European Review of Agricultural Economics Vol **00 (00)** (2021) pp. 1–36
doi:10.1093/erae/jbab001

Competition, price dispersion and capacity constraints: the case of the U.S. corn seed industry

Cornelia Ilin[†,*] and Guanming Shi[‡]

[†]School of Information, UC Berkeley, USA; [‡]Department of Agricultural and Applied Economics, UW-Madison, USA

Received July 2020; final version accepted January 2021

Abstract

In this paper we examine the effect of competition on price dispersion and argue that the effect is contingent on the ability of firms to meet market demand. Our comparative static results show that competition among symmetrically capacity-unconstrained firms, or among firms with asymmetric capacities leads to an overall price increase along the distribution function. To investigate these findings empirically, we use a novel data set from the U.S. corn seed industry with firm and farm level sales information for conventional and genetically modified corn seeds between 2004 and 2009. We estimate the empirical model using the Fixed Effect Instrumental Variable Quantile Regression and find evidence consistent with the theory. The analysis also shows that capacity-unconstrained seed firms charge a price premium, confirming the positive relationship between product availability and pricing found in our theoretical model.

Keywords: Market Structure, Capacity Constraints, Consumer Loyalty, Price Dispersion

JEL classification: L11, L13, L66

1. Introduction

Economists have proposed theoretical models that predict price dispersion as an equilibrium outcome. Examples include search-theoretic models (e.g. Stigler, 1961; Rothschild, 1973; Reinganum, 1979; MacMinn, 1980) and clearinghouse models (e.g. Varian, 1980; Salop and Stiglitz, 1977; Shilony, 1977; Rosenthal, 1980; Narasimhan, 1988; Spulber, 1995; Baye and Morgan, 2001; Baye, Morgan and Scholten, 2006). In search-theoretic models, price dispersion arises from the marginal search cost consumers pay to obtain an additional price quotation (e.g. Stigler, 1961; Rothschild, 1973; Reinganum, 1979; MacMinn, 1980). In clearinghouse models, an information clearinghouse provides price information; thus, consumer search costs are zero.

*Corresponding author: E-mail: cilin@ischool.berkeley.edu

In this case, price dispersion arises due to differences in consumers' decision to access the clearinghouse (e.g. Varian, 1980; Salop and Stiglitz, 1977; Shilony, 1977; Rosenthal, 1980), or from firm heterogeneities attributable to asymmetric consumer loyalty (e.g. Narasimhan, 1988) and/or asymmetric production costs (e.g. Spulber, 1995).

Empirical tests of these models almost always assume that firms have an unlimited capacity to supply a homogeneous product. However, many markets contain *capacity-constrained* firms. In such markets, firms are facing supply constraints (i.e. there is an upper limit of the amount of production in any given period) and may become *capacity-constrained* when the demand is greater than the capacity that the firm can supply. Such markets are not uncommon in the real world. For example, airline companies cannot increase the number of seats in a given plane (unless redesigning and producing a new aircraft); hotel companies are not able to increase the number of guest rooms in a hotel building (unless remodelling or adding extensions to the existing building) and seed companies cannot increase the number of seeds received from the contracted seed farmers (which were produced in the previous season). Besides, customer brand loyalty is also widely observed in such capacity-constrained marketplaces. Consumer loyalty may be linked to network effects (such as the reward point programs in the airline and hospitality businesses) or consumers' heterogeneous tastes (such as utility or disutility from shopping around).

In this study, we aim to offer a description of the role firm capacity constraints and consumer brand loyalty play in determining the price equilibrium. In the theoretical part, we build on previous literature on price dispersion and clearinghouse models of price dispersion when (i) firms have asymmetric capacity constraints to supply a homogeneous product and (ii) consumers are heterogeneous in preference (loyal vs. price sensitive). A clearinghouse model may be more appropriate than a search-theoretic model for such industries because their price dispersion tends to be 'temporal': firms may charge different prices for the same product at different times and their position in the distribution of prices could change over time. As a result, the equilibrium is characterised by firms playing mixed strategies in prices. We show that our model generates temporal price dispersion due to such exogenous consumer heterogeneity and firm heterogeneity. In addition, our clearinghouse model provides a direct interpretation of the effect of competition on price dispersion: the equilibrium distribution of price depends on the number of firms in the market.

