Buyer-Driven Upgrading in GVCs: The Sustainable Quality Program in Colombia

Rocco Macchiavello
London School of Economics

Josepa Miquel-Florensa
Toulouse School of Economics

August 13, 2019*

This paper studies the Sustainable Quality Program in Colombia – a quality upgrading program implemented on behalf of a multinational coffee buyer. The Program is a bundle of contractual arrangements involving farmers, intermediaries, exporters and the multinational buyer. We tackle three questions. First, we investigate the impact of the Program on the supply of quality coffee. Eligible farmers upgraded their plantations, expanded land under coffee cultivation, increased quality and received higher farm gate prices. Second, we quantify how the Program gains are shared between farmers and intermediaries along the chain. In regions in which the Program was rolled out surplus along the chain increased by ≈ 30%. Eligible farmers kept at least half of the gains and their welfare increased by ≈ 20%. Finally, we examine how the Program works conducting counterfactual exercises and comparing the Program price premia along the chain against two prominent non-buyer driven certifications. The Program achieved a better transmission of the export gate price premium for quality to the farm gate and curbed market failures that stifled quality upgrading. Contractual arrangements at the export gate significantly contributed to higher farmers welfare in rural areas.

Keywords: Quality Upgrading, Buyer-Driven Supply Chain, Contracts, Market Structure, Voluntary Standards.

JEL Codes: O12, F61, L23, Q12, Q13

---

*Macchiavello: r.macchiavello@lse.ac.uk; Miquel-Florensa: pepita.miquel@tse-fr.eu.

Without implicating them, we thank many people at Federación Nacional de Cafeteros, and specially Oscar Bernal, Nancy Gonzalez, José Leibovich and Yair Soto. We thank David Atkin, Matteo Bobba, Laura Boudreau, Lorenzo Casaburi, Arnaud Costinot, Marcela Eslava, Michael Kremer, Gianmarco León, Pierre Mérel, Ameet Morjaria, Mounu Prem, Patrick Rey, Ximena Rueda, Chris Udry, Eric Verhoogen, and seminar and workshop participants at ASSA 2019, Banco de la República (CEARN), BREAD, Barcelona Summer Forum, Cambridge, FAO, IGC-PEDL, IGIER-CSEF, Imperial, LSE, Naples, Northwestern, Notre Dame, Roma III, SciencesPo, Southampton, Syracuse, Tilburg, Toronto, TSE, Universidad de Los Andes, World Bank and WUSTL for comments and suggestions. Yusuke Kuroishi and Sergio Montoya provided excellent research assistance. Miquel-Florensa thanks Universidad del Rosario for their hospitality. No data was collected by, obtained from, and the analysis was conducted independently from, the Program buyer. Due to the confidentiality of the data sharing agreement we do not name the Program buyer nor other certifications that appear in the analysis. Funding from IGC is gratefully acknowledged. All errors remain our own.

1. Introduction

Approximately 75% of the world's poor live in rural areas and earn their livelihood in agriculture. Linking small farmers in developing countries to global value chains (GVCs) can lift millions out of poverty but often requires quality upgrading (World Bank (2020)).[1] Multiple challenges frustrate attempts to upgrade quality in agricultural chains in developing countries. Besides well-documented *supply*-side market failures limiting farmers access to inputs and financial services (World Bank (2007)), a recent literature points to high margins and uncompetitive market structures in domestic chains (Antràs and Costinot (2011), Atkin and Donaldson (2015)) and limited contract enforcement (Antràs (2015), Blouin and Macchiavello (2019)) as further challenges on the *demand* side.

Despite the challenges to upgrade quality, between 2006 and 2012 Colombia nearly doubled (from 9% to 17%) the share of coffee exported as high quality *supremo* (as opposed to standard quality). This rather spectacular increase coincided with the roll-out in the country of the *Sustainable Quality Program* - a quality upgrading program implemented on behalf of a multinational coffee buyer. Detailed export records reveal that around 80% of the aggregate increase in *supremo* coffee was exported through the Program to the multinational buyer. Did the Program simply provide a demand outlet for a supply of quality coffee that was expanding for unrelated reasons?, or did the Program actually contribute to quality upgrading?

By better understanding the Program role in this quality upgrading episode this paper offers a window into how global buyers can help farmers in developing countries overcome barriers to participation in GVCs (World Bank (2020)). We explore three questions: **1**) what is the impact of the Program on farms upgrading and the supply of quality?; **2**) how large are the gains generated by the Program in the Colombia chain and how are they shared between farmers and intermediaries?; **3**) how did the Program work? Our results indicate that the Program induced substantial quality upgrading. We estimate that, in regions in which it was rolled out, the Program increased surplus along the chain by around 30%. Farmers kept at least half of this surplus increase. The Program improved the transmission

[1] Richer countries consume higher quality products than developing countries (Linder (1961), Verhoogen (2008)) and the transition from low-quality to high-quality exports is often seen as a step towards economic development (Hausmann et al. (2007)). Conversely, exposure to export markets might foster quality and productivity upgrading (see, e.g., Atkin et al. (2017)).

of the quality price premium from the export gate to the farm gate and curbed market imperfections that stifled quality upgrading. Our analysis suggests that contractual arrangements at the export gate were key drivers of quality upgrading and higher welfare in rural areas.[2]

Section 2 describes the *Sustainable Quality Program* and its implementation in Colombia. Launched by the multinational buyer in 2003, the Program is currently implemented in over a dozen countries with more than 100,000 smallholder farmers participating globally.[3] The Program aims to reliably source large volumes of consistently high quality, sustainable and traceable coffee for the multinational buyer. The Program is a bundle of contractual arrangements involving all actors along the chain: farmers, intermediaries, exporters and the multinational buyer (see Figure 1 for an illustration of the Colombia case). At the farm gate, the Program combines supply- and demand-side interventions. On the *supply* side, the Program provides training, extension services and access to inputs for plot renewal to support quality upgrading. On the *demand* side, the Program commits to purchase from Program farmers all the production that satisfies its quality requirements at a fixed price premium. Program farmers have the option (but not the obligation) to supply the multinational buyer. At the export gate, long-term relationships with exporters underpin contractual arrangements that include provisions on how the exporter must source from farmers and how implementation costs ought to be shared.

In Colombia coffee is cultivated mostly by smallholders and sold to intermediaries (either private buyers or cooperatives). The Program is implemented by the Federación Nacional de Cafeteros (*FNC*) in partnership with regional cooperatives. The Program was rolled out starting in 2006 to more than a thousand villages over a 10 year period.[4] By the end of our sample period in 2016, approximately 40% of the over 100,000 eligible coffee plantations had joined the Program.

Section 3 investigates the Program impact on farm upgrading and the supply

[2]This paper covers the Colombia chain (from farm to export gate) but is not concerned with what happens between the export gate and the final consumers.

[3]The program is a flagship example of buyer-driven (as opposed to traditional NGO-driven, such as Fair Trade, 4C and Rainforest Alliance) Voluntary Sustainability Standard (VSS). VSSs have become increasingly common in global agricultural chains (see, e.g., Giovannucci and Ponte (2005), Nelson and Pound (2009) and Dragusanu et al. (2014) for surveys).

[4]We refer to *veredas*, the smallest administrative unit in Colombia, as villages or localities. On average, a *vereda* comprises 60 to 70 small coffee cultivating plots.

of quality. We are primarily interested in three sets of outcomes: upgrading of coffee plantations; quality; and sales arrangements. On the first dimension, eligible farmers substituted old, unproductive, trees with younger trees of disease-resistant varieties. The Program also increased coffee cultivation on the intensive (expansion of existing plots) and extensive (entry of new coffee plantations) margins. We document these patterns taking advantage of the staggered roll-out of the Program and a georeferenced panel covering the universe of coffee plots in Colombia. The Program selected eligible localities based on *terroir* conditions that are suitable for the production of quality coffee. All plots within selected localities are eligible to receive training, extension services and access to inputs for plot renewal and join the Program, provided they meet the Program standards. We implement a Difference-in-Difference (DID) design. We assuage concerns arising from the Program targeting of localities with specific *terroir* by controlling for plot fixed effects and focus on Intention-to-Treat (ITT) specifications to take into account farmers endogenous take-up decision. The data allow us to investigate the impact of the Program at scale over a ten year horizon, which is critical in light of the multi-year investments required to upgrade coffee plantations.

The investments to upgrade farms increased the quality of coffee produced. We observe quality tests for coffee batches at the mill gate.[5] Coffee batches sourced by the Program are of higher quality than batches sourced from the same narrow origin at the same time. Furthermore, *non*-Program batches sourced from Program localities are of similar quality than batches sourced from non-Program origins. This lessens concerns that the higher quality of Program batches reflects sorting of higher quality beans into the Program. Detailed export data confirm that the higher quality of Program batches carries through to the export gate.

At the farm gate, the Program overcame hold-up and side-selling, two contracting problems common in agricultural chains (see, e.g., Kranton and Swamy (2008)). First, transaction level data from one of the implementing cooperatives reveal that the Program bought all coffee of adequate quality that Program farmers were willing to sell at the pre-announced price premium ($\approx 10\%$). Second, we investigate side-selling using transaction-level data that cover the universe of coffee

[5]We do not observe quality at the farm gate. We do observe, however, the buying point where the batch was sourced from farmers. Buying points are closely matched to villages that are eligible and non-eligible for the Program. Quality is measured as the share of beans with defects and/or diseases, the beans size, the yield-per-bean and the cupping test.

sales from *any* farmer to *any* buyer. To overcome selection concerns, we implement a Spatial Discontinuity (SD) Design that compares the selling behaviour of farmers with plots within 1 km from a border separating eligible and non-eligible villages. The Program increased the share of farmers selling exclusively, and the amount of coffee delivered, to the Program implementer.

Section 4 attempts to quantify the impact of the Program on the Colombia coffee chain. We are interested in quantifying a) the welfare gains generated by the Program; and b) how they are shared between farmers and intermediaries along the chain. To a first approximation, the answer depends on: i) the increase in quality production; ii) the export and farm gate price premia for quality; and iii) the costs to produce quality. The increase in quality (i) is closely related to Program take-up; the Program farm and export gate price premia (ii) are estimated from transaction-level data. The costs of quality upgrading, (iv), however, must be estimated. While we anchor ourselves as much as possible to the data, we are inevitably obliged to make (sometime strong) assumptions. We interpret results as ballpark figures informing us about outcomes that are difficult to quantify from reduced-form estimates alone.

A stylized model of the Colombia coffee chain disciplines our analysis. Farmers characterized by heterogeneous plot size and costs of upgrading supply either standard or quality coffee. The market for standard coffee is perfectly competitive. One multinational buyer sources quality coffee from one monopsonist exporter. The Program is a bundle of contractual arrangements between the multinational buyer, the exporter and the farmers. At the farm gate, the Program provides access to training, extension services and inputs for plot renovation. These interventions increase production. The Program pays a farm gate price premium and provides farmers with additional demand-side benefits, modeled as a reduced-form parameter to be estimated. Farmers decide whether to upgrade, expand and join the Program comparing costs and benefits. Under standard parametric assumptions on the distribution of the idiosyncratic component of the fixed costs, the farmers decision is represented by a multinomial logit.

The model structure identifies both the distribution of the fixed costs *and* the reduced-form value of the Program demand-side benefits. The key assumption is that, conditional on observables, *non-takers* draw fixed costs of investment and upgrading from the same distribution as *takers*. The intuition behind the identification strategy is then to exploit knowledge of the returns from the investment decisions of *non-takers* to identify the distribution of the fixed costs. Given this distribution, the value of the Program demand-side benefits is identified from *takers* take-up in excess of what implied by the observable Program components.

The estimates imply that the Program had a sizable impact on both quality upgrading and farmers welfare. Farmers that join the Program have larger plantations and expand more. The 40% take-up rate translates into ≈ 60% of production in the Program regions being upgraded to high quality. This magnitude is in line with the aggregate export figures mentioned above. The Program increased welfare of eligible farmers by 19%. Higher quality translates into higher revenues at the export gate. Once the higher variable and fixed costs of quality coffee are net out, the Program increased surplus in the Program regions coffee chain by 33%. Farmers keep 56% of the surplus increase, the rest accrues to the exporter.

The three-layered chain between farmers, exporter and the multinational buyer creates potential for double marginalization in which the exporter sets an inefficiently low farm gate premium. The contract between the exporter (*FNC*) and the multinational buyer includes provisions on the farm gate premium and a lump-sum contribution to cover implementation costs. This potentially curbs double marginalization.[6] We thus represent the Program objective function as a weighted average of exporter profits and farmers welfare.[7] Given the structure of demand and supply, the Program farm gate premium of 10% is best rationalized by an objective function that gives significant weight to farmers welfare.

Section 5 attempts to understand how the Program works. To answer this question we would ideally want to unbundle the different components of the Program: the impact on farm production, the demand-side benefits to farmers and the contractual arrangements between *FNC* and the multinational buyer. Unfortunately, such unbundling is not possible.[8] We thus try to make progress in two distinct, but complementary, ways. First, we explore counterfactual scenarios

[6]The logic is akin to (maximum) resale price maintenance (see, e.g., Rey and Tirole (1986) and Tirole (1988)) with, relative to the standard analysis, farmers (buyer) swapping role with consumers (manufacturer).

[7]This reduced-form representation avoids modeling the multinational margins and the bargaining protocol with the exporter. It is also consistent with *FNC* caring about farmers welfare directly but doesn't have adequate instruments to redistribute profits to farmers as desired.

[8]In principle, it would be possible to evaluate (ideally through a set of RCTs) the Program components at the farm gate. We are however also interested in understanding the role of contractual arrangements at the export gate which would be harder to experiment on.

that alter supply, demand and market structure. This approach yields several insights but makes strong assumptions. Second, we compare price premia along the chain between the Program and two prominent *NGO*-driven Voluntary Sustainable Standards (VSSs): an *environmental* label and a *social* label. In Colombia both VSSs are implemented on behalf of international NGOs by *FNC*. The key difference between the Program and the two VSSs is the multinational buyer contractual arrangement with *FNC*. The two different exercises paint a fairly coherent picture with two main take-away points.

The first take-away is that the contractual arrangements between *FNC* and the multinational buyer appear to play a critical role in inducing quality upgrading and the ensuing increase in farmers welfare. Counterfactuals reveal that supply side interventions alone would generate no more than half of the observed gains. The majority of the gains is due to the Program demand side components. The comparison with the two VSSs confirms that the Program improves the transmission of the price premium from the export gate to the farm gate. This is key for quality upgrading and farmers welfare.

The second take-away is that the Program relaxes one or more market failures along the chain. A smoking gun in favour of this interpretation is that the export gate market premium for *supremo* coffee (estimated at $10\% \times 3.75\$ \approx 0.375\$$) is considerably larger than the (marginal) increase in unit costs to produce quality (which we conservatively bound to be below 0.27$). Under perfectly functioning markets, the gap between the two should vanish. This back-of-the-envelope calculation thus suggests that indeed there were constraints to quality upgrading that were relaxed by the Program. The Program did not simply increase demand for quality at the export gate.

Section 6 reviews some of the limitations of our analysis. Three demand most attention: the analysis considers neither *i*) the distributional consequences across farmers; nor *ii*) the impacts of the Program on competing intermediaries, general equilibrium, and environmental preservation; and *iii*) the Program success in Colombia might depend on the local context. Notwithstanding those limitations, Section 6 also presents policy implications of this study.

The paper contributes to our understanding of how supply-chain linkages in GVCs contribute to upgrading in developing countries. The topic has been covered from multimple angles (see, e.g. Gereffi (1999) for a GVC perspective and Harrison and Rodríguez-Clare (2010) for an industrial policy one). The literature is vast and we confine ourselves to review selected contributions.

Atkin et al. (2017) provides experimental evidence showing that connecting rugs producers in Egypt to foreign buyers fosters quality and productivity.[9] We study a quality-upgrading program implemented on behalf of a multinational coffee buyer and find a positive impact on farmers at the other end of the chain.[10] The Program is a prominent example of buyer-driven VSSs and so we complement Boudreau (2019) experimental analysis of a MNEs-driven initiative to improve workers safety in Bangladeshi garments factories.[11]

Moving beyond the export gate, supply-chain linkages between multinationals (MNEs) and domestic suppliers can be an important conduit towards upgrading (see, e.g., Rodríguez-Clare (1996), Alfaro et al. (2004)). A recent paper by Alfaro-Ureña et al. (2019) provides state of the art evidence. Leveraging detailed data on firm-to-firm linkages in Costa Rica, the paper identifies strong and persistent effects on the performance of domestic firms that join a MNE supply chain. We complement their evidence tracing the impact of a MNE-driven quality-upgrading program along an agricultural chain.

Along domestic chains, market structure is a key driver of the transmission of export gate prices to the domestic economy and how gains from trade are shared along supply chains.[12] Deploying a novel methodology, Atkin and Donaldson (2015) quantify how the gains from falling world prices at the import gate are shared between consumers and intermediaries. They find that intermediaries capture the majority of the surplus, particularly so in remote locations. We find that farmers keep a substantial share of the gains from trade and that contractual arrangements at the export gate are key for this result.[13]

[9]More broadly, causal effects of exporting have been documented for wages and ISO 9000 certification in Mexico (Verhoogen (2008)), technology investments in Argentina (Bustos (2011)) and improvements in working conditions in Myanmar (Tanaka (2017)).

[10]An important novelty of Atkin et al. (2017) is the focus on a *demand*-side intervention. At the farm gate the Program is a bundle of both *demand* and *supply* side interventions. An extensive literature has evaluated government and NGO agricultural programs aimed at relaxing *supply* constraints, such as extension services, fertilizer subsidies, and training (see Magruder (2018) for a recent review). We discuss those contributions in Section 5.

[11]Section 5 also relates our findings to the literature on NGO-driven VSSs.

[12]See, e.g., Antràs and Costinot (2011) and Bardhan et al. (2013) for theoretical contributions.

[13]The literature quantifying the gains from trade and market integration is vast (see, e.g., Costinot and Rodríguez-Clare (2014) and Donaldson (2015) for reviews). Unlike much of this literature, we focus on one sector and abstract from general equilibrium effects. Sector specific studies include analysis of price transmission in Uganda coffee (Fafchamps et al. (2004)) and

The paper thus also contributes to a growing literature on contractual arrangements along both domestic and international supply chains. At the export gate, international sourcing is complicated, particularly in developing countries, by difficulties in contracting with local suppliers (Antràs (2015)). Initiatives like the *Sustainable Quality Program* are part of a broader trend in which global buyers reorganize their supply chains to achieve stronger relationships with suppliers. Relational sourcing mitigates contractual difficulties with suppliers and has been studied, among others, by Antràs and Foley (2015), Macchiavello and Morjaria (2015), Blouin and Macchiavello (2019), Startz (2019) and Cajal-Grossi et al. (2019). This paper models the relationship between the multinational buyer and the exporter in a reduced-form way to focus on understanding impact on farmers.

Vertical integration can mitigate contracting problems associated with quality upgrading and supply assurance. Hansman et al. (2019) document that vertical integration facilitates sourcing of quality inputs in the Peru fish-meal industry. Macchiavello and Miquel-Florensa (2018) shows that vertical integration improves supply assurance in the Costa Rica coffee chain. In both cases, the focus is on vertical integration between exporters and domestic suppliers. In our case, the Program contractual arrangements remove double marginalization along the domestic chain and thus mimic a vertically integrated exporter.[14]

2. Setting: The Sustainable Quality Program in Colombia

2.A. The Colombian Coffee Value Chain

Figure 1 describes the Colombia coffee chain. Coffee is cultivated mostly by smallholders. When coffee cherries turn red they are ripe for harvest. Coffee cherries must be processed immediately after harvest to obtain parchment coffee. In Colombia, farmers undertake the first stage processing. Farmers sell parchment coffee to either private buyers or cooperatives that source coffee in the Colombian country-side. Private buyers include independent traders as well as buying agents for milling companies. Regional cooperatives operate buying points scat-

Bangladeshi edible oild imports (Emran et al. (2019)), and evaluations of policy reforms in Mozambique cashew (McMillan et al. (2002)), Zabian cotton (Brambilla and Porto (2006)) and Madagascar Vanilla (Cadot et al. (2008)).

[14]Casaburi and Macchiavello (2019) and Macchiavello and Morjaria (2019) study the interplay between market structure and imperfect contract enforcement in domestic agricultural chains with a rather different focus.

tered around the country-side. The intermediaries deliver the coffee to hulling mills. Most of the coffee is exported.

The *Federación Nacional de Cafeteros* (henceforth FNC) is the para-statal body in charge of designing and, through the affiliated regional cooperatives, implementing sectoral policies. For example, FNC is in charge of extension services and agronomic research. FNC also implements the *Garantía de Compra*, a program that gives all farmers the opportunity to sell coffee at a publicly announced base price that tracks the world price. Through its commercial arm, Almacafé, FNC exports around 30% of Colombian coffee.[15]

FNC also implements the main Voluntary Sustainability Standards (henceforth VSSs), e.g., Fair Trade, Rainforest Alliance and the Sustainable Quality Program studied in this paper. VSSs vary depending on whether they are managed on behalf of NGOs (e.g., Fair Trade, Organic and Rainforest Alliance) or on behalf of large international buyers; and whether eligibility is at the cooperative or farmer association level (e.g., Fair Trade, FLO) or at the individual plot level (e.g., UTZ, 4C or Rainforest Alliance).

2.B. The Sustainable Quality Program in Colombia

i) Overview: Quality and Environmental Sustainability

The Sustainable Quality Program is an individual plot program managed on behalf of a multinational coffee buyer. The buyer targets the high-quality segment of the single-serve global market. The Program, launched in 2003 and currently implemented in several countries, aims to reliably source large volumes of high quality, sustainable and traceable coffee for the multinational buyer.

The Sustainable Quality Program distinguishes itself from other VSSs due to its focus on quality upgrading and environmental sustainability.[16] On the quality front, the buyer exclusively sources *supremo* beans, i.e., those that satisfy stringent quality criteria (see Table A1). On the environmental front, the Sustainable Quality Program was developed in partnership, and shares environmental standards, with a prominent global environmental certification. In Colombia, both programs

[15]FNC was the only trader and exporter of coffee until the collapse of the International Coffee Agreement in 1989 when the market was liberalized (see Leibovich and Ocampo (1985) for a market and institutional description.)

[16]The two dimensions are intimately connected in coffee cultivation. For example, growing coffee trees under the shadow of the forest canopy increases quality and also protects the environment. See, e.g., Bosselmann et al. (2009), Laderach et al. (2011), Oberthur et al. (2011).

are implemented by FNC and also share training and extension services. The main difference between the Sustainable Quality Program and this certification is the buyer commitment to buy the Program farmers production as long as it complies with the quality standards.

ii) Contractual Arrangements Between the Farm Gate and the Export Gate
The Sustainable Quality Program is a bundle of contractual arrangements involving all actors along the chain. In Colombia the program is implemented by the FNC through its commercial arm Almacafé and the relevant regional cooperatives (see Figure 1 for an illustration). We must thus describe contractual arrangements along a four-layered chain: Farmers, Cooperatives, Almacafé and the Multinational Buyer. Given the multi-layered nature of the chain, it is particularly important to discuss double-marginalization.

While we do not know the details, the contract between the Multinational Buyer and FNC includes provisions on *a*) the export gate price premium; *b*) the farm gate price premium; and *c*) a lump-sum contribution to cover the Program implementation costs. Those include costs for training, extension services and plot renewal support offered to eligible farmers as well as to inspect farms conditions and compliance with the Program requirements.

The relationship between FNC and the affiliated cooperatives, suggest that Almacafé and the cooperatives act as a single (quasi-)vertically integrated exporter from the point of view of the Program. Indeed, an analysis of transaction-level data between Almacafé and the cooperatives described below reveals that there is no double-marginalization of Program price premium at this stage of the chain. We will thus assume a single, vertically integrated, exporter.

Despite this simplification, the Program still involves a three-layered chain between Farmers, the Exporter (FNC) and the Multinational Buyer. There is thus potential for double marginalization in which the exporter sets an inefficiently low price premium at the farm gate. The contract between FNC and the Multinational Buyer possibly reduces the extent of such double marginalization. Under a logic akin to (maximum) resale price maintenance, the two parties might agree on a farm gate price premium higher than what would be set by a profit-maximizing exporter that takes the export gate price as given. In Section 4 we thus represent the Program objective function as a weighted average of exporter profits and farmers welfare, with the welfare weight a parameter to be estimated.[17]

iii) The Sustainable Quality Program at the Farm Gate

Figure 2 illustrates the Program from the point of view of eligible farmers. The production of quality coffee depends on *a) Terroir* conditions; *b)* investments on the farm; *c)* and harvesting and processing practices.[18] After describing these aspects, we discuss how eligible farmers might respond to the Program in terms of upgrading and sales.

