The Business School for the World®

Working Paper

2021/57/TOM

When Doing Good May Backfire: Smallholder-Farmer Selection into Yield-Improvement Programs

Utku Serhatlı
Universidade NOVA de Lisboa, utku.serhatli@novasbe.pt

Guillaume Roels
INSEAD, guillaume.roels@insead.edu

Problem Definition: Large buyers of agricultural commodities (e.g., cocoa, coffee, hazelnuts) have spearheaded programs to improve smallholder farmers' yields. However, some farmers, especially those not enrolled in such programs, are concerned that yield improvements could result in a decrease in crop prices and negatively affect their profits. We analyze the implications of such programs on market prices and smallholder farmers' individual and aggregate well-being and identify farmer selection strategies into those programs that mitigate the potential conflict among those objectives.
Methodology/Results: We consider a Cournot competition model where farmers choose their planting areas under yield uncertainty. Farmers are differentiated in terms of their planting cost and yield. We analytically show that a) because yield-improvement programs push market prices down, they may decrease the profits of some farmers; b) the only possible farmer selection into the program that may not be harmful is to enroll all of them; c) selecting the lowest-cost farmers minimizes farmers' individual economic losses, performs well in terms of their aggregate well-being, and either maximizes or minimizes the crop price reduction, depending on whether the yield-improvement program targets a decrease in yield variability or an increase in average yield. Calibrating our model using industry data, we find that farmers' well-being is more sensitive than crop prices to farmer selection. Thus, the lowest-cost farmer selection performs well along the three objectives of crop price reduction, improvement in farmers' aggregate well-being, and minimization of farmers' individual losses.
Managerial Implications: Because this study formalizes smallholder farmers' concerns about the potential downsides of yield-improvement programs, it can help policy-makers assess the various trade-offs involved and guide buyers in their farmer selection strategy.

Keywords: Smallholder Farmers; Cournot Competition; Game Theory; Socially Responsible Operations

Electronic copy available at: http://ssrn.com/abstract=3938055

When Doing Good May Backfire: Smallholder-Farmer Selection into Yield-Improvement Programs

Utku Serhatlı
Nova School of Business and Economics, Universidade NOVA de Lisboa, 2775-405 Carcavelos, Portugal, utku.serhatli@novasbe.pt

Guillaume Roels
INSEAD, Boulevard de Constance, 77305 Fontainebleau, guillaume.roels@insead.edu

Problem Definition: Large buyers of agricultural commodities (e.g., cocoa, coffee, hazelnuts) have spearheaded programs to improve smallholder farmers' yields. However, some farmers, especially those not enrolled in such programs, are concerned that yield improvements could result in a decrease in crop prices and negatively affect their profits. We analyze the implications of such programs on market prices and smallholder farmers' individual and aggregate well-being and identify farmer selection strategies into those programs that mitigate the potential conflict among those objectives.

Methodology/Results: We consider a Cournot competition model where farmers choose their planting areas under yield uncertainty. Farmers are differentiated in terms of their planting cost and yield. We analytically show that *a*) because yield-improvement programs push market prices down, they may decrease the profits of some farmers; *b*) the only possible farmer selection into the program that may not be harmful is to enroll all of them; *c*) selecting the lowest-cost farmers minimizes farmers' individual economic losses, performs well in terms of their aggregate well-being, and either maximizes or minimizes the crop price reduction, depending on whether the yield-improvement program targets a decrease in yield variability or an increase in average yield. Calibrating our model using industry data, we find that farmers' well-being is more sensitive than crop prices to farmer selection. Thus, the lowest-cost farmer selection performs well along the three objectives of crop price reduction, improvement in farmers' aggregate well-being, and minimization of farmers' individual losses.

Managerial Implications: Because this study formalizes smallholder farmers' concerns about the potential downsides of yield-improvement programs, it can help policy-makers assess the various trade-offs involved and guide buyers in their farmer selection strategy.

Key words: Smallholder Farmers; Cournot Competition; Game Theory; Socially Responsible Operations

1. Introduction

Two billion people who grow agricultural commodities (e.g., cocoa, coffee, hazelnuts) depend on small-scale farms (FAO 2015). While small-scale farming is essential to many food supply chains, many smallholder farmers still live below the poverty level (World Bank 2016). One of the major contributors to their low income is their crops' low and variable yield due to adverse weather, poor soil conditions, and pest infestation.

To improve the livelihoods of smallholder farmers and ensure stable supply of the agricultural commodities, numerous food manufacturers and retailers—the main buyers of these commodities—have committed to yield-improvement programs. For instance, Ferrero, a large chocolate manufacturer, conducts field trips to hazelnut growing areas, introduces farmers to modern farming techniques, and provides them with equipment for mechanized farming and drying stations under the scope of the Ferrero Farming Values (FFV) program. Similarly, under its Forever Chocolate program, Barry Callebaut, a major cocoa processor, provides farmers with a productivity package that includes training on tree pruning techniques and the use of fertilizer. Farmers who are enrolled in such yield-improvement programs have reported higher average and less variable yields.

Yield improvement is an increasingly important pillar of various sustainability certifications. Today, many of these certifications are managed directly by the buyers themselves (e.g., Starbucks' C.A.F.E. Practices, Unilever's Sustainable Agriculture Code) to deliver impact and lower cost (Thorlakson 2018) in contrast to the early efforts, which were spearheaded by non-governmental organizations or industry-wide consortia (e.g., Fair Trade, UTZ). Even though yield improvement is, in principle, associated with various Sustainable Development Goals of the United Nations, e.g., No Poverty (SDG #1), this change in the governance structure of farmer certifications has been criticized for shifting their focus on yield improvement and stable supply to the detriment of farmers' well-being (Van Wassenhove and Pot 2021). Are yield-improvement programs hurting smallholder farmers' well-being?

Through extensive interviews of hazelnut growers (see Appendix C), we found that many farmers are quite skeptical about the benefit of the FFV program: Enrolled farmers certainly enjoy improved yields, but these improved yields increase the available crop in the market and drive down market prices. In fact, some yield-improvement practices, like grafting cocoa trees, are prohibited by some governments to avoid a sudden decrease in price and drive farmers out of business (Smith et al. 2014). Moreover, because of their high fixed cost, yield-improvement programs rarely reach every farmer and, therefore, must be selective. In the Black Sea region of Turkey (which produces more than 70% of the global hazelnut), only about $20,000$ farmers are registered into the FFV program per season (Ferrero 2018) out of a population of $440,000$ farmers (ZMO 2018). As another example, Barry Callebaut, who aims to alleviate 500,000 farmers out of poverty by 2025 under the Forever Chocolate program, provided only 71,972 farmers with access to coaching, inputs such as tools and seedlings, and finance in 2019-2020 (Callebaut 2020). Hence, yield-improvement programs, despite their good intents, might backfire by creating disparities among farmers.

In this paper, we study the potential trade-off between low and stable prices (or equivalently, large and stable quantities) and farmer well-being, as well as characterize how the farmer selection into the yield-improvement program affects this trade-off. Assuming that farmers adjust their planting area to their yield and market prices in the long run (Alizamir et al. 2019), we consider a stylized Cournot competition model where farmers produce a homogeneous crop and compete in the face of uncertain yield. Because of the differences in the available land area, access to technology, know-how, or traits of farmlands, farmers have different planting costs. As a result of their differentiated enrollment in the program, they also have different yields. To derive sharp analytical results, we separately consider the effects of a yield mean increase and a yield variance decrease. Our model makes a number of simplifying assumptions, such as ignoring quality differences among farmers, price speculation or political price-setting intervention, and supply chain intermediaries (e.g., governmental export authorities, local processing companies) as a first attempt to understand the nature of that trade-off.

From the perspective of large commodity buyers, we consider the following three objectives: (*i*) minimizing the mean of the market price, or equivalently, maximizing total output (as in Deo and Corbett 2009, Boyabatli et al. 2021, Guda et al. 2021) or minimizing the price variance (as in Hu et al. 2019, Parker et al. 2016), (*ii*) maximizing farmers' total profits (as in Hu et al. 2019, Tang et al. 2018, Zhou et al. 2021), and (*iii*) minimizing the farmers' largest individual economic loss induced by the yield-improvement program (in the same spirit as Tang et al. 2018, Hu et al. 2019). We leave the consideration of other objectives, such as maximizing product quality (Ruben and Zuniga 2011) or minimizing child labor (Cho et al. 2019) for future research. With these objectives in mind, we investigate the following research questions:

1. What is the impact of yield-improvement programs on market prices and farmers' profits?
2. What farmer selection strategy leaves no farmer worse off?
3. For a given number of farmers to be enrolled, which farmers need to be selected to minimize the mean or the variance of the market price, maximize farmers' total profit, or minimize the farmers' largest individual economic loss from the program?
4. To what extent can these three objectives be aligned?

Our study generates the following managerial insights:

1. While yield-improvement programs are often motivated by corporate social responsibility in a particular supply chain, they may backfire once we account for all farmers in a given ecosystem. As argued above, an increase in the total available crop, due to the improved yields of the enrolled farmers, pushes market prices down. This certainly hurts non-enrolled farmers and, potentially, enrolled farmers.
2. Enrolling all farmers is the only selection strategy that may not negatively impact anyone. Moreover, when the program aims at only increasing the mean yield (as opposed to reducing its variance), there is an additional requirement for ensuring that all farmers benefit; namely, that farmers should have similar planting costs.

3. Selecting the lowest-cost farmers into the program minimizes the farmers' largest individual economic loss from the program while performing well in terms of their aggregate well-being. Depending on whether the program aims at reducing the yield variability or increasing its mean, selecting the lowest-cost farmers either minimizes or maximizes the market price.
4. A buyer's multiple objectives might be conflicting. For programs that increase the mean yield, the objectives of minimizing prices and minimizing the farmers' largest individual economic loss are conflicting. For programs that reduce the yield variance, these two objectives are aligned, but they may conflict with the third objective of maximizing farmers' total profit. Thus, buyers may need to be willing to trade off market price for farmer well-being. However, we argue and confirm numerically that prices, unlike farmers' profits, are relatively insensitive to the farmer selection into the yield-improvement program. Hence, for a fixed number of enrolled farmers, selecting the lowest-cost farmers appears to be the best strategy, in terms of all farmers' well-being, while introducing minimal sacrifice in terms of price.

The rest of this paper is organized as follows. After positioning our work relative to the literature in §2, we present our model in §3. Then, we analytically characterize the effect of yield-improvement programs on prices, farmers' total profit, and farmers' individual economic losses and investigate the optimal farmer selection strategy when the program increases the mean yield (§4) or reduces the yield variance (§5). In §6, we check the robustness of our analytical findings using numerical experiments calibrated with industry data. We conclude our study and outline future research directions in §7. Our notation is summarized in Appendix A, the sources of our model calibration are presented in Appendix B, and details on the interviews appear in Appendix C. All proofs appear in an electronic companion.

2. Literature Review

Our research is related to the growing body of work on socially responsible operations (Tang and Zhou 2012). For conciseness, we will limit our review to the operations management literature on smallholder farmers in emerging markets; see Boyabatlı et al. (2021) for a forthcoming overview. The literature can be categorized along two dimensions, namely their scope of analysis and the type of assistance provided. In terms of scope, we distinguish studies that consider only a subset of farmers (e.g., farmers within the same supply chain) from those that consider all farmers in an ecosystem. This distinction matters because prices are typically considered to be exogenous in the former and endogenous in the latter, as is the case in this study. In terms of assistance, we restrict our review to three types of assistance, namely: access to market information, yield improvement, and financial help. Table 1 presents this classification.

First, we review the works that consider only a subset of farmers, e.g., farmers who work with a particular commodity buyer. Farmers can be provided with different types of assistance, including

Type of Assistance	**Subset of Farmers**	**All Farmers in Ecosystem**	
		Identical Farmers	**Non-Identical Farmers**
Access to Market Information	Parker et al. (2016), Levi et al. (2020), Liao et al. (2019)	Tang et al. (2015), Chen and Tang (2015)	Zhou et al. (2021)
Yield Improvement	de Zegher et al. (2019)	Tang et al. (2015), An et al. (2015), Xiao et al. (2020)	**This Study**
Financial Help	de Zegher et al. (2018), Akkaya (2017)	An et al. (2015), Alizamir et al. (2019), Guda et al. (2021)	Tang et al. (2018)

Table 1 Socially responsible operations with respect to smallholders

information to reduce price dispersion and improve market access (Parker et al. 2016, Levi et al. 2020, Liao et al. 2019); yield improvement that is enabled by contracts and new sourcing channels (de Zegher et al. 2019); or financial assistance, such as subsidies (Akkaya 2017) or reward policies (de Zegher et al. 2018). Zhang and Swaminathan (2020) and Boyabatlı et al. (2019) propose using operations research to optimize, respectively, planting schedule and crop rotation and, thus, to improve yields. We differ from these works by considering not only the farmers within the same supply chain, but also other farmers in that ecosystem. Contrary to this literature, we show that operational improvements might hurt farmers' well-being once we account for their negative externality on the rest of the ecosystem.

Next, we review the works that consider all farmers in an ecosystem, thereby assuming that prices are formed endogenously, e.g., through Cournot competition. Most of this literature, which we review next, considers identical farmers. An early reference on the use of Cournot competition under yield uncertainty to guide operational decisions—in the context of vaccine production—is Deo and Corbett (2009). Tang et al. (2015) study farmers' decision to adopt market information (which is free) and yield-improvement advice (which involves a fixed cost, as discussed in §1). They find that market information is beneficial, but yield-improvement advice may not be welfare-maximizing if its cost is too high. Similarly, we find that enrolled farmers may not always benefit from yield-improvement advice; in our study, however, this happens even if it is free to them. Chen and Tang (2015) characterize the value of public and private market information. They find that although both types of information stabilize prices, public information might not always benefit farmers. We identify similar trade-offs between farmers' well-being and the pursuit of low and stable prices. An et al. (2015) assess the benefits and downsides of aggregation of smallholder farmers in terms of yield improvement (increasing/stabilizing process yield) and financial benefit (reducing planting costs), among others. Very much like aggregation, yield-improvement programs create two classes of farmers (enrolled vs. non-enrolled) and might not always be beneficial to them. Xiao et al. (2020) endogenize the process of knowledge sharing and find that it has frictions unless a specific reward mechanism is put in place. Finally, Alizamir et al. (2019) and Guda et al. (2021) inform policy makers on the implications of farming subsidy and guaranteed support price schemes.

In contrast, there is limited literature that specifically accounts for farmer heterogeneity in an ecosystem, as we review next. Zhou et al. (2021) consider heterogeneity in access to market information (private signals). They find that providing assistance (in the form of information provision) may be detrimental to some farmers. We obtain a similar result when the assistance takes the form of yield improvement. In contrast to their work, which assumes that farmers are *ex-ante* statistically identical, we consider farmers with different planting costs. Accordingly, we care not only about *how many* farmers to assist, but also about *which* farmers to assist. Tang et al. (2018) investigate the impact of input- and output-based farm subsidies on both farmer aggregate profit and income inequality when farmers have different yields. Our study is dual to theirs in the sense that they consider farmers who have *ex-ante* different yields and *ex-post* different costs/prices, whereas we consider farmers who have *ex-ante* different costs and *ex-post* different yields. Similar to them, we find that the objectives of aggregate profit and reducing inequalities may not always be aligned and may call for different types of intervention.

Our work differs from this literature on Cournot competition among farmers on two dimensions. First, we assume that farmers are heterogeneous in terms of both their planting costs and their access to the assistance program. Second, we consider the buyer as the primary decision-maker with the multiple objectives of lowering and stabilizing market prices, improving farmers' aggregate well-being, and reducing individual profit loss. Research in sustainability commonly adopts a multi-objective approach (Pacini et al. 2004, Falconer and Hodge 2001, De Koeijer et al. 2002), but often takes a societal perspective, unlike this study, which explicitly associates the different objectives with the different stakeholders in the supply chain or ecosystem (namely, the buyer and the farmers).

3. Model Description

We model the market dynamics between n farmers who sell a homogeneous crop in a Cournot market, as in An et al. (2015), Tang et al. (2015), and Alizamir et al. (2019). We consider a time horizon of one or a couple of seasons, e.g., a year. Prior to the implementation of the program, farmers are differentiated only by their costs per planting area over the time horizon, denoted by c_i. For fruits (e.g., coffee, hazelnut, cocoa), c_i might be defined as the yearly cost of growing a tree. For instance, a coastal hazelnut farm can be harvested with vacuuming harvester machines, whereas a mountainous hazelnut farm has to be harvested by manual labor, which is significantly more costly. Without loss of generality, we assume that farmers are ordered in increasing cost, that is, $c_1 \leq c_2 \leq ... \leq c_n$. In our analytical characterization, we assume that costs are such that all farmers are in business before and after the implementation of the yield-improvement program. We will formalize this assumption later and will relax it in our numerical experiments.