In the empirical analysis, we examine the price dispersion in the U.S. corn seed market. In the agribusiness industry, seed production decisions are made at least one season before marketing, and the quantity produced depends on demand forecasting, the number of contracted seed farmers, land availability, weather and other natural events that may interfere with the growing season. Because of time lags in seed production, the capacity constraint may become binding if a seed firm sells all its products before the end of the marketing season, thus becoming capacity constrained. Farmers, especially those with brand loyalty, may need to decide whether to order early to ensure seed availability, or to shop around for a better deal, but bearing the risk of their preferred products being sold out. These circumstances require the market price to re-equilibrate constantly, leading to a distribution of prices as the optimal solution.

We test our model's predictions relating to the effect of market concentration on price dispersion and the relationship between product availability and pricing, using the fixed effect instrumental variable quantile regression (FEIVQR). We use a novel data set from the U.S. corn seed industry, which provides firm- and farm-level purchase information for conventional and genetically modified corn seeds sold by different firms between 2004 and 2009. The corn seed industry has experienced considerable structural changes since the 1990s, following biotechnological breakthroughs aimed at improving agricultural productivity *via* seed genetic modification, and a series of horizontal mergers between seed firms and vertical mergers between seed and biotechnology firms. Five large biotechnology firms[1] have dominated the corn seed market. This provides an opportunity to empirically test our theory and document the effect of industry consolidation.

Our empirical results confirm the predictions of our theoretical clearinghouse model. An increase in market concentration among symmetrically capacity-unconstrained seed firms leads to price increases at each quantile of the price distribution. Similar results are obtained when competition is between seed firms with asymmetric capacities. Additionally, we find that corn farmers may trade price for product availability, allowing capacity-unconstrained firms to charge a price premium at each quantile of the distribution function.

The remainder of the article is organised as follows: Section 2 characterises the equilibrium in the two-stage clearinghouse model and simulation predictions. Section 3 introduces data relating to the U.S. corn seed industry and descriptive statistics. The econometric model of price dispersion and estimation method are presented in Section 4. We discuss empirical findings in Sections 5 and 6. Section 7 is the conclusion.

2. The model

Our theoretical analysis begins with a two-stage game where firms choose capacities in the first stage following the capacity-constraints model of Davidson and Denekere (1986) and choose prices in the second stage following Varian (1980)'s model of sales and Narasimhan (1988)'s model of asymmetric customer base. We extend the literature by combining the scale of operation decisions on the supply side, with the brand loyalty of consumers on the demand side. Such an extension may be considered marginal theoretically, yet valuable empirically for modelling the actual marketplace interactions where both capacity constraints and customer brand loyalty exist. In this study, we will focus our analysis on firms' decisions on pricing in the second stage, taking stage one decisions on capacity as given.

1 These are: Monsanto, DuPont, Syngenta, Dow Agrosciences and AgReliant.

2.1. The two-stage game

On the supply side, we assume N price-setting firms, which face capacity constraints, but the constraint can be binding (capacity-constrained) or not binding (capacity-unconstrained). On the demand side, a continuum of consumers have unit demand for the product and a maximum willingness to pay, $r > 0$. The market is equipped with a clearinghouse that provides price information. We assume that firms and consumers can list or access price information at minimal or no cost (e.g. *via* newspaper or internet search engine). There are two consumer types: (i) *price-sensitive shoppers* who access the clearinghouse for information about prevailing market prices. They buy product with the lowest listed price, and if no product is listed they will visit a firm randomly; and (ii) *loyal consumers* who always purchase a particular brand regardless of price if less than the reservation value. They do not access the clearinghouse, and they stay uninformed about other firms' pricing practices. Each firm has its own loyal customer base, and we assume a loyal customer will be loyal to only one firm. Let $L_i \geq 0$ denote the number of loyal consumers loyal to firm i.