Eligibility: The Program selects eligible *veredas*, narrow administrative units comprising on average 60-70 small plots, based on *terroir* conditions. *Terroir* conditions depend on the plot location (altitude, slope, orientation and soil). All plots in selected veredas are eligible to receive extension services and training and to join the Program.

Upgrading and Take-Up: Alongside *terroir*, the plot potential for quality depends on farmers investments in the plantation: tree age, density and variety.[19] Farmers decide how to respond to their plot eligibility. The farmer decides whether to take advantage of extension, training and plot renewal support to upgrade and/or expand the plot or not. To join the Program the plot must meet quality and environmental criteria. Conditional on meeting these criteria, the farmer decides whether to join the Program or not.

Upgrading, expanding and joining the Program entail costs. Those depend on the current state of the plot and on the opportunity cost of having new trees with initially lower yields and, in case of expansion, of substituting land away from other activities. These investments result in higher quality and yields per hectare.

Production: Actual quality further depends on appropriate harvesting (cherries must be picked at the optimal time with frequent harvesting rounds), processing

[17] This convenient representation doesn't require knowledge of the multinational margins and bargaining protocol with the exporter. It also encompasses the case in which FNC cares directly about farmers' welfare but lacks the tools to redistribute surplus to farmers as desired.

[18] See, e.g., Leonel and Philippe (2007), Velmourougane et al. (2011) and, specifically for Colombia, Puerta (2001) and Puerta et al. (2016).

[19] The key variety dimension is whether the trees are resistant to the coffee rust (*roya* in Spanish). The *roya* is a fungus that damages the coffee tree and, consequently, the quality of the beans. The Program has stringent quality requirements (i.e., low tolerance) for bean defects. All coffee planted in Colombia is of the arabica type and rust-resistant varieties are the Castillo, Colombia, Tabi and Costa Rica.

(washing, drying and sorting) and storage (to mitigate humidity and insect contamination) practices at the farm. These additional activities result in higher labour costs relative to the production of standard quality.

Sales: Farmers that join the Program have the option (but not the obligation) to sell coffee that satisfies the buyer's quality requirements for a fixed price premium. Over the sample period the Program paid farmers a price on average 10% higher than the FNC base price.[20] Besides the price premium, the Program (might) provide farmers with other demand-side benefits discussed in Section 4.

Potential Opportunism: The Program is potentially vulnerable to sources of opportunism that are common in agricultural chains and contract farming schemes in developing countries (see, e.g., for a theoretical analysis Kranton and Swamy (2008)). On the buyer's side, the Program could *hold-up* farmers by reneging on the promise to buy *all* production of eligible quality at the announced price premium: the buyer could try to pay a lower premium, dispute quality, or simply refuse to buy all eligible supply when its demand is low. On the farmers' side, there is potential for *side-selling*. The Program does not require farmers to deliver coffee in exchange for inputs, extension and training, acknowledging that such provisions would be *de facto* impossible to enforce.

Given these concerns, it is ultimately an empirical question whether farmers *1)* incur the relevant costs to join the Program and upgrade; *2)* whether they produce higher quality; and *3)* whether they (are able to) sell to the Program at the announced price premium. We tackle these three questions in Section 3.

iv) Roll-Out and Data Sources

This paper focuses on the Program in the Cauca and Nariño regions.[21] Table 1 reports descriptive statistics in 2006 (i.e., before the Program launch in 2007) focusing on the universe of plots and farmers in the municipalities in which the

[20]The price premium is fixed at 400 COP/kg over the FNC base price. In the last two years of our sample period the premium increased to 600 COP/Kg for plots that also hold an environmental certification. The Program sources from the relevant cooperatives buying points (see Figure A1 in the Appendix for an illustration). These buying points pre-existed the Program. The Program therefore did not expand farmers physical access to markets.

[21]Cauca and Nariño account for 84.5% of the Program farmers in Colombia. An early pilot was implemented in the Caldas and Huila departments. In 2013 the Program expanded to the Santander department. For consistency, we omit Santander since not all our data cover the relevant period. When data are available, including Santander doesn't change the results.

Program eventually expanded. Farmers have an average of 0.5 Ha. of land cultivated under coffee and own about 1.4 plots with coffee cultivation. About 90% of farmers receive FNC programs, including training and extension services.[22]

From its launch in 2007 until 2016, the last year in our sample period, the Program was progressively rolled out in over 1000 veredas in 33 municipalities (see Figure A2 for a map). There are 85887 plots in the eligible veredas and, by the end of the sample period, 33236 plots (i.e., 38%) had joined the Program (see Table A3). Figure A3 shows the temporal expansion of the Program in terms of eligible plots. Figure A4 shows that about 20% of eligible plots join in the first year after becoming eligible. The take-up rate stabilizes around 40% five years after the vereda became eligible. Upgrading the plot to join the Program takes time. Farmers might also wait to learn the Program's benefits before upgrading.

Data Sources: Our analysis relies on multiple sources of data. First, we track farmers take-up and upgrading decisions through a geo-referenced panel covering *all* coffee plots in Colombia over the 2006-2016 period. This panel is collected by FNC to assist in the provision of extension services throughout the country. It contains detailed information on the conditions of, and work undertaken on, the plots but very limited information on farmers demographics and no information on quality, production and sales. We thus complement it with two additional sources of data with information on farmers production and sales: *a*) detailed data from one of the implementing cooperatives for the 2015 and 2016 harvests, and *b*) transaction-level data on the universe of *all* sales of coffee from *any* farmer to *any* buyer in 2013. In addition to those, we track quality and price premia along the chain through detailed data at both the mill's and the export gate. Both data sources include unusually detailed information on the quality and origin of each coffee batch. The Data Appendix provides further details.

The analysis relies on administrative data collected independently of, and without the purpose of evaluating, the Program. The data are collected as part of FNC's routine functions. Farmers and coffee batches are thus uniformly covered regardless of their relationship with the Program. No data was obtained, and the analysis was conducted independently, from the Program's buyer.

[22]Table A2 compares eligible and non-eligible localities within the Cauca at Nariño departments. Panel A confirms that eligible veredas are different in terms of terrain conditions. Panel B shows that Program and non-program municipalities are relatively similar in terms of socio-economic characteristics. These facts are reported simply to provide context for the Program. The empirical analysis controls for time-varying differences between municipalities.

3. Reduced-Form Supply Response to the Program

This Section explores how eligible farmers responded to the Program. We investigate three questions: A) Did farmers *upgrade* their plots?; B) Did they produce higher *quality*?; and C) Did they sell their quality produce to the Program at the announced price premium?

3.A. Did eligible farmers upgrade their plots?

Specification: Besides time-invariant *terroir* conditions, the quantity and quality of coffee harvested depend on characteristics of the coffee plantation: the average age of trees (old trees produce fewer and worse beans); and the share of trees that are resistant to diseases. We thus define our main outcome of interest as a standardized (z-score) index of the (negative) average age of the trees and the share of rust resistant varieties on the plot.

To estimate the impact of the Program we take advantage of the panel structure of our data and the staggered rolled out of the Program and estimate:

$$Y_{pvmt} = \beta_0 + \beta_1 \times P_{pvmt} + \gamma_p + \gamma_{mt} + \varepsilon_{pvmt} \tag{1}$$

where Y_{pvmt} denotes the outcome of interest for plot p in vereda v of municipality m in season t. Depending on specifications, P_{pvmt} is an indicator for participation or eligibility into the Program. All specifications include plot fixed effects γ_p and municipality-year fixed effects γ_{mt}.[23] We restrict the sample to municipalities in which the Program was rolled out at some point, but results are identical when we include all municipalities in the Cauca and Narinño regions. The error term ε_{pvmt} is arbitrarily correlated across plots and over time in each vereda.

Identification: There are two main potential sources of bias in the estimation of β_1. First, the Program targeted localities with time-invariant *terroir* characteristics suitable to produce high quality. Plot fixed-effects γ_p control for time-invariant *terroir* characteristics. An additional concern is that the progressive roll-out of the Program across veredas might correlate with time-varying conditions. Municipality-year fixed effects γ_{mt} control for time varying factors at the municipality level that may influence both roll-out and upgrading.

[23]Specifications with year fixed effects alone yield very similar results.

A second concern is that whether a plot is upgraded and joins the Program (and if so, when) is a farmer's endogenous decision likely to be driven by unobservable factors correlated with the outcome of interest. In such cases, OLS specifications in which P_{pvmt} takes value equal to one after the plot joins the Program are biased. For example, entrepreneurial farmers might be more likely to join the Program and upgrade their farms, leading to an upward bias in the OLS. On the other hand, all farmers within selected *veredas* are eligible to receive inputs, training and extension services regardless of their decision to join the Program. Some farmers also need to upgrade their plots to the Program's required standards before joining. Both aspects lead to a downward bias in the OLS.

Our preferred estimates are thus intention-to-treat (ITT) specifications in which T_{pvmt} is equal to one for all plots p in years t after the program has been rolled-out in *vereda* v. The ITT specification circumvents the challenges above but still suffers from concerns typical of DID designs. For example, the timing at which the Program is rolled out to specific *veredas* might correlate with factors associated with plot upgrading. Plot upgrading might also feature differential trends correlated with eligibility status across *veredas*.

Figure 3 assuages these concerns (and also previews the main result of this subsection). The Figure investigates dynamic patterns in plot upgrading around the time the *vereda* becomes eligible using the ITT specification. The Figure shows the absence of differential trends in plot upgrading in eligible *veredas* in the years preceding the roll-out of the Program. The Figure also shows a gradual increase in the upgrading index in the years after the Program is rolled-out to the *vereda*. The gradual improvement is consistent with take-up patterns in Figure A4.[24]

Results: Column 1 in Table 2 reports the OLS specification: after joining the Program, plots have a 0.18 standard deviation higher quality index. The corresponding ITT is 0.047 (Column 2). Both coefficients are statistically significant at conventional level.[25]

[24]The Figure displays relatively large standard errors four and five years before the vereda becomes eligible. This is due to the fact that the panel starts in 2006, only two (four) years before the largest waves of Program roll-out in Nariño and Cauca respectively. Figure A5 confirms the absence of pre-trends in the outcomes of interest focusing on the Cauca region only for which more years are available before the Program roll-out. Table A4 in the Appendix reports the corresponding coefficients.

[25]De Chaisemartin and D'Haultfeuille (2019) note that DID designs with period and group fixed effects identify weighted sums of average treatment effects (ATEs) in each group and period with weights that may be negative and propose a correction. The negative weights sum only

Heterogeneity and The Mechanics of Upgrading: The costs of upgrading the plot to the Program required standards are, all else equal, lower for plots that are already in good shape at the time the Program is rolled out in the *vereda*. Columns 3 and 4 in Table 2 split the sample between plots that at the time of eligibility were in the top-quartile of the plot quality index and those who were not. Results indeed confirm that the impact of the program was concentrated on the better plots at the time the vereda becomes eligible.

Table A5 in the Appendix provides a more comprehensive picture of the plot upgrading process. The Table considers a wider vector of outcomes and delves into the mechanics of quality upgrading by parting apart the vereda eligibility indicator (ITT) into *i)* plots already in the Program [*takers*]; *ii)* plots that will join the Program [*will-be takers*]; and *iii)* eligible plots that do not join the Program (by the end of the sample period) [*never takers*]. The results must be interpreted cautiously since the status of each plot at a given point in time is an endogenous choice. Taken together, the results in the Table paint an heterogeneous quality upgrading process in which certain plots must be upgraded *before* they (can) join the Program; other plots enhance their quality potential decreasing tree density; and other plots respond by intensifying coffee cultivation. Results also show some replanting among never takers, possibly reflecting access of these plots to the Program support for plot renewal (see notes to Table A5 for details).

Other Programs (Placebo): The Program implementer is in charge of other support programs, including extension services. A potential concern is thus that the Program might have been rolled out across veredas at the same time of other efforts that also facilitated upgrading. ITT specifications in Table A6 in the Appendix find no impact of the Program on the likelihood that farmers receive extension services; other technical assistance programs; the FNC credit facility; and an identification program that facilitated farmers access to other services offered by the implementing cooperative.

Land Expansion: If the Program increased returns from coffee cultivation, we expect farmers to substitute into coffee and away from other activities. Panel A of Table 3 explores the extent to which the Program induced expansion of land under coffee cultivation. On the *intensive* margin, ITT (OLS) specifications as in

to 0.16 and implementing the correction yields dynamic coefficients nearly identical to those in Table A4. An alternative spatial discontinuity design in Table B4 also confirms the Program's positive impact on plot upgrading.

equation (1) reveal an increase in coffee cultivated area of 2% (7%). A comparison of the OLS and ITT specifications suggests a moderate amount of positive selection with respect to the scope for expanding area under cultivation into the Program, confirmed by (unreported) heterogeneity analysis.

On the *extensive* margin we expect a negative impact on the likelihood that plots abandon coffee cultivation. Joining the Program is associated with a lower likelihood of abandoning coffee cultivation (6%). The ITT is also negative, but not statistically significant at conventional levels. We also expect a positive impact on the *entry* of new plots in eligible veredas. Panel B of Table 3 investigates plot entry and land expansion at the vereda level. These more aggregate specifications control for vereda and municipality-year fixed effects. The Program is associated with a process of land reallocation: 9% increase in aggregate land under coffee cultivation (Column 1), lower exit rates (Columns 2), entry of both new farmers (Column 3) and plots (Columns 4).

3.B. Did eligible farmers produce higher quality?

Eligible farmers invested in plot upgrading: did the quality of their produce increase? Unfortunately we do not observe coffee quality at the plot level. However, we do observe detailed quality tests at the mill gate, with detailed information on the buying point (origin) where the coffee batch was sourced from. At the mill gate, each batch of coffee is subjected to standard quality tests. For each batch of coffee we observe several physical characteristics, including the shares of rotten, low quality and defective beans; the *rendimiento* (the amount of parchment necessary to obtain one kg of green coffee) and the cupping test.

We begin by asking whether batches sourced for the Program are of higher quality than non-program batches sourced from the same buying point (i.e., origin) in the same season. We estimate

$$Q_{bomy} = \beta_0 + \beta_1 \times PB_{bomy} + \gamma_{oy} + \gamma_{my} + \varepsilon_{bomy} \tag{2}$$

where Q_{bomy} denotes a quality outcome of coffee batch b from buying point o in month m of year y and PB_{bomy} is a dummy taking value equal to one for Program batches. The specification includes buying point - season fixed effects (γ_{oy}) to control for time varying buying point characteristics, and month fixed effects (γ_{my}) to account for season and seasonal variation.

Panel A (I) of Table 4 reports the results for the different quality outcomes. Odd numbered Columns compare Program batches against all batches. Even numbered columns compare Program batches against batches sourced under other VSSs and initiatives that might also have quality requirements.

Program batches have 0.42 (0.12) standard deviations higher quality index than non-program batches sourced from the same origin in the same season when compared to all batches (non-standard batches) respectively (Columns 1 and 2). Columns 3 to 8 consider several quality dimensions (the share of healthy beans and defects such as low quality beans, pasilla, and beans affected by the coffee rust). Program batches have higher quality along *all* dimensions.[26]

A potential concern with the specifications in Equation (2) is that mill gate quality tests occur *after* the batch is accepted into the Program. It is thus possible that Program farmers and/or personnel at the buying points simply sort better beans into Program batches. In that case, Program batches have higher quality than non-program batches sourced from the same origin without the Program having actually increased quality. To investigate the relevance of this concern we consider an alternative specification that compares non-program batches across Program and non-program localities. If sorting drives the quality difference, non-program batches from Program origins will have *lower* quality. We estimate

$$Q^{NP}_{bomy} = \beta_0 + \beta_1 \times PO_{oy} + \gamma_o + \gamma_{my} + \varepsilon_{bomy} \tag{3}$$

where Q^{NP}_{bomy} denotes the quality measurement of non-program batches from origin o, and PO_{oy} is a dummy taking value equal to one after the buying point begins sourcing coffee for the Program. Panel B of Table 4 shows that non-program batches from Program origins do not have lower quality than batches from non-program locations. The quality differential in Panel A is thus not due to sorting: the Program increased the supply of quality coffee in the market.[27]

[26]The quality of the beans also depends on storage practices. Unfortunately we do not have direct evidence on those. Program farmers sell their produce earlier (unreported). Although we can't rule out alternative interpretations, this is consistent with farmers reducing quality problems due to poor storage conditions on the farm. The results in Panel A (I), however, are not explained away by the difference in the timing of sales. Panel A (II) of Table 4 shows that the results are robust when we control for month-year-origin fixed effects.

[27]Since quality data are available only for cooperatives buying points, a final concern is that farmers sell low quality beans to other buyers. If that was the case, we expect Program farmers to get lower prices from other buyers. Panel B of Table 6 presented in the next Section rejects this hypothesis. If anything, the Garantía de Compra implies that farmers have incentives to sell the lower quality to the cooperatives buying points.

3.C. Did farmers sell to the Program at the announced price premium?

Agricultural markets are often plagued by contracting problems such as hold-up and side-selling. Did the Program overcome those challenges? In particular, did the Program honor the promise to buy all coffee with suitable quality at the pre-announced price premium (no *hold-up*)? And did farmers increase deliveries to the Program implementer (no *side-selling*)?

i) No Hold-Up: Detailed sourcing records from one of the two implementing cooperatives reveal that the Program did not hold-up farmers. The transaction-level data cover the volumes and prices paid for deliveries from any farmer for the 2015 and 2016 harvest seasons. Figure 4 shows that, on average, over 80% (90%) of Program's farmers deliveries to the Cooperative (conditional on any delivery for the Program) are absorbed by the Program. This stands in contrast to other NGO-driven VSSs discussed in Section 5.

Table 5 estimates regressions of the form

$$P_{sfoy} = \beta_0 + \beta_1 \times PD_{sfoy} + \gamma_{oy} + \gamma_f + \varepsilon_{sfoy} \tag{4}$$

where P_{sfoy} denotes (the log of) the per kilo price paid by the cooperative to farmer f for coffee sold under line s delivered at buying point (origin) o in season y. The specifications include origin-year fixed effects (γ_{oy}). Coffee lines s can either be standard (i.e., receiving the FNC base price) or the various VSSs farmers can sell to, conditional on their certification status. The dummy PD_{sfoy} takes value equal to one for coffee sold under the Program.

Program sales receive 9.6% higher prices. The estimate is remarkably stable across specifications in Table 5: Column 1 doesn't control for farmer fixed effects thus identifying from across-farmers variation; Column 2 controls for farmer fixed effects thus identifying from farmers joining the Program during 2015 and 2016; Column 3 controls for farmer-season fixed effects, thus identifying from Program farmers that, in a given season, do not sell all their coffee to the Program.[28]

Farmers were able to sell coffee to the Program at the announced price premium. The Program price premium of 400 COP/Kg over the FNC base price was

[28] The share of deliveries sold to the program and the price premium received do not vary across cohorts of program farmers (unreported). The Program also pays higher prices than other VSSs (Column 4).

increased to 600 COL/Kg in 2015 and 2016. At that time, the average FNC base price was around 6000 COL/Kg. The announced Program price premium was thus $600/6000 \approx 10\%$, indistinguishable from the estimates in Table 5. Coincidentally, this figure is also similar to the average percentage price premium over the entire sample period (2006 to 2016). In Section 4 we thus calibrate the model using a farm gate price premium of 10%.

ii) No Side-Selling: To explore side-selling we need to observe Program's farmers sales to other buyers. We analyze unique transaction level data covering sales from *any* farmer to *any* buyer - cooperative and private - from the Protección del Ingreso Cafetero (PIC) program.[29]

The PIC data include information on the date, volume, price and buyer for all farmers' sales. Unlike the data from the implementing cooperative analyzed above, however, the PIC data do not distinguish whether a sale between a Program farmer and an implementing cooperative occurs under the Program or not (e.g., because it doesn't satisfy the quality requirements). Furthermore, the PIC data are available only for one harvest. We must thus rely on a cross-sectional identification to examine the impact of the program on the outcomes of interest.

Descriptive evidence in Figure 5 reveals that in Program veredas farmers deliver more coffee to the cooperatives than in non-program veredas. The difference, however, is entirely driven by Program farmers. We subject this evidence to formal testing implementing a *Spatial Discontinuity Design (SDD).* The main idea is to compare sales patterns of farmers that cultivate plots located near a border that separates an eligible and a non-eligible vereda (see Figure A8). Table B1 lends some support to the validity of the approach: there does not appear to be systematic differences between farmers with plots located within 1Km of a border separating a Program vereda from a non-program one.

Table 6 presents estimates from the following specification:

$$Y_{fvm} = \beta_0 + \beta_1 \times PF_{fvm} + \gamma_b + \varepsilon_{fvm} \tag{5}$$

where Y_{fvm} is the outcome of interest and PF_{fvm} denotes the Program indicator for farmer f in vereda v belonging to municipality m. We focus on ITT

[29]The PIC was implemented in 2013 to respond to exceptionally low income for farmers resulting from the combined effect of *i)* a drop in world's prices *and ii)* low volumes and quality of coffee harvested due to El Niño and La Niña (see Figures A6 and A7 for details).

specifications in which P_{fvm} takes value equal to one for farmers with plots in eligible veredas, but also report OLS estimates in which P_{fvm} takes value equal to one for farmers that have joined the Program. The baseline specification restricts the sample to farmers with plots within 1Km from the border and includes border fixed effect γ_b to control for confounders. Spillovers are both a potential source of concern for the design but also of intrinsic interest and are discussed below.

Panel A of Table 6 reports results on farmers selling behaviour. On the extensive margin, the OLS (ITT) estimates in Column 1 (2) reveal a 20% (5%) higher likelihood of selling to the Program implementer. While Program farmers are less likely to sell to other buyers (OLS, Column 3), the Program didn't affect farmers likelihood to sell to other buyers (ITT, Column 4). On the extensive margin, OLS (ITT) estimates in Column 5 (6) show that program farmers sell a 19% (4.5%) higher share of their produce to the program implementer. We thus find no evidence of side-selling behaviour from Program farmers.

3.D. Other Outcomes: Prices, Production, Upgrading

We exploit the SDD to explore the Program impact on prices, production and plot upgrading. This analysis informs the model in Section 4.

No price spillover: Panel B of Table 6 considers prices. Results reveal that farmers in Program veredas received higher prices. The higher prices, both in the OLS and ITT specifications, are entirely due to sales to the Program implementer. In particular, the Program did not impact prices received from other buyers (Columns 5 and 6).[30]

The result that Program farmers receive the same price from other buyers as non-program farmers supports other aspects of our analysis. It is consistent with the quality results in Table 4 (if Program farmers sell lower quality beans to other buyers they should receive lower prices from those buyers). It is also consistent with Program farmers not side-selling (if that was the case, Program farmers sales to other buyers should receive higher prices).

The inclusion of border fixed effects exposes our strategy to spillover across

[30]The PIC data are available for a period in which quality was exceptionally low due to adverse weather. The Program kept quality requirements constant. Under the assumption (consistent with Table 5 and with our conversations in the field) that the Program paid the announced price premium, the OLS estimate in Column 3 and the price premium imply that around 40% of the available production met the Program's quality requirements.

farmers that are located close to each other. Table B2 reports results from specifications that omit border fixed effects and obtain similar results.[31] Opening up the ITT along the lines of Table A5, Table B3 shows that *never-taker* farmers do not receive lower prices from either the Program implementer or from other buyers.

Despite its size, the evidence suggests that the Program did not impact market prices at the farm gate. A possible explanation for this result is that the *Garantía de Compra* gives farmers the option to sell at a price that tracks the world price. This implies that farmers face an horizontal (residual) demand curve that doesn't depend on local demand and supply conditions. These considerations guide important modeling assumptions in Section 4: *i*) we will abstract from other buyers of quality coffee and, therefore, from side-selling considerations; *ii*) we will capture the horizontal (residual) demand curve modeling the market for standard coffee as being perfectly competitive.

Production, Upgrading and Expansion: Table B4 applies the SDD to investigate other outcomes. The Table reveals an ITT estimate of 13% on production, measured from the PIC sales data. Combined with the estimates in Column 3, the ITT estimate is used to calibrate the model in Section 4 (see below for details). The SDD also confirms the results in Table 2 on the program positive impact on plot upgrading and expansion (Columns 4 and 5).[32]

4. The Sustainable Quality Program: A Quantitative Exploration

This Section provides a quantitative exploration of the impact of the Program on the Colombian coffee chain. We *A)* describe the model and motivate its assumptions; *B)* explain the calibration strategy; and *C)* present the results.

[31] The Table also shows that the results are robust to the inclusion of additional controls; to alternative geo-location of farmers with multiple plots; and to the exclusion of administrative borders that separate municipalities. Results are also robust to the inclusion of farmers at varying distances from the border (unreported).

