ and $p_{ms}^{spot} > 0$ are the volume and (expected) price of coffee sold on the spot market. c_{ms} is the (constant) marginal cost of processing coffee for mill m in season s, while F_{mds} captures the cost incurred by mill m

to source coffee from district d in season s. The set $\mathcal{E}_{ms}$ refers to the exporters to whom mill m sells in season s, and $\mathcal{D}_{ms}$ is the set of districts from which mill m sources coffee in season s. While $\mathcal{D}_{ms}$ is endogenously determined by mills' sourcing decisions, $\mathcal{E}_{ms}$ is exogenously given and fixed in the counterfactuals.[16]

The mill's profits formulation captures the ICAFE regulation in two ways. First, we let $\mu = 9\%$ be the share of profits that the mill retains at the end of the season. The remaining share, $1 - \mu$, goes to the farmers in the form of a final liquidation price. Second, as per regulation, the final liquidation price that farmers receive at the end of the season is constant across districts and given by:

$$P_{ms} = z_{ms}(1-\mu)(\bar{p}_{ms} - c_{ms}), \qquad \bar{p}_{ms} = \sum_{e \in \mathcal{E}_{ms}} s_{mes} p_{mes} + \left(1 - \sum_{e \in \mathcal{E}_{ms}} s_{mes}\right) p_{ms}^{spot}. \tag{4}$$

Upstream prices paid to farmers (P_{ms}) thus depend on the average price at which the mill sells parchment ($\bar{p}_{ms}$), audited processing costs (c_{ms}), and the mill's yield factor (z_{ms}). The average output price weights the prices the mill negotiates with downstream exporters in the forward market and the prices it receives in the spot market for any remaining unsold quantities. The weights are given by the share of coffee that mills and exporters trade in the forward (s_{mes}) and in the spot ($1 - \sum_{e \in \mathcal{E}_{ms}} s_{mes}$) markets. For simplicity, we assume that the (expected) spot market price (p_{ms}^{spot}) is such that mills earn zero expected profits in the competitive spot market, i.e., $p_{ms}^{spot} = c_{ms}$. This assumption captures inventory risk: after harvest, parchment not committed to forward contracts is sold, if at all, at a significant discount. Competition in the spot market ensures that the risk a mill is willing to incur is such that expected profits on those sales are equal to zero.

Exporters' Payoff. Exporter e's profits on the forward market in season s are:

$$\pi_{es} = \sum_{m \in \mathcal{M}_{es}} \left(p_{es}^{w} - p_{mes}\right) q_{mes} - \sum_{m \in \mathcal{M}_{es}} \frac{\kappa_{es}}{2} q_{mes}^{2}, \tag{5}$$

where p_{mes} and q_{mes} are defined above. p_{es}^{w} is the exogenous price that exporter e receives in international markets during season s. The term $\frac{\kappa_{es}}{2} q_{mes}^2$ reflects the risks faced by exporters when trading on the forward market (e.g., default risk and price fluctuations).[17] κ_{es} captures the sensitivity of different exporters to such shocks. We denote by $\mathcal{M}_{es}$ the set of mills from which exporter e purchases coffee in season s. Similar to $\mathcal{E}_{ms}$, this set is considered exogenous

[16]This assumption is common in the literature (e.g., Crawford and Yurukoglu, 2012; Crawford et al., 2018) and supported by the data: the median (average) mill in our sample drops 0.1 (0.3) exporters per year.

[17]Convex costs and exogenous downstream prices generate a downward-sloping demand curve in a tractable manner (see, e.g., Leong et al., 2022, for a similar approach in a different context).

and held fixed in counterfactuals.[18]

From equation (5), the quantity of coffee that exporter e demands from mill m in season s in the interior solution case ($p^w_{es} > p_{mes}$) is:

$$q_{mes} = \frac{p^w_{es} - p_{mes}}{\kappa_{es}}. \tag{6}$$

Nash Bargaining. The bargaining between mill m and exporter e in season s yields a forward price p_{mes} defined by:

$$\max_{p_{mes} \in [c_{ms},\ p^w_{es}]} \left(\pi_{ms} - \pi^D_{ms}\right)^{\lambda_{mes}} \left(\pi_{es} - \pi^D_{es}\right)^{1-\lambda_{mes}} \tag{7}$$
$$\text{s.t. equations (3), (5), (6).}$$

$\lambda_{mes} \in [0, 1]$ is the bargaining power of mill m when selling to exporter e in season s. The definition of π_{ms} and π_{es} comes from above. π^D_{ms} and π^D_{es} denote mills' and exporters' disagreement payoffs, respectively. In case of disagreement, we assume that mills do not purchase that quantity from farmers, or, equivalently, sell that quantity on the spot market at an expected profit of zero. Either way, the mill's outside option is given by:

$$\pi^D_{ms} = \mu \sum_{e' \in \mathcal{E}_{ms \setminus \{e\}}} \left(p_{me's} - c_{ms}\right) q_{me's} - \sum_{d \in \mathcal{D}_{ms}} F_{mds}, \tag{8}$$

so that the surplus of mill m when selling to exporter e on the forward market in season s is equal to $\mu \left(p_{mes} - c_{ms}\right) q_{mes}$, the variable profits generated by that sale. Conversely, exporter e's outside option in the event of disagreement with mill m in season s is to source q_{mes} on the spot market, earning zero profits. By the same logic, the surplus of exporter e when sourcing from mill m on the forward market in season s is equal to $\left(p^w_{es} - p_{mes}\right) q_{mes} - \frac{\kappa_{es}}{2} q^2_{mes}$.[19]

We assume that each transaction takes place separately and parties have passive beliefs about all other transactions (including those in which they are a party). Thus, if they fail to reach an agreement, mills and exporters do not renegotiate the contracts they have already

[18] This is also supported by the data: the median (average) exporter drops 0 (0.3) mills per year.

[19] For simplicity, we assume that farmers have passive beliefs about the final prices paid by mill m at the end of the season (P_{ms}) in the event of an off-the-equilibrium-path bargaining breakdown between mill m and exporter e. In other words, farmers do not stop supplying the mill following its negotiation breakdown with a specific buyer.

signed.[20] In the interior solution case, bilateral prices are given by:

$$p_{mes} = \frac{\lambda_{mes} p^w_{es} + (2 - \lambda_{mes}) c_{ms}}{2}. \tag{9}$$

As long as $p^w_{es} > c_{ms}$, the forward price mills receive from downstream exporters increases with the bargaining power of the mill, as do mills' payments to farmers via equation (4). Thus, our model allows us to analyze how mills' bargaining with downstream exporters affects farmers' surplus upstream. Notice that as λ_{mes} approaches zero, exporters have all the bargaining power, and mills earn zero profits.

4.4 Mills' District Choice

Each season, mills decide which districts to source from, depending on whether it is profitable to do so. Let $\mathcal{A}_{ms} = \mathcal{A}^1_{ms} \cup \mathcal{A}^0_{ms}$ be a vector of all the potential districts where each mill can enter across all seasons. Among those, let $\mathcal{A}^1_{ms}$ denote the set of districts in which mill m actually enters in season s, and let $\mathcal{A}^0_{ms}$ be the set of districts where mill m did not enter in season s. We specify the fixed costs that mill m faces when entering district d in seasons s as follows:

$$F_{mds} = F_m(\theta) + \nu_{mds}, \qquad \mathbb{E}[\nu_{mds} | d \in \mathcal{A}_{ms}] = 0, \tag{10}$$

with $\theta \equiv \{\mu, \alpha, \eta_{ds}, \kappa_{es}, \lambda_{mes}, p^w_{es}\}$. Equation (10) states that fixed entry costs depend on a mill-specific component, which we view as a parameter to estimate, and a mean-zero error, which captures idiosyncratic deviations from the mean. Following Eizenberg (2014), we assume that ν_{mds} is observed by mills at the beginning of stage 1 but remains unobserved by the econometrician.

The relevant equilibrium concept of our two-stage game is that of the Subgame Nash Perfect Equilibrium (SPNE): a vector of district choices and pricing strategy for each mill constitutes a Nash Equilibrium if no one has incentive to unilaterally deviate from it. As in Eizenberg (2014), we assume the existence, but not the uniqueness, of a SPNE.[21] A necessary

[20] This is a standard assumption in multilateral oligopoly bargaining problems (Crawford and Yurukoglu, 2012; Crawford et al., 2018), and is commonly referred to as "Nash-in-Nash" bargaining solution (Horn and Wolinsky, 1988; Collard-Wexler et al., 2019).

[21] We assume sequential entry to deal with multiple equilibria in counterfactuals (e.g., Lee and Pakes, 2009; Fan and Yang, 2020).

condition for the existence of an SPNE is:

$$E_{u|\theta}\left[\pi_{ms}(\mathcal{A}_{ms}; u, \theta) - \pi_{ms}(\mathcal{A}_{ms} - 1^{d}_{ms}; u, \theta)\right] \geq F_{mds} \qquad \forall d \in \mathcal{A}^{1}_{ms}, \tag{11}$$

$$E_{u|\theta}\left[\pi_{ms}(\mathcal{A}_{ms} + 1^{d}_{ms}; u, \theta) - \pi_{ms}(\mathcal{A}_{ms}; u, \theta)\right] \leq F_{mds} \qquad \forall d \in \mathcal{A}^{0}_{ms}. \tag{12}$$

1^{d}_{ms} is a vector with entries of zeros apart from a single entry of 1 for district d. In words, if mill m enters district d in season s, the profits earned by entering must be at least as large as the (counterfactual) profits it would have earned by not entering. Conversely, if mill m does not enter district d in season s, the (counterfactual) profits it would have earned by entering cannot exceed the profits it achieves by staying out. The expectation operates over the joint distribution of second-stage shocks $u_{mds} = \{\xi_{mds}, \eta_{md}\}$, conditional on an estimated set of second-stage parameters θ. All shocks in u_{mds} are realized upon entry.

5 Identification and Estimation Results

5.1 Identification

Farmers' Supply. Most farmers always have the option to sell their production to small, often family-owned, mills. In each district-season, we aggregate small mills into one alternative, the outside option, and normalize its utility to zero.[22] We then apply Berry (1994)'s inversion to market shares in equation (2) to obtain the estimating equation:

$$\log\left(S_{mds}\right) = \alpha_0 + \alpha_1 P_{ms} + \eta_{md} + \zeta_{ds} + \xi_{mds}. \tag{13}$$

The inclusion of district-season and mill-district fixed effects implies that the parameters in equation (13) are identified from variations in prices and market shares across mills and seasons, within districts in a given season, and within a mill-district pair over time.[23]

Estimating equation (13) requires addressing the issue of price endogeneity. In a standard environment, one might be concerned that mills with high realizations of the mill-district-season supply shifter (ξ_{mds}) in a certain district, might offer lower prices while still sourcing a large volume from that district. However, in the Costa Rica context mills pay the same price across districts and so that margin of adjustment is not relevant. Furthermore, the price ultimately paid to farmers is computed according to the ICAFE regulation formula in equation (4). These considerations attenuate price endogeneity concerns. It is still possible,

[22] We define small mills as those that (i) source less than 10,000 fanegas of coffee across all seasons, which corresponds to the third percentile of the distribution of quantities sourced by mills (summed across districts and averaged over seasons), and (ii) cumulatively have a market share less than or equal to 1%.

[23] The outside option market share does not divide S_{mds} as it is absorbed by district-season fixed effects.

though, that high realizations of the mill-district-season supply shifters (ξ_{mds}) induce the mill to source more coffee which then lower prices downstream, e.g., by forcing the mill to sell more on the spot market. This would imply $\text{corr}(P_{ms}, \xi_{mds}) < 0$ and the OLS estimate of α_1 to be downward-biased.

We address price endogeneity by leveraging the ICAFE regulation. Equation (4) suggests that the mill yield factor (z_{ms}) provides a natural instrument for the price paid to farmers by mill m in season s (P_{ms}). The yield factor specifies how many kilograms of parchment coffee a mill can produce, and determines the price farmers receive per kilogram of cherries delivered. In practice, the yield factor—or *factor de rendimiento*—used to determine the price paid to farmers is given by $z_{ms} = \max\{z^R_{ms}, z^T_{ms}\}$, where z^R_{ms} is the *rendimiento reportado*, calculated by the mill as the ratio of the actual volume of parchment coffee sold to the total amount of cherry coffee purchased, while z^T_{ms} is the *rendimiento técnico*—the minimum yield factor allowed by ICAFE. Since the *rendimiento* reported by the mill (z^R_{ms}) might correlate with unobserved factors that influence farmers' supply to the mill other than prices, and thus violate the exclusion restriction, we use the *rendimiento técnico* (z^T_{ms}) as an instrument for prices in equation (13).

To compute the *rendimiento técnico*, each season ICAFE samples coffee cherries from multiple buying points across different mills and districts, processes these samples to estimate location-specific yield factors for that season, and finally aggregates this information to estimate a mill-season specific yield factor.[24] From an identification standpoint, this process ensures that the *rendimiento técnico* calculated for a mill-season depends both on samples taken from the mill itself and samples taken from nearby locations, regardless of whether the mill sources there. While a potential concern is that weather realizations in a location correlate with the calculated *rendimiento técnico* and with mills' sourcing in that district, the inclusion of district-season fixed effects controls for those shocks. Furthermore, mill-district fixed effects control for the time-invariant composition of samples used to construct the *rendimiento técnico* for the mill. In sum, aggregate variation in the mill-season specific yield factor is unlikely to correlate with mill-district-season specific shocks (ξ_{mds}) conditional on the fixed effects included in the specification. In terms of monotonicity, the instrument only shifts prices of mills that are reporting a yield below the one that ICAFE estimates.

