Direct Trade Sourcing Strategies for Specialty Coffee

Scott Webster
scott.webster@asu.edu
W.P. Carey School of Business
Arizona State University
Tempe, AZ 85287

Burak Kazaz
bkazaz@syr.edu
Whitman School of Management
Syracuse University
Syracuse, NY 13244

Shahryar Gheibi
sgheibi@siena.edu
School of Business
Siena College
Loudonville, NY 12211

April 18, 2024

Forthcoming in *Manufacturing & Service Operations Management*

Direct Trade Sourcing Strategies for Specialty Coffee

Problem definition: Leading specialty coffee roasters rely on *direct trade* (DT) to source premium coffee beans. We examine a roaster who sells two basic types of roasts: (1) a single-origin roast is sourced from a specific locale, (2) a blend roast uses a mix of beans from sources that vary over the course of a year. The price of blend roasts is lower than single-origin roasts and appeal to a larger market. We study how characteristics of the operating and market environment affect the optimal sourcing strategy for single-origin beans.
Methodology/results: We develop a two-stage stochastic program with recourse that reflects these characteristics. A roaster has the option to allocate some of the single-origin beans for sale under a blend label, known as downward substitution. We identify three distinct optimal sourcing strategies: specialized (no downward substitution), diversified (consistent downward substitution), and mixed (between these extremes) and show that they are robust under different definitions of yield and demand.
Managerial implications: We identify four main insights: (1) Two factors determine which strategy is optimal: mean price of the inferior product (blend label) and the marginal cost of the superior product (single-origin label). (2) When compared with the newsvendor model, we find distinct structural differences across strategies. For example, while the effects of increasing uncertainty on optimal quantity align with the newsvendor model under a mixed strategy, the effects are distinctly different under specialized and diversified strategies (e.g., monotonic decreasing behavior for specialized, no change in quantity under diversified). (3) The weighted average price of an agricultural product is decreasing in negative yield-price correlation. We coin this as "farmer's curse," which carries lessons for direct trade sourcing (e.g., advocating against paying the grower at post-harvest market prices). (4) We find evidence of a virtuous feedback loop wherein the grower-roaster relationship becomes stronger over time. Our findings also point to a simple signal that policymakers may use to identify coffee growing locales where targeted interventions can improve grower welfare.

Keywords: *supply chain management*, *direct trade, sourcing*, *pricing*, *yield uncertainty*, *price uncertainty*

1. Introduction

Background and motivation. Coffee is one of the most popular beverages worldwide. The United States Department of Agriculture estimates the global annual coffee consumption at 175 million bags (60 kilograms per bag) in its 2022 report.[1] While the coffee consumption is expected to continue to experience a steady growth, consumers seem to increasingly demand high quality coffees around the world (Stabiner 2015, Craymer 2015). Some coffee roasters have targeted the specialty coffee market and are reshaping the coffee industry (Strand 2015). Pioneer roasters in the specialty coffee industry include Counter Culture, Intelligentsia, and Stumptown. These roasters, which are referred to as the Big Three, were early drivers of initiatives aimed at enhancing the quality of the produced coffee. The specialty coffee bean market is significant, estimated at 50% of the global value of traded coffee (Rafael 2020). Interest in the segment is reflected in the recent acquisitions of Intelligentsia and Stumptown by Peet's Coffee at high prices.

[1] https://apps.fas.usda.gov/psdonline

Roasters employ a sourcing practice called *direct trade* (DT) to obtain specialty coffee. DT relies on direct communication and close collaboration with growers (e.g., providing guidance and resources to the growers, and monitoring closely the growing and harvesting process, Arellano 2016). The traditional coffee bean supply chain has many layers of entities; DT shortens this supply chain by trading with the grower.

Specialty roaster market segments and operational characteristics. Specialty roasters sell two basic types of roasts: *blend* and *single-origin* (Wernau 2015). The key difference is concisely captured in labeling that appears on each roast sold by Verve, a specialty roaster headquartered in California: "blend – balanced, consistent, year-round; single-origin – vibrant, unique, limited." Blends are available year-round using a mix of beans and a roast profile to achieve consistent flavor throughout the year. The sources of beans vary during the year according to available harvests, and may be supplemented with spot-market purchases. In contrast, roasts under a single-origin label are specific to a particular grower or locale—known as micro-lot single-origin roast. A single-origin label is available for a limited portion of the year and commands a price premium over blend labels. For example, the single-origin price premium is around 30%.[2] Quality is expected to be very high and, by nature of being specific to a grower/locale, single-origin labels offer distinct flavor profiles. Single-origin labels appeal to a segment of consumers who place a high value on quality in their purchases. Blend labels, on the other hand, have wider market appeal and much higher annual volume (e.g., at Counter Culture blend-label sales volume is 5 to 6 times higher than single-origin volume).

Quality of coffee beans is determined by the industry-standard cupping protocol established by the Specialty Coffee Association of America's (SCAA) 100-point scale. A score above 90 is classified in the rarified premium category with one pound of roasted beans retailing for more than $25 (Fischer 2017). By comparison, commodity coffee beans score around 75-80 and commonly retail for less than $10/pound. Coffee beans with a cupping score above 80 are considered specialty grade (e.g., Starbucks beans generally score above 80; Fischer 2017).

Quality influences the cost of green (unroasted) specialty coffee beans. Examining a specialty roaster's 2011-2019 data, we observe that one of their popular blends contained mixes of beans with cupping scores that ranged from 84 to 87.5 whereas single-origin beans range in the upper 80s or 90s, including one lot with an extraordinary score of 95. Over the years 2019 – 2021, we observe that the specialty roaster pays an average of 23% more for single-origin-label beans than blend-label beans.

Research questions, objective, and contributions. Our study considers a specialty coffee roaster's optimal acreage contracting decision for a single-origin label. Specialty coffee roasters make this decision under two forms of uncertainty. The first is the random yield in coffee bean harvest. The second is the price

[2] In January 2021, we collected data on the retail prices of all year-round (nonseasonal) blend and single-origin roasts, excluding decaffeinated roasts. The average percentage price premiums for single-origin roasts over blend labels by roaster: Counter Culture (24%), Intelligentsia (32%), Stumptown (29%).

of commodity-grade coffee bean prices, which influences the roaster's selling price of specialty-coffee beans. The commodity-grade coffee bean prices exhibit remarkable price fluctuations over time due to high variation in global yield. Figure 1 shows that the commodity-grade coffee bean prices exhibit more than a 135% difference between the maximum price of \$2.07 (July 2021) and the minimum price of \$0.87 (May 2019) during this five-year period. Climate change is expected to continue to bring about wide fluctuations in the price of commodity coffee beans and the retail price of (roasted) coffee (Osborn 2015). Recent low yields in Brazil due to extreme weather and COVID-related supply chain snarls have contributed to the recent rapid increase in coffee prices at both wholesale and retail levels (Horner and Lewis 2021, Marcos 2021, Orr 2021). As these observations suggest, the yield and commodity-grade coffee bean prices can be negatively correlated. Therefore, it is essential to account for harvest yield and coffee bean price uncertainties.

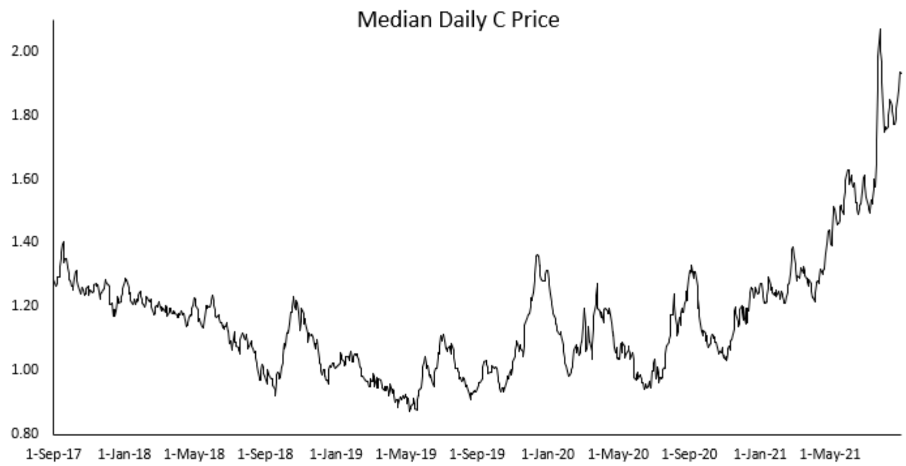


Figure 1. Median commodity-grade coffee bean prices (\$/pound) represented by the futures prices (known as C price) on the New York exchange from 9/1/2017 to 8/31/2021.

Single-origin beans, by virtue of high quality, can be roasted and sold under a single-origin label or mixed with other beans and sold under a blend label at a lower price. This option of shifting a high-quality input for sale in a lower quality output, which is known as *downward substitution*, is potentially valuable to a specialty roaster. Our study identifies when a specialty roaster should make use of the flexibility associated with downward substitution. We provide insights into managing a specialty-coffee supply chain by investigating the following research questions:

1. How should a specialty roaster set DT contract acreage for a single-origin label?
2. After the single-origin harvest yield and market uncertainties are revealed, how should a specialty roaster allocate the harvested beans among single-origin label (SL) and blend labels (BL), and determine the selling price of SL?

3. What is the impact of market and environment characteristics on sourcing strategies of a roaster in a specialty coffee supply chain?

Answers to the above questions underlie our main research objective: We seek to clarify the relationship between characteristics of a specialty roaster's operating environment and effective sourcing strategies.

Our paper shows that there are three distinct sourcing strategies that emerge as optimal:

1. *Specialized sourcing strategy*. This is a conservative sourcing strategy with respect to contract acreage for a single-origin label (SL). The optimal contract acreage is such that there is no chance of downward substitution, i.e., using superior SL beans in inferior BL. Thus, the option value of downward substitution is zero.
2. *Diversified sourcing strategy*. This strategy is an aggressive strategy where the optimal acreage for a single-origin label is high. The flexibility of downward substitution, i.e., using SL beans in BL, is utilized in all realizations of uncertainty.
3. *Mixed sourcing strategy*. This strategy represents a middle ground between the first two sourcing strategies. The probability of downward substitution is between zero and 100%.

We formulate a two-stage decision model under yield and price uncertainty. In the first stage, the specialty roaster decides the amount of DT acreage for the single-origin label. The second-stage decision occurs after random yield is realized and the price that the blend-label can fetch in the market is known: the roaster must decide the quantity to sell as single-origin label (SL) and the quantity to sell as blend label (BL). The second-stage problem is, in effect, a capacity-constrained pricing problem under multinomial logit (MNL) demand. This two-stage stochastic optimization problem with recourse is nontrivial. We show that our first-stage objective function is assured to be concave in the first-stage decision. Consequently, our model is amenable to classical analytical machinery (e.g., implicit function theorem, envelop theorem, etc.) to deduce structural properties of the system, and standard nonlinear optimization procedures are assured to return an optimal solution. While the model is complex, we identify the exact conditions under which each strategy emerges as the optimal choice. We present a comprehensive analysis pertaining to the impact of changing the yield variance. Finally, we show that the three distinct sourcing strategies – specialized, mixed, and diversified – are robust as they continue to appear as the optimal strategies under alternative representations of yield uncertainty and demand.

Our paper develops four main insights. First, we identify that there are two primary factors that determine the optimal sourcing strategy: mean price of inferior product BL and the marginal cost of the superior product SL. An increase in the mean BL price encourages the roaster to increase the optimal acreage. The policy preference switches from the specialized sourcing strategy to mixed, and eventually to, diversified sourcing strategy with increasing values of the mean BL price. The marginal cost has the opposite effect as it encourages the roaster to reduce the acreage commitment. The preference switches from the diversified sourcing to mixed and to specialized sourcing strategies with increasing marginal costs. It is important to note that changes in all other factors have an ambiguous effect on the optimal strategy.

The second insight stems from comparing and contrasting with the well-known newsvendor model where demand is random, rather than the yield. In this comparison, we focus on the impact of increasing volatility in the respective random variable. While the optimal quantity and profit under the mixed strategy follows a behavior similar to the characteristics of the newsvendor model, we identify distinct structural differences. In the newsvendor problem, the optimal order quantity can be increasing or decreasing depending on the problem parameters. Under the specialized sourcing strategy, however, the optimal (acreage) quantity is strictly decreasing as the variance in the yield increases. Moreover, the optimal quantity and the profit under the diversified strategy does not change with yield uncertainty in the absence of yield and blend price correlation. Our study presents a comprehensive discussion about the underlying reasons for the departure in characterization of the optimal quantity and profit in our problem.

Third, the weighted average price of the inferior product BL decreases in negative yield-price correlation. We refer to this effect as the “farmer’s curse.” There are lessons for growers and roasters associated with the farmer’s curse. First, reducing the negative correlation between the grower’s yield and BL price increases the profit. This can be accomplished with investments or actions that reduce the sensitivity of grower’s yield to weather and other environmental conditions relative to other growers (i.e., by reducing negative price-yield covariance for the grower). A second lesson relates to the timing of setting the unit price paid to the grower. Some roasters pay growers according to the market price at the time of harvest, which is known as differential pricing. Other roasters set the unit price paid to the grower prior to the growing season arguing that this benefits the grower by reducing risk due to price uncertainty. However, an additional, perhaps less known, advantage is that setting the purchase price well before yields are realized mitigates negative price-yield correlation (and the farmer’s curse) leading higher profit for the grower, all else equal. This second lesson may encourage more specialty roasters to move away from differential pricing practices and adopt fixed prices that are set prior to the growing season.

Fourth, our study presents evidence of a virtuous feedback loop wherein the grower-roaster relationship becomes stronger over time. Our comprehensive examination of the impact of yield variance shows that the roaster operating under a specialized or mixed sourcing strategy is incentivized to work with the grower to reduce the volatility of yields (via a combination of improved cultivation methods, harvesting practices, post-harvest processing and storage). These actions shift the strategy more towards specialized sourcing, which further increases the incentive of the roaster to invest and work closely with the grower. Our conversations with specialty roaster executives suggest that specialized sourcing may be relatively common in practice (e.g., SL beans sell out well before the shelf life of the green beans), though we leave this as a question for possible empirical investigation. In addition, our findings also point to a simple signal that policymakers may use to identify coffee growing locales where targeted interventions can improve grower welfare.

Our analyses and results extend beyond specialty coffee to other settings (e.g., wine production, and semiconductor manufacturing) where the following characteristics are present: (1) a firm's input quantity decision is made in the presence of uncertain supply and uncertain demand, (2) the firm offers products that compete with one another and differ by quality, and (3) during the production process, the firm has flexibility to divert high-quality inputs to a lower quality product in the product line.

2. Related Literature

This study examines sourcing and pricing decisions of an agricultural firm under random yield and market-price for a similar (and inferior) product. The impact of supply uncertainty on sourcing strategies is extensively examined in supply chain literature. Yano and Lee (1995) and Tang (2006) provide a comprehensive review of this literature in manufacturing environments. Our paper is related with the stream of research that explores operational hedging strategies for managing supply uncertainty. In this literature, supply diversification is proven to be an effective approach toward mitigating supply risk: Gerchak and Parlar (1990), Parlar and Wang (1993), Anupindi and Akella (1993), and more recently, Tomlin and Wang (2005), Dada et al. (2007), Burke et al. (2009), Jain et al. (2014), and Tan et al. (2016) demonstrate the benefits of dual/multiple sourcing in managing supply uncertainty. However, in these studies the firm (buyer) does not have pricing power, and thus, the selling price is considered as exogenous. Our work characterizes sourcing strategies that differ from supply diversification.

There are a few studies that consider a firm's price-setting ability in the presence of supply uncertainty. Li and Zheng (2006), and Feng (2010) investigate joint inventory and pricing policies in multi-period manufacturing settings. Tang and Yin (2007) consider pricing and purchase quantity decisions under supply uncertainty. Kazaz and Webster (2015) consider a single-period newsvendor model under supply and demand uncertainty and explore how the source of uncertainty influence the tractability of the problem and the optimal decisions. Our study departs from these papers by considering the consumer choice between two products (single-origin and blend) and the effect of uncertainty in the competing market on sourcing, production, and pricing decisions of an agricultural processor.

Sourcing via DT resembles agricultural supply chain literature where agricultural firms lease farmland in order to grow products subject to yield uncertainty. Kazaz (2004), Kazaz and Webster (2011) examine the impact of yield-dependent trading cost on sourcing and production planning of an agricultural firm which sells a single product. Our study departs from these publications by considering two products—one superior and one inferior product—with two associated consumer segments. While our main focus is exploring the sourcing and pricing decisions of the superior product, our analysis captures the consumer's choice between the two products, which affects the demand for the superior product. Noparumpa et al. (2011) examine the interrelationships among three flexibilities: downward substitution, price setting, and fruit trading under supply and quality uncertainty. While we do not incorporate quality uncertainty into our model, we capture the impact of uncertainty in the price of an inferior product determined by the global market on sourcing and pricing of a superior product by an agricultural processor. Our model's structure is similar to that of Noparumpa et al. (2015) where a winemaker has to determine how much of its wine to be sold in advance in the form of wine futures, the price of wine futures, and the amount of wine to be sold later in the retail chain. Our study differs from their work in terms of the decisions, and the environmental factors influencing sourcing and pricing decisions: Our paper explores the impact of the interaction between two different consumer segments on allocation decisions of an agricultural processor.

There is a stream of literature that considers simultaneous sourcing of agricultural inputs (Boyabatli et al. 2011, Boyabatli 2015). However, this body of research investigates multi-product supply management where the crop yield is assumed to be deterministic using two or more different agricultural inputs. Our study is structurally different from these papers as we emphasize developing effective policies for sourcing a single agricultural input under yield and market-price uncertainty.

In sum, motivated by an emerging agribusiness practice in the coffee industry, our study features a new agricultural supply chain framework. To target the fast-growing segment of quality-sensitive consumers, a number of coffee roasters have drastically changed their sourcing policy and are engaged in DT, which in turn has led to exposure to supply risk. Our study investigates the sourcing and pricing decisions of such agricultural processors and provides insights into managing the supply chain of the associated agricultural products.

3. Model

3.1. Model Assumptions and Justifications

We first summarize the process leading up to the sale of roasted coffee beans sold under the single-origin label (SL). The roaster has established a DT relationship and payment terms with a grower for a particular SL in advance of the growing season. After harvesting and milling (pulping, fermenting, washing, drying, hulling), the quality of harvested beans is assessed. Beans that meet contracted quality specifications are purchased by the roaster, and the lot is shipped to the roaster's warehouse. Green beans can be stored for up to one year in cool/dry conditions prior to roasting without significant degradation in quality. The roasting schedule is set to keep the supply of the roast a little ahead of demand, up to the point when the lot is exhausted. The roaster can mix any desired portion of the single-origin-quality beans with other beans to be sold under a blend label.

As discussed in Section 1, roasted beans sold under blend labels (BLs) serve a larger market than SL beans. We capture this feature in our model by assuming that the market for the roaster's products is comprised of two distinct segments – a *connoisseur segment* and a *mass-market segment*. Consumers in the connoisseur segment, who place a high value on quality, choose between purchasing an SL roast and a BL roast from the specialty roaster, and an outside option. For example, the outside option choice may be the purchase of a roast from another specialty roaster, or no purchase at all. Consumers in the mass-market segment, which is notably larger, choose between purchasing a BL roast from the roaster and an outside option. Compared to the connoisseur segment, the mass-market segment is more price sensitive and less concerned about quality. The outside option for this segment is likely to be broader, including commercial coffee alternatives.

The real-world elements surrounding the roaster's supply quantity and pricing decisions of SL beans are many and nuanced. As a first step toward gaining insight into sourcing and pricing strategies, we study a stylized model of a single-origin label for a particular harvest cycle. Demand for the SL depends upon the SL price and the BL price as governed by the MNL model. The MNL model is widely used for modeling demand of differentiated products. It exhibits strong theoretical support because each individual chooses the alternative with highest utility. Furthermore, Guadagni and Little (1983) find remarkable predictive accuracy of the MNL model applied to coffee purchases. The authors calibrate the model using 32-weeks of ground coffee purchases from 100 households and evaluate accuracy against a holdout sample of 20 weeks.

Both the harvest yield and the blend price are uncertain at the time the DT contract is established with the grower. The determination of the BL price is dominated by considerations related to the mass-market segment (e.g., state of the mass-market segment, retail price of commercial coffees, etc.), and is not affected by downward substitution volume, e.g., downward substitution from a particular harvest of SL beans is a small fraction of annual blend volume.

Stage one in our model is the contracting stage. The roaster determines the amount of acreage for the SL product. The upfront payment per acre and the price to be paid for each pound of beans are exogenous to the model, e.g., driven by broader considerations that are reflected in the roaster's DT contracting policies. The price per pound to be paid to the grower(s) is fixed at the time of contracting.[3]

Stage two in our model occurs after harvest yield is realized and the BL price is known. The roaster determines the quantity of the harvested beans to be sold under the SL (at the market clearing price) and the quantity, if any, to be sold as blend.[4] The timeline of decisions and events for a focal SL are illustrated in Figure 2 (depending on the SL locale, there is either one or two harvests per year).

[3] Differential pricing in which the price paid to the grower depends on the market price at the time of harvest (e.g., tied to C price) has been criticized by roasters like Intelligentsia arguing that it is harmful to growers because it exposes the grower to price risk.

[4] Equivalently, the roaster determines the selling price for the SL, which determines sales volume, and any leftover is roasted and sold at the BL price.

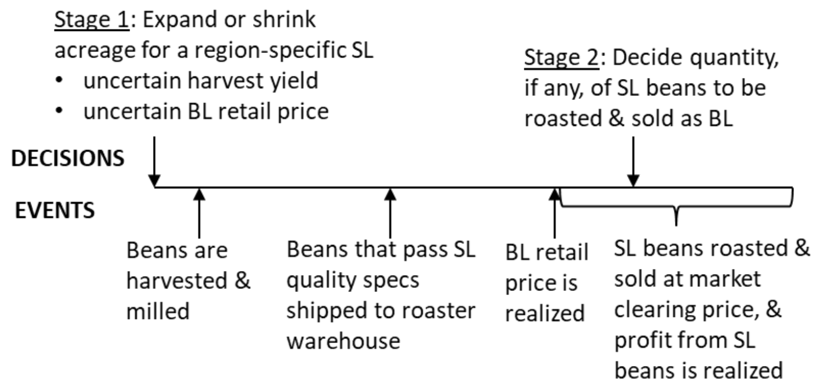


Figure 2. Timeline of decisions and events for a particular SL.