Farmers' outputs are subject to random yield due to variable climate conditions. As is common in the agricultural operations literature (Yano and Lee 1995, Kazaz 2004, Alizamir et al. 2019, Zhou et al. 2021), we employ a multiplicative random yield model. Specifically, farmer i observes yield θ_i, which is randomly distributed with mean $E[\theta_i]$ and variance $V[\theta_i]$. In the case of fruits, yield can be defined as the quantity obtained per tree. Hence, a farmer i planting q_i (e.g., fruit trees) harvests over the time horizon production quantity $q_i\theta_i$.

Let Y (Yes) and N (No) respectively denote the sets of enrolled and non-enrolled farmers in the yield-improvement program, with respective sizes n_Y and n_N. Because most crops are concentrated in particular geographical regions (e.g., Black Sea region of Turkey for hazelnuts), they tend to be exposed to similar yield risks (e.g., spring frosts, floods). Accordingly, we assume that all non-enrolled farmers are subject to the same yield, i.e., $\theta_i = \theta_N$, $\forall i \in N$. Moreover, because yield-improvement programs generally follow standard processes for yield-improving activities, we assume that all enrolled farmers are subject to the same yield, i.e., $\theta_i = \theta_Y$, $\forall i \in Y$. The yield-improvement program leads to a (weakly) higher average yield (e.g., through the adoption of fertilizers), i.e., $E[\theta_Y] \geq E[\theta_N]$, and a (weakly) lower yield variance (e.g., through the set up of irrigation systems to protect against droughts), i.e., $V[\theta_Y] \leq V[\theta_N]$. To understand the impact of these two effects, we analyze them separately. In this study, we ignore the quality differentiation that may come from the yield-improvement program, other than that which has a direct impact on quantity (e.g., higher shell-to-nut ratio), and leave this for future research.

Given Y, each farmer i strategically decides on planting area q_i.[1] Let $Q_Y = \sum_{i\in Y} q_i$ and $Q_N = \sum_{i\in N} q_i$ and $Q = Q_Y + Q_N$. Similarly, let $C_Y = \sum_{i\in Y} c_i$ and $C_N = \sum_{i\in N} c_i$ and $C = C_N + C_Y$. We denote the market-clearing price by $p(Q_Y, Q_N)$, which is assumed to linearly depend on the aggregate production, i.e.,

$$p(Q_Y, Q_N) = \alpha - \beta\theta_Y Q_Y - \beta\theta_N Q_N; \tag{1}$$

here, α denotes the upper limit of the price and β denotes the sensitivity of price to production quantity. (We ignore any price fluctuation that may originate from factors other than yield, e.g., speculation and political price setting.) Assuming that farmers receive the market price (or a fraction thereof, see Smith et al. (2014)), their profit is given by

$$\pi_i(Y) = E\left[p(Q_Y, Q_N)\theta_i q_i - c_i q_i\right] \qquad \forall i. \tag{2}$$

For brevity, we use the shorthand notation π_i when there is no ambiguity on the set Y. Moreover, farmer i's expected revenue per unit area is defined as $E[p(Q_Y, Q_N)\theta_i]$.

The sequence of events is shown in Figure 1, and the notation is summarized in Table A-1 in Appendix A. Throughout the paper, the overline denotes the average of the value within the set

[1] This assumption is valid in steady state, despite the short-term frictions associated with these strategic planting adjustments.

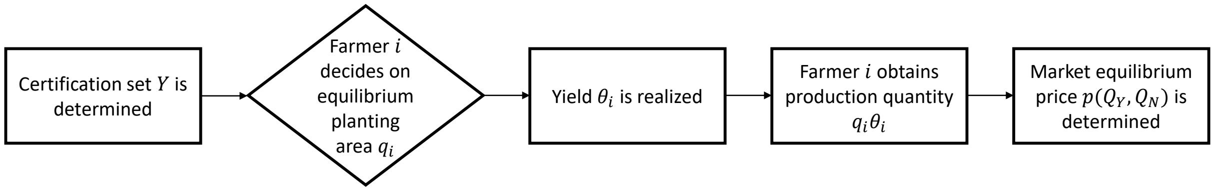

Figure 1 Sequence of the Cournot game under yield uncertainty

denoted in the subscript (For instance, $\overline{c_Y} = \frac{C_Y}{n_Y}$). For any set $Y \subseteq \{1,2,...,n\}$, the complement of set Y is defined as $Y^c \equiv \{1,2,...,n\} \backslash Y$. Furthermore, let Y^{LC} and N^{LC} respectively denote the set of enrolled and non-enrolled farmers when the enrolled farmers are selected to be the lowest-cost ones, i.e., $Y^{LC} = \{1,2,...,n_Y\}$, $N^{LC} = (Y^{LC})^c$. Similarly, let Y^{HC} and N^{HC} respectively denote the set of enrolled and non-enrolled farmers when the enrolled farmers are selected to be the highest-cost ones, i.e., $Y^{HC} = \{n-n_Y+1, n-n_Y+2, ..., n\}$, $N^{HC} = (Y^{HC})^c$. We use $\varnothing$ to denote the empty set.

We separately consider two aspects of yield improvement: increase of the mean yield in §4 and decrease of its variance in §5.

4. Increasing the Mean Yield

In this section, we consider programs that improve the mean yield. Without loss of generality, we consider an additive yield improvement. To keep the notation parsimonious, we assume that the yield variance is equal to zero for all farmers. Assumption 1 formalizes these two statements.

Assumption 1. *$E[\theta_Y] = \mu + b$ with $b \geq 0$, $E[\theta_N] = \mu$ and $V[\theta_Y] = V[\theta_N] = 0$.*

Given this setting, we first characterize the farmers' equilibrium planting areas for a given selection of farmers into the yield-improvement program. Then, we explore the impact of different farmer selections on the market price and on farmers' profits.

4.1. Equilibrium Characterization

We impose a sufficient condition to ensure that all farmers remain in business after the yield-improvement program is implemented.

Assumption 2. *For any sets Y and $N \equiv Y^c$, $\alpha\mu(\mu+b) + C_N(\mu+b) + C_Y\mu \geq c_n(\mu+b)(1+n)$.*

Assumption 2 is more likely to hold when farmers have similar planting costs, so that c_n (the highest cost) is not too different from the average cost and when their planting costs are low. In particular, when $c_i = c \;\; \forall i$, Assumption 2 simplifies to $\frac{\alpha\mu(\mu+b)}{\mu+b(1+n_Y)} \geq c$.

Lemma 1 characterizes the equilibrium planting areas and profits.

Lemma 1. *Suppose Assumptions 1 and 2 hold. Then, the equilibrium planting areas are equal to*

$$q_i^* = \frac{\alpha\mu(\mu+b)+C_N(\mu+b)+C_Y\mu - c_i(\mu+b)(n+1)}{\beta\mu^2(\mu+b)(n+1)} \qquad i\in N \tag{3}$$

$$q_i^* = \frac{\alpha\mu(\mu+b)+C_N(\mu+b)+C_Y\mu - c_i\mu(n+1)}{\beta\mu(\mu+b)^2(n+1)} \qquad i\in Y, \tag{4}$$

and the equilibrium profits are equal to

$$\pi_i^* = (q_i^*)^2\beta\theta_i^2 \qquad \forall i. \tag{5}$$

By Lemma 1, farmer i's equilibrium planting area decreases with its own cost c_i; hence, farmers who have lower costs choose to cultivate greater farm space. Moreover, because farmers' profits are quadratic in the planting costs, whereas planting quantities (and the market price) are linear, farmers' profits will be more sensitive than price to different farmer selection strategies. In the following subsections, we investigate the impact of improving the mean yields on market prices and farmers' economic well-being.

4.2. Impact on Price

From a buyer's perspective, one of the natural objectives of improving yields is to increase output, which, in turn, might lower market prices. First, we analyze the directional effect of these improvements on the market price. Then, motivated by the fixed nature of the cost of yield-improvement programs, we characterize the farmer selection strategy that minimizes the market price for a given number of enrollments.

4.2.1. Effect on Price. To assess the effect of yield improvement on price, we compare the market price before and after the program is implemented. Using Lemma 1, we derive the equilibrium market price by substituting the equilibrium production quantities, (3) and (4), into (1), which results in

$$p(Q_Y,Q_N) = \frac{\alpha\mu(\mu+b)+C_N(\mu+b)+C_Y\mu}{\mu(\mu+b)(n+1)}. \tag{6}$$

Proposition 1 characterizes the change in the equilibrium market price following the yield improvement of one additional farmer.

Proposition 1. *Suppose Assumptions 1 and 2 hold. For any sets Y and $N\equiv Y^c$, $|N|\geq 1$, improving the yield of farmer $j\in N$ leads to a decrease in the equilibrium market price, i.e. $p(Q_{Y\cup\{j\}},Q_{N\setminus\{j\}})\leq p(Q_Y,Q_N)$.*

Proposition 1 shows that the market price reduces as more farmers enroll in the yield-improvement program. Naturally, enrolled farmers obtain higher yield, so for any planting quantity, their production quantity is higher; this results in a decrease in the market price. However, farmers will strategically adjust their planting quantity (in the long run); in particular, non-enrolled farmers will clearly reduce their planting areas because the decrease in the market price reduces their marginal return per area planted. Even after accounting for all farmers' adjustments in planting areas, the total production is higher, and, in turn, the net effect of the yield-improvement program is to decrease prices.

4.2.2. Optimal Farmer Selection. Now that we have established that yield-improvement programs decrease the market price, Proposition 2 characterizes the farmer selection into the program, for a given number of enrollments, that minimizes the equilibrium market price.

PROPOSITION 2. *Suppose Assumptions 1 and 2 hold. For any given n_Y, selecting the highest-cost (resp., lowest-cost) farmers minimizes (resp., maximizes) the equilibrium market price, i.e., $p(Q_{Y^{HC}}, Q_{N^{HC}}) \leq p(Q_Y, Q_N) \leq p(Q_{Y^{LC}}, Q_{N^{LC}})$, $\forall Y$ and $N \equiv Y^c$ such that $|Y| = n_Y$.*

To interpret this result, we express the expected equilibrium production quantity of farmer i as

$$E[q_i\theta_i] = \frac{1}{\beta}\underbrace{\left(p(Q_Y, Q_N) - \frac{c_i}{E[\theta_i]}\right)}_{\text{unit profit margin}}. \tag{7}$$

Since farmer i's production quantity decreases in unit cost $\left(\frac{c_i}{E[\theta_i]}\right)$, the largest total production quantity (and, therefore, the lowest market price) is obtained by associating the highest yields with the highest costs, i.e., by selecting the highest-cost farmers.

Selecting the highest-cost farmers might also make farmers' production quantities more homogeneous. Although higher-cost farmers initially produce less, they may end up (after their yield has improved) planting and producing more than lower-cost, non-enrolled farmers. Also, since the highest-cost farmers face the most financial distress, prioritizing them into the yield-improvement program raises the minimum income level across all farmers. Raising the minimum income level is crucial to protect smallholder farmers against market price volatility. In the late 1990s, coffee prices plummeted, due to global increase in coffee production, deeply affecting high-cost coffee farmers in Central America (Kilian et al. 2006).

4.3. Impact on Farmers' Profits

In addition to reducing market prices, buyers might also be concerned about farmer well-being. In this section, we focus on two objectives: (i) minimizing farmers' individual potential profit losses and (ii) maximizing their total profits. First, we analyze the effect of the yield-improvement program on each farmer's profit. Then, again motivated by the fixed-cost nature of yield-improvement programs, we investigate the optimal farmer selection into the program with respect to these two objectives. Finally, we discuss whether or not these objectives can be aligned.

4.3.1. Effect on Farmers' Profits. We start by analyzing the effect of yield improvement on farmers' individual profits. The following proposition shows the impact of improving the yield of an additional farmer on farmers' profits.

PROPOSITION 3. *Suppose Assumptions 1 and 2 hold. For any sets Y and $N \equiv Y^c$, $|N| \geq 1$, and any $j \in N$, $\pi_i(Y) \geq \pi_i(Y \cup \{j\})$, $\forall i \neq j$.*

Proposition 3 shows that individual farmers do not want other farmers to have an improved yield. Each additional enrollment in the program leads to a price drop (by Proposition 1), which hurts the profit of all farmers, with the exception of the one being enrolled.

Next, Proposition 4 shows the total effect of yield improvement on farmers' profits.

PROPOSITION 4. *Suppose Assumptions 1 and 2 hold. For any sets Y and $N \equiv Y^c$, $|N| \geq 1$,*
(i) $\pi_i(Y) \leq \pi_i(\varnothing)$, $\forall i \in N$,
(ii) $\pi_i(Y) \leq \pi_i(\varnothing)$ if and only if $c_i(n+1) \leq C_Y$, $\forall i \in Y$.

The total effect of the yield-improvement program on farmers' individual profits depends on whether they are enrolled. On the one hand, non-enrolled farmers are worse off after the program is implemented because their output suffers price drop. On the other hand, the profit of an enrolled farmer can go either way, depending on the following factors:

- An increase in yield. The higher the farmer's cost, the higher the relative increase in its production quantity by (7).
- A decrease in the equilibrium market price. By (6), this decrease is smaller if (i) few farmers, out of the entire pool of farmers, are being enrolled and (ii) enrolled farmers have a relatively low cost, so that their increase in production quantity is relatively small.

Combining these two effects, an enrolled farmer is better off only if $c_i \geq \frac{C_Y - c_i}{n}$, i.e., if there are few other farmers enrolled, and they have a relatively low cost.

Hence, even an enrolled farmer might not benefit from the yield-improvement program if they have a low cost. Because the enrollment rate is currently low (e.g., 25% for cocoa, see Van Wassenhove et al. (2021)), we suspect that, unless there is a big disparity in planting costs (which might be the case for coffee and hazelnut, as we discuss in §6), most enrolled farmers benefit from an improved yield. But it may not be so in the future as these programs become more widespread.

Corollary 1 explores the possibility of an enrollment strategy that doesn't leave farmers worse off.

COROLLARY 1. *Suppose Assumptions 1 and 2 hold. No farmer is worse off if and only if (i) all farmers are enrolled and (ii) $c_i(n+1) \geq C$, $\forall i$.*

Corollary 1 states that it is very difficult, if not impossible, for a yield-improvement program to improve the well-being of all farmers. Since non-enrolled farmers are always worse off, enrolling every farmer is the only potential enrollment strategy that leaves no farmer worse off. And yet, this will be the case only if $c_i(n+1) > C$ for all i, i.e., no farmer should have a cost that is lower than $\frac{n}{n+1}$ of the average cost. Hence, even enrolling everyone does not guarantee higher well-being for all farmers—farmers' planting costs also need to be similar.

4.3.2. Optimal Farmer Selection. So far, we have established that yield improvement might hurt some farmers, namely, all non-enrolled farmers and the enrolled farmers that have a low cost ($c_i(n+1) < C_Y$). For any number of enrollments n_Y, Proposition 5 characterizes the farmer selection that minimizes the economic loss of those farmers.

PROPOSITION 5. *Suppose Assumptions 1 and 2 hold. For any given n_Y, selecting the lowest-cost farmers minimizes all farmers' individual economic loss from the yield-improvement program, i.e.,* $\max\{0, \pi_i(\varnothing) - \pi_i(Y^{LC})\} \leq \max\{0, \pi_i(\varnothing) - \pi_i(Y)\}$, $\forall i$, $\forall Y$ *and* $N \equiv Y^c$ *such that* $|Y| = n_Y$.

Because the production quantity of farmer i is linearly decreasing in its unit cost $\frac{c_i}{E[\theta_i]}$ by (7), selecting the lowest-cost farmers causes the smallest increase in the total production quantity and the smallest decrease in price (Proposition 2), which leads to the smallest drop in profit for all the farmers who are being hurt. As a result, the farmer selection that minimizes their largest individual economic loss (the lowest cost) is conflicting with the selection that minimizes the market price (the highest cost).

Considering the objective of maximizing the farmers' total profit, Proposition 6 shows, using an interchange argument, that selecting the lowest-cost farmers is locally optimal. Specifically, it shows that selecting a high-cost farmer, rather than a low-cost farmer, hurts every other farmer. Hence, if the yield-improvement program is gradually expanded, at any point in time, all currently enrolled and all non-enrolled farmers that are not considered for yield improvement would prefer a low-cost farmer to be selected into the program.

PROPOSITION 6. *Suppose Assumptions 1 and 2 hold. For any given n_Y, deviating from the lowest-cost farmer enrollment strategy by enrolling farmer k, for any $k \in N^{LC}$, instead of farmer j, for any $j \in Y^{LC}$, results in a drop in profit of any other farmer i, i.e., $\pi_i(Y^{LC} \cup \{k\} \backslash \{j\}) \leq \pi_i(Y^{LC})$, for any $i \neq k, j$.*

Despite this local optimality condition, the lowest-cost farmer selection may not (globally) maximize total farmer profits. (For a counter-example, consider a case where there are only two farmers in the market, and only one of them is to be enrolled: Total profit $\pi_1 + \pi_2$ is maximized by selecting farmer 1 if $2\alpha\mu(\mu + b) \leq (c_1 + c_2)(10\mu + 5b)$, and by selecting farmer 2 otherwise.) Therefore, the objective of maximizing farmers' total profit may not necessarily be aligned with the objective of minimizing the farmers' largest individual economic loss (which would prescribe to select the lowest-cost farmers) and certainly not aligned with the objective of minimizing the market price (which would prescribe to select the highest-cost farmers).