In Section 5.2, we experiment with heterogeneous price sensitivities by mills' organizational form and farmers' characteristics. We also correlate answers in the mill surveys with the estimated mill-district fixed effect η_{md}. These correlations help interpret which mill differentiation factors influence farmers' supply decisions.

[24]Multiple samples collected at different buying points ensures the accuracy of the estimated minimum price. Mills can appeal the minimum price proposal, making potential inaccuracies costly for ICAFE.

Bilateral Oligopoly. Since mills' and exporters' prices and marginal costs are observed in the data, we are able to recover the empirical distribution of the mill bargaining power (λ_{mes}) by inverting equation (9):

$$\lambda_{mes} = \frac{2(p_{mes} - c_{ms})}{p^w_{es} - c_{ms}}. \tag{14}$$

A high margin for mills relative to exporters in the forward market implies a high λ_{mes}. Equation (14) highlights a key advantage of applying our model to the Costa Rica coffee industry. Marginal costs are typically unobservable, and the existing literature relies on numerical methods to estimate parties' bargaining power (e.g., Crawford and Yurukoglu, 2012; Crawford et al., 2018). Instead, we directly recover the empirical distribution of bargaining parameters at the mill-exporter-season level consistent with observed margins.

Similarly, we recover the empirical distribution of the parameter governing exporters' risk sensitivity on the forward market (κ_{es}) by inverting equation (6):

$$\kappa_{es} = \frac{1}{|\mathcal{M}_{es}|} \sum_{m \in \mathcal{M}_{es}} \frac{p^w_{es} - p_{mes}}{q_{mes}}. \tag{15}$$

Since we observe all the variables on the right-hand side, we can directly infer κ_{es} from the data. Equation (15) implies that exporters with higher margins per quantity traded can sustain higher risks on the forward market. In Section 5.2, we explore how κ_{es} and λ_{mes} vary across mills' and exporters' organizational forms and characteristics. Notice that we estimate κ_{es} without requiring information on λ_{mes}, and vice versa.

Entry Costs. Following Eizenberg (2014), we use mills' revealed preferences to identify a range of fixed entry costs that support the observed distribution of prices and district choices as an equilibrium of our two-stage game.[25] This approach is appealing because it requires minimal assumptions about the true data-generating process, with the only condition being that agents are profit-maximizing, a realistic assumption for private mills in our context. As for cooperatives, we do not explicitly model their entry decisions across districts since they are bound to source from districts where their members are located.

We use the necessary conditions for an SPNE in equations (11) and (12) to derive upper

[25]Mills' fixed costs of entering local districts cannot be point-identified. Multiple equilibria prevent constructing a well-behaved likelihood function (Bresnahan and Reiss, 1991), so this holds even under stronger functional form assumptions for the fixed costs in equation (10).

and lower bounds on the fixed costs mills incur to source coffee from a district in each season:

$$F_{mds} \leq E_{u|\theta}\left[\pi_{ms}(\mathcal{A}_{ms}; u, \theta) - \pi_{ms}(\mathcal{A}_{ms} - 1^{d}_{ms}; u, \theta)\right] \equiv \overline{F}_{mds}(\theta) \quad \forall d \in \mathcal{A}^{1}_{ms}, \tag{16}$$

$$F_{mds} \geq E_{u|\theta}\left[\pi_{ms}(\mathcal{A}_{ms} + 1^{d}_{ms}; u, \theta) - \pi_{ms}(\mathcal{A}_{ms}; u, \theta)\right] \equiv \underline{F}_{mds}(\theta) \quad \forall d \in \mathcal{A}^{0}_{ms}. \tag{17}$$

To recover the upper bound of the fixed cost of entry into district d for mill m in season s, we simulate a counterfactual scenario where the mill does not enter that district while keeping all other mills' location decisions fixed. The upper bound is defined as the difference between the baseline and counterfactual expected variable profits. Similarly, we use the observed choices of districts where the mill did not source to derive a lower bound on the fixed cost of entering each district. See Online Appendix C.1 for computational details on the estimation of fixed costs.

This strategy identifies upper bounds for $d \in \mathcal{A}^{1}_{ms}$ (the set of districts in which mill m actually enters in season s) and lower bounds for $d \in \mathcal{A}^{0}_{ms}$ (the set of districts where mill m did not enter in season s). This is because mills make district choices knowing the realization of entry cost shocks (ν_{mds}), so the observed set of mills in district d is a selected sample with a favorable draw. A similar argument applies for lower bounds. To find bounds for all districts, we follow Eizenberg (2014) and assume the fixed cost bounds are constrained by the expected change in variable profits from removing or adding a district. Specifically, we assume the fixed cost of entering district d is bounded above by the largest change from removing a district, and below by the smallest change from adding one. In Section 5.2, we explore how fixed costs of entry F_{mds} vary with mills' organizational form and characteristics.

5.2 Estimation Results

Farmers' Supply. Table 1 reports the estimates from equation (13). For comparison, columns (1) to (3) report OLS estimates that correlate the market share of mill m in district d in season s with the price paid by the mill in that season, conditional on mill-district and district-season fixed effects, while columns (4) to (6) report IV estimates, with the price paid by the mill instrumented as described above. Our preferred specification is the IV in column (6), which allows for heterogeneous price sensitivity across mills with different organizational forms. Estimates from this specification are used to estimate the fixed costs of entry and to conduct the policy counterfactuals described in the next section.

Before turning to the results, we describe our instrument. Online Figure B.3 shows the distribution of the *rendimiento técnico* (z^{T}_{ms})—our instrument for prices—and the *rendimiento reportado* (z^{R}_{ms}). Both variables are standardized to have zero mean and unit variance in the sample. The *rendimiento técnico* is lower than the *rendimiento reportado* in 55% of the observations, indicating that observed prices are at the minimum allowed in around 45% of mill-season observations. The first-stage estimates in Online Table A.5 show that the instrument is relevant (Kleibergen-Paap F-statistic of 26.07).[26]

The comparison of OLS and IV baseline specifications in columns (1) and (4) reveals that the OLS estimate of α_1 is downward-biased. The average farmers' supply price elasticity implied by the OLS estimate in column (1) is 0.2, while the one implied by the IV estimate in column (4) is 3.1.[27] Our IV estimate of the elasticity of supply is consistent with that estimated by Dominguez-Iino (2024) across several agricultural commodities (maize, soybean, beef) in Brazil, but is higher than the elasticity estimated by Rubens (2023) for Chinese tobacco leaf producers using a production-function approach.[28]

The remaining columns of Table 1 explore the heterogeneity in farmers' supply sensitivity to prices across mills with different organizational forms. Both OLS and IV estimates find that farmers are *more* price-sensitive when selling to private integrated mills and cooperatives than to private non-integrated mills. A potential explanation is that private non-integrated mills are often local family businesses with long-standing relationships with farmers, while vertically integrated mills tend to source from more districts, often changing sourcing locations across seasons, and thus sourcing from farmers that are less loyal to the mill. At first, this explanation contrasts with the higher estimated price sensitivity towards cooperatives which, in theory, buy from loyal members. However, evidence from both Costa Rica (Hopfensitz and Miquel-Florensa, 2017) and other contexts (e.g., Banerjee et al., 2001; Casaburi and Macchiavello, 2015; Montero, 2022), finds that cooperatives often struggle with curbing members' side-selling due to their relatively fragile governance.

A critical aspect of the supply equation (13) is the inclusion of mill-district fixed effects (η_{md}) that capture non-price factors potentially driving farmers' supply. The top-left panel of Figure 2 illustrates the distribution of the estimated η_{md}, while the top-right panel shows its average by organizational form. Two features stand out. First, these factors shape farmers' supply decisions to a significant extent. On average, a one-standard-deviation increase in

[26] Using $z_{ms} = \max\{z^R_{ms}, z^T_{ms}\}$ as an instrument for prices we do not reject the null hypothesis that the second-stage coefficients are equal to those reported in Table 1. The point estimates are somewhat closer to OLS ones, consistent with our earlier observation that the *rendimiento reportado* z^R_{ms} is more likely to violate the exclusion restriction. We thus use the *rendimiento técnico* as instrument in our preferred specification.

[27] We calculate the own price elasticity of farmers' supply as $\alpha_1 P_{ms}(1 - S_{mds})$ and report the unweighted average elasticity across all mills, districts, and seasons. We also computed weighted elasticities using the total quantity of coffee sourced by mill m in district s during season s as weights. The weighted OLS elasticity is 0.13, while the weighted IV elasticity is 2.2.

[28] In the appendix, Rubens (2023) finds that a discrete choice framework à la Berry (1994)—which is closer to our model—produces results similar to his baseline production-function approach. In addition to differences in context, the different sources of variation used to identify farmers' supply may also contribute to the differences in estimated elasticities between our papers.

the mill-district fixed effect is equivalent to a 2.5 standard deviations higher price. At the district level, regressing farmers' surplus—computed from equation (2)—on mill-district fixed effects and the average prices paid by mills to farmers reveals that mill-district fixed effects account for approximately 84% of the variation in farmers' surplus across districts in the average season. In other words, non-price factors that drive farmers' supply choices across differentiated mills are extremely important for farmers. Second, decomposing the mill-district fixed effects into separate mill and district components reveals that the mill component accounts for approximately 60% of the variance of η_{md}. Mill-level characteristics, such as management practices and relationships with downstream buyers, drive farmers' choices.

Figure 3 leverages the mill survey to interpret estimated mill-district supply shifters (η_{md}). The survey elicited mills' offering of services potentially valuable to farmers, e.g., loans and cash advances before and during the harvest season (as in, e.g., Macchiavello and Morjaria, 2020), and agricultural inputs and participation in certification programs (as in, e.g., Macchiavello and Miquel-Florensa, 2024). Estimated supply shifters positively correlate with (*i*) the share of farmers receiving advance payments (top left), (*ii*) the share of farmers receiving loans from the mill (top right), (*iii*) the share of farmers receiving agricultural inputs (bottom left), and (*iv*) mills' participation in voluntary sustainability standards such as 4C, Starbucks Cafe Practices, and Nespresso AAA (bottom right). In sum, mills' provision of a wide range of services to farmers positively correlates with farmers' supply to the mill, conditional on prices.[29]

Table 2 explores differences across organizational forms in the estimated mill-district specific supply shifters, aggregated at the mill level. Column (1) finds that private integrated mills and cooperatives have higher estimated supply shifters than private non-integrated mills on average. This premium, however, is explained by their size. Column (2) controls for the (log of the) average total volume sourced by the mill across seasons. The estimated organizational form dummies are halved relative to column (1), and are no longer statistically significant. While mill's market power might depress prices paid to farmers, larger mills provide services that farmers value.

Finally, Online Table A.6 provides a preliminary exploration of heterogeneity in supply elasticity across farmers. The table reports IV results that allow the estimated α_1 to vary with the share of large farmers (column (1)) and male farmers (column (2)) in the district. Defining large farms as those that produce more than the average farm in Costa Rica,

[29]To preserve mills' anonymity, the figure aggregates the estimated average mill fixed effects and the survey responses by deciles of the mills' size distribution, measured by total coffee production. Survey responses were recorded anonymously in compliance with the ICAFE regulation.

column (1) finds some evidence that farmers are less sensitive to prices in districts in which there is a higher share of large farms.[30] The implied economic magnitude, however, is relatively modest. Column (2) finds no evidence of heterogeneity in price sensitivity by gender. Therefore, in the reminder of the paper, we abstract from heterogeneity across *farmers*, and focus on heterogeneity across *mills*.

Bilateral Oligopoly. A central feature of our framework is that mills bargain prices on the forward markets with exporters. Since we observe both the prices negotiated between the mills and the exporters, as well as the mills' marginal costs and the exporters' prices downstream, the estimation of bargaining power in equation (14) is straightforward. The middle-left panel of Figure 2 illustrates the empirical distribution of mills' bargaining power when trading on the forward market (λ_{mes}). The middle-right panel shows the average mill's bargaining power by organizational form. Three facts stand out. First, the average estimated λ_{mes} is equal to 0.61, i.e., on average mills have significant bargaining power in their negotiations with exporters. Second, bargaining power is highly dispersed. The standard deviation is 0.20, the 75^{th} and 25^{th} percentiles are 0.42 and 0.76 respectively, implying an inter-quartile range of 0.33. On average, private non-integrated mills negotiate higher prices than cooperatives and private integrated mills. Finally, a variance decomposition exercise reveals that in the average season, mill and exporter fixed effects account for 60% and 40% of the estimated variation, respectively.

Given the persistent differences in bargaining power across both mills and exporters, Table 3 explores the correlation between bargaining power and the characteristics of mills and exporters. There are two takeaways. First, mills are generally more powerful than exporters. Second, when trading outside their organization, integrated mills are less powerful and pay a price penalty, while integrated exporters are more powerful and receive a price premium. This pattern holds even when controlling for mill and exporter size. It can be explained by the fact that exporters trade more frequently outside their organization than mills (see Online Tables A.3 and A.4). Both patterns are consistent with vertical integration providing supply assurance to exporters, as documented in Macchiavello and Miquel-Florensa (2018).

Finally, after averaging both variables at the mill level, the correlation between the mill-district fixed effect (η_{mds}) and the mill's bargaining power in the forward market (λ_{mes}) is 15%. This correlation is statistically significant at the 5% level and remarks the importance of accounting for mills' bargaining power with downstream exporters to understand market shares and farmers' surplus upstream.