3.2. Model of Second-Stage Problem

We present the second-stage problem in this section. The optimal solution to this problem is rolled back into the first-stage problem that is formulated in the next subsection.

To simplify notation, and without loss of generality, we normalize the output unit (unit of beans) to the size of the connoisseur segment. For example, suppose the size of the connoisseur segment is one million pounds of roasted beans, e.g., the MNL choice model specifies the percentage splits of this volume across the three consumer choices: SL from the roaster, BL from the roaster, no-purchase from the roaster. In this example, the normalized output unit is one-million pounds. We similarly normalize the input unit (unit of farm space); one unit of farm space is equal to the number of acres that produce an average yield equal to the size of the connoisseur segment. We refer to the output unit as *normalized pound* and the input unit as *normalized acre*.

Let φ denote the realized BL price at the beginning of Stage 2 and let p denote the SL price. SL demand is denoted q and is given by the MNL discrete-choice model. The random utilities within the connoisseur segment of SL, BL, and the outside option are $\tilde{U} = v - bp + \tilde{\varepsilon}$, $\tilde{U}_B = v_B - b\varphi + \tilde{\varepsilon}_B$, $\tilde{U}_0 = \tilde{\varepsilon}_0$,

$$q = \Pr\left(\tilde{U} > \max\left\{\tilde{U}_B, \tilde{U}_0\right\}\right) = \frac{e^{v-bp}}{1 + e^{v_B - b\varphi} + e^{v-bp}} = (1-q)\frac{e^{v-bp}}{1 + e^{v_B - b\varphi}} \tag{1}$$

where $(\tilde{\varepsilon}, \tilde{\varepsilon}_0, \tilde{\varepsilon}_B)$ are i.i.d. Gumbel random variables, b is the connoisseur price-sensitivity parameter, and $v - bp$ and $v_B - b\varphi$ are nominal utilities of SL and BL, respectively (nominal utility of the outside option is normalized to $v_0 = 0$ without loss of generality).

Note that v and v_B are measures of the quality of SL and BL product as perceived by the connoisseur segment, $v > v_B$. Inverting the above yields the SL clearing price

$$p(q) = \frac{1}{b}\left(v - \ln\left(\frac{q\left(1 + e^{v_B - b\varphi}\right)}{1 - q} \right) \right). \tag{2}$$

Let Q denote normalized acres under contract for SL and let $k(Q)$ denote the cost prior to harvest (e.g., payment to the grower to help with expenses of tending to the crop). Function $k(Q)$ is continuous, nondecreasing, and convex. Let c denote the post-harvest variable cost per normalized pound (that satisfies quality specifications), a portion of which is paid to the grower and a portion that is incurred by the roaster (e.g., cost of transportation, storage, roasting, packaging). The post-harvest variable cost for the SL green beans is the same regardless of whether the beans are ultimately sold as SL or BL. Finally, let y denote the realized yield per normalized acre. Then the roaster's SL profit at the end of Stage 2 can be expressed as

$$\pi(q \mid Q, \varphi, y) = \left(p(q) - c\right)q + (\varphi - c)(Qy - q) - k(Q), \tag{3}$$

i.e., q normalized pounds are sold at the SL price p and $Qy - q$ normalized pounds are sold at the BL price φ. Hereafter we suppress the arguments Q, φ, y in the profit function $\pi(\cdot)$ to simplify the presentation. The notation for the model is listed in the appendix.

The second-stage optimization problem can be expressed as

$$\pi^*(Q, \varphi, y) = \max_q \left\{ \pi(q) : 0 \le q \le 1, q \le Qy \right\}$$

and the optimal solution is denoted $q^*(Q, \varphi, y)$ (we suppress the arguments of q^* unless necessary for clarity). The first constraint ensures that q is a valid fraction (i.e., percent of connoisseur segment) and the second constraint ensures that SL demand is not higher than available supply. The proposition below characterizes the optimal second-stage solution. Note that $W(\cdot)$ is the Lambert W function, which is the inverse of $x = We^W$.

Proposition 1. *For contract acreage Q, realized yield y, and realized BL price φ, the optimal quantity sold as SL is*

$$q^* = \min\left\{ W\left(\frac{e^{v-1}}{e^{b\varphi} + e^{v_B}} \right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\varphi} + e^{v_B}} \right) \right)^{-1}, Qy \right\}$$

with quantity $Qy - q^*$ sold as BL. The optimal SL selling price is

$$p^* = \begin{cases} \varphi + \dfrac{1}{b}\left(1 + \ln\left[\left(\dfrac{e^{v-1}}{e^{b\varphi} + e^{v_B}} \right)\left(\dfrac{1 - Qy}{Qy} \right) \right] \right), & \text{if } q^* = Qy \\ \varphi + \dfrac{1}{b}\left(1 + W\left(\dfrac{e^{v-1}}{e^{b\varphi} + e^{v_B}} \right) \right), & \text{if } q^* < Qy \end{cases}$$

and the optimal profit is

$$\pi^*(Q,\varphi,y)=\begin{cases}\frac{1}{b}\left(1+\ln\left[\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right]\right)Qy+(\varphi-c)Qy-k(Q), & \text{if } q^*=Qy\\ \frac{1}{b}W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)+(\varphi-c)Qy-k(Q), & \text{if } q^*<Qy\end{cases}.$$

From Proposition 1, we see that the optimal price is nonincreasing in yield and contract acreage, whereas the optimal SL quantity and optimal second-stage profit are nondecreasing in these values.

3.3. Model of First-Stage Problem

In stage one, the roaster sets the value of Q to maximize expected profit, accounting for optimal decisions in stage two. The first-stage problem is

$$\Pi^*=\max_{Q\geq 0}\left\{\Pi(Q)=E\left[\pi^*(Q,\tilde{\varphi},\tilde{y})\right]\right\} \tag{4}$$

and the optimal solution is denoted Q^*. We analyze sourcing strategies associated with the optimal solution in the next section.

4. Optimal Sourcing Strategies

We first clarify several points regarding the structure of our cost model. Note that (3) can be rewritten as

$$\pi(q)=(p(q)-\varphi)q+(\varphi-c)Qy-k(Q).$$

The cost term in the above expression, $k(Q)+cQy$, does not affect the second-stage optimal solution; it only affects the first-stage objective function only through its expectation, i.e., $E\left[k(Q)+cQ\tilde{y}\right]=k(Q)+cQ$. The upfront portion of expected cost, $k(Q)$, is a convex increasing function, and the after-harvest portion, cQy, is linear. This aligns with practice. In particular, $k(Q)$ tends to be linear over intervals of Q, i.e., piecewise linear. Increases in SL marginal cost tend to occur at points where expansion requires contracting with a new grower in the region.

4.1. Contract-Acreage Benchmarks

To facilitate characterization of optimal sourcing strategies, we define three contract-acreage benchmarks:

$$Q_D=W\left(\frac{e^{v-1}}{e^{b\mu_\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\mu_\varphi}+e^{v_B}}\right)\right)^{-1}$$

$$Q_L=\frac{1}{y_h}W\left(\frac{e^{v-1}}{e^{b\varphi_h}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi_h}+e^{v_B}}\right)\right)^{-1} \tag{5}$$

$$Q_H=\frac{1}{y_l}W\left(\frac{e^{v-1}}{e^{b\varphi_l}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi_l}+e^{v_B}}\right)\right)^{-1}. \tag{6}$$

The value of Q_D is the optimal contract acreage when variance of both yield and blend price is zero, i.e., its value is equivalent to the optimal acreage in the deterministic model. The values of Q_L and Q_H delineate the boundaries of optimal sourcing strategies. We use the notation $\tilde{\cdot}$, $\mu_{\cdot}$, $\sigma_{\cdot}^2$ to denote random variable, mean, and variance, respectively.

Lemma 1. $\Pr\left(q^*\left(Q,\tilde{\varphi},\tilde{y}\right) < Q\tilde{y}\right) = 0$ *if and only if* $Q \leq Q_L$, *and* $\Pr\left(q^*\left(Q,\tilde{\varphi},\tilde{y}\right) < Q\tilde{y}\right) = 1$ *if and only if* $Q \geq Q_H$.

Lemma 2. *Suppose yield and/or blend price is uncertain at the time of contracting* (*i.e.,* $\sigma_y^2 > 0$ *and/or* $\sigma_\varphi^2 > 0$). *Then* $Q_L < Q_D < Q_H$.

Lemma 1 identifies limits on contract acreage consistent with specialized, mixed, and diversified sourcing strategies. If $Q^* \leq Q_L$, then there is 100% probability that all harvested beans will be sold as SL, i.e., a specialized sourcing strategy is optimal. This sourcing strategy of limiting supply in the first stage so that there is never sufficient harvest to maximize profit in the second stage is unique to settings with uncertainty; such a strategy would never be optimal in a deterministic setting. At the opposite extreme, $Q^* \geq Q_H$ implies that there is 100% probability that some of the harvested beans will be sold as BL, i.e., a diversified sourcing strategy is optimal. If $Q^* \in (Q_L, Q_H)$, then a mixed sourcing strategy is optimal.

Lemma 2 shows that contracted acreage under the specialized (diversified) sourcing strategy is more conservative (aggressive) than the optimal acreage for a deterministic model. At the extreme of no uncertainty, the benchmark expressions are identical, i.e.,

$$\text{if } \sigma_y^2 = \sigma_\varphi^2 = 0, \text{ then } Q^* = Q_L = Q_D = Q_H. \tag{7}$$

We show through (9) that, when there is no uncertainty, the optimal acreage amount becomes equal to Q_D. In the absence of uncertainty, the firm contracts exactly the amount of acreage that would produce a quantity equivalent to the demand for SL. In effect, there would be no downward substitution in the absence of uncertainty in BL price and/or SL yield, resembling the behavior observed under the specialized sourcing strategy. This result also indicates that, in the absence of uncertainty, the mixed and diversified strategies do not exist. Thus, mixed and diversified sourcing strategies arise as a result of uncertainty.

4.2. Conditions for Sourcing Strategy Dominance: Specialized, Mixed, Diversified

A natural next question is what types of environments favor each sourcing strategy under uncertainty (in yield and blend price). We begin to shed light on this question by characterizing the optimal first-stage decision in Proposition 2. We then build on this result by presenting and interpreting three propositions on conditions for optimality of different sourcing strategies. But first, we require additional notation and clarification of assumptions.

The optimal unconstrained second-stage SL quantity decision is

$$q^{o}(\varphi) = W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1} \tag{8}$$

(see Proposition 1). We use the convention of $k'(Q)$ as the right-derivative at Q if left- and right-derivatives differ. For example, in practice there is a finite upper limit $Q_{\max}$ on contract acreage in a region for a single-origin label, which implies the right derivative at $Q_{\max}$ is $k'(Q_{\max}) = \infty$ (and the left derivative is finite). We denote the price-yield correlation coefficient of $(\tilde{\varphi}, \tilde{y})$ as ρ.

The first-stage acreage decision Q influences the marginal returns in stage 2. If the realized yield Qy from this acreage decision is less than the unconstrained optimal second-stage SL quantity in (10), then the firm charges a higher price to clear this limited supply. Let $m_c(Q, \varphi, y)$ denote the constrained optimal markup over the realized BL price φ per unit of input (normalized acre) instead of output unit (normalized pound):

$$m_c(Q,\varphi,y) = \frac{y}{b}\ln\left[\left(\frac{e^{v}}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right].$$

When the realized yield is greater than or equal to the value designated in (10), i.e., $Qy \geq q^{o}(\varphi)$, then the firm has sufficient supply and the realized yield of SL beans is not binding. Let $m_u(Q, y)$ denote the unconstrained optimal markup over the realized BL price φ per unit of input:

$$m_u(Q,y) = \frac{y}{b}\left(\frac{1}{1-Qy}\right).$$

These two optimal markup expressions enable us to write the expected second-stage return from the first-stage acreage decision Q. We express the expected value of the second-stage shadow price (Lagrangian multiplier) as follows:

$$\lambda(Q) = E\left[\left(m_c(Q,\tilde{\varphi},\tilde{y}) - m_u(Q,\tilde{y})\right)^{+}\right]. \tag{9}$$

Note that, if (φ, y) is such that $Qy \geq q^{o}(\varphi)$ (i.e., supply constraint in the second stage is not binding), then $m_c(Q,\varphi,y) - m_u(Q,y) < 0$, and the above shadow price becomes $\left(m_c(Q,\varphi,y) - m_u(Q,y)\right)^{+} = 0$, explaining the presence of the truncation operation in (11). The expected shadow price $\lambda(Q)$ becomes positive only in the constrained scenario when $Qy < q^{o}(\varphi)$.

The expected shadow price in (11) helps determine the optimal acreage decision in stage 1 which is obtained from the comparison of the marginal revenue and marginal cost. This is formalized in the next proposition.

Proposition 2. *The optimal contract acreage is*

$$Q^{*} = \max\left\{Q : \mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} + \lambda(Q) \geq k'(Q) + c\right\}. \tag{10}$$

The optimality condition in (12) compares two terms. The left-hand side of the inequality is the marginal revenue with respect to the input unit and the right-hand side is the marginal cost. Both terms require the expectation operator over the blend price and yield, e.g., marginal cost is $E\left[k'(Q)+\tilde{y}c\right]=k'(Q)+c$. The expression for marginal revenue can be interpreted as the sum of two basic elements, $E\left[\tilde{\varphi}\tilde{y}\right]+\lambda(Q)=$ $\left(\mu_{\varphi}+\rho\sigma_{\varphi}\sigma_{y}\right)+\lambda(Q)$. The first term in parentheses, $\mu_{\varphi}+\rho\sigma_{\varphi}\sigma_{y}$, is the expected revenue per input unit if sold as BL, e.g., compared to the mean BL price, a negative (positive) correlation between SL yield and BL price leads to a decrease (increase) in the expected BL revenue per input unit. The second term in the marginal revenue, $\lambda(Q)$, is the expected shadow price $\lambda(Q)$ from the first-stage acreage decision Q as expressed in (11) . Both the marginal revenue and the marginal cost terms are transparent (i.e., closed-form expressions of model primitives) in the above proposition.

Let us examine the elements and intuition that underlie the value of the less-transparent term $\lambda(Q)$ to gain additional insights (the supporting details are provided in the proof of Proposition 2). The value of $m_c(Q,\varphi,y)/y$ is equal to the supply-constrained SL markup over the BL price φ, i.e., the SL market-clearing markup per output unit (e.g., $m_c(Q,\varphi,y)$ expresses the same markup per input unit, as noted above). At values of φ and y that satisfy $Qy = q^{o}(\varphi)$ (i.e., SL supply is equal to the optimal unconstrained SL supply), $m_u(Q,y)/y$ is equal to the unconstrained optimal SL markup over the BL price φ, e.g., $m_u\left(q^{o}(\varphi)/y,y\right)$ is the unconstrained optimal SL markup per input unit. It is important to highlight that the value of $m_c(Q,\varphi,y)$ is decreasing in φ and y, and it captures the gain in profit from increasing Q at markup $m_c(Q,\varphi,y)$. In contrast, the value of $m_u(Q,y)$ is increasing in y, and it captures the loss in profit from the consequent reduction in markup to clear the market as Q increases.

The difference between $m_c(Q,\varphi,y)$ and $m_u(Q,y)$ is equivalent to the contribution of the downward substitution flexibility to the marginal revenue at realization (φ, y). To formalize this notion, let us consider the scenario where the firm does not have the flexibility to sell SL beans as BL, i.e., all SL beans are sold as SL at the market-clearing price. Expressing

$$\bar{m}_c(Q)=E\left[m_c(Q,\tilde{\varphi},\tilde{y})\right]$$

$$\bar{m}_u(Q)=E\left[m_u(Q,\tilde{y})\right],$$

it is straightforward to verify that marginal revenue in the absence of the downward substitution flexibility is

$$\mu_{\varphi}+\rho\sigma_{\varphi}\sigma_{y}+E\left[\left(m_c(Q,\tilde{\varphi},\tilde{y})-m_u(Q,\tilde{y})\right)\right]=\mu_{\varphi}+\rho\sigma_{\varphi}\sigma_{y}+\bar{m}_c(Q)-\bar{m}_u(Q).$$

Then, the flexibility to downward substitute increases the marginal revenue by

$$V_F(Q) = E\left[\left(m_c\left(Q,\tilde{\varphi},\tilde{y}\right) - m_u\left(Q,\tilde{y}\right)\right)^+\right] - \left(\bar{m}_c\left(Q\right) - \bar{m}_u\left(Q\right)\right) = \lambda\left(Q\right) - \left(\bar{m}_c\left(Q\right) - \bar{m}_u\left(Q\right)\right).$$

The value of marginal revenue in the absence of downward substitution helps determine when each sourcing strategy – specialized, mixed, diversified – is optimal. Using Q_L and Q_H developed earlier, the optimal sourcing strategy decision is determined in the next proposition.

Proposition 3. (*a*) *The specialized sourcing strategy is optimal if and only if*

$$\mu_\varphi + \rho\sigma_\varphi\sigma_y + \left(\bar{m}_c\left(Q_L\right) - \bar{m}_u\left(Q_L\right)\right) \le k'(Q_L) + c \tag{11}$$

(*b*) *The diversified sourcing strategy is optimal if and only if*

$$\mu_\varphi + \rho\sigma_\varphi\sigma_y \ge k'(Q_H) + c. \tag{12}$$

(*c*) *The mixed strategy is optimal if and only if* (13) *and* (14) *do not hold.*

Proposition 3 identifies conditions for optimality of different sourcing strategies. Recall that Q_H is such that some SL beans are sold as BL for any realization of $\left(\tilde{\varphi},\tilde{y}\right)$. Thus, any increase in normalized acres above Q_H will be sold as BL yielding expected revenue per input unit of $\mu_\varphi + \rho\sigma_\varphi\sigma_y$. That is, $\lambda(Q) = 0$ if and only if $Q \ge Q_H$, which explains why $\lambda(Q)$ does not appear in (14). On the other hand, if $Q \le Q_L$, then all SL beans will be sold as SL. As alluded to above, the shadow price, $m_c\left(Q,\varphi,y\right) - m_u\left(Q,y\right)$, is positive for every realization of $\left(\tilde{\varphi},\tilde{y}\right)$ if and only if $Q < Q_L$. Equivalently, $\lambda(Q) = \bar{m}_c\left(Q\right) - \bar{m}_u\left(Q\right)$ if and only if $Q \le Q_L$, which explains the simpler form of $\lambda(Q)$ (i.e., without truncation) that appears in (13) for the optimality condition of specialized sourcing.

In the discussion following expression (9), we have identified that randomness in BL price and/or SL yield must be present for the mixed and diversified sourcing strategies to emerge. The necessary and sufficient conditions for the optimality of the specialized and mixed sourcing strategies (identified in Proposition 3) cannot be expressed in closed form. At a high level, Proposition 3 indicates that specialized (diversified) sourcing is preferred when marginal revenue is low (high) relative to marginal cost. The next result develops a sufficient condition for the two pure sourcing strategies—specialized and diversified.

Corollary 1. (*a*) *A sufficient condition for the specialized sourcing strategy to be optimal is*

$$\varphi_h + \frac{1}{b}\left(\ln\left(\left(\frac{y_h}{y_l}\right)\left(\frac{1-Q_L y_l}{1-Q_L y_h}\right)\right) + \frac{1}{1-Q_L y_h} - \frac{1}{1-Q_L y_l}\right) \le k'(Q_L) + c \tag{13}$$

which, if there is no yield uncertainty ($\sigma_y^2 = 0$), *reduces to*

$$\varphi_h \le k'(Q_L) + c. \tag{14}$$

(*b*) *A sufficient condition for the diversified sourcing strategy to be optimal is*

$$\varphi_l \ge k'(Q_H) + c. \tag{15}$$

Corollary 1 sheds light on *relatively low or high* marginal revenue in terms of a few model primitives. Corollary 1(*a*) shows that specialized sourcing dominates when the upper limit of BL price is high compared to SL cost and range of yield ($y_h - y_l$) is low; Corollary 1(*b*) shows that diversified sourcing dominates when the lower limit of BL price is high compared to SL cost.

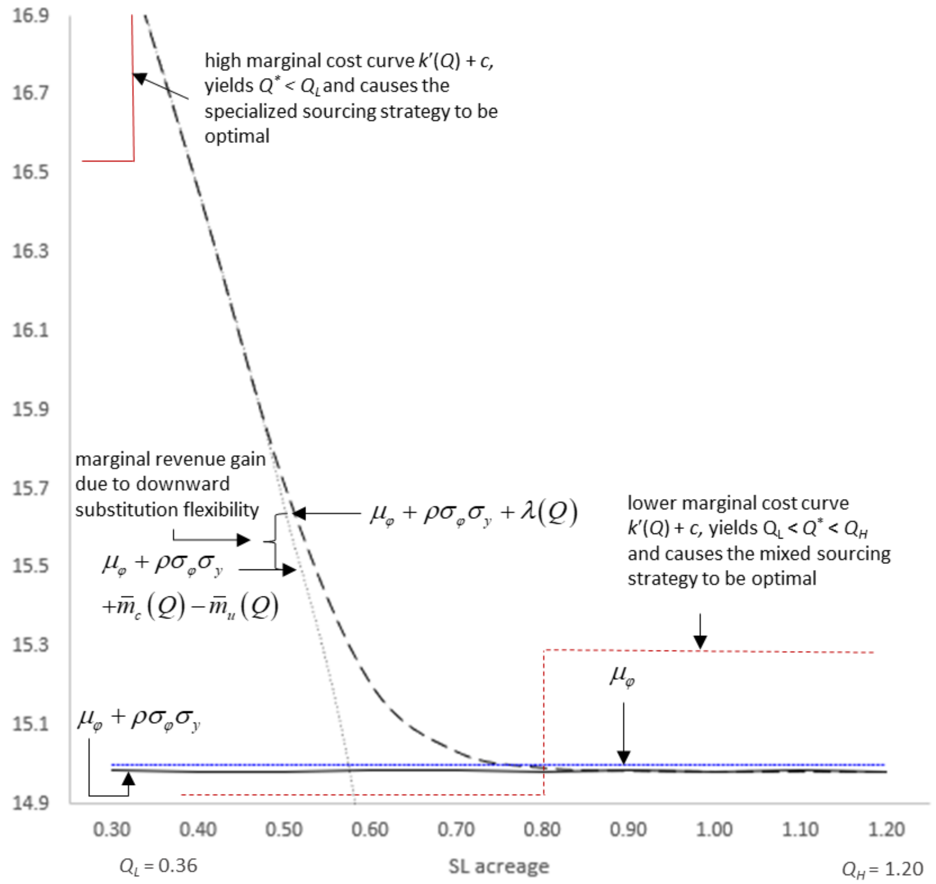


Figure 3. Illustrations of functions and relationships identified in propositions 2 – 3. The marginal revenue at $Q_L = 0.36$ is \$16.76 and the upper bound given in Corollary 1 (see (15)) is \$19.61.