5. Decreasing Yield Variability

In this section, we consider programs that reduce yield variability while keeping the mean yield unchanged. For notational simplicity, we assume that the yield variance of the enrolled farmers is completely eliminated, that is:

Assumption 1' *$E[\theta_N] = E[\theta_Y] = \mu$ and $V[\theta_N] = \sigma \geq V[\theta_Y] = 0$.*

Next, we derive the equilibrium planting areas under Assumption 1'.

5.1. Equilibrium Characterization

Similar to Assumption 2, we set a sufficient condition to ensure that all farmers remain in business after the implementation of the program.

Assumption 2' *For any set Y, $\alpha\mu + C + C_N n_Y - c_n(1+n_Y)(1+n_N) \geq 0$.*

Similar to Assumption 2, Assumption 2' holds when farmers have relatively similar costs (so that the highest cost, c_n, is not too different from the average cost), and when their costs are low. In particular, when $c_i = c \ \ \forall i$, Assumption 2' simplifies to $\frac{\alpha\mu}{1+n_Y(n_N-1)} \geq c$. Lemma 2 characterizes the equilibrium planting areas and profits.

LEMMA 2. *Suppose Assumptions 1' and 2' hold. The equilibrium planting areas are equal to*

$$q_i^* = \frac{\sigma^2(-c_i(1+n_N)(1+n_Y) + C + C_N n_Y + \alpha\mu) + \mu^2(-c_i(1+n) + C + \alpha\mu)}{\beta(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y) + (1+n)\mu^2)} \qquad i \in N \tag{8}$$

$$q_i^* = \frac{\sigma^2(1+n_N)(-c_i(1+n_Y) + C_Y + \alpha\mu) + \mu^2(-c_i(1+n) + C + \alpha\mu)}{\beta\mu^2(\sigma^2(1+n_N)(1+n_Y) + (1+n)\mu^2)} \qquad i \in Y \tag{9}$$

and the equilibrium profits are equal to

$$\pi_i^* = (q_i^*)^2 \beta E[\theta_i^2] \qquad \forall i. \tag{10}$$

Similar to Lemma 1, we obtain that farmers' profits are quadratic in their planting costs while their equilibrium planting quantities and prices are linear. Therefore, profits will be more sensitive than prices to the farmers selected into the program. Next, we investigate the impact of yield improvement on market prices and farmers' well-being.

5.2. Impact on Price

We analyze the impact of yield improvement on both the mean and variance of the market price.

5.2.1. Effect on Price. From (1) and Lemma 2, the mean and variance of the market price can be expressed as

$$E[p(Q_Y,Q_N)] = \frac{\sigma^2(1+n_N)(\alpha\mu+C_Y)+\mu^2(\alpha\mu+C)}{\mu\left(\sigma^2(1+n_N)(1+n_Y)+\mu^2(1+n)\right)}, \quad \text{and} \tag{11}$$

$$V[p(Q_Y,Q_N)] = \left[\frac{\sigma n_N(-\overline{c_N}(1+n_Y)+C_Y+\alpha\mu)}{\sigma^2(1+n_N)(1+n_Y)+\mu^2(1+n)}\right]^2. \tag{12}$$

Proposition 7 shows the effect of improving the yield of an additional farmer on the mean and variance of the market price.

Proposition 7. *Suppose Assumptions 1′ and 2′ hold. For any sets Y and $N \equiv Y^c$, $|N| \geq 1$, improving the yield of farmer $j \in N$ leads to (i) a decrease in the expected equilibrium market price, i.e. $E[p(Q_{Y\cup\{j\}},Q_{N\setminus\{j\}})] \leq E[p(Q_Y,Q_N)]$ and (ii) a decrease in its variance, i.e., $V[p(Q_{Y\cup\{j\}},Q_{N\setminus\{j\}})] \leq V[p(Q_Y,Q_N)]$.*

To get intuition into the effect of yield improvement on the mean price, consider first the farmer whose yield has just been improved. Because this farmer faces no yield variability, they choose to plant more, which has the effect of dropping the market price. Next, consider all other farmers. Since they face lower prices, on average, they choose to plant less. Still, the net effect is an increase in the total planting area, resulting in a drop in price on average.

For the variance of price, there are two effects. On the one hand, the yield uncertainty of the focal farmer, whose yield has just been improved, is eliminated. On the other hand, non-enrolled farmers (who still face yield uncertainty) plant less, whereas enrolled farmers (who do not face yield uncertainty) plant more on aggregate, resulting in a shift in production toward the farmers that face no uncertainty. Each of these two effects contributes to a drop in the variance of the market price.

5.2.2. Optimal Farmer Selection. So far, we have established that yield improvement decreases both the expected value and the variance of the price. Next, we characterize the farmer selection that minimizes the mean and the variance of the market price for a given number of enrollments.

Proposition 8. *Suppose Assumptions 1′ and 2′ hold. For any given n_Y, selecting the lowest-cost (resp., highest-cost) farmers minimizes (resp., maximizes) the mean and variance of the equilibrium market price, i.e., $E\left[p(Q_{Y^{LC}},Q_{N^{LC}})\right] \leq E\left[p(Q_Y,Q_N)\right] \leq E\left[p(Q_{Y^{HC}},Q_{N^{HC}})\right]$ and $V\left[p(Q_{Y^{LC}},Q_{N^{LC}})\right] \leq V\left[p(Q_Y,Q_N)\right] \leq V\left[p(Q_{Y^{HC}},Q_{N^{HC}})\right]$, $\forall Y$ and $N \equiv Y^c$ such that $|Y| = n_Y$.*

Hence, in contrast to Proposition 2, selecting the lowest-cost farmers leads to the lowest prices when the program aims at reducing yield variability, as opposed to increasing the mean yield. To gain more insight into this result, we can express by (8), (9), and (11) the expected equilibrium production quantity of farmer i as

$$E[q_i\theta_i] = \frac{E[p(Q_Y, Q_N)\theta_i]E[\theta_i]}{\beta E[\theta_i^2]} - \frac{c_i E[\theta_i]}{\beta E[\theta_i^2]}. \tag{13}$$

Because the first term is identical across farmers in N and across farmers in Y, the total production quantity (summed across all farmers) is decreasing in the sum of ratios $\frac{c_i E[\theta_i]}{E[\theta_i^2]}$. Hence, the largest total production quantity is obtained by associating the highest second moment of the random yield ($E[\theta_i^2]$) with the highest planting costs (c_i); this is in contrast to the programs that increase the mean yield, where the association was determined in terms of $\frac{c_i}{E[\theta_i]}$ by (7). Selecting the lowest-cost farmers yields the lowest expected market price. Moreover, under the lowest-cost farmer selection, it is the largest farmers who have no yield variability; hence this selection also minimizes the variance of the market price.

Contrasting Propositions 2 and 8, we conclude that the farmer selection that achieves the most favorable market prices depends on the type of yield improvement. A program that improves the mean yield should target the highest-cost farmers, whereas one that reduces the yield variance should target the lowest-cost farmers.

5.3. Impact on Farmers' Profits

Next, we consider the impact of the yield-improvement program on the well-being of farmers, with the two objectives of minimizing the farmers' largest individual economic loss and maximizing farmers' total profits. We first discuss the effect of the yield-improvement program on farmers' individual profits and then characterize the design of yield-improvement programs to achieve these objectives.

5.3.1. Effect on Farmers' Profits. To assess the effect of yield improvement on farmer profit, we need to understand how farmers adjust their planting areas. Proposition 9 shows the change in farmer profits following the yield improvement of an additional farmer.

PROPOSITION 9. *Suppose Assumptions 1′ and 2′ hold. For any sets Y and $N \equiv Y^c$, $|N| \geq 1$,*
(i) $\pi_i(Y) \geq \pi_i(Y \cup \{j\})\ \forall i \in Y$.
(ii.1) $\pi_i(Y) \geq \pi_i(Y \cup \{j\})\ \forall i \in N \backslash \{j\}$ *if* $n_N \geq n_Y + 1$ *and* $c_j \geq \overline{c_N}$.
(ii.2) $\pi_i(Y) \leq \pi_i(Y \cup \{j\}),\ \forall i \in N \backslash \{j\}$ *if* $n_N \leq n_Y + 1$ *and* $c_j \leq \overline{c_N}$.

By Proposition 9(i), all enrolled farmers ($i \in Y$) see their profits decrease with each additional enrollment. Because enrolled farmers do not face yield variability,

$$\begin{aligned} E[p(Q_Y, Q_N)\theta_Y] &= \mu E[p(Q_Y, Q_N)] \\ &= \frac{\sigma^2(1+n_N)(\alpha\mu + C_Y) + \mu^2(\alpha\mu + C)}{\sigma^2(1+n_N)(1+n_Y) + \mu^2(1+n)}, \end{aligned} \tag{14}$$

and they care only about the mean market price. Since the mean market price decreases with each incremental enrollment by Proposition 8, enrolled farmers suffer from the addition of other farmers into the yield-improvement program.

By Proposition 9(ii.1)-(ii.2), the effect of the yield-improvement program on non-enrolled farmers ($i \in N$) is more convoluted since their expected revenue per unit area is affected by both the mean and the variance of the market price. By the law of total covariance, we obtain:

$$E[p(Q_Y, Q_N)\theta_N] = \mu E[p(Q_Y, Q_N)] - \sigma\sqrt{V[p(Q_Y, Q_N)]} \tag{15}$$

$$= \frac{\sigma^2(\alpha\mu + C + C_N n_Y) + \mu^2(\alpha\mu + C)}{\sigma^2(1+n_N)(1+n_Y) + \mu^2(1+n)}. \tag{16}$$

Similar to enrolled farmers, non-enrolled farmers suffer from lower market prices due to the incremental enrollment. However, contrary to enrolled farmers, they enjoy lower price uncertainty. The net result depends on the relative strength of these two effects. To assess it, notice that their expected unit revenue (16) increases with the average cost of the non-enrolled farmers ($\overline{c_N}$) and the number of enrolled farmers (n_Y) if $n_Y \geq n_N - 1$. In particular, if more than half of the farmers are enrolled ($n_Y \geq n_N - 1$), enrolling a relatively low-cost farmer j, with $c_j \leq \overline{c_N}$, benefits all non-enrolled farmers. In contrast, if the majority of the farmers are not enrolled ($n_Y \leq n_N - 1$), enrolling a relatively high-cost farmer j, with $c_j \geq \overline{c_N}$, hurts all non-enrolled farmers. This explains the two cases in Proposition 9(ii).

Proposition 10 shows the total impact of yield improvement on farmers' profits. Specifically, it shows that all enrolled farmers benefit from the yield-improvement program, whereas all non-enrolled farmers experience a profit decrease.

PROPOSITION 10. *Suppose Assumptions 1′ and 2′ hold. For any sets Y and $N \equiv Y^c$, $\pi_i(Y) \leq \pi_i(\varnothing)$ $\forall i \in N$ and $\pi_i(Y) \geq \pi_i(\varnothing)$, $\forall i \in Y$.*

To understand the driver of this result, consider the change in expected revenue per unit area before and after farmers in Y are enrolled. For enrolled farmers, by (14), it is equal to

$$E[p(Q_Y, Q_N)\theta_i] - E[p(0, Q)\theta_i] = \left[\frac{\sigma^2 n_N(1+n_N)}{1+n}\right]\left[\frac{\alpha\mu + C_Y - \overline{c_N}(1+n_Y)}{\sigma^2(1+n_N)(1+n_Y) + \mu^2(1+n)}\right]$$

$$= \frac{\sigma(1+n_N)}{1+n}\sqrt{V[p(Q_Y, Q_N)]}. \tag{17}$$

Because enrolled farmers have no yield uncertainty, their expected revenue is higher for every planting decision they make. Even if the other farmers adjust their planting decisions (resulting in a decrease in price), the yield uncertainty reduction effect dominates. In other words, the enrolled farmers' expected revenue per unit area becomes greater (given that (17) is non-negative). Hence, unlike programs that increase the mean yield (Proposition 4), enrolled farmers always benefit from programs that reduce the yield variance without any additional condition.

For non-enrolled farmers, by (16), the change in expected revenue per unit area is equal to

$$E[p(Q_Y,Q_N)\theta_i]-E[p(0,Q)\theta_i] = -\left[\frac{\sigma^2 n_N n_Y}{1+n}\right]\left[\frac{\alpha\mu + C_Y - \overline{c_N}(1+n_Y)}{\sigma^2(1+n_N)(1+n_Y)+\mu^2(1+n)}\right] = -\frac{\sigma n_Y}{1+n}\sqrt{V[p(Q_Y,Q_N)]}. \tag{18}$$

Non-enrolled farmers face lower, but less volatile prices on average. On aggregate, the effect of the price decrease dominates the effect of decrease in its variance (because (18) is non-positive). Therefore, non-enrolled farmers earn less because for any planting decision they make, their expected revenue is lower. In turn, they plant less and their profit decreases.

Since non-enrolled farmers are always worse off and enrolled farmers are always better off, enrolling everyone is the only strategy that benefits all farmers.

COROLLARY 2. *Suppose Assumptions 1′ and 2′ hold. No farmer is worse off if and only if all farmers are enrolled.*

Unlike programs programs that increase the mean yield (Corollary 1), there is no required condition on the cost of the enrolled farmers to ensure that no farmer is hurt. However, it remains improbable to implement a yield-improvement program that includes every farmer.

5.3.2. Optimal Farmer Selection. Proposition 11 characterizes the farmer selection, for a given n_Y, that minimizes the farmers' largest individual economic loss.

PROPOSITION 11. *Suppose Assumptions 1′ and 2′ hold. For any given n_Y, selecting the lowest-cost farmers minimizes all farmers' individual economic loss from the yield-improvement program, i.e.,* $\max\{0,\pi_i(\varnothing)-\pi_i(Y^{LC})\} \le \max\{0,\pi_i(\varnothing)-\pi_i(Y)\},\ \forall i,\ \forall Y$ *such that* $|Y| = n_Y$.

By Proposition 10, non-enrolled farmers are the only farmers hurt by the yield-improvement program. As discussed in our interpretation of their expected unit revenue (15), two effects are at work: (i) a decrease in the average price and (ii) a decrease in the variance of the price. While selecting the lowest-cost farmers leads to the greatest decrease in the mean price, it also leads to the greatest decrease in its variance (Proposition 8). It turns out that the latter effect dominates, given that the change in expected unit revenue of the non-enrolled farmers (18) is maximized when the lowest-cost farmers are selected. Hence, even if non-enrolled farmers are hurt by the yield-improvement program overall, selecting the lowest-cost farmers minimizes their individual economic loss because they benefit from a more stable (albeit lower) market price.

Proposition 12 characterizes the performance of the lowest-cost farmer selection with respect to the objective of maximizing total farmer profit. Similar to Proposition 6, we use an interchange argument to demonstrate that a low-cost farmer selection will perform reasonably well. Specifically, replacing the enrolled farmer with the largest cost with any other non-enrolled farmer in the enrollment set would result in (a) a profit drop for all other non-enrolled farmers, (b) an aggregate profit drop for the two farmers who are swapped, and (c) a potential profit increase—bounded from above by 16/9—for the other enrolled farmers.

PROPOSITION 12. *Suppose Assumptions 1′ and 2′ hold. For any given n_Y, deviating from the lowest-cost farmer selection by enrolling farmer k, for any $k \in N^{LC}$, instead of farmer j, the n_Yth smallest-cost farmer, i.e., $j = \arg\max_{i \in Y^{LC}} c_i$, results in*

(a) a drop in profit for the non-enrolled farmer i, i.e., $\pi_i(Y^{LC}) \geq \pi_i(Y^{LC} \cup \{k\} \backslash \{j\})$, for any $i \in N^{LC} \backslash \{k\}$,

(b) a drop in the sum of profits of farmers j and k, i.e., $\pi_j(Y^{LC}) + \pi_k(Y^{LC}) \geq \pi_j(Y^{LC} \cup \{k\} \backslash \{j\}) + \pi_k(Y^{LC} \cup \{k\} \backslash \{j\})$,

(c) a potential increase in profit for the enrolled farmer i, but by no more than 16/9 percent, i.e., $\pi_i(Y^{LC}) \geq \frac{9}{16} \pi_i(Y^{LC} \cup \{k\} \backslash \{j\})$, for any $i \in Y^{LC} \backslash \{j\}$.

Hence, $\sum_i \pi_i(Y^{LC}) \geq \frac{9}{16} \sum_i \pi_i(Y^{LC} \cup \{k\} \backslash \{j\})$.