The timing of the game is as follows. In the first stage, firms simultaneously and independently choose a capacity level $K_i \geq 0, i = 1, 2, \ldots N$. Note that when $K_i = 0$, the firm chooses not to enter the market. We assume that each firm will be able to serve its own loyal customers, $K_i > L_i$. The cost to install capacity K_i is normalised to zero. After the first stage, capacity decisions are common knowledge to all firms. In the second stage, firms simultaneously and independently choose a distribution of prices following the probability density function $f(p_i)$ and supply the demand they face, q_i, for price p_i, at total cost $c(q_i)$.[2]

Let the total market demand $D(p)$ given a single market price p be defined as follows:

$D(p) = D_s(p) + \sum_i D_{L_i}$, where $D_s(p) > 0$ is the total demand of price-sensitive shoppers and $D_{L_i} \geq 0$ is the demand of consumers loyal to firm i.

We examine how total market demand is allocated among firms. We assume that each firm first serves its loyal customer base. Then, when the lower-priced firm *cannot* meet the entire demand of price-sensitive shoppers, the sales of the remaining firms occur following Beckmann (1965)'s contingent demand rationing.[3] With this specification, all price-sensitive shoppers have the same probability of being rationed by a non-lowest-priced firm.[4] Kreps and Scheinkman (1983) show that a two-stage game with capacity choices

2 We assume that firms' production functions exhibit decreasing returns to scale, thus $c(q_i)$ is increasing and convex: $c'(q_i) > 0$, $c''(q_i) \geq 0$ for $q_i > 0$.

3 Also called the proportional-rationing rule.

4 Note that loyal consumers are not rationed, they purchase only from the company they are loyal to. The demand of loyal consumers is independent of the price level, provided it is below the reservation price, r.

and Bertrand price competition is the same as the one-stage Cournot outcome. However, Davidson and Deneckere (1986) show that only the efficient rationing rule can provide Kreps and Scheinkman results. They argue that just by imposing the Beckman (1965) rationing rule, the equilibrium is a distribution of prices that results from the mixed strategies firms play. Kreps and Scheinkman (1983)'s efficient rationing rule and Beckmann (1965)'s contingent demand rationing occupy the two opposite extremes. We chose the latter in our study to better fit the empirical setting in the U.S. corn seed industry: farmers arrive at the firm's 'shop' at random and are allowed to make unlimited purchases on a first come, first served basis (as long as capacity is available). Those that arrive late may not be able to purchase any product. This seems to be the typical behaviour we observe in the seed industry.

Assuming firm j charges the lowest price in the market, the probability of not being served by firm j is:

$$\begin{cases} \left[1-\frac{K_j-L_j}{D_s(p_j)}\right], & \textit{if } K_j-L_j \le D_s(p_j), \text{firm } j \text{ is capacity} - \text{constrained;} \\ 0, & \text{otherwise.} \end{cases}$$

Then, the residual demand faced by any firm i charging $p_i > p_j$ is given by:

$$\begin{cases} D_s(p_i)\left[1-\frac{K_j-L_j}{D_s(p_j)}\right]\left(\frac{K_i-L_i}{\sum_{r\neq j}(K_r-L_r)}\right), & \textit{if } K_j-L_j \le D_s(p_j); \\ 0, & \text{otherwise,} \end{cases}$$

implying that residual demand is shared among all firms charging a price greater than p_j. The sharing proportion depends on each firm's relative residual capacity after serving their loyal customers. To simplify notation, let $x_i = \left(1-\frac{K_j-L_j}{D_s(p_j)}\right)\frac{K_i-L_i}{\sum_{r\neq j}(K_r-L_r)}$. We then formulate the demand for firm i's product as a function of its own price p_i and for any given value of competitors' prices p_{-i} as follows:

$$D(p_i|p_{-i}) = \begin{cases} \min\left[K_i, L_i+D_s(p_i)\right], & p_i<p_{-i} \\ \min\left[K_i, \max\left(L_i+\frac{D_s(p_i)}{n}\right), L_i \right. & \\ \left. \qquad +D_s(p_i)-\sum_{-i}(K_{-i}-L_{-i})\right], & p_i<p_{-i} \\ \min\left[K_i, L_i+D_s(p_i)\cdot x_i\right], & p_i>p_j. \end{cases}$$

Finally, each firm's strategy space is continuous, ranging from the firm's corresponding average cost, denoted as p_i^* as it is also the firm's reservation selling price, to the consumer's reservation buying price, r. The maximum number of consumers a firm can serve is $\min\left[K_i, L_i+D_s(p_i)\right]$. Then p_i^* is defined as: $p_i^* = \frac{c\left(\min\left[K_i, L_i+D_s(p_i)\right]\right)}{\min\left[K_i, L_i+D_s(p_i)\right]}$. Next, we focus on the second stage of the game and compute the equilibrium prices for given capacity levels.