[30] It might be easier for mills to provide services to larger farmers, making them less price sensitive. Macchiavello and Miquel-Florensa (2024) and Abouaziza et al. (2024) find higher take-up rates of specific sustainable sourcing programs among larger coffee farmers in the Colombia and Rwanda contexts respectively.

The exporters' (linear) demand for forward coffee from each mill in equation (15) is also transparently identified, from direct observation of the exporters' prices in foreign markets and negotiated prices with the mills. Online Figure B.4 reports the empirical distribution of exporters' estimated sensitivity to risk when trading on the forward market (κ_{es}). The Figure illustrates the significant heterogeneity in exporters' volumes purchased on the forward market. Online Table A.7 shows that unconditionally (column (1)), as well as conditionally on size (column (2)), vertically integrated exporters demand larger volumes of coffee on the forward market, consistent with Macchiavello and Miquel-Florensa (2018)'s analysis of vertical integration as a strategy to achieve supply assurance.

Entry Costs. The bottom-left panel of Figure 2 shows the distribution of the upper and lower bounds of the fixed costs mills incur when sourcing coffee from districts in a given season, while the bottom-right panel shows the average entry cost bound by organizational form. Our estimation procedure recovers 3,575 upper bounds and 1,235 lower bounds. Online Table A.8 shows that the average cost ranges from approximately 425 USD to 673 USD, estimates that are in line with the actual cost of operating seasonal buying points in the country. Finally, Table 4 shows that private integrated mills face higher entry costs. The estimates are consistent with the evidence that vertically integrated mills' source larger volumes per district—suggesting higher fixed costs to establish a denser network of buying points that remain open throughout the harvest season—but also source in more districts, so that higher fixed costs are incurred to operate further away from the mills' premises.

6 Regulating the Coffee Value Chain

In this section, we use the model to evaluate the effects of two common policy interventions that aim at fostering competition and protecting farmers in agricultural chains: (i) revenue-sharing regulations, such as the one implemented by ICAFE, and (ii) bans on vertical integration between mills and exporters. We investigate the effects of these regulatory changes along the chain, i.e., on farmers, mills, and exporters, as well as their heterogeneous effects across local markets (districts). For each intervention, we analyze two scenarios: one in which mills do not endogenously adjust the districts where they source from, and one in which they do. The comparison highlights the role played by endogenous entry/exit responses in shaping the distributional consequences of each regulatory change.

6.1 Revenue-Sharing Regulation

Policy Design and Trade-Off. As described in section 2, the ICAFE regulation is designed to achieve an equitable functioning of the coffee chain. In particular, the ICAFE regulation prescribes that mills retain $\mu = 9\%$ of the difference between the revenues generated from sales and the incurred processing costs (see equation (4)). The rest is transferred to farmers as a liquidation price at the end of each harvest season. This type of regulation is by no means unique to the Costa Rica coffee chain. Panel A in Online Table A.9 reports other contexts in which similar revenue-sharing rules are in place, such as coffee in Burundi and tea in Kenya and Rwanda.[31]

To better understand the trade-offs faced by the regulator, we investigate how the market equilibrium would look under different values of μ. On one hand, reducing μ directly increases farmers' surplus and reduces mills' profits. On the other hand, a lower μ could also lead some mills to stop sourcing from certain districts. This would benefit incumbent mills but indirectly harm farmers. The regulator faces the opposite trade-off if μ increases. We consider small changes around the 9% value currently mandated by the regulation. We focus on small changes because the counterfactual analysis depends on structural parameters estimated with data generated under the existing regulation.[32]

We focus on the direct effects of changes in μ on farmers' welfare and mills' profits. By altering the total quantity of coffee sourced by each mill, changes in μ could also affect exporters' profits. Our model, however, abstracts from these effects. First, given our assumptions on exporters' demand and the bargaining protocol, μ does not alter the quantity or price negotiated between mills and exporters in equations (6) and (9). Second, for tractability reasons, we follow the existing literature (Lee et al., 2021) and hold the bilateral network of negotiations between mills and buyers fixed in the counterfactual. Any counterfactual change in quantity sourced by a mill is absorbed by the spot market. As a result, the policy has no direct impact on exporters.

Implementation. We simulate the effects of increasing or decreasing μ under two scenarios: when mills cannot change the districts they source from, and when they can. Implementing the counterfactual scenario without mills' entry and exit responses is straightforward. After changing μ, we update equation (4) and compute farmers' surplus and mills' profits using equations (2) and (3), respectively.

Implementing the scenario where mills can adjust entry and exit across districts is more

[31] Other countries in Central America attempted to adopt a revenue-sharing system similar to the Costa Rica one but encountered political resistance from powerful elites (Paige, 1997).

[32] Furthermore, completely removing the revenue-sharing might alter farmers' exposure to price risk and mills' working capital needs, aspects which lie outside the scope of our model.

challenging. As discussed by Eizenberg (2014), set identification of mills' entry costs keeps estimation computationally efficient. However, in the counterfactual, it delivers a set of feasible allocations of mills across districts rather than a single allocation. To determine whether an allocation can be sustained as a counterfactual equilibrium, one must guess an initial configuration of mills' entry choices across districts and then check if any mill has a profitable unilateral deviation, given the estimated set of entry costs. This process is computationally demanding, as the number of configurations to consider grows exponentially with the number of mills and districts. For example, with 5 mills and 5 districts, there are $2^{25} = 33,554,432$ potential equilibria to check.

We address this computational challenge by building on previous literature (e.g., Lee and Pakes, 2009; Fan and Yang, 2020; Montag, 2023). Specifically, we allow mills to adjust their sourcing strategies sequentially, starting from the observed allocation in the data and taking rivals' sourcing choices as given, until no mill has a unilateral profitable deviation. To determine the sequence in which mills are allowed to move, we average the estimated mill-district fixed effects from equation (13) at the mill level, denoting with $\overline{\eta}_m$ the resulting average, and rank mills accordingly. After a policy change, mills enter districts sequentially in decreasing order of this ranking.

We consider two cases: decreasing or increasing μ. When μ decreases, mills' margins shrink, potentially incentivizing them to exit districts. To compute the counterfactual allocation, we first replicate the steps of the scenario without responses in sourcing choices. Then, we check if any mills make negative profits in their districts. If so, these mills exit, and we recompute the equilibrium. We iterate this process in each district and season until all mills earn non-negative profits in all districts. When raising μ, mills' margins increase, and mills may have an incentive to start sourcing from new districts. In this case, we let the highest-ranked mill that has not entered a district enter that district. We iteratively let mills enter until an incumbent would make negative profits. We repeat this procedure for each district and season in the data.[33]

Results. The top-left panel of Figure 4 shows the distribution of farmers' surplus across districts (averaged over seasons) when mills cannot adjust their sourcing choices. The top-right panel shows the same distribution when mills endogenously change the set of districts they source from. The bottom-left and right panels display the distribution of mills' profits (summed across districts and averaged over seasons) when mills keep sourcing choices fixed and when they change them, respectively. Online Figure B.5 shows the change in the number of mills, both in levels and relative to the number of incumbents.

[33]In all cases, we compute counterfactual equilibria using 1,000 bootstrap draws from the joint empirical distribution of ξ_{mds}, η_{md}, and F_{mds}. See Online Appendix C.2 for additional details.

In the first scenario, when mills do not adjust their sourcing strategies, increasing μ harms all farmers and benefits all mills, whereas decreasing μ has the opposite effect. The magnitude of these direct effects increases with the absolute value of μ. In the second scenario, while these direct effects still dominate on average, mills' endogenous sourcing responses introduce two additional effects. First, they dampen both gains and losses. For instance, when μ increases, mills expand into new districts, partially offsetting the direct losses experienced by the average farmer. Second, these responses create heterogeneity in policy outcomes across mills and districts. In some districts, where pro-competitive effects are strong enough, farmers may even be better off than under the initial equilibrium. Thus, failing to account for mills' endogenous adjustments leads to an overestimation of both gains and losses in absolute terms and overlooks the distributional consequences of policy changes.

Focusing on the scenario with endogenous entry and exit of mills, Figure 5 explores further efficiency and distributional aspects of the regulation. By reallocating market shares across heterogeneous mills, the policy change also affects allocative efficiency. The top-left panel of Figure 5 shows a negative correlation between changes in market share and mills' average appeal ($\overline{\eta}_m$) when μ decreases, indicating that the exit of some mills from certain districts worsens allocative efficiency. In contrast, the top-right panel shows a positive correlation when μ increases. This occurs because increasing μ triggers the entry of new, more appealing mills.[34] The bottom panels of Figure 5 show that, under both scenarios, changes in farmers' surplus are more pronounced in districts with a higher initial share of large farmers. In contrast, there is no evidence of heterogeneous effects based on the initial share of female farmers in a district.

Overall, the results of this counterfactual analysis indicate a need for caution regarding policies designed to redistribute surplus along the supply chain. This is particularly important if the strategic entry and exit behaviors of mills could undermine the original objectives of these policies. In fact, reallocating resources may distort the equilibrium allocation and decrease allocative efficiency.

6.2 Banning Vertical Integration

Policy Design and Trade-Off. A common concern in agricultural chains is that vertical integration between traders, or first-stage processors, and exporters may facilitate market foreclosure and depress payments to farmers. Indeed, it is not unusual for vertical integration across stages of the chain to be regulated, if not banned altogether. Panel B in the Online

[34] At first, it may seem counterintuitive that allocative efficiency moves in the opposite direction of farmers' surplus. However, as shown in equation (2), farmers' surplus depends on the characteristics of the available options, rather than changes in market shares. As a result, the two patterns can coexist within our model.

Table A.9 reports several contexts in which bans on vertical integration are in place. For example, Kenya requires coffee to be exported through centralized auctions. Exporters do not have licenses to trade coffee before it reaches the auctions, *de facto* banning vertical integration. Ethiopia operated a similar auction system, replaced in 2008 with the creation of the commodity exchange, *ECX*, under which vertical integration was still banned. Only in 2017, the Ethiopian Coffee and Tea Authority overseeing *ECX* loosened the rules to allow for vertical integration, but only for some specialty-grade coffee under a strict licensing regime. Looking beyond coffee, cocoa in Ghana and Ivory Coast, cotton in Tanzania, rice in Vietnam and Thailand, palm oil in Indonesia, rubber in both Malaysia and Thailand, are all contexts in which exporters cannot vertically integrate to source directly from farmers.

Our model captures concerns about a negative impact of vertical integration on farmers' welfare through the estimated bargaining weights in Figure 2 (middle panels) and Table 3: vertically integrated mills negotiate lower prices than comparable non-integrated private mills, and all mills obtain lower prices when supplying vertically integrated exporters, resulting in lower payments to farmers. Notwithstanding these concerns, Table 2 revealed that vertically integrated mills have higher supply-shifter estimates than other non-integrated private mills, likely because their more secure marketing channel downstream enables them to offer services valued by farmers upstream. The effect of vertical integration on farmers' welfare is thus potentially ambiguous. Given the widespread adoption of bans on vertical integration, we use our model to quantify the effects of simulated forced divestitures between mills and exporters in the Costa Rica coffee chain.

Implementation. We simulate counterfactual scenarios in which vertical integration between mills and exporters is prohibited. To perform this counterfactual, we must make assumptions about what would happen to integrated mills and exporters after the forced divestiture. We assume that following the regulatory change, both integrated mills and exporters would stay in the market and be "transformed" into private non-integrated ones.

To convert private integrated mills and exporters, we assign them the supply shifters (η_{md}), marginal costs (c_{ms}), entry costs $[\underline{F}_{mds}, \overline{F}_{mds}]$, price sensitivity ($\alpha_1$), yield ($z_{ms}$), forward risk sensitivity (κ_{es}), bargaining power (λ_{mes}), and export prices (p^w_{es}) of non-integrated ones. We do so by regressing each variable on a constant and a dummy equal to one if the mill or the exporter are integrated. The resulting integration premium is then subtracted from the observed, or estimated, values of each variable for each integrated mill and exporter.

As before, we consider two scenarios. In the first, mills are not allowed to change the set of districts they source from. In the second, they can change their sourcing districts. Also in this case, we let mills move in decreasing order of their average appeal. However, we modify the algorithm as follows. First, we simulate divestitures by discounting the relevant variables for integrated mills and exporters. Next, we check if all mills make positive profits in each district. If all mills are profitable, we simulate entry until some mill would make negative profits. Conversely, if some mills make negative profits, we let them exit the district until all remaining mills are profitable. We repeat this procedure for all districts and seasons.

Results. Figure 6 shows the distribution of farmers' surplus, mills' profits, and exporters' profits following the implementation of divestitures. When mills cannot change the districts they source from, banning integration makes all farmers worse off. The positive effect on prices obtained by mills in the negotiations with exporters is more than offset by the negative effect stemming from the reduction in services provided by mills, as estimated by the supply shifters. Formerly integrated private mills experience profit losses, while their competitors (i.e., cooperatives and non-integrated private mills) gain market shares and experience profit gains. While the average formerly integrated private exporter incurs a loss, some are better off due to a reduction in forward risk costs (κ_{es}) following the divestiture.

Allowing mills to change the set of districts from which they source reduces the losses of the losers and the gains of the winners. This occurs because formerly integrated private mills face lower entry costs, and some begin sourcing from new districts despite lower appeal and higher marginal costs. These pro-competitive effects of divestitures slightly reduce the drop in farmers' surplus and substantially erode the profit gains of incumbents.

The top-left panel of Figure 7 shows a positive correlation between changes in market share and mills' average appeal ($\overline{\eta}_m$) following divestitures, suggesting a modest improvement in efficiency.[35] In contrast to Section 6.1, the top-right panel finds no statistically significant correlation between changes in farmers' surplus and their initial characteristics.