Figure 3 provides several insights in addition to illustrating functions and values that appear in propositions 2 and 3. Dollar values in this figure are reported in dollars per acreage that yields one pound of roasted beans on average (or dollars per pound). Input quantity values (Q) are reported as fraction of acreage yielding 6 million pounds of SL beans on average. We generate the marginal revenue elements in the figure using the following calibration: marginal distributions of BL price and SL yield are truncated normal with the bivariate distribution obtained by the Gaussian copula at correlation ρ = -0.1, the mean BL price per pound is μ_φ = \$15 (blue dotted line in Figure 3) with range [\$13, \$17], the yield range is [0.5, 1.5] (see the appendix for additional detail on the calibration). For this calibration, the mean optimal SL price per pound ranges from \$18.67 to \$17.44 as Q ranges from 0.3 to 1.2 normalized acres. We next elaborate on the major insights from Figure 3.

Marginal cost plays a significant role in influencing which sourcing strategy is optimal and Q^* (weakly) decreases as the marginal cost increases. This was already established in Proposition 3 and Corollary 1, and it is demonstrated in Figure 3 which features two different marginal cost curves, $k'(Q_L) + c$, both designated with red lines. When the marginal cost curve is as represented in upper left red line (\$16.50 for $Q < 0.34$), increasing acreage beyond $Q = 0.34$ increases the marginal cost above the marginal revenue. In this case, $Q^* = 0.34 < Q_L = 0.36$, making the specialized sourcing optimal. In the lower red line, the marginal cost is \$14.90 up to $Q = 0.8$, then jumps to \$15.30 and above the marginal revenue, making $Q^* = 0.8 \in [Q_L, Q_H]$; this means the mixed sourcing strategy is optimal for this marginal cost curve. While not in the figure, we may envision an alternative marginal cost curve that jumps from \$14.90 to \$15.30 at $Q \geq Q_H = 1.2$, which is an example where the diversified sourcing strategy is optimal. In conclusion, as the marginal cost increases, the optimal sourcing strategy moves from the diversified to the mixed and then to the specialized sourcing strategy.

In coffee, the weighted average BL price decreases with the negative price-yield correlation. In Figure 3, the horizontal blue dotted line designates the mean BL price (\$15) in a season. The horizontal black line shows the mean BL price (\$14.98) per unit harvested, i.e., the average BL price is weighted by seasonal yields. As is common in agriculture, seasons of high yield tend to be associated with lower prices than seasons with low yield (i.e., negative price-yield correlation). Consequently, the volume-weighted price per season is lower than the unweighted price per season, and the difference is the price-yield covariance $\rho\sigma_\varphi\sigma_y$. This price difference can be interpreted as the *farmer's curse*.

The farmer's curse stemming from the negative price-yield covariance in BL price carries lessons for growers and roasters. First, reducing the negative correlation between the grower's yield and BL price increases profit. This can be accomplished, for example, with investments or actions that reduce the sensitivity of grower's yield to weather and other environmental conditions relative to other growers (i.e., by reducing negative price-yield covariance for the grower). Second, some roasters use differential pricing for green coffee beans wherein the price paid to the grower is determined by the market price of green beans *at the time of harvest*. As noted earlier, Intelligentsia sets the price to be paid to the grower prior to the growing season for the stated reason that this reduces the grower's exposure to risk (due to price uncertainty). However, the use of non-differential pricing offers an additional and perhaps less obvious advantage, all else equal – it reduces the negative price-yield covariance for the grower which translates to higher grower profit per acre. This is an important insight because awareness of this additional effect may encourage more specialty roasters to move away from differential pricing.

The shadow price decreases in *Q* under the specialized and mixed strategies and it becomes zero under the diversified strategy. Under the specialized and mixed strategies each additional unit of acreage leads to higher SL yield (on average) and a lower price to clear the SL bean supply. This leads to a reduction in the value of $\lambda(Q)$ and consequently the marginal revenue. Figure 3 indicates the marginal revenue with the black dashed curve and demonstrates how each unit of additional acreage reduces the marginal revenue (in a convex manner) under the specialized and mixed strategies. The marginal revenue eventually becomes equal to the weighted mean BL price (black horizontal line) when the firm switches to the diversified sourcing strategy, making the shadow price equal to zero; in this case, each additional unit of acreage does not bring any additional revenue from SL beans (and only the revenue from BL beans).

The value from the downward substitution flexibility, V_F, adds no value under the specialized sourcing strategy and is increasing in Q under the mixed and diversified sourcing strategies. The dashed black curve represents marginal revenue curve in the absence of the downward substitution flexibility where the firm sets a market-clearing price for its excess SL beans (rather than using them in BL). For any given acreage Q, the difference between the black dashed curve and the dotted black curve in Figure 3 is the value gained from downward substitution. Under the specialized strategy, there is no downward substitution, i.e., $V_F = 0$. At $Q = 0.52$ that corresponds to a mixed strategy, for example, V_F is identified to be equal to \$0.20. Figure 3 demonstrates that V_F (the difference between the black dashed and black dotted curves) is increasing with higher levels of acreage.

We draw on the above results and intuition to summarize conditions under which each strategy is likely to dominate (see Figure 4). We summarize the key insights in the following discussion.

Two primary factors that determine the sourcing strategy choice are: (1) mean BL price, (2) marginal cost. Let us begin with the more subtle factor: the mean BL price. An increase in the BL price effectively makes the firm gravitate towards the diversified sourcing strategy by increasing Q^* and decreasing Q_H. Suppose that each possible realization of BL price is increased by amount x (e.g., the mean BL price changes from μ_φ to $\mu_\varphi + x$). Since $Q^* \geq Q_H$ before the change, marginal revenue, $\mu_\varphi + \rho\sigma_\varphi\sigma_y$, is increasing in x, and thus Q^* is increasing in x. We note that Q_H is the unconstrained optimal second-stage supply at realization (φ_h, y_l), and thus Q_H is decreasing in x. From this intuition, we may suspect that decreasing the BL price generally moves the preferred strategy in the opposite direction. However, the effects of decreasing BL price when the current strategy is either specialized or mixed is more nuanced because of the effect of decreasing BL price on $\lambda(Q)$. Nevertheless, as we prove in the appendix (see Proposition A2), this directional result continues to hold. This enables us to identify distinct regions in the blend price and marginal cost dimensions for which each sourcing strategy is preferred, which are depicted in Figure 4 (created using the same calibration as Figure 5 below).

The second factor that influences the choice of the sourcing strategy is the marginal cost. An increase in the marginal cost $k'(Q) + c$ (weakly) decreases Q^*, moving the firm's preference towards the mixed and even further to the specialized sourcing strategy. This is also depicted in Figure 4. All other factors such as the perceived quality of SL means (measured by v) and BL beans (measured by v_B) do not have a clear directional effect on the choice of the strategy.[5] For example, while Q_L, Q^*, and Q_H are all increasing in v, the relative effects on marginal revenue and marginal cost can vary depending on parameter values. This ambiguous effect also holds for an increase in the price sensitivity parameter b (see Proposition A3 in the appendix for details). Consequently, these parameters do not neatly fit within a strategy grid. We examine the effects of changes in the yield volatility of SL yield on sourcing strategies in the next section.

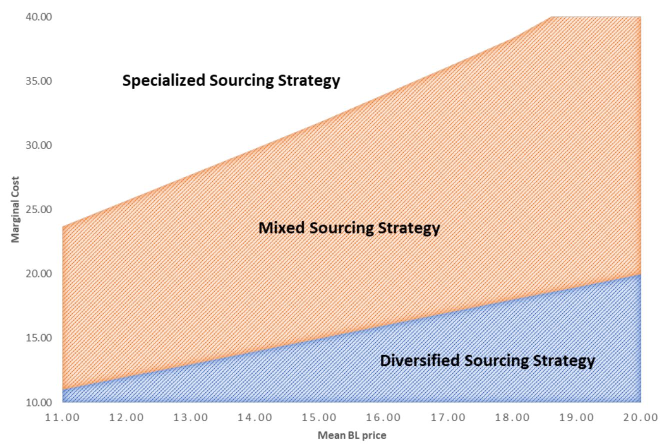


Figure 4. Strategy grid: regimes where a particular sourcing strategy is optimal.

4.3. Impact of Changing Yield Variance

This section examines the impact of changing yield variance on the firm's optimal sourcing strategy, acreage decisions and profit. Specifically, we analyze how improvements in yield variance influence a roaster's incentive to further invest in a grower. It is important to remember that the volatility in coffee bean yield (i.e., quantity of harvested beans that satisfy the roaster's quality specifications) is not only affected by weather, but also through factors that can be influenced by the grower and/or roaster, e.g., farming practices to mitigate pests/disease and promote robust growth, post-harvest processing, and storage.

Before presenting results on the effects of increasing yield variance, we clarify a few technical details that underlie our analysis. The clarifications pertain to (1) how the probability distribution changes as yield variance changes and (2) cost function $k(Q)$.

[5] More specifically, the following strategy shifts are possible as v increases relative to v_B (shift direction depends on parameter values): specialized to mixed, mixed to specialized, and diversified to mixed.

Recall that $F(\varphi, y)$ is the joint distribution of $(\tilde{\varphi}, \tilde{y})$. Let $F_y(y \mid \varphi)$ and $F_y(y)$ denote the conditional and marginal distribution functions of $\tilde{y}$. To isolate the effect of increasing σ_y, we keep all else equal, i.e., all other parameters associated with $F(\varphi, y)$ remain fixed. For example, if $(\tilde{\varphi}, \tilde{y})$ is bivariate normal, then F can be described by five parameters reflecting means, variances, and Pearson correlation coefficient, e.g., μ_φ, μ_y, σ_φ, σ_y, ρ. More generally, as σ_y increases from $\sigma_y = \sigma_1$ to $\sigma_y = \sigma_2 > \sigma_1$ with corresponding random variables $(\tilde{\varphi}, \tilde{y}_1)$ and $(\tilde{\varphi}, \tilde{y}_2)$, we assume $E[\tilde{y}_2] = E[\tilde{y}_1] = 1$ and

$$\sigma_2(\varphi) = V[\tilde{y}_2 \mid \tilde{\varphi} = \varphi]^{1/2} \geq V[\tilde{y}_1 \mid \tilde{\varphi} = \varphi]^{1/2} = \sigma_1(\varphi) \text{ for all } \varphi \in [\varphi_l, \varphi_h] \tag{16}$$

(the inequality is strict for some φ due to $\sigma_2 > \sigma_1$). For some of our results we will assume that random blend price and yield are independent. Under independence, we have

$$E[\tilde{y}_2 \mid \tilde{\varphi} = \varphi] = E[\tilde{y}_1 \mid \tilde{\varphi} = \varphi] = 1 \text{ for all } \varphi \in [\varphi_l, \varphi_h],$$

which implies that $\tilde{y}_2$ is a mean-preserving spread of $\tilde{y}_1$.

The cost of acreage $k(Q)$ is often a piecewise linear convex increasing function in practice. This is because the marginal cost of acreage is constant over initial values of Q, however, it jumps up to a higher value when expansion requires the roaster to contract with a new grower in the region (as depicted in Figure 3). In other words, the real-world k function may not be differentiable everywhere (i.e., not differentiable at breakpoints in the piecewise function). However, any piecewise linear function can be replaced with an arbitrarily close approximation that is differentiable everywhere (e.g., by inserting a smooth convex function about each kink). For purposes of ensuing analysis, we assume that $k(Q)$ is twice differentiable everywhere.

4.3.1. Yield variance comparative statics

We next present the impact of increasing yield variance on the roaster's profit and acreage decisions under each sourcing strategy. We divide our examination into two parts: first, we present our findings under independent random variables representing the yield and the blend price, and then when the two random variables are correlated. Later, we compare these findings with the literature pertaining to the newsvendor problem where demand is the source of uncertainty (rather than yield).

Proposition 4. (*a*) *Suppose that blend price and yield are independent.*

(*i*) *Specialized sourcing is optimal*: Π^* *is decreasing in* σ_y

Q^* *is decreasing in* σ_y

(*ii*) *Mixed sourcing is optimal*: Π^* *decreasing in* σ_y

Q^* *is decreasing or increasing in* σ_y

(*iii*) *Diversified sourcing is optimal*: Π^* *remains fixed as* σ_y *changes*

Q^* *remains fixed as* σ_y *changes.*

(*b*) *Suppose that blend price and yield are dependent.*

(*i*) *Diversified sourcing is optimal and* $\rho < 0$: Π^* *is decreasing in* σ_y

Q^* *is decreasing in* σ_y

(*ii*) *Diversified sourcing is optimal and* $\rho > 0$: Π^* *is increasing in* σ_y

Q^* *is increasing in* σ_y

The impact of increasing yield variance on optimal acreage aligns with the newsvendor problem under the mixed sourcing strategy but departs from the newsvendor problem under the specialized and diversified sourcing strategies. To compare the effects of increases in variance on optimal decisions with the newsvendor problem we focus on the case where price and yield are uncorrelated. To see the impact of yield variance on the optimal acreage decision Q^*, we must examine the marginal profit function and compare it with that of the newsvendor problem. The marginal profit in the newsvendor problem is both concave and convex in random demand; we see this when the marginal profit is written as $pE\left[\mathbf{1}_{\tilde{d}>q}\right]-c$, which has a step-function form (e.g., roughly speaking, convex in the neighborhood of the vertical step up and concave in the neighborhood of the horizontal step right). As a result, increases in demand uncertainty may either increase or decrease the marginal profit, leading to the classic newsvendor comparative statics result wherein the optimal order quantity in the newsvendor problem can be either increasing or decreasing in demand variance depending on model parameters (Feng and Shanthikumar 2022). In particular, the newsvendor quantity decreases in variance when the critical ratio ($1 - c/p$) is low (e.g., concave structure of the marginal profit dominates) and increases in variance when the critical ratio is high (e.g., convex structure of the marginal profit dominates). Aligned with the newsvendor structure, the marginal profit under a mixed strategy exhibits a similar step-function form (see the truncation operator in equation (11)), which is what drives the ambiguous result (*aii*) for Q^*.

Interestingly, the mixed convex/concave structure disappears under specialized and diversified sourcing. Marginal profit is concave in random yield under the specialized sourcing strategy. Consequently, marginal profit decreases in variance, which causes Q^* to decrease as yield variance increases (result (*ai*)). On the other hand, marginal profit is linear in random yield under the diversified sourcing strategy, which leads to Q^* being unaffected by changes in variance (result (*aiii*)).

We note that the introduction of correlation between random price and yield creates ambiguity in the effect of increased yield variance on Q^* under the specialized and mixed sourcing strategies. The fact that marginal profit is linear in yield under diversified sourcing allows for the full characterization effects of increasing variance given in part (*b*) of Proposition 4.

The impact of increasing yield variance on optimal profit aligns with the newsvendor problem under the specialized and mixed sourcing strategies but departs from the newsvendor problem under the diversified sourcing strategy. Feng and Shantikumar (2022) show that the optimal profit in the newsvendor problem is decreasing with higher degrees of demand variance; this is because the profit function $p\min\left\{\tilde{d},q\right\}-cq$ is concave in $\tilde{d}$ due to the min operator. For our problem, the first-stage profit function is concave in yield under specialized and mixed sourcing (see Proposition A4 in the appendix). Consequently, like the newsvendor problem, optimal profit is decreasing in yield variance under specialized and mixed sourcing. On the other hand, expected profit is linear in random yield under the diversified strategy (see Proposition A4 in the appendix), which explains why changes in yield variance do not affect optimal profit.

From a robustness perspective, it is important to note that the results in Proposition 4 continue to hold under an alternative model with additive yield uncertainty, e.g., random yield is $Q+\tilde{y}$ instead of $Q\tilde{y}$. Proposition A4 in the appendix proves this result. It shows that the basic structure (e.g., concave) of the relevant functions is robust to a change from multiplicative yield to additive yield.

Figures 5 and 6 illustrate how optimal profit per unit changes with yield volatility under each sourcing strategy. Figure 5 reports results for the case where blend price and yield are independent. The profit under the specialized and mixed sourcing strategy is decreasing in yield variation, confirming the findings in parts (*ai*) and (*aii*) of Proposition 4. Moreover, the profit under the diversified sourcing strategy remains the same with higher variation in yield as shown in part (*aiii*) of Proposition 4. Figure 6 presents the impact of yield variation on profits of each sourcing strategy under a negative yield-blend price correlation. As noted in Section 1, the price of green coffee beans tends to be lower when global yields are high, and higher when global yields are low. This implies a negative price-yield correlation, as is common among agricultural products in general. There are three primary coffee growing regions located along the equatorial zone – Central/South America, Africa/Middle East, and Asia. Yields in subregions tend to be positively correlated because of weather. The overall effect is that the correlation between random yield for a particular single-origin label and the random blend price is likely to be either negative or insignificant (Kazaz 2004, Kazaz and Webster 2011). Figure 6 reports results for the case where blend price and yield are negatively correlated at ρ = -0.1 (see the appendix for additional detail on the calibration). Under a negative yield-blend price correlation, profits under all three sourcing strategies decrease in this example. As shown in part (*bi*) of Proposition 4, the profit under the diversified strategy is decreasing in variation in the yield when the correlation is negative; and this result differs from the behavior when yield and blend price random variables are independent.

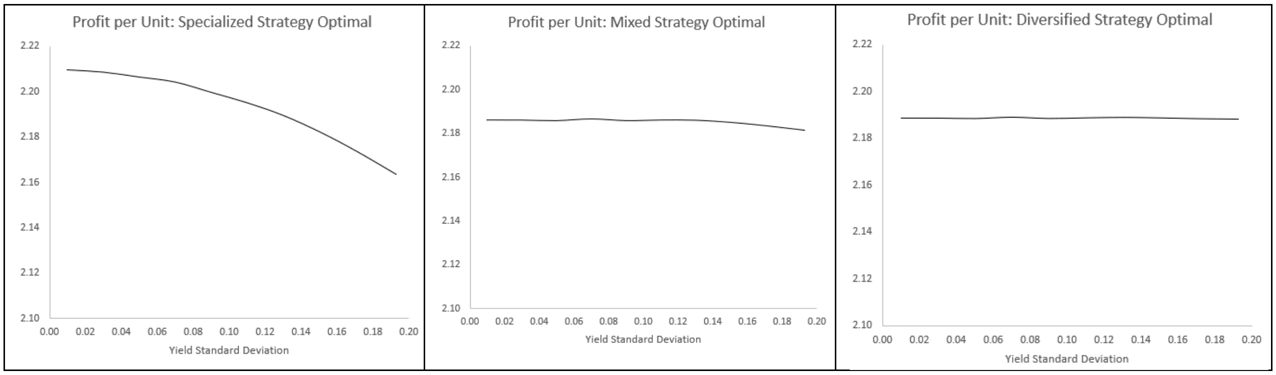


Figure 5. Optimal profit per unit Π^*/Q^* as σ_y changes for specialized sourcing (left panel), mixed sourcing (middle panel), diversified sourcing (right panel). Blend price and yield are independent ($\rho = 0$).

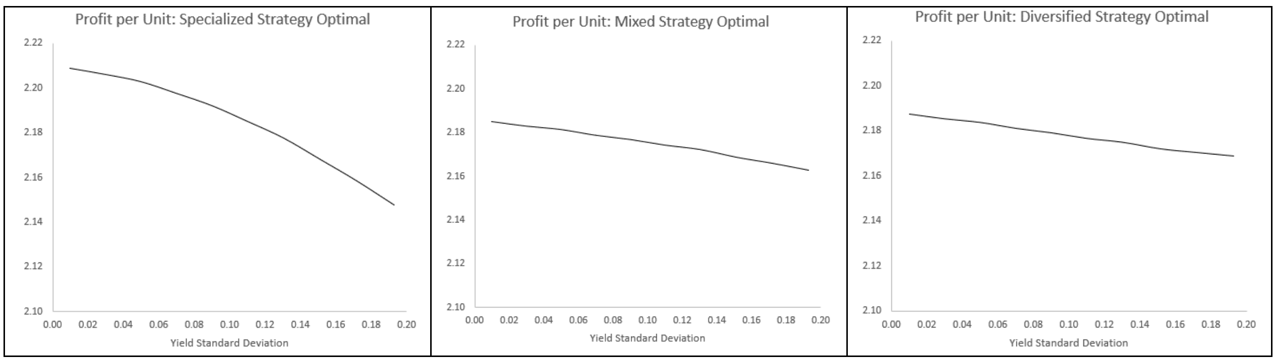


Figure 6. Optimal profit per unit Π^*/Q^* as σ_y changes for specialized sourcing (left panel), mixed sourcing (middle panel), diversified sourcing (right panel). Blend price and yield are negatively correlated ($\rho = -0.1$).

4.3.2. Yield variance comparative dynamics: positive feedback loop

In this section, we build on the above insights developed through Proposition 4 and examine the roaster's incentives to invest in growers. Recall that profit is concave in yield under specialized sourcing and linear in yield under diversified sourcing. Relative to the linear structure observed under the diversified sourcing strategy, the concave functional form induces higher profit from investments that lower volatility of yield (e.g., via improved cultivation, harvesting, and storage). Mixed sourcing combines the character of specialized and diversified sourcing. Profit is strictly concave at realizations where the supply constraint is binding and linear at realizations where the supply constraint is not binding.