2.2. The price subgame

Let K_i, K_{-i} be capacities chosen in the first stage by firm i and its competitors $-i$. Let L_i, L_{-i} be their corresponding loyal customer base. Assume without loss of generality that firms may be symmetric or asymmetric in either feature.[5] For any price p, we assume that the market demand function $D(p) : \mathbb{R}_{++} \to \mathbb{R}_{++}$ is differentiable and strictly decreasing in p. We further assume that the market demand function $D(p) = 0$ if $p > r$ and $D(p) > 0$ if $p < r$, and $\lim_{p \to 0} D(p) = +\infty$ and $\lim_{p \to \infty} D(p) = 0$.

The market revenue function, $p * D(p) : \mathbb{R}_{++} \to \mathbb{R}_{++}$ is single peaked and attains a unique maximum at the consumer's reservation price, r, and is strictly concave in p for $p < r$. We also assume a quantity q^*such that average cost p_i^* takes on the minimum value: for $q \leq q^*$, $\frac{\partial p_i^*(q)}{\partial q} \leq 0$, and for $q \geq q^*$, $\frac{\partial p_i^*(q)}{\partial q}$ ≥ 0. The equilibrium pricing strategy is static.[6]

Since capacity decisions (including entry decisions) are already made in stage one, the number of firms, N, is considered fixed in stage two.

Theorem 1. *For each pair (K_i, K_{-i}) and (L_i, L_{-i}), there is a unique set of Nash equilibrium in prices:*

(1) If all firms are capacity-constrained, then the equilibrium is a pure strategy with each firm charging the consumer's reservation price r when $\mathrm{K_i} + \sum_{-i} \mathrm{K_{-i}} \leq \mathrm{D}(\mathrm{p}^*)$, and a mixed strategy otherwise.
(2) If all firms are capacity-unconstrained, the equilibrium is a mixed strategy.
(3) If there is a mix of capacity-constrained and capacity-unconstrained firms, then the equilibrium is a mixed strategy.

Proof. See Appendix A (Appendix in supplementary data at *ERAE* online). □

Given the predictions stated in Theorem 1, we now establish a mixed strategy pricing equilibrium for firm i. Let $f(p_i)$ denote the probability density function for firm i's price p_i. In each time period, the firm randomly draws a price out of $f(p_i)$.

When firm i happens to draw the lowest price in the market, the event is considered a **win (*w*)** and the number of consumers being served by firm i is: $\min[K_i, L_i + D_s(p_i)]$.

When firm i fails to draw the lowest price, the event is considered a **loss (*l*)** and the number of consumers the firm will serve is: $x_i \cdot \min[K_i, L_i + D_s(p_i)]$.

5 When $K_i = K_{-i}$ and $L_i = L_{-i}$, the capacity level and the loyal customer base are symmetric. When $K_i \neq K_{-i}$ and $L_i \neq L_{-i}$, they are asymmetric.

6 The static equilibrium assumes a repeated game with infinite horizon. However, history (past prices) does not matter in the firms' equilibrium actions.

When firm i and one or more other firms draw the same lowest price, the event is considered a **tie** and each lowest-priced firm gets:

$$\min\left[K_i, \max\left(L_i + \text{ } \textit{equal share of } D_s(p_i) \text{ } \textit{consumers}, \text{ } L_i + D_s(p_i) - \sum_{-i}(K_{-i} - L_{-i})\right)\right].$$

Proposition 1. *The equilibrium pricing strategy has a continuous probability distribution.*

Proof. Similar to Varian (1980), assume that at the equilibrium firm i may charge some $\tilde{p}_i$ with positive probability.[7] Then, given a tie at $\tilde{p}_i$, firm i can deviate and charge a lower price $(\tilde{p}_i - \varepsilon)$. Firm i will trade an ε portion of its existing profit margin for additional profits from sales attracted from its tied competitors. This outcome is contradictory to the equilibrium concept. Therefore, in the equilibrium, there is no mass point along the price density function. The equilibrium pricing strategy has a continuous probability distribution. □

Let $F(p_i)$ denote firm i's cumulative distribution function, which is continuous on $[p^*, r]$.

