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

January 2025

Abstract

This paper develops a structural model of agricultural chains to assess how common market frictions and organizational structures shape welfare along the chain. We estimate the model using administrative data that cover all transactions between farmers, intermediaries (mills), and downstream exporters in the Costa Rica coffee industry over a 10-year period. Through counterfactuals, we show that policies aimed at preserving competition among mills without distorting allocative efficiency are beneficial to farmers. In contrast, banning integrated organizations (that is, mills owned by downstream exporters) makes farmers worse off by preventing them from accessing valuable services provided by these organizations.

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, 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.

1 Introduction

Approximately 75% of the world's poor live in rural areas and earn their livelihood in agriculture (World Bank, 2007, 2017). Due to market frictions, including imperfect access to finance, market power, and weak contract enforcement, regulators often support farmers by increasing their payments or limiting the market power of intermediaries (mills) they sell to. However, evaluating the impact of these regulatory interventions is challenging for two main reasons. First, data are often scarce due to the informality of many agricultural transactions. Second, these policies have ambiguous effects on market structure and welfare along the chain, as they may drive some intermediaries out of the market, reduce competition among them, and ultimately harm farmers.

In this paper, we develop a structural model of agricultural chains to evaluate these trade-offs. Through counterfactuals, we quantify how common policy interventions affect market structure and welfare along the chain. We overcome the challenge of data limitation applying the model to the Costa Rica coffee industry, an ideal setting in terms of data availability and reliability. Moreover, the industry features frictions and regulations commonly observed in agricultural markets, which makes our framework portable to other contexts as well.

We provide three contributions. First, we collect data on the three main actors in the Costa Rica coffee chain: farmers, who grow and harvest coffee; mills, which process it; and exporters, who sell it to international markets. We observe the universe of transaction-level data along the chain, as well as the characteristics of the parties involved, for ten harvest seasons. Second, we develop a structural model that captures key features of agricultural chains, such as interlinked transactions between mills and farmers, mills' imperfect competition and choice of local sourcing districts, and bilateral bargaining between mills and exporters. Finally, we use the model to assess the effects of two policy interventions. We find that farmers benefit from the current rent-sharing rule that mandates mills to transfer to them 91% of their profit margin, and that banning vertical integration between mills and exporters would instead harm farmers, as they would lose valuable services that integrated entities are able to provide.

We begin our analysis by providing background on the Costa Rica coffee chain. Coffee cherries are harvested by smallholder farmers and delivered to mills within a few hours of harvest. Then, mills process the coffee and sell it to downstream exporters, who consolidate the coffee before selling it to international markets. The Costa Rica coffee chain is regulated by the local coffee board (*Instituto del Café de Costa Rica*, or ICAFE). A key aspect of the regulation is that, at the end of each harvest season, farmers are paid a (minimum) final price which is proportional to mills' profit margin.

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 farmers' characteristics, mills' marginal costs of processing coffee, and the price that exporters obtain in international markets. Information about organizational forms is also available. We know if mills are owned by farmers (cooperatives), by downstream exporters (vertical integration), or operate as independent private entities. To complement these data, we conducted a survey among a sub-sample of representative coffee mills in Costa Rica to gather evidence about services that they offer to farmers, such as advance payments, loans and training programs.

Using these data, we document two facts. First, in any given season, farmers sell their coffee to one mill only within the district where they operate. Second, different types of organizations behave differently along the chain. Vertically integrated mills tend to source larger amounts of coffee and are active in more districts than cooperatives and non-integrated mills. Integrated mills and exporters also trade more coffee than others using forward contracts, which provide demand and supply assurance.[1]

These facts inform a structural model of agricultural value chains. The model consists of three building blocks: farmers' choice regarding which mills to sell their coffee, bilateral bargaining between mills and exporters over prices and quantities of coffee to be traded, and mills' sourcing strategies across local coffee markets.

We model farmers' supply decisions as a discrete choice among a set of differentiated mills. To do so, we "invert" 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. Even conditional on prices, farmers may have preferences for certain mills, such as those that provide credit, inputs, and other services.

We model trade between mills and exporters as a bilateral-oligopoly bargaining process.[2] To pin down bilateral prices and quantities, we rely on the "Nash-in-Nash"

[1]Mills and exporters can trade coffee through spot or forward contracts. "Forward trade" refers to contracts with at least a 30-day interval between signing and delivery. While these contracts provide demand and supply assurance, they also carry default risk and exposure to price fluctuations. Large, often vertically integrated mills and exporters are the primary users of forward contracts.

[2]This assumption is supported by the fact that the average mill in the data sells to 2 exporters,

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 their outside options if bargaining collapses. We rationalize forward trade as a bilateral-oligopoly bargaining game. If mills and exporters fail to agree, we assume that mills do not buy that quantity from farmers, while exporters source on the spot market at a competitive price.

Following Eizenberg (2014), we model mills' sourcing decisions as a two-stage game. 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 characterize a set of sufficient conditions for the existence of the equilibrium of this game.

We estimate the model and recover measures of farmers', mills', and exporters' welfare. Estimating farmers' supply requires addressing price endogeneity: even conditional on mills' characteristics, the price that mills offer to farmers may be correlated with unobserved shocks that explain why some mills receive more coffee than others. We address this issue using ICAFE's minimum price regulation. Each season, ICAFE sets a mill-specific minimum price that mills must pay to farmers based on weather conditions, geography, and agronomic factors. This minimum price is unrelated to market competition, and serves as an instrument for the actual prices mills pay.

Unlike much of the existing literature, we observe mills' and exporters' margins, not just who trades with whom. This allows us to efficiently estimate the structural parameters of the bargaining game, as margins provide direct information about the relative bargaining power of mills and exporters. We use a revealed preference approach to estimate bounds on the fixed costs mills incur to reach each district, based on their observed sourcing decisions.

The model estimates suggest that the Costa Rica coffee chain is dominated by vertically integrated mills and exporters. On the sourcing side, mills owned by downstream exporters are particularly appealing to farmers, as revealed by their large market shares conditional on prices. These integrated mills provide more services to farmers and can pay higher fixed entry costs to source from each district. However, integrated mills pay on average lower prices to farmers than private non-integrated mills. On the down-

while the average exporters sources from 5 mills.

stream sales side, relative bargaining power is tilted in favor of integrated mills that trade most of the coffee produced in a season.

Based on the regulatory experiences in Costa Rica and in other countries, we use the estimated model to evaluate the impact of two policy interventions.

First, we evaluate the ICAFE regulation, which requires mills to transfer 91% of their profit margin to farmers as payment. We investigate how changing this share impacts farmers' surplus and mills' profits. On the one hand, increasing the share directly reduces payments to farmers. On the other hand, a higher share increases mills' profits and potentially allows them enter more districts, indirectly increasing total resources available to share with farmers. Our results show that a tight share at 9% preserves competition and allocative efficiency. This suggests that the regulator assigns a larger weight to farmers over mills in its objective function.[3]

Second, we consider the impact of policies that ban vertical integration between mills and exporters. Many regulators around the world are concerned that vertical integration in agricultural chains depresses prices paid to farmers. However, vertical integration also facilitates mills' access to finance and a more secure marketing channel downstream, which benefits farmers. Our exercise shows that banning vertical integration between mills and exporters increases competition at the farm gate but, while beneficial for non-integrated mills and exporters, ultimately reduces farmers' surplus.

Related Literature. A large literature has focused on market failures that directly affect farming activities, e.g., in input (Duflo et al., 2008, 2011; Ashraf et al., 2009; Bold et al., 2017) and credit and insurance (Karlan et al., 2014; Cai et al., 2015) markets. The experimental literature has investigated potential failures in output markets (Maitra et al., 2017; Bergquist and Dinerstein, 2020; Casaburi and Reed, 2022). Imperfections in input markets interact with downstream market structure in recent experimental work on interlinked transactions (Casaburi and Willis, 2018; Casaburi and Macchiavello, 2019). Non-experimental papers also explore market imperfections along agricultural chains (Blouin and Macchiavello, 2019; Macchiavello and Morjaria, 2015, 2020; Macchiavello and Miquel-Florensa, 2018). Emran et al. (2021) study pass-through of import prices to domestic chains in a context with credit constraints. Porto et al. (2011) and Dillon and Dambro (2017) review evidence on the degree of competition in sub-Saharan agricultural markets.

[3]This result aligns with the view that Costa Rica's egalitarian society facilitated the adoption of such regulations. Similar efforts in more unequal Central America countries have failed (Paige, 1997).

This body of evidence allows to understand market failures in agricultural chains. However, because market failures are identified one at a time and by focusing on individual farmers' or buyers' decisions, the evidence is not sufficient to evaluate (and design) policies that tackle the interaction between multiple market failures along the chain. This paper contributes by developing and estimating a structural model of an agricultural chain that allows for market imperfections at multiple stages of the chain.

Moreover, this paper adapts tools developed in the industrial organization literature to a new context, agricultural value chains.[4] We estimate an "inverted" demand framework for differentiated products (Berry, 1994; Berry et al., 1995; Nevo, 2000) to model farmers' supply. A similar approach is used by Zavala (2022) and Rubens (2023) to model input suppliers' choice of manufacturers.[5] The key contribution of our paper is that we also model the downstream bargaining between mills and coffee exporters, as well as mills' entry decision in local districts. 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. Last, we contribute to the recent literature on bilateral bargaining (Crawford and Yurukoglu, 2012; Crawford et al., 2018; Collard-Wexler et al., 2019; Alviarez et al., 2023) by modeling price negotiations between upstream mills and downstream exporters, both within and outside vertical relationships. While the existing literature tends to focus on aspects of vertical integration that are pro-competitive (no double marginalization) and anti-competitive (foreclosure, rising rivals' costs), we shed light on novel mechanisms that are more common in agricultural chains. We quantify differences across mills' organizational structures (private integrated or not, and cooperatives) in processing costs, fixed entry costs, bargaining power parameters, and farmers' preferences.

The paper unfolds as follows. Section 2 provides background on the Costa Rica coffee industry. Section 3 introduces our data. Section 4 shows the empirical facts that motivate our model in Section 5. Section 6 discusses identification and estimation. Section 7 introduces the counterfactual exercises. Section 8 concludes.

[4] Recent papers bridging these two literatures are Dominguez-Iino (2021) and Laoprapassorn (2022). See Macchiavello et al. (2022) for a survey of the literature using similar tools to study food markets.