[32] *Never taker* farmers do not have lower production nor different selling behaviour (Table B3). Distance from the border does not correlate with the share sold to the Program implementer for Program farmers (unreported). These results confirm field conversations suggesting that non-program farmers were not able to sell to the Program through Program farmers.

4.A. Model

The Standard Coffee Market: There is a unit mass of farmers i with coffee plantations of size L_i and a vector of characteristics Z_i. Potential production $Q_i = Q(L_i, Z_i)$ is distributed according to a cumulative distribution $\Omega()$. Farmers incur constant unit harvesting and processing cost c^S.

Standard coffee is exported at the exogenous world price p^W. A perfectly competitive market intermediates standard coffee between the farm gate and the export gate. Intermediation, transport and processing costs denoted by τ yield the farm gate price for standard coffee $p^S = (1-\tau)p^W$, with $c^S < p^S$.

An Overview of The Sustainable Quality Program: The Sustainable Quality Program is a bundle of contractual arrangements involving all actors along the coffee chain. We model the Program following the description in Section 2. We model *FNC* as a single, vertically integrated, exporter that also implements the Program (i.e., we bundle Almacafé and the cooperatives together). Table A7 shows that there is no double-marginalization between Almacafé and the implementing cooperatives for the Program coffee thus supporting the assumption.[33] We assume that the *Buyer* and *FNC* are the only buyer and exporter of quality coffee.[34]

We represent the Program as a vector of attributes at the farm and export gates. At the farm gate, the Program provides access to training, extension services and inputs for plot renovation. These interventions increase plot production by a factor ω. The Program also pays a farm gate price premium π for quality coffee and, as explained below, provides farmers with additional demand-side benefits that are not explicitly modeled, denoted α. The Program cannot discriminate and so ω, π and α are constant across farmers that join the Program.

At the export gate, the Program involves contractual arrangements between the *Buyer* and *FNC*. These contractual arrangements involve a price premium η at the export gate and result in a *Program Objective Function* that gives weight

[33]The Table estimates the price premium paid by Almacafé to the cooperatives using batch-level specifications as in Equation (2). Almacafé pays 4-5% price premium for Program coffee. At the corresponding mill gate price level this translates into a $\approx$ 450 Col/Kg premium, nearly identical to the Program farm gate premium.

[34]While the assumption is a rather extreme simplification of the market, we feel that from a quantitative point of view it is unlikely to significantly distort our analysis. Besides failing to detect any side-selling behaviour, two further pieces of evidence support our approach. First, at baseline, i.e., before the Program roll-out, less than 10% of production was of *supremo* quality. Second, the figures at the export gate and the results below suggest that, in the regions in which it was rolled out, the Program created its own supply of quality coffee.

λ to the exporter profit and $(1-\lambda)$ to farmers welfare.

In sum, the endogenous variables are the Program price premium π and farmers responses to Program eligibility (described momentarily). The Program is a vector of parameters that, directly and indirectly through farmers responses, shift the *supply* (ω and α) and the *demand* (η and λ) of quality coffee at the farm gate.

Farmers' Choice Set: Figure 2 illustrates the Program from the farmer's point of view. Let D_i denote farmer's i response to Program's eligibility. Farmers decide whether to upgrade the plot to the required standards and join the Program ($P_i = 1$) or not ($P_i = 0$). They also decide whether and, if so, how to invest in their plot: $I_i \in \{N, R, RE\}$. Farmers can either do nothing (N); renew the plantation (R); or renew *and* expand the area under coffee cultivation (RE). Farmers thus chose among *six* alternatives, $D_i = I_i \times P_i$:

$$D_i = \begin{cases} N0 & \text{Do nothing \& } \textit{not join Program} \\ R0 & \text{Renew \& } \textit{not join Program} \\ ER0 & \text{Renew and expand \& } \textit{not join Program} \\ N1 & \text{Do nothing \& } \textit{join Program} \\ R1 & \text{Renew \& } \textit{join Program} \\ ER1 & \text{Renew, expand \& } \textit{join Program} \end{cases}$$

Costs and Benefits: Farmers compare the costs and benefits of each alternative. Each decision D_i entails farmer specific fixed costs F_i^D drawn from a cumulative distribution $\Phi^D()$. We normalize $F_i^{N0} = 0$.

On the benefit's side, decision $D_i = D$ results in production $Q_i^D = (1+\omega^D) \times Q_i$. We denote $\vec{\omega}$ the vector of production increases associated with each decision and normalize $\omega^{N0} = 0$. To calibrate the model, we will assume $\omega^{R1} = \omega^{N1} + \omega^{R0}$ and $\omega^{ER1} = \omega^{N1} + \omega^{ER0}$. The parameter $\omega = \omega^{N1}$ will then denote the direct impact of the Program on plot production. As further discussed below, empirical estimates of ω provide an upper bound to the impact of the Program supply-side components (training, extension services and inputs for plot renovation) on production.

If the farmer doesn't join the Program ($P_i = 0$) she sells her production of standard coffee at farm gate price p^S incurring unit cost c^S. If the farmer joins the Program ($P_i = 1$) she has the option (but not the obligation) to sell high quality coffee at price $p^Q = (1+\pi) \times p^S$, with π being the Program price premium.

This static and deterministic framework does not capture other demand-side benefits that farmers might derive from the Program. For example, farmers might value the option to be able to sell to the Program. The Program also provides a degree of price insurance, in addition to the Guarantía de Compra available to all farmers.[35] The value of these benefits depends on the curvature of the farmers' utility function, their expectations and discount factor, which are not observable. We thus model these valuable program attributes as a positive *wedge* α in the utility that farmers derive from joining the Program, regardless of their sales decision. In estimating α and interpreting counterfactuals it will be important to bear in mind that it is a reduced-form representation of demand-side benefits that are not explicitly modeled.

Farmers that join the Program must also be given incentives to produce high quality coffee and sell it to the Program. Producing quality coffee requires higher harvesting and processing costs, $c^Q = (1+\gamma) \times c^S$. Denote with $\mu^C = p^C - c^C$ the margin for coffee $C \in \{S, Q\}$. Farmers produce high quality coffee and sell it to the Program if $\mu^Q \geq \mu^S$ (given evidence discussed in Section 3, we ignore side-selling constraints). Only farmers that join the Program produce quality coffee.

Farmers' Decision: Let $\mathbf{I_P}$ be an indicator function taking value equal to one if the farmer joins the Program. Farmer i payoff from decision D is given by:

$$W_i^D = (1 + \mathbf{I_P}\alpha)[(1 + \mathbf{I_P}\pi)p^S - (1 + \mathbf{I_P}\gamma)c^S)] \times (1 + \omega^D)Q_i - F_i^D \qquad (6)$$

Farmers take the decision that maximizes their payoff, $\mathbf{D}_i \in \text{argmax}\, W_i^D$. Farmers decisions are thus function of the farmer's potential production Q_i, prices and costs (p^S, c^S and γ), the Program's attributes (π, α and the vector ω) and the vector of fixed costs F_i^D.

Coffee Supply: The aggregate coffee supply function is obtained aggregating farmers' individual decisions. The aggregation over farmers i is both with respect to potential quantity Q_i and fixed costs F_i. The supply of quality coffee, $\mathbf{Q^Q}$, is given by aggregating over farmers that join the Program ($\mathbf{I_P} = 1$):

[35]To see why, note that the base price in the Guarantía de Compra p^S tracks the world price p^W and is thus random. Denote $\bar{p}^S$ its expected value. The program pays a fixed amount $\mathbf{p} = \pi \times \bar{p}^S$ on top of the basis price p^S. All else equal, this fixed premium is more valuable to a risk-averse farmer when p^S is low. The Program might also mitigate weather-related risk by helping farmers planting disease resistant-varieties. Given the perfectly elastic demand, this insurance might also be valuable.

$$\mathbf{Q}^Q = \int_Q \int_F \mathbf{I_P} \times \mathbf{I_D} Q_i^{\mathbf{D}_i} d\Omega() d\Phi() \tag{7}$$

Similarly, the supply of standard coffee, $\mathbf{Q}^S$, is given by aggregating over farmers that do not join the program replacing $\mathbf{I_P}$ with $(1 - \mathbf{I_P})$ in Equation 7. When farm gate demand conditions change, the supply of quality coffee responds through farmers decisions on the extensive margin (i.e., join, upgrade and expand).[36]

Program's Objective Function: The *FNC* exports quality coffee to the *Buyer* at a price $(1+\eta) \times p^W$ incurring the same unit transport and intermediation costs $\tau \times p^W$ as standard coffee. If *FNC* exports standard coffee, it makes zero profits on those sales. The *FNC* profits are then given by

$$\mathbf{\Pi}^{FNC} = \mathbf{Q}^Q \times ((1+\eta)p^W - (1+\pi)p^S - \tau p^W) \tag{8}$$

The Program involves a three-layered chain between the Farmers, *FNC* and the *Buyer*. There is thus potential for double marginalization in which *FNC* sets an inefficiently low price premium π at the farm gate from the point of view of the chain. Besides the export gate premium η, the non-linear contract between *FNC* and the *Buyer* includes provisions on the farm gate price premium π and a lump-sum contribution and thus reduces double marginalization. The two parties agree on a π higher than what would be set by a profit-maximizing *FNC* that takes the export-gate price as given. We thus represent the program's objective function as setting the farm gate price premium to maximize a weighted average of the *FNC* profits and farmers' welfare.[37] We thus have:

$$\pi \in \arg\max \lambda \mathbf{\Pi}^{FNC} + (1-\lambda) \int_Q \int_F W_i^{\mathbf{D_i}} d\Omega() d\Phi(). \tag{9}$$

[36]In reality, farmers might respond to changes in demand conditions on the intensive margin as well by adjusting complementary inputs (e.g., effort) not provided by the Program. Conditional on plot characteristics, scope for such adjustments are relatively small. Nevertheless, due to this possibility, our preferred interpretation is that empirical estimates of the parameter ω provide an upper bound to the impact of the Program supply-side components (training, extension services and inputs for plot renovation) on production.

[37]The logic is similar to resale price maintenance (see, e.g., Tirole (1988) and Rey and Tirole (1986)) with farmers (buyer) and consumers (manufacturer) swapping roles. The reduced-form representation avoids modeling the *Buyer* margins and the bargaining protocol and is consistent with the possibility that *FNC* cares about the farmers' welfare directly but doesn't have adequate instruments to redistribute profits to farmers as desired through an appropriate set of transfers.

4.B. Calibration, Identification and Assumptions

We calibrate the *demand* and *supply* side of both the standard and quality coffee markets. The parameters for the standard coffee market (p^W, p^S and thus τ) are observed in the data. The distribution of land size L_i and plot characteristics Z_i (and thus Q_i) are also observed in the data. Costs parameters, c^S and γ, are obtained from detailed FNC agronomists costs sheets. The vector $\vec{\omega}$ is estimated based on spatial discontinuity specifications as those in Section 3. The farm gate price premium π was estimated in Section 3. We estimate the export-gate premium η from transaction-level export data. The parameters α and the fixed costs F_i^D are estimated from farmers' responses to Program's eligibility. Finally, given knowledge of demand and supply functions, the Program's objective function parameter λ is estimated as the welfare weight that rationalizes the observed farm gate price premium π. Table 8 summarizes the model parameters. We explore robustness of our conclusions to parameters that are not estimated (c^S, γ, τ, p^W).

Standard Coffee Market: The model assumes a perfectly competitive market for standard coffee. The existence of the *Garantía de compra* and further evidence lend some support for this assumption. The *Garantía de compra* implies that the farm gate price p^S perfectly tracks the exogenous world price p^W.[38] The assumption of a competitive market is also supported by the lack of a Program impact on farm gate prices for non-program coffee (see Table 6). We set $p^W = 3.75\$$ at the average world price for Colombian milds and $p^S = 2.68\$$ at the average FNC base price over the sample period. This yields transport and intermediation cost $\tau = 1.4$.[39]

Price Premium at the Export Gate (η): We estimate a Program's price premium at the export gate $\eta \approx 19\%$. Table 7 reports results using detailed data on coffee export transactions. The data allow to trace coffee batches to the export gate. Column 1 shows that the foreign *Buyer* pays a price premium of around 19% at the export gate relative to standard quality coffee.[40]

[38]The assumption of an exogenous world price is justified by Colombia accounting for less than 9% of the world's production and our analysis focusing on two regions that account for less than a third of Colombia's exports.

[39]Without affecting any of the results we abstract from small differences in intermediation costs across buying points and from Program's roll out across different cohorts.

[40]Column 2 finds a nearly identical premium restricting the comparison to coffee sourced from the Program regions. This suggests that the regional markets for standard coffee are integrated further corroborating the assumption that the Program did not have general equilibrium effects.

The Program's price premium at the export gate reflects both the market premium for quality coffee and the *Buyer*'s willingness to pay for other attributes, e.g., supply assurance and traceability. Columns 3 to 5 assess the relative importance of these components. The Program exclusively sources *supremo* coffee beans (see Table A1). Restricting the sample to exports of supremo coffee (Column 3) and including further quality controls (Column 4) reveals that the *Buyer* pays 9% more than the market premium for the corresponding quality. The market premium for coffee of the quality sourced by the program is thus around 10%. Column 5 confirms this estimate showing that *supremo* coffee (the quality imported by the Program) commanded a 10% price premium over standard quality on the sample of coffee batches exported *before* the roll-out of the Program.

Price Premium at the farm gate (π): The Program price premium at the farm gate is estimated in Table 5 at $\pi = 10\%$.

The Program Welfare Weight (λ): Given knowledge of η, τ, p^W and the supply parameters derived below, λ is identified by "inverting" the Program maximization problem in Equation (9). The parameter λ is the welfare weight that rationalizes the observed farm gate price premium $\pi = 10\%$.

Farmers Decisions $\mathbf{D}_i$: The supply of standard and quality coffee is given by the aggregation of farmers decisions $\mathbf{D}_i$. Farmers decisions $\mathbf{D}_i$ are defined in the three years after the vereda becomes eligible for the Program. A farmer is considered to take up ($P_i = 1$) if she has joined the Program within three years of becoming eligible. Similarly, a farmer renews (expands) if she undertakes any upgrading work on (expands the area under coffee cultivation in) the plot within three years of becoming eligible. With these definitions, 39% of farmers take-up the Program ($P_i = 1$). Conditional on taking-up, 41% (21%) of farmers renew (renew and expand) their plots. Conditional on not taking up, 18% (12%) of farmers renew (renew and expand).[41]

Farmers Potential Production Q_i: We parametrize production to be a function of plot size L_i (in hectares) and average plantation's age A_i (in years): $Q_i = \theta(A_i) \times L_i^{\chi}$. We estimate a regression of log production on log area under coffee

[41] Joining the Program is a discrete choice. Although renewal and expansion are in principle continuous choices the data reveal that many farmers did not renew nor expand thus justifying our discrete choice framework. To expand land farmers must plant new/younger trees. No farmer expands land under coffee cultivation without also renewing.

cultivation and dummies for the plantation's age combining the PIC sales data and the plot census. The estimation yields an inverted-U shaped function $\theta()$ with highest production at around 6 years of age and nearly constant returns to scale $\chi = 0.9$. *FNC* agronomists confirmed these estimates to be reasonable.[42] We use estimates of $\theta()$ and χ to compute potential production Q_i for each eligible plot at the time the Program was rolled out in the vereda.

Production (ω vector): We estimate the vector $\vec{\omega}$ building upon ITT estimates from the spatial discontinuity design in Table B4. Column 2 reveals an ITT impact of the Program on production equal to 13%. The ITT is a weighted average of *i)* the ωs associated with each investment decision and with joining the Program; *ii*) the Program shifting the share of farmers investing. The OLS specification in Column 6 of Table B4 shows that, among takers, farmers renewing (resp. expanding) have 12% (resp. 37%) higher production relative to farmers that simply join the program. The assumption $\omega^{R1} = \omega^{N1} + \omega^{R0}$ (resp. $\omega^{ER1} = \omega^{N1} + \omega^{ER0}$) then allows us to recover the direct impact of the Program on production $\omega = \omega^{N1}$ from the ITT estimate using the observed share of farmers' decisions as weights. Estimates yield the program $\omega = \omega^{N1} = 25\%$, a figure also consistent with reports from *FNC* agronomists.[43] As discussed above, we interpret $\omega = 25\%$ as an upper bound to the impact of the Program supply-side components (training, extension services and inputs for plot renovation) on production.

Variable Costs (c^S and γ): Our data do not contain information on variable costs. We thus asked *FNC* agronomists who shared detailed cost sheets. Over the sample period, the average (deflated) harvesting and processing unit costs for standard coffee were estimated to be $c^S = 0.68\$$ per Kg.

Producing quality coffee eligible for the Program entails higher harvesting and processing cost. The agronomist estimated these additional costs to be $\gamma = 16.7\%$ of the unit costs to produce standard quality. In the baseline scenario thus the Program increases farmers' margin by $((p^Q - c^Q)/(p^S - c^S) - 1 \approx 8\%)$. We perform robustness checks to alternative costs scenarios.[44]

[42]The regression is estimated on farmers *not* in eligible veredas since the PIC data is available towards the end of the sample period, i.e., when eligible farmers's production has already been altered by the Program.

[43]The full vector $\vec{\omega}$ is given by $\omega^{R0} = 12\%$, $\omega^{ER0} = 38\%$ for non-takers, and $\omega^{R1} = 37\%$, $\omega^{ER1} = 63\%$ for takers.

[44]Variable costs could be backed-out imposing more structure on the model. We refrain from doing so hoping that acknowledging the data limitations and relying on information from

Fixed Costs (F_i): We also do not have data on the costs for renewing, expanding and joining the Program.[45] Even if we did, we would still estimate fixed costs F_i^D since they include opportunity costs typically not covered in surveys. The fixed costs of decisions *not* taken are also needed to perform counterfactuals.

We estimate the fixed costs F_i^D from farmer's decisions. A descriptive analysis of take-up patterns guides the parametrization of fixed costs. Figure A9 shows an inverted-U relationship between plot's size and take-up which depends on the age of the plantation and substantial heterogeneity in farmers' take-up across municipalities (see Table A8, Column 1, for the corresponding estimates). These facts motivate a specification

$$F_i^D = \Phi^D(L_i, A_i, X_i, \delta_t, \delta_m) + \epsilon_i^D \tag{10}$$

in which fixed costs F_i^D are a decision-specific function $\Phi^D()$ of farmer's plot size L_i, plantation's age A_i, farmer's characteristics X_i, cohort δ_t and municipality δ_m fixed effects (see Table A8 for further details).

The random term ϵ_i^D is i.i.d. across farmers and decisions and type-I GEV distributed with scale parameter σ. Under this assumption, the parameters of the farmer's discrete choice can be estimated from a multinomial logit model in which the farmer's decision $\mathbf{D}_i$ is regressed on the fixed costs in Equation (10) and on farmer's potential production, Q_i (see Equation (6)).

Identification of σ and α: The model's structure and functional form assumptions identify σ and α. With regard to functional forms, the multinomial logit is a standard model to estimate discrete choice problems. With regard to the model structure, the key assumption is that, conditional on observables, *non-takers* draw fixed costs of investment and upgrading from the same distribution as *takers*. The intuition behind the identification strategy is then to exploit knowledge of the returns from the investment decisions of *non-takers* to identify the scale parameter

agronomists yields a more transparent and reliable analysis. This choice comes at the cost of assuming identical variable costs across plots. While this is restrictive, production volumes Q_i and estimated fixed costs F_i^D allow for plot heterogeneity. We refrain from statements about distributional impact of the Program *across* farmers which would be sensitive to misspecification of the correlation between variable costs, fixed costs and plot characteristics across plots. We focus instead on the average impact of the program *along* stages of the chain. Our conclusions are then robust to wide ranges of assumed c^S and γ, as detailed below.

[45] Note that, unlike other common NGO-driven certifications, there are no fees charged to farmers to join the Program. The fixed costs of joining are the (opportunity cost of) resources invested to upgrade the plot and meet the Program standards.

of the fixed costs distribution σ. Given σ, the value of α is identified as the one that rationalizes *takers*'s investment decisions in excess of what would be implied solely by the program's higher production ω and margins μ^Q.[46]

Denote with β^D the estimated coefficient for Q_i in decision D in the multinomial logit. For *non-takers*, i.e., for $D \in \{R0, RE0\}$, we have $\beta^D = \frac{\mu^S \times \omega^D}{\sigma}$. Knowledge of $\mu^S = p^S - c^S$ and ω^D thus identifies σ. Given σ, the value of α is identified from *takers* decisions. For $D \in \{N1, R1, RE1\}$ we have $\beta^D = \frac{\mu^Q(1+\alpha)(1+\omega^D)-\mu^S}{\sigma}$. This gives $1+\alpha = \frac{\beta^D \times \sigma + \mu^S}{\mu^Q \times\ (1+\omega^D)}$.[47]

4.C. Results

Table 9 reports the results. Column 1 reports the baseline estimates. We estimate $\lambda = 0.695$. This represents a fairly sizable deviation from the benchmark in which the farm gate price premium is set to maximize exporter's profits. While the Program reduced double marginalization, it did not completely eliminate it since the Program price premium at the export gate is larger than the price premium at the farm gate, $\eta \times p^W > \pi \times p^S$.[48]

We also estimate $\alpha = 21\%$, suggesting that the Program gives farmers substantial demand-side benefits beyond the price premium π. As a benchmark, these benefits are similar in magnitude to the increase in production ω and larger than the increase in margins $(\mu^Q/\mu^S) - 1 \approx 8\%$ from the Program.

The estimates imply that the Program had a sizable impact on both quality upgrading and farmers welfare. Nearly 40% of farmers take-up the Program and, since farmers joining the Program are larger and expand more, 59% of the aggregate production is estimated to be high quality. This estimate is in line with the increase in the share of coffee exported as *supremo* (from 9% to 17%) mentioned in the introduction. To see why, note that by 2012 the Program had expanded to around half of the veredas in the Cauca and Nariño regions, two regions that account for around a third of Colombia's coffee exports. The estimates therefore

[46]The investment decision of non-takers identifies σ and is thus critical to identify α. Unreported results show that a simpler model based on the binary take-up decision yields quantitatively similar results. The simpler model doesn't separately identify σ and α and thus precludes some of the counterfactual analysis.

[47]The model is overidentified. The footnotes to Tables 8 and 9 discuss further the economic intuition behind our identification strategy and the model estimation.

[48]In fact, the farm gate price premium is smaller than the market premium at the export gate, $10\% \times p^W > \pi \times p^S$. This suggests the existence of constraints, possibly due to limited contract enforcement, to the lump-sum transfer between the multinational buyer and FNC.

imply an increase in the share of Colombia coffee exported as *supremo* equal to $(59\% - 9\%) \times 1/2 \times 1/3 \approx 8\%$.

Conditional on take-up, the Program increased farmers profits by 17% (without, by assumption, reducing the profits of non-takers). Given the value of demand-side benefits α, this yields a 19% welfare increase for eligible farmers. Higher quality translates into higher prices and revenues at the export gate. Once the higher variable and fixed cost to produce quality coffee are net out, the Program increased the surplus generated by the chain in Colombia (*FNC* plus farmers) by 33%. Farmers are estimated to keep at least 56% of this increase in surplus.[49]

Robustness: Table A9 explores the robustness of these estimates to alternative baseline parameters. We consider higher increases in the additional variable costs needed to produce quality (from $\gamma = 16.7\%$ at baseline to $\gamma = 22.5\%$); higher unit costs for standard coffee (from $c^S = 0.68$ to $c^S = 1.02$); higher intermediation costs for Program coffee (from $\tau = 1.4$ to $\tau = 1.6$); and alternative scenarios of world prices p^W (from 3.25 to 4.25, with baseline of 3.75).

The results are robust to these changes to the baseline scenario. Estimated λ ranges from 0.685 to 0.765, α from 21% to 24%. Estimated increases in takers profits range from 15% to 18%; the increase in the chain surplus from 32% to 37% and the share of that increase accruing to farmers from 52% to 57%.

The results are robust to alternative scenarios because the variables that drive the outcomes of interest, namely take-up and price premia along the chain, are all anchored to the data. First, the model is estimated targeting, and thus almost exactly replicates, the observed vector of farmers' investment and take-up decisions (Panel B, Table 9). Changes in baseline parameters are thus compensated by changes in estimated fixed costs so that actual take-up and quality upgrading match those observed in the data.

Alongside take-up and volumes, price premia at the farm and export gate drive the increase in surplus created by the Program and how it is shared along the chain. These price premia are also anchored to the data, and estimated to be $\pi = 10\%$ and $\eta = 19\%$. This explains why the main results are robust to alternative scenarios for costs (c^S, γ and τ) and world prices p^W.

[49] This figure might be a lower bound. Farmers profits might be *under*stated as the estimated fixed costs might include wedges arising from constraints to upgrade that are not costs incurred by farmers. We also do not account for the costs of implementing the Program, and thus *over*state *FNC*'s profits. On the other hand, a share of these costs was covered by the lump-sum payment from the multinational buyer. We do not have details of the contractual terms.