Then, for p_i, the expected profit of firm i is given by:

$$\int_{p^*}^{r}\left\{\Pi_w(p_i)\cdot[1-F(p_i)]^{N-1} + \Pi_l(p_i)\cdot\left[1-(1-F(p_i))^{N-1}\right]\right\}f(p_i)\,dp_i. \tag{1}$$

where $\Pi_w(p_i) = \{(p_i - c)\cdot\min[K_i, L_i + D_s(p_i)]\}$,
and $\Pi_l(p_i) = \{(p_i - c)\cdot\min[K_i, L_i + D_s(p_i)\cdot x_i]\}$.

In equation (1), $[1-F(p_i)]^{N-1}$ is the probability that firm i charges the lowest price among the N firms, and $\left[1-[1-F(p_i)]^{N-1}\right]$ is the probability that there is at least one other firm with a lower price than firm i.

The objective of firm i is to maximise expected profits (as shown in equation (1)) by choosing the density function $f(p_i)$ subject to the constraints: $f(p_i) \geq 0$ and $\int_{p^*}^{r} f(p_i)\,dp_i = 1$, given the strategies of other firms and consumer behaviour. A mixed pricing strategy is a Nash equilibrium if and only if all prices charged with positive probability density $(f(p_i) > 0)$ yield the same expected profit. Without entry in the short run, each firm may expect at least the profits in the *loss* event: $\Pi_l(r) = \{(r-c)\cdot\min[K_i, L_i + D_s(p_i)\cdot x_i]\}$. Then, the equilibrium density function for prices $f(p_i)$ is the solution to the following

7 Note that $f(p^*) = 0$ because when p^* is the lowest price, profits are zero, and if there is a tie at p^*, profits will be negative.

problem:

$$\Pi_w(p_i) * [1 - F(p_i)]^{N-1} + \Pi_l(p_i) \cdot \left[1 - (1 - F(p_i))^{N-1}\right] = \Pi_l(r).$$

Rearranging terms and solving for the cumulative distribution function, one obtains:

$$1 - \mathrm{F}(p_i) = \left[\frac{\Pi_l(p_i) - \Pi_l(r)}{\Pi_l(p_i) - \Pi_w(p_i)}\right]^{\frac{1}{N-1}}. \tag{2}$$

The denominator in the right-hand side is negative for any $p_i \in [p^*, r]$. Hence, the numerator must be negative so that profits in the event of *loss* under a price equal with the consumer's reservation buying price r are definitely greater than profits in the event of *loss* under any other price less than r. To guarantee a proper cumulative distribution function, $F(p_i)$ has to be an increasing function of p_i. This is true whenever:

Proposition 2. $\left[\frac{\Pi_l(p_i) - \Pi_l(r)}{\Pi_l(p_i) - \Pi_w(p_i)}\right]$ *is strictly decreasing in* p_i.

Proof. Taking the derivative with respect to $p_i$we obtain $\left[\Pi_l(p_i) - \Pi_w(p_i)\right]\left\{\min\left[K_i, L_i + D_s(p_i) \cdot x_i\right]\right\} + \left[\Pi_l(p_i) - \Pi_l(r)\right]\left\{\min\left[K_i, L_i + D_s(p_i)\right] - \min\left[K_i, L_i + D_s(p_i) \cdot x_i\right]\right\} < 0$, when $\left[\Pi_l(p_i) - \Pi_w(p_i)\right] < 0$ and $\left[\Pi_l(p_i) - \Pi_l(r)\right] < 0$. Therefore, the expression derived in equation (2) is a legitimate candidate for a cumulative distribution function. □