Similar to Proposition 6, this is only a local optimality condition, and selecting the lowest-cost farmers is, in general, not optimal. (For a counter-example, consider a case with the following parameters: $\alpha = 300$, $\beta = 1$, $n = 3$, $n_Y = 2$, $\mu = 1.5$, $\sigma = 10$, $c_1 = 10$, $c_2 = 100$, and $c_3 = 101$: Selecting the first and third farmers yields a higher profit than selecting the first and second farmers.)

In fact, we conjecture that the optimal selection is a disconnected set, consisting of some lowest- and highest-cost farmers. Our intuition is motivated by the comparison of the changes in unit revenues for enrolled and non-enrolled farmers, (17) and (18), which indicated misaligned preferences regarding the configuration of the selection set. On the one hand, each non-enrolled farmer's profit loss increases in market price variability by (18); thus, they prefer that low-cost farmers be selected. On the other hand, each enrolled farmer's profit gain increases in market price variability by (17); thus, they prefer that high-cost farmers be selected. When n_Y is small, most farmers are non-enrolled. In that case, it makes sense to prioritize the selection of the lowest-cost farmers to maximize total profit, as non-enrolled farmers are more numerous. However, when n_Y is large, most farmers are enrolled, so yield-improvement programs should prioritize the selection of highest-cost farmers.

Hence, when the program aims at reducing the yield variance, the objectives of minimizing the average price and its variability, and of minimizing the farmers' potential economic loss from the yield-improvement program, are aligned (as they would both prescribe selecting the lowest-cost farmers); however, they may not be aligned with the objective of maximizing farmers' total profits.

Nevertheless, there is reason to believe that selecting the lowest-cost farmers performs well in terms of aggregate profit. In fact, in our numerical experiments in §6 (which are calibrated with industry data), selecting the lowest-cost farmers turned out to be optimal in every single instance.

6. Calibrated Numerical Experiments

To get a better sense of the magnitude of the trade-offs involved in practice and to test the robustness of our analytical results with respect to the assumption that all farmers remain in business after the implementation of the yield-improvement program (Assumptions 2 and 2'), we resort to calibrated numerical experiments. We consider three industries in which yield-improvement programs are common: cocoa, coffee, and hazelnut. Table 2 outlines the base values of the parameters, and sources are provided in Appendix B. Given that our analytical characterization of the equilibrium predicts that the effect of the yield-improvement program on price depends on the type of yield improvement (i.e., targeted to an increase in its mean as described in §4 or to a decrease in its variance, as described in §5), we consider them separately. For each type of yield improvement, we analytically characterized the farmer selection that minimizes the market price (Propositions 2 and 8) and the one that minimizes the farmers' individual economic loss (Propositions 5 and 11), but only partially the one that maximizes farmers' total profit (Propositions 6 and 12). In our simulations, we identify the latter by solving an integer optimization problem using Gurobi's non-linear solver in a Python environment. (The code is available upon request.)

In the following paragraphs, we first evaluate the benefit of yield improvement using the base values of estimated parameters, depicted in Table 2. Second, we assess the sensitivity of prices and profits to variation in planting costs. Third, we evaluate the magnitude of the trade-offs between the various performance objectives (price, farmers' individual profit losses, and farmers' total profits). Fourth, we check the robustness of our results across the precise specification of the parameters. In all our numerical experiments, we relax Assumptions 2 and 2'.

Benefit of Yield Improvement. Using the base parameter values reported in Table 2, and assuming that all farmers have the same costs corresponding to the mean value (e.g., for cocoa, $c_i = 679\ \forall i$), we evaluate the benefit of the yield-improvement program, assuming that 20% of the farmers are enrolled, which is in line with certification programs in practice (Meier et al. 2020). We consider $n = 20$ farmers[2] and compare the null case, where no farmer is enrolled (i.e., $n_Y = 0$), to one where 20% of them are enrolled (i.e., $n_Y = 4$) in terms of our three performance metrics; namely, prices, farmers' individual economic losses, and farmers' total profits.

We find that increasing the enrolled farmers' mean yield leads to a decrease in the average prices by 3-4% across all three commodities. Given the thin profit margins of these industries, this is a rather substantial benefit for the buyer. Decreasing the enrolled farmers' yield variance leads to only a 0-2% reduction in the mean price, but more importantly, a 23-28% reduction in the standard deviation of the price across all three commodities. Even though the buyers may not substantially benefit from a decrease in mean prices, they may greatly benefit from increased price stability.

[2] Restricting n to 20 is justified later when we investigate the magnitude of trade-offs.

	Parameters	Base Values
Cocoa	μ (ton per ha)	0.45
	σ (ton per ha)	0.005
	b (ton per ha)	0.1125
	α (\$ per ton)	8106
	β (\$ per ton-squared)	1.09
	c_i (\$ per ha per year)	$\sim N(679, 81)$
Coffee	μ (ton per ha)	0.85
	σ (ton per ha)	0.057
	b (ton per ha)	0.2125
	α (\$ per ton)	11647
	β (\$ per ton-squared)	0.84
	c_i (\$ per ha per year)	$\sim N(2463, 542)$
Hazelnut	μ (ton per ha)	1.21
	σ (ton per ha)	0.23
	b (ton per ha)	0.3025
	α (\$ per ton)	5448
	β (\$ per ton-squared)	1.95
	c_i (\$ per ha per year)	$\sim N(3471, 382)$

Table 2 Numerical experiments' parameter values

We now consider the effect of the yield-improvement programs on farmers' profits. Programs that increase the mean yield (resp., reduce its variance) lead to an increase in farmers' total profits by 17% (1.2%) for cocoa, 34% (0.3%) for coffee, and 188% (6.4%) for hazelnut; however, not every farmer benefits from the yield-improvement program. The farmers who are the most negatively affected by the program experience a profit decrease by 33% (7%) for cocoa, 46% (3%) for coffee, and 87% (18%) for hazelnut, respectively. Hence, programs that aim to increase the mean yield appear to be effective at both reducing the average price and increasing total farmer profits, but could also significantly hurt some farmers. In contrast, programs that aim to reduce the yield variability seem to have less impact across the board, with the exception of greater price stability.

Sensitivity of Performance Metrics to Planting Costs. Now that we have quantified the benefits of yield-improvement programs, we assess the sensitivity of the performance metrics to the farmer selection. In particular, we compare the lowest-cost farmer selection with the highest-cost farmer selection, since they are the best and worst strategies to minimize prices by Propositions 2 and 8. By Lemmas 1 and 2, we know that prices evolve linearly with farmers' costs, whereas profits evolve quadratically. As before, we use the parameters in Table 2, $n = 20$, and $n_Y = 4$. However, we no longer keep the farmers' costs equal to their mean: We consider infinitesimal variations in costs (while assuming that all farmers remain in business before and after implementing the program), i.e., a 1%-coefficient of variation around the mean planting cost estimate. For instance, in the case of cocoa, we assume farmers' costs are normally distributed with mean 679 and standard deviation 679×0.01. For each commodity, we simulate 100 runs.

In both the mean and the variance model, we obtain that the difference in prices between the lowest-cost and the highest-cost farmer selections is less than 1% for all commodities. Hence, farmer selection appears to have little effect on prices if farmers have similar planting costs. However, the same cannot be said for profits. In the mean model, selecting the lowest-cost farmers, instead of the highest-cost ones, results in a higher total profit by 4% for cocoa, 6% for coffee, and 24% for hazelnut; resulting in a lower maximum individual profit loss by 8% for cocoa, 11% for coffee, and 22% for hazelnut. Similarly, in the variance model, the same comparison results in higher total profits by 1% for cocoa and coffee and 12% for hazelnut. It results in lower maximum individual profit losses by 7% for cocoa, 9% for coffee, and 18% for hazelnut. As a result, even if farmers are quite (but not completely) similar, choosing the right farmer selection seems to matter in terms of their economic well-being.

Magnitude of Trade-Offs. To further get a sense of the magnitude of the trade-offs involved, this time calibrated on farmers who might have significantly different costs, we now consider the cost distribution presented in Table 2. For instance, for cocoa, we assume that farmers' costs are normally distributed with mean 679 and standard deviation 81. This standard deviation reflects the disparity in the planting costs of farmers from different regions and with different levels of access to technology and farming practices. For instance, the planting cost of cocoa is estimated to be 760$ per hectare in Côte d'Ivoire and 535$ per hectare in Nigeria. For each commodity, we simulated 100 samples of such cost distribution.

With such a large disparity in planting costs, some farmers may go out of business if it is no longer profitable to plant. (This possibility tends to be assumed away in the literature on Cournot agricultural markets (An et al. 2015, Tang et al. 2015, Alizamir et al. 2019), which typically assumes identical farmers.) Although we wanted to allow for the possibility of bankruptcy (to test the robustness of our results with respect to Assumptions 2 and 2'), we also wanted to consider it within reasonable limits. Accordingly, in our experiments, we used $n = 20$ to ensure that in the mean model (the variance model), 95% (97%) of cocoa, 69% (69%) of coffee, and 62% (71%) of hazelnut farmers remain in business, on average. (Choosing a larger value of n would result in higher fractions of farmers going bankrupt.) Given that our estimates of planting costs are global, n should be interpreted as the number of farming regions.

To assess the magnitude of trade-offs, we compute the relative performance of the lowest-cost vs. the highest-cost farmer selection into the yield-improvement program along our three objectives (market prices, farmers' total profits, and farmers' largest individual profit loss): $\frac{p(Q_{YLC}, Q_{NLC})}{p(Q_{YHC}, Q_{NHC})}$, $\frac{\sum_i \pi_i(Y^{LC})}{\sum_i \pi_i(Y^{HC})}$, and $\frac{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{LC})\}}{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{HC})\}}$, respectively. Table 3 presents the simulation output where "Mean" refers to the mean value of the performance metric, "5%" refers to the 5%-quantile, and "95%" refers to the 95%-quantile.

	Objective	Cocoa 5%	Cocoa Mean	Cocoa 95%	Coffee 5%	Coffee Mean	Coffee 95%	Hazelnut 5%	Hazelnut Mean	Hazelnut 95%
Higher Mean Yield	$\frac{p\left(Q_{YLC}, Q_{NLC}\right)}{p\left(Q_{YHC}, Q_{NHC}\right)}$	1.01	1.01	1.01	1.01	1.01	1.02	0.99	1.00	1.01
	$\frac{\sum_i \pi_i(Y^{LC})}{\sum_i \pi_i(Y^{HC})}$	1.25	1.34	1.42	1.29	1.44	1.59	1.58	1.80	2.02
	$\frac{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{LC})\}}{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{HC})\}}$	0.37	0.53	0.70	0.18	0.38	0.58	0.13	0.43	0.74
Lower Yield Variability	$\frac{p\left(Q_{YLC}, Q_{NLC}\right)}{p\left(Q_{YHC}, Q_{NHC}\right)}$	0.99	0.99	0.99	0.99	1.00	1.00	0.97	0.98	0.99
	$\frac{\sum_i \pi_i(Y^{LC})}{\sum_i \pi_i(Y^{HC})}$	1.07	1.08	1.10	1.02	1.03	1.03	1.10	1.19	1.27
	$\frac{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{LC})\}}{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{HC})\}}$	0.35	0.51	0.67	0.17	0.38	0.59	0.18	0.49	0.80

Table 3 Relative performance of the lowest-cost (LC) and the highest-cost (HC) prioritization strategies in terms of price, total profit, and the farmers' individual economic loss when all parameters are set to their base value in Table 2.

	Objective	Cocoa 5%	Cocoa Mean	Cocoa 95%	Coffee 5%	Coffee Mean	Coffee 95%	Hazelnut 5%	Hazelnut Mean	Hazelnut 95%
Higher Mean Yield	$\frac{p\left(Q_{YLC}, Q_{NLC}\right)}{p\left(Q_{YHC}, Q_{NHC}\right)}$	1.00	1.00	1.02	1.00	1.01	1.03	0.97	1.00	1.02
	$\frac{\sum_i \pi_i(Y^{LC})}{\sum_i \pi_i(Y^{HC})}$	1.07	1.39	1.70	1.17	1.46	1.75	1.24	1.66	2.07
	$\frac{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{LC})\}}{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{HC})\}}$	0.27	0.51	0.74	0.12	0.36	0.60	0.01	0.39	0.79
Lower Yield Variability	$\frac{p\left(Q_{YLC}, Q_{NLC}\right)}{p\left(Q_{YHC}, Q_{NHC}\right)}$	0.99	0.98	1.00	0.99	1.00	1.00	0.97	0.98	0.99
	$\frac{\sum_i \pi_i(Y^{LC})}{\sum_i \pi_i(Y^{HC})}$	1.00	1.08	1.16	1.00	1.03	1.06	1.09	1.19	1.28
	$\frac{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{LC})\}}{\max_i\{\pi_i(\varnothing)-\pi_i(Y^{HC})\}}$	0.37	0.55	0.74	0.25	0.45	0.65	0.21	0.56	0.90

Table 4 Relative performance of the lowest-cost (LC) and the highest-cost (HC) prioritization strategies in terms of price, total profit, and the farmers' individual economic loss when all parameters are randomly drawn within 20% of their base value given in Table 2

In all of our simulation runs, it turned out that the farmer selection that maximized the farmers' total profits (obtained from the integer optimization problem) was to select the lowest-cost farmers. This result complements our analytical findings in Propositions 6 and 12, which indicate that the lowest-cost farmer selection would perform well. Moreover, Table 3 reveals that both farmers' total profits and individual profit loss are significantly improved by selecting the lowest-cost (vs. the highest-cost) farmers. From Propositions 5 and 11, we know that selecting the lowest-cost farmers into the yield-improvement program minimizes the farmers' largest individual economic loss under Assumptions 2 and 2' (which prevents bankruptcy). Table 3 not only confirms this holds even in the presence of bankruptcy, but that the gain in performance can be significant. In contrast, the average equilibrium price is relatively insensitive to the farmer selection strategy.

As an illustration, consider cocoa: When the improvement involves increasing the mean yield (resp., decreasing yield variability), the lowest-cost farmer selection, relative to the highest-cost farmer selection, generates, on average, a 34% (8%) increase in farmers' total profits, and a 47% (49%) reduction in the farmers' largest individual economic loss, while leading to an increase (decrease) in prices of less than 1%.[3] Therefore, our experiments confirm that, when the improvement is to reduce yield variability, the lowest-cost farmer selection has potential to align all three objectives; when the improvement is to increase the mean yield, the lowest-cost farmer selection achieves a large profit benefit, while slightly sacrificing price.

Robustness. To test the robustness of our results to the precise specification of the parameter values presented in Table 2, we randomly draw parameter values within 20% of their base values using a uniform distribution (As before, farmers' planting costs are drawn from the normal distribution specified in Table 2.). We consider 100 random samples and present the results in Table 4. Comparing Tables 3 and 4 shows that while the mean values of the relative performance metrics remain similar, drawing the parameter values randomly (as opposed to having them fixed at their base values) leads to larger 90% confidence intervals. We conclude that our key results and insights remain robust to the exact specification of these parameters.

7. Conclusion

The potential flip side of yield-improvement programs is that they may result in lower market prices. Higher yield means higher production quantity, which in turn leads to lower prices. Through extensive interviews, we discovered that some commodity growers are concerned about the potential negative impact of such programs on their profits. When farmers are differentiated in terms of their planting cost and access to yield-improvement programs, some win, but others might lose. To better understand the potential negative externalities of such yield-improvement programs, we develop a Cournot competition model that accounts for differences in planting costs and captures the impact of yield improvement not only on enrolled farmers, but also on non-enrolled farmers. Although large buyers who typically spearhead such yield-improvement programs could have numerous objectives, we focus on minimizing market prices, maximizing farmers' total profits, and minimizing the farmers'

[3] When the mean yield of hazelnuts is increased, the 5%-quantile of the relative price drops below 1, which appears to contradict Proposition 2. This is because we allow for the possibility of going bankrupt. In these cases, the lowest-cost farmer selection leads some high-cost farmers to go bankrupt, whereas the highest-cost farmer selection results in some low-cost farmers going bankrupt. Therefore, the average farmers' planting cost is lower under the lowest-cost farmer selection, resulting in a higher equilibrium planting quantity, and thus in a lower market price. This effect is amplified when the discrepancy in planting costs is high, which explains why we see this occurring for hazelnuts and not for the other commodities.

	Higher Mean Yield		Lower Yield Variability	
	Enrolled	**Non-enrolled**	**Enrolled**	**Non-Enrolled**
Price	− (Proposition 1)		− (Proposition 7)	
Profit (Incremental)	−	−	−	+/−
	(Proposition 3)		(Proposition 9)	
Profit (Aggregate)	+/−	−	+	−
	(Proposition 4)		(Proposition 10)	

Table 5 Effect of the yield-improvement program on price and farmers' incremental and aggregate profits. A + refers to an increase, − refers to a decrease, and a +/− refers to either, depending on the parameters and the scope of the enrollment strategy.

largest individual economic loss. We summarize our results on the impact of yield-improvement program on these performance metrics in Table 5 and then on the optimal farmer selection into the program in Table 6.