[5] An alternative approach to measuring monopsonony power in agricultural markets involves estimating the processors' production function (Morlacco, 2019; Avignon and Guigue, 2022). See De Loecker and Scott (2022) for a discussion about differences and complementarities between these two approaches in the context of markup estimation.

2 Industry Background

This section provides background information on the Costa Rica coffee industry.

2.1 The Coffee Chain

Figure 1 summarizes the three main stages of the Costa Rican coffee chain, known as the *Sistema de Liquidación Final*. In the first stage, farmers produce coffee cherries and sell them to mills. Mills may or may not provide valuable services to farmers, such as advance payments, loans and training programs. Mills process the cherries into parchment coffee.[6] In the second stage, mills sell the processed 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. In the last stage, mills pay to farmers a liquidation price.

Figure 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.

Three features of the coffee chain are worth noting. First, given the wet processing, coffee farmers must deposit the harvested cherries to a mill within twenty-four hours of

[6] Coffee beans are obtained by removing the pulp from the cherries within hours of harvest. After being washed and dried, beans become storable.

harvest to preserve quality. The volume and fragility of the cherries, together with the mountainous terrain conditions that make transport challenging, restrict the choice of the mills where to deposit the harvest. Farmers deposit their harvest either at mills located in their district or at mobile buying points from mills located in neighboring districts. Otherwise, farmers (mills) are free to sell to (buy from) whoever they want.

A second feature of the chain is the existence of a variety of organizational forms. This paper focuses on the distinction 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 local institutions collectively owned by local farmers. Membership in a cooperative involves an implicit albeit non-binding commitment to sell cherries exclusively to that cooperative. Cooperatives do not purchase cherries from non-members or neighboring districts. While prices are an important factor in farmers' decisions, non-price factors—such as the ability of different types of mills to offer advance payments, loans, and training programs—also play a role.

A final feature of the chain is that mills and exporters trade coffee either on the spot market or through forward contracts. Spot market transactions occur at prevailing international prices. Forward contracts, signed before the harvest season begins, offer supply and demand assurance but carry risks, including potential defaults and fluctuations in international exchange rates and prices.

2.2 The Coffee Board

The coffee sector in Costa Rica is regulated through the *Instituto del Café de Costa Rica* (ICAFE), a non-government institution that represents the interests of farmers, mills, and exporters. All contracts between farmers and mills, as well as between mills and exporters, are registered with ICAFE, which is also responsible for enforcing contracts along the chain.[7]

Two features of the regulation are worth noting. First, the liquidation price that mills must pay to farmers at the end of the harvest season is directly tied to the mill's margin. By law, mills retain 9% of their profit margin, calculated as the difference

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

between the price received from downstream exporters and their processing costs, both of which must be approved by the regulator.[8] 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.

A second feature of the regulation is the existence of a minimum price that mills must pay to farmers. This price is determined as follows. Based on coffee production from previous harvests, ICAFE identifies a set of representative districts. During the current 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 calculate the minimum price. Mills with higher estimated yield are required to pay a higher minimum price. At the same time, mills calculate their own yield (*rendimiento informado*) and submit a price proposal to ICAFE. The final price paid by each mill is the higher of the self-reported price or the minimum price set by the regulator. This feature of the regulation allows us to identify farmers' supply decisions.

3 Data

This section describes our data and provides summary statistics.

3.1 Data Sources

From ICAFE, we obtain transaction-level data among all farmers and mills and among all mills and exporters active in a given season. We use three different data sources.

Farmer-Mill Transactions. We observe the quantity of coffee that each active farmer sells to each active mill in a given district and season. We use this information to construct a measure of farmer size. Additionally, we have information on the gender of farmers. Concerning mills, we observe the number of districts they source from and the

[8]ICAFE's mandate to approve all prices between mills and exporters enables a meaningful interpretation of prices within vertically integrated organizations, which cannot fall below a legal minimum.

quantity of coffee purchased in each of them in each season. We also observe the final liquidation price that each mill pays to farmers and mills' (marginal) cost of processing coffee. We convert all the nominal quantities (prices and costs) in USD per kilogram using mill-season specific exchange rates provided by ICAFE. We 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 in our sample. Information about farmer-mill transactions is available for ten harvest seasons, from 2010/2011 to 2019/2020.

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 days 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. Information about mill-exporter transactions is available for ten harvest seasons, from 2010/2011 to 2019/2020.

Mill Survey. In 2024, we conducted a joint survey with ICAFE to collect information about the services that mills provided to farmers during the 2022/2023 harvest season. We selected a representative sample of 70 mills, stratified by geographical area and organizational form. Among them, 40 responded to the survey. All responses were recorded anonymously, as per ICAFE regulation. The survey covered topics such as mill size (coffee sourced and number of employees and farmers), the provision of advance payments, agricultural inputs, and loans to farmers, and participation in sustainability programs (Rainforest Alliance, 4C, Starbucks, or Nespresso initiatives).

3.2 Summary Statistics

Farmers and Mills. Our sample includes 105 districts where mills compete to source coffee from farmers.[9] There are 365 mills in total. Among them, 9 are vertically inte-

[9]There are 174 administrative districts in Costa Rica where coffee is grown and harvested every season. 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 Figure B.1.

grated, and collectively source between 35% and 55% of all coffee produced, depending on the season. In the average district-season, about 460 farmers (28% of which are female) and 6 mills are active. The average mill sources coffee from 3 districts.

Table A.1 reports summary statistics by organizational form. Private integrated mills and cooperatives source coffee from more farmers and have higher market shares than private non-integrated mills. Private integrated mills also source from more districts and have lower processing costs than the others. However, they pay on average lower prices to farmers than private non-integrated mills. Table A.2 shows the summary statistics of the survey. The average mill pays about 20% of its farmers in advance. It also provides agricultural inputs to about 25% of its farmers and offers loans to about 19% them. Around 60% of mills participate in at least one voluntary sustainability program.[10]

Mills and Exporters. In an average season, 165 mills and 71 exporters are active. On average, each mill sells to 2 exporters, while each exporter sources from 5 mills. Among the exporters, 7 are vertically integrated with their mills.[11] Table A.3 presents summary statistics from the perspective of mills, while Table A.4 provides them from the exporters' perspective. Both tables distinguish between integrated mills and exporters and those trading at arm's length (e.g., cooperatives or private non-integrated entities). Integrated mills and exporters tend to maintain longer relationships, trade larger volumes of coffee, and offer lower prices compared to their competitors.

4 Stylized Facts

This section introduces two facts about the Costa Rica coffee industry. These facts motivate two of our modeling assumptions in the next section.

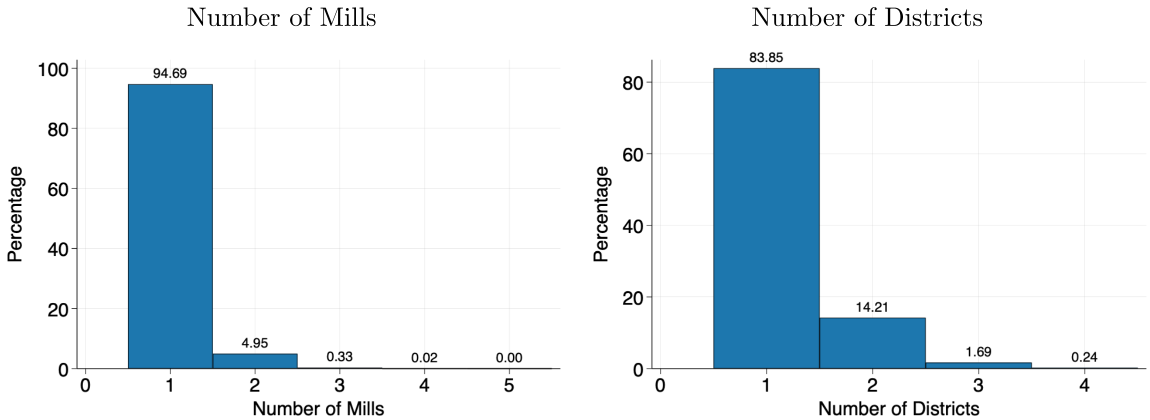
Fact 1. *In any given district-season, each farmer sells to one mill only.*

The left panel of Figure 2 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. The right panel reports the share of farmers that sell in a given number of districts. About 85% of farmers sell in a single districts.

[10] We do not include programs such as Fair Trade, which are exclusively available to cooperatives or farmer associations.

[11] The discrepancy between the number of integrated mills and exporters arises because some mills operate under different names and firm identifiers, despite being part of the same group.

Figure 2. Farmers' Supply



Note: The left panel shows the share of farmers that sell to a given number of mills within a district-season. The right panel shows the share of farmers that sell in a given number of districts.

Two main reasons help explaining Figure 2. 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 product. Second, it is expensive for mills to collect cherries from multiple districts, which is why an average mill typically operates in only 3 districts. As a result, farmers have a limited number of mills to choose from. From a modeling perspective, this suggests that farmers' supply can be represented as a discrete choice among a finite set of differentiated mills sourcing in their district.