Overall, even if private integrated mills pay lower prices than non-integrated ones, as shown in Online Table A.1, they have higher appeal due to the services they provide, as highlighted in Table 2 and Figure 3. Therefore, while divestitures may raise the average prices mills pay to farmers, the loss of services previously offered by integrated mills ultimately leaves farmers worse off. In terms of policy implications, our results suggest that in contexts where transactions are interlinked, fostering price competition among mills alone is insufficient to ensure farmers' welfare.

In interpreting these results, it is important to recall that the regulation in Costa Rica sets a minimum price that mills must pay to farmers. While the estimated effects of a ban could potentially be less negative in a context with weaker institutional protections for farmers, two considerations suggest that the trade-offs identified in this paper are likely relevant in

[35] The improvement in allocative efficiency may appear to conflict with the reduction in farmers' surplus. However, as shown in equation (2), farmers' surplus depends on the characteristics of the available options, rather than changes in market shares.

a variety of contexts. First, minimum price regulations are also extremely common both in contexts that ban vertical integration (e.g., cocoa in Ghana and Ivory Coast, cotton in Tanzania, and rice in Vietnam) and in contexts that do not (e.g., coffee in Brazil and Colombia and across multiple crops in India). Second, both counterfactuals highlight how policies that curb profits of intermediaries can ultimately hurt farmers by reducing entry in local markets.

7 Conclusions

Linking smallholder farmers to markets, both domestic and foreign, has the potential to reduce poverty and support structural transformation. Several market imperfections, however, plague agricultural chains in developing countries. Not only are markets for farming inputs often poorly functioning, but output markets are also concentrated and monopsonistic. Governments regulate the processing and marketing of agricultural products, often with the explicit goal of protecting farmers.

This paper develops a structural model that captures the oligopolistic structure of agricultural chains and farmers' valuation of non-price factors in choosing between available marketing channels. The structural parameters of the model are estimated in the Costa Rica coffee chain and used to quantify the trade-offs involved in, and welfare effects of, common regulatory interventions. We provide insights into the structural reasons behind the dominance of large, vertically integrated chains: by securing market access downstream, these vertically integrated organizations can offer better services to farmers upstream and source coffee from a broader range of locations. Through counterfactual analysis, we show that common pro-competitive policy interventions, such as revenue-sharing rules and bans on vertical integration, may have limited effects on farm-gate prices and even make farmers worse off by curbing access to valuable services provided by processors and intermediaries.

This paper only scratches the surface of a promising, and important, area for further work, and we hope that it will spur further applications of industrial organization tools to policy evaluation in agricultural chains. Counterfactual analysis along the lines developed in this paper should be used to evaluate other common policy interventions, such as minimum farm-gate prices and zoning regulations. This paper has also taken farmers' investments as given. Investments in farms, however, are important for quality upgrading (Macchiavello and Miquel-Florensa, 2024)—a margin we abstracted from by focusing on commodity-grade coffee—and, increasingly, for both mitigation and adaptation to climate change (Abouaziza et al., 2024). Understanding how regulatory interventions downstream shape farmers' investments is a priority area for future work.

References

Abouaziza, Mohamed, Rocco Macchiavello, Ameek Singh, and Iris Steenkamp, "Building Relationships for Sustainable Sourcing with Rwanda Smallholder Coffee Farmers," 2024.

Alviarez, Vanessa I, Michele Fioretti, Ken Kikkawa, and Monica Morlacco, "Two-sided Market Power in Firm-to-Firm Trade," Technical Report, National Bureau of Economic Research 2023.

Ashraf, N., X. Ginè, and D. Karlan, "Finding Missing Markets (and a Disturbing Epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya," *American Journal of Agricultural Economics*, 2009, *91* (4), 973–990.

Avignon, Rémi and Etienne Guigue, "Markups and Markdowns in the French Dairy Market," Technical Report, Technical report, mimeo 2022.

Banerjee, Abhijit, Dilip Mookherjee, Kaivan Munshi, and Debraj Ray, "Inequality, Control Rights, and Rent Seeking: Sugar Cooperatives in Maharashtra," *Journal of Political Economy*, 2001, *109* (1), 138–190.

Bergquist, Lauren F, Benjamin Faber, Thibault Fally, Matthias Hoelzlein, Edward Miguel, and Andrés Rodríguez-Clare, "Scaling Agricultural Policy Interventions," Technical Report, National Bureau of Economic Research 2022.

Bergquist, Lauren Falcao and Michael Dinerstein, "Competition and Entry in Agricultural Markets: Experimental Evidence from Kenya," *American Economic Review*, 2020, *110* (12), 3705–3747.

Berry, S., J. Levinsohn, and A. Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

Berry, Steven T, "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, pp. 242–262.

Blouin, Arthur and Rocco Macchiavello, "Strategic Default in the International Coffee Market," *Quarterly Journal of Economics*, 2019, *134* (2), 895–951.

Bold, T., K. C. Kaizzi, J. Svensson, and D. Yanagizawa-Drott, "Lemon Technologies and Adoption: Measurement, Theory, and Evidence from Agricultural Markets in Uganda," *Quarterly Journal of Economics*, 2017, *132* (3), 1055–1100.

Boudreau, Laura, Julia Cajal-Grossi, and Rocco Macchiavello, "Global Value Chains in Developing Countries: a Relational Perspective from Coffee and Garments," *Journal of Economic Perspectives*, 2023, *37* (3), 59–86.

Bresnahan, Timothy F and Peter C Reiss, "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 1991, *99* (5), 977–1009.

Brugues, Felipe, "Take the Goods and Run: Contracting Frictions and Market Power in Supply Chains," *Working Paper*, 2024.

Burke, Marshall, Lauren Falcao Bergquist, and Edward Miguel, "Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets," *The Quarterly Journal of Economics*, 2019, *134* (2), 785–842.

Cai, H., Y Chen, H. Fang, and L. Zhou, "The Effect of Microinsurance on Economic Activities: Evidence from a Randomized Field Experiment," *Review of Economics and Statistics*, 2015, *97* (2), 287–300.

Casaburi, L. and J. Willis, "Time vs. State in Insurance: Experimental Evidence from Contract Farming in Kenya," *American Economic Review*, 2018, *108* (12), 3778–3813.

_ and R. Macchiavello, "Demand and Supply of Infrequent Payments as a Commitment Device: Evidence from Kenya," *American Economic Review*, 2019, *109* (2), 523–555.

Casaburi, Lorenzo and Jack Willis, "Value Chain Microfinance," *Oxford Review of Economic Policy*, 2024, *40* (1), 160–175.

_ and Rocco Macchiavello, "Loyalty, Exit, and Enforcement: Evidence from a Kenya Dairy Cooperative," *American Economic Review*, 2015, *105* (5), 286–290.

_ and Tristan Reed, "Using Individual-Level Randomized Treatment to Learn about Market Structure," *American Economic Journal: Applied Economics*, 2022, *14* (4), 58–90.

Collard-Wexler, Allan, Gautam Gowrisankaran, and Robin S Lee, ""Nash-in-Nash" Bargaining: a Microfoundation for Applied Work," *Journal of Political Economy*, 2019, *127* (1), 163–195.

Crawford, Gregory S and Ali Yurukoglu, "The Welfare Effects of Bundling in Multichannel Television Markets," *American Economic Review*, 2012, *102* (2), 643–685.

_ , Robin S Lee, Michael D Whinston, and Ali Yurukoglu, "The Welfare Effects of Vertical Integration in Multichannel Television Markets," *Econometrica*, 2018, *86* (3), 891–954.

Dominguez-Iino, Tomas, "Efficiency and Redistribution in Environmental Policy: An Equilibrium Analysis of Agricultural Supply Chains," 2024. Working Paper.

Duflo, E., M. Kremer, and J. Robinson, "How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya," *American Economic Review*, 2008, *98* (2), 482–488.

_ , _ , and _ , "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," *American Economic Review*, 2011, *101* (6), 2350–2390.

Eizenberg, Alon, "Upstream Innovation and Product Variety in the US Home PC Market," *Review of Economic Studies*, 2014, *81* (3), 1003–1045.

Fan, Y. and C. Yang, "Estimating Discrete Games with Many Firms and Many Decisions: An Application to Merger and Product Variety," 2022. Working Paper.

Fan, Ying and Chenyu Yang, "Competition, Product Proliferation, and Welfare: A Study of the US Smartphone Market," *American Economic Journal: Microeconomics*, 2020, *12* (2), 99–134.

Garg, Shresth and Sagar Saxena, "Redistribution Through Prices in Indian Agricultural Markets," Technical Report, working paper 2022.

Hopfensitz, Astrid and Josepa Miquel-Florensa, "Mill Ownership and Farmer's Cooperative Behavior: the Case of Costa Rica Coffee Farmers," *Journal of Institutional Economics*, 2017, *13* (3), 623–648.

Horn, Henrick and Asher Wolinsky, "Bilateral Monopolies and Incentives for Merger," *The RAND Journal of Economics*, 1988, pp. 408–419.

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry, "Agricultural Decisions after Relaxing Credit and Risk Constraints," *Quarterly Journal of Economics*, 2014, *129* (2), 597–652.

Lee, Robin S and Ariel Pakes, "Multiple Equilibria and Selection by Learning in an Applied Setting," *Economics Letters*, 2009, *104* (1), 13–16.

_ , Michael D Whinston, and Ali Yurukoglu, "Structural Empirical Analysis of Contracting in Vertical Markets," *Handbook of Industrial Organization*, 2021, *4* (1), 673–742.

Leong, Kaiwen, Huailu Li, Marc Rysman, and Christoph Walsh, "Law Enforcement and Bargaining over Illicit Drug Prices: Structural Evidence from a Gang's Ledger," *Journal of the European Economic Association*, 2022, *20* (3), 1198–1230.

Loecker, Jan De and Paul Scott, "Markup Estimation using Production and Demand Data. An Application to the US Brewing Industry," *Unpublished paper*, 2022.

Macchiavello, R. and A. Morjaria, "Competition and Relational Contracts: Evidence from Rwanda's Coffee Mills," *Quarterly Journal of Economics*, 2020, *136* (2), 1089–1143.

Macchiavello, Rocco, "Relational Contracts and Development," *Annual Review of Economics*, 2022, *14* (1), 337–362.

_ and Josepa Miquel-Florensa, "Vertical Integration and Inter-Firm Relationships in the Costa Rica Coffee Chain," *London School of Economics, Mimeo*, 2018.

_ and _ , "The Sustainable Quality Program in the Colombia Coffee Chain," 2024.

_ , Thomas Reardon, and Timothy J Richards, "Empirical Industrial Organization Economics to Analyze Developing Country Food Value Chains," *Annual Review of Resource Economics*, 2022, *14* (1), 193–220.

Montag, Felix, "Mergers, Foreign Competition, and Jobs: Evidence from the US Appliance Industry," *George J. Stigler Center for the Study of the Economy & the State Working Paper*, 2023, (326).

Montero, Eduardo, "Cooperative Property Rights and Development: Evidence from Land Reform in El Salvador," *Journal of Political Economy*, 2022, *130* (1), 48–93.

Nevo, A., "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 2000, *9* (4), 513–548.

Paige, Jeffery M, *Coffee and Power: Revolution and the Rise of Democracy in Central America*, Harvard University Press, 1997.

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii, "Moment Inequalities and Their Application," *Econometrica*, 2015, *83* (1), 315–334.

Rubens, Michael, "Market Structure, Oligopsony Power, and Productivity," *American Economic Review*, 2023, *113* (9), 2382–2410.

Technoserve, "Coffee Wet Mill Processing Guide," Technical Report 2020.