As summarized in Table 5, while yield improvements lead to a decrease in market price (Propositions 1 and 7), their impact on profits depend on their type. When yield variability is reduced, on aggregate, enrolled farmers benefit from the program while non-enrolled farmers are hurt (Proposition 10). When the mean yield is improved, some enrolled farmers might be hurt as well (Proposition 4). Considering the effect of improving the yield of an additional (incremental) farmer, we found that all other farmers lose when the mean yield is increased (Proposition 3), but some non-enrolled farmers might benefit when yield variability is reduced (Proposition 9).

Moreover, the buyer's multiple objectives might not necessarily be aligned. Accordingly, the farmer selection strategy into the yield-improvement program depends on both the buyer's primary objective and the type of yield improvement, as summarized in Table 6. Specifically, when the mean yield is increased, selecting the highest-cost farmers leads to the lowest market price; selecting the lowest-cost farmers leads to the smallest individual economic loss from the yield-improvement program and is likely to perform well in terms of farmers' total profits (our numerical experiments indeed suggest that it is often optimal). When yield variability is reduced, selecting the lowest-cost farmers results in the lowest mean market prices and the lowest variability in market prices. This farmer selection also leads to the smallest individual economic loss from the yield-improvement program and is likely to perform well in terms of farmers' total profits (which is supported by our numerical experiments).

This study provides theoretical groundings behind the concerns of the farmers we interviewed. Because many smallholder farmers live on the edge of poverty, any negative impact on their well-being can be extremely detrimental. Our analysis indicates that selecting the lowest-cost farmers minimizes the largest economic loss imposed on farmers because it either has the smallest impact on market prices (increase in mean yield) or reduces the market price volatility the most (reduction in

	Mean Model	**Variance Model**
Lowest Price	HC (Proposition 2)	LC (Proposition 8)
Lowest Individual Profit Reduction	LC (Proposition 5)	LC (Proposition 11)
Highest Total Farmer Profit	LC near optimal (Proposition 6)	LC near optimal (Proposition 12)

Table 6 Optimal farmer selection strategies into the yield-improvement program. HC refers to selecting the highest-cost and LC refers to selecting the lowest-cost farmers into the yield-improvement program.

yield variance). Overall, the lowest-cost farmer selection strategy performs reasonably well, according to our objectives.

There are several directions in which to extend our study. First, one could relax Assumptions 2 and 2'; without them (and as captured in our numerical experiments), some farmers may be driven out of business due to the low prices resulting from the yield-improvement program. Second, one could capture other correlation structures in yield variability to assess the impact of crop diversification. Third, to extend our results to other types of certifications, one could incorporate other economic aspects such as quality differentiation and price premiums (UTZ), or guaranteed minimum prices (FairTrade), and explore why farmers capture so little value from such price premiums (Van Wassenhove and Pot 2021), as well as non-economic benefits (e.g., preventing child labor).

One caveat of our study is that it assumes that farmers adjust their planting quantities to changing market prices and yields. While this likely holds in the long run on aggregate, we should also keep in mind that many individual smallholders may be limited in their decisions, since expansion requires funding and shrinkage below a certain level may not be viable. In particular, the highest-cost farmers may also be the smallest (in equilibrium), so not helping them improve their yield may seem unethical. Yet, our study warns about the long-term consequences of this strategy, as it might depress prices and severely affect the non-enrolled farmers. Accounting for these frictions would further inform the potential conflict among the various performance objectives.

We hope our study will help commodity buyers assess the potential negative externalities of their yield-improvement programs on broader ecosystems, rather than on their supply chains, as well as to understand the various trade-offs associated with such programs and how they are affected by the type of yield improvement.

References

Akkaya D (2017) *Agricultural Supply Chains under Government Interventions*. Ph.D. thesis, Stanford University.

Alizamir S, Iravani F, Mamani H (2019) An analysis of price vs. revenue protection: Government subsidies in the agriculture industry. *Management Science* 65(1):32–49.

An J, Cho SH, Tang CS (2015) Aggregating smallholder farmers in emerging economies. *Production and Operations Management* 24(9):1414–1429.

Boyabatlı O, Kazaz B, Tang C, eds. (2021) *Agricultural Supply Chain Management Research: Operations and Analytics in Planting, Selling, and Government Interventions*. Series in Supply Chain Management (Springer), forthcoming.

Boyabatlı O, Nasiry J, Zhou Y (2019) Crop planning in sustainable agriculture: Dynamic farmland allocation in the presence of crop rotation benefits. *Management Science* 65(5):2060–2076.

Boyabatli O, Shao L, Zhou YH (2021) Integrated optimization of farmland cultivation and fertilizer application: Implications for farm management and food security. Working paper, Singapore Management University, Lee Kong Chian School of Business.

Callebaut B (2020) Forever chocolate progress. Technical report, Barry Callebaut, URL `https://www.barry-callebaut.com/sites/default/files/2020-12/Forever-Chocolate_Report-%202019_20.pdf`.

Chen YJ, Tang CS (2015) The economic value of market information for farmers in developing economies. *Production and Operations Management* 24(9):1441–1452.

Cho SH, Fang X, Tayur S, Xu Y (2019) Combating child labor: Incentives and information disclosure in global supply chains. *Manufacturing & Service Operations Management* 21(3):692–711.

Coppola G, Costantini M, Orsi L, Facchinetti D, Santoro F, Pessina D, Bacenetti J (2020) A comparative cost-benefit analysis of conventional and organic hazelnuts production systems in center Italy. *Agriculture* 10(9):409.

De Koeijer T, Wossink G, Struik P, Renkema J (2002) Measuring agricultural sustainability in terms of efficiency: The case of Dutch sugar beet growers. *Journal of Environmental Management* 66(1):9–17.

de Zegher JF, Iancu DA, Lee HL (2019) Designing contracts and sourcing channels to create shared value. *Manufacturing & Service Operations Management* 21(2):271–289.

de Zegher JF, Iancu DA, Plambeck EL (2018) Sustaining smallholders and rainforests by eliminating payment delay in a commodity supply chain—it takes a village. *Submitted to Management Science* .

Demir İ (2018) Findik tariminda çiftçi bakiş açisindan maliyetler ve etkinlik: Bağlak sayisi üzerine stokastik sinir analizi. *Electronic Turkish Studies* 13(22):619–639.

Deo S, Corbett CJ (2009) Cournot competition under yield uncertainty: The case of the US influenza vaccine market. *Manufacturing & Service Operations Management* 11(4):563–576.

Fairtrade (2020a) About cocoa. URL `https://www.fairtrade.org.uk/farmers-and-workers/cocoa/about-cocoa/`, accessed 2021-07-07.

Fairtrade (2020b) Coffee farmers. URL `https://www.fairtrade.org.uk/farmers-and-workers/coffee/`, accessed 2021-07-07.

Falconer K, Hodge I (2001) Pesticide taxation and multi-objective policy-making: farm modelling to evaluate profit/environment trade-offs. *Ecological Economics* 36(2):263–279.

FAO (2015) The economic lives of smallholder farmers. URL http://www.fao.org/3/a-i5251e.pdf.

FAO (2020) Crops production, food and agriculture organization of the United Nations. URL http://www.fao.org/faostat/en/#data/QC.

Ferrero (2018) Sharing values to create value. URL https://s3-eu-west-1.amazonaws.com/ferrero-static/globalcms/documenti/3733.pdf.

Guda H, Dawande M, Janakiraman G, Rajapakshe T (2021) An economic analysis of agricultural support prices in developing economies. *Production and Operations Management* Articles in Advance.

Hu M, Liu Y, Wang W (2019) Socially beneficial rationality: The value of strategic farmers, social entrepreneurs, and for-profit firms in crop planting decisions. *Management Science* 65(8):3654–3672.

IndexMundi (2021a) Cocoa beans monthly price. URL https://www.indexmundi.com/commodities/?commodity=cocoa-beans, accessed 2020-12-11.

IndexMundi (2021b) Coffee, other mild arabicas monthly price. URL https://www.indexmundi.com/commodities/?commodity=other-mild-arabicas-coffee, accessed 2020-12-11.

Kazaz B (2004) Production planning under yield and demand uncertainty with yield-dependent cost and price. *Manufacturing & Service Operations Management* 6(3):209–224.

Kilian B, Jones C, Pratt L, Villalobos A (2006) Is sustainable agriculture a viable strategy to improve farm income in Central America? A case study on coffee. *Journal of Business Research* 59(3):322–330.

Levi R, Rajan M, Singhvi S, Zheng Y (2020) The impact of unifying agricultural wholesale markets on prices and farmers' profitability. *Proceedings of the National Academy of Sciences* 117(5):2366–2371.

Liao CN, Chen YJ, Tang CS (2019) Information provision policies for improving farmer welfare in developing countries: Heterogeneous farmers and market selection. *Manufacturing & Service Operations Management* 21(2):254–270.

Meier C, Sampson G, Larrea C, Schlatter B, Voora V, Dang D, Bermudez S, Wozniak J, Willer H (2020) The state of sustainable markets 2020: Statistics and emerging trends. Technical report, International Trade Center.

Montagnon C (2017) Coffee production costs and farm profitability: Strategic literature review. Technical report, Specialty Coffee Association.

Neale B (2016) *Evaluating the True Cost of Cocoa Production & the Viability of Mondelēz International's Farm of the Future Model*. Master's thesis, Duke University Nicholas School of the Environment.

Oseni JO, Adams AQ (2013) Cost benefit analysis of certified cocoa production in Ondo State, Nigeria. *2013 Fourth International Conference* (African Association of Agricultural Economists (AAAE)).

OTB (2021) Hazelnut prices by years. URL https://www.ordutb.org.tr/en/#, accessed 2020-12-11.

Pacini C, Wossink A, Giesen G, Huirne R (2004) Ecological-economic modelling to support multi-objective policy making: a farming systems approach implemented for Tuscany. *Agriculture, Ecosystems & Environment* 102(3):349–364.

Parker C, Ramdas K, Savva N (2016) Is it enough? Evidence from a natural experiment in India's agriculture markets. *Management Science* 62(9):2481–2503.

Ruben R, Zuniga G (2011) How standards compete: comparative impact of coffee certification schemes in Northern Nicaragua. *Supply Chain Management: An International Journal* 16(2):98–109.

Siray E, Erol, Fatih Ö, Ömür D, Halil, Sayili M, Akcay Y (2015) Fındık yetiştiren işletmelerin ekonomik analizi: Giresun ili örneği. *Gaziosmanpaşa Üniversitesi Ziraat Fakültesi Dergisi (Journal of Agricultural Faculty of Gaziosmanpasa University)* 32(2):64–78.

Smith C, Soonieus R, Duke L (2014) Barry Callebaut: To follow the chair? INSEAD Case Study 6440.

Tang CS, Wang Y, Zhao M (2015) The implications of utilizing market information and adopting agricultural advice for farmers in developing economies. *Production and Operations Management* 24(8):1197–1215.

Tang CS, Wang Y, Zhao M (2018) The impact of input-vs. output-based farm subsidies on farmer welfare and income inequality in developing economies. Working paper, University of California, Los Angeles.

Tang CS, Zhou S (2012) Research advances in environmentally and socially sustainable operations. *European Journal of Operational Research* 223(3):585–594.

Thorlakson T (2018) A move beyond sustainability certification: The evolution of the chocolate industry's sustainable sourcing practices. *Business Strategy and the Environment* 27(8):1653–1665.

Van Wassenhove LV, Pot V (2021) Tony's Chocolonely: The road to 100% slave-free chocolate? URL `https://publishing.insead.edu/case/tonys-chocolonely`, INSEAD Case Study 6616.

Van Wassenhove LV, Pot V, Dewilde S, Breugem T (2021) Are chocolate eaters really SDG smart? URL `https://publishing.insead.edu/case/chocolate-eaters`, INSEAD Case Study 6604.

World Bank (2016) A year in the lives of smallholder farmers. URL `https://www.worldbank.org/en/news/feature/2016/02/25/a-year-in-the-lives-of-smallholder-farming-families`.

Xiao S, Chen YJ, Tang CS (2020) Knowledge sharing and learning among smallholders in developing economies: Implications, incentives, and reward mechanisms. *Operations Research* 68(2):435–452.

Yahaya AM, Karli B, Gül M (2015) Economic analysis of cocoa production in Ghana: The case of eastern region. *Custos e Agronegócio* 11(1):336–352.

Yano CA, Lee HL (1995) Lot sizing with random yields: A review. *Operations Research* 43(2):311–334.

Zhang Y, Swaminathan JM (2020) Improved crop productivity through optimized planting schedules. *Manufacturing & Service Operations Management* 22(6):1165–1180.

Zhou J, Fan X, Chen YJ, Tang CS (2021) Information provision and farmer welfare in developing economies. *Manufacturing & Service Operations Management* 23(1):230–245.

ZMO (2018) Fındık raporu — 2018. URL https://www.zmo.org.tr/genel/bizden_detay.php?kod=30070&tipi=, accessed 2021-07-07.

Appendix A: Notation

Variable	Description
Parameters:	
Y, N	Set of enrolled and non-enrolled farmers, respectively
n_S	Number of farmers in set $S \in \{Y, N\}$, respectively
θ_i	Yield of farmer i
α, β	Parameters of the linear demand function
c_i	Farmer i's planting cost per area over the time horizon
n	Total number of farmers in the market
Decision Variables:	
q_i	Planting area used by farmer i
Sums and Averages:	
C_Y, C_N	Sum of costs in set Y and N, respectively, i.e., $\sum_{i \in Y} c_i$ and $\sum_{i \in N} c_i$
$\overline{c_Y}, \overline{c_N}$	Average cost of farmer in set Y and N, respectively, i.e., C_Y/n_Y and C_N/n_N
Q_Y, Q_N	Sum of planting areas in set Y and N, respectively, i.e., $\sum_{i \in Y} q_i$ and $\sum_{i \in N} q_i$
Q	Sum of planting areas of all farmers, i.e., $Q = Q_Y + Q_N$.
Functions:	
$p(Q_Y, Q_N)$	Market price as a function of aggregate planting areas
π_i	Farmer i's profit over the time horizon

Table A-1 Summary of Notation

Appendix B: Simulation Parameters

B.1. Planting Cost

To estimate the planting cost for each crop in $ per hectare (ha) per year, we used references for multiple countries that produce a majority of the crop and applied the following procedure.

1. We reviewed various streams of literature (listed below) to estimate the costs of planting hazelnut, cocoa, and coffee. When we encountered different estimates for the same country from different sources, we took an average to obtain one estimate per country.

2. We converted the reported estimates to USD using the currency conversion rate in the corresponding year.[4]

3. Given that we are considering longitudinal data, we adjusted the commodity price estimates by the inflation rates using historical consumer price index data.[5]

[4] `https://www.exchangerates.org.uk/`

[5] `https://inflationdata.com/Inflation/Consumer_Price_Index/HistoricalCPI.aspx?reloaded=true`

4. We computed the mean and standard deviation of the cost distribution as follows. Let w_i be country i's share of the average farming land (total quantity produced divided by the country's average yield) during the horizon of our study (2008-2018) using FAO (2020) website data and c_i be its estimated cost. The mean cost is computed as $\sum_i w_i c_i$, and its variance as $\sum_i w_i (c_i - \sum_i w_i c_i)^2$.

For cocoa, we used three sources: Côte d'Ivoire (Neale 2016), Ghana (Yahaya et al. 2015), and Nigeria (Oseni and Adams 2013). The cost range that appears in these studies is $[\$535, \$760]$ per ha per year, with a weighted average equal to \$679 per ha per year.

For coffee, we used data for Honduras, Nicaragua, Brazil, Colombia, Costa Rica, Ethiopia, Kenya, Tanzania, and El Salvador (Montagnon 2017). The cost range that appears in these studies is $[\$435, \$4121]$ per ha per year (excluding Ethiopia, which was an outlier), with a weighted average equal to \$2463 per ha per year.

For hazelnut, we obtained data for Italy (Coppola et al. 2020) and Turkey (Demir 2018, Siray et al. 2015). The cost range that appears in these studies is $[\$3328, \$5428]$ per ha per year, with a weighted average equal to \$3471 per ha per year. These numbers are consistent with the cost estimates of the farmers we interviewed.

B.2. Yield Parameters

To estimate yield parameters μ and σ, we used the 2008-2018 annual production data from the statistics tool of the Food and Agriculture Organization of the United Nations (FAO 2020). The yield mean, μ, is equal to 0.45 ton per ha for cocoa; 0.85 ton per ha for coffee; and 1.21 ton per ha for hazelnut. Similarly, the yield standard deviation, σ, is 0.005 ton per ha for cocoa; 0.057 ton per ha for coffee; and 0.23 ton per ha for hazelnut.

As for the yield increase, b, Ferrero (2018) reports that orchards that participated in the Ferrero Farming Values yield-improvement program saw an average of 25-30% increase in yield. In our interviews, we confirmed a yield increase of 25%. Accordingly, we used $b = 0.25\mu$ in our simulations.