5. Mechanisms: How did the program work?

The reduced form results in Section 3 show a positive impact of the Program on quality upgrading and farmers sales. The calibration in Section 4 suggests that the gains were sizable and quite equitably shared along the chain. So, how did the Program work?

The Program is a bundle of contractual arrangements involving all actors along the chain. To answer the question we would ideally want to unbundle the different components of the program. Unfortunately, such unbundling is not possible.[50] We thus attempt to make progress towards understanding how the Program works in two indirect ways. First, we explore counterfactual scenarios based on the model estimates in which we alter the different components of the Program. This approach yields numerous insights at the cost of making many assumptions. Second, we complement the counterfactual exercise with a comparison of price (and quality) premia along the chain between the Sustainable Quality Program and two other VSSs that share only certain aspects of the bundle.

5.A. Unbundling the Program: Counterfactual Analysis

The Program changed the supply and demand for quality coffee at the farm gate. The two endogenous variables are the share of quality coffee and the farm gate price premium π. The Program shifted the *supply* curve out by drawing farmers to upgrade quality through its impact on production (ω) and demand-side (α) benefits. On the *demand* side, the Program alters the price premium for quality at the export gate (η) and its transmission to the farm gate (λ). We are interested in three sets of counterfactuals: on the *supply* side, on the *demand* side, and on *market structure*. Results are in Columns 2 to 10 in Panel A of Table 9.

Supply-Side Counterfactuals: *Supply*-side counterfactuals in Columns 2 to 4 set $\omega = 0$ and $\alpha = 0$ (first one at a time, then simultaneously) holding demand parameters η and λ constant. Setting $\omega = 0$ (from $\omega = 25\%$) has a sizable effect on the Program's overall impact: quality upgrading and the increase in chain surplus would both drop significantly (from 59% to 35% and from 33%

[50] In principle, it might be possible to evaluate (ideally through a set of RCTs) the various components of the Program at the farm gate: Program's attributes impacting production (ω), the price premium (π) and other demand-side benefits (α). We are however also interested in understanding the role played by the contractual arrangements between the exporter and the foreign buyer (η and λ). Those would be harder to experiment on.

to 15% respectively). The exporter would need to increase the farm gate price premium to 14% to source coffee. Farmers would get a larger share (from 56% to 61%) of a smaller pie and their welfare increase would be smaller (from 19% to 9%). Demand-side benefits to farmers $\alpha = 21\%$ are estimated to be of a similar magnitude as ω. Setting $\alpha = 0$ yields quantitatively similar results as setting $\omega = 0$. Removing *both* components simultaneously results in even lower quality upgrading (17%) and increase in chain surplus (6%).

Demand-Side Counterfactuals: *Demand*-side counterfactuals in Columns 5 to 7 set $\lambda = 1$ and $\eta = 10\%$ (first one at a time, then simultaneously) holding supply parameters ω and α constant. Replacing the Program contractual arrangements with a profit maximizing exporter ($\lambda = 1$) or with a buyer paying the market quality premium ($\eta = 10\%$) results in a lower farm gate price premium $\pi = 4.5\%$. This is the lowest price premium that must be paid to induce farmers to produce quality.[51] Given the common minimum farm gate price premium, both scenarios yield lower quality upgrading and increase in farmers welfare (from 59% to 49% and from 19% to 13% respectively). The two scenario, however, differ in how much surplus is created and how it is shared along the chain. A profit maximizing exporter would reduce both the increase in surplus along the chain (to 28%) and the share of that increase accruing to farmers (to 46%). An export gate price premium at the market level would reduce even more the increase in chain surplus (to 19%) but would leave a higher share of it to farmers (66%).

Market Structure Counterfactuals: We compare the Program against a monopsonist benchmark that features none of the Program's attributes. The monopsonist benchmark pays the market quality premium at the export gate ($\eta = 10\%$ instead of 19%); maximizes profits ($\lambda = 1$ instead of 0.695); does not provide inputs, extension and training ($\omega = 0$ instead of 25%) nor other demand benefits ($\alpha = 0$ instead of 21%) to farmers.

Figure 6 illustrates the comparison at the farm gate. The Figure reports the share of quality coffee (on the x-axis) and the farm gate price premium π (on the y-axis) under different scenarios. Relative to the monopsonist benchmark, the Program shifts out both the supply curve and the demand curve. The monopsonist benchmark yields limited quality upgrading (9%, an estimate nearly identical to

[51] The supply of quality coffee is zero unless farmers earn higher margins on quality relative to standard coffee ($\pi \geq c \times \gamma / p^S \approx 4.5\%$).

the actual share of *supremo* coffee exported before the Program roll-out), limited welfare gains for farmers (2%) and increase in the chain surplus (3%).

Figure 6 and the results of the *Supply-Side* and *Demand-Side* counterfactuals might give the impression that the supply side attributes played the lyon's share in the Program's success. We would like to put a caveat to this interpretation. First, α captures demand-side benefits that farmers derive from the Program. Although α shifts the *supply* of quality coffee at the farm gate, α would be substantially lower in the absence of *demand*-side contractual arrangements between the foreign buyer and the FNC. For example, without a firm demand commitment from the foreign buyer, the exporter might not honour the promise to buy all eligible supply at a fixed farm gate premium. Comparing Columns 2 and 8 reveals that simply removing ω accounts for $(59\% - 35\%)/(59\% - 9\%) \approx 48\%$ of the reduction in quality associated with the monopsonist benchmark. As noted above, ω provides an upper bound to the impact of the Program supply-side components (training, extension services and inputs for plot renovation). The counterfactual analysis thus suggests that the Program's *supply-side* components accounted for less than half of the gains from the Program with contractual arrangements on the *demand-side* accounting for the remaining share.

Table 9 also compares the monopsonist benchmark (Column 8) to an scenario in which contracts with farmers are enforceable (Column 9) and to a competitive market for quality coffee (Column 10, also illustrated in Figure 6). Enforcing contracts with farmers removes the constraint $\pi \geq c \times \gamma/p^S$. While the constraint is not binding under the Program's configuration, it is binding once either η and/or λ are set to the monopsonistic benchmark. Removing the constraint results in even lower farm gate prices, quality upgrading and surplus in the chain. In our model, the lack of contract enforcement disciplines the monopsonist attempt to extract rents from farmers. This is an illustration of the well-known idea that, in the presence of multiple distortions (lack of contract enforcement and monopsony power), removing one distortion doesn't necessarily improve efficiency.

A perfectly competitive market that pays the market quality premium at the export gate $\eta = 10\%$ would translate into a higher farm gate price premium $\pi \approx 14\%$. Relative to the monopsonist benchmark, a competitive market would generate more quality upgrading (from 9% to 15%), a larger increase in chain surplus (from 3% to 4%) and, by definition, would give 100% of those gains to farmers. The estimates however suggest that the gains from a competitive market structure would be small relative to those generated by the Program's bundle.

5.B. Price Premia Along the Chain: A Comparison Across Programs

We complement the counterfactual analysis by comparing the Program price premia along the chain with two other (non *buyer-driven*) VSSs: an *environmental* certification and a *social* certification (henceforth *EC* and *SC* respectively).[52] *EC* is an individual plot environmental VSS which requires very similar environmental standards and shares extension services with the Sustainable Quality Program. *SC* is a group-level VSS that pays a price premium and a guaranteed minimum price for certified coffee. Part of the price premium is then earnmarked to social programs that benefit farmers.

In Colombia, FNC implements both *EC* and *SC* on behalf of international NGOs. Recall also that *all* farmers in Program veredas are eligible to receive training, extension and support for plot renewal from the Sustainable Quality Program, regardless of their take-up decision. Although the two VSSs differ from the Sustainable Quality Program in several ways, the main difference is the Program contractual arrangements between FNC and the multinational buyer.

Table 10 reports the results. The comparison is descriptive and results have to be interpreted cautiously. Neverthless, the Table reveals notable differences in the structure of price premia between the two NGO-driven VSSs and the Program.

At the farm gate, Program sales pay a 10% farm gate premium relative to standard coffee (Column 1, see also Table 5). Farmers that hold *EC* receive a 1% premium on their sales.[53] This is a notable difference given that *EC* shares many aspects of the extension services with the Program. Finally, *SC* being a group-level certification, a farm gate price premium cannot be directly estimated.

At the mill gate, we observe the prices paid by *FNC* to the cooperatives for coffee sourced under the three programs. Relative to standard coffee, all three programs pay a positive price premium (Column 2 of Table 10). The Program

[52] To comply with the data sharing agreement, we do not explicitly name these VSSs.

[53] Note that the Program premium is estimated at the *sale* level, while the *EC* price premium is estimated at the *farmer* level. This reflects the difference between the Program (which pays a premium for quality) and *EC* (which is an individual certification unattached to any particular buyer). As noted in Section 3, however, Program farmers sell around 90% of their production to the Program. The Program price premium at the *farmer* level ($\approx 8.5 - 9\%$) is thus also larger than the *EC* premium.

premium is largest and, as noted in Table A7, its magnitude is consistent with no double-marginalization. The price premium for *EC* is 3%, roughly half of the Program premium. Given the small price premium at the farm gate, there appears to be double-marginalization for coffee sold under the *EC* label. The price premium for *SC* is 1.5%, roughly a quarter of the Program premium.

At the export gate, Column 4 shows that all three programs pay a positive price premium. The premium is $\approx$ 20% for the Sustainable Quality Program; $\approx$ 5% for *EC*; and $\approx$ 9% for *SC*. Two considerations are noteworthy. First, the Program transmits $(27c/75c) \approx 36\%$ ($(27c/37.5c) \approx 72\%$ if evaluated at market prices) of the export gate premium to farmers. The figure is $(2.7c/19c) \approx 14\%$ for *EC* and, given the even larger gap between the mill and export gate premia, presumably even lower for *EC*. Second, the mill and export gate premia are estimated conditional on the coffee being sold as certified. The three programs however differ substantially in the share of eligible coffee sold as certified. As discussed in Section 3, around 85-90% of the coffee produced by Program farmers is sold under the Program. We estimate the corresponding figure to be only 26% for *EC*. In Colombia, the share is even lower for *SC*.[54]

These facts are not specific to Colombia. The lack of guaranteed demand and the gap between volumes of certifi*able* versus actually sold as certifi*ed* coffee in the market are well documented in the literature (see, e.g., Samper and Quiñones Ruiz (2017)). Panhuysen and Pierrot (2014) note that in 2013 only 20-35% of worldwide production eligible for 4C, FT, RA and UTZ certifications was sold as such. de Janvry et al. (2015) provides a detailed analysis of how the floor price associated with Fair Trade contributes to this gap. They show that Fair Trade fails to deliver stated price premia because buyers do not committ to demand: when world prices are low and the Fair Trade premium binds, a lower share of eligible coffee is sold as FT certified. Minten et al. (2018) find that for Fair Trade and Organic certifications, there is very limited transmission of export premiums to coffee producers in Ethiopia.

[54]The Program has an explicit focus on quality while *EC* and *SC* do not. At the mill gate, Program batches have significantly higher quality, even relative to other VSSs (as already shown in Table 4). *EC* and *SC* have quality higher than standard coffee (Column 3, Table 10) but not different from other VSSs. At the export gate, Program batches also have significantly higher quality while both *EC* and *SC* have quality no different than standard coffee and *lower* than other VSSs (Column 5). A common explanation offered by practitioners in conversations with the authors is that the price premium associated with those labels does not depend on quality. This gives exporters an incentive to sell lower quality coffee under the label.

5.C. Summing Up: How did the Program Work?

We have tried to understand the role played by the different Program components complementing a counterfactual analysis based on the model's estimates with a descriptive comparison against two prominent NGO-driven VSSs. The two exercises paint a fairly coherent picture with two main take away points.

First, the contractual arrangements between *FNC* and the multinational buyer appear to play an important role in inducing quality upgrading and the resulting increase in farmers welfare. Several observations point towards this conclusion. The counterfactual analysis suggests that even a highly succesful program yielding a $\omega = 25\%$ increase in production would generate less than half of the overall gains from the Program. Our preferred interpretation is that ω provides an upper bound to the gains that can be achieved by programs that relax *supply*-side constraints alone. An extensive literature evaluating both government and NGO supported agricultural programs (such as extension services, fertilizer subsidies, and training) has provided a limited catalogue of successes (Magruder (2018)). At the top range of existing estimates, Deutschmann et al. (2019) evaluation of the One Acre Fund (1AF) Program finds a 24% increase in output associated with the program. Similarly to the Program, 1AF provides training and high-quality inputs but also crop insurance. We conjecture that it might be difficult to achieve a comparable impact on output without the multinational buyer's commitment. This conjecture is supported by the descriptive comparison: in Colombia, the Program *supply*-side interventions (extensions, training and inputs for plot renewal) and implementer are similar to those of the environmental label.

The importance of the contractual arrangements between *FNC* and the multinational buyer is further supported by results on the transmission of price premia from the export gate to the farm gate. The two NGO-driven VSSs appear to have limited to no transmission. In contrast, the Program transmits $\approx 70\%$ of the market quality premium at the export gate to the farm gate. The counterfactual analysis confirms that the contractual arrangements between *FNC* and the multinational buyer play a quantitatively important role in the transmission.

Second, the Program appears to have curbed one or more market failures along the chain. A smoking gun in favour of this interpretation is that the market price premium for *supremo* coffee at the export gate, estimated at $10\% \times 3.75\$ \approx 0.375\$$ is considerably larger than the (marginal) increase in unit costs to produce quality.

In our model farmers have homogeneous unit costs and so the marginal and average increase in unit costs coincide at $16.7\% \times 0.68\$ \approx 0.11\$$. We can, however, bound the marginal unit cost in a model in which farmers had heterogeneous costs of production. In such case, the marginal increase in unit costs to produce quality would be bounded above by the Program farm gate premium, $\pi \times p^S \approx 0.27\$$. Under perfectly functioning markets the gap between the two should vanish. The Program did not simply expand demand for quality coffee at the export gate.

This back-of-the-envelope calculation thus suggests that indeed there were constraints to quality upgrading in the sector that were relaxed by the Program. Our analysis provides evidence that the Program relaxed distortions arising from market power and limited contract enforcement along the chain.

The Program might also relax market failures preventing farmers from accessing inputs (including credit and information) to upgrade quality. Our analysis however is not conclusive since the estimated (reduced-form) impact on production ω could be entirely due to farmers responding to changed demand conditions and/or lower costs of accessing inputs. While disentangling these two aspects is left for future research, existing evidence suggests that farmers response to changes in demand might play a critical role. Few studies unbundle *demand* and *supply*-side interventions. Ashraf et al. (2009) evaluates a program that combines credit for inputs with increased access to export markets. Bundling the two components did not increase farmer incomes relative to only offering the export program. More recently, Arouna et al. (2019) vary the terms of a contract farming scheme. They find that giving farmers a price guarantee achieves the same impact as a contractual bundle that also includes training and input loans. Karlan et al. (2014) find that farmers provided with rainfall-index insurance are able to find resources to expand production on their farms. Casaburi and Willis (2018) find high take-up of a crop insurance product offered as part of a contract-farming scheme with sugar cane farmers in Kenya. This suggests that once *demand* uncertainty is resolved, farmers are able to address supply constraints on their own.[55]

[55] Bold et al. (2017) highlight how widespread fake and low quality agricultural inputs contribute to low technology adoption and upgrading in agriculture. In our context *FNC* has conducted numerous successful campaigns to improve inputs (e.g., with respect to tree varieties) and these concerns do not appear to be relevant.

6. Conclusion

Linking smallholder farmers in developing countries to global value chains (GVCs) has the potential to lift millions out of poverty but often necessitates quality upgrading. Foreign buyers can play an important role in this process (World Bank (2020)). This paper studies the *Sustainable Quality Program* - a buyer-driven quality upgrading program - in the Colombia coffee chain. The case provides a window into how multinational buyers can help farmers in developing countries overcome barriers to participation in GVCs.

We ask three questions. First, we show that the Program induced quality upgrading, expansion of land under coffee cultivation and higher farm gate prices among eligible farmers. Second, we quantify how the Program gains are shared between farmers and intermediaries along the chain. In regions in which it was rolled out the Program increased surplus along the chain by $\approx 30\%$ with farmers keeping at least half of these gains. Finally, counterfactual exercises and a comparison of Program price premia along the chain against two prominent non-buyer driven VSS reveal that the Program improved the transmission of the export gate price premium to the farm gate and curbed market failures that stifled quality upgrading. Contractual arrangements at the export gate significantly contributed to the increase in welfare in rural areas.

In future work, we plan to turn our attention to aspects that we left out from this analysis. First, we do not investigate the Program impact on inequality across farmers. There is growing interest in understanding how trade impacts inequality (see Goldberg and Pavcnik (2007) for a review with a focus on developing countries). Through its impact on quality upgrading, export opportunities can increase inequality. For example, Verhoogen (2008) shows that exporting increased inequality among Mexican manufacturing workers due to its higher skill content. In our context, the Program take-up rates are higher among farmers that had larger and better coffee plantations at the time of becoming eligible. The Program thus might have increased inequality in eligible communities.[56]

Second, our attempt to quantify the impact of the Program abstracts from a number of channels. Our analysis focuses on the welfare impact of the Program

[56] Dragusanu and Nunn (2014) find positive impact of Fair Trade certification on the income of Costa Rican coffee farmers. Looking at the distributional impact of Fair Trade, they find that the benefits are not evenly distributed: skilled coffee growers benefit, intermediaries are hurt, and unskilled workers are unaffected.

on coffee farmers in eligible localities and on the intermediaries directly involved in the Program supply chain. Within the Colombia coffee chain, we abstract from potential losses for competing exporters and intermediaries. To a first approximation, the export figures in the introduction and the results of our calibration analysis suggest that the Program *did not* divert quality coffee away from direct competitors. The Program, however, diverted production away from standard coffee. Since we have assumed zero-profits for intermediaries of standard coffee, such diversion has no negative impact in our framework.

The Program expanded production and land under coffee cultivation with impacts that might go beyond the coffee chain. Labour accounts for the majority of the costs of expanding production and upgrading quality (e.g., through more careful and frequent harvest). The Program might have expanded labour demand particularly for seasonal workers which are often among the poorest in rural areas. The overall Program impact on inequality in rural communities is therefore ambiguous. By diverting resources away from other crops, the Program might impact welfare in other chains and, potentially, food security.

Our analysis also abstracts from environmental benefits associated with the Program. The Program sets stringent environmental requirements on coffee plantations. Several studies have documented environmental benefits associated with the kind of practices required by the Program (see Ibañez and Blackman (2016), Rueda et al. (2015), Rueda and Lambin (2013)). By not including those, our calculations might *under*state the welfare benefits of the Program.

Finally, the Program success in Colombia might depend on the local context. Two aspects might play a particular important role: *i*) the local implementer capacity and relationship with farmers; *ii*) the existence of the Garantía de Compra. The former might have been key to set up successful extension and training services, gain farmers trust and develop the relationship with the multinational buyer. The Garantía de Compra might have eliminated negative spillover on eligible farmers that did not join the Program. Exploring the impact of the Program in other contexts is an important avenue for future research.

Notwithstanding these limitations, policy implications can be harvested from the analysis. The evidence points at the critical role of guaranteeing a stable demand, with adequate price transmission from the export gate to the farm gate, to harness the potential for quality upgrading. The long-term relationship between the multinational buyer and the exporter was critical in overcoming constraints to quality upgrading and resulted in contractual arrangements that bypassed monopsony distortions in the domestic chain. To increase farmers income, governments and regulators should focus on strengthening exporters capabilities to initiate, develop and sustain long-term relationships with large buyers involved in GVCs.

References

Alfaro, L., Rodríguez-Clare, A., Hanson, G. H., and Bravo-Ortega, C. (2004). Multinationals and linkages: an empirical investigation [with comments]. *Economia*, 4(2):113–169.

Alfaro-Ureña, A., Manelici, I., and Vasquez, J. (2019). The effects of joining multinational supply chains: New evidence from firm-to-firm linkages. *Working paper.*

Antràs, P. (2015). *Global production: Firms, contracts, and trade structure.* Princeton University Press.

Antràs, P. and Costinot, A. (2011). Intermediated trade. *The Quarterly Journal of Economics*, 126(3):1319–1374.

Antràs, P. and Foley, C. F. (2015). Poultry in motion: a study of international trade finance practices. *Journal of Political Economy*, 123(4):853–901.

Arouna, A., Michler, J. D., and Lokossou, J. C. (2019). Contract farming and rural transformation: Evidence from a field experiment in benin. *National Bureau of Economic Research Working Paper.*

Ashraf, N., Giné, X., and Karlan, D. (2009). Finding missing markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in kenya. *American Journal of Agricultural Economics*, 91(4):973–990.

Atkin, D. and Donaldson, D. (2015). Who's getting globalized? the size and implications of intra-national trade costs. *Working paper.*

Atkin, D., Khandelwal, A. K., and Osman, A. (2017). Exporting and firm performance: Evidence from a randomized experiment. *The Quarterly Journal of Economics*, 132(2):551–615.

Bardhan, P., Mookherjee, D., and Tsumagari, M. (2013). Middlemen margins and globalization. *American Economic Journal: Microeconomics*, 5(4):81–119.

Blouin, A. and Macchiavello, R. (2019). Strategic default in the international coffee market. *The Quarterly Journal of Economics*, 134(2):895–951.

Bold, T., Kaizzi, K., Svensson, J., and Yanagizawa-Drott, D. (2017). Lemon technologies and adoption: Measurement, theory and evidence from agricultural markets in uganda. *The Quarterly Journal of Economics*, 132:1055–1100.

Bosselmann, A. S., Dons, K., Oberthur, T., Olsen, C. S., Raebild, A., and Usma, H. (2009). The influence of shade trees on coffee quality in small holder coffee agroforestry systems in southern colombia. *Agriculture, ecosystems & environment*, 129(1):253–260.

Boudreau, L. (2019). Multinational enforcement of labor law: Experimental evidence from bangladesh's apparel sector. *Working paper*.

Brambilla, I. and Porto, G. G. (2006). *Farm productivity and market structure: Evidence from cotton reforms in Zambia*. The World Bank.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms. *American economic review*, 101(1):304–40.

Cadot, O., Dutoit, L., and De Melo, J. (2008). The elimination of madagascar's vanilla marketing board, 10 years on. *Journal of African Economies*, 18(3):388–430.

Cajal-Grossi, J., Macchiavello, R., and Noguera, G. (2019). International buyers' sourcing and suppliers' markups in bangladeshi garments. *CEPR Discussion Paper No. DP13482*.

Casaburi, L. and Macchiavello, R. (2019). Demand and supply of infrequent payments as a commitment device: evidence from kenya. *American Economic Review*, 109(2):523–55.

Casaburi, L. and Willis, J. (2018). Time versus state in insurance: Experimental evidence from contract farming in kenya. *American Economic Review*, 108(12):3778–3813.

Costinot, A. and Rodríguez-Clare, A. (2014). Trade theory with numbers: Quantifying the consequences of globalization. In *Handbook of international economics*, volume 4, pages 197–261. Elsevier.

De Chaisemartin, C. and D'Haultfeuille, X. (2019). Two-way fixed effects estimators with heterogeneous treatment effects. *National Bureau of Economic Research Working Paper.*

de Janvry, A., McIntosh, C., and Sadoulet, E. (2015). Fair trade and free entry: Can a disequilibrium market serve as a development tool? *The Review of Economics and Statistics*, 97(3):567–573.

Deutschmann, J. W., Duru, M., Siegal, K., and Tjernström, E. (2019). Can smallholder extension transform african agriculture? *National Bureau of Economic Research Working Paper.*

Donaldson, D. (2015). The gains from market integration. *economics*, 7(1):619–647.

Dragusanu, R., Giovannucci, D., and Nunn, N. (2014). The economics of fair trade. *Journal of Economic Perspectives*, 28(3):217–36.

Dragusanu, R. and Nunn, N. (2014). The impacts of fair trade certification: Evidence from coffee producers in costa rica. *Working paper.*

Echavarria, J. J., Gamboa-Arbelaez, J., Hirs-Garzon, J., and Villamizar-Villegas, M. (2017). Impacto de subsidios en el agro: Una mirada desde el sector cafetero. *Borradores de Economia - Banco de la Republica de Colombia.*

Emran, M. S., Mookherjee, D., Shilpi, F., and Uddin, M. H. (2019). Credit rationing and pass-through in supply chains: Theory and evidence from bangladesh. *Working Paper.*

Fafchamps, M., Hill, R., Kaudha, A., et al. (2004). The transmission of international commodity prices to domestic producers. *Working paper, University Library of Munich, Germany.*

Gereffi, G. (1999). International trade and industrial upgrading in the apparel commodity chain. *Journal of international economics*, 48(1):37–70.