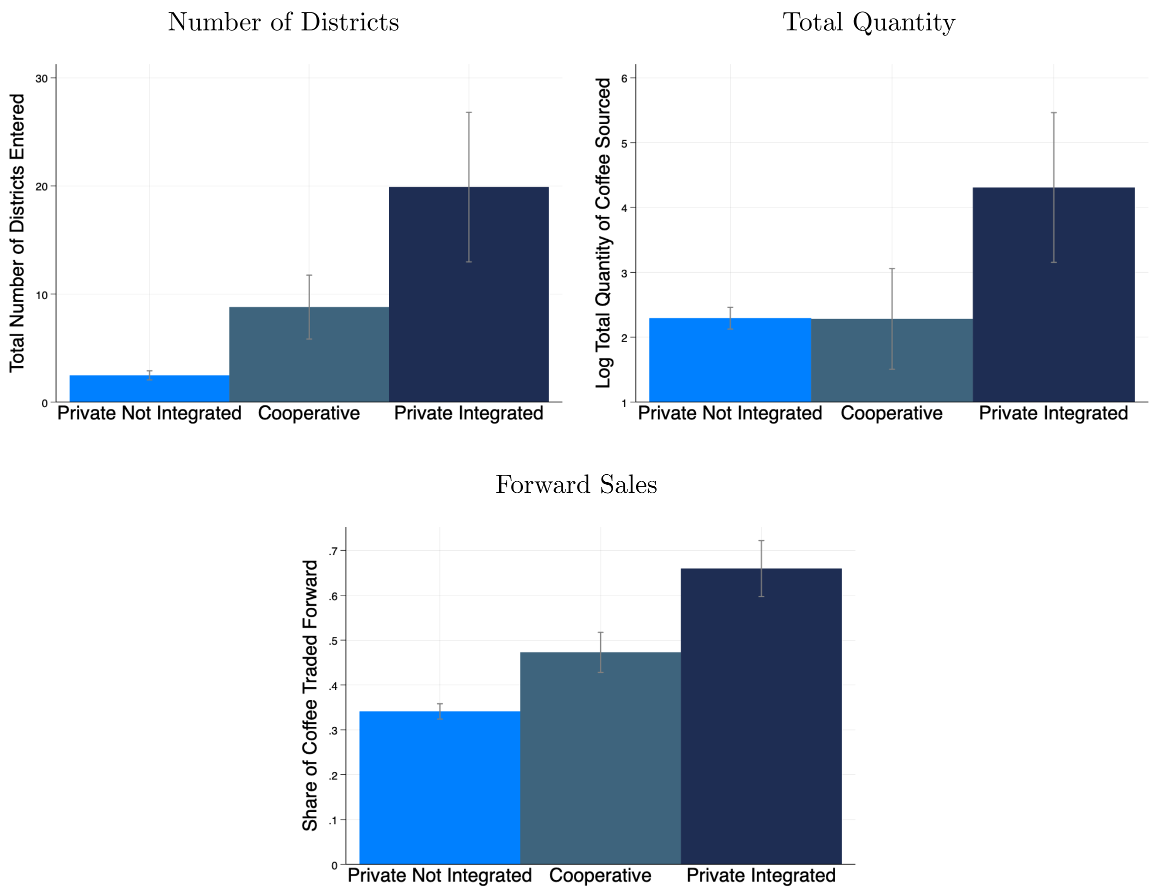
Fact 2. *Mills with different organizational forms behave differently along the chain.*

The top left panel of Figure 3 shows the number of districts that mills belonging to different organizations (vertically private integrated, private not integrated, cooperatives) source coffee from. The top right panel shows the log of the total quantity of coffee they source. The average private integrated mill sources coffee from 20 districts, whereas private non-integrated and cooperatives enter less than 10 districts across all seasons. Additionally, private integrated mills source more coffee than the other two types of organization.

The bottom panel shows the share of total mills' coffee sales to exporters using forward contracts. Private integrated mills sustain a higher share of trade through forward contracts (about 65%) than the others. This fact is explained by the ability of integrated organizations to prevent default on forward contracts.[12]

[12] Macchiavello and Miquel-Florensa (2018) show that forward contracts ensure supply and demand

Figure 3. Mills' Sourcing Activities and Forward Sales



Note: The top left panel shows the total number of districts that mills belonging to different organizations enter. The top right panel shows the log of the total quantity (in kilograms) of coffee sourced by mills belonging to different organizations. The bottom panel shows the share of coffee that mills belonging to different organization trade using forward contracts.

Overall, there is evidence that private integrated mills are more mobile across districts, are able to source more coffee, and can sustain more forward trade than private non-integrated mills and cooperatives. These features make them the largest mills in the Costa Rica coffee industry. To account for this heterogeneity, we allow the structural parameters of the model to vary across different mills' organizational forms.

---

assurance, reducing inventory risks. In Costa Rica, mills can default on forward contracts when market conditions change. However, this very rarely happens when mills are integrated with exporters, favoring stronger reliance on forward sales that are enforced by the organizational structure.

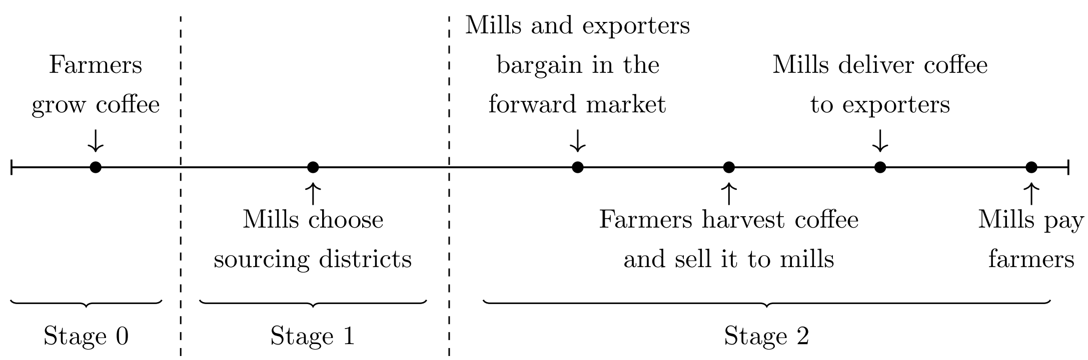
5 Model

This section introduces an empirical model of the coffee value chain. The structural framework reflects the two stylized facts described above, as well as the institutional and regulatory features of the Costa Rica coffee chain, but it could easily be generalized to other settings. It consists of three parts. First, we model farmers' supply as a discrete choice over mills that are sourcing in their district. Second, we model the bargaining between mills and exporters. Last, we model mills' entry decisions across districts.

5.1 Timeline

Figure 4 illustrates the timeline of the events within each harvest season in the model.

Figure 4. Timeline of the Events



Note: The figure shows the timeline of the events in the model.

In stage 0, farmers grow coffee in their districts. While the amount of coffee farmed is assumed to be exogenous, farmers have rational expectations about the quantity of coffee demanded by mills.[13] In stage 1, mills compare the cost of entering each district against expected profits and decide which districts to source coffee from.

The following events take place in stage 2. Before receiving coffee from farmers, mills and exporters negotiate a forward price for a future quantity of coffee to be exchanged

[13]The assumption that coffee grown is exogenous is motivated by the fact that coffee requires approximately three years to produce after planting. Therefore, in each season, the actual coffee harvest in a given season is the result of past planting decisions and current weather conditions. For simplicity, we do not model farmers' investment choices. The rational expectation assumption ensures that the total quantity of coffee sold by farmers to mills equals the total amount of coffee that mills sell to exporters on the forward and spot markets.

after the harvest. Once harvest takes place, mills receive coffee from farmers and fulfill their contractual obligations with downstream exporters. Any remaining coffee unsold on the forward market by mills (or unpurchased by exporters) is traded on the spot market at an exogenous price determined by international coffee markets. Both parties make zero profits on the spot market. Finally, mills pay farmers in compliance with the regulation. Because coffee is perishable, no mills brings coffee from one season the next, and we solve the model separately for each season.

5.2 Farmers' Choice of Mills

There are $f = 1, \dots, F$ farmers located in $d = 1, \dots, D$ districts, producing coffee each season $s = 1, \dots, S$. Each season, each mill $m = 1, \dots, M$ decides whether to enter each local district. Farmers sell coffee to one of the mills $\mathcal{M}_{ds}$ that operates in their district during that season. Each farmer sells to the mill that maximizes her indirect utility, which reads:

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

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 and captures interlinked transactions. It explains why some mills systematically receive more coffee than others, even conditional on prices. For example, mills offering more services (e.g., cash advances, loans, and inputs) may have a higher η_{md}. 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 that mill m has in district d in season s and farmers' surplus in district d in season s as follows:

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

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 also positively depends on the number of mills active in a district.

5.3 Mills and Exporters

Mills' Payoff. 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 are set by international coffee markets. The profits that mill m makes when trading on the forward market in season s are:

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

p_{mes} denotes the forward price that mill m charges to exporter e in season s, and q_{mes} is the corresponding quantity traded. c_{ms} is the 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. $\mathcal{E}_{ms}$ is exogenously given and is held fixed in the counterfactual scenarios.[14] $\mathcal{D}_{ms}$ is endogenously determined by mills' sourcing decisions.

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:

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

Upstream prices are a weighted average of the prices that each mill negotiates with its downstream exporters in the forward market and the prices it receives in the spot market (p_s^{spot}) for any remaining unsold quantities. The weights are given by the share of coffee that mills and exporters trade in the forward market (s_{mes}) and in the spot market ($1 - \sum_{e \in E_{ms}} s_{mes}$).

[14] This assumption is common in the literature (e.g., Crawford and Yurukoglu, 2012; Crawford et al., 2018), and is consistent with focusing on short-run counterfactuals in which firms neither establish new relationships nor drop existing ones. Our data support this assumption: the median (average) mill in our sample drops 0.1 (0.3) exporters over time.

Exporters' Payoff. The profits of exporter e when trading on the forward market in season s are:

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

The definitions of p_{mes} and q_{mes} are provided above. p^{w}_{es} is the exogenous price that exporter e receives in international markets during season s. The term $\frac{\kappa_{es}}{2} q^{2}_{mes}$ reflects the risks faced by exporters when trading on the forward market (e.g., default risk and price fluctuations).[15] κ_{es} captures the sensitivity of different exporters to such shocks. We denote by $\mathcal{M}_{ms}$ the set of mills from which exporter e purchases coffee in season s. Similar to $\mathcal{E}_{ms}$, this set is considered exogenous and held fixed in counterfactuals.[16]

From equation (5), the quantity of coffee that exporter e demands from mill m in season s in the interior solution case is:

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

Nash Bargaining. Every season, mills and exporters bargain over forward prices p_{mes} to solve the following problem:

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

$\lambda_{mes} \in [0, 1]$ 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 marginal quantity from farmers, and exporters source it on the spot market at a competitive price. Therefore, $\pi^{D}_{ms} = \pi^{D}_{es} = 0$, and parties always have an incentive to participate in the bargaining process.

We assume that each transaction takes place separately and parties have passive beliefs about all other transactions (including those to which they are party). Thus, if

[15] Assuming convex costs and exogenous downstream prices is a tractable way to generate a downward-sloping demand curve for forward quantities. Leong et al. (2022) use a similar formulation in a different context.