B.3. Price-Demand Curve Parameters

For the market prices, we used local commodity exchange data (OTB 2021) for hazelnut, and the data portal IndexMundi, which contains data from the International Cocoa Organization and the International Coffee Organization, for cocoa and coffee (IndexMundi 2021a,b). Similar to the cost of farming, we also adjusted the prices for inflation. Using the same production data as in our estimation of the yield parameters, we estimated the demand parameters, α and β, using linear regression by expressing the average market prices as a function of the annual production quantity. From the regression, we obtained that α is equal to $8,106\$$ per ton for cocoa; $11,647\$$ per ton for coffee; and $5,448\$$ per ton for hazelnut. Similarly, β is equal to $1.09\$$ per ton-squared for cocoa; $0.84\$$ per ton-squared for coffee; and $1.95\$$ per ton-squared for hazelnut.

B.4. Internal Consistency Check

To test the internal consistency of our model, we compared our model's predicted equilibrium planting quantities (Lemma 1) to the average reported harvested area of FAO during the time period we considered (2008-2018), using the base parameter estimates in Table 2, assuming that no yield is improved, and assuming that all farmers' planting costs are equal to the mean estimate. We considered $n = 6$ million cocoa farmers

(Fairtrade 2020a), $n = 25$ million coffee farmers (Fairtrade 2020b), and $n = 440,000$ hazelnut farmers (ZMO 2018).

For cocoa, our model predicts an equilibrium total planting quantity of $Q^* = 13,400$ ha, whereas the reported planting quantity is $10,919$ ha (19% lower). For coffee, the model's prediction is $12,210$ ha, whereas the reported planting quantity is $10,606$ ha (13% lower). For hazelnut, the model's prediction is $1,093$ ha, whereas the reported hazelnut harvested area is equal to $759,000$ ha (30% lower). Overall, it appears that our model only slightly overestimates the planting quantity, at least without accounting for the effect of yield improvement.

Appendix C: Details of Interviews

We conducted ten 30-minute interviews during the summer of 2018 in Ordu, Turkey with the following stakeholders:

- Five smallholder farmers (three in person and two over the phone)
- Two managers of hazelnut processors (in person)
- One hazelnut trader (over the phone)
- One fertilizer/equipment seller (over the phone)
- The general secretary of Ordu hazelnut commodity exchange (in person)

The following questions were asked:

1. What is the cost of producing hazelnut?
2. Are farmers getting fair prices for your hazelnut? What do you think influences hazelnut prices?
3. Which channel do farmers prefer to use to sell your hazelnut?
4. Do farmers require financing? If so, how do they acquire it?
5. What do you think about the yield-improvement programs, such as Ferrero's FFV program?
   (a) What is the benefit of being enrolled in a yield-improvement program?
   (b) What is the cost of being enrolled in a yield-improvement program?
6. What are other problems in the hazelnut farming sector?

Interview notes are available upon request.

Proofs of Statements

Proof of Lemma 1. Under Assumption 1, $\pi_i = p(Q_Y, Q_N)(\mu + b)q_i - q_i c_i$ for $i \in Y$ and $\pi_i = p(Q_Y, Q_N)\mu q_i - q_i c_i$ for $i \in N$ where $p(Q_Y, Q_N) = \alpha - \beta(\mu + b)Q_Y - \beta\mu Q_N$. Since $\frac{\partial^2 \pi_i}{\partial q_i^2} = -2\beta(\mu + b)^2 < 0$ for $i \in Y$ and $\frac{\partial^2 \pi_i}{\partial q_i^2} = -2\beta\mu^2 < 0$ for $i \in N$, $\pi_i(q_i; q_{-i})$ is strictly concave. Solving the first-order optimality conditions while ignoring the non-negativity constraints, we obtain:

$$q_i = \frac{\alpha\mu - \beta\mu^2 Q_N - \beta\mu(\mu + b)Q_Y - c_i}{\beta\mu^2} \qquad i \in N$$
$$q_i = \frac{\alpha(\mu + b) - \beta\mu(\mu + b)Q_N - \beta(\mu + b)^2 Q_Y - c_i}{\beta(\mu + b)^2} \qquad i \in Y.$$

Solving these equations with $Q_N = \sum_{i \in N} q_i$ and $Q_Y = \sum_{i \in Y} q_i$ yields (3) and (4). By Assumption 2, $q_i^* \geq 0\ \forall i$. Plugging (3) and (4) into $\pi_i = p(Q_Y, Q_N)(\mu + b)q_i - q_i c_i$, $\forall i \in Y$ and $\pi_i = p(Q_Y, Q_N)\mu q_i - q_i c_i$, $\forall i \in N$, respectively, yields (5). □

Proof of Proposition 1. By Lemma 1,

$$Q_N^* = \sum_{i \in N} q_i^* = \frac{n_N[\alpha\mu(\mu + b) + C_Y\mu] - C_N(\mu + b)(1 + n_Y)}{\beta\mu^2(\mu + b)(n + 1)}$$
$$Q_Y^* = \sum_{i \in Y} q_i^* = \frac{n_Y[\alpha\mu(\mu + b) + C_N(\mu + b)] - C_Y\mu(1 + n_N)}{\beta\mu(\mu + b)^2(n + 1)},$$

which results in

$$Q^* = \frac{\alpha\mu(\mu + b)(\mu n + b n_N) + C_N(\mu + b)(-\mu - b(1 + n_Y)) + C_Y\mu(b n_N - \mu)}{\beta\mu^2(\mu + b)^2(n + 1)}. \tag{EC.1}$$

Hence,

$$p^*(Q_Y, Q_N) = \frac{\alpha\mu(\mu + b) + C\mu + C_N b}{\mu(\mu + b)(n + 1)}. \tag{EC.2}$$

Enrolling $j \in N$ results in a decrease in C_N, and, in turn, a decrease in $p^*(Q_Y, Q_N)$. □

Proof of Proposition 2. By Proposition 1, the equilibrium market price is equal to (EC.2), which is minimized when C_N is the smallest. Hence, selecting the highest-cost farmers minimizes the equilibrium market price. □

Proof of Proposition 3. Let q_i^+ denote the equilibrium planting area of farmer i after farmer j is enrolled. Then, using (3) and (4), we obtain

$$q_i^+ - q_i = \frac{-bc_j}{\beta\mu^2(\mu + b)(n + 1)} \leq 0 \qquad i \in N,$$
$$q_i^+ - q_i = \frac{-bc_j}{\beta\mu(\mu + b)^2(n + 1)} \leq 0 \qquad i \in Y.$$

Using (5), we then obtain that $\pi_i(Y) \geq \pi_i(Y \cup \{j\})$, $\forall i \neq j$. □

Proof of Proposition 4. First, consider any farmer $i \in Y$. By Lemma 1,

$$\pi_i(\varnothing) - \pi_i(Y) = \frac{[\alpha\mu(\mu+b) + C(\mu+b) - c_i(\mu+b)(n+1)]^2}{\beta\mu^2(\mu+b)^2(n+1)^2} - \frac{[\alpha\mu(\mu+b) + C\mu + C_N b - c_i\mu(n+1)]^2}{\beta\mu^2(\mu+b)^2(n+1)^2}$$
$$= \frac{b[C_Y - c_i(n+1)][2\alpha\mu(\mu+b) + 2C(\mu+b) - C_Y b - c_i(2\mu+b)(n+1)]}{\beta\mu^2(\mu+b)^2(n+1)^2}. \tag{EC.3}$$

Under Assumption 2, the second term in brackets in the numerator of (EC.3) is non-negative. Therefore, $\pi_i(Y) \le \pi_i(\varnothing)$ if and only if the first term in brackets in the numerator of (EC.3) is also non-negative, i.e., $c_i(n+1) \le C_Y$. Second, consider any farmer $i \in N$. By Lemma 1,

$$\pi_i(\varnothing) - \pi_i(Y) = \frac{[C_Y b][2\alpha\mu(\mu+b) + 2C(\mu+b) - C_Y b - 2c_i(\mu+b)(n+1)]}{\beta\mu^2(\mu+b)^2(n+1)^2} \ge 0 \tag{EC.4}$$

since the second term in brackets in the numerator of (EC.4) is non-negative under Assumption 2.

□

Proof of Proposition 5. First, consider farmer $i \in Y$. By Proposition 4, if $C_Y - c_i(n+1) \le 0$, $\pi_i(\varnothing) \le \pi_i(Y)$, and, hence, $\max\{0, \pi_i(\varnothing) - \pi_i(Y)\} = 0$. Else, if $C_Y - c_i(n+1) > 0$, $\pi_i(\varnothing) \ge \pi_i(Y)$ and, hence, $\max\{0, \pi_i(\varnothing) - \pi_i(Y)\} = \pi_i(\varnothing) - \pi_i(Y)$. Because

$$\frac{d(\pi_i(\varnothing) - \pi_i(Y))}{dC_Y} = 2b\frac{\alpha\mu(\mu+b) + C\mu + C_N b - c_i\mu(n+1)}{\beta\mu^2(\mu+b)^2(n+1)^2} \ge 0$$

under Assumption 2, $\pi_i(\varnothing) - \pi_i(Y)$, as expressed by (EC.3) in the proof of Proposition 4, is minimized when C_Y is the smallest, i.e., when $Y = Y^{LC}$.

Second, consider any $i \in N$. Because

$$\frac{d(\pi_i(\varnothing) - \pi_i(Y))}{dC_Y} = 2b\frac{\alpha\mu(\mu+b) + C\mu + C_N b - c_i(\mu+b)(n+1)}{\beta\mu^2(\mu+b)^2(n+1)^2} \ge 0$$

under Assumption 2, $\pi_i(\varnothing) - \pi_i(Y)$, as expressed by (EC.4) in the proof of Proposition 4, is minimized when C_Y is the smallest, i.e., when $Y = Y^{LC}$. □

Proof of Proposition 6. Keeping C fixed, the equilibrium planting areas, (3) and (4), are both maximized when C_N is the largest. Therefore, by (5), and because $c_k \ge c_j$ under the lowest-cost farmer selection strategy, $\pi_i(Y^{LC} \cup \{k\} \backslash \{j\}) \le \pi_i(Y^{LC})\ \forall j \in Y^{LC},\ \forall k \in N^{LC},\ \forall i \ne j, k$. □

Lemma EC.1. *Under Assumption 2′,* $\alpha\mu + C - c_i(1+n) \ge 0\ \ \forall i$.

Proof. Under Assumption 2' and because $c_n \geq c_i \ \forall i$,

$$\begin{aligned}
&\alpha\mu + C + C_N n_Y - c_n(1+n_Y)(1+n_N) \geq 0 \\
\iff &\alpha\mu + C - c_n(1+n) + \underbrace{n_Y(C_N - c_n n_N)}_{\leq 0} \geq 0 \\
\implies &\alpha\mu + C - c_n(1+n) \geq 0 \\
\implies &\alpha\mu + C - c_i(1+n) \geq 0 \quad \forall i \in \{1,2,...,n\}.
\end{aligned}$$

□

Lemma EC.2. *Under Assumption 2',* $\alpha\mu + C_Y - c_i(1+n_Y) \geq 0 \ \ \forall i.$

Proof. By Lemma EC.1 and because $c_n \geq c_i \ \forall i$,

$$\begin{aligned}
&\alpha\mu + C - c_n(1+n) \geq 0 \\
\iff &\alpha\mu + C_Y - c_n(1+n_Y) + \underbrace{C_N - c_n n_N}_{\leq 0} \geq 0 \\
\implies &\alpha\mu + C_Y - c_n(1+n_Y) \geq 0 \\
\implies &\alpha\mu + C_Y - c_i(1+n_Y) \geq 0 \ \ \forall i \in \{1,2,...,n\}.
\end{aligned}$$

□

Proof of Lemma 2. Under Assumption 1', $\pi_i = E\left[p(Q_Y,Q_N)\theta_Y q_i - q_i c_i\right]$ for $i \in Y$ and $\pi_i = E\left[p(Q_Y,Q_N)\theta_N q_i - q_i c_i\right]$ for $i \in N$ where $p(Q_Y,Q_N) = \alpha - \beta\theta_Y Q_Y - \beta\theta_N Q_N$. Since $\frac{\partial^2 \pi_i}{\partial q_i^2} = -2\beta(\mu^2 + \sigma^2) < 0$ for $i \in N$, and $\frac{\partial^2 \pi_i}{\partial q_i^2} = -2\beta\mu^2 < 0$ for $i \in Y$, $\pi_i(q_i;q_{-i})$ is strictly concave. Solving the first-order optimality conditions while ignoring the non-negativity constraints, we obtain:

$$\begin{aligned}
q_i &= \frac{\alpha\mu - c_i - \beta\left[Q_N(\mu^2+\sigma^2) + Q_Y(\mu^2)\right]}{\beta(\mu^2+\sigma^2)} && i \in N, \\
q_i &= \frac{\alpha\mu - c_i - \beta(Q_N + Q_Y)\mu^2}{\beta(\mu^2)} && i \in Y.
\end{aligned}$$

Solving these equations with $Q_N = \sum_{i\in N} q_i$ and $Q_Y = \sum_{i \in Y} q_i$ yields (8) and (9). By Lemmas EC.1 and EC.2, $q_i^* \geq 0 \ \forall i$. Plugging (8) and (9) into $\pi_i = E\left[p(Q_Y,Q_N)\theta_Y q_i - q_i c_i\right] \ \forall i \in Y$ and $\pi_i = E\left[p(Q_Y,Q_N)\theta_N q_i - q_i c_i\right] \ \forall i \in N$, respectively, yields (10).

Proof of Proposition 7. Throughout the proof, n_N, n_Y, C_N and C_Y refer to the base case; that is, before farmer j is added to the yield-improvement program. By Lemma 2,

$$Q_N = \sum_{i\in N} q_i^* = \frac{-C_N(1+n_Y)+(C_Y+\alpha\mu)n_N}{\beta(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)} \tag{EC.5}$$

$$Q_Y = \sum_{i\in Y} q_i^* = \frac{\sigma^2(1+n_N)(-C_Y+\alpha\mu n_Y)+\mu^2(-C_Y(1+n_N)+(C_N+\alpha\mu)n_Y)}{\beta\mu^2(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}. \tag{EC.6}$$

Then,

$$Q = Q_Y + Q_N = \frac{\sigma^2(1+n_N)(-C_Y+\alpha\mu n_Y)+\mu^2(-C+\alpha\mu n)}{\beta\mu^2(\sigma^2(1+n_N)(1+n_Y)+\mu^2(1+n))}. \tag{EC.7}$$

Let $Q_Y^+ = Q_{Y\cup\{j\}}$ and $Q_N^+ = Q_{N\setminus\{j\}}$. Further, let $Q^+ = Q_Y^+ + Q_N^+$. Therefore,

$$\begin{aligned} Q^+ - Q = {} & \mu^2\sigma^2 \frac{-c_j n_N(1+n) - C_N + n(2C_Y+C_N+2\alpha\mu) - 2n_Y(C_Y+C_N+\alpha\mu)}{\beta\mu^2((1+n)\mu^2+(2+n_Y)n_N\sigma^2)((1+n)\mu^2+(1+n_Y)(1+n_N)\sigma^2)} \\ & + \sigma^4 \frac{(1+n_N)n_N(-c_j(1+n_Y)+C_Y+\alpha\mu)}{\beta\mu^2((1+n)\mu^2+(2+n_Y)n_N\sigma^2)((1+n)\mu^2+(1+n_Y)(1+n_N)\sigma^2)}. \end{aligned}$$

The coefficient of σ^4 is non-negative by Lemma EC.2. The coefficient of $\mu^2\sigma^2$ is also non-negative because by Lemma EC.1,

$$\begin{aligned} & -c_j n_N(1+n) - C_N + n(2C_Y+C_N+2\alpha\mu) - 2n_Y(C_Y+C_N+\alpha\mu) \\ & = 2n_N(C+\alpha\mu) - C_N(n+1) - c_j n_N(n+1) \\ & \geq 2n_N(1+n)c_n - \overline{c_N} n_N(n+1) - c_j n_N(n+1) \\ & = n_N(1+n)\underbrace{(c_n - c_j)}_{\geq 0} + n_N(1+n)\underbrace{(c_n - \overline{c_N})}_{\geq 0} \geq 0; \end{aligned}$$

here, the last inequality is because $c_n \geq \overline{c_N}$ and $c_n \geq c_j$. Therefore, $Q^+ \geq Q$. Because $E[p(Q_Y, Q_N)] = [\alpha - \beta\mu Q]$, $E[p(Q_Y^+, Q_N^+)] \leq E[p(Q_Y, Q_N)]$.