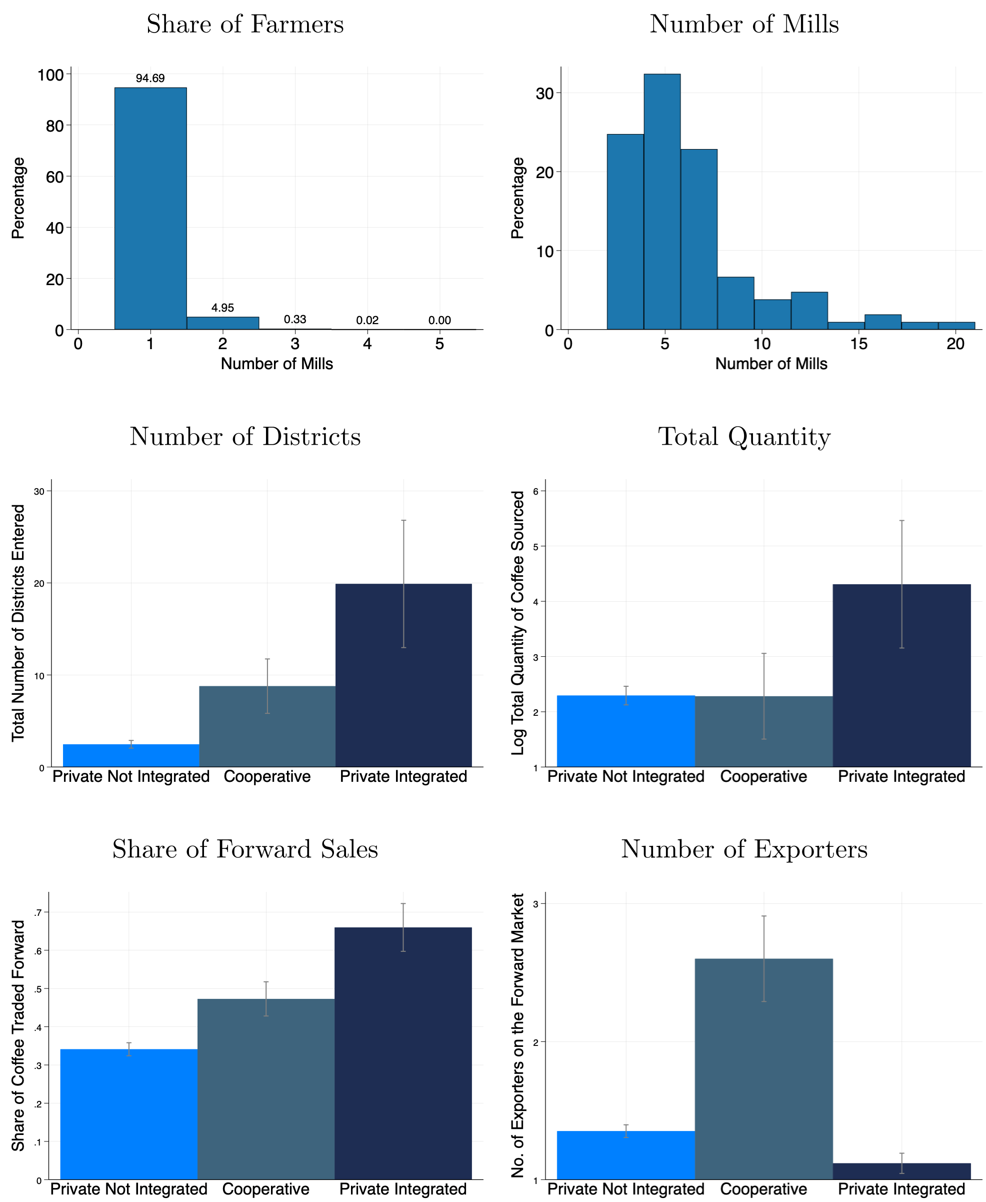
Wollmann, T. G., "Trucks Without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles," *American Economic Review*, 2018, *108* (6), 1364–1406.

World Bank, *World Development Eeport 2008: Agriculture for Development*, The World Bank, 2007.

_ , "Enabling the Business of Agriculture 2017," 2017.

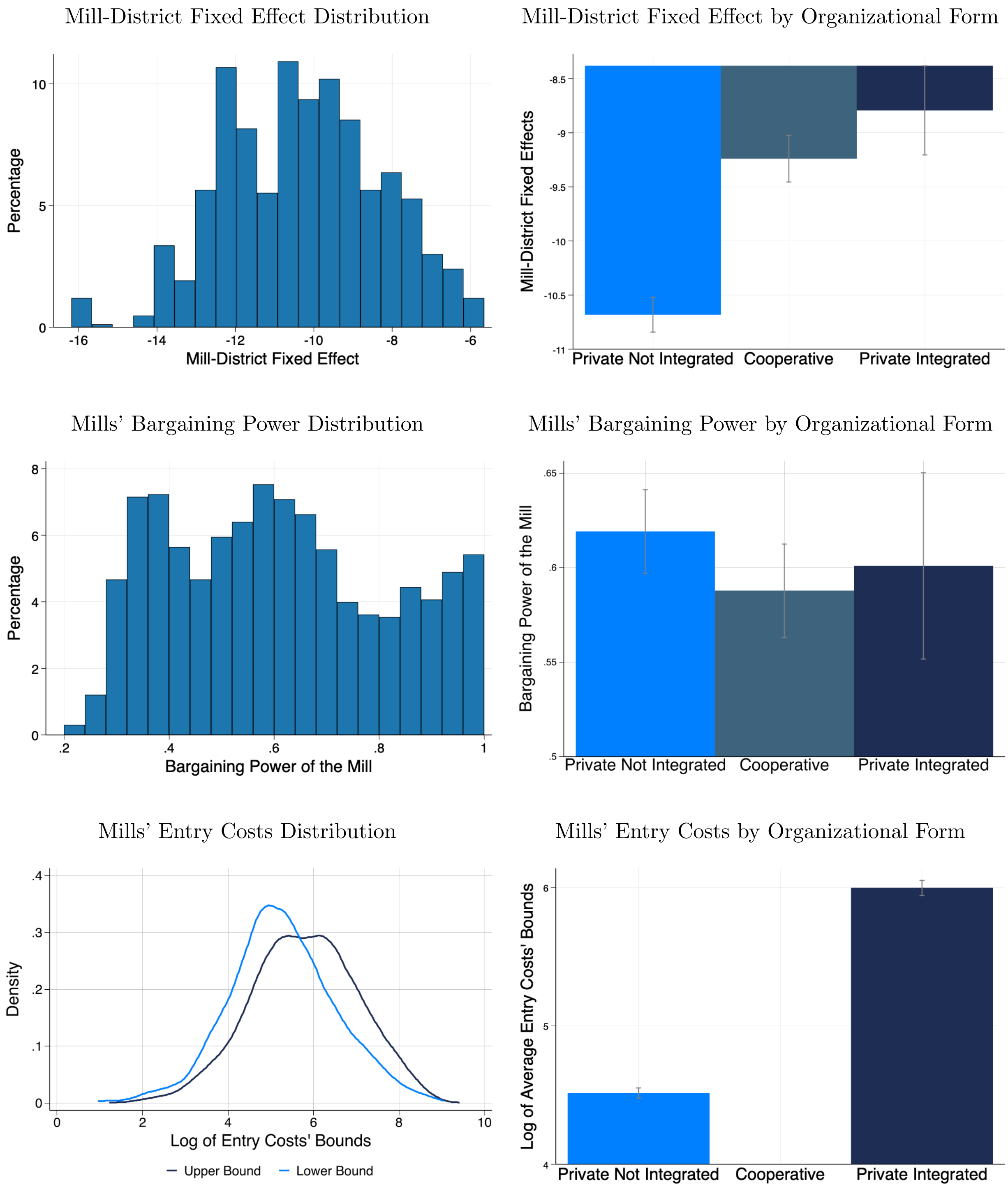
Figures and Tables

Figure 1. Motivating Facts



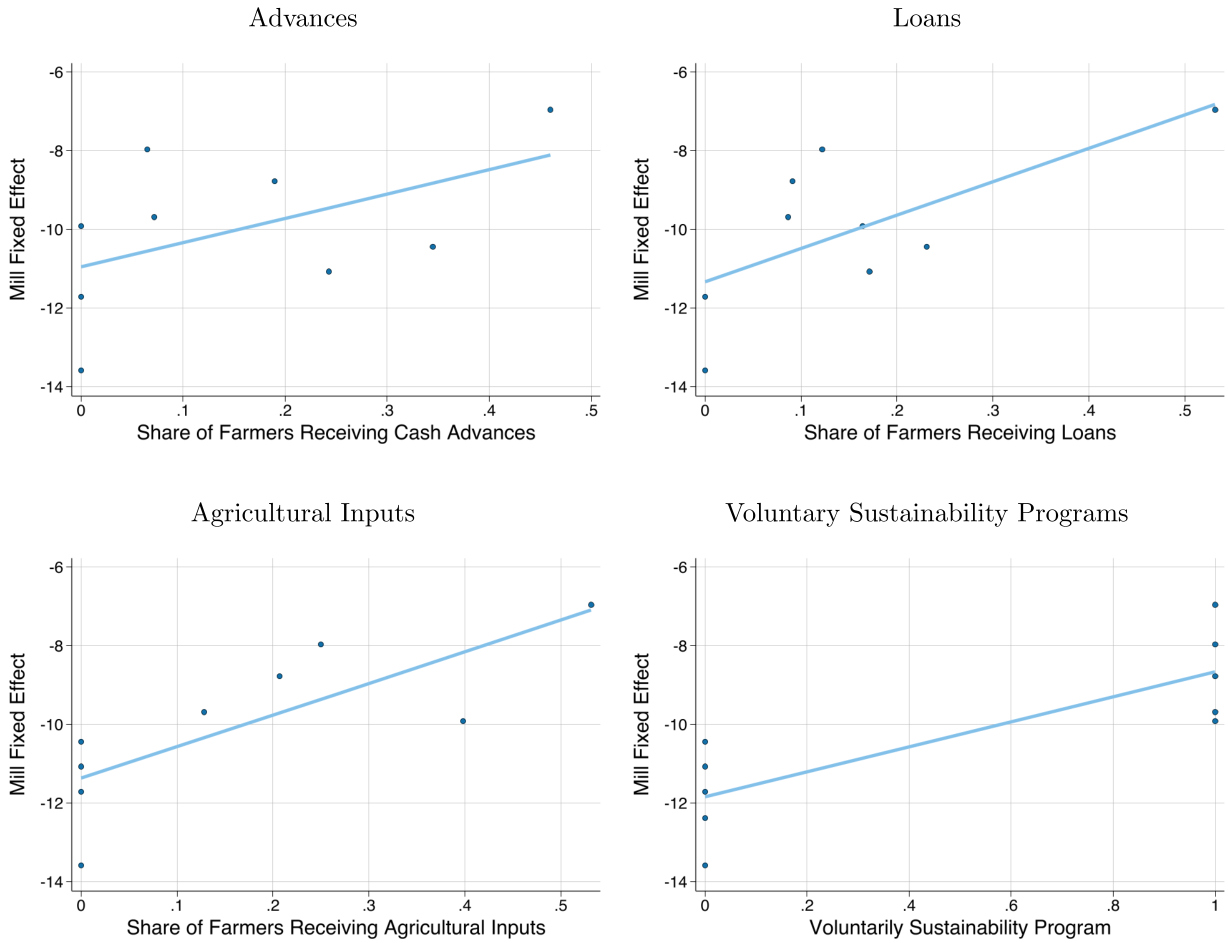
Note: The top panel contains the figures corresponding to Fact 1. The top-left panel shows the share of farmers that sell to a given number of mills within a district-season. The top-right panel shows the distribution of the number of mills in a given district (averaged across seasons). The middle panel contains the figures corresponding to Fact 2. The middle-left panel shows the total number of districts that mills belonging to different organizations enter. The middle-right panel shows the log of the total quantity of coffee (in kilograms) sourced by mills belonging to different organizations. The bottom panel contains the figures corresponding to Fact 3. The bottom-left panel shows the share of coffee that mills belonging to different organizations trade using forward contracts. The bottom-right panel shows the number of exporters that mills belonging to different organizations sell to using forward contracts.

Figure 2. Estimation Results



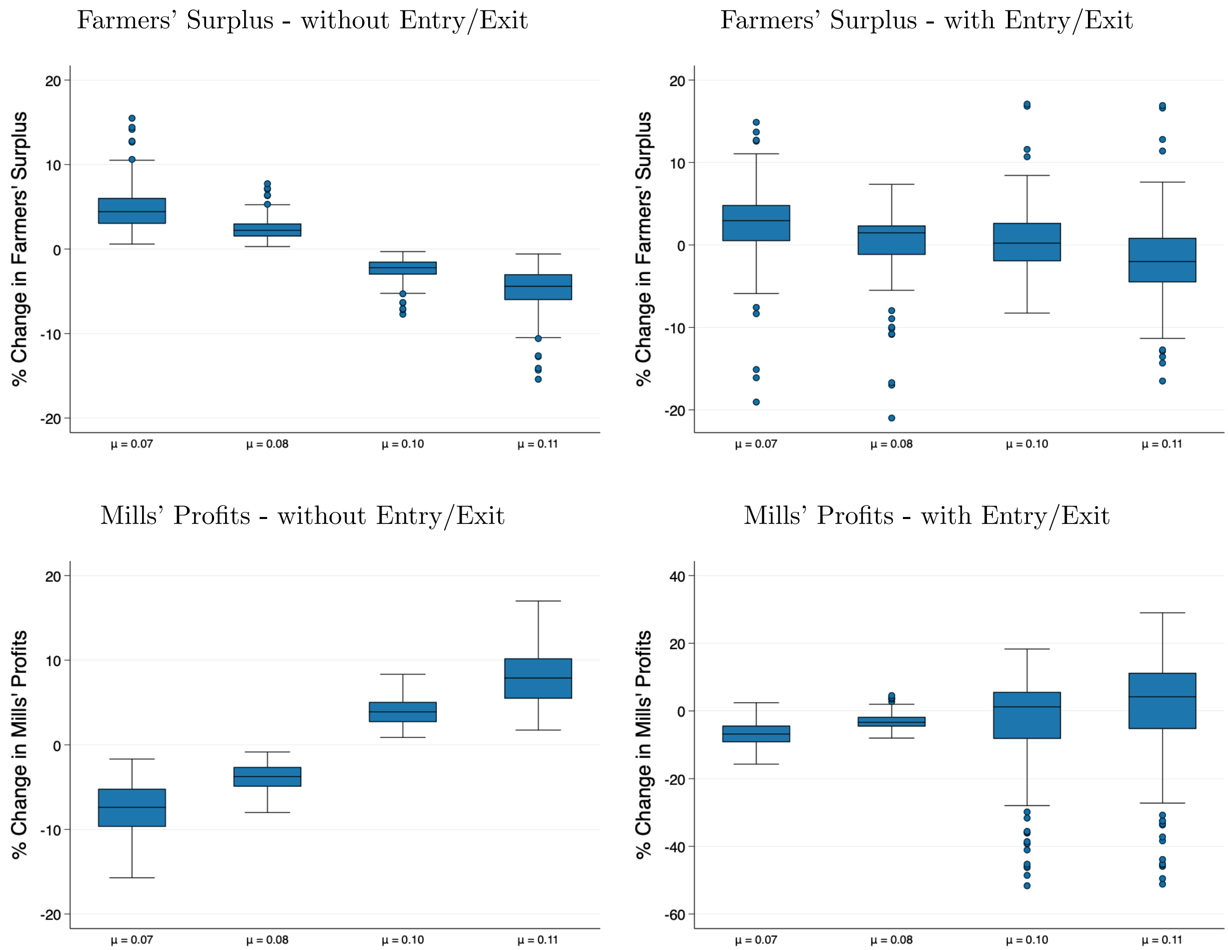
Note: The left panels display the distributions the mill-district fixed effects from equation (13) (top), the mills' bargaining power when trading with a given exporter in a given season on the forward market from equation (14) (middle), and the log of the upper and lower bounds of the fixed costs mills incur when entering districts in a given season from equations (16) and (17) (bottom). The right panels show the average values of these variables by organizational form. Since cooperatives operate as non-strategic firms, we do not model their entry decisions.