Proposition 3. *In equilibrium, for* $p_i \in [p^*, r]$ *the cumulative distribution function of firm i's pricing is:*

$$\mathrm{F}(p_i) = 1 - \left[\frac{(r - p_i) \cdot \min\left[K_i, L_i + D_s(p_i) \cdot x_i\right]}{(p_i - c)\left\{\min\left[K_i, L_i + D_s(p_i)\right] - \min\left[K_i, L_i + D_s(p_i) \cdot x_i\right]\right\}}\right]^{\frac{1}{N-1}}. \tag{3}$$

Proof. Deviating to a price lower than the firm's reservation selling price p^* is not advantageous since only negative profits can be obtained. Similarly, pricing above the consumer's reservation buying price r is not a profitable deviation because there is zero demand at any such price. Finally, since $\lim_{p_i \to r} F(p_i) = 1$ and $\lim_{p_i \to p^*} F(p_i) = 0$ and $F(p_i)$ is increasing in p_i, F is a well-defined cumulative distribution function. □

2.3. The simulation results

Equation (3) can be used to derive predictions about distributional effects of competition (via N) on price and about relationships between product availability (via K) or loyal customer base (via L) and pricing. Instead of deriving analytical results, we chose to rely on simulation to document the sensitiveness of the derived price equilibrium to changes in the number of firms in a local market, capacity choices in the previous season, and size of consumer brand loyalty. By doing so, our analysis may provide insights that are not subject to errors in functional form assumptions and thus could be generalised to a broader extent. We are interested in examining the effects on pricing (decision in stage two) associated with firms' entry or exit in a local market (decision in stage one) and whether or not capacity constraints and the relative size of loyal customer base would influence such effects. Note that in our model, entry/exit and capacity decisions are made prior to pricing decisions. Thus, at the pricing decision stage, entry/exit and capacity are taken as given.

In particular, we focus on the following two scenarios: (i) all firms are capacity unconstrained with similar loyal customer base $\left(K_i = K_{-i} > 1, L_i = L_{-i}\right)$ and (ii) incumbent firms $(i1, i2)$ can be capacity constrained or unconstrained $\left(K_{i1} < 1, K_{i2} \geq 1\right)$ with different loyal customer base $\left(L_{i1} \neq L_{i2}\right)$, while the entrant firm is unconstrained $\left(K_e \geq 1\right)$. We assume the loyal customer base for the entrant firm is similar to one of the incumbent firms, say firm $i2$ $\left(L_{i2} = L_e\right)$.

By varying the number of firms N for given K's, L's and other parameter values in equation (3), we simulate the price distribution under different scenarios. We further assume 'no business stealing effect': when a new firm enters the market, its loyal customer base comes only from price-sensitive consumers, not from loyal customers of incumbent firms.

We obtain the following results:

Result 1. Under scenario (1)*, the effects of competition on the distribution function of prices is positive at all quantiles.*

Result 1 is illustrated in Figure 1. Both panels a and b (with varying loyal customer base) suggest that an increase in the number of firms has a *uniform positive effect* on the distribution function, leading to an increase in prices for

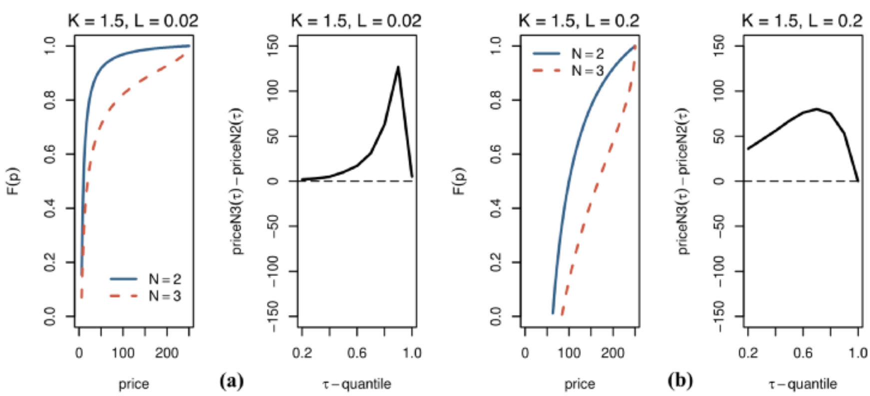

Fig. 1. Effect of competition and loyalty for capacity-unconstrained firms. Panels (a)–(b): cumulative distribution function (left) and its corresponding quantile function (right), $c = 0$, $r = 250$. The results are robust to changing from 1.5 to 1.