Similarly,

$$\begin{aligned} Q_N^+ - Q_N = {} & -\mu^2 \frac{(n+1)(-c_j(n+1)+C_Y+C_N+\alpha\mu)}{\beta((n+1)\mu^2+(2+n_Y)n_N\sigma^2)((n+1)\mu^2+(1+n_Y)(1+n_N)\sigma^2)} \\ & - \sigma^2 \frac{(-c_j(1+n_Y)(1+n_N)(n+1)+(1+n_Y+n_N^2)(C_Y+\alpha\mu)+(1+n_Y)(2+n_Y)C_N)}{\beta((n+1)\mu^2+(2+n_Y)n_N\sigma^2)((1+n)\mu^2+(1+n_Y)(1+n_N)\sigma^2)}. \end{aligned}$$

Because

$$\frac{d(Q_N^+ - Q_N)}{dc_j} = \frac{1+n}{\beta((n+1)\mu^2+(2+n_Y)n_N\sigma^2)} \geq 0, \tag{EC.8}$$

$Q_N^+ - Q_N$ is maximized when c_j is the largest. Under Assumption 2', $c_j \leq c_n \leq \frac{\alpha\mu+C_N+C_Y+n_Y C_N}{(1+n_Y)(1+n_N)}$. Considering the right-hand side, we obtain

$$Q_N^+ - Q_N \leq \frac{(C_N - n(C_Y+\alpha\mu)+n_Y(C_Y+C_N+\alpha\mu))(n_Y(1+n)\mu^2+(1+n_Y)(n_N^2-1)\sigma^2)}{(1+n_Y)(1+n_N)\beta((1+n)\mu^2+(1+n_Y)(1+n_N)\sigma^2)((1+n)\mu^2+(2+n_Y)n_N\sigma^2)}. \tag{EC.9}$$

Because the derivative of the right-hand side of (EC.9) with respect to C_N equals

$$\frac{n_Y(n+1)\mu^2+(1+n_Y)(n_N^2-1)\sigma^2}{(1+n_N)\beta((n+1)\mu^2+(1+n_Y)(1+n_N)\sigma^2)((n+1)\mu^2+(2+n_Y)n_N\sigma^2)}\geq 0,$$

the right-hand side of (EC.9) is maximized when C_N takes its maximum value. By Assumption 2', $C_N \leq n_N \frac{\alpha\mu+C_N+C_Y+n_Y C_N}{(1+n_Y)(1+n_N)}$, i.e., $C_N \leq n_N \frac{C_Y+\alpha\mu}{1+n_Y}$. Replacing C_N with its upper bound, we obtain $Q_N^+ - Q_N \leq 0$. Because $V[p(Q_Y,Q_N)] = \beta^2\sigma^2 Q_N^2$, $V[p(Q_Y^+,Q_N^+)] \leq V[p(Q_Y,Q_N)]$ holds. □

Proof of Proposition 8. Since $E[p(Q_Y,Q_N)] = \alpha - \beta\mu Q$, the expected price is minimized when Q is the largest. Using (EC.7) in the proof of Proposition 7, we obtain that Q is maximized when C_Y is the smallest. Similarly, since $V[p(Q_Y,Q_N)] = \beta^2\sigma^2 Q_N^2$, the variance of price is minimized when Q_N is the smallest. Using (EC.5) in the proof of Proposition 7, we obtain that Q_N is minimized when C_Y is the smallest. □

Proof of Proposition 9. Throughout the proof, n_N, n_Y, C_N, and C_Y refer to the base case; that is, before farmer j is added to the yield-improvement program. For any $i \neq j$, let $q_i(Y)$ denote the equilibrium planting area of farmer i when the set of enrolled farmers is Y. For any $i \in Y\backslash\{j\}$, farmer i is worse off if and only if $\frac{\pi_i(Y)}{\pi_i(Y\cup\{j\})} \geq 1$, which by Lemma 2 is equivalent to $\frac{q_i(Y)}{q_i(Y\cup\{j\})} \geq 1$, i.e.,

$$\frac{[(1+n)\mu^2+n_N(2+n_Y)\sigma^2]}{[(1+n)\mu^2+(1+n_N)(1+n_Y)\sigma^2]} \geq \frac{[\sigma^2 n_N(-c_i(2+n_Y)+C_Y+c_j+\alpha\mu)+\mu^2(-c_i(1+n)+C+\alpha\mu)]}{[\sigma^2(1+n_N)(-c_i(1+n_Y)+C_Y+\alpha\mu)+\mu^2(-c_i(1+n)+C+\alpha\mu)]} \tag{EC.10}$$

or

$$\left(\frac{\sigma}{\mu}\right)^2 \geq \frac{-\left[(-(1+n)c_j+C+\alpha\mu)+(-(1+n)\overline{c_N}+C+\alpha\mu)\right]}{(1+n_N)\left(-(1+n_Y)c_j+C_Y+\alpha\mu\right)}.$$

By Lemmas EC.1 and EC.2, the right-hand side is negative because its nominator is negative and its denominator is positive. Since the left-hand side is always positive, the above inequality always holds. Thus, $\pi_i(Y) \geq \pi_i(Y\cup\{j\})$.

For any $i \in N\backslash\{j\}$, farmer i is worse off if and only if $\frac{\pi_i(Y)}{\pi_i(Y\cup\{j\})} \geq 1$, which by Lemma 2 is equivalent to $\frac{q_i(Y)}{q_i(Y\cup\{j\})} \geq 1$, i.e.,

$$\frac{[(1+n)\mu^2+n_N(2+n_Y)\sigma^2]}{[(1+n)\mu^2+(1+n_N)(1+n_Y)\sigma^2]} \geq \frac{[\sigma^2(-c_i n_N(2+n_Y)+C+(C_N-c_j)(1+n_Y)+\alpha\mu)+\mu^2(-c_i(1+n)+C+\alpha\mu)]}{[\sigma^2(-c_i(1+n_N)(1+n_Y)+C+C_N n_Y+\alpha\mu)+\mu^2(-c_i(1+n)+C+\alpha\mu)]}. \tag{EC.11}$$

Dividing both sides by μ^2 and rearranging terms, we obtain that $\frac{\pi_i(Y)}{\pi_i(Y\cup\{j\})} \geq 1$ if and only if

$$\left(\frac{\sigma}{\mu}\right)^2\left[\overbrace{(1+n_Y-n_N)}^{\text{positive}\iff n_N\le n_Y+1}\overbrace{(c_j(1+n_N)(1+n_Y)-C-C_Nn_Y-\alpha\mu)}^{\le 0 \text{ under Assumption 2'}}+n_N\overbrace{(1+n_N)(1+n_Y)(c_j-\overline{c_N})}^{\text{positive}\iff c_j\ge\overline{c_N}}\right]$$

$$+\left[\underbrace{(1+n_Y-n_N)}_{\text{positive}\iff n_N\le n_Y+1}\underbrace{(c_j(1+n)-C-\alpha\mu)}_{\le 0 \text{ by Lemma EC.1}}+\underbrace{n_N(1+n)(c_j-\overline{c_N})}_{\text{positive}\iff c_j\ge\overline{c_N}}\right]\ge 0,$$

where the inequalities follow by Assumption 2' and Lemma EC.1. Therefore, $(ii.1)$ if $n_N\ge n_Y+1$ and $c_j\ge\overline{c_N}$, then $\pi_i(Y)\ge\pi_i(Y\cup\{j\})$; and $(ii.2)$ if $n_N\le n_Y+1$ and $c_j\le\overline{c_N}$, then $\pi_i(Y)\le\pi_i(Y\cup\{j\})$. □

Proof of Proposition 10. Let $q_i(Y)$ denote the equilibrium planting area of farmer i when the set of enrolled farmers is Y. First, consider any $i\in Y$. By Proposition 9, $\pi_i(Y)\ge\pi_i(Y\cup\{k\})$ for any Y and $k\notin Y$. Moreover by Lemma 2,

$$\begin{aligned}\pi_i(\varnothing)-\pi_i(Y\cup N)&=\frac{(-c_i(1+n)+C+\alpha\mu)^2}{\beta(\mu^2+\sigma^2)(1+n)^2}-\frac{(-c_i(1+n)+C+\alpha\mu)^2}{\beta\mu^2(1+n)^2}\\&=-\frac{\sigma^2(-c_i(1+n)+C_N+C_Y+\alpha\mu)^2}{\beta\mu^2(\mu^2+\sigma^2)(1+n)^2}\le 0.\end{aligned}\tag{EC.12}$$

Combining these two results and applying the first iteratively for all $j\in N$, we obtain $\pi_i(\varnothing)\le\pi_i(Y\cup N)\le\pi_i(Y\cup N\backslash\{j\})\le...\le\pi_i(Y)$. Hence, $\pi_i(Y)\ge\pi_i(\varnothing)$.

Second, consider any $i\in N$. By Lemmas 2 and EC.1,

$$\begin{aligned}q_i(\varnothing)-q_i(Y)=&\frac{(-c_i(1+n)+C+\alpha\mu)}{\beta(\mu^2+\sigma^2)(1+n)}-\frac{\sigma^2(-c_i(1+n_N)(1+n_Y)+C+C_Nn_Y+\alpha\mu)+\mu^2(-c_i(1+n)+C+\alpha\mu)}{\beta(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}\\=&\frac{(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)(-c_i(1+n)+C+\alpha\mu)}{\beta(1+n)(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}\\&-\frac{(1+n)\left(\sigma^2(-c_i(1+n_N)(1+n_Y)+C+C_Nn_Y+\alpha\mu)+\mu^2(-c_i(1+n)+C+\alpha\mu)\right)}{\beta(1+n)(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}\\=&\mu^2\frac{\overbrace{(1+n)(-c_i(1+n)+C+\alpha\mu)-(1+n)(-c_i(1+n)+C+\alpha\mu)}^{=0}}{\beta(1+n)(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}\\&+\sigma^2\frac{(1+n_N)(1+n_Y)(-c_i(1+n)+C+\alpha\mu)-(1+n)(-c_i(1+n_N)(1+n_Y)+C+C_Nn_Y+\alpha\mu)}{\beta(1+n)(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}\\=&\sigma^2\frac{n_Yn_N(-(1+n)\overline{c_N}+C+\alpha\mu)}{\beta(1+n)(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}\ge 0\end{aligned}\tag{EC.13}$$

after we rearrange the terms. Using (10), we obtain $\pi_i(\varnothing)\ge\pi_i(Y)$. □

Proof of Proposition 11. Let $q_i(Y)$ denote the equilibrium planting area of farmer i when the set of enrolled farmers is Y. By Proposition 10, $\max\{0, \pi_i(\varnothing) - \pi_i(Y)\} = 0, \forall i \in Y$ and $\max\{0, \pi_i(\varnothing) - \pi_i(Y)\} = \pi_i(\varnothing) - \pi_i(Y)$, $\forall i \in N$. Using (10), $\forall i \in N$, $\pi_i(\varnothing) - \pi_i(Y) = [q_i(\varnothing) - q_i(Y)][q_i(\varnothing) + q_i(Y)]\beta(\mu^2 + \sigma^2)$. Then, using (EC.13) in the proof of Proposition 10, we obtain

$$\frac{d(q_i(\varnothing) - q_i(Y))}{dC_N} = -\frac{\sigma^2 n_Y(1+n)}{\beta(1+n)(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}$$
$$\frac{d(q_i(\varnothing) + q_i(Y))}{dC_N} = \frac{\sigma^2 n_Y(1+n)}{\beta(1+n)(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}.$$

Hence,

$$\begin{aligned}\frac{d(\pi_i(\varnothing) - \pi_i(Y))}{dC_N} =& \beta(\mu^2+\sigma^2)\left[\frac{(q_i(\varnothing)-q_i(Y))\sigma^2 n_Y(1+n) - (q_i(\varnothing)+q_i(Y))\sigma^2 n_Y(1+n)}{\beta(1+n)(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}\right] \\ =& -\beta(\mu^2+\sigma^2)\left[\frac{2q_i(\varnothing)\sigma^2 n_Y(1+n)}{\beta(1+n)(\mu^2+\sigma^2)(\sigma^2(1+n_N)(1+n_Y)+(1+n)\mu^2)}\right] \le 0.\end{aligned} \tag{EC.14}$$

Therefore, $\pi_i(\varnothing) - \pi_i(Y)$ is minimized when C_N is the largest. Hence, for any fixed n_Y, setting $Y = Y^{LC}$ yields the lowest drop in profits from the yield-improvement program. □

Proof of Proposition 12. Let $Y = Y^{LC}$ and $N = N^{LC}$ and $Y^- = Y \backslash \{j\}$ and $N^- = N \backslash \{k\}$. Further, let $Y' = Y^- \cup \{k\}$ and $N' = N^- \cup \{j\}$ denote alternative sets with corresponding profits $\{\pi'_1, \pi'_2, ...\pi'_{j-1}, \pi'_k\}$ and $\{\pi'_{j+1}, ..., \pi'_{k-1}, \pi'_j, \pi'_{k+1}, ..., \pi'_n\}$.

a) By Lemma 2, we obtain that for any $i \in N^-$,

$$\frac{\pi_i}{\pi'_i} = \left(\frac{\mu^2(-c_i(1+n)+C+\alpha\mu)+\sigma^2(-c_i(1+n_N)(1+n_Y)+C_{Y^-}+c_j+(C_{N^-}+c_k)(n_Y+1)+\alpha\mu)}{\mu^2(-c_i(1+n)+C+\alpha\mu)+\sigma^2(-c_i(1+n_N)(1+n_Y)+C_{Y^-}+c_k+(C_{N^-}+c_j)(n_Y+1)+\alpha\mu)}\right)^2.$$

Since $c_k \ge c_j$, $\pi_i \ge \pi'_i$. The bound becomes tight as $c_j \to c_k$.

b) We denote, $\Delta\Pi_{j,k} = \pi_j + \pi_k - \pi'_j - \pi'_k$. We derive the lower bound by minimizing this function across all problem instances by considering one parameter at a time:

$$\min_{\alpha,\beta,n_Y,n,\mu,\sigma,c_j,c_k,c_n,C_{Y^-}\ge 0} \; \min_{C_{N^-}\ge 0} \; \Delta\Pi_{j,k}$$

$$\text{subject to } -c_n(n-n_Y+1)(n_Y+1)+C_{Y^-}+c_j+(C_{N^-}+c_k)(n_Y+1)+\alpha\mu \ge 0 \tag{EC.15}$$
$$C_{Y^-} \le c_j(n_Y-1) \tag{EC.16}$$
$$c_j \le c_k \tag{EC.17}$$
$$c_k \le c_n \tag{EC.18}$$
$$c_j(n-n_Y-1) \le C_{N^-} \tag{EC.19}$$
$$C_{N^-} \le c_n(n-n_Y-1) \tag{EC.20}$$
$$n_Y \le n-1 \tag{EC.21}$$
$$n_Y \ge 1. \tag{EC.22}$$

Here, constraint (EC.15) ensures that Assumption 2' holds; constraints (EC.16)-(EC.20) guarantee that farmer j has a higher cost than all other farmers in Y but a lower cost than all farmers in N, including k, while farmer n has the highest cost over all other farmers; constraints (EC.21)-(EC.22) restrict the sizes of the sets so that the size of the set N is greater than or equal to 1 (since $k \in N$), and the size of the set Y is greater than or equal to 1 (since $j \in Y$). Because

$$\frac{d\Delta\Pi_{j,k}}{dC_{N^-}} = -\frac{2(c_k - c_j)\sigma^2((1+n_Y(n-1)+n)\mu^2 + \sigma^2 + (1+n_Y+n_Y^2)(n-n_Y)\sigma^2)}{\beta(\mu^2+\sigma^2)((1+n)\mu^2 + (1+n_Y)(n-n_Y+1)\sigma^2)^2} \leq 0,$$

C_{N^-} should be increased as much as possible, i.e., until (A.8) is tight. At that point, C_{N^-} is equal to

$$C_{N^-} = c_n(n - n_Y - 1).$$

Next, we substitute this optimal value of C_{N^-} in the objective function and the remaining constraints and pursue the optimization problem with the other variables:

$$\begin{aligned}
\min_{\alpha,\beta,n_Y,n,\mu,\sigma,c_j,c_k,c_n \geq 0} \quad \min_{C_{Y^-} \geq 0} \quad & \Delta\Pi_{j,k} \\
\text{subject to} \quad & -2c_n(n_Y+1) + C_{Y^-} + c_j + c_k(n_Y+1) + \alpha\mu \geq 0 && \text{(EC.23)} \\
& C_{Y^-} \leq c_j(n_Y - 1) && \text{(EC.24)} \\
& c_j \leq c_k && \text{(EC.25)} \\
& c_k \leq c_n && \text{(EC.26)} \\
& n_Y \leq n - 1 && \text{(EC.27)} \\
& n_Y \geq 1. && \text{(EC.28)}
\end{aligned}$$

Because

$$\frac{d\Delta\Pi_{j,k}}{dC_{Y^-}} = \frac{2(c_k - c_j)\sigma^2((n+n_Y+n(n-n_Y))\mu^2 + n_Y(n-n_Y+1)^2\sigma^2)}{\beta((1+n)\mu^3 + (n_Y+1)(n-n_Y+1)\mu\sigma^2)^2} \geq 0,$$