Figure 3. Interlinked Transactions and Mills' Appeal to Farmers



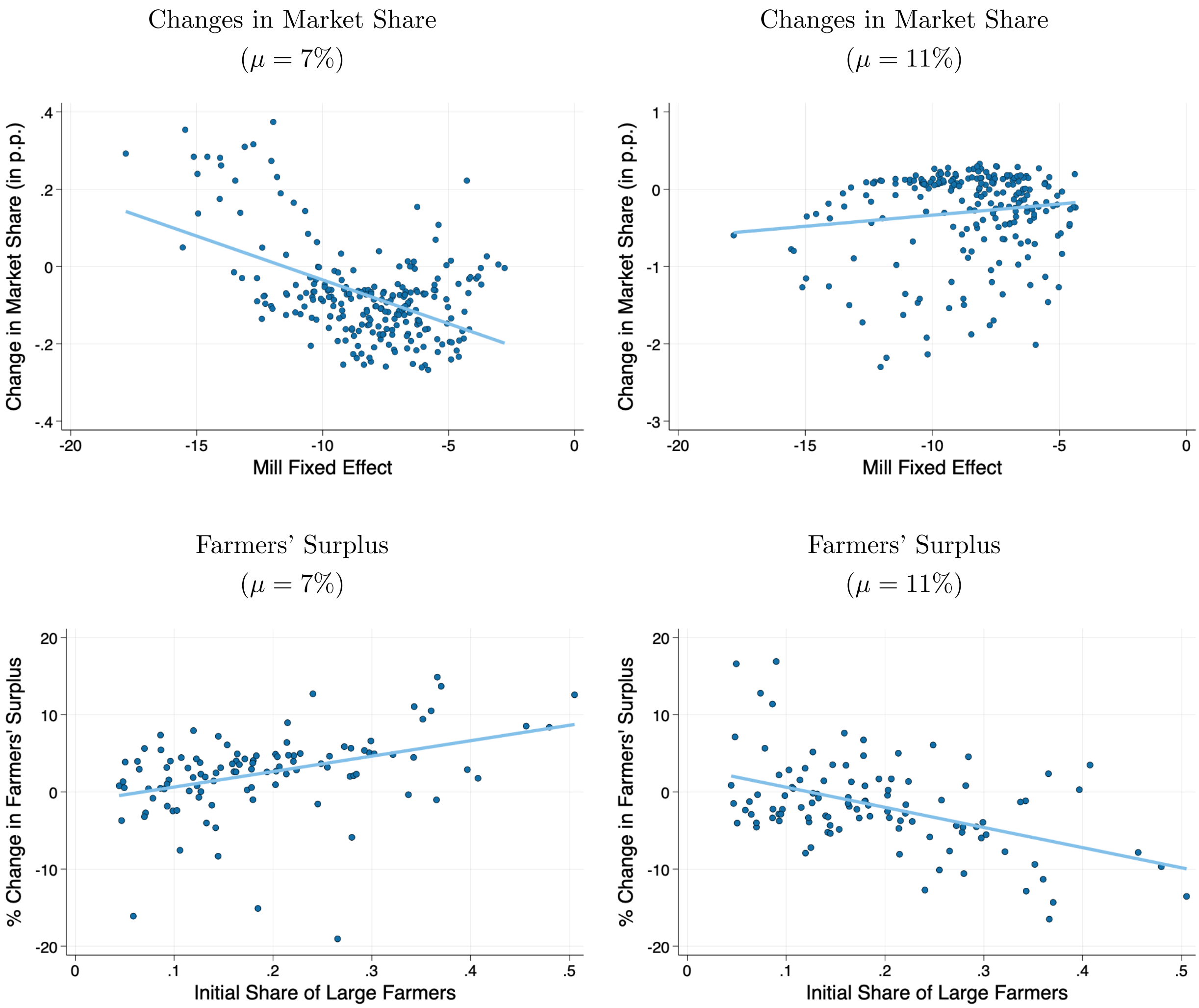
Note: We correlate mill fixed effects from equation (13) with survey responses, aggregated by mill size deciles due to survey anonymity. The top panel show the correlation with the share of advance payments to farmers (left) and loans (right). The bottom panel covers agricultural inputs (left) and mills' participation in voluntary sustainability programs, including Rainforest Alliance, 4C, Starbucks, or Nespresso (right).

Figure 4. Rent Sharing Counterfactual - Farmers' Surplus and Mills' Profits



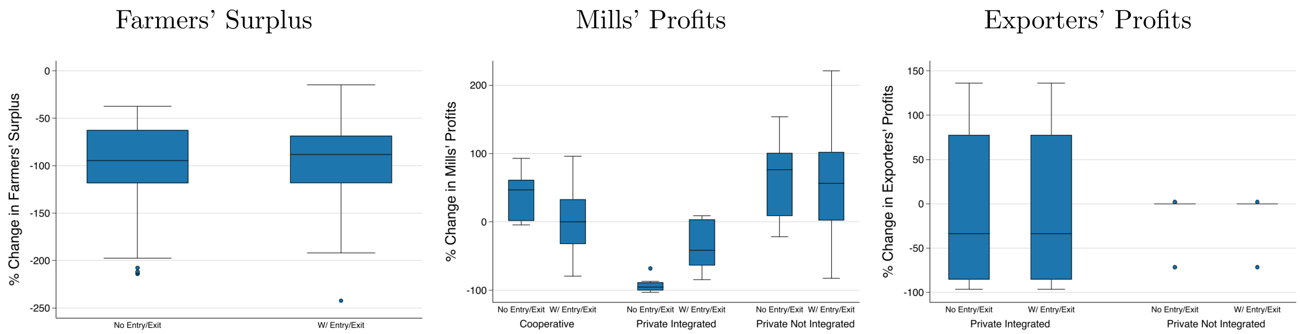
Note: The top-left panel shows the distribution of changes in farmers' surplus across districts (averaged over seasons) when μ is decreased or increased and mills do not change their sourcing choices. The top-right panel displays the same distributions when mills are allowed to endogenously change the set of districts they source from. The bottom-left panel shows the distribution of changes in mills' profits (summed over districts and averaged over seasons) when μ is decreased or increased and mills do not change their sourcing choices. The bottom-right panel displays the same distributions when mills are allowed to endogenously change the set of districts they source from. All changes are relative to the initial equilibrium with $\mu = 9\%$.

Figure 5. Rent Sharing Counterfactual - Additional Results



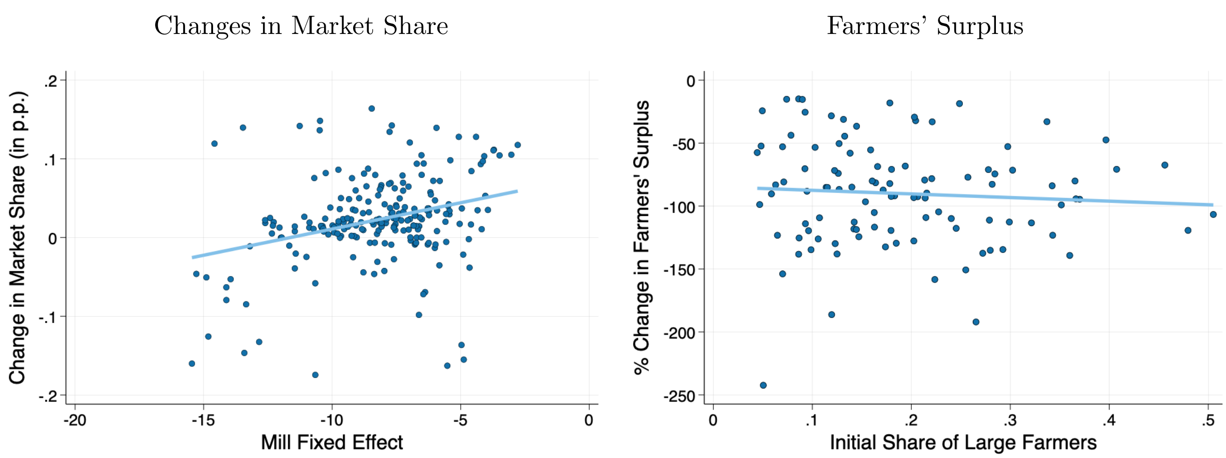
Note: The top panels display the correlation between changes in market share (measured in percentage points, p.p.) for each mill (averaged across districts and seasons) and the average mill appeal (calculated by averaging the mill-district fixed effects across districts). The left panel shows changes in market shares when μ is lowered from 9% to 7%. The right panel shows changes in market shares when μ is increased from 9% to 11%. The bottom panels show the correlation between changes in farmers' surplus in each district (averaged over seasons) as μ changes and the initial share of large farmers in that district (also averaged over seasons). The left panel shows correlations when μ is lowered from 9% to 7%. The right panel when μ is increased from 9% to 11%. In all panels, mills are allowed to endogenously adjust their sourcing decisions. We define the share of large farmers aw the share of farmers in a district who produce more coffee than the national average in a given season.

Figure 6. Divestitures Counterfactual - Farmers' Surplus & Mills' and Exporters' Profits



Note: The left panel shows the distribution of changes in farmers' surplus across districts (averaged over seasons) when integrated mills and exporters are divested. The middle panel shows the distribution of changes in mills' profits (summed across districts averaged over seasons). We show the results for cooperatives, formerly private integrated mills, and non-integrated mills. The right panel shows the distribution of changes in exporters' profits (averaged over seasons) when integrated mills and exporters are divested. We show the results for formerly private integrated exporters and non-integrated exporters. We report the results both when mills are allowed to adjust their sourcing choices and when they are not. All changes are relative to the initial equilibrium with integrated mills and exporters.

Figure 7. Divestitures Counterfactual - Additional Results



Note: The left panel shows the correlation between changes in market share (measured in percentage points, p.p.) for each mill (averaged across districts and seasons) and the average mill appeal (calculated by averaging the mill-district fixed effects across districts). The right panel shows the correlation between changes in farmers' surplus in each district (averaged over seasons) after simulating divestitures between mills and exporters and the initial share of large farmers in that district (also averaged over seasons). In both panels, mills are allowed to endogenously adjust their sourcing decisions after the policy change. We define the share of large farmers as the share of farmers in a district who produce more coffee than the national average in a given season.

Table 1. Farmers' Supply - Second Stage

VARIABLES	(1) log(S_{mds})	(2) log(S_{mds})	(3) log(S_{mds})	(4) log(S_{mds})	(5) log(S_{mds})	(6) log(S_{mds})
P_{ms}	0.08	0.06		1.26***	1.03***	
	(0.07)	(0.06)		(0.46)	(0.37)	
P_{ms} × Cooperative			0.27**			3.73***
			(0.13)			(1.07)
P_{ms} × Private Not Integrated			0.04			1.17***
			(0.06)			(0.45)
P_{ms} × Private Integrated		0.24***	0.34***		0.87**	3.45***
		(0.09)	(0.09)		(0.41)	(0.95)
Observations	4,835	4,835	4,835	4,835	4,835	4,835
Estimator	OLS	OLS	OLS	IV	IV	IV
Mill-District FE	Yes	Yes	Yes	Yes	Yes	Yes
District-Season FE	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-stat				26.07	16.38	9.337

Note: The unit of observation is a mill-district-season tuple. The first three columns show OLS estimates. The last three show IV estimates. The dependent variable is the log of the market share of mill m in district d in season s. P_{ms} is the price per kilogram that mill m pays to farmers in season s. Cooperative is a binary variable equal to one if mill m belongs to a cooperative. Private Not Integrated is a binary variable equal to one if mill m is private non-integrated. Private Integrated is a binary variable equal to one if mill m is vertically integrated with a downstream exporter. In the last three columns, we instrument prices with the *rendimiento técnico* computed by ICAFE. Cluster standard errors by mill-district and district-season are shown in parentheses. Significance levels: *** 0.01, ** 0.05, * 0.1.

Table 2. Mill-District Fixed Effects

VARIABLES	(1) $\overline{\eta}_m$	(2) $\overline{\eta}_m$
Cooperative	1.40***	0.75
	(0.40)	(0.52)
Private Integrated	1.89***	0.78
	(0.71)	(0.85)
Log Mill Size		0.42**
		(0.16)
Observations	84	84
Estimator	OLS	OLS

Note: The unit of observation is a mill. We focus on the subset of mills that are active in all ten harvest seasons in our sample. The dependent variable is the average fixed effect of mill m across the districts it enters. Cooperative is a binary variable equal to one if mill m belongs to a cooperative. Private Integrated is a binary variable equal to one if mill m is vertically integrated with a downstream exporter. Private non-integrated mills are the omitted category. Log Mill Size is the log of the total quantity sourced by mill m. Heteroscedasticity-robust standard errors are shown in parentheses. Significance levels: *** 0.01, ** 0.05, * 0.1.