[16] We impose this assumption for the same reason as we do for mills. Our data support it: the median (average) exporter in our sample drops 0 (0.3) mills over time.

they fail to reach an agreement, mills and exporters do not renegotiate the contracts they have already signed.[17] In the interior solution case, bilateral prices are given by:

$$p_{mes} = \frac{\lambda_{mes} p^w_{es} + (2 - \lambda_{mes}) c_{ms}}{2}. \tag{8}$$

Notice that as λ_{mes} approaches zero, exporters have all the bargaining power, and mills earn zero profits. The opposite argument holds when λ_{mes} approaches one.

5.4 Mills' District Choice

Each season, mills decide which districts to source from, depending on whether it is profitable to do so. Following Eizenberg (2014), let $\mathcal{A}_{ms}$ 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:

$$F_{mds} = F_m(\theta) + \nu_{mds}, \qquad \mathbb{E}[\nu_{mds} | d \in \mathcal{A}_{ms}] = 0. \tag{9}$$

We let $\theta \equiv \{\mu, \alpha, \eta_{ds}, \kappa_{es}, \lambda_{mes}, p^w_{es}\}$. Equation (9) 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.

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.[18] A necessary condition for the existence of an SPNE is:

$$E_{u|\theta}\left[\pi_{ms}(\mathcal{A}_{ms}; u, \theta) - \pi_{ms}(\mathcal{A}_{ms} - 1^d_{ms}; u, \theta)\right] \geq F_{mds} \qquad \forall d \in \mathcal{A}^1_{ms}, \tag{10}$$

$$E_{u|\theta}\left[\pi_{ms}(\mathcal{A}_{ms} + 1^d_{ms}; u, \theta) - \pi_{ms}(\mathcal{A}_{ms}; u, \theta)\right] \leq F_{mds} \qquad \forall d \in \mathcal{A}^0_{ms}. \tag{11}$$

1^d_{ms} is a vector with entries of zeros apart from a single entry of 1 for district d. In

[17] 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).

[18] In Section 7, we will assume sequential entry to deal with multiple equilibria.

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 $u_{mds} = \{\xi_{mds}, \eta_{md}, \nu_{mds}\}$, conditional on a estimated set of second-stage parameters θ. We assume that u_{mds} is realized between stage 1 and stage 2 in Figure 4.

6 Identification and Estimation Results

This section describes our identification strategy and estimation results.

6.1 Identification

Farmers' Supply. In each district-season, we aggregate small mills into one alternative, the outside option, and normalize its utility to zero.[19] We do so because most farmers always have the option to sell their production to small, often family-owned, mills. We then apply Berry (1994)'s inversion to market shares in equation (2) to obtain the following estimating equation:[20]

$$\log\left(S_{mds}\right) = \alpha_0 + \alpha_1 P_{ms} + \eta_{md} + \zeta_{ds} + \xi_{mds}. \tag{12}$$

The inclusion of district-season and mill-district fixed effects implies that the parameters in equation (12) are identified from variations in prices and market shares around these means.

Estimating equation (12) requires addressing the issue of price endogeneity. Mills with high realizations of the mill-district-season demand shifter (ξ_{mds}) are able to source large quantities of coffee from farmers even when offering lower prices. Therefore, we expect $\text{corr}(P_{ms}, \xi_{mds}) < 0$ and the OLS estimate of α_1 to be downward-biased. Because mills must pay the same price to all farmers in a given season, the correlation between prices and the error term is a concern only along the mill-season dimension.

[19] In each district-season, small mills are defined 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%.

[20] We do not divide S_{mds} by the market share of the outside option, as it is absorbed by the district-season fixed effect.

We address price endogeneity by leveraging the ICAFE regulation. As explained in Section 2.2, each season, ICAFE determines a mill-specific coffee yield, known as the *rendimiento técnico*, which is used to calculate the minimum price mills must pay farmers at the end of the harvest season. We use the *rendimiento técnico* as an instrument for prices in equation (12).

This instrument is uncorrelated with market competition conditions. ICAFE computes the *rendimiento técnico* based on the physical properties of coffee beans, which are determined by weather, geography, and agronomic aspects of the processing process. Unobserved mill-season-specific shocks are unlikely to systematically affect the correlation between the *rendimiento técnico* and prices. ICAFE collects coffee samples from multiple buying stations across different districts for each mill, ensuring the accuracy of the minimum price.[21] In terms of monotonicity, the instrument only shifts prices of mills that are reporting a yield below the one that ICAFE estimates.

In Section 6.2, we experiment with heterogeneous price sensitivities by mills' organizational form and farmers' characteristics. We also use the survey answers to inspect the determinants of the estimated mill-district fixed effect.

Bilateral Oligopoly. We recover an empirical distribution of the parameter governing exporters' risk sensitivity (κ_{es}) by inverting equation (6):

$$\kappa_{es} = \frac{1}{|\mathcal{M}_{es}|} \sum_{m \in \mathcal{M}_{es}} \frac{p^w_{es} - p_{mes}}{q_{mes}}. \tag{13}$$

Exporters with higher margins per quantity traded are those who can sustain higher risks on the forward market. Similarly, since mills' and exporters' prices and marginal costs are observed in the data, we recover en empirical distribution of the mill bargaining power (λ_{mes}) by inverting equation (8):[22]

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

A high margin for mills relative to exporters in the forward market implies a high λ_{mes}.

In Section 6.2, we further explore sources of variation in sensitivity to forward risk

[21] Mills can appeal the minimum price proposal, making potential inaccuracies costly for ICAFE.

[22] Information about marginal costs is usually hard to observe in the data, and the existing literature typically relies on numerical solutions to recover parties' bargaining power (e.g., Crawford and Yurukoglu, 2012; Crawford et al., 2018; Leong et al., 2022).

and bargaining power by examining how κ_{es} and λ_{mes} change with mills' and exporters' organizational forms and characteristics.

Entry Costs. Since we do not require a unique equilibrium, mills' fixed costs of entering local districts cannot be point-identified.[23] 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. 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.[24]

We use the necessary conditions for an SPNE in equations (10) and (11) to derive upper and lower bounds on the fixed costs mills incur to source coffee from a district in each season:

$$F_{mds} \leq E_{u|\theta}\left[\pi_{ms}(\mathcal{A}_{ms}; u, \theta) - \pi_{ms}(\mathcal{A}_{ms} - 1^d_{ms}; u, \theta)\right] \equiv \overline{F}_{mds}(\theta) \quad \forall d \in \mathcal{A}^1_{ms}, \tag{15}$$

$$F_{mds} \geq E_{u|\theta}\left[\pi_{ms}(\mathcal{A}_{ms} + 1^d_{ms}; u, \theta) - \pi_{ms}(\mathcal{A}_{ms}; u, \theta)\right] \equiv \underline{F}_{mds}(\theta) \quad \forall d \in \mathcal{A}^0_{ms}. \tag{16}$$

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 Appendix C.1 for computational details on the estimation of fixed costs.

This strategy identifies upper bounds for $d \in \mathcal{A}^1_{ms}$ and lower bounds for $d \in \mathcal{A}^0_{ms}$. This is because mills make district choices knowing the realization of u_{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

[23] This holds even with stronger functional form assumptions for the fixed costs of entry in equation (9), as multiple equilibria prevent constructing a well-behaved likelihood function, as in Bresnahan and Reiss (1991).

[24] We do not model the entry decisions of cooperatives, as they are local entities that do not to source coffee from districts without members.

district, and below by the smallest change from adding one.

In Section 6.2, we explore sources of variation in fixed costs of entry by examining how F_{mds} changes with mills' organizational form and characteristics.

6.2 Estimation Results

Farmers' Supply. Table A.5 shows the estimates we obtain from equation (12). As expected, the OLS estimate of α_1 is downward-biased. The average own-price elasticity implied by the OLS estimate in column (1) is 0.2. The one implied by the IV estimate in column (3) is 3.1.[25] The instrument is relevant, as indicated by a Kleibergen-Paap F-statistic of 26.07 in column (4) and the first-stage estimates in Table A.6.

There is heterogeneity in price sensitivity across organizational forms. Farmers are more price-sensitive when selling to private integrated mills and cooperatives than to private non-integrated mills. One explanation for this pattern is that private non-integrated mills are often local family businesses with long-standing relationships with farmers, making some farmers less sensitive to prices when selling to them.

Larger farmers appear to be less price sensitive than smaller ones, as shown in Table A.7. One possible reason is that larger farmers are less financially constrained and may prioritize non-price factors. In contrast, there is no evidence of heterogeneity in price sensitivity by gender.

Column (1) of Table A.8 shows that private integrated mills and cooperatives have higher average fixed effects than private non-integrated mills. This premium can be explained by their size, as shown in column (2). To further investigate the role of non-price factors in farmers' choice of mills, we correlate the estimated average mill fixed effects with mills' survey responses.[26] Figure B.2 presents the results. Mills offering cash advances, loans, and agricultural inputs to farmers tend to have higher fixed effects. Mills participating in voluntary sustainability programs also show higher fixed effects.

[25] Because of logit supply, we calculate the own-price elasticity of mill m in district d in season s as $\alpha_1 P_{ms}(1 - S_{mds})$. In the text, we report the average elasticity across all mills, districts, and seasons, based on the OLS and IV estimates of α_1. Notation comes from equation (1) and equation (2).

[26] Since survey responses were recorded anonymously in compliance with the ICAFE regulation, we proceed in two steps. First, we estimate the mill-district fixed effect using the sub-sample of mills that were active throughout the sample period. We then compute the average fixed effects across districts for each mill. Second, we aggregate both the estimated average mill fixed effects and survey responses by deciles of the mills' size distribution, measured by total coffee production (available in both ICAFE's data and the survey). We report correlations at this level of aggregation.

Overall, the evidence suggests that farmers prefer large, private integrated mills that offer them services beyond price payments.

Bilateral Oligopoly. Figure B.3 shows the empirical distribution of exporters' estimated sensitivity to risk when trading on the forward market (κ_{es}). Table A.9 shows that vertically integrated exporters can sustain higher risk, both unconditionally and conditionally on size. This is consistent with their ability to trade larger volumes of coffee across all seasons.

Figure B.4 shows the empirical distribution of mills' bargaining power when trading on the forward market (λ_{mes}). Table A.10 shows 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 Tables A.3 and A.4). This behavior likely reflects integrated exporters' ability to better exploit favorable market conditions.