Giovannucci, D. and Ponte, S. (2005). Standards as a new form of social contract? sustainability initiatives in the coffee industry. *Food policy*, 30(3):284–301.

Goldberg, P. K. and Pavcnik, N. (2007). Distributional effects of globalization in developing countries. *Journal of economic Literature*, 45(1):39–82.

Hansman, C., Hjort, J., León, G., and Teachout, M. (2019). Vertical integration, supplier behavior and quality upgrading among exporters. *Working paper*.

Harrison, A. and Rodríguez-Clare, A. (2010). Trade, foreign investment, and industrial policy for developing countries. In *Handbook of development economics*, volume 5, pages 4039–4214. Elsevier.

Hausmann, R., Hwang, J., and Rodrik, D. (2007). What you export matters. *Journal of economic growth*, 12(1):1–25.

Ibañez, M. and Blackman, A. (2016). Is eco-certification a win–win for developing country agriculture? organic coffee certification in colombia. *World Development*, 82:14–27.

Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, 129(2):597–652.

Kranton, R. and Swamy, A. V. (2008). Contracts, hold-up, and exports: Textiles and opium in colonial india. *American Economic Review*, 98(3):967–89.

Laderach, P., Oberthur, T., Cook, S., Iza, M. E., Pohlan, J. A., Fisher, M., and Lechuga, R. R. (2011). Systematic agronomic farm management for improved coffee quality. *Field Crops Research*, 120(3):321–329.

Leibovich, J. and Ocampo, J. A. (1985). La comercialización externa de café colombiano. *Coyuntura Económica*, 15(3):177–206.

Leonel, L.-E. and Philippe, V. (2007). Effects of altitude, shade, yield and fertilization on coffee quality (coffea arabica l. var. caturra) produced in agroforestry systems of the northern central zones of nicaragua. *Journal of Food Science*, 68(7):2356–2361.

Linder, S. B. (1961). *An essay on trade and transformation*. Almqvist & Wiksell Stockholm.

Macchiavello, R. and Miquel-Florensa, J. (2018). Vertical Integration and Inter-Firm Relationships in the Costa Rica Coffee Chain. *Working paper.*

Macchiavello, R. and Morjaria, A. (2015). The value of relationships: evidence from a supply shock to kenyan rose exports. *American Economic Review*, 105(9):2911–45.

Macchiavello, R. and Morjaria, A. (2019). Competition and Relational Contracts: Evidence from Rwanda's Coffee Mills. *Working paper.*

Magruder, J. R. (2018). An assessment of experimental evidence on agricultural technology adoption in developing countries. *Annual Review of Resource Economics*, 10:299–316.

McMillan, M., Rodrik, D., and Welch, K. H. (2002). When economic reform goes wrong: cashews in mozambique. *National Bureau of Economic Research Working Paper.*

Minten, B., Dereje, M., Engida, E., and Tamru, S. (2018). Tracking the quality premium of certified coffee: evidence from ethiopia. *World Development*, 101:119–132.

Nelson, V. and Pound, B. (2009). The last ten years: A comprehensive review of the literature on the impact of fairtrade. *Natural Resources Institute*, (1-48).

Oberthur, T., Laderach, P., Posada, H., Fisher, M. J., Samper, L. F., Illera, J., Collet, L., Moreno, E., Alarcon, R., Villegas, A., et al. (2011). Regional relationships between inherent coffee quality and growing environment for denomination of origin labels in nariño and cauca, colombia. *Food policy*, 36(6):783–794.

Panhuysen, S. and Pierrot, J. (2014). Coffee barometer. *HIVOS, The Hague, Netherlands.*

Puerta, G., Gonzalez, F., Correa, A., Alvarez, I., Ardilla, J., Girón, O., Ramírez, C., Baute, J., Sanchez, P., Santamaria, M., et al. (2016). Diagnóstico de la calidad del café según altitud suelos y beneficio en varias regiones de colombia. *Cenicafé.*

Puerta, G. I. (2001). Como garantizar la buena calidad de la bebida del cafe y evitar los defectos. Technical report.

Rey, P. and Tirole, J. (1986). The logic of vertical restraints. *The American Economic Review*, pages 921–939.

Rodríguez-Clare, A. (1996). Multinationals, linkages, and economic development. *The American Economic Review*, pages 852–873.

Rueda, X. and Lambin, E. F. (2013). Linking globalization to local land uses: How eco-consumers and gourmands are changing the colombian coffee landscapes. *World Development*, 41:286–301.

Rueda, X., Thomas, N. E., and Lambin, E. F. (2015). Eco-certification and coffee cultivation enhance tree cover and forest connectivity in the colombian coffee landscapes. *Regional environmental change*, 15(1):25–33.

Samper, L. F. and Quiñones Ruiz, X. F. (2017). Towards a balanced sustainability vision for the coffee industry. *Resources*, 6(2).

Startz, M. (2019). The value of face-to-face: Search and contracting problems in nigerian trade. *Working paper.*

Tanaka, M. (2017). Exporting sweatshops? evidence from myanmar. *Review of Economics and Statistics*, pages 1–44.

Tirole, J. (1988). *The theory of industrial organization.* MIT press.

Velmourougane, K., Bhat, R., Gopinandhan, T. N., and Panneerselvam, P. (2011). Impact of delay in processing on mold development, ochratoxin a and cup quality in arabica and robusta coffee. *World Journal of Microbiology and Biotechnology*, 27(8):1809–1816.

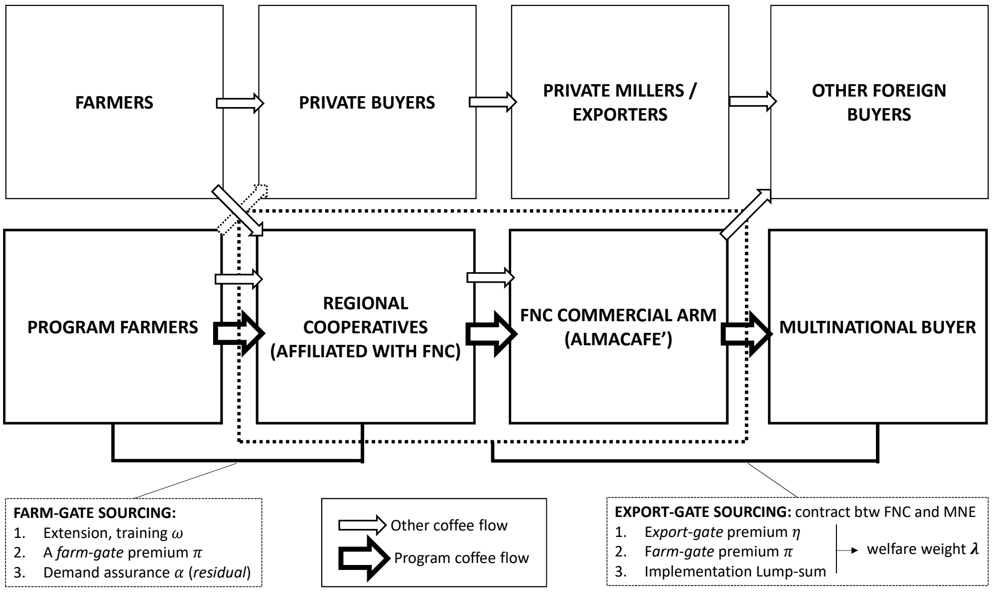
Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the mexican manufacturing sector. *The Quarterly Journal of Economics*, 123(2):489–530.

World Bank (2007). *World development report 2008: Agriculture for development.* World Bank.

World Bank (2020). World development report on global value chains. *In Preparation.*

Main figures

Figure 1: **The Colombian Coffee Chain and The Sustainable Quality Program**



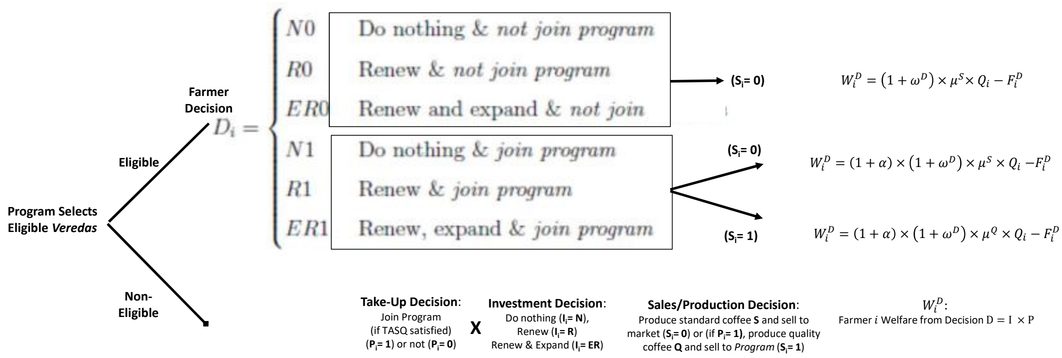
The Figure illustrates both the Colombia coffee chain (*The Chain*) as well as the Sustainable Quality Program (*The Program*). The Figure also introduces the notation used in the model in Section 4.

The Chain: The Colombia coffee chain is as follows. Farmers, mostly smallholders, growth coffee. When coffee cherries change color from green to red they are ripe for harvest. The timing of the harvest season depends on location and weather patterns. Coffee cherries must be processed immediately after harvest to obtain parchment coffee. In Colombia, farmers undertake the first stage processing. Farmers sell parchment coffee to intermediaries: either private buyers or cooperatives. Private buyers include independent farm-gate traders as well as buying agents for milling companies. Regional cooperatives affiliated with the *Federación Nacional de Cafeteros* (henceforth FNC) operate buying points scattered around the country-side. Intermediaries deliver coffee to one of around 60 hulling mills in the country. Hulling mills can be either owned by private companies (mainly exporters) or by the *FNC*. The FNC is a para-statal body in charge of regulating and implementing, through the regional cooperatives, sectoral policies. Through its commercial arm (Almacafé), FNC exports around 30% of Colombian coffee. Finally, (most) coffee is exported to foreign buyers.

The Program: The Sustainable Quality Program is a bundle of contractual arrangements involving all actors along the chain. In Colombia, like other Voluntary Sustainability Standards, the Sustainable Quality Program is implemented by FNC through the regional cooperatives operating in the localities where the Program is rolled out. At the *farm gate*, the Program involves extension services training and access to inputs aiming to increase both yields and quality. The Program selects areas based on quality potential. All farmers in the selected areas are eligible to receive extension services, training and inputs. However, in order to become a Program farmer, the farmer must upgrade their farms and satisfy certain quality requirements. Program farmers have the option, but not the obligation, to sell coffee that meets the program's standard at a fixed price premium. Only the regional cooperatives involved source coffee for the Program. The cooperatives then deliver the coffee to Almacafé, the commercial arm of FNC, which prepares it for export to the multinational buyer. The cooperatives and Almacafé have a close and long-standing relationship and, as shown in Table A7, they act as a vertically integrated actor from the point of view of the Program.
At the *export gate*, the Program involves a contract between the multinational buyer (MNE) and FNC. To ensure the reliable and traceable supply of high quality coffee, the Program pays a significant price premium at the export gate (denoted η). In addition, the contract between the multinational buyer and the FNC includes provisions regarding the price premium to be paid at the farm gate (π) and a(n undisclosed) lump-sum transfer from the multinational buyer to the FNC to cover the costs of implementing the Program. Regardless of the distribution of bargaining power between FNC and the multinational buyer, the non-linear contract between the FNC and the multinational buyer results in a farm gate premium π higher than what would be chosen by a profit-maximizing FNC that sets it after having negotiated η at the export gate. It is thus convenient to represent the contract as choosing π to maximize an objective function with weight λ over FNC profits and $(1-\lambda)$ over farmers' welfare. This reduced-form representation avoids modeling the multinational margins and the bargaining protocol.

The Figure also illustrates the flow of coffee sold through the Sustainable Quality Program. Program coffee is *exclusively* sourced by the implementing cooperatives from plots that have joined the program in eligible localities. Program farmers, however, can sell their production to any buyer and can also sell to the cooperative as non-program coffee (e.g., if the quality doesn't meet the Program's standards). Traceability, however, requires that program coffee can only be sold, and sourced, through the implementing cooperatives. The cooperatives also source from non-Program farmers. Similarly, Almacafé also sources non-program coffee from the cooperative and from other intermediaries.

Figure 2: **The Program at the farm gate**



The Figure illustrates the program's key elements and timing of events, from the farmer's point of view. The Figure also introduces the notation used in the model in Section 4.

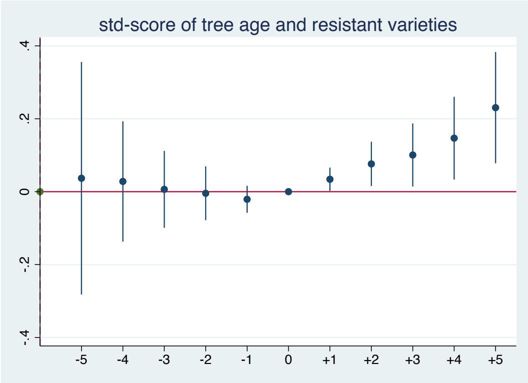
Eligibility: First, based on *terroir* conditions, the Program selects eligible *veredas*. All plots in selected veredas are eligible to receive extension services and training and to join the Program.

Joining the Program and Upgrading: To join the Program the plot must meet certain quality and environmental criteria. The farmer decides how to respond to the eligibility status of her plot (and thus to better access to inputs, training and extension) by taking both an investment $\mathbf{I_i}$ and a joining $\mathbf{P_i}$ decision, i.e., $D_i = \mathbf{I_i} \times \mathbf{P_i}$. The farmer could *i)* do nothing ($D_i = N0$); *ii)* renew the plot without joining ($D_i = R0$); *iii)* renew and expand the plot without joining ($D_i = E0$); *iv)* join the Program and do nothing ($D_i = N1$); *v)* renew the plot and join the Program ($D_i = R1$) or *vi)* renew, expand and join the Program ($D_i = E1$). Each decision entails farmer-specific fixed costs (e.g., depending on the current state of the plantation, uprooting and replanting might be needed, in addition to the opportunity cost arising from the fact that new trees take two years before they are productive). We denote the farmer-specific fixed costs of each decision d_i as F_i^d. Upgrading and/or expanding the plantation result in higher productivity and yield per hectare (ω^d).

Production and Sales: To sell to the Program, farmers that have joined ($\mathbf{P_i}$) must incur $\gamma = 17\%$ higher unit costs relative to standard quality (which has constant unit cost c^S). Program farmers have the option, but not the obligation, to sell coffee to the Program. If they sell to the program, they are paid a price premium π over the market price for standard coffee, denoted p^S. The unit margin on standard coffee is denoted $\mu^S = p^S - c^S$ and the margin on quality (and Program) coffee μ^Q is defined accordingly.

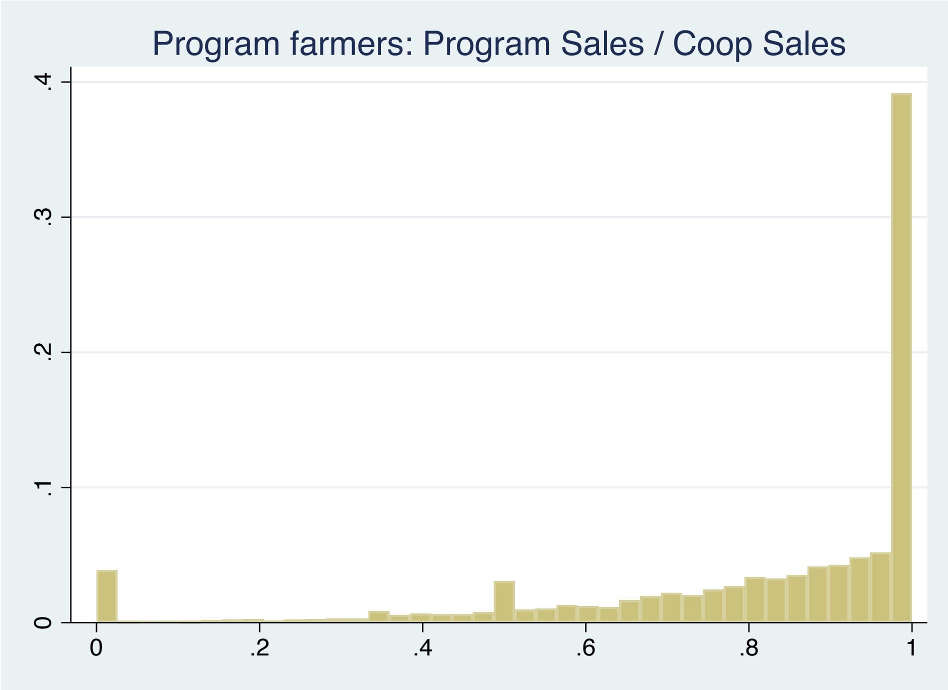
Farmers might derive other demand-side benefits from the program. For example, farmers might value the option to be able to sell to the program. The program also provides a degree of price insurance, in addition to the Guarantía de Compra available to all farmers. To see why, note that the base price in the Guarantía de Compra p^S tracks the world price p^W and is thus random. Denote $\bar{p}^S$ its expected value. The Program pays a fixed amount $\mathbf{p} = \pi \times \bar{p}^S$ on top of the basis price p^S. All else equal, this fixed premium is more valuable to a risk-averse farmer when p^S is low. The program might also mitigate weather-related risk by helping farmers planting disease resistant-varieties. Given the perfectly elastic demand, this insurance might also be valuable. The value of these benefits depend on the curvature of the farmer's utility function, their expectations and discount factor, which are not observable. We thus model these valuable Program attributes as a positive *wedge* α in the utility that farmers derive from joining the Program, regardless of their sales decision.

Figure 3: **Plot Upgrading: Parallel Trends Before Program Roll-Out**



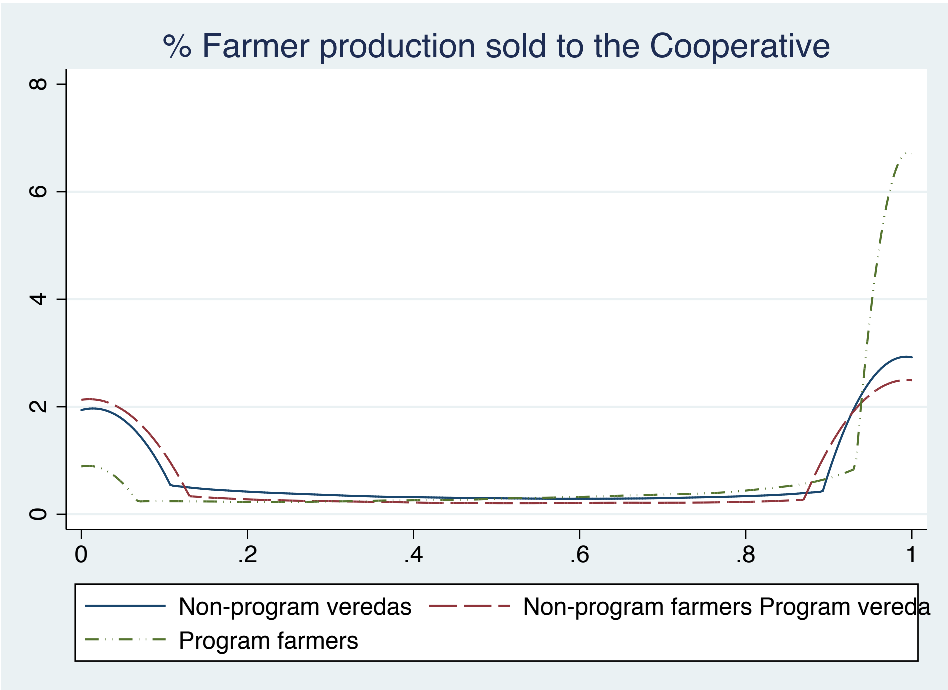
The Figure investigates pre-trends for the plot quality index adapting the baseline ITT specification in Column 2 of Table 2. The preferred ITT specification suffers from typical concerns arising in DID designs, including endogenous timing of program roll-out and differential trends. To assuage these concerns the Figure investigates dynamic patters in plot upgrading around the time the *vereda* becomes eligible. The Figure shows the absence of differential trends in plot upgrading in eligible *veredas* in the years preceding the roll-out of the Program. The Figure also shows a gradual increase in the quality of coffee plantations in the years after the Program is rolled out in the *vereda*, consistent with the take-up patterns in Figure A3.
A word of caution on the Figure is that the estimated coefficients 4 to 5 years before the Program roll-out display relatively large standard errors. This is a consequence of the fact that data are available starting in 2006, only two (four) years before the largest waves of program roll-out in Nariño and Cauca respectively. Figure A5 in the Appendix confirms the absence of pre-trends in the main outcomes of interest through more precise estimates by focusing on the Cauca region only for which more years are available before the program roll-out. Table A4 presents the coefficients in the figure.
De Chaisemartin and D'Haultfeuille (2019) note that DID designs with period and group fixed effects identify weighted sums of average treatment effects (ATEs) in each group and period with weights that may be negative and propose a correction. Unreported results show that the negative weights sum only to 0.16. The correction identifies dynamic coefficients close to those in Table A4. An alternative spatial discontinuity design in Table B4 also confirms the Program's positive impact on plot plot upgrading.

Figure 4: **Program Sourcing: No Hold-Up**



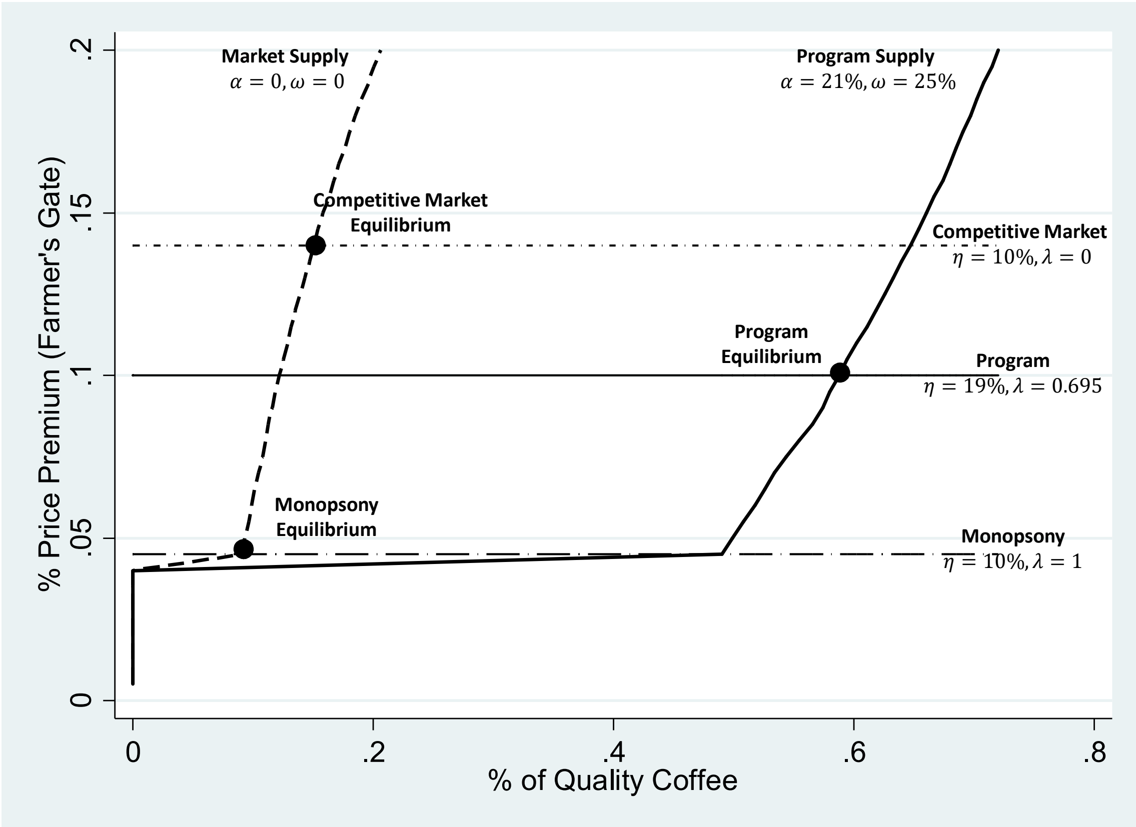
The Figure shows the distribution of the share of Program farmer's deliveries that actually occur under the Program. The data covers volumes and prices for any coffee soured by one of the two Program implementing cooperatives during the 2015 and 2016 harvest seasons. On average, over 80% of Program's farmers deliveries to the cooperative are absorbed by the Program. A small percentage of Program farmers do not sell their produce under the Program. Most likely these farmers have *de facto* abandoned the Program. Conditional on delivering through the Program, over 90% of Program farmers deliveries are accepted by the Program. This Figure stands in stark contrast to those for other non-buyers driven Voluntary Sustainability Standards, as discussed in Section 5.