Table 3. Bargaining Power

VARIABLES	(1) λ_{mes}	(2) λ_{mes}	(3) λ_{mes}	(4) λ_{mes}	(5) λ_{mes}
Integrated Relationship		-0.07***	-0.02		
		(0.01)	(0.02)		
Log Relationship Size			-0.02***		
			(0.00)		
Integrated Mill				-0.13***	-0.13***
				(0.01)	(0.01)
Integrated Exporter				-0.06***	-0.06***
				(0.01)	(0.01)
Log Exporter Size					0.00
					(0.03)
Log Mill Size					0.00
					(0.36)
Constant	0.61***	0.61***	0.79***	0.63***	0.63
	(0.01)	(0.01)	(0.03)	(0.01)	(0.56)
Observations	1,328	1,328	1,328	1,269	1,269
Estimator	OLS	OLS	OLS	OLS	OLS
Season FE	Yes	Yes	Yes	Yes	Yes
Sample	Full	Full	Full	Arms' Length	Arms' Length

Note: The unit of observation is a mill-exporter-season tuple. The dependent variable is the bargaining power of mill m when trading with exporter e in season s. Integrated Relationship is a binary variable equal to one if mill m and exporter e are vertically integrated. Log Relationship Size is the log of the total quantity exchanged by a mill-exporter pair on the forward market. Integrated Mill is a binary variable equal to one if mill m is vertically integrated with a downstream exporter. Integrated Exporter is a binary variable equal to one if exporter e is vertically integrated with an upstream mill. Log Exporter Size is the log of the total quantity purchased by exporter e in season s on the forward market. Log Mill Size is the log of the total quantity sold by mill m in season s on the forward market. Cluster standard errors by mill-exporter are shown in parentheses. Significance levels: *** 0.01, ** 0.05, * 0.1.

Table 4. Differences in Entry Costs Across Organizational Forms

VARIABLES	(1) Log $\overline{F}_{mds}$	(2) Log $\overline{F}_{mds}$	(3) Log $\underline{F}_{mds}$	(4) Log $\underline{F}_{mds}$
Private Integrated	1.51***	1.88***	1.34***	1.55***
	(0.18)	(0.14)	(0.30)	(0.20)
Observations	3,575	3,430	1,235	882
Estimator	OLS	OLS	OLS	OLS
District-Season FE	No	Yes	No	Yes

Note: The unit of observation is a mill-district-season tuple. In the first two columns, the dependent variable is the log of the upper bound that mill m faces when sourcing coffee from district d in season s. In the last two columns, the dependent variable is the log of the lower bound that mill m faces when sourcing coffee from district d in season s. Private Integrated is a binary variable equal to one if mill m is vertically integrated with a downstream exporter. Since we do not estimate entry costs for cooperatives, private non-integrated mills are the omitted category. Cluster standard errors by mill are shown in parentheses. Significance levels: *** 0.01, ** 0.05, * 0.1.

Online Appendix

"Market Structure, Vertical Integration and Farmers' Welfare in the Costa Rica Coffee Industry"

Fabrizio Leone, Rocco Macchiavello, Pepita Miquel-Florensa, Nicola Pavanini

A Tables

Online Table A.1. Summary Statistics by Organizational Form - Farmer-Mill Transactions

Organizational Form	No. of Farmers	No. of Districts	Price	Costs	Market Share
Cooperative	155.16	5.36	2.73	0.50	0.29
	(140.40)	(5.14)	(0.53)	(0.15)	(0.20)
Private Integrated	119.04	13.19	2.97	0.45	0.33
	(64.99)	(8.84)	(0.39)	(0.07)	(0.18)
Private Not Integrated	17.53	1.63	3.87	0.64	0.07
	(56.99)	(2.21)	(1.51)	(0.25)	(0.13)

Note: The first column reports the number of farmers that mills source from. The second column reports the number of districts that mills enter. The third column reports the price per kilogram in USD that mills pay to farmers. The fourth column reports the processing costs per kilogram in USD that mills sustain. The last column shows mills' market share. We report averages first and standard errors in parentheses below. We distinguish between cooperatives, private integrated mills that are vertically integrated with downstream exporters, and private non-integrated mills.

Online Table A.2. Survey Summary Statistics

Agricultural Inputs	Cash Advances	Loans	Sustainability Program
0.25	0.19	0.19	0.59
(0.34)	(0.29)	(0.24)	(0.50)

Note: The first column reports the average share of farmers to which mills supply agricultural inputs. The second column reports the the average share of farmers to which mills provide cash advances. The third column shows the average share of farmers to which mills offer loans. The last column shows the share of mills that participate in a voluntary sustainability program (Rainforest Alliance, 4C, Starbucks, or Nespresso initiatives). We report averages first and standard errors in parentheses below. All numbers are computed using the responses of the 40 representative mills that answered our survey, and are relative to the 2022-2023 harvest season.

Online Table A.3. Mills' Summary Statistics - Mill-Exporter Transactions

	No. of Exporters	Relationship Age	Average Price	Log Total Quantity
Cooperative	2.27	4.72	4.09	11.23
	(1.52)	(2.02)	(0.59)	(1.12)
Private Integrated	1.11	8.28	4.03	13.78
	(0.12)	(1.42)	(0.45)	(1.08)
Private Not Integrated	1.24	4.13	5.04	9.93
	(0.48)	(2.92)	(1.51)	(1.37)

Note: The first column reports the average number of exporters that mills sell to. The second column reports the average number of consecutive periods that mills trade with exporters. The third columns shows the average price per kilogram in USD at which mills sell coffee to exporters. The last column shows the log of the total quantity that mills sell to exporters in kilograms (averaged across seasons). We report averages first and standard errors in parentheses below. We distinguish between mills that are vertically integrated with downstream exporters and non-integrated mills trading at arms' length (cooperatives or private).

Online Table A.4. Exporters' Summary Statistics - Mill-Exporter Transactions

	No. of Mills	Relationship Age	Average Price	Log Total Quantity
Non-Integarted Exporters	2.38	3.24	4.93	10.71
	(5.58)	(2.81)	(1.44)	(1.58)
Integrated Exporter (Outside Suppliers)	16.97	4.45	4.63	10.64
	(10.13)	(2.11)	(0.58)	(0.99)
Integrated Exporter (Intra-Firm Trade)	1.93	6.46	4.23	13.26
	(1.03)	(3.78)	(0.53)	(1.28)

Note: The first column reports the average number of mills exporters source from. The second column reports the average number of consecutive periods exporters trade with a mill. The third column shows the average price per kilogram paid by the exporter to mills. The last column shows the log of the total quantity (in kilograms) sourced by the exporter. We report averages first and standard errors in parentheses below. We distinguish between exporters that are vertically integrated with upstream mills and non-integrated exporters sourcing at arms' length. When looking at integrated exporters, we further distinguish between their trade with their integrated mills (intra firm) and with independent mills (outside suppliers).

Online Table A.5. Farmers' Supply - First Stage

VARIABLES	(1) P_{ms}	(2) P_{ms}	(3) P_{ms}
z_{ms}	0.03***	0.04***	
	(0.01)	(0.01)	
z_{ms} × Cooperative			0.03***
			(0.01)
z_{ms} × Private Not Integrated			0.04***
			(0.01)
z_{ms} × Private Integrated		-0.07***	-0.04***
		(0.01)	(0.01)
Observations	4,835	4,835	4,835
Estimator	OLS	OLS	OLS
Mill-District FE	Yes	Yes	Yes
District-Season FE	Yes	Yes	Yes

Note: The unit of observation is a mill-district-season tuple. The dependent variable is the price that mill m pays to farmers in season s. z_{ms} is the yield of mill m in season s estimated by ICAFE (*rendimiento técnico*). Cooperative is a binary variable equal to one if mill m belongs to a cooperative. Private Not Integrated is a binary variable equal to one if mill m is private non-integrated. Private Integrated is a binary variable equal to one if mill m is vertically integrated with a downstream exporter. Cluster standard errors by mill-district and district-season are shown in parentheses. Significance levels: *** 0.01, ** 0.05, * 0.1.

Online Table A.6. Farmers' Supply - Second Stage - Heterogeneity

VARIABLES	(1) $\log(S_{mds})$	(2) $\log(S_{mds})$
P_{ms}	1.10***	1.24***
	(0.43)	(0.46)
P_{ms} × Share of Large Farmers$_{ds}$	-0.45*	
	(0.27)	
P_{ms} × Share of Female Farmers$_{ds}$		0.05
		(0.10)
Observations	4,835	4,835
Estimator	IV	IV
Mill-District FE	Yes	Yes
District-Season FE	Yes	Yes
Kleibergen-Paap F-stat	13.32	12.67

Note: The unit of observation is a mill-district-season tuple. The dependent variable is the log of the market share of mill m in district d in season s. P_{ms} is the price per kilogram that mill m pays to farmers in season s. Share of Large Farmers$_{ds}$ is the share of farmers in district d who produce more coffee than the national average in season s. Share of Female Farmers$_{ds}$ is the share of female farmers in district d in season s. Both shares are standardized to have mean zero and unit variance in the sample. We instrument prices with the *rendimiento técnico* computed by ICAFE. Cluster standard errors by mill-district and district-season are shown in parentheses. Significance levels: *** 0.01, ** 0.05, * 0.1.

Online Table A.7. Exporters' Forward Risk Sensitivity

VARIABLES	(1) κ_{es}	(2) κ_{es}
Integrated Exporter	0.09	1.17**
	(0.27)	(0.52)
Log Exporter Size		-0.45**
		(0.21)
Observations	347	347
Estimator	OLS	OLS

Note: The unit of observation is an exporter-season pair. The dependent variable is the sensitivity of exporter e in season s to risk on the forward market. Integrated Exporter is a binary variable equal to one if exporter e is vertically integrated with an upstream mill. Log Exporter Size is the log of the total quantity purchased by exporter e in season s on the forward market. Cluster standard errors by exporter are shown in parentheses. Significance levels: *** 0.01, ** 0.05, * 0.1.

Online Table A.8. Entry Costs in Level

	Number	Mean	Q25	Median	Q75
Upper Bound	3,575	673.25	139.63	331.62	788.41
Lower Bound	1,235	425.80	89.96	181.79	428.86

Note: The table presents summary statistics for the estimated upper and lower bounds of the fixed costs mills face when entering districts in a given season. Entry costs are expressed in USD.

Online Table A.9. Regulation Examples on Sharing Revenues and Banning Vertical Integration

Panel A: Revenue Sharing Regulation

Country	Sector	Source
Costa Rica	Coffee	ICAFE
Burundi	Coffee	Coffee Code
Kenya	Tea	The Tea Act
Rwanda	Tea	NAEB

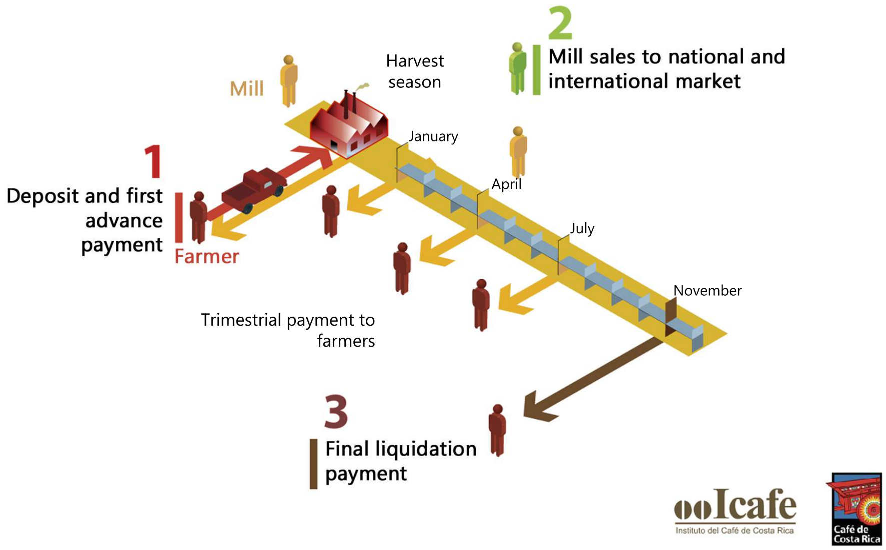
Panel B: Ban of Vertical Integration

Country	Sector	Source
Kenya	Coffee	The Coffee Act
Ethiopia	Commodities	ECX
Ghana	Cocoa	Cocoa Board Law
Tanzania	Cotton	The Cotton Industry Act
Thailand	Rice	Rice Trading Act

Note: All links accessed on February, 25th 2025. Click on ***Source*** to explore regulation documents. ***Panel A: Revenue sharing.*** **Costa Rica, Coffee** ensures that mills retain 9% of revenues' profits; the rest is transferred to farmers. **Burundi, Coffee** establish that 10%-15% of revenues are taken by mills, farmers take the major part of the sales revenue. **Kenya, Tea** sets that farmers take 60-70% of sales revenue. (link here). **Rwanda, Tea** establish a trade mechanism that ensures that 50% of total sales must be paid to farmers (link here). ***Panel B: Banning vertical integration.*** **Kenya, Coffee** sets a state-owned board that grants licenses to supply chain actors that ensures control on vertical integration practices **Ethiopia, Commodities**. prevent integration through usage of limited licenses per company and public auction system. (link here) **Ghana, Cocoa** sets a owned-state board that participates in the supply chain to prevent integration practices, grants licenses and public auctions (link here). **Tanzania, Cotton** establishes a government board that oversees the industry using licenses for supply chain links, price regulations, and foreign investment (link here). **Thailand, Rice** through the Rice Department and Trade Competition Act prevents integration of supply actors through licenses, public international auctions and separation of supply chain roles (link here). Commodities include coffee, cereals, oils, etc.