C_{Y^-} should be made as small as possible, i.e., until either (EC.23) becomes tight or C_{Y^-} becomes non-positive. When (EC.23) is binding, C_{Y^-} is equal to

$$C^*_{Y^-} \equiv 2c_n(n_Y+1) - c_j - c_k(n_Y+1) - \alpha\mu.$$

Thus,

$$C_{Y^-} = \max\{0, C^*_{Y^-}\}.$$

Suppose first that $C^*_{Y^-} \geq 0$. After substituting C_{Y^-} with $C^*_{Y^-}$ into the objective function and setting $n_N = n - n_Y$, the objective function becomes

$$\Delta\Pi_{j,k} = \frac{\sigma^2(c_k - c_j)z(c_i, \ldots)}{(\beta(\mu^2+\sigma^2)((1+n_N+n_Y)\mu^3 + (1+n_N)(1+n_Y)\mu\sigma^2)^2)} \geq 0$$

where

$$\begin{aligned}z(c_i,...) = {} & 2(c_n-c_k)(((1+n_N)^2+(1+n_N)(4+n_N)n_Y+(5+n_N)n_Y^2)\mu^4 \\ & +(1+n_N)(1+n_N+6n_Y(1+n_Y)+n_Nn_Y(3+n_Y))\mu^2\sigma^2+2(1+n_N)^2n_Y(1+n_Y)\sigma^4) \\ & +(c_k-c_j)((1+n_N+n_Y)(1+n_N+3n_Y)\mu^4 \\ & +((1+n_N)^2+2(1+n_N)(3+n_N)n_Y+(3+4n_N)n_Y^2)\mu^2\sigma^2+(1+n_N)^2n_Y(2+n_Y)\sigma^4).\end{aligned}$$

Therefore, $\Pi_{j,k}\geq 0$ if $C_Y^*\geq 0$. Suppose next that $C_Y^*\leq 0$. After substituting C_{Y-} with 0, we obtain:

$$\begin{aligned}\min_{\alpha,\beta,n_Y,n,\mu,\sigma,c_j,c_k\geq 0}\quad \min_{c_n\geq 0}\quad & \Delta\Pi_{j,k} \\ \text{subject to } & -2c_n(n_Y+1)+c_j+c_k(n_Y+1)+\alpha\mu\geq 0 && \text{(EC.29)} \\ & c_j\leq c_k && \text{(EC.30)} \\ & c_k\leq c_n && \text{(EC.31)} \\ & n_Y\leq n-1 && \text{(EC.32)} \\ & n_Y\geq 1, && \text{(EC.33)}\end{aligned}$$

where (EC.29) turns out to be equal to $C_{Y-}^*\leq 0$. Because

$$\frac{d\Delta\Pi_{j,k}}{dc_n}=-\frac{2(-1+n_N)(c_k-c_j)\sigma^2((1+n_N+n_Nn_Y+n_Y^2)\mu^2+(1+n_N(1+n_Y+n_Y^2))\sigma^2)}{\beta(\mu^2+\sigma^2)((1+n_N+n_Y)\mu^2+(1+n_N)(1+n_Y)\sigma^2)^2}\leq 0,$$

(here, $n_N=n-n_Y$), c_n should be increased as much as possible, i.e., until (EC.29) is tight. At that point, c_n is equal to

$$c_n=\frac{c_j+c_k(n_Y+1)+\alpha\mu}{2(n_Y+1)}.$$

We note that the minimizing c_n satisfies $C_{Y-}^*=0$, which recovers the earlier case. Therefore, in either case $\Delta\Pi_{j,k}\geq 0$.

c) Let $C_{Y=}=C_Y-c_j-c_i$ denote the sum of costs of farmers in Y, except the costs of farmers j and i. We consider minimizing $\sqrt{\frac{\pi_i}{\pi_i'}}$, for $i\in Y^-$ across all problem instances, considering one parameter at a time.

$$\begin{aligned}\min_{\alpha,\beta,\mu,\sigma,n_Y,n,c_i,c_j,c_n,C_{Y=},C_{N-}\geq 0}\quad \min_{c_k\geq 0}\quad & \sqrt{\frac{\pi_i}{\pi_i'}} \\ \text{subject to } & -c_n(n-n_Y+1)(n_Y+1)+C_{Y=}+c_j+c_i+(C_{N-}+c_k)(n_Y+1)+\alpha\mu\geq 0 && \text{(EC.34)} \\ & C_{Y=}\leq c_j(n_Y-2) && \text{(EC.35)} \\ & c_i\leq c_j && \text{(EC.36)} \\ & c_j\leq c_k && \text{(EC.37)} \\ & c_k\leq c_n && \text{(EC.38)}\end{aligned}$$

$$c_j(n-n_Y-1) \le C_{N^-} \tag{EC.39}$$

$$C_{N^-} \le c_n(n-n_Y-1) \tag{EC.40}$$

$$n_Y \le n-1 \tag{EC.41}$$

$$n_Y \ge 2. \tag{EC.42}$$

Here, constraint (EC.34) ensures that Assumption 2' holds; constraints (EC.35)-(EC.40) guarantee that farmer j has a higher cost than all other farmers in Y but a lower cost than all farmers in N, including farmer k, and that farmer n has the highest cost over all other farmers; constraints (EC.41)-(EC.42) restrict the sizes of the sets so that the size of N is greater than or equal to 1 (since $k \in N$) and the size of Y is greater than or equal to 2 (since j and $i \in Y$).

The derivative of $\sqrt{\frac{\pi_i}{\pi_i'}}$ with respect to c_k is equal to

$$\frac{d\sqrt{\frac{\pi_i}{\pi_i'}}}{dc_k} = -\frac{(n-n_Y+1)(\mu^2\sigma^2(2c_j - c_i n + C_{Y^=} + C_{N^-} + \alpha\mu) + \sigma^4(n-n_Y+1)(c_j - c_i n_Y + C_{Y^=} + \alpha\mu))}{(\mu^2(c_j + c_k - c_i n + C_{Y^=} + C_{N^-} + \alpha\mu) + (1-n_Y+n)(c_k - c_i n_Y + C_{Y^=} + \alpha\mu)\sigma^2)^2} \le 0.$$

This inequality is satisfied since Lemma EC.1 ensures non-negativity of the coefficient of $\mu^2\sigma^2$ in the numerator as

$$\begin{aligned} 2c_j - c_i n + C_{Y^=} + C_{N^-} + \alpha\mu &= c_j - c_k - c_i(n+1) + C_Y + C_N + \alpha\mu \\ &\ge c_j - c_k - c_i(n+1) + c_k(n+1) = \underbrace{(c_j - c_i)}_{\ge 0} + n\underbrace{(c_k - c_i)}_{\ge 0} \ge 0; \end{aligned}$$

here, the last inequality is because $c_k \ge c_j \ge c_i$, and Lemma EC.2 ensures non-negativity of the coefficient of σ^4 in the numerator as

$$c_j - c_i n_Y + C_{Y^=} + \alpha\mu = -c_i(n_Y+1) - C_Y + \alpha\mu \ge 0.$$

We note that (EC.34) is a sufficient condition for Lemma EC.1 and EC.2 to hold. Thus, c_k must be increased as much as possible, i.e., until (EC.38) becomes tight. Therefore, the minimizing value of c_k is equal to c_n.

After substituting the minimizing value of c_k into the objective function and constraints, the optimization problem reduces to

$$\min_{\alpha,\beta,\mu,\sigma,n_Y,n,c_j,c_n,C_{Y^=},C_{N^-}\ge 0} \quad \min_{c_i \ge 0} \quad \sqrt{\frac{\pi_i}{\pi_i'}}$$

$$\text{subject to } -c_n(n-n_Y)(n_Y+1) + C_{Y^=} + c_j + c_i + C_{N^-}(n_Y+1) + \alpha\mu \ge 0 \tag{EC.43}$$

$$C_{Y^=} \le c_j(n_Y-2) \tag{EC.44}$$

$$c_i \le c_j \tag{EC.45}$$

$$c_j \le c_n \tag{EC.46}$$

$$c_j(n-n_Y-1) \le C_{N^-} \tag{EC.47}$$

$$C_{N^-} \le c_n(n-n_Y-1) \tag{EC.48}$$

$$n_Y \le n-1 \tag{EC.49}$$

$$n_Y \ge 2. \tag{EC.50}$$

Because

$$\frac{\sqrt{\frac{\pi_i}{\pi_i'}}}{dc_i} = -\frac{(c_n-c_j)(n-n_Y+1)\sigma^2(n\mu^2+n_Y(n-n_Y+1)\sigma^2)}{(\mu^2(c_j+c_n-c_in+C_{Y^=}+C_{N^-}+\alpha\mu)+(n-n_Y+1)(-c_in_Y+c_n+C_{Y^=}+\alpha\mu)\sigma^2)^2} \le 0,$$

c_i must be increased as much as possible, i.e., until (EC.45) becomes tight. Therefore, the minimizing value of c_i is equal to c_j. After substitution of the optimizing value of c_i into the objective function and the constraints, the optimization problem reduces to

$$\min_{\alpha,\beta,\mu,\sigma,n_Y,n,c_j,C_{Y^=},C_{N^-}\ge 0} \quad \min_{c_n\ge 0} \sqrt{\frac{\pi_i}{\pi_i'}}$$

$$\text{subject to } -c_n(n-n_Y)(n_Y+1)+C_{Y^=}+2c_j+C_{N^-}(n_Y+1)+\alpha\mu \ge 0 \tag{EC.51}$$

$$C_{Y^=} \le c_j(n_Y-2) \tag{EC.52}$$

$$c_j \le c_n \tag{EC.53}$$

$$c_j(n-n_Y-1) \le C_{N^-} \tag{EC.54}$$

$$C_{N^-} \le c_n(n-n_Y-1) \tag{EC.55}$$

$$n_Y \le n-1 \tag{EC.56}$$

$$n_Y \ge 2. \tag{EC.57}$$

The derivative of $\sqrt{\frac{\pi_i}{\pi_i'}}$ with respect to c_n is equal to

$$\frac{\sqrt{\frac{\pi_i}{\pi_i'}}}{dc_n} = -\frac{(n-n_Y+1)(\sigma^2\mu^2(-c_j(n-2)+C_{Y^=}+C_{N^-}+\alpha\mu)+\sigma^4(n-n_Y+1)(-c_j(n_Y-1)+C_{Y^=}+\alpha\mu))}{(\mu^2(c_j+c_n-c_jn+C_{Y^=}+C_{N^-}+\alpha\mu)+(n-n_Y+1)(-c_jn_Y+c_n+C_{Y^=}+\alpha\mu)\sigma^2)^2} \le 0.$$

This inequality is satisfied since Lemma EC.1 ensures non-negativity of the coefficient of $\mu^2\sigma^2$ in the numerator as

$$\begin{aligned}&-c_j(n-2)+C_{Y^=}+C_{N^-}+\alpha\mu = -c_j(n+1)-c_k+c_j+C_Y+C_N+\alpha\mu\\ &\ge -c_j(n+1)-c_k+c_j+c_k(n+1) = n\underbrace{(c_k-c_j)}_{\ge 0} \ge 0;\end{aligned}$$

here the last inequality is because $c_k \ge c_j$, and Lemma EC.2 ensures non-negativity of the coefficient of σ^4 in the numerator as

$$-c_j(n_Y-1)+C_{Y^=}+\alpha\mu = -c_j(n_Y+1)+C_Y+\alpha\mu \ge 0.$$

We note that (EC.51) is a sufficient condition for Lemma EC.1 and EC.2 to hold. Thus, c_n must be increased as much as possible, i.e., until (EC.51) becomes tight. Therefore, the minimizing value of c_n is equal to

$$c_n = \frac{C_{Y=} + 2c_j + C_{N-}(n_Y+1) + \alpha\mu}{(n-n_Y)(n_Y+1)}.$$

After substituting the optimizing value of c_n in the objective function and the constraints, the optimization problem reduces to

$$\begin{aligned}
&\min_{\alpha,\beta,\mu,\sigma,n_Y,n,c_j,C_{Y=}\geq 0} \quad \min_{C_{N-}\geq 0} \sqrt{\frac{\pi_i}{\pi_i'}} && \\
&\text{subject to } C_{Y=} \leq c_j(n_Y-2) && \text{(EC.58)}\\
&\qquad c_j[(n-n_Y)(n_Y+1)-2] \leq C_{Y=} + C_{N-}(n_Y+1) + \alpha\mu && \text{(EC.59)}\\
&\qquad c_j(n-n_Y-1) \leq C_{N-} && \text{(EC.60)}\\
&\qquad C_{N-} \leq (C_{Y=} + 2c_j + \alpha\mu)\frac{(n-n_Y-1)}{n_Y+1} && \text{(EC.61)}\\
&\qquad n_Y \leq n-1 && \text{(EC.62)}\\
&\qquad n_Y \geq 2. && \text{(EC.63)}
\end{aligned}$$

The derivative of $\sqrt{\frac{\pi_i}{\pi_i'}}$ with respect to C_{N-} is equal to

$$\frac{d\sqrt{\frac{\pi_i}{\pi_i'}}}{dC_N^-} = -\frac{(n-n_Y+1)(-c_j(n_Y-1)+C_{Y=}+\alpha\mu)\sigma^2(n_Y\mu^2+(1+n_Y)(n-n_Y+1)\sigma^2)}{(n_Y+1)(n-n_Y)[f(c_j,...)]^2} \leq 0$$

where

$$\begin{aligned}
f(c_j,...) = {}& \mu^2\left(c_j - c_j n + C_{Y=} + C_{N-} + \alpha\mu + \frac{2c_j + C_{Y=} + C_{N-} + n_Y C_{N-} + \alpha\mu}{(1+n_Y)(n-n_Y)}\right)\\
&+ \sigma^2(1-n_Y+n)\left(-c_j n_Y + C_{Y=} + \alpha\mu + \frac{2c_j + C_{Y=} + C_{N-} + n_Y C_{N-} + \alpha\mu}{(1+n_Y)(-n_Y+n)}\right).
\end{aligned}$$

The numerator is non-negative since (EC.61) can be rearranged as $-c_j(n_Y-1) + C_{Y=} + \alpha\mu \geq (n_Y+1)(\overline{c_N} - c_j)$; and the right-hand side is non-negative because $\overline{c_N} \geq c_j$. Therefore, C_{N-} should be maximized. The minimizing C_{N-} makes (EC.61) binding, i.e.,

$$C_{N-} = (2c_j + C_{Y=} + \alpha\mu)\frac{(n-n_Y-1)}{n_Y+1}.$$

After substituting the optimizing C_{N-} in the objective function and the constraints, the optimization problem reduces to

$$\begin{aligned}
&\min_{\mu,\sigma,n\geq 0} \quad \min_{n_Y\geq 0} \frac{(1+n)\mu^2 + (1-n_Y^2+n+n_Y n)\sigma^2}{(1+n)\mu^2 + (2+n_Y)(n-n_Y+1)\sigma^2} && \\
&\text{subject to } n_Y \leq n-1 && \text{(EC.64)}\\
&\qquad n_Y \geq 2. && \text{(EC.65)}
\end{aligned}$$

The derivative of the objective function with respect to n_Y is equal to

$$\frac{(1+n)\mu^2\sigma^2+(1-n_Y+n)^2\sigma^4}{((1+n)\mu^2+(2+n_Y)(n-n_Y+1)\sigma^2)^2}\geq 0.$$

Therefore, n_Y needs to be minimized. Minimizing n_Y makes (EC.65) binding, i.e., $n_Y=2$. After substituting the optimizing n_Y in the objective function and the constraints, the optimization problem reduces to

$$\min_{\mu,\sigma\geq 0} \quad \min_{n\geq 0} \frac{(n+1)\mu^2+3(n-1)\sigma^2}{(1+n)\mu^2+4(n-1)\sigma^2}$$
$$\text{subject to } n_Y\leq n-1.$$

The derivative of the objective function with respect to n is equal to

$$-\frac{2\mu^2\sigma^2}{((1+n)\mu^2+4(-1+n)\sigma^2)^2}\leq 0.$$

Therefore, n needs to be maximized. As n approaches to infinity, the objective function approaches to $\frac{\mu^2+3\sigma^2}{\mu^2+4\sigma^2}$. Consequently, if we set $\xi=\frac{\sigma^2}{\mu^2}$, $\sqrt{\frac{\pi_i}{\pi_i'}}$ takes its minimum value as ξ goes to infinity, making the objective function equal to $\frac{3}{4}$. Hence, $\frac{\pi_i}{\pi_i'}\geq\frac{9}{16}$.

Combining these results, we obtain

$$\sum_i \pi_i=\sum_{i\in Y^-}\pi_i+\sum_{i\in N^-}\pi_i+\pi_j+\pi_k\geq\frac{9}{16}\sum_{i\in Y^-}\pi_i'+\sum_{i\in N^-}\pi_i'+\pi_j'+\pi_k'\geq\frac{9}{16}\sum_i\pi_i'.$$

□