Figure 5: **Program Sourcing: No Side-Selling**



The Figure provides descriptive evidence on farmers' sales patterns across eligible and non-eligible veredas. The Figure relies on data from the Protección del Ingreso Cafetero (PIC) program. The PIC data covers the universe of coffee transactions between *any* farmer and *any* buyer - private and cooperatives - in the Colombian country side for the 2012-2013 season. The sample includes all farmers in the municipalities where the Program eventually expanded. The Figure shows the share of farmer's sales delivered to the cooperative implementing the Program (with any other buyers being the alternative). Farmers in Program veredas deliver significantly more coffee to the Program implementer than farmers in non-eligible ones. The difference, however, is entirely driven by Program farmers. In fact, the Figure shows that a significantly higher share of Program farmers (79% vs. 64 %) sell almost all their coffee to the cooperative implementing the Program.

Figure 6: **The Program: Demand and Supply at the farm gate**



The Figure illustrates how the Program changed the supply and demand for quality coffee at the farm gate. The two endogenous variables are the share of quality coffee produced (on the x-axis) and the farm gate price premium (on the y-axis).
Relative to the monopsonistic benchmark, the Program shifted the *supply* curve out by incentivizing farmers to join the Program and upgrade. Besides the endogenous price premium π, the Program incentivized farmers to join through its impact on production volumes $\omega = 25\%$ and through valuable demand-side components $\alpha = 21\%$ (e.g., option value and price insurance). The supply curve under the monopsonistic market benchmark is given by setting $\omega = \alpha = 0$. Note that under both scenarios, the supply of quality coffee is zero unless the price premium gives farmers higher margins for quality coffee relative to standard coffee ($\mu^Q \geq \mu^S$), i.e., $\pi \geq c \times \gamma / p^S \approx 4.5\%$.
On the *demand* side, the Program alters both the buyer's willingness to pay for quality coffee at the export gate (from $\eta = 10\%$ to $\eta = 19\%$) as well as the price transmission mechanism (from $\lambda = 1$ to $\lambda = 0.695$). This results in a farm gate price premium of 10% instead of the (minimum) 4.5%.
As a benchmark, the Figure also reports the farm gate price premium that would result from a competitive market ($\pi = 10\% \times \tau \approx 14\%$).

Main tables

Table 1: **Descriptive Statistics (by Region)**

	(1) *Nariño*		(2) *Cauca*	
	N	Mean/SE	N	Mean/SE
Part A: Plot Characteristics (2006 Census)				
Total Plot Area (Ha.)	32755	0.666	27427	0.513
		(0.004)		(0.003)
Coffee Cultivated Area (Ha.)	32755	0.568	27427	0.491
		(0.003)		(0.003)
Area plot with shade	32755	0.189	27427	0.187
		(0.002)		(0.002)
Average Tree Age	32755	9.601	27427	13.768
		(0.037)		(0.051)
Tree Density (per Ha.)	32755	5445.485	27427	5117.767
		(8.356)		(10.000)
Share Plot Resistant Varieties	32755	0.168	27427	0.188
		(0.002)		(0.002)
Part B: Farmer Characteristics (2006 Census)				
Total Land (Ha.)	22946	0.949	20271	0.695
		(0.007)		(0.005)
Coffee Cultivated Area (Ha.)	22946	0.811	20271	0.665
		(0.005)		(0.004)
Number of Plots	22946	1.427	20271	1.353
		(0.005)		(0.005)
Part C: Farmer Participation in FNC Programs (Average 2007-2013)				
Individual Extension	38354	0.949	35972	0.948
		(0.001)		(0.001)
Credit Program	38354	0.485	35972	0.637
		(0.002)		(0.002)
Extension program	38354	0.907	35972	0.911
		(0.001)		(0.001)
ID program	38354	0.886	35972	0.883
		(0.001)		(0.001)

Notes: The table compares the plots/farmers from the Program municipalities.
Part A: Plot weighted age and density denote the plot age and density (trees per Ha.) weighed across the different subplots in the plot. The *Share resistant varieties* denotes the share of the plantation in rust-resistant varieties.
Part C: The *"Individual extension"* dummy takes value 1 when the farmer had a one-to-one activity with the extension services. The *"Extension program"* dummy takes value one when farmer has participated in any group or individual extension program. *"ID program"* refers to the FNC program to ensure all farmers had an ID that allowed them to do monetary transactions with the cooperative and keep track of the programs they are involved and their benefits.

Table 2: **The Program - Plot Upgrading**

	(1)	(2)	(3)	(4)
	Standardized Plot Upgrading Score (Tree Age and Share Resistant Varieties)			
	OLS	ITT	Heterogeneity (ITT)	
			Good plots	Other plots
Program Plot	0.1862***			
	(0.009)			
Program Vereda		0.0478**	0.1506***	-0.0213
		(0.020)	(0.034)	(0.022)
Observations	775,263	775,263	196,985	578,278
Number of plots	91,766	91,766	32,060	59,706
Plot FE	Yes	Yes	Yes	Yes
Mun-Year FE	Yes	Yes	Yes	Yes

Notes: Robust standard errors (cluster vereda) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Time period is 2006 to 2016. Unit of observation is plot-year.
The *Plot Upgrading Score* is an index constructed with equal weight from the yearly standardized (negative) average age of coffee trees on the plot and the share of planted area with trees that are of roya resistant varieties.
The variable *"Program Plot"* is a dummy taking value 1 after the plot joins the Program. The variable *"Program Vereda"* is a dummy taking value 1 for all plots in a vereda after the vereda becomes eligible for the Program. The sample includes all plots in the municipalities where the Program eventually expanded. At the end of the panel 97% of the plots in these municipalities were eligible.
The sample *good plots* is defined as plots in the top quartile of the Plot Upgrading Score at the moment the vereda becomes eligible.

Table 3: **The Program - Expansion of Land under Coffee Cultivation**

Panel A: Plot Level

	(1)	(2)	(3)	(4)
	Coffee Planted Area (Ln)		Plot Exit	
Program Plot	0.0720***		-0.0609***	
	(0.005)		(0.001)	
Program Vereda		0.0173***		-0.0048
		(0.006)		(0.004)
Observations	775,263	775,263	775,263	775,263
R^2	0.080	0.077	0.087	0.078
Number of plots	91,766	91,766	91,766	91,766
Plot FE	Yes	Yes	Yes	Yes
Mun-Year FE	Yes	Yes	Yes	Yes

Panel B: Vereda Level

	(1)	(2)	(3)	(4)
	Coffee Area (Ln)	N of Plots Exiting	N of Farmers	N of Plots
Program Vereda	0.0989***	-0.0155**	2.0397**	2.6247**
	(0.030)	(0.008)	(0.843)	(1.063)
Observations	10,797	10,797	10,797	10,797
R^2	0.559	0.325	0.430	0.401
Number of veredas	1,046	1,046	1,046	1,046
Vereda FE	Yes	Yes	Yes	Yes
Mun-Year FE	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses (cluster vereda in Panel A and cluster municipality in Panel B). *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. Time period is 2006 to 2016. Unit of observation is plot-year in Panel A and vereda-year in Panel B. The variable *"Program plot"* takes value 1 after the plot joins the Program. The variable *"Program vereda"* takes value 1 for any plot in a vereda after the vereda becomes eligible for the Program. The sample includes all plots in the municipalities where the Program eventually expanded. At the end of the panel 97% of the plots in these municipalities were eligible.

Table 4: **The Program - Coffee Quality**

Panel A (I): All Batches, *Within* Origin-Season

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Quality index		**% healthy beans**		**% pasilla beans**		**% broca affected beans**	
Program Batches	0.4051***	0.1348***	0.0181***	0.0064***	-0.0133***	-0.0040***	-0.0016***	-0.0005**
	(0.037)	(0.029)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Sample	all	non-std	all	non-std	all	non-std	all	non-std
Mean dependent var.	–	–	0.936	0.956	0.0533	0.0378	0.0134	0.00780
Observations	113,210	66,283	113,210	66,283	113,210	66,283	113,210	66,283
R^2	0.654	0.630	0.675	0.662	0.668	0.622	0.747	0.709
Origin - Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel A (II): All Batches, *Within* Origin-Season-Month

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Program Batches	0.4236***	0.1184***	0.0193***	0.0055***	-0.0140***	-0.0037***	-0.0018***	-0.0008***
	(0.037)	(0.025)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Observations	113,210	66,283	113,210	66,283	113,210	66,283	113,210	66,283
R^2	0.472	0.445	0.488	0.469	0.501	0.454	0.598	0.556
Origin - Year - Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Non-Program Batches, Program Origins vs. Non-Program Origins

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Quality index		**% healthy beans**		**% pasilla beans**		**% broca affected beans**	
Program Origin	0.142	0.132	0.004	0.005	-0.007*	-0.007*	-0.001	-0.0017
	(0.089)	(0.091)	(0.004)	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)
Sample	all	non-std	all	non-std	all	non-std	all	non-std
Mean dependent var.	–	–	0.919	0.946	0.0645	0.0455	0.0155	0.00802
Observations	19,095	12,465	19,095	12,465	19,095	12,465	19,095	12,465
R^2	0.502	0.488	0.489	0.513	0.473	0.486	0.642	0.622
Origin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors (cluster year-origin) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Time period is 2009-2014. Differences on time span for the quality measures are due to changes in FNC recording policies. The unit of observation is a coffee batch entering the mill. The *quality index* is the z-score of yearly standardized grams of healthy beans in sample and the negative of the grams of beans qualifying as pasilla (subproduct) in sample. Index time span is 2010-2014. The shares of broca affected beans, pasilla and healthy beans are available for the 2010-2014 period. Rendimiento is the measure of conversion of parchment into green coffee (Kg. of parchment needed to obtain 1 Kg. of green coffee after mill processing). Samples refer to all and to non-standard batches from the region where the Program was implemented.
Panels A (I) and (II) compare batches of coffee sourced for the Program against non-program batches sourced from the same origin (buying point) in the same season (*Within* Origin-Season). Panel B compares non-program batches sourced sourced from Program origins and non-program origins. A buying point becomes a Program origin after at least one vereda supplying the buying point becomes eligible for the Program. Panel B shows that non-program batches from Program origins do not have lower quality than non-program batches from non-program locations. If anything, the quality of non-program batches is slightly higher after the origin becomes eligible for the Program. The quality differentials presented in Panel A are thus not due to bean sorting: the Program increased the aggregate supply of quality coffee in the market.
Quality testings are available only for the cooperatives buying points. A final concern is thus that farmers sort low quality beans and sell them to other buyers. If that was the case, we expect Program farmers to get lower prices from these other buyers. Panel B of Table 6 provides direct evidence against this hypothesis. Furthermore, the existence of the Garantía de Compra implies that, if anything, farmers have incentives to sell the lower quality to the cooperatives buying points.

Table 5: **The Program - Price Premium at the Farm Gate**

	(1)	(2)	(3)	(4)
		ln price per kg.		
Program sales	0.0960***	0.0950***	0.0950***	0.0573***
	(0.008)	(0.009)	(0.011)	(0.016)
Sale types included	All	All	All	Non-standard
Observations	35,103	35,103	35,103	29,306
R^2	0.677	0.777	0.833	0.881
Buying point -Year FE	Yes	Yes	Yes	Yes
Farmer FE	No	Yes	–	–
Farmer-Year FE	–	–	Yes	Yes

Notes: Robust standard errors (cluster buying point year) in parentheses *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. The data for this table comes from farmer sales to one of the implementing cooperatives. Unit of observation is farmer/program/year/buying point. Time period is 2015-2016. Columns (1) to (3) compare Program sales with all other sale types of the farmer. Column (4) compares only across non-standard sales. The sample includes 10134 farmers. From the five programs considered (Standard, The Sustainable Quality Program, Environmental, Specialty and Regional) 37.51% of the farmers sell only to one program (mean 1.9, median 2), and only 1.13 % sell to more than 3 programs. Farmers sell on average to 1.5 different buying points (median 1), 52% of farmers are exclusive to one buying point, 93% sell to one or two. From the 10134 farmers, 9,275 own Program plots in 2016, and 8492 own Program plots in 2015.

Table 6: **Spatial Discontinuity Design: (No) Side-Selling and Prices**

Panel A: Farmer Sales

	(1)	(2)	(3)	(4)	(5)	(6)
	Sells to Cooperative		**Sells to Other Buyers**		**Share Sold to Cooperative**	
Program farmer	0.2004***		-0.1040***		0.1967***	
	(0.027)		(0.016)		(0.032)	
Program vereda		0.0516*		-0.0376+		0.0455*
		(0.027)		(0.023)		(0.025)
Observations	5,829	5,829	5,829	5,829	5,829	5,829
Farmer controls	Yes	Yes	Yes	Yes	Yes	Yes
Border (1km) FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Prices

	(1)	(2)	(3)	(4)	(5)	(6)
	ln price (all buyers)		**ln price (Cooperative)**		**ln price (other buyers)**	
Program farmer	0.0333***		0.0390***		0.0043	
	(0.003)		(0.004)		(0.004)	
Program vereda		0.0082***		0.0150***		-0.0010
		(0.002)		(0.003)		(0.003)
Observations	5,829	5,829	4,581	4,581	3,023	3,023
Farmer controls	Yes	Yes	Yes	Yes	Yes	Yes
Border (1km) FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors (cluster municipality) in parentheses *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. Unit of observation is a farmer. Cross section created with first 12 months of the PIC data. For (the relatively few) farmers that have more than one plot, the location is determined by the coordinates of the biggest plot. Farmer controls include plot area, age and elevation. The dummy *Program Farmer* takes value one for farmers that have at least one plot being part of the Program at the time for which the PIC data used in the table are available. The dummy *Program vereda* takes value one for farmers that have their biggest plot in an eligible vereda. Note that information of sales from the PIC data does not allow to track whether the coffee is sold under the Program and, for (the relatively few) farmers with multiple plots, the plot where the coffee is grown. Very few farmers have Program and non-program plots.

Table 7: **The Program - Quality Price Premium at the Export Gate**

	(1)	(2)	(3)	(4)	(5)
	Export Gate Price per Kg (log)				
Program Batch	0.1862***	0.2002***	0.1059***	0.0870***	
	(0.020)	(0.023)	(0.029)	(0.027)	
High Quality Batch					0.1088***
					(0.009)
Observations	53,218	13,118	3,341	3,341	22,309
Sample	I	II	III	III	IV
Contract conditions	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes
Origin - Year FE	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses (clustered at the year-month-destination) *** p<0.01, ** p<0.05, * p<0.1. The Table estimates the Program premium at the export gate (η). The unit of observation is a coffee batch at the export gate. The sample includes the universe of export transactions from FNC over the period 2006-2013. Those account for over a third of all coffee exports for Colombia. The data allow us to trace the coffee to the export gate and to precisely measure quality. Program batch is a dummy taking value equal to one for batches exported under the Program. High Quality Batch is a dummy taking value equal to one for batches of *supremo* coffee, i.e., the coffee exported by the Program. Column (1) includes all the coffee transactions (Sample I). Column (2) includes only the coffee exported from the two program regions, Cauca and Nariño (Sample II). The program only sources supremo coffee (see Table A1 for a description). Column (3) includes only exports of supremo coffee from those regions and Column (4) includes additional quality controls on that sample (Sample III). Contract controls include quantity, exchange rate, port of departure and terms of payment. Quality controls include cup test, bean size and defects. Finally, Column (5) explores the price premium associated with *supremo* coffee exported during 2006-2008, i.e., before the beginning of the Program (Sample IV). This sample is used to estimate the market price premium for quality used in the calibration exercise.

Table 8: **Model Parameters**

	Reduced Form Estimates	
ω	Increases in Production	
	ω^P	25%
	ω^U	12%
	ω^E	38%
θ_i, χ		[0.66 − 1] , 0.9
π	Program farm gate price premium	10%
η	Program FOB price premium	20%
	Directly Observed in the Data	
L_i	Farm Size Distribution	Ha
p_w	FOB price for standard coffee	3.75 USD/Kg
τ	Transport cost	1.4
	Information from Agronomists	
c	Variable cost for standard coffee	0.68 USD/Kg
γ	Additional program variable cost	16.7%
	Estimates from Multinomial Logit	
σ	Fixed Cost Scale Parameter	**0.888**
α	Demand-Side Value	**21%**

Notes: The Table reports the parameters used to calibrate the model. The vector ω is estimated from the RDD in Table B4. The ITT impact in Column (2) is a weighted average of *i)* the ωs associated with each investment decision and *ii)* the Program shifting the share of farmers investing. Regressing farmers' production on farmers' investment decisions among takers reveals that farmers renewing $D = R1$ (resp. expanding) have 12% (resp. 37%) higher production relative to farmers that simply join the program (Column 6). We assume $\omega^{R1} = \omega^{N1} + \omega^{R0}$ (resp. $\omega^{ER1} = \omega^{N1} + \omega^{ER0}$). Using the observed share of farmers' decisions as weights yields $\omega^{N1} = 25\%$.
The potential production function $Q_i = \theta(A_i) \times L_i^n$ is estimated from (unreported) OLS regressions on the sample of non-eligible farmers.
The price premium $\pi = 10\%$ is estimated in Section 3, Table 5.
The price premium at the Export Gate $\eta \approx 19\%$ is estimated from transaction-level data at the export gate in Table 7.
The average world price for Colombian milds during the sample period was $p^W = 3.75\$$. The average FNC base price over the sample period was $p^S = 2.68\$$. This yields transport and intermediation cost $\tau = 1.4$.
The cooperatives agronomists shared detailed information on variable cost figures. Over the sample period, the average (deflated) harvesting and processing unit costs for standard coffee were estimated by the agronomist to be $c^S = 0.68\$$. The additional harvesting and processing unit costs to produce program coffee is $\gamma = 0.167\%$ higher.
The identification of fixed costs scale σ and demand value α stems from both the model's structure and from functional form assumptions. Both parameters are estimated from a multinomial logit model of farmers' take-up and investment choices (see Table A8 for details). The coefficient of Q_i on the likelihood of choosing $D \in \{R0, RE0\}$ is the increase in returns relative to the baseline decision $D = N0$, i.e., $\beta^D = \frac{\mu^S \times \omega^D}{\sigma}$ for $D \in \{R0, RE0\}$. Knowledge of $\mu^S = p^S - c^S$ and ω^D thus identifies σ from β^D. Knowledge of σ in turn identifies α. For $D \in \{N1, R1, RE1\}$ we have $1 + \alpha = \frac{\beta^D \times \sigma + \mu^S}{\mu^Q \times (1+\omega^D)}$.
The model is overidentified. In principle we could thus estimate (some of) the ωs from farmers investment decisions. This is however not needed since we can anchor the estimates for ωs to the ITT estimates in the spatial discontinuity design in Section 3. We therefore estimate the scale parameter σ imposing the linear constraint $\beta^{RE0} = \beta^{R0} \times \omega^{R0} / \omega^{R0}$. Given the numerous fixed effects, imposing the corresponding constraints for decisions $D \in \{N1, R1, RE1\}$ often yields a non-concave log-likelihood function. We thus estimate the model approximating the linear constraints and estimate separate αs for each decision in $D = N1, R1, RE1$. We iterate until we find an approximation that yields αs within one-digit from each other and then set α to be the (weighted) average of the estimated ones.
The investment decision of non-takers identifies σ and is thus critical to identify α. Unreported results show that a simpler model based on the binary take-up decision alone yields quantitatively similar results. The simpler model doesn't separately identify α and σ and thus precludes some of the counterfactual analysis.
An alternative identification strategy illustrates the economic intuition. Exogenous increases in world prices shift returns from cultivation and could be used to identify the costs of renewal and expansion. Given these estimated costs α can be identified from program take-up. Such a strategy requires that the estimated fixed costs carry over to the population of interest: eligible farmers at the time of eligibility. Our identification strategy gives up exogenous world price increases but estimates fixed costs directly from the population of interest.

Table 9: **A Quantitative Exploration of the Sustainable Quality Program**

Panel A: Calibration & Counterfactuals

	Baseline Estimates	**Counterfactuals**								
		Supply Side			*Demand Side*			*Market Structure*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Parameters		ω=0	α=0	ω=0 α=0	λ=1	η=0.1	λ=1 & η=0.1	Monopsonist	Monopsonist + Enforceable Contracts	Competitive Market
λ (estimated)	0.695 [0.685,0.705]	0.695	0.695	0.695	1	0.695	1	1	1	0
α (estimated)	0.21 [0.185,0.215]	0.21	0	0	0.21	0.21	0.21	0	0	0
ω (observed)	25%*	0%	25%	0%	25%	25%	25%	0%	0%	0%
η (observed)	19%*	19%	19%	19%	19%	10%	10%	10%	10%	10%
Outcomes										
π (observed)	10%*	14%	13.5%	16%	4.5^{+}%	4.5^{+}%	4.5^{+}%	4.5^{+}%	0.5%	14%
% Quality	59%	35%	38%	17%	49%	48%	48%	9%	7%	15%
Take-Up (T=1)	41%	28%	28%	16%	34%	34%	34%	11%	9%	15%
$\triangle$ Farmers Π T=1	17%	12%	32%	23%	14%	14%	14%	17%	16%	21%
$\triangle$ Farmers W	19%	9%	9%	4%	13%	13%	13%	2%	1.5%	4%
$\triangle$ Chain Surplus	33%	15%	17%	6%	29%	20%	20%	3%	2.5%	4%
% Surplus Farmers	56%	61%	53%	62%	46%	66%	66%	62%	54%	100%

Panel B: Fit

	Take-Up (P = 1)	**% in D | (P = 1)**			**% in D | (P = 0)**		
		Nothing	Upgrade	Expand & Upgrade	Nothing	Upgrade	Expand & Upgrade
Data	39%	36%	42%	21%	63%	24%	13%
Model	41%	32%	45%	23%	64%	23%	13%

Notes: Panel A of this Table presents both the baseline estimates and the counterfactual results. In Column 1 (*) refers to program parameters that are estimated outside the multinomial logit model. Those are the Program's impact on production $\omega = 25\%$ (estimated from results in Table B4); the Program premium at the export gate $\eta = 19\%$ (estimated in Table 7); and the Program's price premium at the farm gate $\pi = 10\%$ (estimated in Table 5).
The scale parameter σ and the Program's demand-side benefits α are estimated in the multinomial logit model. The model is then used to simulate farmer-specific fixed costs F_i^D for each different investment decisions D. The aggregation of farmer choices derives the supply function. Given the supply function, λ is then identified as the parameter that rationalizes the estimated farm gate price premium $\pi = 10\%$. Confidence intervals for both λ and α are obtained by bootstrapping the entire model 100 times. In each round we exclude 5% of the observations to estimate the multinomial logit.
In the remaining Columns 2 - 10, ω and η are exogenous parameters that are changed across counterfactuals, while π is an endogenous outcome. In the counterfactual, the symbol (+) next to the endogenous value for π indicates that the farmer incentive constraint is binding. In Column 9, *enforceable contracts* refers to a counterfactual in which the farmer incentive constraint is removed.
The Program is represented as a vector of four parameters. On the *supply* side, the Program increases production by $\omega = 25\%$ and provides demand-side benefits to farmers $\alpha = 21\%$. On the *demand* side, the Program entails an objective function with weight $\lambda = 0.695$ on exporter's profits and an estimated export gate price premium $\eta = 19\%$. We are interested in three sets of counterfactuals: on the *supply* side, on the *demand* side, and on the *market structure*. Simulated fixed costs are held constant across counterfactuals.
The Program farm gate price premium π is endogenously chosen and thus varies across counterfactuals alongside the other Program's outcomes. The main outcomes of interest are the % of quality production, the Program's take-up, the increase in takers profits and farmers welfare, the increase in the chain surplus and how the increase is shared between farmers and the exporter.
Panel B reports the multinomial model fit. The Panel reports both the overall fit in matching observed Program's take-up rate as well as the entire vector of investment decisions by takers and non-takers. **P** = 1 indicates Program take-up. Conditional on the take-up decision, D refers to the farmers investment choices in the multinomial logit model.
The model is overidentified. In principle we could thus estimate (some of) the ωs from farmers investment decisions. This is however not needed since we can anchor the estimates for ωs to the ITT estimates in the spatial discontinuity design in Section 3. We therefore estimate the scale parameter σ imposing the linear constraint $\beta^{RE0} = \beta^{R0} \times \omega^{R0}/\omega^{R0}$. Given the numerous fixed effects, imposing the corresponding constraints for decisions $D \in \{N1, R1, RE1\}$ often yields a non-concave log-likelihood function. We thus estimate the model approximating the linear constraints and estimate separate αs for each decision in $D = N1, R1, RE1$. We iterate until we find an approximation that yields αs within one-digit from each other and then set α to be the (weighted) average of the estimated ones.