Entry Costs. Figure B.5 shows the empirical distribution of the upper and lower bounds of the fixed costs mills incur when sourcing coffee from districts in a given season. Our estimation procedure recovers 3,575 upper bounds and 1,235 lower bounds. Table A.11 shows that the average cost ranges from approximately 425 USD to 673 USD, which corresponds to about 10% to 36% of mills' revenues, as shown in Table A.12. Finally, Table A.13 shows that private integrated mills face higher entry costs. This is consistent with the fact that they are the largest type of mills in the country and can afford entering districts even when entry costs are high.

7 Regulating the Coffee Value Chain

In this section, we use the model to assess the impact of policy interventions by comparing various outcomes between the counterfactual and baseline equilibrium. Our primary focus is on the effects of regulatory changes on welfare along the chain, which are captured by changes in farmers' surplus and in mills' and exporters' profits.

7.1 Rent Sharing Regulation

Policy Design and Trade-Off. As specified in equation (4), the ICAFE regulation prescribes that mills retain a share $\mu = 9\%$ of their profit margin. The rest is transferred to farmers as a liquidation price at the end of each harvest season. We investigate how the market equilibrium would look like under different values of μ. Reducing the share of margins retained by the mill generates the following trade-off for the regulator. On one hand, it directly increases farmers' surplus and reduces mills' profits. On the other hand, a lower μ could lead to the exit of some mills from some districts. This would benefit incumbent mills but indirectly harm farmers. The regulator faces the opposite trade-off if μ increases. Notice that changing μ does not influence the price mills and exporters negotiate in equation (8). As a result, this policy has no impact on exporters.

Implementation. We simulate the effects of increasing or decreasing μ under two scenarios. In the first, mills cannot change their sourcing choices. In the second, while keeping holding the total number of mills constant, they are free to enter new districts or exit old ones. Comparing counterfactual changes across these scenarios highlights the role of mills' entry and exit choices in shaping the distribution of welfare changes along the chain. 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 their sourcing choices is more 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 a distribution of mills across districts and 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 allocations 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.

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 (12) at the mill level 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 proceed as follows: we let the highest-ranked mill that has not entered a district enter that district. We iteratively let mills enter until until an incumbent would make negative profits. We repeat this procedure for each district and season in the data. 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 Appendix C.2 for additional details.

Results. The top left and right panels of Figure B.6 show the distribution of farmers' surplus across districts (averaged over seasons) and mills' profits (summed over districts and averaged over seasons), respectively, when mills are not allowed to change the set of districts from which they source. As expected, increasing μ harms farmers and benefits mills, while decreasing μ has the opposite effect. The magnitude of these direct effects increases with the absolute value of μ.

The bottom panels show the changes in farmers' surplus and mills' profits when mills can endogenously change the set of districts they source from. While the direct effects described in the top panels still dominate on average, mills' endogenous entry and exit responses dampen both gains and losses. For example, increasing μ leads mills to enter new districts, which mitigates the direct loss that farmers experience. In some districts, where the pro-competitive effects are strong enough, farmers are even better off than under the initial equilibrium. Overall, ignoring mills' endogenous changes in sourcing decisions leads to overestimate gains and losses in absolute terms. Figure B.7 shows the change in the number of mills, both in levels and relative to the number of incumbents.

Figure B.8 shows that when mills are allowed to change their sourcing strategies, increasing or decreasing μ makes mills with a lower estimated mill-district fixed effect source relatively more coffee, thereby reducing allocative efficiency. Finally, Figure B.9 shows 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.

7.2 Banning Vertical Integration

Policy Design and Trade-Off. A common concern in agricultural chains is that vertical integration between mills and exporters may facilitate market foreclosure and depress payments to farmers. This has led some to advocate for divesting mills and exporters.[27] However, as we also highlight in Section 6.2, vertically integrated mills and exporters tend to be more efficient and have a more secure marketing channel downstream, which may benefit farmers upstream. We use our model to simulate divestitures between mills and exporters and quantify which forces prevail in the counterfactual equilibrium.

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

To convert private integrated mills and exporters, we assign them the appeal (η_{md}), marginal costs (c_{ms}), entry costs $[\underline{F}_{mds}, \overline{F}_{mds}]$, price sensitivity ($\alpha_1$), 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 whether a mill or an exporter is integrated. The resulting integration premium is then subtracted from the values of each variable for the integrated mills and exporters.

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 set. 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

[27]Similar concerns and proposals have been expressed, among others, in industries such as cocoa in Ghana, cotton in Tanzania, and coffee in Ethiopia and Kenya.

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 out the market until all remaining mills are profitable. We repat this procedure for all districts and seasons.

Results. Figure B.10 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. 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. On average, formerly integrated private exporters also suffer losses.

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 erode the profit gains of incumbents.

Unlike in Section 7.1, there is no evidence that changes in farmers' surplus correlate with the initial characteristics of farmers, as shown in Figure B.11.

8 Conclusions

Enhancing the efficiency of agricultural chains has the potential to improve living standards in many emerging economies. However, it is challenging to determine which policies are effective and which are not, and to quantify their welfare and distributional impact along the chain. In the absence of experimental regulatory changes across countries, structural models can help make progress.

In this paper, we develop a structural model of agricultural chains to serve this purpose. We apply the model to the coffee chain in Costa Rica, which provides an ideal context due to its data availability and the regulatory features in place. We demonstrate how to estimate the structural parameters of the model and quantify the welfare effects of widespread regulatory interventions in agricultural markets.

From a methodological point of view, we show how tools from the empirical industrial organization literature can be used to study market frictions and policy interventions along agricultural value chains. Our model consists of three key building blocks common to many contexts: transactions between farmers and mills, transactions between mills and exporters, and the sourcing choices that mills make across local markets. Properly adapted, this framework can be used in other settings.

From an empirical point of view, we provide insights into the structural reasons behind the dominance of large, vertically integrated mills and exporters in the Costa Rica coffee chain. By securing market access downstream, these vertically integrated organizations can offer better services to farmers upstream and source coffee from a broader range of districts. Through counterfactual analysis, we show that policy interventions aimed at preserving competition among mills without distorting allocative efficiency benefit farmers. Conversely, banning privately integrated organizations—a common proposal by regulators worldwide—prevents farmers from accessing the services provided by these organizations and ultimately makes them worse off.

References

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.

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.

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

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 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.

Dillon, B. and C. Dambro, "How Competitive are Crop Markets in Sub-Saharan Africa?," *American Journal of Agricultural Economics*, 2017, *99* (5), 1344–1361.

Dominguez-Iino, Tomas, "Efficiency and Redistribution in Environmental Policy: An Equilibrium Analysis of Agricultural Supply Chains," *Cited on*, 2021, p. 6.

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.

Emran, M Shahe, Dilip Mookherjee, Forhad Shilpi, and M Helal Uddin, "Credit Rationing and Pass-Through in Supply Chains: Theory and Evidence from Bangladesh," *American Economic Journal: Applied Economics*, 2021, *13* (3), 202–236.

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.

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.

Laoprapassorn, Bunyada, "Entry and Spatial Competition of Intermediaries: Evidence from Thailand's Rice Market," Technical Report, Puey Ungphakorn Institute for Economic Research 2022.

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

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 and Ameet Morjaria, "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports," *American Economic Review*, 2015, *105* (9), 2911–45.

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

— , **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.

Maitra, P., S. Mitra, D. Mookherjee, A. Motta, and S. Visaria, "Financing Smallholder Agriculture: An Experiment with Agent-Intermediated Microloans in India," *Journal of Development Economics*, 2017, *127*, 306–337.

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).

Morlacco, Monica, "Market Power in Input Markets: Theory and Evidence from French Manufacturing," *Unpublished, March*, 2019, *20*, 2019.

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.

Porto, G., N. Depetris Chauvin, and M. Olarreaga, "Supply Chains in Export Agriculture, Competition, and Poverty in Sub-Saharan Africa," 2011. The World Bank and Centre for Economic Policy Research.

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

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.

Zavala, Lucas, "Unfair Trade? Monopsony Power in Agricultural Value Chains," *The World Bank*, 2022.

Appendices

A Tables

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
Private Integrated	119.04	13.19	2.97	0.45	0.33
Private Not Integrated	17.53	1.63	3.87	0.64	0.07

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

Table A.2. Survey Summary Statistics

Agricultural Inputs	Cash Advances	Loans	Sustainability Program
0.25	0.19	0.19	0.59

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). 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.

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

	No. of Exporters	Relationship Age	Average Price	Log Total Quantity
Arms' Length	1.34	4.14	4.93	10.06
Integrated Mill	1.11	8.28	4.03	13.78

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 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 distinguish between mills that are vertically integrated with downstream exporters and non-integrated mills (cooperatives or private).

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

	No. of Mills	Relationship Age	Average Price	Log Total Quantity
Arms' Length	2.38	3.24	4.93	10.71
Integrated Exporter	13.76	3.97	4.56	11.43

Note: The first column reports the average number of mills that exporters source form. The second column reports the average number of consecutive periods that mills trade with exporters. The third columns shows the average price 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 distinguish between mills that are vertically integrated with downstream exporters and non-integrated mills (cooperatives or private).

Table A.5. 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 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 A.6. Farmers' Supply - First Stage

VARIABLES	(1) P_{ms}	(2) P_{ms}	(3) P_{ms}
$\texttt{yield}_{ms}$	0.03***	0.04***	
	(0.01)	(0.01)	
$\texttt{yield}_{ms}$ × Cooperative			0.03***
			(0.01)
$\texttt{yield}_{ms}$ × Private Not Integrated			0.04***
			(0.01)
$\texttt{yield}_{ms}$ × 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. $\texttt{yield}_{ms}$ 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.

Table A.7. Farmers' Supply - Second Stage - Heterogeneity

VARIABLES	(1) $\log(S_{mds})$	(2) $\log(S_{mds})$
P_{ms}	1.10***	1.24***
	(0.43)	(0.46)
$P_{ms} \times$ Share of Large Farmers$_{ds}$	-0.45*	
	(0.27)	
$P_{ms} \times$ 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 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.