We next discuss a positive feedback loop, and even a virtuous cycle, this insight brings to the lives of the roaster and grower. Figures 5 and 6 illustrate a core insight related to Proposition 4—gains from investments to improve the consistency of yield satisfying SL quality standards become amplified as the sourcing strategy moves from diversified to mixed to specialized. To formalize this observation, let $\bar{\Pi}_s^*$ denote the optimal profit per unit with optimal strategy $s \in \{SS, MS, DS\}$ for specialized, mixed and diversified sourcing strategies, respectively. We then have $\left|d\bar{\Pi}_{SS}^* / d\sigma_y\right| > \left|d\bar{\Pi}_{MS}^* / d\sigma_y\right| > \left|d\bar{\Pi}_{DS}^* / d\sigma_y\right|$. In this sense, there is potential for a virtuous cycle to arise in a grower-roaster DT relationship: a roaster dedicates more attention and resources on a single-origin-label grower over time. However, this also causes the marginal cost of SL beans to increase, and in turn, this behavior pushes the character of the sourcing strategy toward specialized sourcing (see Figure 4). This behavior is further exacerbated by greater incentives for the roaster to further invest in and work more closely with the grower, resulting in a positive feedback loop and the virtuous cycle depicted in Figure 7.

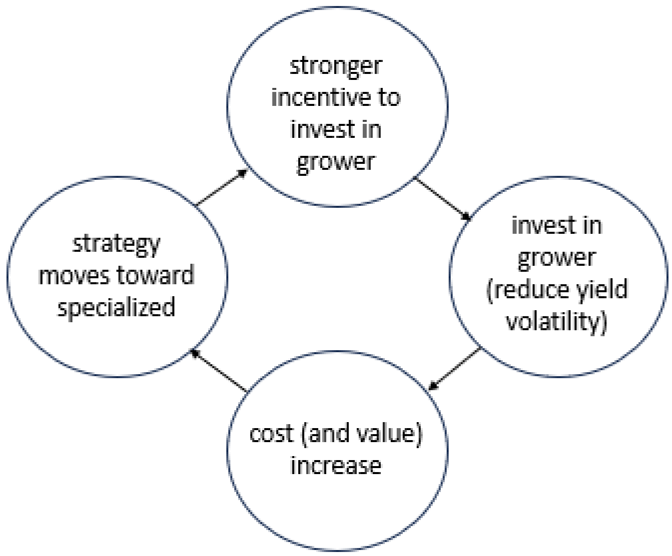


Figure 7. Virtuous cycle that may arise under DT with single-origin-label growers.

Let us next elaborate on implications of this virtuous cycle for geographic comparative advantage, consequent economic stratification, and opportunities for interventions to improve grower welfare. By definition, a particular single-origin label is specific to a locale. This can create natural barriers to expansion of single-origin-label contract acreage, e.g., large increases in marginal cost $k'(Q) + c$ as Q passes through thresholds associated with adding a new grower or pushing the limits of available acreage in the locale. Such barriers amplify the incentives of a roaster to invest in growers, leading to some locales where grower welfare is very high relative to other locales. In essence, the "relief valve" in the form of contract-acreage expansion is very costly for the roaster with the effect akin to sky-high real estate prices in a landlocked desirable area. Large regional differences in grower welfare signal an opportunity for policy-maker interventions that promote wider geographical investment by specialty roasters.

4.4. Robustness of Optimal Sourcing Strategies

In this section, we examine the robustness of the three sourcing strategies – specialized, mixed, diversified – under alternative demand models. We show that the existence of the three distinct optimal sourcing strategies is robust to any alternative model of demand that admits a unimodal second-stage profit function. For example, the result continues to hold for the nested logit discrete choice model that yields a concave profit function (Li and Huh 2011).

Proposition 5. *If second-stage profit function* $\hat{\pi}\left(q\,|\,\varphi\right)=\left(p\left(q\right)-\varphi\right)q$ *is unimodal for all* $\varphi \in [\varphi_l, \varphi_h]$, *then one of three distinct strategies is optimal: specialized, mixed, diversified.*

The mixed multinomial logit (MMNL) model exhibits the desirable theoretical property of arbitrarily closely approximating any random utility maximization discrete choice model (Feng et al. 2022, McFadden and Train 2000). One question is whether, or under what conditions, the three optimal strategies arise under MMNL demand. Adapting the MMNL model to our setting, the nominal SL utility of customer type k is v_k and the nominal BL utility is v_{Bk} where we assume $v_k - v_{Bk}$ is constant across segments, i.e., for n customer types, $v_k - v_{Bk} = \ln(1/\alpha)$ for $k = 1, \ldots, n$, equivalently,

$$\alpha = e^{v_{Bk}} / e^{v_k}.$$

The choice probability for a type k customer is

$$q_k = \frac{e^{v_k - bp}}{1 + e^{v_{Bk} - b\varphi} + e^{v_k - bp}} = \frac{e^{v_k - bp}}{1 + \alpha e^{v_k - b\varphi} + e^{v_k - bp}} = \left(1 - q_k\right)\frac{e^{v_k - bp}}{1 + \alpha e^{v_k - b\varphi}}$$

and the clearing price is

$$p(q_k) = \frac{1}{b}\ln\left(\left(\frac{1 - q_k}{q_k}\right)\frac{e^{v_k}}{1 + \alpha e^{v_k - b\varphi}}\right)$$

(see (1) and (2)). The probability that a customer belongs to segment k is denoted w_k. The unconstrained (with respect to capacity) second-stage optimization problem is

$$\max_{\mathbf{q}\in[0,1]^n}\left\{\hat{\pi}\left(\mathbf{q}\,|\,\varphi\right) = \sum_{k=1}^{n}\left(p\left(q_k\right) - \varphi\right)w_k q_k : p\left(q_1\right) = \ldots = p\left(q_n\right)\right\}$$

(the constraint assures price p is fixed across segments).

Following Li et al. (02019), the above n-dimensional optimization problem can be transformed to an equivalent univariate optimization problem. Let $A_k\left(\varphi\right) = e^{v_k}\left(1 + \alpha e^{v_k - b\varphi}\right)^{-1}$. Then

$$q_k = \left(1 - q_k\right)A_k\left(\varphi\right)e^{-bp} = \frac{A_k\left(\varphi\right)e^{-bp}}{1 + A_k\left(\varphi\right)e^{-bp}},\ k = 1, \ldots, n. \tag{17}$$

From (19), it follows that

$$e^{-bp} = \frac{q_1}{A_1\left(\varphi\right)\left(1 - q_1\right)}$$

and we express q_k for $k = 2, \ldots, n$ as a function of q_1

$$f_k(q_1) = \frac{A_k(\varphi)\left(\dfrac{q_1}{(1-q_1)A_1(\varphi)}\right)}{1+A_k(\varphi)\left(\dfrac{q_1}{(1-q_1)A_1(\varphi)}\right)},\ k = 2, \ldots, n$$

to obtain the following univariate optimization problem:

$$\max_{q_1 \in [0,1]} \left\{ \overline{\pi}(q_1 \mid \varphi) := \left(p(q_1) - \varphi\right) w_1 q_1 + \sum_{k=2}^{n} \left(p(q_1) - \varphi\right) w_k f_k(q_1) \right\}.$$

One consequence of the generality of the MMNL model is that the profit function is not unimodal (Hansen and Martin 1990), even for the case of a single product (e.g., see Figure 1 in Li et al. 2019). However, the optimality of specialized, mixed, and diversified sourcing strategies persists under MMNL demand if differences in nominal valuations across customer segments are not too large. This is formalized in the following corollary.

Corollary 2. *If*

$$\frac{\max_k \left\{ e^{v_k} \left(1 + \alpha e^{v_k - b\varphi_h}\right)^{-1} \right\}}{\min_k \left\{ e^{v_k} \left(1 + \alpha e^{v_k - b\varphi_h}\right)^{-1} \right\}} \le 2,$$

then $\overline{\pi}(q_1 \mid \varphi)$ *is concave on* $[0, 1]$ *for all* $\varphi \in [\varphi_l, \varphi_h]$, *and one of three distinct sourcing strategies is optimal: specialized, mixed, diversified.*

Proposition A4 in the appendix shows that our results continue to hold under an additive form of yield uncertainty, rather than the multiplicative form. The above results show that the three sourcing strategies – specialized, mixed, diversified – are optimal under more complex demand representations. In sum, we find that the optimal sourcing strategies are robust over alternative definitions of yield and demand.

5. Conclusion

In this paper, we study how characteristics of the coffee bean grower environment and the specialty roast market environment affect a specialty roaster's DT sourcing strategy for an SL roast. We describe the real-world details of the business environment, which are nuanced and complex. We define a stylized model that takes a first step in the literature toward capturing the essence of the system for the purposes of our research questions. Key features that we incorporate into our model include: uncertainty in yield, uncertainty in BL price, two distinct products that compete with one another – a high-priced SL and a lower-priced BL, two market segments – a connoisseur segment that prefers specialty roasts and a mass-market segment that considers both specialty and commercial roasts, a decision prior to harvesting on the amount of DT contract acreage for the single-origin label, a decision after harvesting on the allocation of harvested beans between the single-origin label and the blend label.

We find that a specialty roaster's optimal sourcing strategy for single-origin beans is one of three distinct types: specialized, diversified, and mixed. When a specialized sourcing strategy is optimal, there is virtually no chance that any of the beans will be allocated for sale under the blend label; all the harvested beans are consistently sold under the single-origin label. When a diversified sourcing strategy is optimal, the specialty roaster always uses some of the beans for sale under the single-origin label and some for sale under the blend label. A mixed sourcing strategy occupies the middle ground between these two extremes.

Our study develops the conditions under which each sourcing strategy is optimal. The optimality of these three strategies is robust because they continue to appear as the optimal choices under alternative representations of yield uncertainty and demand. Our study also presents a comprehensive analysis and discussion pertaining to the impact of changing the yield variance.

Our paper develops a rich set of insights. First, there are two primary factors that determine the optimal sourcing strategy: mean price of inferior product BL and the marginal cost of the superior product SL. An increase in the mean BL price moves the policy preference from the specialized sourcing strategy to mixed, and to diversified sourcing strategies. An increase in the marginal cost has the opposite effect, moving the strategy choice from diversified sourcing to mixed and to specialized sourcing strategies.

Second, the comparison with the newsvendor model reveals distinct insights regarding the impact of increasing uncertainty. While the optimal quantity and profit under the mixed strategy follows a behavior similar to the characteristics of the newsvendor model, we identify distinct structural differences. In the newsvendor problem, the optimal order quantity can be increasing or decreasing in variance depending on the problem parameters; however, the optimal quantity is strictly decreasing under the specialized sourcing strategy and it does not change under the diversified strategy in the absence of yield-price correlation.

Third, our study identifies a "farmer's curse" where the weighted average price of an agricultural product is decreasing in negative yield-price correlation. Farmer's curse leads to several lessons for growers and roasters: (1) reducing the negative correlation between the grower's yield and BL price increases the profit, encouraging investments in actions that reduce the sensitivity of grower's yield to weather and other environmental conditions relative to other growers (i.e., by reducing negative price-yield covariance for the grower); (2) discouraging roasters to pay growers with differential pricing (i.e., based on market prices at the time of harvest).

Fourth, we present evidence of a virtuous feedback loop wherein the grower-roaster relationship becomes stronger over time. This encourages the roaster to invest in the grower to improve the consistency of high-quality yields, which in turn can help improve living standards for growers through higher-value crops.

There are several opportunities for future research. The modeling approach in our study can be extended in two ways. First, our model can be extended to include multiple single-origin labels that compete with one another (e.g., a roaster may have multiple SLs for sale at the same time). We caution that such a model is more complex, raising the possibility of a first-stage objective function that is not well behaved (e.g., not unimodal). Second, we assume that, under mixed and diversified strategies, the amount of the single-origin beans allocated to the blend label does not affect the BL price. This assumption is motivated by large differences in SL versus BL volumes. However, the implications of this assumption for our conclusions could be investigated.

In addition to the above modeling extensions, there are opportunities for empirical investigation. Our model provides two basic predictions that could be tested empirically. First, we describe the character of the optimal sourcing strategy under different business environments. Second, we describe how the DT relationship and sourcing strategy may evolve over time with impacts on grower welfare. Deviations between empirical observation and our predictions point to opportunities to improve our understanding of coffee supply chains through new models that better fit reality.

6. References

Anupindi R, Akella R (1993) Diversification under supply uncertainty. *Management Sci*ence 39(8):994–963.

Boyabatlı O, Kleindorfer PR, Koontz SR (2011) Integrating long-term and short-term contracting in beef supply chains. *Management Science* 57(10):1771–1787.

Boyabatlı O (2015) Supply Management in multiproduct firms with fixed proportions technology. *Management Science* 61(12):3013–3031.

Burke G, Carrilo JE, Vakharia AJ (2009) Sourcing decisions with stochastic supplier reliability and stochastic demand. *Production and Operations Management* 18(4):475-484.

Corless RM, Gonnet GH, Hare DEG, Jeffrey DJ, Knuth DE (1996) On the Lambert W function. *Advances in Computational Mathematics* 5(1):329–359.

Craymer L (2015) China's changing tastes offer upside for coffee. *Wall Street Journal* (September 16).

Dada M, Petruzzi NC, Schwarz LB (2007) A newsvendor's procurement problem when suppliers are unreliable. *Manufacturing and Service Operations Management* 9(1):9–32.

Dong L, Kouvelis P, Tian Z (2009) Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management* 11(2):317–339.

Feng Q (2010) Integrating dynamic pricing and replenishment decisions under supply capacity uncertainty. *Management Science* 56(12):2154-2172.

Feng Q, Shanthikumar (2022) Applications of stochastic orders and stochastic functions in inventory and pricing problems. *Production and Operations Management* 31(4):1433-1453.

Feng Q, Shanthikumar JG, Xue M (2022) Consumer choice models and estimation: a review and extension. *Production and Operations Management* 31(2):847-867.

Gerchak Y, Parlar M (1990) Yield variability, cost tradeoffs and diversification in the EOQ model. *Naval Research Logistics* 37(3):341-354.

Guadagni PM, Little JDC (1983) A logit model of brand choice calibrated on scanner data. *Marketing Science* 2(3):203–238.

Hanson W, Martin RK (1990) Optimal bundle pricing. *Management Science* 36(2):155-174.

Horner W, Lewis JT (2021) Coffee prices soar after bad harvests and insatiable demand. *Wall Street Journal* (July 12).

Jain N, Gitrotra K, Netessine S (2014) Managing global sourcing: inventory performance. *Management Science* 60(5):1202-1222.

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

Kazaz B, Webster S (2011) The impact of yield-dependent trading costs on pricing and production planning under supply uncertainty. *Manufacturing & Service Operations Management* 13(3):404-417.

Kazaz B, Webster S (2015) Technical note—price-setting newsvendor problems with uncertain supply and risk aversion. *Operations Research* 63(4):807-818.

Li H, Huh WT (2011) Pricing multiple products with multinomial logit and nested logit models: concavity and implications. *Manufacturing & Service Operations Management* 13(4):549-563.

Li H, Webster S, Mason N, Kempf K (2019) Product-line pricing under discrete mixed multinomial logit demand. *Manufacturing & Service Operations Management* 21(1):14-28.

Li Q, Zheng S (2006) Joint inventory replenishment and pricing control for systems with uncertain yield and demand. *Operations Research* 54(4):696-705.

Marcos CM (2021) Your daily coffee habit is about to get more expensive. *New York Times* (August 12).

McFadden D, Train K (2000) Mixed MNL models for discrete response. *Journal of Applied Econometrics* 15(5):447-470.

Noparumpa T, Kazaz B, Webster S (2011) Production planning under supply and quality uncertainty with two customer segments and downward substitution. Working paper, Syracuse University.

Noparumpa T, Kazaz B, Webster S (2015) Wine futures and advance selling under quality uncertainty. *Manufacturing & Service Operations Management* 17(3):411-426.

Orr M (2021) Coffee could soon cost more at grocery stores and cafes. Here’s why. *Associates Press* (September 27).

Osborn K (2015) Why El Niño might make your latte cost more. *Money* (September 3).

Parlar M, Wang D (1993) Diversification under yield randomness in two simple inventory models. *European Journal of Operational Research* 66(1):52-64.

Shaked M, Shanthikumar JG (2007) *Stochastic Orders* (Springer, New York, NY).

Stabiner K (2015) Europe (finally) wakes up to superior coffee. *Wall Street Journal* (September 29).

Strand O (2015) Coffee's next generation of roasters. *Wall Street Journal* (May 22).

Tan B, Feng Q, Chen W (2016) Dual sourcing under random supply capacities: the role of the slow supplier. *Production and Operations Management* 25(7):1232-1244.

Tang CS (2006) Perspectives in supply chain risk management. *International Journal of Production Economics* 103(2):451-488.

Tang C, Yin R (2007) Responsive pricing under supply uncertainty. *European Journal of Operational Research* 182(1):239-255.

Tomlin B, Wang Y (2005) On the value of mix flexibility and dual-sourcing in unreliable newsvendor networks. *Manufacturing & Service Operations Management* 7(1):37-57.

Wernau J (2015) Coffee disconnect is brewing. *Wall Street Journal* (August 6).

Yano AY, Lee HL (1995) Lot sizing with random yields: a review. *Operations Research* 43(2):311-334.

7. Appendix

1. List of Notation

Q = normalized acres under DT contract for the single-origin label

$k(Q)$ = cost prior to harvest for Q normalized acres under contract for the single-origin label (e.g., payment to the grower(s) to help with expenses of tending to the crop); $k(Q)$ is continuous and the marginal cost $k'(Q) \geq 0$ is nondecreasing in Q

c = variable cost after harvest per normalized pound (that satisfies quality specifications), a portion of which is paid to the grower(s) and a portion that is incurred by the roaster (e.g., cost of transportation, storage, roasting, packaging)

q = connoisseur demand for the roaster's single-origin label in normalized pounds

q_B = connoisseur demand for the roaster's blend label in normalized pounds

q_0 = fraction of the connoisseur segment who choose not to purchase coffee beans from the roaster

p = selling price per normalized pound for the roaster's single-origin label

b = price sensitivity coefficient in MNL consumer utility

v = intercept of single-origin label consumer utility, i.e., nominal quality measure of SL beans

v_B = intercept of blend label consumer utility, i.e., nominal quality measure of BL beans, $v_B > v$

$\tilde{y}$ = random yield per normalized acre with positive support $[y_l, y_h]$, mean $\mu_y = 1$ (due to normalized units), and variance σ_y^2, i.e., $\tilde{y}$ is the harvested normalized pounds per normalized acre that satisfy the quality specs

$\tilde{\varphi}$ = random selling price per normalized pound for the roaster's blend label with positive support $[\varphi_l, \varphi_h]$, mean μ_φ, and variance σ_φ^2

$F(\varphi, y)$ = probability distribution function for $(\tilde{\varphi}, \tilde{y})$ with support $\Omega = [\varphi_l, \varphi_h] \times [y_l, y_h]$, covariance $\sigma_{\varphi y}$, and correlation coefficient ρ

$\tilde{\boldsymbol{\varepsilon}} = (\tilde{\varepsilon}, \tilde{\varepsilon}_0, \tilde{\varepsilon}_B)$ = i.i.d. normalized Gumbel random variables

$\tilde{U} = v - bp + \tilde{\varepsilon}$ = random single-origin label consumer utility

$\tilde{U}_B = v_B - b\varphi + \tilde{\varepsilon}_B$ = random blend label consumer utility

$\tilde{U}_0 = \tilde{\varepsilon}_0$ = random no-purchase utility, e.g., v_0 is normalized to zero without loss of generality

2. Calibrations for Numerical Examples

The marginal distributions of blend price and yield are both truncated normal; the marginal density functions for random price and yield are

$$f_\varphi(x) = \phi\left(\frac{x-\mu_\varphi}{\bar{\sigma}_\varphi}\right) / \left[\bar{\sigma}_\varphi\left(\Phi\left(\frac{\varphi_h-\mu_\varphi}{\bar{\sigma}_\varphi}\right) - \Phi\left(\frac{\varphi_l-\mu_\varphi}{\bar{\sigma}_\varphi}\right)\right)\right], \; x \in [y_l, y_h] \text{ and } \mu_y = 0.5(y_l + y_h)$$

$$f_y(x) = \phi\left(\frac{x-\mu_y}{\bar{\sigma}_y}\right) / \left[\bar{\sigma}_y\left(\Phi\left(\frac{y_h-\mu_y}{\bar{\sigma}_y}\right) - \Phi\left(\frac{y_l-\mu_y}{\bar{\sigma}_y}\right)\right)\right], \; x \in [y_l, y_h] \text{ and } \mu_y = 0.5(y_l + y_h)$$

where ϕ and Φ are the pdf and cdf of the standard normal distribution (the blend-label price marginal density is similarly defined). The standard deviation of random price and yield are given by

$$\sigma_\varphi = \bar{\sigma}_\varphi\left(1 - \frac{2\left(\frac{\varphi_h-\mu_\varphi}{\bar{\sigma}_\varphi}\right)\phi\left(\frac{\varphi_h-\mu_\varphi}{\bar{\sigma}_\varphi}\right)}{\Phi\left(\frac{\varphi_h-\mu_\varphi}{\bar{\sigma}_\varphi}\right) - \Phi\left(\frac{\varphi_l-\mu_\varphi}{\bar{\sigma}_\varphi}\right)}\right)^{1/2}.$$

$$\sigma_y = \bar{\sigma}_y\left(1 - \frac{2\left(\frac{y_h-\mu_y}{\bar{\sigma}_y}\right)\phi\left(\frac{y_h-\mu_y}{\bar{\sigma}_y}\right)}{\Phi\left(\frac{y_h-\mu_y}{\bar{\sigma}_y}\right) - \Phi\left(\frac{y_l-\mu_y}{\bar{\sigma}_y}\right)}\right)^{1/2}.$$

Note that the density function approaches the uniform distribution as $\bar{\sigma}_\varphi$ or $\bar{\sigma}_y$ standard goes to infinity.

We use a Gaussian copula to generate random variables consistent with correlation ρ. The mean BL price per pound is μ_φ = \$15 with range [\$13, \$17] and yield range is [0.5, 1.5]. The price sensitivity is $b = 1$, the nominal perceived quality of SL by the connoisseur segment is $v = 20$, and the nominal quality of BL by the connoisseur segment is $v_B = 17$. The variable cost per output unit is c = \$13.25. We consider various input-unit-marginal-cost functions $k'(Q)$.