Table 10: **Comparing VSSs - Price and Quality Along the Chain**

	(1)	(2)	(3)	(4)	(5)
	Farm Gate	Mill Gate		Export Gate	
	Price (ln)	Price (ln)	Quality	Price (ln)	Quality
Program	0.0959***	0.0660***	0.4586***	0.2186***	1.2512***
	(0.008)	(0.004)	(0.035)	(0.015)	(0.088)
Environmental Certification *EC*	0.0117***	0.0317***	0.3727***	0.0472***	0.0155
	(0.003)	(0.003)	(0.026)	(0.006)	(0.028)
Social Certification *SC*		0.0149***	0.3239***	0.0900***	0.0009
		(0.004)	(0.046)	(0.006)	(0.021)
Sample		All	All	All	All
Observations	34,888	213,252	113,210	53,218	53,218
R^2	0.676	0.930	0.483	0.911	0.435

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Column (1) replicates the specification of Column (1) in Table 5, where unit of observation are farmer sales in a program-year-buying point. In that specification, the dummy Program takes value equal to one for *sales* under program. The dummy "*EC*", instead, takes value equal to one for *farmers* that hold the Environmental Certification. As noted in Section 3, however, Program farmers sell more than 90% of their produce to the Program and so the difference in farm gate price premium is still substantial. Sample difference with Table 5 is due to farmers missing information on certification status.
In Columns (2) and (3) the unit of observation is a coffee batch entering an Almacafé mill. Column (2) replicates the specification of Columns (1) in Table A7, and Column (3) follows the specification of Column (1) in Table 4. The quality index is computed as the average of the (negative) of a standardized index of bean defects and bean size. Here the identifiers are at the batch level.
Columns (4) and (5) replicate the specification of Table 7, where the unit of observation is a batch of coffee at the export gate. The quality measure is the screen size of export beans.

Appendix - Tables

Table A1: **Commercial Classification of Colombian Beans**

Supremo	Screen 17, high grade washed arabica, often specified with further details.
Excelso	Type 'Klauss': Screen 16.5 for Germany Type 'Europa': Screen 15 for France, Spain, Italy (Tolerance: 2.5% of beans between screens 12 and 15) Type 'Scandinavia': Screen 14 for Nordic countries
Usually Good Quality (UGQ)	'Usually Good Quality': Screen 14 for the US (Tolerance: 1.5% of beans between screens 12 and 14)
Caracol	Screen 12 (Tolerance: 10% of flat beans)

Notes: The Table reports a description of Quality Classification for Colombian exports of green coffee. The Sustainable Quality Program only sources **Supremo** coffee. *Source*: FAO Food and Quality Standards.

Table A2: **Program vs. Non Program Localities**

Panel A: Veredas

	(1) *Non-program ver.*		(2) *Program ver.*		*T-test*
	N	Mean/SE	N	Mean/SE	(1)-(2)
Agro-climatically attainable yield 1961-90, kg DW/ha	286	1036.185 (21.875)	312	1070.715 (22.898)	-34.529
Altitude (mean), metres	306	1794.774 (31.394)	335	1874.720 (29.756)	-79.945*
Altitude (std), metres	299	154.820 (6.164)	342	171.118 (5.042)	-16.299**
Slope (mean), degrees	301	16.541 (0.408)	340	18.300 (0.285)	-1.759***
Slope (std), degrees	302	6.444 (0.119)	339	7.009 (0.101)	-0.565***
Terrain ruggedness index (mean)	299	77.883 (1.894)	342	86.955 (1.380)	-9.072***
Terrain ruggedness index (std)	301	26.796 (0.628)	340	30.072 (0.602)	-3.276***

Panel B: Municipalities

	(1) *Non-program mun.*		(2) *Program mun.*		*T-test*
	N	Mean/SE	N	Mean/SE	(1)-(2)
Official Area (Km^2)	36	455.250 (89.469)	33	374.788 (63.005)	80.462
Altitude	36	1747.139 (124.334)	33	1683.727 (112.956)	63.412
Distance to district capital	36	53.229 (4.365)	33	53.062 (3.585)	0.167
Rurality Index (Rural/ Total Population)	36	0.750 (0.031)	33	0.767 (0.029)	-0.017
Poverty Index (SISBEN)	36	93.295 (2.721)	32	97.587 (0.430)	-4.292
Land Gini Index	36	0.769 (0.015)	32	0.739 (0.013)	0.029
Land Gini Index (Ownership)	36	0.754 (0.013)	32	0.716 (0.009)	0.037**
Literacy rate in 2005	36	84.589 (1.842)	32	85.435 (0.794)	-0.847
Index of soil agricultural suitability	35	2.895 (0.277)	32	2.393 (0.176)	0.502
Coffee cultivation 1997 (thsds. hectares)	32	1.484 (0.307)	29	1.181 (0.155)	0.303
Presence of coca cultivation	36	0.944 (0.222)	33	0.879 (0.212)	0.066
Presence indigenous population (1535-1540)	36	0.361 (0.081)	33	0.758 (0.076)	-0.396***
Spanish occupied land (1510 - 1561)	36	0.361 (0.081)	33	0.273 (0.079)	0.088
Presence of land conflicts (1901 - 1917)	36	0.056 (0.039)	33	0.061 (0.042)	-0.005
Presence of land conflicts (1918 - 1931)	36	0.083 (0.047)	33	0.121 (0.058)	-0.038
Violence 1948 to 1953	36	0.139 (0.058)	33	0.061 (0.042)	0.078
Presence of ELN	36	0.194 (0.078)	33	0.152 (0.063)	0.043
Presence of FARC	36	1.167 (0.146)	33	0.424 (0.115)	0.742***
Guerrilla Massacres	32	0.125 (0.059)	29	0.069 (0.048)	0.056
Paramilitary Massacres	32	0.875 (0.317)	29	0.172 (0.100)	0.703**

Notes: Panel A reports information on *terroir* conditions at the vereda level. There are two definitions of veredas: the DANE (Colombian Government) definition and the FNC definition. Each DANE vereda includes roughly 1.5 FNC ones. Shape files for FNC veredas borders are not available. The analysis in this table uses the DANE borders definition. The main analysis in the paper, however, uses the finer FNC definition (with no material impact on any of the results). Panel B reports information on socio-economic characteristics at the municipality level (Source: CEDE Database, Universidad d los Andes). Corresponding information at the vereda level is not available. The variables on land distribution, poverty, coca presence and armed groups presence are the mean for the 2012-2014 period. The incidence of conflict (Masacres) is the average for the 2000-2005 period. Differences in numbers of municipalities across variables are due to missing information. Program status defined as municipalities where the program had expanded by 2014.

Table A3: **Program Expansion in Cauca and Nariño**

	Municipalities	Veredas	Plots (Takers)	Farmers (Takers)
2006	–	–	–	–
2007	10	106	1082	621
2008	26	655	12112	7631
2009	27	672	13002	8151
2010	31	935	21230	14244
2011	32	954	22459	14932
2012	32	970	25084	16640
2013	32	990	27571	17998
2014	33	1011	28296	18240
2015	33	1015	28711	18748
2016	33	1027	29629	19862

Notes: The Table shows the expansion of the Program in the Cauca and Nariño regions. In these two regions the Program started being rolled out in 2007 in initially 106 veredas belonging to 10 municipalities. By the end of the sample period in 2016, the Program had been rolled out to 1027 veredas in 33 municipalities. By 2016, 29629 plots, owned by 19862 individual farmers, had joined the Program. Plots with missing values for any of the main variables (2%) and/or that appear for one year only during the sample period (0.3%) have been excluded.

Table A4: **Parallel Trends and Dynamic ITT - Estimates**

	(1)	(2)	(3)	(4)
	Cauca and Nariño Departments		**Cauca Department Only**	
	Plot Upgrading Index	Coffee Planted Area (Ln)	Plot Upgrading Index	Coffee Planted Area (Ln)
Lag 5+	0.0368	-0.0148	0.0106	-0.0025
	(0.163)	(0.048)	(0.189)	(0.061)
Lag 4	0.0281	0.0150	0.0379	0.0123
	(0.084)	(0.030)	(0.101)	(0.040)
Lag 3	0.0064	-0.0112	0.0166	-0.0056
	(0.054)	(0.021)	(0.064)	(0.027)
Lag 2	-0.0044	-0.0149	0.0141	-0.0099
	(0.038)	(0.015)	(0.045)	(0.019)
Lag 1	-0.0209	-0.0127	0.0048	0.0016
	(0.019)	(0.007)	(0.018)	(0.008)
Post 1	0.0341**	0.0175***	0.0424**	0.0253***
	(0.016)	(0.006)	(0.019)	(0.008)
Post 2	0.0762***	0.0292***	0.0962***	0.0416***
	(0.031)	(0.010)	(0.036)	(0.013)
Post 3	0.1006**	0.0324**	0.1326***	0.0438**
	(0.044)	(0.013)	(0.051)	(0.017)
Post 4	0.1468***	0.0411**	0.1904***	0.0570***
	(0.058)	(0.016)	(0.067)	(0.021)
Post 5+	0.2305***	0.0496**	0.3371***	0.0802***
	(0.078)	(0.021)	(0.097)	(0.029)
Observations	775,263	775,263	343,841	343,841
Number of plots	91,766	91,766	39,788	39,788
Plot and Mun-Year FE	Yes	Yes	Yes	Yes

Notes: Robust standard errors (cluster vereda) are reported in parentheses *** p<0.01, ** p<0.05, * p<0.1. The unit of observation is a plot. The dependent variables are the plot upgrade index (the average of yearly standardized (negative) of average tree age on the plot and share planted trees with rush resistant varieties) and the area of the plot planted with coffee (in logs). The lag and post variables are dummies taking value equal to 1 in the corresponding year relative to when the plot vereda becomes eligible to the Program.
The first two columns report the results including both Cauca and Nariño departments and correspond to Figure 3 in the main text. The Program expansion started in 2007, one year after the beginning of our panel in 2006. Only 5.3% of the veredas are observed five years before becoming eligible for the Program. The vast majority of those are in the Cauca department, where most of the Program expansion started at a later date. Columns (3) and (4) repeat the exercise restricting the sample to plots in the Cauca department only and correspond to Figure A5 in the Appendix.
De Chaisemartin and D'Haultfeuille (2019) note that DID designs with period and group fixed effects identify weighted sums of average treatment effects (ATEs) in each group and period with weights that may be negative and propose a correction. Unreported results show that the negative weights sum only to 0.16. The correction identifies dynamic coefficients close to those in this Table.

Table A5: **The Mechanisms of Plot Upgrading**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Plot Quality Score and its Individual Components								
	Plot Quality Score			**Average Tree Age**			**% Resistant Varieties**		
Program Plot (OLS, Takers)	0.1862***		0.2422***	-0.9948***		-1.7230***	0.0840***		0.0921***
	(0.009)		(0.022)	(0.078)		(0.195)	(0.004)		(0.007)
Will be Takers			0.1315***			-1.3846***			0.0347***
			(0.020)			(0.185)			(0.007)
Never Takers (Eligible)			0.0066			-0.3121*			-0.0100+
			(0.020)			(0.177)			(0.007)
Program Vereda (ITT)		0.0478**			-0.5996***			0.0067	
		(0.020)			(0.179)			(0.007)	
Plot and Mun-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Plot controls	No	No	No	No	No	No	No	No	No
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel B: Other plot characteristics and improvement works								
	Tree Density (per Ha)			**% Area Plot Shaded**			**Improvement activity**		
Program Plot (OLS, Takers)	-54.0122***		-5.4679	0.0051*		0.0094+	-0.0144***		-0.0066**
	(8.107)		(15.478)	(0.003)		(0.006)	(0.002)		(0.003)
Will be Takers			44.6253***			0.0108*			0.0289***
			(15.326)			(0.006)			(0.004)
Never Takers (Eligible)			54.0242***			0.0000			-0.0065**
			(13.233)			(0.005)			(0.003)
Program Vereda (ITT)		46.3852***			0.0023			-0.0021	
		(13.174)			(0.005)			(0.003)	
Plot and Mun-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Plot controls	No	No	No	Yes	Yes	Yes	No	No	No
Observations	775,263	775,263	775,263	775,263	775,263	775,263	775,263	775,263	775,263
Number of plots	91,766	91,766	91,766	91,766	91,766	91,766	91,766	91,766	91,766

Notes: Robust standard errors (cluster vereda) in parentheses. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. Time period is 2006 to 2016. Unit of observation is plot-year. The variable *"Program plot"* takes value 1 at the moment a plot enters the Program. The variable *"Program vereda"* is a treatment indicator that takes value 1 for any plot in the vereda when the first plot in the vereda joins the Program, i.e. when the vereda becomes eligible. The variable *"Will be takers"* takes value 1 when the plot has not yet joined the Program but does before the end of the sample (2016). The variable *"Never takers eligible"* takes value one for plots that are in Program veredas but do not join the Program before the end of the data available. Plot characteristics included in Columns (4) to (6) of Panel B are plot age and density. Plot improvement activity is an indicator of the flow of improvement activities, that denotes the share of farmer land that is improved in a given year.

The Table provides a comprehensive picture of the plot upgrading process. The Table considers a wider vector of outcomes and delves into the mechanics of quality upgrading by parting apart the vereda eligibility indicator (ITT) into *i)* plots already in the Program [*takers*]; *ii)* plots that will join the Program [*will-be takers*]; and *iii)* eligible plots that do not join the Program before the end of the data available [*never takers*]. The results must be interpreted cautiously since the status of each plot at a given point in time is an endogenous choice. Taken together, the results in the Table paint a highly heterogeneous quality upgrading process in which certain plots must be upgraded *before* they (can) join the Program; other plots that are already intensively cultivated switch towards quality by decreasing coffee tree density while other plots respond by intensifying coffee cultivation. Results also show some limited replanting activities among never takers, possibly reflecting access of these plots to the program support for plot renewal.

Panel A considers the plot quality index and its components. Columns (1) and (2) replicate the OLS and ITT specifications while Column (3) shows that the upgrading index is higher for takers and will-be takers alike. Our preferred interpretation is that plots must be upgraded *before* they (can) join the Program. Columns (4) to (9) consider the two components of the index (age and share of rust resistant varieties) separately. Besides confirming the results from the index analysis, Column (6) and (9) show some limited replanting activities among never takers, possibly reflecting access of these plots to the Program support for plot renewal. For both components of the index, however, the estimated coefficient for never takers is only one fifth of the coefficients estimated for takers and will-be takers.

Panel B considers additional outcomes. A higher density of trees increases yield per hectare but, beyond a certain point, compromises quality as trees compete for nutrients. The appropriate direction of upgrading, then, depends on plot specific characteristics. Plots that are already intensively cultivated and are switching towards quality should *decrease* density, as found by the OLS specification in Column (1). The program, however, also provided support for plot renewal and incentivized farmers' to cultivate more intensely. Accordingly, the ITT estimates in Column (2) show a positive effect on tree density. Column (3) reconciles the two estimates showing that increases in density are concentrated in will be takers and never takers plots. Estimated coefficients are however small relative to average (and, according to extensionists, optimal for the local conditions and varieties) tree density which is around 5000 trees/Ha. Columns (4) to (6) show a modest increase in the share of shadow cultivated coffee for takers. Besides characteristics of the capital stock invested on the plot, Columns (7) to (9) focus on upgrading work. The results confirm that upgrading activity mostly takes place after the vereda becomes eligible but before the plot joins the Program.

Table A6: **Participation in FNC programs**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Individual extension		Extension program		Credit program		ID program	
Program farmer	0.0283***		-0.0365***		0.0042		-0.0872***	
	(0.003)		(0.004)		(0.006)		(0.002)	
Program vereda		-0.0023		0.0044		0.0131		-0.0020
		(0.008)		(0.006)		(0.011)		(0.004)
Mean dep. var.	0.9462		0.9134		0.5623		0.9072	
Observations	300,133	300,133	300,133	300,133	300,133	300,133	300,133	300,133
Number of Farmers	55,604	55,604	55,604	55,604	55,604	55,604	55,604	55,604
Farmer and Mun-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses (cluster vereda) *** p<0.01, ** p<0.05, * p<0.1. The specifications are as in Equation (1) but at the farmer, rather than plot, level and, due to data availability, cover a different sample period (2007-2012). The sample includes all plots in the municipalities where the Program eventually expanded. At the end of the panel 97% of the plots in these municipalities were eligible. The Dependent Variable is a dummy that takes value 1 if the farmer participated in a Program in a given year. The *"Individual extension"* dummy takes value 1 when the farmer had a one-to-one activity with the extension services. The *"Extension program"* dummy takes value one when farmer has participated in any group or individual extension Program. *"ID program"* refers to the FNC program to ensure all farmers had an ID that allowed them to do monetary transactions with the cooperative and keep track of the programs they are involved and their benefits.

Table A7: **The Program - Price Premium at the mill gate**

	(1)	(2)	(3)	(4)	(5)
	Price per Kg (log)				
Program batch	0.0629***	0.0460***	0.0463***	0.0447***	0.0382***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
Sample	All	(I)	(II)	(III)	(III)
Quality control	No	No	No	No	Yes
Observations	213,252	122,481	44,808	26,025	26,025
R^2	0.929	0.938	0.950	0.939	0.945
Origin - Year FE	Yes	Yes	Yes	Yes	Yes
Mill and Coop FE	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses (cluster origin-year). *** p<0.01, ** p<0.05, * p<0.1. The Table estimates the price premium paid by Almacafé mill to the implementing cooperative using batch-level specifications as in equation (2). The unit of observation is a transaction, a coffee batch entering the mill. Almacafé pays to the implementing Cooperatives 4 to 5% price premium for the Program coffee. At the corresponding mill gate price level these estimates imply a premium of $\approx$ 450 Col per Kg, nearly identical to the premium paid to farmers. There is thus no double marginalization along the domestic chain. Sample All includes all batches sourced by any Almacafé mill. Sample (I) only includes non-standard batches, i.e., coffee sourced under any Voluntary Sustainability Standard. Sample (II) only includes non-standard batches sourced from the Cauca and Nariño regions in which the Program was implemented. Sample (III) further restrict the sample to batches sourced from the implementing cooperatives. Quality controls include the *quality index*, the z-score of grams of healthy beans in sample and the negative of the grams of beans with broca and qualifying as pasilla (subproduct) in sample. The quality index is available only for the 2009-2014 and, therefore, for clarity of comparison the sample is restricted and held constant in Columns (4) and (5).

Table A8: **Farmer's Take-Up and Investment Decisions**

	(1) $\mathbf{I_{P=1}}$		(2) $\mathbf{I_{D=R0}}$	(3) $\mathbf{I_{D=E0}}$	(4) $\mathbf{I_{D=N1}}$	(5) $\mathbf{I_{D=R1}}$	(6) $\mathbf{I_{D=E1}}$
L_i	1.8331***		2.6753***	0.2206*	1.8295***	3.6204***	2.0510***
	(0.065)		(0.096)	(0.123)	(0.102)	(0.096)	(0.119)
A_i	-0.2559***		-0.7655***	-0.3291***	-0.3641***	-0.6798***	-0.3753***
	(0.030)		(0.043)	(0.052)	(0.044)	(0.044)	(0.054)
$A_i \times L_i^2$	0.0215		0.0688**	-0.0605	0.0318	0.0330	-0.0538+
	(0.020)		(0.032)	(0.044)	(0.039)	(0.030)	(0.037)
L_i^2	-0.3696***		-0.5818***	0.0013	-0.4324***	-0.7029***	-0.3452***
	(0.027)		(0.043)	(0.055)	(0.051)	(0.041)	(0.050)
Constrained	0.0876***		-0.0000	-20.0557	0.0112	0.4948***	-19.9833
	(0.027)		(0.038)	(570.273)	(0.039)	(0.038)	(587.031)
Multi	0.7020***		0.0887**	-0.1424***	0.9741***	0.6420***	0.3550***
	(0.025)		(0.037)	(0.048)	(0.037)	(0.036)	(0.047)
Observations	32,413	32,427	32,427	32,427	32,427	32,427	32,427

Notes: Robust standard errors in parentheses (cluster vereda) are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1, + p<0.15.

The Table explores farmers take-up and investment decision. The cross-sectional specifications cover the sample of all farmers in eligible veredas and are informed by patterns documented in Figure A9. All independent variables are defined at the time the farmer's vereda becomes eligible for the Program: A_i is a dummy for young plantation (average age $\leq$ 6 years) and L_i plantation size. *Multi* (resp. *Constrained*) is a dummy taking value = 1 if the farmer owns multiple plots (resp. coffee is grown on $\geq$ 90% of the plot's area). The specifications include municipality fixed effects and regional-specific cohort effects. Column (1) reports results from a logistic regression of the take-up decision. Columns (2) - (6) report the results from a multinomial logit regression of the six decisions D_i. Decision D_i is defined considering farmers choices in the three years after the farmer's vereda becomes eligible for the Program. A farmer is considered to take up ($P_i = 1$) if she has joined the program within three years of becoming eligible. Similarly, we use the plot panel and define a farmer to have renewed (expanded) if she has undertaken any upgrading work on (expanded the area under coffee cultivation in) the plot within three years of becoming eligible. The model is actually calibrated using estimates from a more flexible specification in which the fixed costs F_{ivmr}^D for decision D for farmer i in vereda v in municipality m of region r are given by $F_{ivmr}^D = \Phi^D(L_{ivrm}, A_{ivrm}) + \sum_v \gamma_v^{D,L} \times L_{ivrm} + \sum_v \gamma_v^{D,A} \times A_{ivmr} + \sum_v \gamma_v^{D,R} \times \gamma_r + \gamma_{mr} + \delta X_{ivrm} + \epsilon_{ivrm}^D$ where $\Phi^D()$ is a decision-specific quadratic function of plot's size L_{ivrm} interacted with dummies for the age's plantation A_{ivrm}; $\gamma_v^{D,\cdot}$ are cohort- and decision-specific dummies, γ_{mr} (γ_r) are dummies for the municipality (region) and X_{ivrm} are the *multi* and *constrained* dummies defined above. In addition, the model substitutes L_i with potential production Q_i.

Table A9: **Model Calibration: Robustness to Alternative Scenarios**

	Baseline	ROBUSTNESS							
	γ=0.167, c=0.68, τ=1.4, pw=3.75	γ=0.20	γ=0.225	c=0.75	c=1.02	τ=1.5	τ=1.6	pw=3.25	pw=4.25
λ (estimated)	0.695	0.705	0.71	0.71	0.745	0.725	0.765	0.71	0.685
α (estimated)	21%	23%	24%	22%	24%	22%	22%	22%	22%
π (observed)	10%	10%	10%	10%	10%	10%	10%	10%	10%
% Quality	59%	59%	59%	59%	59%	59%	59%	59%	59%
Take-Up (T=1)	41%	41%	41%	41%	41%	41%	41%	41%	41%
△ Farmers T=1	17%	16%	15%	17%	16%	17%	17%	17%	18%
△ Farmers W	19%	18%	18%	18%	18%	18%	18%	18%	18%
△ Chain Surplus	33%	33%	33%	33%	36%	35%	37%	34%	32%
% Surplus Farmers	56%	56%	56%	55%	52%	53%	50%	55%	57%

Notes: The Table explores the robustness of the model's estimates to alternative baseline parameters. We consider higher increases in the variable costs to produce quality (from $\gamma = 16.7\%$ at baseline to $\gamma = 22.5\%$); higher variable costs (from $c^S = 0.68$ to $c^S = 1.02$); higher intermediation costs for Program coffee (from $\tau = 1.4$ to $\tau = 1.6$); and alternative scenarios of world prices p^W (from 3.25 to 4.25, with baseline of 3.75).

The results are robust to these broad changes to the baseline scenario. The estimated λ ranges from 0.685 to 0.765 while α ranges from 21% to 24%. The estimated increases in takers' profits ranges from 15% to 18%; the increase in the chain surplus from 32% to 37% and the share of that increase accruing to farmers from 52% to 57%.

The results are robust to alternative scenarios because the variables that drive the outcomes of interest, namely take-up and price premia along the chain, are all anchored to the data. First, the model is estimated targeting, and almost exactly replicates, the observed vector of farmers' investment and take-up decisions (Panel B, Table 9). Changes in baseline parameters are thus compensated by changes in estimated fixed costs so that actual take-up and quality upgrading match those observed in the data.

Alongside take-up and volumes, price premia at the farm's and export gate drive the increase in surplus created by the program and how it is shared between the exporter and the farmers. These price premia are also anchored to the data, and estimated to be $\pi = 10\%$ and $\eta = 19\%$. This explains why the main results are robust to alternative scenarios for costs (c^S, γ and τ) and world prices p^W.