Table A.8. 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)
Constant	-10.68***	-14.00***
	(0.26)	(1.37)
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 A.9. Exporters' Forward Risk Sensitivity

	(1)	(2)
VARIABLES	κ_{es}	κ_{es}
Integrated Relationship	-0.78***	-0.67***
	(0.20)	(0.18)
Integrated Exporter	-0.15	0.39
	(0.26)	(0.31)
Log Exporter Size		-0.33***
		(0.11)
Observations	1,930	1,930
Estimator	OLS	OLS

Note: The unit of observation is a mill-exporter-season tuple. The dependent variable is the sensitivity of exporter e in season s to risk on the forward market. Integrated Relationship is a binary variable equal to one if mill m and exporter e are vertically integrated. 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-season are shown in parentheses. Significance levels: *** 0.01, ** 0.05, * 0.1.

Table A.10. 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 A.11. 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.

Table A.12. Entry Costs as a Share of Mills' Revenues

	Number	Mean	Q25	Median	Q75
Upper Bound	3,575	0.36	0.03	0.10	1.00
Lower Bound	1,235	0.10	0.02	0.05	0.11

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 as a share of mills' revenues in that season.

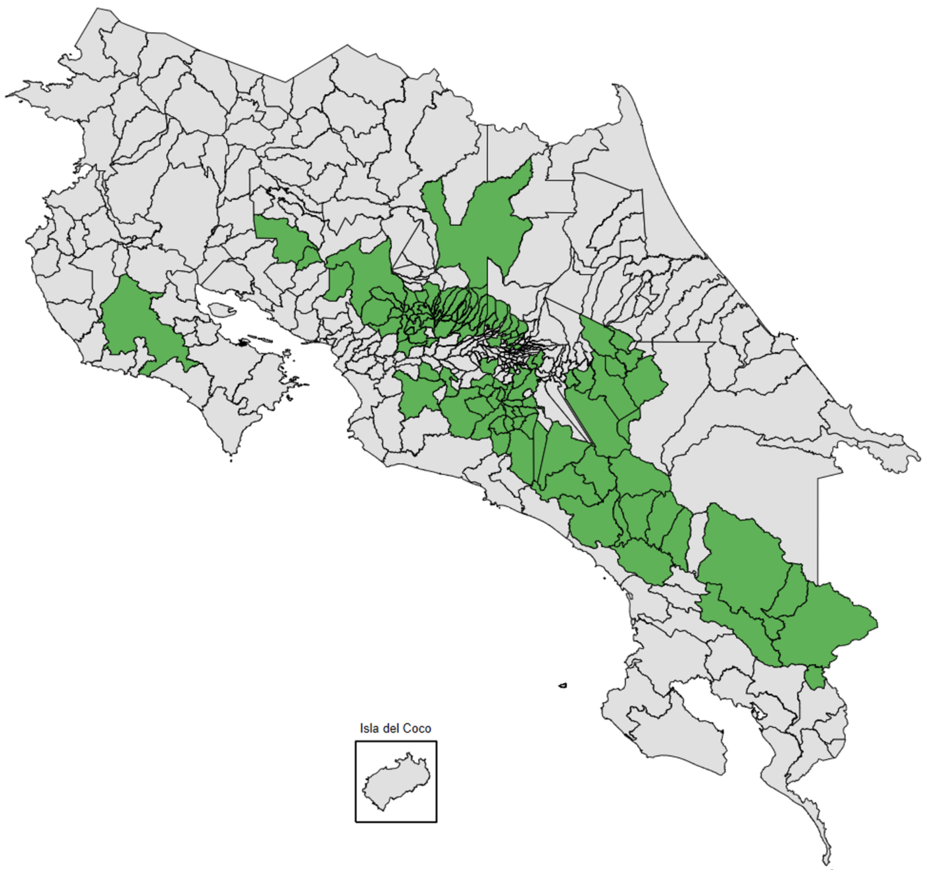
Table A.13. Entry Costs' Determinants

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.

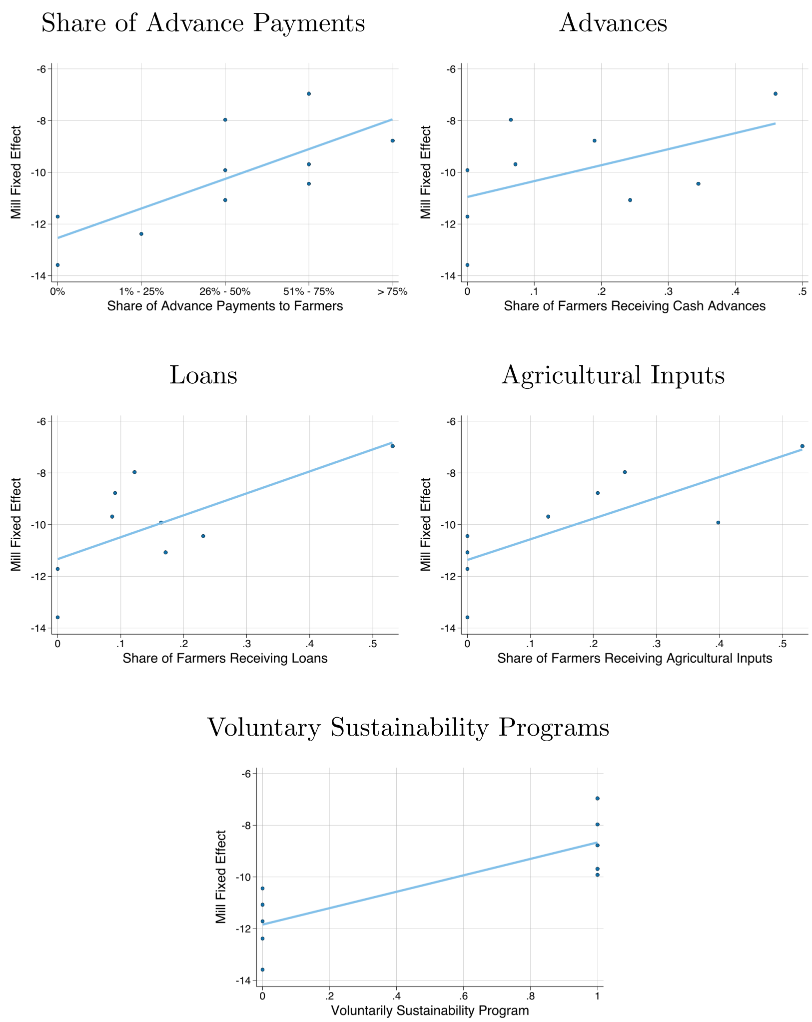
B Figures

Figure B.1. Merged Districts



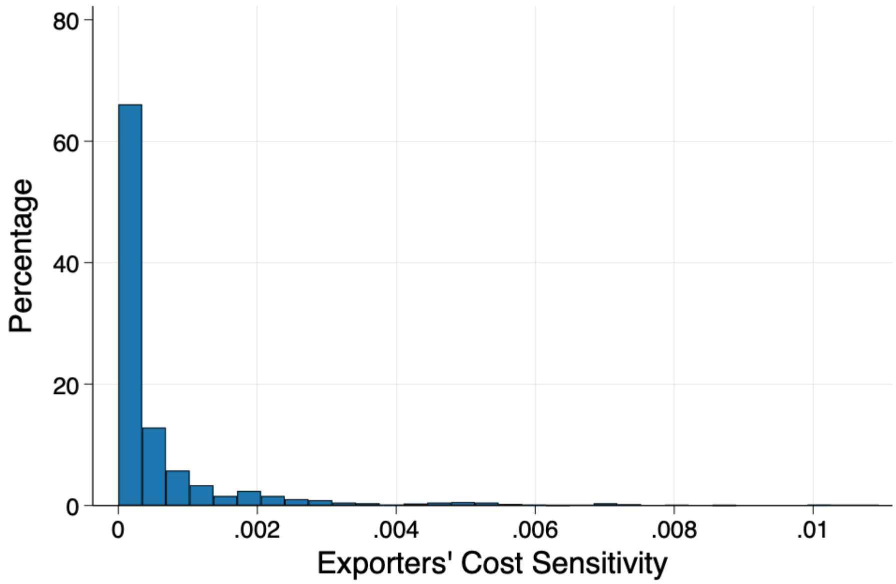
Note: The figure shows the 105 (merged) districts that we use to define the local markets within which mills compete to source coffee from farmers in each harvest season. Green districts are those where coffee is grown and harvested every season in the data. They account for 98% of coffee production in any given season. 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.

Figure B.2. The Determinants of Mills' Appeal to Farmers



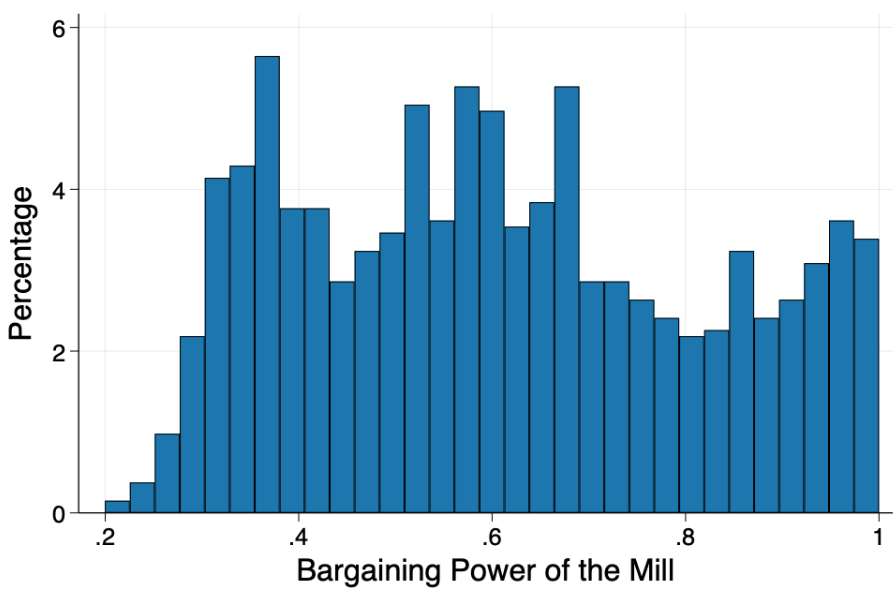
Note: We correlate the mill fixed effects estimated from equation (12) with the mills' responses in the survey. Since the survey is anonymous, both variables are collapsed by deciles of mill size, and correlations are reported at this level of aggregation. The top left panel shows the correlation between the mill fixed effect and the share of advance payments mills offer to farmers. The top right panel shows the correlation with the share of farmers receiving any cash advances. The middle left panel correlates the mill fixed effects with the share of farmers receiving loans, while the middle right panel shows the correlation with the share of farmers receiving agricultural inputs. Finally, the bottom panel correlates the mill fixed effect with whether mills participate in any voluntary sustainability programs (Rainforest Alliance, 4C, Starbucks, or Nespresso initiatives).