For Figure 3, $\rho = -0.1$, $\bar{\sigma}_\varphi = 2$, and $\bar{\sigma}_y = 0.15$, yielding $\sigma_\varphi = 1.08$ and $\sigma_y = 0.15$. The solid red curve in Figure 3 shows $k'(Q)$ = \$3.3 for $Q \leq 0.34$, then jumps to a level more than \$3.7 as Q passes above 0.34. The dotted red curve in Figure 3 shows $k'(Q)$ = \$1.65 for $Q \leq 0.8$, then jumps to \$2.05 Q passes above 0.8. For this calibration, the mean optimal SL price per pound ranges from \$18.67[6] to \$17.44[7] as Q ranges from 0.3 to 1.2 normalized acres.

[6] The 0.2 and 0.8 SL price percentiles are \$17.71 and \$19.62.
[7] The 0.2 and 0.8 SL price percentiles are \$16.46 and \$18.44.

For figures 5 and 6, $\rho \in \{0, -0.1\}$, $\overline{\sigma}_\varphi = 2$, and $\overline{\sigma}_y \in [0.01, 0.20]$, yielding $\sigma_\varphi = 1.08$ and $\sigma_y \in [0.01, 0.19]$.[8] For the scenario where specialized sourcing is optimal, $k' = \$2.93$, and for the scenario where diversified sourcing is optimal, $k' = \$1.40$, and for the scenario where diversified sourcing is optimal, $k' = \$0.75$. For this calibration, we calculate that $Q_L = 0.358$ and $Q_H = 1.214$. For the mixed strategy, we set the optimal quantity at the midpoint between Q_L and Q_H. The average blend-label price is \$15, the average single-origin label price varies between \$18.31 and \$18.39 under specialized and mixed strategies, and remains stable at \$17.44 under a diversified strategy as σ_y and ρ vary.

We note that closed-form expressions are not available for the curves in the figures. We use Analytical Solver Platform by Frontline Systems with a sample size of 100,000 to generate the summary statistics that are necessary for the figures.

3. Supplemental Propositions

Proposition A1. $\Pi(Q)$ *is strictly concave on* $[0, \infty)$.

Proof. The first-stage profit function can be expressed as

$$\Pi(Q) = \int_{\Omega_c(Q)} \pi^*\left(Q,\varphi,y \mid q^* = Qy\right) dF + \int_{\Omega_u(Q)} \pi^*\left(Q,\varphi,y \mid q^* = q^{\circ}\right) dF$$

where $\Omega_c(Q) = \left\{(\varphi, y) : q^* = Qy, (\varphi, y) \in \Omega\right\}$ and $\Omega_u(Q) = \left\{(\varphi, y) : q^* = q^{\circ}, (\varphi, y) \in \Omega\right\}$, i.e., $\Omega_c(Q)$ and $\Omega_u(Q)$ partition Ω into realizations for which the constraint is binding and nonbinding in the second-stage problem. Note that

$$\pi^*\left(Q,\varphi,y \mid q^* = q^{\circ}\right) = \frac{1}{b} W\left(\frac{e^{v-1}}{e^{b\varphi} + e^{v_B}}\right) + (\varphi - c)Qy - k(Q)$$

is concave in Q (because $(c - \varphi)Qy + k(Q)$ is convex. Now consider

$$\pi^*\left(Q,\varphi,y \mid q^* = Qy\right) = \left(\frac{1}{b}\left(v - \ln\left(\frac{Qy\left(1 + e^{v_B - b\varphi}\right)}{1 - Qy}\right)\right) - \varphi\right)Qy + (\varphi - c)Qy - k(Q).$$

As noted in the proof of Proposition 1,

$$\pi(q) = \left(p(q) - \varphi\right)q + (\varphi - c)Qy - k(Q) = \left(\frac{1}{b}\left(v - \ln\left(\frac{q\left(1 + e^{v_B - b\varphi}\right)}{1 - q}\right)\right) - \varphi\right)q + (\varphi - c)Qy - k(Q)$$

is strictly concave in q, i.e.,

[8] For reference, the standard deviation of a uniform random variable on [0.5, 1.5], the upper limit for our distribution, is $12^{-1/2} \approx 0.29$, e.g., the probability density function is relatively flat on [0.5, 1.5] when σ_y is set at its maximum value of $\sigma_y = 0.19$.

$$g''(q) = \frac{\partial^2}{\partial q^2}\left(\frac{1}{b}\left(v - \ln\left(\frac{q\left(1+e^{v_B - b\varphi}\right)}{1-q}\right)\right) - \varphi\right)q < 0.$$

Therefore, $\pi^*\left(Q,\varphi,y \mid q^* = Qy\right)$ is strictly concave in Q, i.e.,

$$\frac{\partial^2}{\partial Q^2}\left[\left(\frac{1}{b}\left(v - \ln\left(\frac{Qy\left(1+e^{v_B - b\varphi}\right)}{1-Qy}\right)\right) - \varphi\right)Qy + (\varphi - c)Qy - k(Q)\right] = g''(q)y^2 - k''(Q) < 0.$$

Thus, $\Pi(Q)$ is the sum of up to two integrals for which each integrand is strictly concave in Q for any realization, which implies that $\Pi(Q)$ is strictly concave. □

Without loss of generality, we may express random blend price as the sum of a reference random variable, denoted $\tilde{\varphi}_0$, and some constant x, i.e., $\tilde{\varphi} = \tilde{\varphi}_0 + x$. As x increases, the endpoints of the support and the mean of $\tilde{\varphi}$ increase while variance and correlation remain fixed. The next proposition characterizes the effect of increasing random blend price through x.

Proposition A2. *Optimal acreage Q^* is increasing in x, and Q_L and Q_H are decreasing in x.*

Proof. Note that $\mu_\varphi = \mu_{\varphi_0} + x$, $\varphi_l = \varphi_{0l} + x$, $\varphi_h = \varphi_{0h} + x$, $\sigma_\varphi = \sigma_{\varphi_0}$, and $(\tilde{\varphi}, \tilde{y})$ correlation is unaffected by changes in x. Furthermore,

$$Q_L = \frac{1}{y_h} W\left(\frac{e^{v-1}}{e^{b\varphi_h} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\varphi_h} + e^{v_B}}\right)\right)^{-1} = \frac{1}{y_h} W\left(\frac{e^{v-1}}{e^{b(\varphi_{0h}+x)} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b(\varphi_{0h}+x)} + e^{v_B}}\right)\right)^{-1}$$

$$Q_H = \frac{1}{y_l} W\left(\frac{e^{v-1}}{e^{b\varphi_l} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\varphi_l} + e^{v_B}}\right)\right)^{-1} = \frac{1}{y_l} W\left(\frac{e^{v-1}}{e^{b(\varphi_{0l}+x)} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b(\varphi_{0l}+x)} + e^{v_B}}\right)\right)^{-1}.$$

The above can be rewritten as

$$\frac{y_h Q_L}{1 - y_h Q_L} = W\left(\frac{e^{v-1}}{e^{b(\varphi_{0h}+x)} + e^{v_B}}\right) \quad \Rightarrow \quad \frac{y_h Q_L}{1 - y_h Q_L} e^{\frac{y_h Q_L}{1 - y_h Q_L}} = \frac{e^{v-1}}{e^{b(\varphi_{0h}+x)} + e^{v_B}} \tag{18}$$

$$\frac{y_l Q_H}{1 - y_l Q_H} = W\left(\frac{e^{v-1}}{e^{b(\varphi_{0l}+x)} + e^{v_B}}\right) \quad \Rightarrow \quad \frac{y_l Q_H}{1 - y_l Q_H} e^{\frac{y_l Q_H}{1 - y_l Q_H}} = \frac{e^{v-1}}{e^{b(\varphi_{0l}+x)} + e^{v_B}}, \tag{19}$$

from which it is clear that an increase in x requires a decrease in Q_L and Q_H to maintain the identities (i.e., the LHS expressions are increasing in Q_L and in Q_H and the RHS expressions are decreasing in x). Therefore, Q_L and Q_H are decreasing in x.

The effect of increasing x on Q^* is more subtle because there are opposing directional forces. From Proposition 2,

$$Q^* = \max\left\{Q : \left(\mu_{\varphi_0} + x\right) + \rho\sigma_\varphi\sigma_y + \lambda(Q) \geq k'(Q) + c\right\} \tag{20}$$

$$\lambda(Q) = E\left[\left(m_c\left(Q,\tilde{\varphi},\tilde{y}\right) - m_u\left(Q,\tilde{y}\right)\right)^+\right] = E\left[\left(\frac{\tilde{y}}{b}\ln\left[\left(\frac{e^v}{e^{b(\tilde{\varphi}_0+x)} + e^{v_B}}\right)\left(\frac{1-Q\tilde{y}}{Q\tilde{y}}\right)\right] - \frac{\tilde{y}}{b}\left(\frac{1}{1-Q\tilde{y}}\right)\right)^+\right]$$

and thus, for any $\delta > 0$,

$$\left(\mu_{\varphi_0} + x\right) + \rho\sigma_\varphi\sigma_y + \lambda\left(Q^*\right) \geq k'\left(Q^*\right) + c$$

$$\left(\mu_{\varphi_0} + x\right) + \rho\sigma_\varphi\sigma_y + \lambda\left(Q^* + \delta\right) < k'\left(Q^* + \delta\right) + c.$$

Note that

$$\frac{\partial}{\partial x}\left(\frac{y}{b}\ln\left[\left(\frac{e^v}{e^{b(\varphi_0+x)} + e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right] - \frac{\tilde{y}}{b}\left(\frac{1}{1-Qy}\right)\right) = -\left(\frac{ye^{b(\varphi_0+x)}}{e^{b(\varphi_0+x)} + e^{v_B}}\right),$$

and thus

$$-\frac{\partial}{\partial x}\lambda(Q) = E\left[\frac{\tilde{y}e^{b(\tilde{\varphi}_0+x)}}{e^{b(\tilde{\varphi}_0+x)} + e^{v_B}}\middle| m_c\left(Q,\tilde{\varphi},\tilde{y}\right) > m_u\left(Q,\tilde{y}\right)\right]\Pr\left(m_c\left(Q,\tilde{\varphi},\tilde{y}\right) > m_u\left(Q,\tilde{y}\right)\right)$$

$$\leq E\left[\frac{\tilde{y}}{1 + e^{v_B - b(\tilde{\varphi}_0+x)}}\right] \qquad \text{(because the ratio is positive for all realizations)}$$

$$\leq E\left[\tilde{y}\right] \qquad \text{(because the denominator is positive for all realizations)}$$

$$= 1,$$

i.e., the LHS of the optimality condition (22) is decreasing in x. Therefore, since the LHS of the optimality condition (22) is decreasing in Q and the RHS is increasing in Q, if follows that Q^* is increasing in x. □

The corollary below follows from Proposition A2.

Corollary A1. *As blend price decreases, the optimal sourcing strategy may shift, but only in one direction—from diversified to mixed, or from mixed to specialized, not the other way around.*

Proposition A3. (*a*) *Q_L and Q_H increase as v increases, v_B decreases, b decreases.* (*b*) *If $Q^* > Q_H$, then Q^* is unchanged as v increases, v_B decreases, b decreases.* (*c*) *If $Q^* < Q_H$, then Q^* increases as v increases, v_B decreases, b decreases.*

Proof. Part (*a*). From the proof of Proposition A2 and expressions (20) and (21), it is clear that Q_L and Q_H are increasing in v and decreasing in v_B and b, i.e., the RHS of (20) and (21) is increasing in v and decreasing in v_B and b, and the LHS is increasing in Q_L (in (20)) and in Q_H (in (21)). Thus, the result follows.

Part (*b*). If $Q^* > Q_H$, then $\lambda(Q^*) = 0$. Parameters v, v_B, and b are not present in the optimality condition (see (22)), then thus Q^* is unaffected by changes in v, v_B, and b. ('>' appears instead of '≥' in the proposition because the left and right partial derivatives of $\lambda(Q^*)$ with respect to v, v_B, and b differ; in particular, the right derivative with respect to v is positive and the left derivative is zero; vice-versa for v_B and b.)

Part (*c*). Recall

$$\lambda(Q)=E\left[\left(m_c\left(Q,\tilde{\varphi},\tilde{y}\right)-m_u\left(Q,\tilde{y}\right)\right)^+\right]=\frac{1}{b}E\left[\tilde{y}\left(\ln\left[\left(\frac{e^v}{e^{b(\tilde{\varphi}_0+x)}+e^{v_B}}\right)\left(\frac{1-Q\tilde{y}}{Q\tilde{y}}\right)\right]-\left(\frac{1}{1-Q\tilde{y}}\right)\right)^+\right]$$

and that $\tilde{y}$ is nonnegative. Thus, it is clear that $\lambda(Q)$ increases (in the weak sense) as v increases and as v_B and b decrease. Recall also that the RHS of the inequality in the optimality condition (22) in increasing in Q. Therefore, it follows from (22) that the result holds. □

Note that even if $Q^* < Q_H$, optimal acreage Q^* may remain fixed even though $\lambda(Q)$ may strictly increase as v increases, v_B decreases, b decreases. This is because there may be a jump in marginal cost when Q increases above Q^*, i.e., for $\delta = 0^+$ (infinitesimal positive value), $k'(Q^* + \delta) + c$ is much larger than $k'(Q^*) + c$ (see Figure 3 that illustrates marginal cost step functions). Thus, an increase in Q_L and Q_H as v increases, v_B decreases, b decreases can cause the optimal sourcing policy to shift from mixed to specialized or from diversified to mixed (e.g., as Q_L or Q_H increase from below Q^* to above Q^*). Since Q^* does not change as v increases, v_B decreases, b decreases once it reaches the threshold Q_H whereas Q_H is increasing, it is not possible for the strategy to shift from mixed to diversified as v increases, v_B decreases, b decreases. On the other hand, if $Q^* \le Q_L$ (i.e., specialized sourcing is optimal), then Q^* and Q_L may increase as v increases, v_B decreases, b decreases with the relate rates of increase dependent on parameters, implying that the possibility of strategy shift from specialized to mixed cannot be excluded.

The model for yield presented in the manuscript is based on multiplicative uncertainty, which is the dominant model in the literature for yield uncertainty. However, an alternative model is additive uncertainty, which is relatively common in the literature for modeling uncertain demand.

Multiplicative yield uncertainty: random yield $= Q\tilde{y}$, normalized such that $E[\tilde{y}] = 1$

Additive yield uncertainty: random yield $= Q + \tilde{y}$, normalized such that $E[\tilde{y}] = 0$

One may wonder whether our main conclusions are robust to a shift from multiplicative to additive yield uncertainty. As a step to addressing this question, we present the following proposition to be used in subsequent proofs.

Proposition A4. *For multiplicative yield uncertainty and for additive yield uncertainty*:

(*a*) *Under specialized and mixed sourcing, first-stage profit* Π *is strictly concave in yield.*

(*b*) *Under diversified sourcing, first-stage profit* Π *is linear in yield.*

Proof. We first present the proof for the case of multiplicative yield uncertainty. After this is complete, we prove that the results extend to the case of additive yield uncertainty.

Under multiplicative yield uncertainty, it follows from Proposition 1 that realized first-stage profit can be expressed as

$$\hat{\Pi}(Q,\varphi,y)=\pi^*(Q,\varphi,y)=h(\varphi,y)+(\varphi-c)Qy-k(Q) \tag{21}$$

where

$$h(\varphi,y)=\frac{1}{b}\left(Qyv-g_1(\varphi,y)-g_2(y)\right)\times\mathbf{1}_{q^*=Qy}+\frac{1}{b}W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\times\mathbf{1}_{q^*>Qy}$$

$$\frac{g_1(\varphi,y)}{Q}=y\ln\left(e^{b\varphi}+e^{v_B}\right) \tag{22}$$

$$\frac{g_2(y)}{Q}=y\ln\left(\frac{Qy}{1-Qy}\right) \tag{23}$$

$$q^*=\min\left\{W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1},Qy\right\}$$

(see Proposition 1).

Part (*a*). First suppose that $Q\le Q_L$ so that specialized sourcing is in effect. From Lemma 1, it follows that, under specialized sourcing, (23) reduces to

$$\hat{\Pi}(Q,\varphi,y)=\pi^*(Q,\varphi,y)=\frac{1}{b}\left(Qyv-g_1(\varphi,y)-g_2(y)\right)+(\varphi-c)Qy-k(Q)\,. \tag{24}$$

Function Qyv, g_1 and function $(\varphi-c)Qy$ are linear in y, and

$$\frac{g_2{}'(y)}{Q}=\ln\left(\frac{Qy}{1-Qy}\right)+y\left(\frac{1-Qy}{Qy}\right)\left(\frac{Q(1-Qy)+Q^2y}{(1-Qy)^2}\right)=\ln\left(\frac{Qy}{1-Qy}\right)+\frac{1}{1-Qy}$$

$$\frac{g_2{}''(y)}{Q}=\frac{Q}{Qy(1-Qy)}+\frac{Q}{(1-Qy)^2}=\frac{1}{y(1-Qy)}+\frac{Q}{(1-Qy)^2}>0$$

(as shown in the proof of Corollary 1, $0<Qy<1$ for $Q\le Q_L$; e.g., see (35)). Therefore, $\hat{\Pi}(Q,\varphi,y)$ is strictly concave in y.

Now suppose that $Q\in(Q_L,Q_H)$ so that mixed sourcing is in effect. We will show that

$$h_1(\varphi,y)=\frac{1}{b}\left(Qyv-g_1(\varphi,y)-g_2(y)\right)$$

is concave and maximized at the boundary point y^{o} satisfying

$$Qy^{\text{o}}=W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1},$$

which assures that $h(\varphi, y) = h_1(\varphi, y) \times \mathbf{1}_{q^*=Qy} + \frac{1}{b} W\left(\frac{e^{v-1}}{e^{b\varphi} + e^{v_B}}\right) \times \mathbf{1}_{q^*>Qy}$ is concave in *y*.

The above analysis shows $h_1(\varphi, y)$ is concave in *y* for $y < y^{\text{o}}$. Note that

$$\pi^*(Q,\varphi,y) = \begin{cases} \frac{1}{b}\left(\ln\left[\left(\frac{e^v}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right]\right)Qy + (\varphi - c)Qy - k(Q), & \text{if } q^* = Qy \\ \frac{1}{b} W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right) + (\varphi - c)Qy - k(Q), & \text{if } q^* < Qy \end{cases}$$

$$= \max_q \left\{\hat{\pi}(q) = (p(q) - \varphi)q : 0 \le q \le 1, q \le Qy\right\} + (\varphi - c)Qy - k(Q)$$

$$p(q) = \frac{1}{b}\left(v - \ln\left(\frac{q\left(1 + e^{v_B - b\varphi}\right)}{1-q}\right)\right)$$

(see Proposition 1). As shown in the proof of Proposition 1, function $\hat{\pi}(q) = (p(q) - \varphi)q$ has a unique unconstrained optimum

$$\max_q \left\{\hat{\pi}(q)\right\} = \frac{1}{b} W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right) = \frac{1}{b}\left(\ln\left[\left(\frac{e^v}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy^{\text{o}}}{Qy^{\text{o}}}\right)\right]\right)Qy^{\text{o}}$$

$$\ge \frac{1}{b}\left(\ln\left[\left(\frac{e^v}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right]\right)Qy$$

for all *Qy*. Therefore

$$h(\varphi, y) = h_1(\varphi, y) \times \mathbf{1}_{q^*=Qy} + \frac{1}{b} W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right) \times \mathbf{1}_{q^*>Qy} \text{ is concave in } y.$$

Part (*b*). Under diversified sourcing, we have $Q \ge Q_H$, and from Lemma 1 it follows that

$$\hat{\Pi}(Q,\varphi,y) = \pi^*(Q,\varphi,y) = \frac{1}{b} W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right) + (\varphi - c)Qy - k(Q), \tag{25}$$

which is linear in *y*.

Now suppose that we have additive yield uncertainty, i.e.,

random yield $= Q + \tilde{y}$, normalized such that $E[\tilde{y}] = 0$

(we assume the distribution is such that negative yield is not possible over the range of possible values of *Q* in the optimization problem). It is straightforward to verify that when *Qy* in Proposition 1 is replaced by *Q* + *y*, realized first-stage profit (given the optimal second-stage decision) can be expressed as

$$\hat{\Pi}(Q,\varphi,y)=\frac{1}{b}\max\left\{(Q+y)v-g_1(\varphi,y)-g_2(y),W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right\}+(\varphi-c)(Q+y)-k(Q) \tag{26}$$

where g_1 and g_2 are redefined as

$$g_1(\varphi,y)=(Q+y)\ln\left(e^{b\varphi}+e^{v_B}\right) \tag{27}$$

$$g_2(y)=(Q+y)\ln\left(\frac{Q+y}{1-(Q+y)}\right). \tag{28}$$

In effect, under multiple yield uncertainty, we have functions of Qy whereas under additive uncertainty, we have functions of $Q+y$. This difference does not affect the properties of convexity and concavity. For example, consider function $\hat{g}_2(z)=z\ln\left(\frac{z}{1-z}\right)$ where $z=Qy$ in one case, and $z=Q+y$ in another case.