Appendix - Tables BDD robustness

Table B1: **Balance of Farmers' Characteristics Across Borders**

VARIABLES	(1) Total area	(2) Num. plots	(3) Ownership	(4) Altitude	(5) Longitude	(6) Latitude
	Panel A: With Border Fixed Effects					
Program vereda	0.0753	0.0098	0.0195	19.2162	0.1828	0.0367
	(0.045)	(0.052)	(0.024)	(14.301)	(0.170)	(0.173)
Observations	5,696	5,696	5,696	5,696	5,696	5,696
R^2	0.171	0.129	0.208	0.751	0.998	0.997
Border FE	Yes	Yes	Yes	Yes	Yes	Yes
	Panel B: Without Border Fixed Effects					
Program vereda	0.0416	0.0768	-0.0560	-34.5020	3.7527	1.0626
	(0.077)	(0.069)	(0.052)	(35.069)	(9.957)	(12.999)
Observations	5,696	5,696	5,696	5,696	5,696	5,696
R^2	0.000	0.001	0.002	0.004	0.001	0.000
Border FE	No	No	No	No	No	No

Notes: Robust standard errors (cluster municipality) in parentheses *** p<0.01, ** p<0.05, * p<0.1. The unit of observation is a farmer. The table compares farmers that have plots on both sides of a border that separates an eligible from a non-eligible vereda. The Table shows that there is no significant difference between farmers with plots within 1 Km of the border. "Plot ownership" is a dummy that takes value equal to one if the farmers has a legal property right on the plot.

Table B2: **(No) Side-Selling and Prices: Robustness**

	(1)	(2)	(3)	(4)	(5)	(6)
	Robustness 1: Do not include border FE					
	Sell Coop	Sell others	Share Coop	ln price all	ln price Coop	ln price others
Program vereda	0.0329+	-0.0483+	0.0416*	0.0150***	0.0186***	0.0042
	(0.022)	(0.033)	(0.025)	(0.005)	(0.006)	(0.006)
Observations	5,829	5,829	5,829	5,829	4,581	3,023
Farmer controls	Yes	Yes	Yes	Yes	Yes	Yes
Border FE	No	No	No	No	No	No
	Robustness 2: Exclude between-municipalities borders					
	Sell Coop	Sell others	Share Coop	ln price all	ln price Coop	ln price others
Program vereda	0.1293**	-0.0433	0.1100*	0.0088*	0.0183***	-0.0012
	(0.056)	(0.059)	(0.059)	(0.005)	(0.005)	(0.006)
Observations	2,398	2,398	2,398	2,398	1,953	1,041
Farmer controls	Yes	Yes	Yes	Yes	Yes	Yes
Border 1Km FE	Yes	Yes	Yes	Yes	Yes	Yes
	Robustness 3: Different Farmer Controls					
	Sell Coop	Sell others	Share Coop	ln price all	ln price Coop	ln price others
Program vereda	0.0444+	-0.0359+	0.0401+	0.0076***	0.0142***	-0.0012
	(0.028)	(0.023)	(0.026)	(0.003)	(0.003)	(0.003)
Observations	5,829	5,829	5,829	5,829	4,581	3,023
Farm area and elevation	Yes	Yes	Yes	Yes	Yes	Yes
Border 1Km FE	Yes	Yes	Yes	Yes	Yes	Yes
	Robustness 4: Multi-plot farmer - Assign to centroid					
	Sell Coop	Sell others	Share Coop	ln price all	ln price Coop	ln price others
Program vereda	0.0436+	-0.0372*	0.0419*	0.0085***	0.0150***	0.0019
	(0.027)	(0.022)	(0.023)	(0.002)	(0.003)	(0.003)
Observations	5,706	5,706	5,706	5,706	4,479	2,952
Farmer controls	Yes	Yes	Yes	Yes	Yes	Yes
Border 1Km FE	Yes	Yes	Yes	Yes	Yes	Yes
	Robustness 5: Multi-plot farmer - Assign to plot closer to border					
	Sell Coop	Sell others	Share Coop	ln price all	ln price Coop	ln price others
Program vereda	0.0520*	-0.0341+	0.0502**	0.0083***	0.0141***	0.0014
	(0.026)	(0.021)	(0.024)	(0.002)	(0.003)	(0.003)
Observations	5,590	5,590	5,590	5,590	4,402	2,902
Farmer controls	Yes	Yes	Yes	Yes	Yes	Yes
Border 1Km FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors (cluster municipality, except in Robustes 1, cluster border) in parentheses *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. Unit of observation is a farmer. Cross section created with first 12 months of the PIC data. All specifications include farmers within 1K from the border. Results are robust to including farmers at different distances (unreported). In robustness 4, farmers that have more than one plot are located at the coordinates of the biggest plot. Farmer controls include plot area, share resistant varieties, sun exposure, plantation density, age and elevation.

Table B3: **(No) Farmer Spillovers**

	(1) **ln Quantity**	(2) **Share Coop.**	(3) **ln price (all)**	(4) **ln price (Coop)**
Program Farmer (OLS, Takers)	0.7495***	0.1760***	0.0299***	0.0381***
	(0.081)	(0.028)	(0.003)	(0.004)
Will be Takers	0.0373	0.0088	-0.0080	-0.0014
	(0.132)	(0.044)	(0.005)	(0.007)
Never Takers (Eligible)	-0.1204	-0.0283	-0.0039	-0.0011
	(0.073)	(0.030)	(0.003)	(0.004)
Observations	5,829	5,829	5,829	4,581
R^2	0.381	0.287	0.316	0.297
Farmer Controls	Yes	Yes	Yes	Yes
Border FE	Yes	Yes	Yes	Yes
Border Distance	1K	1K	1K	1K

Notes: Robust standard errors (cluster municipality) in parentheses *** p<0.01, ** p<0.05, * p<0.1. Unit of observation is a farmer. Cross section created with first 12 months of the PIC data. The variable *"Program Farmer"* is an indicator that takes value 1 for farmers that have joined the Program. The variable *"Will be Takers"* takes value 1 for eligible farmers that will join the Program at a later date before the end of the sample period in 2016. The variable *"Never Taker"* takes value 1 for farmers in eligible veredas that had not joined the Program by 2016. Farmer controls include log plot area, average age of tree, density, share planted with resistant varieties, share of trees grown under shadow, and altitude.

Table B4: **Other Outcomes: Production and Upgrading**

	(1) **Quantity**	(2) **Produced**	(3) **(log)**	(4) **Plot Upgrade Index**	(5) **ln Land Under Coffee**
Program Farmer	0.6205***				
	(0.054)				
Program Vereda		0.1358***		0.1060**	0.0494**
		(0.048)		(0.043)	(0.023)
Expand			0.3831***		
			(0.089)		
Renew			0.1337+		
			(0.085)		
Observations	5,829	5,829	790	8157	8157
Farmer controls	Yes	Yes	Yes	N.A.	N.A.
Border FE	Yes	Yes	No	Yes	Yes
Border Distance	1Km	1Km	N.A.	1Km	1Km
Vereda FE	N.A.	N.A	Yes	N.A	N.A
Takers only	N.A.	N.A.	Yes	N.A.	N.A.

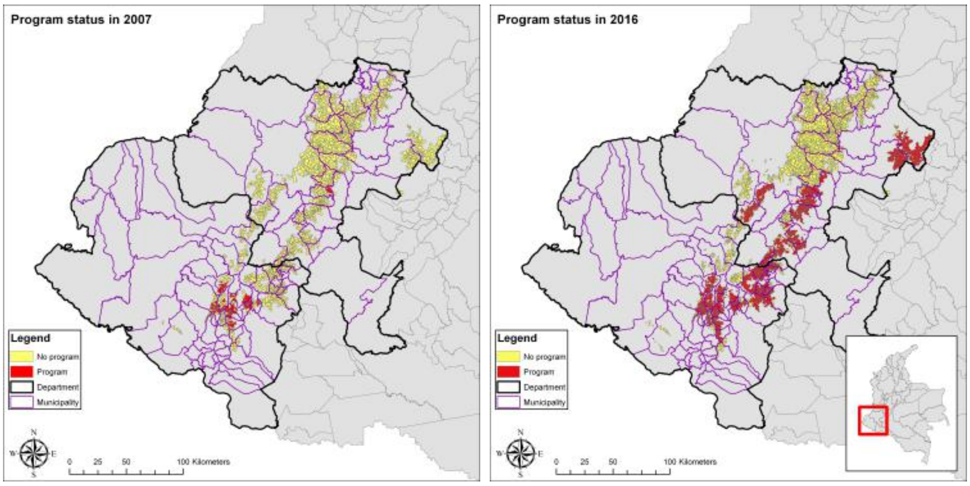
Notes: Robust standard errors (cluster municipality for columns 1 to 5) in parentheses *** p<0.01, ** p<0.05, * p<0.1, + p<0.15. Unit of observation is a farmer (Columns (1) to (3)) and plot in Columns (4) and (5). The sample in Columns (1) to (3) is based on the PIC data. In Column (3) the sample is restricted to program takers within the border. In Columns (4) and (5) the data are from the panel of plots.

Figures appendix

Figure A1: **Prices at Buying Points: an Illustration**

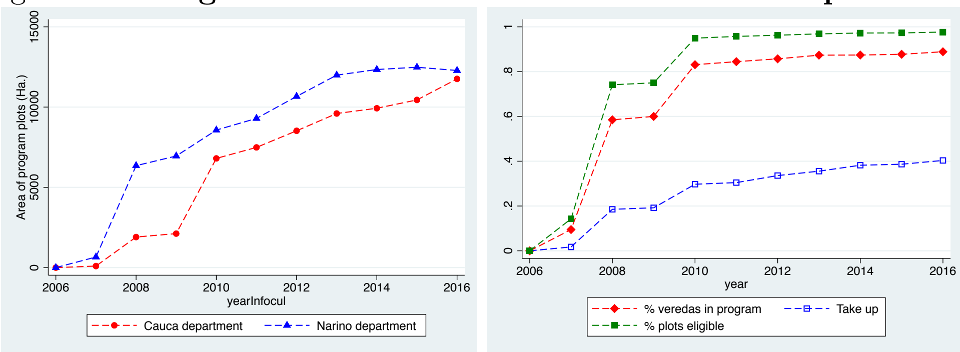
The Figures illustrate how the base price and the Program's price premium are announced to farmers. All cooperatives buying points around the country post the weekly base price (left panel). The base price is established for humidity <12% and conversion factor from parchment to excelso $\leq$ 94. These characteristics are well known to farmers and provide a very minimal quality standard met by essentially all coffee produced. The base price is adjusted to take into account regional differences in transport costs. Buying points located in regions that source for the Program add to the base price information on the prices paid by the program (right panel). In the right panel, the Program's price per kilo is 5320, relative to the FNC base price in that week of 4920. This fixed price premium of 50000 COL per carga (125kg) remained stable for most of the sample period. In the last two years of the sample, the price premium was increased to 600 Col per Kg for farmers that also have the Environmental label. Over the sample period, the Program's price premium represented a 10% price premium over the FNC base price.

Figure A2: **Program Expansion**



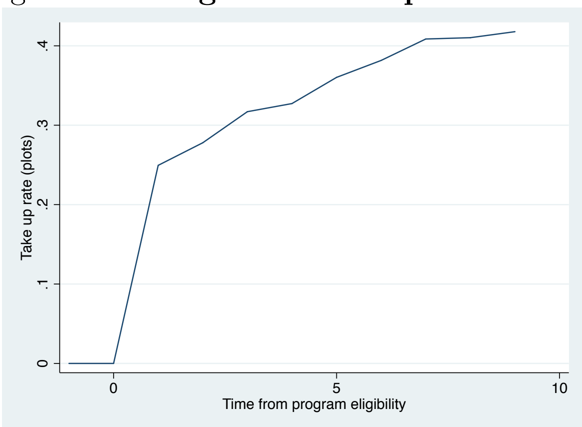
The maps show the Program veredas in 2007 (top) and 2016 (bottom). The Program expanded in the southern and oriental part of Cauca Department and the Nariño Department. As showed in Table A2, the program area is similar in terms of terrain characteristics and orientation, with potential to produce homogeneous quality and organoleptic properties coffee. The analysis focuses on the municipalities that enter the program. At the end of the panel, 88% of the veredas in municipalities where the program expanded are eligible, what represents 97.67% of the plots in these municipalities.

Figure A3: **Program Roll-Out in Cauca and Nariño Departments**



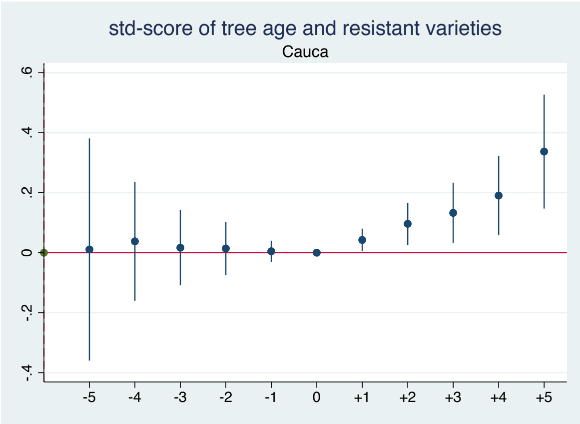
The Figure illustrates the expansion of the Program in the Cauca and Nariño departments between 2006 and 2016. The left figure plots the evolution of the land under the Program in each department. The variation in the Figure corresponds to the OLS specifications in the plot-panel analysis. The right Figure presents the expansion of the Program in terms of eligibility and take-up. The Figure thus captures the variation used in the ITT specifications on the plot-level panel. At the end of the panel, 88% of the veredas in municipalities where the Program expanded are eligible, what represents 97.67% of the plots in these municipalities.

Figure A4: **Program Take-Up Over Time**



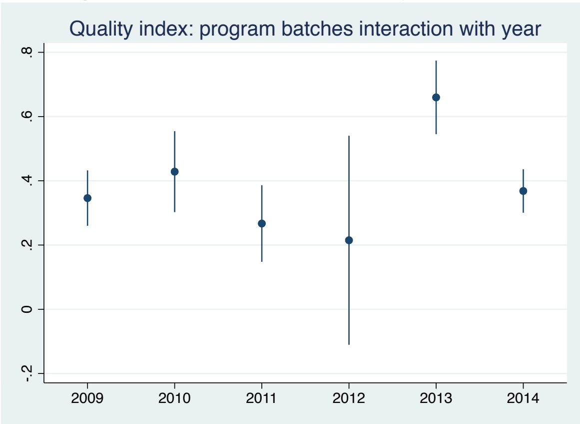
The Figure illustrates the Program take-up rate over time. The year zero is defined as the last year before the plot's vereda becomes eligible for the Program. The Figure shows that in the first year after becoming eligible approximately 25% of eligible plots take up the Program. The take-up rate keeps increasing and it stabilizes around 40% five years after eligibility. The dynamic patterns reflects *i)* the fact it might take some time for farmers to upgrade the plot to the required standards, and *ii)* some farmers might "wait and see" and learn from the experience of others how the Program works before incurring the costs of joining.

Figure A5: **Parallel Trends in Plot Upgrading (Cauca)**



The Figure repeats the exercise in Figure 3 and investigates pre-trends for the main outcomes of interest focusing on the Cauca region only. The bulk of the Program roll-out in Cauca happened in 2010. For this region, then, the available data cover more years in the pre-period. This allows for a more precise test for (the absence of) pre-trends.

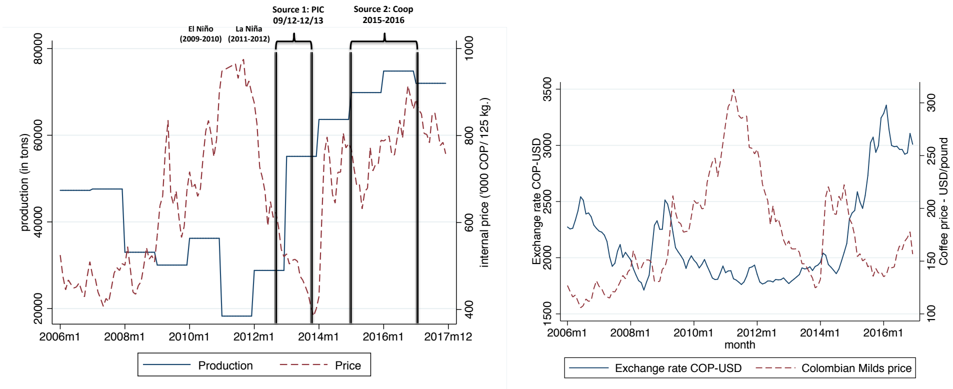
Figure A6: **Program Batches Quality Premia Over Time**



The Figure explores the stability of the Program's quality premium over time. The Figure reports coefficients β_y from the regression $Q_{bomy} = \beta_0 + \beta_y \times PB_{bomy} + \gamma_{oy} + \gamma_{my} + \varepsilon_{bomy}$ where Q_{bomy} is the quality index and PB_{bomy} a dummy taking value equal to 1 for coffee batches sourced under the Program. The unit of observation is a coffee batch tested at the mill gate. The Figure shows that the Program quality premium has remained constant over time with the notable exception of 2013, a year in which adverse meteorological conditions due to La Niña significantly compromised the availability of quality coffee in the market. Despite the lower availability of quality coffee, the Program didn't compromise on the required quality standard. This resulted in a *higher* quality of Program batches relative to the average coffee available in the market and in lower volumes sourced under the Program.

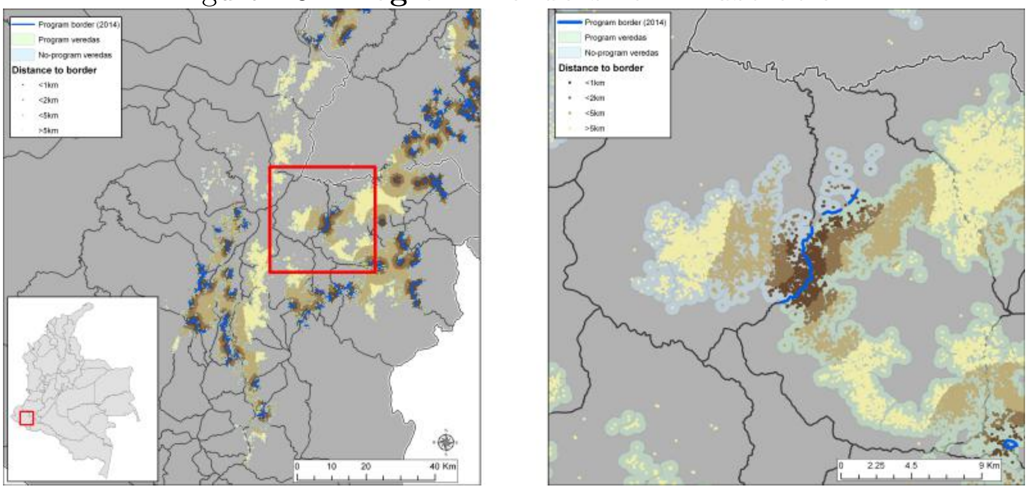
.

Figure A7: **Price and Quantity Volatility**



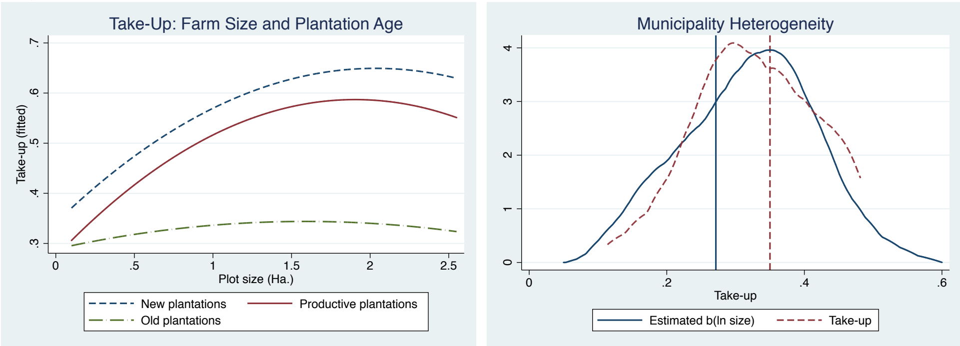
The left Figure presents prices and quantities for the Cauca department, and shows the periods of the data available on farmer sales. The right figure shows the exchange rate and Colombian Milds monthly prices (Source: International Coffee Organization).

Figure A8: **Program Borders: an Illustration**



The map shows the border of the Program veredas, and the bandwidth of plots at 1, 2 and 5km. from the border.

Figure A9: **Take-Up: by Municipality and Plot Characteristics**



The two Figures motivate the parametrization of fixed costs assumed ($\mu_{im} = \alpha_m + \theta_{a_i}(q_i)^2$). The left Figure shows that the relationship between plot size and take-up is inverted-U, and that it is rather different depending on whether the plantation is new (25th. younger quartile of the age distribution), in the productive age, or old (over 75th quartile). This motivates our distribution of the fixed costs to take a quadratic form in plot size and with coefficients that vary depending on the age of the plantation. The right Figure reports the distribution of take-up rates across municipalities and the municipality specific estimated β_{land} in the regression Take $\text{up}_{pm} = \alpha + \beta_{land} * \text{Land Size}_{pm} + \beta_{age} * \text{Age}_{pm} + \epsilon_{pm}$ for each program municipality. The Figure shows that there is substantial heterogeneity across municipalities in take-up rates as well as in the relationships between farm size and take-up rates. This heterogeneity motivates our approach in which we let the distribution of fixed costs to vary across municipalities.

Appendix A: Data sources

Our analysis relies on multiple sources of data that track coffee at different stages of the chain: at the plot and farmer level; at the point of sale between the farmer and the first intermediary in the chain; at the subsequent quality inspection points at the mill; finally, at the export gate.

Plot-Farmer level:

At the plot level, we exploit information from the Coffee Information System (SICA - *Sistema de Información Cafetera*). This is a continuously updated geo-referenced census of *all* plots cultivating coffee in Colombia. The census contains information on plots characteristics (e.g., location and size) together with information on the coffee plantation (number of trees, average age of trees, cultivated varieties) as well as on improvement that were made to the plot over the course of the year. For each plot we know the year in which it joined the Sustainable Quality Program. Our analysis focuses on the universe of coffee plots in the municipalities in which the program was implemented in the Cauca and Nariño departments over the 2006-2016 sample period.

The annual plot census contains a unique identifier for the farmer cultivating the plot. We can thus merge the panel with farmer level information, e.g., the farmer's participation in different FNC programs (technical training, credit, and other socio-economic programs). At any point in time, most farmers only farm one coffee plot. The average share of farmers with one plot only is 67%, and 91% of farmers have two plots or less. When needed, however, we aggregate plots to the farmer level. For example, data on coffee sales and on certain other FNC programs participation are available at the farmer (not the plot) level. In this case, we attribute to the farmer the location of his/her largest plot.[57]

Transactions along the chain

Farmer Sales: We match the plot and farmer level information to transaction-level data on farmers's sales. We obtain data on *all* sales of coffee from *all* farmers to *any* buyers in Colombia from the Farmer's Income Protection Program (PIC -

[57] In the Appendix we have shown that results of the farmer-level analysis are robust to allocating the farmer to the centroid of his/her plots, or to the plot closer to the vereda border.

Protección del Ingreso Cafetero, see Echavarria et al. (2017) for program details). The main advantage of this data is its comprehensive coverage. The data, however, has two drawbacks for our analysis. First, the PIC program was only implemented, and thus data availability is restricted to, the 2012-2013 harvest season.[58] Second, the dataset does not contain information on sales linked to any specific VSS, including the Sustainable Quality one. We are thus restricted to conduct the analysis at the farmer level. We complement the PIC data with detailed sourcing records from one of the two cooperatives implementing the Program. These records cover the period 2015-2016 and include sales by VSS, thus enabling us to explore panel specifications and comparing within-farmer sales across programs.

Batches at mill entry: We complement this information with records on coffee transactions at the mill level, i.e., one step down the chain. For all the Almacafé mills, including those that source the Program coffee as well as coffee for other VSSs, we have information on all purchases and sales of coffee at the transaction level for the 2007 - 2016 period. On the sourcing side, the data includes detailed information on the origin of the coffee, its price, quantity and, crucially, extremely detailed product description (including whether the coffee is part of any VSS and quality testing). So, while we do not observe quality directly at the farmer level, we can compare the quality of batches of coffee sourced at the same time from the same narrow locality.

Batches at export port Finally, we use transaction-level data at the export gate. Relative to standard transaction-level customs records, these data also include detailed product characteristics, quality testing and contractual terms (e.g., payment conditions). Note however that, due to the nature of coffee export contracts in Colombia, price information is available only for batches of coffee exported by FNC.

[58] The PIC program monitored sales to determine eligibility for income support transfers. To be conservative, we only use the first 12 months of the program since subsequent changes in its implementation raise concerns about the accuracy from private buyers records. Results are however robust to the inclusion of the complete program period.