Figure B.3. Empirical Distribution of Exporters' Forward Risk Sensitivity



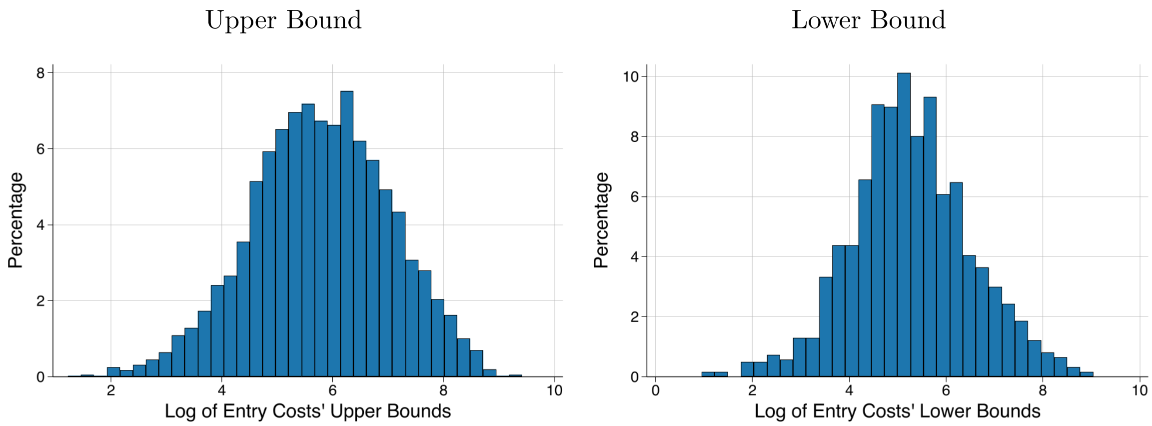
Note: The figure shows the empirical distribution of the exporters' sensitivity to risk when trading on the forward market. We estimate this distribution from equation (13).

Figure B.4. Empirical Distribution of Mills' Bargaining Power



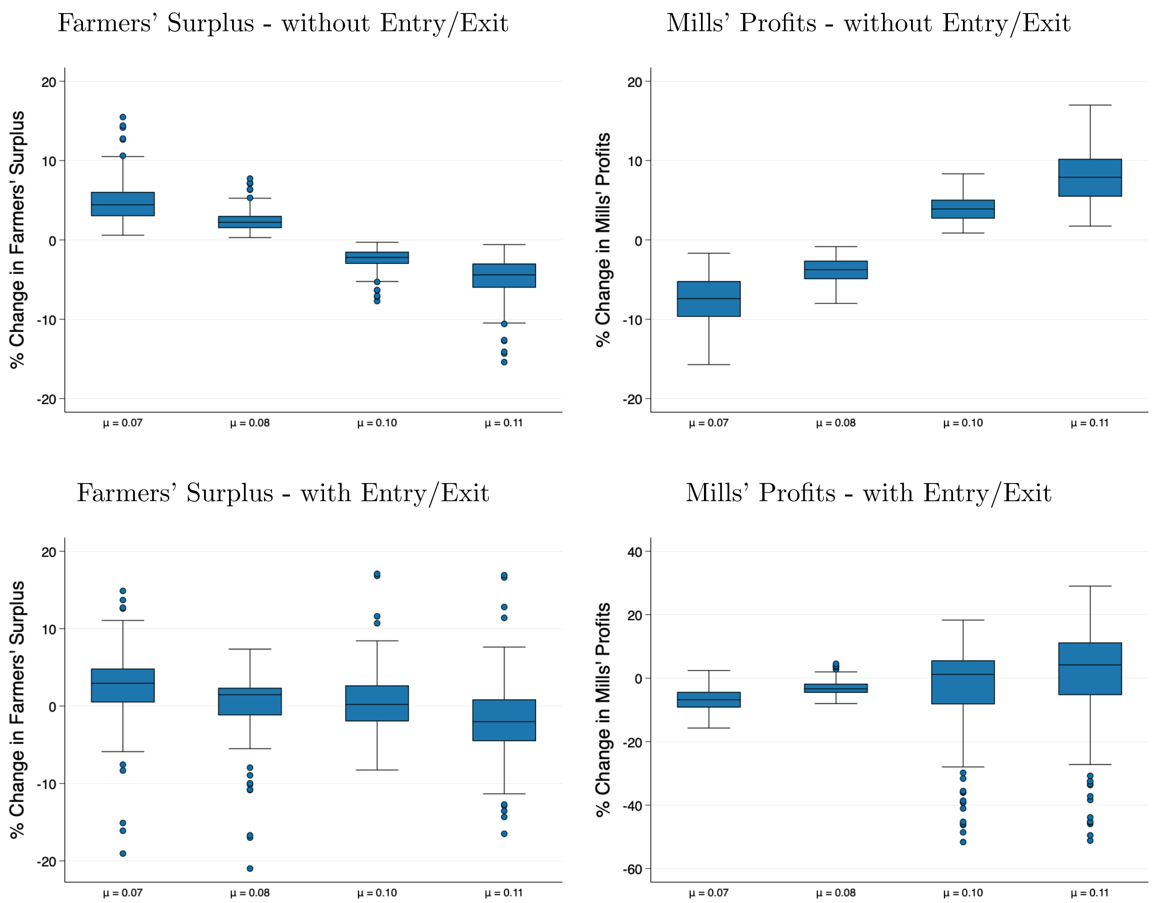
Note: The figure shows the empirical distribution of the bargaining power of the mill on the forward market. We estimate this distribution from equation (14).

Figure B.5. Entry Costs' Bounds



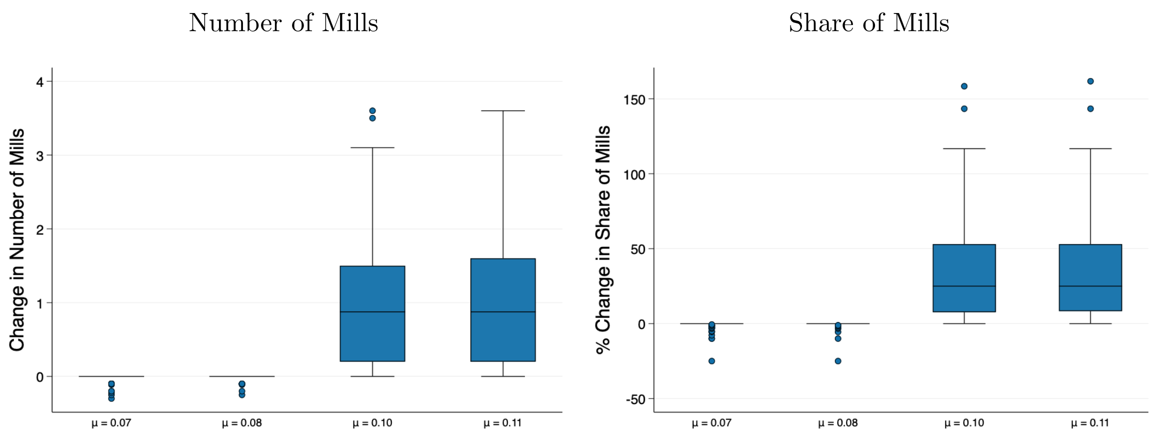
Note: The left panel shows the empirical distribution of the log of the upper bound of the fixed costs that mills face when entering districts in a given season. The right panel shows the empirical distribution of the log of the lower bound of the fixed costs that mills face when entering districts in a given season.

Figure B.6. Rent Sharing Counterfactual - Farmers' Surplus and Mills' Profits



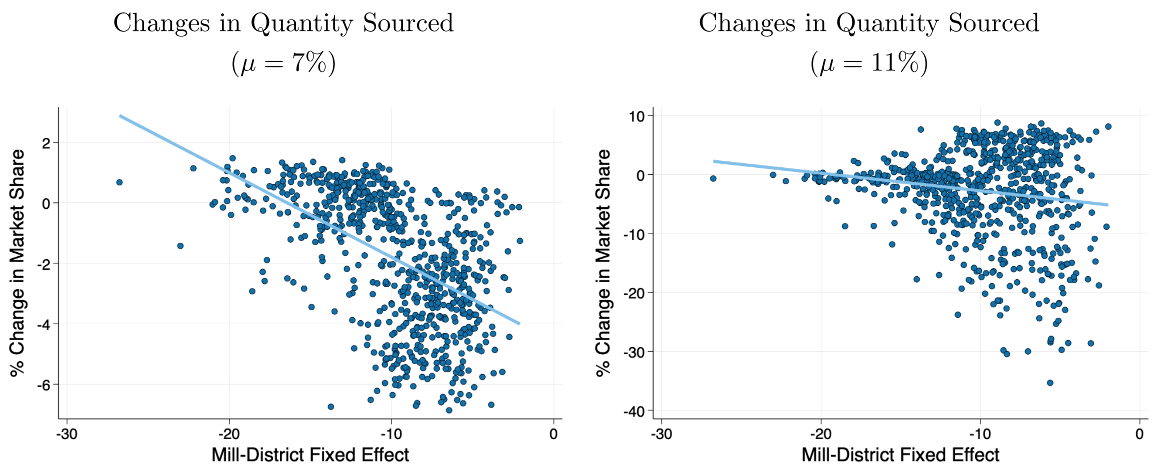
Note: The top left panel shows the distribution of changes in farmers' surplus across districts (averaged over seasons) when μ is decreased or increased. The top right panel displays the distribution of changes in mills' profits (summed over districts and averaged over seasons) with the same change in μ. In both cases, mills are not allowed to adjust their sourcing choices. The bottom panels show changes in farmers' surplus and mills' profits when mills are allowed to endogenously adjust their sourcing decisions. All changes are relative to the initial equilibrium with $\mu = 9\%$.

Figure B.7. 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 $\mu = 9\%$.