Then

$$\frac{d\hat{g}_2(Qy)}{dy}=Q\hat{g}_2'(Qy)$$

$$\frac{d^2\hat{g}_2(Qy)}{dy^2}=Q^2\hat{g}_2''(Qy) \tag{29}$$

$$\frac{d\hat{g}_2(Q+y)}{dy}=\hat{g}_2'(Q+y)$$

$$\frac{d^2\hat{g}_2(Q+y)}{dy^2}=\hat{g}_2''(Q+y), \tag{30}$$

i.e., the sign of (31) matches the sign of (32). Thus, following the arguments of the proof for the case of multiplicative yield uncertainty, the results also hold for the case of additive yield uncertainty. □

4. Proofs of Propositions

Proof of Proposition 1. Define $\hat{\pi}(q)=\left(p(q)-\varphi\right)q$ and note that

$$q^*=\arg\max_q\left\{\pi(q):0\le q\le 1,q\le Qy\right\}=\arg\max_q\left\{\hat{\pi}(q):0\le q\le 1,q\le Qy\right\}.$$

Note that

$$p(q)=\frac{1}{b}\left(v-\ln\left(\frac{q\left(1+e^{v_B-b\varphi}\right)}{1-q}\right)\right)=\frac{1}{b}\left(v-\ln\left(1+e^{v_B-b\varphi}\right)-\ln(q)+\ln(1-q)\right)$$

$$p'(q)=-\frac{1}{b}\left(\frac{1}{q}+\frac{1}{1-q}\right).$$

Thus,

$$\hat{\pi}'(q)=p'(q)q+p(q)-\varphi=\frac{1}{b}\left(v-\ln\left(\frac{q\left(1+e^{v_B-b\varphi}\right)}{1-q}\right)-q\left(\frac{1}{q}+\frac{1}{1-q}\right)-b\varphi\right)$$

$$=\frac{1}{b}\left(v-b\varphi-1-\ln\left(\frac{q}{1-q}e^{\frac{q}{1-q}}\right)-\ln\left(1+e^{v_B-b\varphi}\right)\right)=\frac{1}{b}\left(\ln\left(\frac{e^{v-b\varphi-1}}{1+e^{v_B-b\varphi}}\right)-\ln\left(\frac{q}{1-q}e^{\frac{q}{1-q}}\right)\right).$$

It is straightforward to verify that $\hat{\pi}(q)$ is strictly concave on [0, 1] (i.e., $\hat{\pi}''(q)<0$). Setting $\hat{\pi}'(q)=0$ and solving for q yields a unique stationary point q° satisfying

$$\frac{q^{\circ}}{1-q^{\circ}}=W\left(\frac{e^{v-b\varphi-1}}{1+e^{v_B-b\varphi}}\right)\Rightarrow q^{\circ}=W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1}, \tag{31}$$

which implies that the constrained optimal solution is

$$q^{*}=\min\left\{W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1},Qy\right\}.$$

The expressions for p^* and $\pi^*(Q,\varphi,y)$ are obtained by substituting $q^*=Qy$ (binding constraint) and $q^*=q^{\circ}$ (nonbinding constraint) into the price and profit expressions. □

Proof of Lemma 1. From Proposition 1, the unconstrained optimal stage-two decision is

$$q^{\circ}(\varphi)=W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1}.$$

Note that that $W(z)$ is positive and monotonic increasing on (0, ∞), and that $\frac{W(z)}{1+W(z)}=\frac{1}{1/W(z)+1}$, is increasing in W for $W>0$, and thus $\frac{d}{dz}\left(\frac{W(z)}{1+W(z)}\right)>0$. Furthermore, $\frac{d}{d\varphi}\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)<0$. Therefore,

$$\frac{d}{d\varphi}W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1}=\left(\frac{d}{dz}\frac{W(z)}{1+W(z)}\right)\left(\frac{d}{d\varphi}\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)<0.$$

From the above inequality, it follows that $q^{\circ}{}'(\varphi)<0$. Therefore,

$$\min_{\varphi,y}\left\{\frac{q^{\circ}(\varphi)}{y}:(\varphi,y)\in\Omega\right\}=Q_L=\frac{1}{y_h}W\left(\frac{e^{v_1-1}}{e^{b\varphi_h}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v_1-1}}{e^{b\varphi_h}+e^{v_B}}\right)\right)^{-1}$$

$$\max_{\varphi,y}\left\{\frac{q^{\circ}(\varphi)}{y}:(\varphi,y)\in\Omega\right\}=Q_H=\frac{1}{y_l}W\left(\frac{e^{v-1}}{e^{b\varphi_l}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi_l}+e^{v_B}}\right)\right)^{-1}$$

where $\Omega = [\varphi_l, \varphi_h] \times [y_l, y_h]$. Thus, the supply constraint in stage two will always be binding if $Q \le Q_L$, and never binding if $Q \ge Q_H$. Therefore,

$$\Pr\left(q^*(Q,\tilde{\varphi},\tilde{y}) < Q\tilde{y}\right) = 0 \text{ iff } Q \le Q_L \text{ and } \Pr\left(q^*(Q,\tilde{\varphi},\tilde{y}) < Q\tilde{y}\right) = 1 \text{ iff } Q \ge Q_H. \square$$

Proof Lemma 2. Note that $y_l \le 1 \le y_h$, and the inequalities are strict if $\sigma_y^2 > 0$. Similarly, $\varphi_l \le \mu_\varphi \le \varphi_h$ with strict inequalities when $\sigma_\varphi^2 > 0$. As shown in the proof of Lemma 1,

$$\frac{d}{d\varphi} W\left(\frac{e^{v-1}}{e^{b\varphi} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\varphi} + e^{v_B}}\right)\right)^{-1} < 0.$$

Therefore, it follows that

$$Q_L = \frac{1}{y_h} W\left(\frac{e^{v-1}}{e^{b\varphi_h} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\varphi_h} + e^{v_B}}\right)\right)^{-1} \le W\left(\frac{e^{v-1}}{e^{b\mu_\varphi} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\mu_\varphi} + e^{v_B}}\right)\right)^{-1} = Q_D$$

$$\le \frac{1}{y_l} W\left(\frac{e^{v-1}}{e^{b\varphi_l} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\varphi_l} + e^{v_B}}\right)\right)^{-1} = Q_H$$

and the inequalities are strict if and only if $\sigma_y^2 > 0$ and/or $\sigma_\varphi^2 > 0$. $\square$

Proof of Proposition 2. Note that

$$\Pi(Q) = E\left[\pi^*(Q,\tilde{\varphi},\tilde{y})\right] = E\left[\left(p^*(Q,\tilde{\varphi},\tilde{y}) - \tilde{\varphi}\right) q^*(Q,\tilde{\varphi},\tilde{y}) + (\tilde{\varphi} - c) Q\tilde{y}\right] - k(Q)$$

$$= E\left[\hat{\pi}^*(Q,\tilde{\varphi},\tilde{y})\right] + \left(E[\tilde{\varphi}\tilde{y}] - c\right)Q - k(Q) = \int_\Omega \hat{\pi}^*(Q,\varphi,y)\, dF + \left(\mu_\varphi + \rho\sigma_\varphi\sigma_y - c\right)Q - k(Q)$$

$$\hat{\pi}^*(Q,\varphi,y) = \begin{cases} \pi_c(Q,\varphi,y) = \dfrac{1}{b} \ln\left[\left(\dfrac{e^v}{e^{b\varphi} + e^{v_B}}\right)\left(\dfrac{1 - Qy}{Qy}\right)\right] Qy, & q^{\circ}(\varphi) \ge Qy \\ \pi_u(\varphi) = \dfrac{1}{b} W\left(\dfrac{e^{v-1}}{e^{b\varphi} + e^{v_B}}\right), & q^{\circ}(\varphi) \le Qy \end{cases}$$

$\Pi(Q)$ is a continuous and strictly concave function (see Proposition A1). However, $\Pi(Q)$ is not assured to be differentiable everywhere because $k(Q)$ may not be differentiable everywhere. Therefore, optimal acreage is given by $Q^* = \max\left\{Q : \Pi'(Q) \ge 0\right\}$.

We partition the support Ω of $(\tilde{\varphi},\tilde{y})$ into two sets that depend on Q:

$$\Omega_c(Q) = \left\{(\varphi,y) : q^*(Q,\varphi,y) = Qy,\ (\varphi,y) \in \Omega\right\}$$

$$\Omega_u(Q) = \left\{(\varphi,y) : q^*(Q,\varphi,y) = q^{\circ}(\varphi),\ (\varphi,y) \in \Omega\right\}.$$

Set $\Omega_c(Q)$ contains realizations for which the supply constraint in the second stage is binding, and $\Omega_u(Q)$ realizations for which the supply constraint is not binding. Function $\hat{\pi}^*(Q,\tilde{\varphi},\tilde{y})$ is piecewise continuous on Ω, and thus

$$\begin{aligned}
\Pi'(Q) &= \int_{\Omega}\left(\frac{\partial}{\partial Q}\hat{\pi}^*(Q,\varphi,y)\right)dF + \mu_\varphi + \rho\sigma_\varphi\sigma_y - k'(Q) - c \\
&= \int_{\Omega_c(Q)}\left(\frac{\partial}{\partial Q}\pi_c(Q,\varphi,y)\right)dF + \mu_\varphi + \rho\sigma_\varphi\sigma_y - k'(Q) - c \\
&= \int_{\Omega_c(Q)}\left(\frac{\partial}{\partial Q}\left(\frac{y}{b}\ln\left[\left(\frac{e^v}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right]Q\right)\right)dF + \mu_\varphi + \rho\sigma_\varphi\sigma_y - k'(Q) - c \\
&= \frac{1}{b}\int_{\Omega_c(Q)} y\left(\ln\left[\left(\frac{e^v}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right] - \frac{1}{1-Qy}\right)dF + \mu_\varphi + \rho\sigma_\varphi\sigma_y - k'(Q) - c \\
&= \mu_\varphi + \rho\sigma_\varphi\sigma_y + \int_{\Omega_c(Q)}\left[m_c(Q,\varphi,y) - m_u(Q,y)\right]dF - k'(Q) - c
\end{aligned}$$

where

$$m_c(Q,\varphi,y) = \frac{y}{b}\ln\left[\left(\frac{e^v}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right]\times\frac{\partial}{\partial Q}Q = \frac{y}{b}\ln\left[\left(\frac{e^v}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right]$$

$$m_u(Q,y) = Q\times\frac{\partial}{\partial Q}\frac{y}{b}\ln\left[\left(\frac{e^v}{e^{b\varphi}+e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right] = \frac{y}{b}\left(\frac{1}{1-Qy}\right).$$

Recall

$$p^* = \begin{cases} \varphi + \dfrac{1}{b}\left(\ln\left[\left(\dfrac{e^v}{e^{b\varphi}+e^{v_B}}\right)\left(\dfrac{1-Qy}{Qy}\right)\right]\right), & \text{if } q^* = Qy \\ \varphi + \dfrac{1}{b}\left(W\left(\dfrac{e^v}{e^{b\varphi}+e^{v_B}}\right)\right), & \text{if } q^* < Qy \end{cases}$$

(see Proposition 1). We see that the value of $m_c(Q,\varphi,y)/y$ is equal to the optimal supply-constrained SL markup over the BL price φ, i.e., the SL market-clearing markup. Accordingly, $m_c(Q,\varphi,y)$ expresses the markup per input unit (normalized acre) instead of output unit (normalized pound). The value $m_c(Q,\varphi,y)$ captures the gain in profit from increasing Q at markup $m_c(Q,\varphi,y)$. On the flipside, value of $m_u(Q,y)$ captures the loss in profit from the consequent reduction in markup to clear the market as Q increases.

Note that $m_c(Q,\varphi,y)$ is decreasing in φ and y, and $m_u(Q,y)$ is increasing in y. We rewrite (10) as $q^{o}(\varphi)/\left(1-q^{o}(\varphi)\right)=W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)$ and observe that the unconstrained second-stage SL price decision can be expressed as

$$p^{o}(\varphi)=\varphi+\frac{1}{b}\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)=\varphi+\frac{1}{b}\left(\frac{1}{1-q^{o}(\varphi)}\right).$$

Thus, at $Qy = q^{o}(\varphi)$ (i.e., (φ, y) is a point on boundary between $\Omega_c(Q)$ and $\Omega_u(Q)$), we have

$$m_c(Q,\varphi,y)=m_c\left(\frac{q^{o}(\varphi)}{y},\varphi,y\right)=m_u(Q,y)=m_u\left(\frac{q^{o}(\varphi)}{y},y\right)=\frac{y}{b}\left(\frac{1}{1-q^{o}(\varphi)}\right)$$

i.e., at $Qy = q^{o}(\varphi)$, both m_c and m_u are equal to the optimal unconstrained markup per input unit. Therefore,

$$\lambda((Q)=\int_{\Omega_c(Q)}\left[m_c(Q,\varphi,y)-m_u(Q,y)\right]dF=E\left[\left(m_c(Q,\tilde{\varphi},\tilde{y})-m_u(Q,\tilde{y})\right)^{+}\right]$$

and

$$\Pi'(\mathrm{Q})=\mu_{\varphi}+\rho\sigma_{\varphi}\sigma_{y}+\lambda(Q)-k'(Q)-c$$

where $\lambda(Q)$ is the expected value of the Lagrange multiplier λ in the Lagrangian formulation of the second-stage optimization problem:

$$\pi^{*}(Q,\varphi,y)=\max_{q\in[0,1],\lambda\geq 0}\left\{L(q,\lambda)=\left(p(q)-\varphi\right)q+(\varphi-c)Qy-k(Q)-\lambda\left(q-Qy\right)\right\}.\ \square$$

Proof of Proposition 3. Since $\Pi(Q)$ is concave (see Proposition A1), it follows that a specialized sourcing strategy is optimal (i.e., $Q^{*}\leq Q_L$) if and only if $\Pi'(Q_L)=\mu_{\varphi}+\rho\sigma_{\varphi}\sigma_{y}+\lambda(Q_L)-k'(Q_L)-c\leq 0$, which is (13). □

Proof of Corollary 1. We begin by identifying properties of functions that we use in the proof. We rearrange (7) to identify an identity for the market-clearing price:

$$Q_L=\frac{1}{y_h}W\left(\frac{e^{v-b\varphi_h-1}}{1+e^{v_B-b\varphi_h}}\right)\left(1+W\left(\frac{e^{v-b\varphi_h-1}}{1+e^{v_B-b\varphi_h}}\right)\right)^{-1}\quad\Rightarrow\quad Q_L y_h\left(1+W\left(\frac{e^{v-b\varphi_h-1}}{1+e^{v_B-b\varphi_h}}\right)\right)=W\left(\frac{e^{v-b\varphi_h-1}}{1+e^{v_B-b\varphi_h}}\right)$$

$$\Rightarrow\quad\frac{Q_L y_h}{1-Q_L y_h}=W\left(\frac{e^{v-b\varphi_h-1}}{1+e^{v_B-b\varphi_h}}\right)\qquad\Rightarrow\quad\frac{Q_L y_h}{1-Q_L y_h}e^{\frac{Q_L y_h}{1-Q_L y_h}}=\frac{e^{v-b\varphi_h-1}}{1+e^{v_B-b\varphi_h}}$$

$$\Rightarrow\quad\ln\left(\frac{Q_L y_h}{1-Q_L y_h}e^{\frac{Q_L y_h}{1-Q_L y_h}}\right)=v-b\varphi_h-1-\ln\left(1+e^{v_B-b\varphi_h}\right)$$

and thus, the market-clearing price at Q_L and realization (φ_h, y_h) satisfies

$$p^*\left(Q_L,\varphi_h,y_h\right)=\frac{1}{b}\left(v-\ln\left(\frac{Q_L y_h\left(1+e^{v_B-b\varphi_h}\right)}{1-Q_L y_h}\right)\right)=\varphi_h+\frac{1}{b}\left(\frac{1}{1-Q_L y_h}\right) \tag{32}$$

(see (2)). The unconstrained optimal stage-two decision $q^{\circ}\left(\varphi\right)=W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1}$

satisfies $q^{\circ}\left(\varphi\right)\in(0, 1)$ and $q^{\circ}{}'\left(\varphi\right)<0$. Therefore,

$$0 < Q_L y_l \leq Q_L y_h < 1 \tag{33}$$

$$p^{\circ}{}'\left(\varphi\right)>0 \tag{34}$$

where $p^{\circ}(\varphi)$ is the unconstrained optimal price given blend price φ. From (35), it follows that $\ln\left(\frac{Q_L y}{1-Q_L y}\right)$

exists, and it is straightforward to verify

$$\frac{\partial}{\partial y}\ln\left(\frac{Q_L y}{1-Q_L y}\right)>0 \tag{35}$$

$$\frac{\partial}{\partial\varphi}\ln\left(1+e^{v_B-b\varphi}\right)<0 \tag{36}$$

$$\frac{1+e^{v_B-b\varphi_h}}{1+e^{v_B-b\varphi_l}}\geq\frac{e^{v_B-b\varphi_h}}{e^{v_B-b\varphi_l}}. \tag{37}$$

Note that

$$p^*\left(Q_L,\varphi,y\right)=\frac{1}{b}\left(v-\ln\left(\frac{Q_L y}{1-Q_L y}\right)-\ln\left(1+e^{v_B-b\varphi}\right)\right)$$

$$Q_L\frac{\partial p^*\left(Q_L,\varphi,y\right)}{\partial Q}=Q_L\frac{\partial}{\partial Q}\frac{1}{b}\left(\ln\left(1-Q_L y\right)-\ln\left(Q_L y\right)\right)=-\frac{1}{b}\left(\frac{1}{1-Q_L y}\right).$$

From the above numbered expressions, it follows that for all $(\varphi, y)\in\Omega$,

$$p^*\left(Q_L,\varphi,y\right)\leq p^*\left(Q_L,\varphi_h,y_l\right) \tag{38}$$

$$Q_L\frac{\partial p^*\left(Q_L,\varphi,y\right)}{\partial Q}\leq Q_L\frac{\partial p^*\left(Q_L,\varphi,y_l\right)}{\partial Q}. \tag{39}$$

$$p^*\left(Q_L,\varphi,y\right)\geq p^*\left(Q_L,\varphi_l,y_h\right) \tag{40}$$

$$Q_L\frac{\partial p^*\left(Q_L,\varphi,y\right)}{\partial Q}\geq Q_L\frac{\partial p^*\left(Q_L,\varphi,y_h\right)}{\partial Q}. \tag{41}$$

Marginal revenue (with respect to contract acreage) at quantity Q_L can be expressed as

$$MR = E\left[\frac{\partial}{\partial Q}\left(p^*\left(Q_L,\tilde{\varphi},\tilde{y}\right)Q_L\tilde{y}\right)\right] = E\left[\tilde{y}\left(p^*\left(Q_L,\tilde{\varphi},\tilde{y}\right)+Q_L\frac{\partial p^*\left(Q_L,\tilde{\varphi},\tilde{y}\right)}{\partial Q}\right)\right]. \tag{42}$$

Part (*a*). Substituting upper bounds (40) and (41) into (44),

$$MR \le E\left[\tilde{y}\left(p^*\left(Q_L,\varphi_h,y_l\right)+Q_L\frac{\partial p^*\left(Q_L,\tilde{\varphi},y_l\right)}{\partial Q}\right)\right]$$

$$= E\left[\tilde{y}\frac{1}{b}\left(v-\ln\left(\frac{Q_L y_l}{1-Q_L y_l}\right)-\ln\left(1+e^{v_B-b\varphi_h}\right)-\left(\frac{1}{1-Q_L y_l}\right)\right)\right]$$

$$= \varphi_h+\frac{1}{b}\left(\ln\left(\frac{Q_L y_h}{1-Q_L y_h}\right)-\ln\left(\frac{Q_L y_l}{1-Q_L y_l}\right)+\frac{1}{1-Q_L y_h}-\frac{1}{1-Q_L y_l}\right) \qquad \text{(from (34))}$$

$$= \varphi_h+\frac{1}{b}\left(\ln\left(\left(\frac{y_h}{y_l}\right)\left(\frac{1-Q_L y_l}{1-Q_L y_h}\right)\right)+\frac{1}{1-Q_L y_h}-\frac{1}{1-Q_L y_l}\right)$$

Thus, (15) is a sufficient condition for inequality (13), which implies $Q^* \le Q_L$, i.e., a specialized sourcing strategy is optimal. If $\sigma_y^2 = 0$, then $y_l = y_h = 1$, and (15) simplifies to (16).

Part (*b*). Substituting lower bounds (42) and (43) into (44),

$$MR \ge E\left[\tilde{y}\left(p^*\left(Q_L,\varphi_l,y_h\right)+Q_L\frac{\partial p^*\left(Q_L,\tilde{\varphi},y_h\right)}{\partial Q}\right)\right]$$

$$= E\left[\tilde{y}\frac{1}{b}\left(v-\ln\left(\frac{Q_L y_h}{1-Q_L y_h}\right)-\ln\left(1+e^{v_B-b\varphi_l}\right)-\left(\frac{1}{1-Q_L y_h}\right)\right)\right] \tag{43}$$

$$= \varphi_h+\frac{1}{b}\ln\left(\frac{1+e^{v_B-b\varphi_h}}{1+e^{v_B-b\varphi_l}}\right) \qquad \text{(from (34))}$$

$$\ge \varphi_h+\frac{1}{b}\ln\left(\frac{e^{v_B-b\varphi_h}}{e^{v_B-b\varphi_l}}\right)=\varphi_l\,. \qquad \text{(from (39))}$$

Thus, (17) is a sufficient condition for inequality (13) to be reversed, which implies $Q^* > Q_L$, i.e., a diversified sourcing strategy is optimal. □

Proof of Proposition 4. Let

$$\Pi_c^* = \frac{Q^*}{b}E\left[\tilde{y}\left(\ln\left[\left(\frac{e^v}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\left(\frac{1-Q^*\tilde{y}}{Q^*\tilde{y}}\right)\right]\right)\middle|\left(\tilde{\varphi},\tilde{y}\right)\in\Omega_c\left(Q^*\right)\right]\Pr\left(\left(\tilde{\varphi},\tilde{y}\right)\in\Omega_c\left(Q^*\right)\right)$$

$$\Pi_u^* = \frac{1}{b}E\left[W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\middle|\left(\tilde{\varphi},\tilde{y}\right)\in\Omega_u\left(Q^*\right)\right]\Pr\left(\left(\tilde{\varphi},\tilde{y}\right)\in\Omega_u\left(Q^*\right)\right)$$

$$\Pi_0^* = E\left[(\tilde{\varphi}-c)Q^*\tilde{y}\right]-k\left(Q^*\right)=\left(\mu_\varphi+\rho\sigma_\varphi\sigma_y-c\right)Q^*-k\left(Q^*\right).$$

It follows from Proposition 1 and (4), that optimal profit can be expressed as

$$\Pi^* = \Pi_c^* + \Pi_u^* + \Pi_0^*.$$

Note that under specialized sourcing (SS), $\Omega_u\left(Q^*\right)=\varnothing$, and under diversified sourcing (DS), $\Omega_c\left(Q^*\right)=\varnothing$, and thus,

$$\text{SS: } \Pi^* = \Pi_c^* + \Pi_0^*$$

$$\text{DS: } \Pi^* = \Pi_u^* + \Pi_0^*.$$

The proposition presents comparative-statics results for optimal profit and for optimal acreage. In the following, we first prove results for optimal profit and then prove results for optimal acreage.