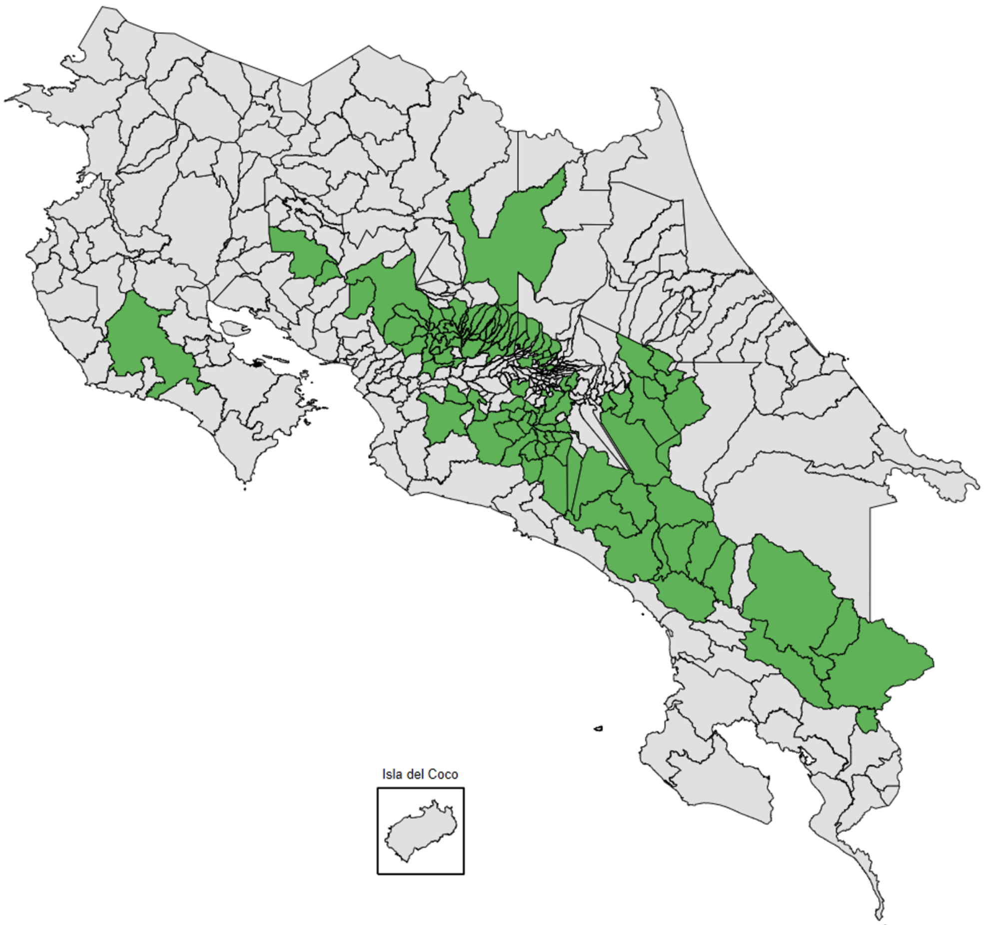
B Figures

Online Figure B.1. Structure of the Liquidation Procedure



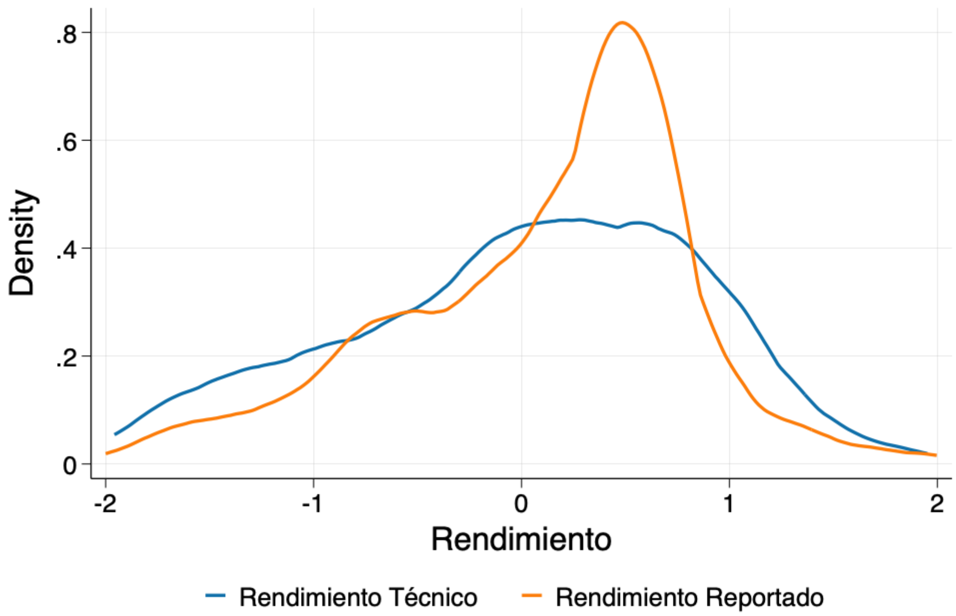
Note: The figure illustrates the structure of the coffee liquidation procedure. Farmers deliver their cherries to a mill within twenty-four hours of harvest to preserve quality. Mills process and sell coffee to national and international exporters. At the end of the season, the final price that mills must pay to farmers is established.

Online Figure B.2. Merged Districts



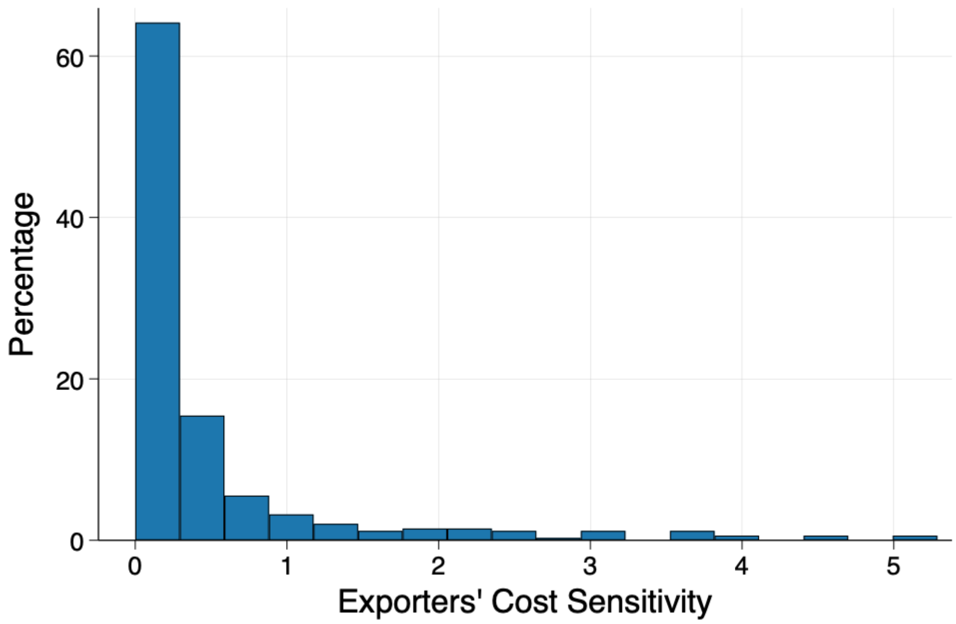
Note: The figure shows the 105 (merged) districts that we use to define the local markets within which mills compete to source coffee from farmers in each harvest season. Green districts produce coffee every season in our data, and account for 98% of total production. These districts are created by merging smaller, adjacent geographical units within the same province. Gray districts are those in which no coffee production is ever registered or where small amounts are registered in some seasons only. We exclude gray districts from our sample.

Online Figure B.3. *Rendimiento Técnico* versus *Rendimiento Reportado*



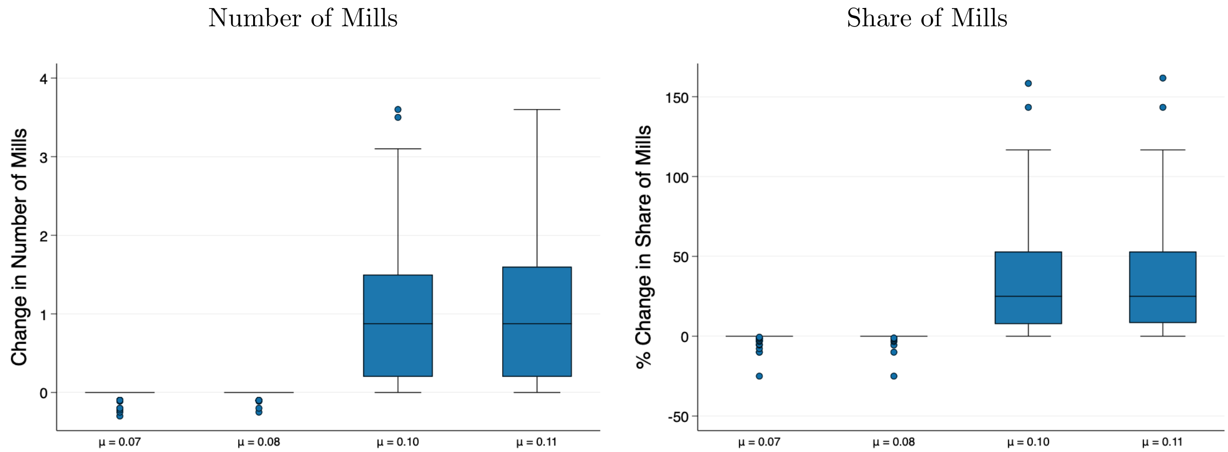
Note: The figure shows the density of the *rendimiento técnico* (z^T_{ms}) computed by ICAFE and *rendimiento reportado* (z^R_{ms}), which is obtained and reported by each mill. The yield factor used to determine the price paid to farmers is defined as $z_{ms} = \max\{z^R_{ms}, z^T_{ms}\}$. The yield factor captures the kilos of parchment coffee (output) produced from one kilo of coffee cherries (input).

Online Figure B.4. Empirical Distribution of Exporters' Forward Risk Sensitivity



Note: The figure shows the empirical distribution of the exporters' sensitivity to risk when trading on the forward market. We estimate this distribution from equation (15).

Online Figure B.5. Rent Sharing Counterfactual - Mills' Entry and Exit



Note: The left panel shows the distribution of changes in the number of mills active across districts (average over seasons) when μ is decreased or increased. The right panel expresses these changes as a share relative to the initial number of mills in a district. All changes are relative to the initial equilibrium with $\mu = 9\%$.

C Computational Appendix

This section provides computational details of our model.

C.1 Fixed Cost Estimation

Let M be the number of mills, D the number of districts, and S the number of seasons. We estimate entry costs in equation (16) and (17) as follows:

1. We estimate equation (13) and store the estimated ξ_{mds} and η_{md};

2. We draw a $((M \times D \times S) \times 1000)$ matrix of mill-district-season-specific demand shifters from the empirical distribution of ξ_{mds} and a $((M \times D) \times 1000)$ matrix of mill-district fixed effects from the empirical distribution of η_{md};

3. For each draw, we compute mill-season profits under the observed allocation of mills across districts. When computing market shares, we discard the outside option as it does not appear in equation (13);

4. For each draw, we compute mill-season profits by sequentially removing a district $d \in \mathcal{A}^1_{ms}$ and estimate the upper bounds as an average across all draws, as shown in equation (16). We also compute mill-season profits by sequentially adding a district $d \in \mathcal{A}^0_{ms}$ and estimate the lower bounds as an average across all draws, as shown in equation (17).

We impute missing lower bounds for districts $d \in \mathcal{A}^1_{ms}$ and missing upper bounds for districts $d \in \mathcal{A}^0_{ms}$ using the support of the estimated upper and lower bounds. Finally, we estimate the mill-specific component of fixed costs in equation (10) as an average over upper and lower bounds for that mill, respectively.

C.2 Implementation of the Counterfactual Exercises

Let M be the number of mills, D the number of districts, and S the number of seasons. We compute counterfactual changes in farmers' surplus and mills' and exporters profits as follows:

1. We use the same $((M \times D \times S) \times 1000)$ matrix of mill-district-season demand shifters and $((M \times D) \times 1000)$ matrix of mill-district fixed effects of Section C.1;

2. For each mill, we compute an average upper and lower bound to entry costs as:

$$\overline{F}_m = \sum_{d \in \mathcal{D}_{ms}, s \in \mathcal{S}_m} \frac{\overline{F}_{mds}}{|\mathcal{D}_{ms}| \times |\mathcal{S}_m|} \quad \text{and} \quad \underline{F}_m = \sum_{d \in \mathcal{D}_{ms}, s \in \mathcal{S}_m} \frac{\underline{F}_{mds}}{|\mathcal{D}_{ms}| \times |\mathcal{S}_m|}.$$

   $\mathcal{D}_{ms}$ is the set of districts entered by mill m in season s. $\mathcal{S}_m$ is the set of seasons in which mill m is active;

3. We draw a $(M \times 1000)$ matrix of mill-specific entry costs with support $[\underline{F}_m, \overline{F}_m]$;

4. For each draw of mill-district-season demand shifters, mill-district fixed effects, and mill-specific entry costs, we estimate farmers' surplus using equation (2), mills' profits using equation (3), and exporters' profits using equation (5) under the initial equilibrium and counterfactual scenario being analyzed. We then compute average outcome changes across draws and between scenarios.