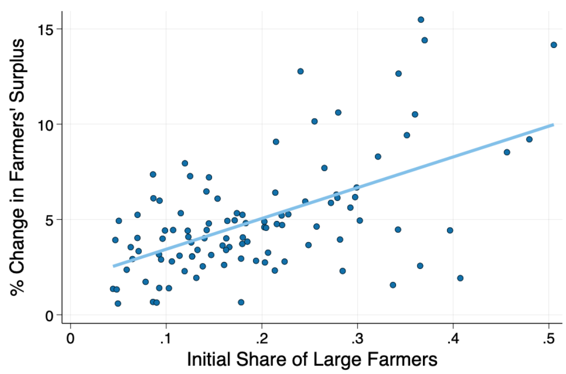
Figure B.8. Rent Sharing Counterfactual - Allocative Efficiency



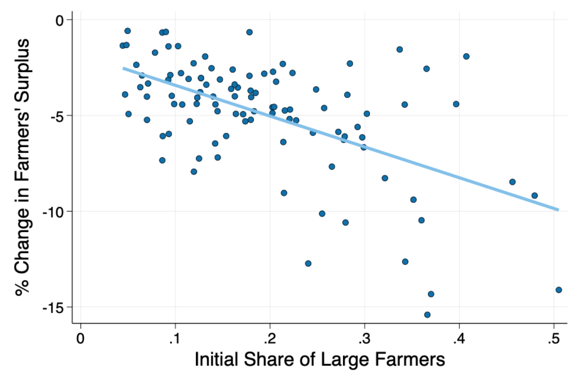
Note: The figure shows the correlation between changes in the market share each mill in her districts (averaged over seasons) and the estimated mill-district fixed effect. The left panel shows changes in market shares when μ is lowered from 9% to 7%. The right panel shows changes in market shares when μ is increased from 9% to 11%.

Figure B.9. Rent Sharing Counterfactual - Farmers' Surplus and Characteristics

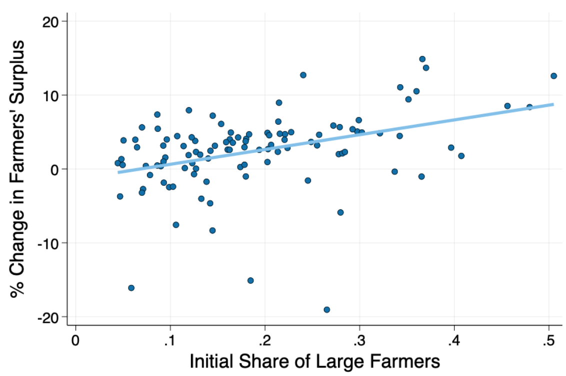
Farmers' Surplus - without Entry/Exit
($\mu = 7\%$)



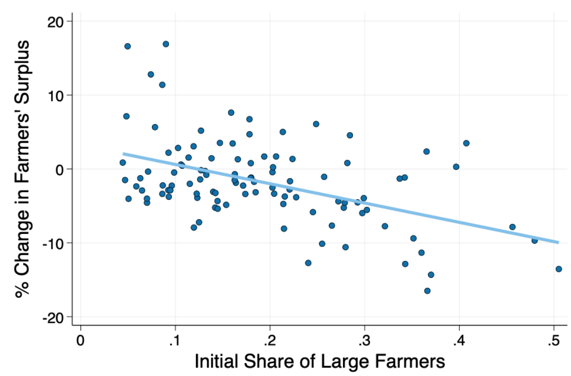
Farmers' Surplus - without Entry/Exit
($\mu = 11\%$)



Farmers' Surplus - with Entry/Exit
($\mu = 7\%$)

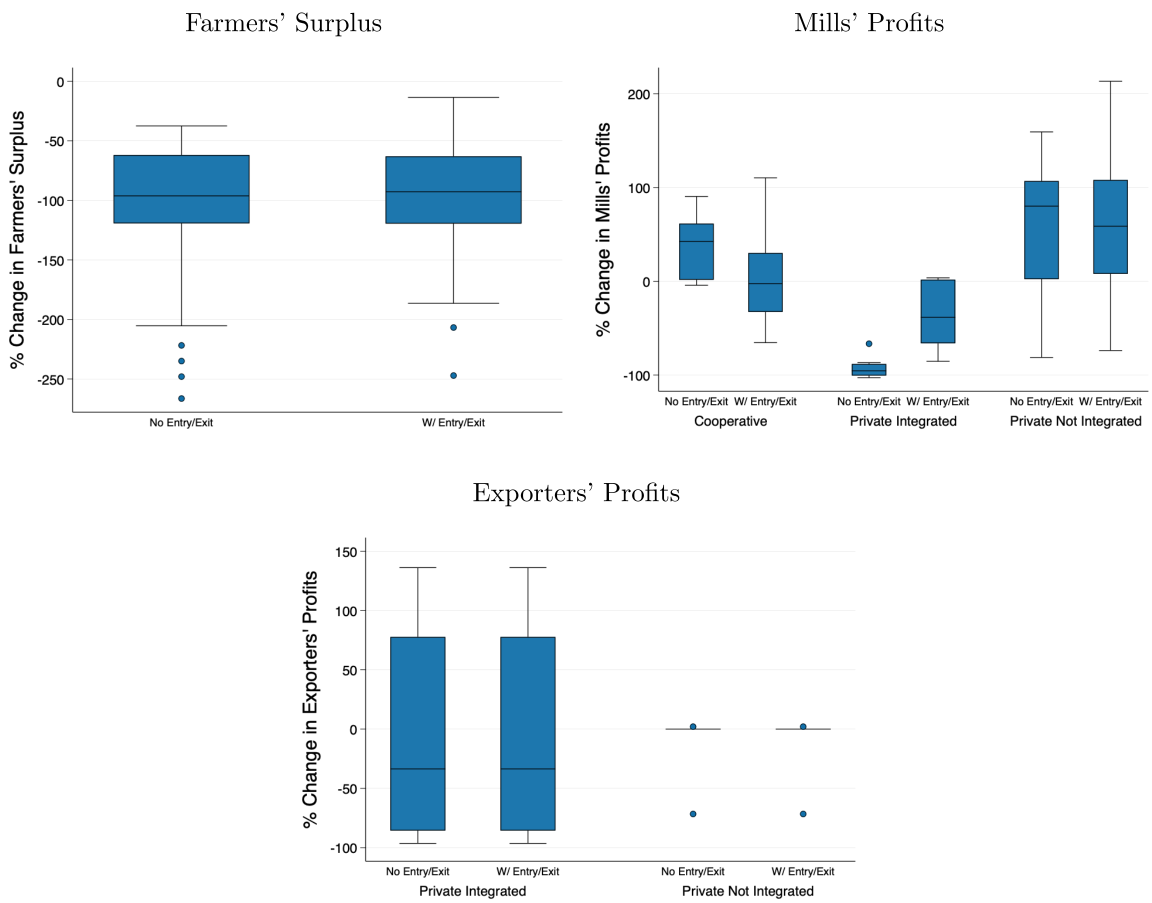


Farmers' Surplus - with Entry/Exit
($\mu = 11\%$)



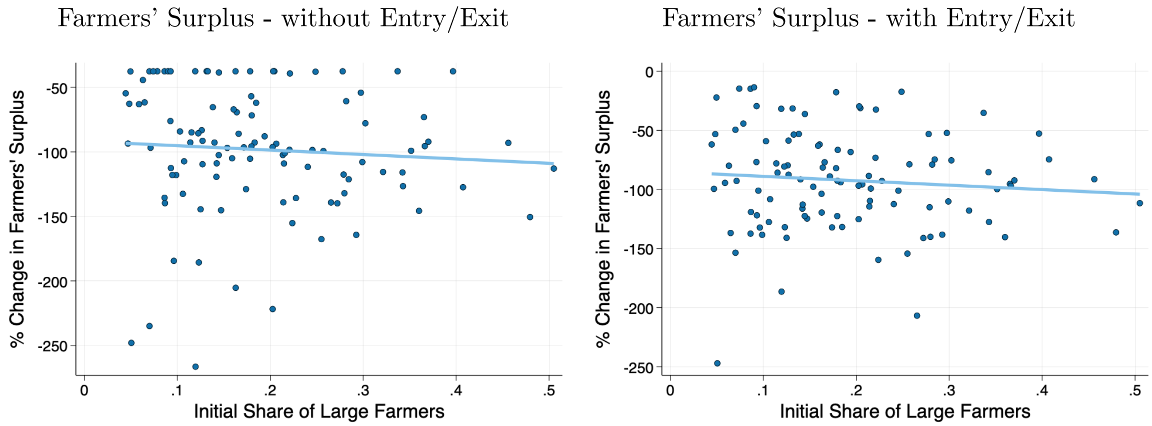
Note: The figure shows the correlation between changes in farmers' surplus in each district (averaged over seasons) as μ changes and the initial share of large farmers in that district (also averaged over seasons). The top left and right panels show these correlations for $\mu = 7\%$ and $\mu = 11\%$, respectively, assuming mills cannot change the districts they source coffee from. The bottom left and right panels show the same correlations but allowing mills to adjust their sourcing decisions. We define the share of large farmers are the share of farmers in a district who produce more coffee than the national average in a given season.

Figure B.10. Divestitures Counterfactual - Farmers' Surplus & Mills' and Exporters' Profits



Note: The top left panel shows the distribution of changes in farmers' surplus across districts (averaged over seasons) when integrated mills and exporters are divested. The top right panel shows the distribution of changes in mills' profits (summed across districts averaged over seasons) when integrated mills and exporters are divested. We show the results for cooperatives, formerly private integrated mills, and non-integrated mills. The bottom 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 B.11. Divestitures Counterfactual - Farmers' Surplus and Characteristics



Note: The figure 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). The left panel shows the results assuming mills cannot change the districts they source coffee from. The right panel shows the same correlation but allowing mills to adjust their sourcing decisions. We define the share of large farmers are the share of farmers in a district who produce more coffee than the national average in a given season.

C Computational Appendix

This section provides computational details of our model.

C.1 Fixed Cost Estimation

We estimate entry costs in equation (15) and (16) as follows:

1. We estimate equation (12) 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;
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 (15). 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 (16).

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 (9) as an average over upper and lower bounds for that mill, respectively.

C.2 Implementation of the Counterfactual Exercises

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:

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

$\mathcal{S}_m$ 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.