Impact of changes in σ_y on Π^*. Since concave function $\Pi(Q)$ is differentiable everywhere (because $k(Q)$ is differentiable everywhere), we apply the envelope theorem:

$$\frac{d\Pi^*}{d\sigma_y}=\frac{\partial\Pi\left(Q^*,\sigma_y\right)}{\partial\sigma_y}=\frac{\partial}{\partial\sigma_y}\left(\Pi_c\left(Q^*,\sigma_y\right)+\Pi_u\left(Q^*,\sigma_y\right)+\Pi_0\left(Q^*,\sigma_y\right)\right)$$

$$\frac{\partial\Pi_0\left(Q^*,\sigma_y\right)}{\partial\sigma_y}=\frac{\partial}{\partial\sigma_y}\left[\left(\mu_\varphi+\rho\sigma_\varphi\sigma_y-c\right)Q^*-k\left(Q^*\right)\right]=\rho\sigma_\varphi Q^*$$

$$\frac{\partial\Pi_u\left(Q^*,\sigma_y \mid Q^*\geq Q_H\right)}{\partial\sigma_y}=\frac{\partial}{\partial\sigma_y}\frac{1}{b}E\left[W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\right]=0$$

$$\frac{\partial\Pi_c\left(Q^*,\sigma_y \mid Q^*\geq Q_H\right)}{\partial\sigma_y}=0 \qquad \text{(because } \Omega_c\left(Q^*\right)=\varnothing\text{)}$$

$$\frac{\partial\Pi_u\left(Q^*,\sigma_y \mid Q^*\leq Q_L\right)}{\partial\sigma_y}=0 \qquad \text{(because } \Omega_u\left(Q^*\right)=\varnothing\text{)}$$

$$\frac{\partial\Pi_c\left(Q^*,\sigma_y \mid Q^*\leq Q_L\right)}{\partial\sigma_y}=\frac{\partial}{\partial\sigma_y}\frac{Q^*}{b}E\left[\tilde{y}\left(\ln\left[\left(\frac{e^{v}}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\left(\frac{1-Q^*\tilde{y}}{Q^*\tilde{y}}\right)\right]\right)\right]$$

$$=\frac{Q^*}{b}\frac{\partial}{\partial\sigma_y}E\left[\tilde{y}\left(v-\ln\left(e^{b\tilde{\varphi}}+e^{v_B}\right)-\ln\left(\frac{Q^*\tilde{y}}{1-Q^*\tilde{y}}\right)\right)\right]$$

$$=-\frac{Q^*}{b}\frac{\partial}{\partial\sigma_y}E\left[\tilde{y}\left(\ln\left(e^{b\tilde{\varphi}}+e^{v_B}\right)+\ln\left(\frac{Q^*\tilde{y}}{1-Q^*\tilde{y}}\right)\right)\right] \qquad \text{(because } \partial E\left[\tilde{y}\right]/\partial\sigma_y=0\text{)}$$

$$=-\frac{1}{b}\frac{\partial}{\partial\sigma_y}E\left[g_1\left(\tilde{\varphi},\tilde{y}\right)+g_2\left(\tilde{y}\right)\right]$$

$$g_1(\varphi, y) = Q^* y \ln\left(e^{b\varphi} + e^{v_B}\right)$$

$$g_2(y) = Q^* y \ln\left(\frac{Q^* y}{1 - Q^* y}\right).$$

Therefore, if DS is optimal, then

$$d\Pi^*_{Q^* \geq Q_H} / d\sigma_y = \partial\Pi_0\left(Q^*, \sigma_y \mid Q^* \geq Q_H\right) / \partial\sigma_y = \rho\sigma_\varphi Q^*, \tag{44}$$

which yields results (*aiii*) and (*b*) for Π^*.

Now suppose that SS is optimal. We first examine $\partial E\left[g_1(\tilde{\varphi}, \tilde{y})\right] / \partial\sigma_y$, which includes interaction among the random variables. However, g_1 is the product of two independent random variables, and thus

$$E\left[g_1(\tilde{\varphi}, \tilde{y})\right] = Q^* E[\tilde{y}] E\left[\ln\left(e^{b\tilde{\varphi}} + e^{v_B}\right)\right] = Q^* E\left[\ln\left(e^{b\tilde{\varphi}} + e^{v_B}\right)\right],$$

which implies $\partial E\left[g_1(\tilde{\varphi}, \tilde{y})\right] / \partial\sigma_y = 0$ and

$$\frac{\partial\Pi_c\left(Q^*, \sigma_y \mid Q^* \leq Q_L\right)}{\partial\sigma_y} = -\frac{1}{b}\frac{\partial}{\partial\sigma_y} E\left[g_2(\tilde{y})\right] = -\frac{Q^*}{b}\frac{\partial}{\partial\sigma_y} E\left[\tilde{y}\ln\left(\frac{Q^*\tilde{y}}{1 - Q^*\tilde{y}}\right)\right].$$

Function $-g_2(y)$ is strictly concave in y (and g_2 is strictly convex in y; see Proposition A4). Furthermore, recall that $\tilde{y}_2$ is a mean-preserving spread of $\tilde{y}_1$, i.e., $\int_{-\infty}^{x} F_{y_1}(y \mid \varphi)\,dy \leq \int_{-\infty}^{x} F_{y_2}(y \mid \varphi)\,dy$ for all x and for all $\varphi \in [\varphi_l, \varphi_h]$, and the inequality is strict for some x and φ. This implies that $\tilde{y}_1$ is smaller than $\tilde{y}_2$ in convex order (see Theorem 3.A.1 in Shaked and Shanthikumar 2007). By the definition of convex order, $E\left[g(\tilde{y}_1)\right] \leq E\left[g(\tilde{y}_2)\right]$, equivalently $E\left[-g(\tilde{y}_1)\right] \geq E\left[-g(\tilde{y}_2)\right]$, for all convex functions g (with strict inequality for strictly convex g), Therefore, $\frac{\partial}{\partial\sigma_y} E\left[g_2(\tilde{y})\right] > 0$ and

$$\frac{\partial\Pi_c\left(Q^*, \sigma_y \mid Q^* \leq Q_L\right)}{\partial\sigma_y} < 0. \tag{45}$$

Consolidating the above,

$$\begin{aligned} d\Pi^*_{Q^* \leq Q_L} / d\sigma_y &= \partial\Pi_c\left(Q^*, \sigma_y \mid Q^* \leq Q_L\right) / \partial\sigma_y + \partial\Pi_0\left(Q^*, \sigma_y \mid Q^* \leq Q_L\right) / \partial\sigma_y \\ &= \partial\Pi_c\left(Q^*, \sigma_y \mid Q^* \leq Q_L\right) / \partial\sigma_y + \rho\sigma_\varphi Q^* \\ &= \partial\Pi_c\left(Q^*, \sigma_y \mid Q^* \leq Q_L\right) / \partial\sigma_y < 0 \end{aligned} \tag{46}$$

This proves (*ai*) for Π^*.

Now suppose that mixed sourcing MS is optimal. Then $Q^* \in (Q_L, Q_H)$, and

$$\Pi^{*}=E\left[\pi^{*}\left(Q^{*},\tilde{\varphi},\tilde{y}\right)\right]=E\left[h(\tilde{\varphi},\tilde{y})+(\tilde{\varphi}-c)Qy-k(Q)\right]$$

where

$$h(\varphi,y)=\frac{Qy}{b}\left(v-\ln\left[\left(e^{b\varphi}+e^{v_B}\right)\left(\frac{Qy}{1-Qy}\right)\right]\right)\times\mathbf{1}_{q^*=Qy}+\frac{1}{b}W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\times\mathbf{1}_{q^*>Qy}$$

$$q^{*}=\min\left\{W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1},Qy\right\}$$

(see Proposition 1). As shown in the proof of Proposition A4, $h(\varphi,y)$ is concave in y. Thus, following the arguments given for the case of SS, we have

$$d\Pi^{*}_{Q^{*}\in(Q_L,Q_H)}/d\sigma_y<0,$$

which is result (*aii*) for Π^*.

Impact of changes in σ_y on Q^*. Since Π is concave in Q, we apply the implicit function theorem to obtain

$$Q^{*}{}'(\sigma_y)=\frac{\Pi_{Q\sigma_y}\left(Q^{*}\right)}{-\Pi_{QQ}\left(Q^{*}\right)}$$

(we use subscripts to denote argument in the partial derivative). Since $\Pi_{QQ}\left(Q^{*}\right)<0$, the sign of $Q^{*}{}'(\sigma_y)$ is determined by the sign of $\Pi_{Q\sigma_y}\left(Q^{*}\right)$.

As shown in the proof of Proposition 2,

$$\Pi_Q(Q)=\mu_\varphi+\rho\sigma_\varphi\sigma_y+E\left[\left(m_c(Q,\tilde{\varphi},\tilde{y})-m_u(Q,\tilde{y})\right)^{+}\right]-k'(Q)-c$$

$$=\mu_\varphi+\rho\sigma_\varphi\sigma_y+\frac{1}{b}E\left[\left(v\tilde{y}-\frac{g_1(\tilde{\varphi},\tilde{y})}{Q}-\frac{g_2(\tilde{y})}{Q}-g_3(\tilde{y})\right)^{+}\right]-k'(Q)-c \qquad (47)$$

where $g_3(y)=m_u(Q,y)=\frac{y}{b}\left(\frac{1}{1-Qy}\right)$ and g_1 and g_2 are defined as above, i.e.,

$$\frac{g_1(\varphi,y)}{Q}=y\ln\left(e^{b\varphi}+e^{v_B}\right)$$

$$\frac{g_2(y)}{Q}=y\ln\left(\frac{Qy}{1-Qy}\right).$$

Suppose $Q^*\geq Q_H$ (i.e., diversified sourcing is optimal). Then

$$E\left[\left(v\tilde{y}-\frac{g_1(\tilde{\varphi},\tilde{y})}{Q^*}-\frac{g_2(\tilde{y})}{Q^*}-g_3(\tilde{y})\right)^+\right]=0, \text{ and}$$

$$\Pi_Q(Q^*)=\mu_\varphi+\rho\sigma_\varphi\sigma_y-k'(Q)-c$$

(see Proposition 3). Thus,

$$\Pi_{Q\sigma_y}(Q^*)=\rho\sigma_\varphi,$$

which yields results (*aiii*) and (*b*) for Q^*.

Now suppose that $Q^* < Q_H$ (i.e., specialized or mixed sourcing is optimal). Then

$$E\left[\left(v\tilde{y}-\frac{g_1(\tilde{\varphi},\tilde{y})}{Q^*}-\frac{g_2(\tilde{y})}{Q^*}-g_3(\tilde{y})\right)^+\right]>0.$$

As shown above, $\partial E\left[g_1(\tilde{\varphi},\tilde{y})\right]/\partial\sigma_y=0$ (because blend price and yield are independent). Furthermore, as shown in the proof of Proposition A4, g_2 is convex in y for any realization for which

$$m(y)=vy-\frac{g_1(\varphi,y)}{Q^*}-\frac{g_2(y)}{Q^*}-g_3(y)>0.$$

Note also that

$$g_3''(y)=\frac{2Q^*}{\left(1-Q^*y\right)^3}>0 \text{ for all } Q^*y<1,$$

and $m(y) > 0$ if and only if the supply constraint it binding, which implies that the optimal second-stage quantity satisfies $q^* = Q^*y < 1$ (see Proposition 1). Therefore, $m(y)$ is concave in y for all y such that $m(y) > 0$.

Suppose specialized sourcing is optimal. Then $Q^* \le Q_L$, which implies

$$E\left[\left(m(\tilde{y})\right)^+\right]=E\left[m(\tilde{y})\right].$$

Recall from the first part of the proof that $\tilde{y}_1$ is smaller than $\tilde{y}_2$ in convex order, equivalently, $\tilde{y}_1$ is larger than $\tilde{y}_2$ in concave order, which implies

$$\Pi_Q(Q^*\mid\tilde{y}_2)=\mu_\varphi+\frac{1}{b}E\left[m(\tilde{y}_2)\right]-k'\left(Q^*_{\tilde{y}=\tilde{y}_2}\right)-c$$
$$\le\mu_\varphi+\frac{1}{b}E\left[m(\tilde{y}_1)\right]-k'\left(Q^*_{\tilde{y}=\tilde{y}_2}\right)-c\le\mu_\varphi+\frac{1}{b}E\left[m(\tilde{y}_1)\right]-k'\left(Q^*_{\tilde{y}=\tilde{y}_1}\right)-c=\Pi_Q\left(Q^*\mid\tilde{y}_1\right)$$

and thus $Q^*_{\tilde{y}=\tilde{y}_1}\ge Q^*_{\tilde{y}=\tilde{y}_2}$. This proves result (*ai*) for Q^*.

Suppose mixed sourcing is optimal. Then $Q^* \in (Q_L, Q_H)$, which implies

$$E\left[\left(m(\tilde{y})\right)^{+}\right] > E\left[m(\tilde{y})\right]$$

i.e., $\left(m(y)\right)^{+}$ is concave at some points and convex at some points (convexity derives from the max operator). Consequently, the effect of increasing σ_y on Q^* is indeterminate. This proves result (*aii*) for Q^*. □

5. Robustness of Sourcing Strategies to Other Demand Models

Proof of Proposition 5. Recall $q^{o}(\varphi) = \arg\max_q \{\pi(q \mid \varphi)\} = \arg\max_q \{\hat{\pi}(q \mid \varphi) = (p(q) - \varphi)q\}$, i.e., for analysis of the optimal second-stage solution, it is sufficient to maximize $\hat{\pi}(q \mid \varphi)$. We define

$$q^{o}(\varphi) = \inf\left\{\arg\max_q \{\hat{\pi}(q \mid \varphi)\}\right\}$$

(i.e., the *inf* operator assures unique $q^{o}(\varphi)$ if unimodal function $\hat{\pi}(q \mid \varphi)$ admits an interval of values at its mode). Then it follows that $\hat{\pi}_{qq}\left(q^{o}(\varphi) \mid \varphi\right) < 0$ for any $\varphi \in [\varphi_l, \varphi_h]$. We apply the implicit function theorem to obtain

$$q^{o}{}'(\varphi) = \frac{\hat{\pi}_{q\varphi}\left(q^{o}(\varphi) \mid \varphi\right)}{-\hat{\pi}_{qq}\left(q^{o}(\varphi) \mid \varphi\right)} = \frac{\frac{\partial}{\partial \varphi}\left(p(q) - \varphi + p'(q)q\Big|_{q=q^{o}(\varphi)}\right)}{-\hat{\pi}_{qq}\left(q^{o}(\varphi) \mid \varphi\right)} = \frac{-1}{-\hat{\pi}_{qq}\left(q^{o}(\varphi) \mid \varphi\right)} < 0.$$

Therefore,

$$\min_{\varphi, y}\left\{\frac{q^{o}(\varphi)}{y} : (\varphi, y) \in \Omega\right\} = \frac{q^{o}(\varphi_h)}{y_h} =: Q_L$$

$$\max_{\varphi, y}\left\{\frac{q^{o}(\varphi)}{y} : (\varphi, y) \in \Omega\right\} = \frac{q^{o}(\varphi_l)}{y_l} =: Q_H$$

and

$\Pr\left(q^*(Q, \tilde{\varphi}, \tilde{y}) < Q\tilde{y}\right) = 0$ if and only if $Q \leq Q_L$

$\Pr\left(q^*(Q, \tilde{\varphi}, \tilde{y}) < Q\tilde{y}\right) = 1$ if and only if $Q \geq Q_H > Q_L$.

Finally, since there are no restrictions on the first-stage cost function $k(Q) + cQ$ other than convexity, there exist cost functions for which the optimal first-stage solution ranges beyond the interval $[Q_L, Q_H]$. Therefore, $Q^* < Q_L$, $Q^* \in (Q_L, Q_H)$, and $Q^* > Q_H$ are possible, which implies that any one of the three strategies may emerge as optimal. □

Proof of Corollary 2. Recall that $A_k(\varphi) = e^{v_k}\left(1 + \alpha e^{v_k - b\varphi}\right)^{-1}$ and

$$\overline{\pi}(q_1 | \varphi) = (p(q_1) - \varphi) w_1 q_1 + \sum_{k=2}^{n} (p(q_1) - \varphi) w_k f_k(q_1).$$

Li et al. (2019) study a special case of $\alpha = 0$ (i.e., $A_k(\varphi) = e^{v_k}$) and prove that $\overline{\pi}(q_1 | \varphi, \alpha = 0)$ is concave if

$$\frac{\max_k e^{v_k}}{\min_k e^{v_k}} = \frac{\max_k A_k(\varphi | \alpha = 0)}{\min_k A_k(\varphi | \alpha = 0)} \leq 2,$$

While the value of $A_k(\varphi)$ differs from Li et al. (2019) due to $\alpha > 0$, equation (19) and corresponding profit function are structurally identical to the model in Proposition 1 in Li et al. (2019), and thus the analysis and result passes through, i.e., the value of α in the computation of A_k is immaterial. Thus, for given φ, $\overline{\pi}(q_1 | \varphi)$ is concave if $\max_k A_k(\varphi) / \min_k A_k(\varphi) \leq 2$.

Suppose $v_j \geq v_k$. Then $e^{v_j}\left(1 + \alpha e^{v_j - b\varphi}\right)^{-1} \geq e^{v_k}\left(1 + \alpha e^{v_k - b\varphi}\right)^{-1}$. Note that

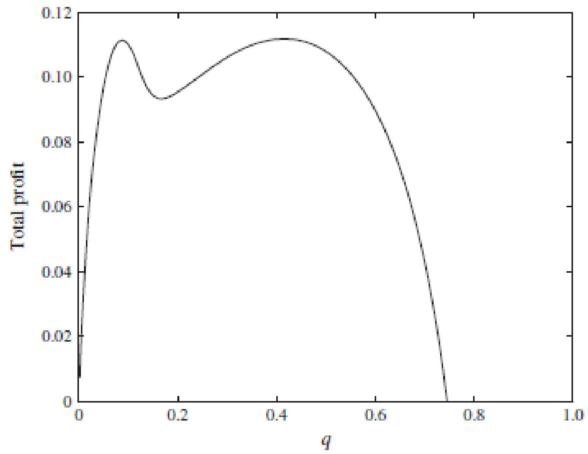
$$\frac{d}{d\varphi}\left(\frac{e^{v_j}\left(1 + \alpha e^{v_j - b\varphi}\right)^{-1}}{e^{v_k}\left(1 + \alpha e^{v_k - b\varphi}\right)^{-1}}\right) = \frac{e^{v_j}}{e^{v_k}} \frac{d}{d\left(e^{b\varphi}\right)}\left(\frac{e^{b\varphi} + \alpha e^{v_k}}{e^{b\varphi} + \alpha e^{v_j}}\right) \times \frac{de^{b\varphi}}{d\varphi} = \frac{e^{v_j}}{e^{v_k}}\left(\frac{\alpha e^{v_j} - \alpha e^{v_k}}{\left(e^{b\varphi} + \alpha e^{v_j}\right)^2}\right) be^{b\varphi} \geq 0.$$

Therefore

$$\max_{\varphi \in [\varphi_l, \varphi h]}\left\{\frac{\max_k A_k(\varphi)}{\min_k A_k(\varphi)}\right\} = \frac{\max_k A_k(\varphi_h)}{\min_k A_k(\varphi_h)} = \frac{\max_k \left\{e^{v_k}\left(1 + \alpha e^{v_k - b\varphi_h}\right)^{-1}\right\}}{\min_k \left\{e^{v_k}\left(1 + \alpha e^{v_k - b\varphi_h}\right)^{-1}\right\}}. \square$$

Temp content below for reference (to be deleted)

Figure 1. Profit by Market Share (Not Concave or Quasiconcave in Market Share)



Proposition 1. *For a single-product MMNL model, if* $\frac{\max_k A_{1k}}{\min_k A_{1k}} \leq 2$ *and* $b_{1k} = b$ *for all* k*, then* $\hat{\Pi}(\mathbf{q}^1)$ *is concave on* Ω_1*.*

Proof. To simplify presentation, we suppress the product subscript in our notation (e.g., q_k in place of q_{1k}). Note that the purchase probability of the product by a segment k customer is $\mathbf{q}^k = q_k$. Accordingly,

$$f_k(q_1) = \frac{A_k\left(\frac{q_1}{A_1(1-q_1)}\right)}{1 + A_k\left(\frac{q_1}{A_1(1-q_1)}\right)}$$

$$g_k(q_k) = \frac{1}{b}\log\left(\frac{A_k(1-q_k)}{q_k}\right)$$

and the profit contribution from a segment k customer as a function of q_1 simplifies to

$$\hat{R}_k(q_1) = (g_k(f_k(q_1)) - c)\, f_k(q_1) = \left(\frac{\log A_k - \log(\lambda_k q_1) + \log(1-q_1) - bc}{b}\right)\frac{\lambda_k q_1}{1 - q_1 + \lambda_k q_1}$$

where $\lambda_k = A_k/A_1$. We can derive that

$$-bz\frac{\partial^2 \hat{R}_k}{\partial q_1^2} = \frac{\lambda_k^2}{x} + \frac{2\lambda_k}{y} + \frac{x}{y^2} + 2L(\lambda_k - 1)^2\left(\frac{1}{z} - \frac{x}{z^2}\right) + \frac{2}{z}(\lambda_k - 1)(L - \lambda_k) - \frac{2x}{yz}(\lambda_k - 1)$$

where $x := \lambda_k q_1$, $y := q_{01}$, $z := \lambda_k q_1 + q_{01}$ and $L := \log A_k - \log(\lambda_k q_1) + \log q_{01} - bc$. Assume without loss of generality that $\lambda_k \geq 1$. From $\frac{\max_k A_k}{\min_k A_k} \leq 2$, we have $\lambda_k \leq 2$, equivalently, $\lambda_k \geq 2(\lambda_k - 1)$. Therefore,

$$\begin{aligned}
-bz\frac{\partial^2 R_k}{\partial {q_1}^2} &\geq \frac{\lambda_k^2}{x} + \frac{2\lambda_k}{y} + \frac{x}{y^2} + \frac{2}{z}L(\lambda_k - 1)^2 - \frac{2x}{z^2}L(\lambda_k - 1)^2 - \frac{2}{z}\lambda_k(\lambda_k - 1) - \frac{2x}{yz}(\lambda_k - 1) \\
&\geq \frac{\lambda_k^2}{x} + \frac{2\lambda_k}{y} + \frac{x}{y^2} - \frac{2}{z}\lambda_k(\lambda_k - 1) - \frac{2x}{yz}(\lambda_k - 1) \\
&= \lambda_k\left[\frac{\lambda_k}{x} - \frac{2}{z}(\lambda_k - 1)\right] + \frac{2}{y}\left[\lambda_k - \frac{x}{z}(\lambda_k - 1)\right] + \frac{x}{y^2} \\
&\geq \frac{x}{y^2} \geq 0
\end{aligned}$$

where the first inequality holds due to $L > 0$ (i.e., $L = b(p - c) > 0$), the second inequality holds because $\frac{x}{z} \leq 1$, and the third inequality holds due to $\lambda_k \geq 2(\lambda_k - 1)$ and $x \leq z$. Therefore, $\hat{R}_k$ is concave on Ω_1. The weighted sum of concave functions is concave, and thus $\hat{\Pi}$ is concave on Ω_1. □

NOTES FOR COMPUTATION OF FIGURES

We are plotting optimal profit and optimal contract acreage as yield uncertainty varies, for specialized, mixed, and diversified sourcing strategies. We identify values of $\kappa = k'(Q) + c$ where each strategy is optimal.[9] Relevant expressions for characterizing optimal profit and acreage are listed below.

$$\kappa = c + k'(Q)$$

$$Q_L = \frac{1}{y_h} W\left(\frac{e^{v-1}}{e^{b\varphi_h} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\varphi_h} + e^{v_B}}\right)\right)^{-1}$$

$$Q_H = \frac{1}{y_l} W\left(\frac{e^{v-1}}{e^{b\varphi_l} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\varphi_l} + e^{v_B}}\right)\right)^{-1}$$

$$Q_D = W\left(\frac{e^{v-1}}{e^{b\mu_\varphi} + e^{v_B}}\right)\left(1 + W\left(\frac{e^{v-1}}{e^{b\mu_\varphi} + e^{v_B}}\right)\right)^{-1}$$

$$m_c(Q,\varphi,y) = \frac{y}{b}\ln\left[\left(\frac{e^{v}}{e^{b\varphi} + e^{v_B}}\right)\left(\frac{1-Qy}{Qy}\right)\right] = \frac{y}{b}\left(v - \ln\left[\left(e^{b\varphi} + e^{v_B}\right)\left(\frac{Qy}{1-Qy}\right)\right]\right)$$

$$m_u(Q,y) = \frac{y}{b}\left(\frac{1}{1-Qy}\right)$$

$$\lambda(Q) = E\left[\left(m_c(Q,\tilde{\varphi},\tilde{y}) - m_u(Q,\tilde{\varphi},\tilde{y})\right)^{+}\right] = E\left[\frac{\tilde{y}}{b}\left(v - \ln\left(\frac{Q\tilde{y}\left(e^{b\tilde{\varphi}} + e^{v_B}\right)}{1-Q\tilde{y}}\right) - \frac{1}{1-Q\tilde{y}}\right)^{+}\right]$$

$$\lambda(Q) = E\left[m_c(Q,\tilde{\varphi},\tilde{y}) - m_u(Q,\tilde{\varphi},\tilde{y})\right] = \frac{1}{b}\left(v - E\left[\tilde{y}\ln\left(\frac{Q\tilde{y}\left(e^{b\tilde{\varphi}} + e^{v_B}\right)}{1-Q\tilde{y}}\right) + \frac{1}{1-Q\tilde{y}}\right]\right) \text{ if } Q \le Q_L$$

$$\lambda(Q) = E\left[\left(m_c(Q,\tilde{\varphi},\tilde{y}) - m_u(Q,\tilde{\varphi},\tilde{y})\right)^{+}\right]$$

$$= \frac{1}{b}E\left[\tilde{y}\left(v - \ln\left(\frac{Q\tilde{y}\left(e^{b\tilde{\varphi}} + e^{v_B}\right)}{1-Q\tilde{y}}\right) - \frac{1}{1-Q\tilde{y}}\right)^{+}\right] \text{ if } Q_L < Q < Q_H$$

$$\lambda(Q) = 0 \text{ if } Q \ge Q_H$$

Specialized Sourcing

Specialized sourcing is optimal if and only if

[9] The value of $k + c$ is an increasing step function, so the value of κ in the calibration is valid in the neighborhood of the relevant Q value, but can increase arbitrarily at some larger value of Q.

$$\mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} + \lambda(Q_L) \leq \kappa.$$

The optimal acreage is

$$Q^* = \max\left\{Q : \mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} + \lambda(Q) \geq \kappa\right\}.$$

The optimal profit is

$$\Pi^* = E\left[\pi^*\left(Q^*, \tilde{\varphi}, \tilde{y}\right)\right] = E\left[\frac{1}{b}\left(1 + \ln\left[\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}} + e^{v_B}}\right)\left(\frac{1 - Q^*\tilde{y}}{Q^*\tilde{y}}\right)\right]\right)Q^*\tilde{y} + (\tilde{\varphi} - c)Q^*\tilde{y} - k'\left(Q^*\right)\right]$$

$$= E\left[\frac{1}{b}\left(1 + \ln\left[\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}} + e^{v_B}}\right)\left(\frac{1 - Q^*\tilde{y}}{Q^*\tilde{y}}\right)\right]\right)Q^*\tilde{y} + Q^*\tilde{\varphi}\tilde{y}\right] - \kappa Q^*$$

$$= \left(E\left[\frac{\tilde{y}}{b}\left(v - \ln\left(\frac{\left(e^{b\tilde{\varphi}} + e^{v_B}\right)Q^*\tilde{y}}{1 - Q^*\tilde{y}}\right)\right)\right] + \mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} - \kappa\right)Q^*$$

$$= \left(E\left[m_c\left(Q^*, \tilde{\varphi}, \tilde{y}\right)\right] + \mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} - \kappa\right)Q^*$$

and the optimal profit per input unit is

$$\Pi^*/Q^* = E\left[m_c\left(Q^*, \tilde{\varphi}, \tilde{y}\right)\right] + \mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} - \kappa.$$

The marginal revenue is

$$\Pi^*_Q = \lambda\left(Q^*\right) + \mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} = E\left[m_c\left(Q^*, \tilde{\varphi}, \tilde{y}\right) - m_u\left(Q^*, \tilde{\varphi}, \tilde{y}\right)\right] + \mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y}$$

Diversified Sourcing

Diversified sourcing is optimal if and only if

$$\mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} \geq \kappa,$$

The optimal acreage is

$$Q^* = \max\left\{Q : \mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} \geq \kappa\right\},$$

i.e., implies $Q^* \geq Q_H$. The optimal profit is

$$\Pi^* = E\left[\pi^*\left(Q^*, \tilde{\varphi}, \tilde{y}\right)\right] = E\left[\frac{1}{b}W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}} + e^{v_B}}\right) + (\tilde{\varphi} - c)Q^*\tilde{y} - k'\left(Q^*\right)\right]$$

$$= E\left[\frac{1}{b}W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}} + e^{v_B}}\right) + Q^*\tilde{\varphi}\tilde{y}\right] - \kappa Q^*$$

$$= \left(E\left[\frac{1}{bQ^*}W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}} + e^{v_B}}\right)\right] + \mu_{\varphi} + \rho\sigma_{\varphi}\sigma_{y} - \kappa\right)Q^*$$

and the optimal profit per input unit is

$$\Pi^* = E\left[\frac{1}{bQ^*}W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\right] + \mu_\varphi + \rho\sigma_\varphi\sigma_y - \kappa\ .$$

The marginal revenue is

$$\Pi^*_Q = \lambda\left(Q^*\right) + \mu_\varphi + \rho\sigma_\varphi\sigma_y = \mu_\varphi + \rho\sigma_\varphi\sigma_y$$

Mixed Sourcing

A necessary and sufficient condition for mixed sourcing to be optimal is

$$\mu_\varphi + \rho\sigma_\varphi\sigma_y + \lambda\left(Q_L\right) > c + k'(Q_L) \text{ and } \mu_\varphi + \rho\sigma_\varphi\sigma_y < c + k'(Q_H).$$

Said differently, mixed sourcing is optimal if and only if

$$Q_L < Q^* = \max\left\{Q : \mu_\varphi + \rho\sigma_\varphi\sigma_y + \lambda\left(Q\right) \geq c + k'\left(Q\right)\right\} < Q_H.$$

In numerical analysis, we can identify a particular value for $Q^* \in (Q_L, Q_H)$, and compute the value of $\lambda(Q^*)$ at this value, e.g., we may select the midpoint $Q = (Q_L + Q_H)/2$. The second-stage optimal SL quantity is

$$q^* = \min\left\{W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\left(1+W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)\right)^{-1}, Qy\right\}$$

The optimal profit is

$$\begin{aligned}
\Pi^* &= E\left[\pi^*\left(Q^*,\tilde{\varphi},\tilde{y}\right)\right] \\
&= E\left[\frac{1}{b}\left(Q^*\tilde{y}\ln\left[\left(\frac{e^{v}}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\left(\frac{1-Q^*\tilde{y}}{Q^*\tilde{y}}\right)\right]\times\mathbf{1}_{q^*=Qy} + W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\times\mathbf{1}_{q^*>Qy}\right) + \left(\tilde{\varphi}-c\right)Q^*\tilde{y} - k'\left(Q^*\right)\right] \\
&= E\left[\left(Q^* m_c\left(Q^*,\tilde{\varphi},\tilde{y}\right)\times\mathbf{1}_{q^*=Q^*\tilde{y}} + \frac{1}{b}W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\times\mathbf{1}_{q^*<Q^*\tilde{y}}\right) + \left(\tilde{\varphi}-c\right)Q^*\tilde{y} - k'\left(Q^*\right)\right] \\
&= E\left[\left(m_c\left(Q^*,\tilde{\varphi},\tilde{y}\right)\times\mathbf{1}_{q^*=Q^*\tilde{y}} + \frac{1}{bQ^*}W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\times\mathbf{1}_{q^*<Q^*\tilde{y}}\right)Q^* + \left(\tilde{\varphi}-c\right)Q^*\tilde{y} - k'\left(Q^*\right)\right]
\end{aligned}$$

and the optimal profit per input unit is

$$\Pi^*/Q^* = E\left[m_c\left(Q^*,\tilde{\varphi},\tilde{y}\right)\times\mathbf{1}_{q^*=Q^*\tilde{y}} + \frac{1}{bQ^*}W\left(\frac{e^{v-1}}{e^{b\tilde{\varphi}}+e^{v_B}}\right)\times\mathbf{1}_{q^*<Q^*\tilde{y}}\right] + \mu_\varphi + \rho\sigma_\varphi\sigma_y - \kappa\ .$$

The marginal revenue is

$$\Pi^*_Q = \lambda\left(Q^*\right) + \mu_\varphi + \rho\sigma_\varphi\sigma_y = E\left[\left(m_c\left(Q^*,\tilde{\varphi},\tilde{y}\right) - m_u\left(Q^*,\tilde{\varphi},\tilde{y}\right)\right)^+\right] + \mu_\varphi + \rho\sigma_\varphi\sigma_y$$

OPTIMAL PROFIT ANALYSIS, SPECIALIZED STRATEGY

$$\Pi(Q^*) = E\left[\left(p^*\left(Q^*,\tilde{\varphi},\tilde{y}\right)-c\right)\tilde{y}-k\right]Q^* = \left(E\left[\tilde{y}p^*\left(Q^*,\tilde{\varphi},\tilde{y}\right)\right]-c-k\right)Q^*$$

$$MR = E\left[\frac{\partial}{\partial Q}\left(p^*\left(Q^*,\tilde{\varphi},\tilde{y}\right)Q^*\tilde{y}\right)\right] = E\left[\tilde{y}p^*\left(Q^*,\tilde{\varphi},\tilde{y}\right)\right] + E\left[\tilde{y}Q^*\frac{\partial p^*\left(Q^*,\tilde{\varphi},\tilde{y}\right)}{\partial Q}\right] = c+k$$

$$E\left[\tilde{y}Q^*\frac{\partial p^*\left(Q^*,\tilde{\varphi},\tilde{y}\right)}{\partial Q}\right] = -E\left[\tilde{y}\frac{1}{b}\left(\frac{1}{1-Q^*\tilde{y}}\right)\right] \text{ (given optimal specialized strategy)}$$

$$\Rightarrow \quad E\left[\tilde{y}p^*\left(Q^*,\tilde{\varphi},\tilde{y}\right)\right] = c+k+E\left[\tilde{y}\frac{1}{b}\left(\frac{1}{1-Q^*\tilde{y}}\right)\right] = \text{optimal expected price per acre}$$

$$\Rightarrow \quad \Pi(Q^*) = \left(c+k+E\left[\tilde{y}\frac{1}{b}\left(\frac{1}{1-Q^*\tilde{y}}\right)\right]-c-k\right)Q^* = \frac{1}{b}E\left[\frac{\tilde{y}}{1-Q^*\tilde{y}}\right]Q^* = \frac{1}{b}E\left[\frac{1}{1/\tilde{y}-Q^*}\right]Q^*$$

$$= \frac{1}{b}E\left[\frac{1}{\frac{1}{Q^*\tilde{y}}-1}\right] = \frac{1}{b}E\left[\frac{Q^*\tilde{y}}{1-Q^*\tilde{y}}\right] = \frac{1}{b}\left(1+E\left[\frac{1}{1-Q^*\tilde{y}}\right]\right) \qquad \text{(shows up in multiple places)}$$

I may want to introduce the above into a proposition.

$$g(y) = \left(y^{-1}-Q^*\right)^{-1}$$

$$g'(y) = -\left(y^{-1}-Q^*\right)^{-2}y^{-2} = -\frac{1}{\left[y\left(y^{-1}-Q^*\right)\right]^2} = -\frac{1}{\left(1-Q^*y\right)^2}$$

$$g''(y) = -\frac{Q^*}{\left(1-Q^*y\right)^3} < 0$$

Taking the expected value of a concave function, e.g., adding yield variance puts downward pressure profit. However, Q^* changes as variance changes as well.

OPTIMAL PROFIT ANALYSIS, DIVERSIFIED STRATEGY

Should try to derive a similar simple profit expression under diversified strategy. The changes is that the price derivative has different forms depending on the realization. However, price doesn't depend on Q_1 when nonbinding, so this piece disappears. Will probably end up with the following:

$$\Pi\left(Q_1^*\right) = \frac{1}{b}\left(1+E\left[\frac{1}{1-Q_1^*\tilde{y}_1} \mid q_1^*\left(Q_1^*,\varphi,y_1\right)=Q_1^*\tilde{y}_1\right]\Pr\left(q_1^*\left(Q_1^*,\varphi,y_1\right)=Q_1^*\tilde{y}_1\right)\right).$$

So, while Q_1^* is larger under a diversified strategy (all else equal) the probability is smaller, which is 1 for specialized.

ANALYSIS RELATED TO PROPOSITION 4

$$\pi^*(Q,\varphi,y)=\begin{cases}\left(\frac{1}{b}\left(v-\ln\left(\frac{Qy\left(1+e^{v_B-b\varphi}\right)}{1-Qy}\right)\right)-c\right)Qy-kQ, \text{ if } q^*=Qy\\ \frac{1}{b}W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)+(\varphi-c)Qy-kQ, \qquad \text{if } q^*<Qy\end{cases}$$

$$=\begin{cases}\left(\frac{1}{b}\left(v-\ln\left(\frac{Qy}{1-Qy}\right)\right)-\ln\left(e^{b\varphi}+e^{v_B}\right)\right)Qy+(\varphi-c)Qy-kQ, \text{ if } q^*=Qy\\ \frac{1}{b}W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)+(\varphi-c)Qy-kQ, \qquad \text{if } q^*<Qy\end{cases}$$

The first term in each of the two lines can be roughly interpreted as the profit lift in the second stage from single-origin (these terms would be zero if all single-origin was sold as blend. When capacity is not binding, then yield variance has no effect on the profit lift. A guess is that the lift function in the first line is concave in y, which would suggest that increase yield variance has a negative effect with the supply constraint is binding, but not so when not binding (e.g., a distinct difference).

<u>Constraint not binding in the second stage</u>

For a given Q, when the constraint is not binding, <u>profit is linear in yield</u>, so increasing yield variance does not affect profit (keeping Q fixed). Note that $W(z)$ is concave in z. Let $z(\varphi) = [v + e^{b\varphi}]^{-1}$.

$$z'(\varphi)=\frac{-be^{b\varphi}}{\left(v+e^{b\varphi}\right)^2}<0$$

$$z''(\varphi)=\frac{-b^2e^{b\varphi}}{\left(v+e^{b\varphi}\right)^2}+\frac{2b^2e^{2b\varphi}}{\left(v+e^{b\varphi}\right)^3}=\frac{2b^2e^{2b\varphi}-b^2e^{b\varphi}\left(v+e^{b\varphi}\right)}{\left(v+e^{b\varphi}\right)^3}=\frac{b^2\left(e^{2b\varphi}-ve^{b\varphi}\right)}{\left(v+e^{b\varphi}\right)^3} \qquad \text{(unknown sign)}$$

In this case $b = e^{vB}$. We can't say much about whether $W\left(\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}\right)$ is convex or concave in φ. The term

$\frac{e^{v-1}}{e^{b\varphi}+e^{v_B}}$ is $e^{-b\varphi}$ times the market share is sold at price $\varphi + 1/b$.

<u>Constraint binding in the second stage</u>

Given $Q_1^* \leq Q_1^L$:

$$\Pi\left(Q_1^*\right)=E\left[\pi^*\left(Q_1^*,\tilde{\varphi},\tilde{y}_1 \mid q_1^*=Q_1^*\tilde{y}_1\right)\right]$$

$$=E\left[\left(\frac{1}{b}\left(v_1-\ln\left(\frac{Q_1^*\tilde{y}_1\left(1+e^{v_B-b\tilde{\varphi}}\right)}{1-Q_1^*\tilde{y}_1}\right)\right)-\tilde{\varphi}\right)Q_1^*\tilde{y}_1+(\tilde{\varphi}-c_1)Q_1^*\tilde{y}_1\right]-k_1Q_1^*$$

$$=\left(\frac{v_1}{b}-c_1-k_1-\frac{1}{b}E\left[\left(\ln\left(\frac{Q_1^*\tilde{y}_1}{1-Q_1^*\tilde{y}_1}\right)+\ln\left(1+e^{v_B-b\tilde{\varphi}}\right)\right)\tilde{y}_1\right]\right)Q_1^*$$

$$=\left(\frac{v_1}{b}-c_1-k_1+\frac{1}{b}E\left[\left(b\tilde{\varphi}-\ln\left(\frac{Q_1^*\tilde{y}_1}{1-Q_1^*\tilde{y}_1}\right)-\ln\left(e^{b\tilde{\varphi}}+e^{v_B}\right)\right)\tilde{y}_1\right]\right)Q_1^*$$

$$=\left(E[\tilde{\varphi}\tilde{y}_1]-c_1-k_1+\frac{1}{b}\left(v_1-E\left[\left(\ln\left(\frac{Q_1^*\tilde{y}_1}{1-Q_1^*\tilde{y}_1}\right)+\ln\left(e^{b\tilde{\varphi}}+e^{v_B}\right)\right)\tilde{y}_1\right]\right)\right)Q_1^*$$

The second expected value can be interpreted as the lift in price per acre over blend. The function $\ln\left(e^{b\tilde{\varphi}}+e^{v_B}\right)$ is convex in φ, so increasing variance in φ reduces profit. Will be partially offset by reoptimizing, but profit is still dropping. Covariance is the product of standard deviations and correlation, so $E[\tilde{\varphi}\tilde{y}_1]$ also decreases as variance in φ increases (keeping negative correlation fixed). It seems likely that $y_1\ln\left(\frac{Q_1^*y_1}{1-Q_1^*y_1}\right)$ is generally convex in y_1, but not clear (e.g., profit harmed by increasing yield variance).

$$MR=E\left[\frac{\partial}{\partial Q_1}\left(p_1^*\left(Q_1^*,\tilde{\varphi},\tilde{y}_1\right)Q_1^*\tilde{y}_1\right)\right]=E\left[\tilde{y}_1\left(p_1^*\left(Q_1^*,\tilde{\varphi},\tilde{y}_1\right)+Q_1^*\frac{\partial p_1^*\left(Q_1^*,\tilde{\varphi},\tilde{y}_1\right)}{\partial Q_1}\right)\right]$$

$$=\frac{1}{b}\left(v_1-E\left[\tilde{y}_1\left(\ln\left(\frac{Q_1^*\tilde{y}_1}{1-Q_1^*\tilde{y}_1}\right)+\ln\left(1+e^{v_B-b\tilde{\varphi}}\right)+\frac{1}{1-Q_1^*\tilde{y}_1}\right)\right]\right)$$

$$E[\tilde{\varphi}\tilde{y}_1]=\mu_\varphi+\sigma_{\varphi y_1}$$

If v_B is large, then $1+e^{v_B-b\tilde{\varphi}}\approx e^{v_B-b\tilde{\varphi}}$ and

$$\Pi\left(Q_1^*\right)\approx\left(\frac{v_1}{b}-c_1\right)Q_1^*-k_1Q_1-\frac{Q_1^*}{b}E\left[\left(\ln\left(\frac{Q_1^*\tilde{y}_1}{1-Q_1^*\tilde{y}_1}\right)+\ln\left(e^{v_B-b\tilde{\varphi}}\right)\right)\tilde{y}_1\right]$$

$$=\left(E[\tilde{\varphi}\tilde{y}_1]+\frac{v_1}{b}-c_1-k_1-\frac{1}{b}E\left[\tilde{y}_1\ln\left(\frac{Q_1^*\tilde{y}_1}{1-Q_1^*\tilde{y}_1}\right)\right]\right)Q_1^*$$

$$MR\approx E[\tilde{\varphi}\tilde{y}_1]+\frac{1}{b}\left(v_1-E\left[\tilde{y}_1\left(\ln\left(\frac{Q_1^*\tilde{y}_1}{1-Q_1^*\tilde{y}_1}\right)+\frac{1}{1-Q_1^*\tilde{y}_1}\right)\right]\right)$$