all quantiles. Consumers pay a higher price after the entry of the new firm. The residual demand from price-sensitive consumers decreases as some price-sensitive consumers convert to the loyal customer base of the entrant firm. Therefore, the incentive to price low to attract the price-sensitive shoppers decreases. Instead, firms focus on extracting surplus from their loyal customer bases. Comparing panel b to panel a, we see that the size of the loyal customer base matters: the magnitude of the competition effect increases in the lower tail of price distribution but decreases in the upper tail when the size of loyal consumers increases.

Result 2. Under scenario (2)*, the effects of competition on the distribution function of prices differ by incumbent firms' capacities:*

1. Response of the capacity-constrained incumbent firm $i1$: Increased competition from capacity-unconstrained firms leads to an overall price increase in all quantiles.
2. Response of the capacity-unconstrained incumbent firm $i2$: Increased competition from capacity-unconstrained firms leads to an overall price increase in all quantiles when firm capacity $i1$ and loyal consumer base are sufficiently large.

Figure 2 shows *the response of the capacity-constrained incumbent firm* $i1$. We observe that entry by a capacity-unconstrained firm triggers higher prices paid by all consumers in the market. It seems that incumbent firm $i1$ may focus on extracting more surplus from its loyal customers and withdraw from competing for price-sensitive shoppers. When varying the loyal customer base, we found qualitatively consistent results, yet the magnitude of the competition effect will be greater when the loyal consumer base is larger. Panel (b) vs. (a) also show similar patterns as in Result 1: the price increase effect is larger in the lower tail of the price distribution with a larger loyal consumer base in the market.

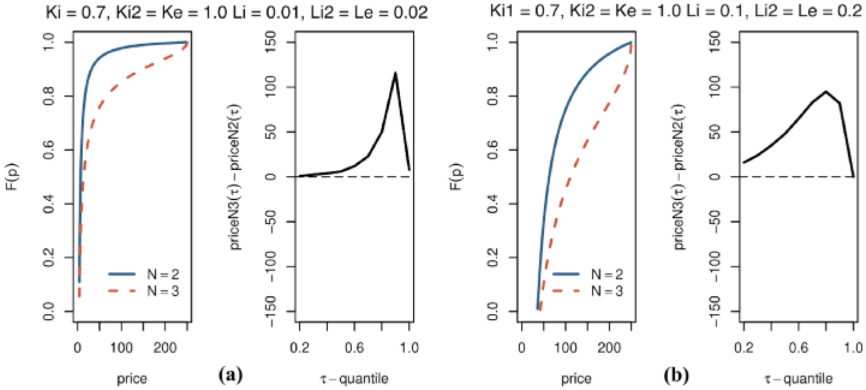

Fig. 2. Effect of competition and loyalty: response of the capacity-constrained firm $i1$. Panels (a)–(b): cumulative distribution function (left) and its corresponding quantile function (right).

Figure 3 shows *response of the capacity-unconstrained firm i2*. When incumbent firm *i*1 is small relative to incumbent firm *i*2, the entrant of another capacity-unconstrained firm leads to a lower price charged by firm *i*2 to all consumers in the market (illustrated in panels a and b). Comparing panel b to panel a, an increase in the loyal customer base would narrow the price jump but does not alter the overall trend.

However, when the incumbent firm *i*1 is larger than incumbent firm *i*2, then effects of the entry of another capacity-unconstrained firm will depend critically on the size of loyal customer base. When the loyal customer base is relatively small, the competition effect will be such that firm *i*2 charges a higher price in the upper quantiles only (illustrated in panel c). When size of loyal customer base increases, then firm *i*2 would charge a higher price to all consumers in the market upon the entry of the new firm (see panel d). Comparing panel b to panel d suggests a switch of pricing strategy from competing for price-sensitive shoppers to extracting surplus from loyal customers only.

The results above generate another implication:

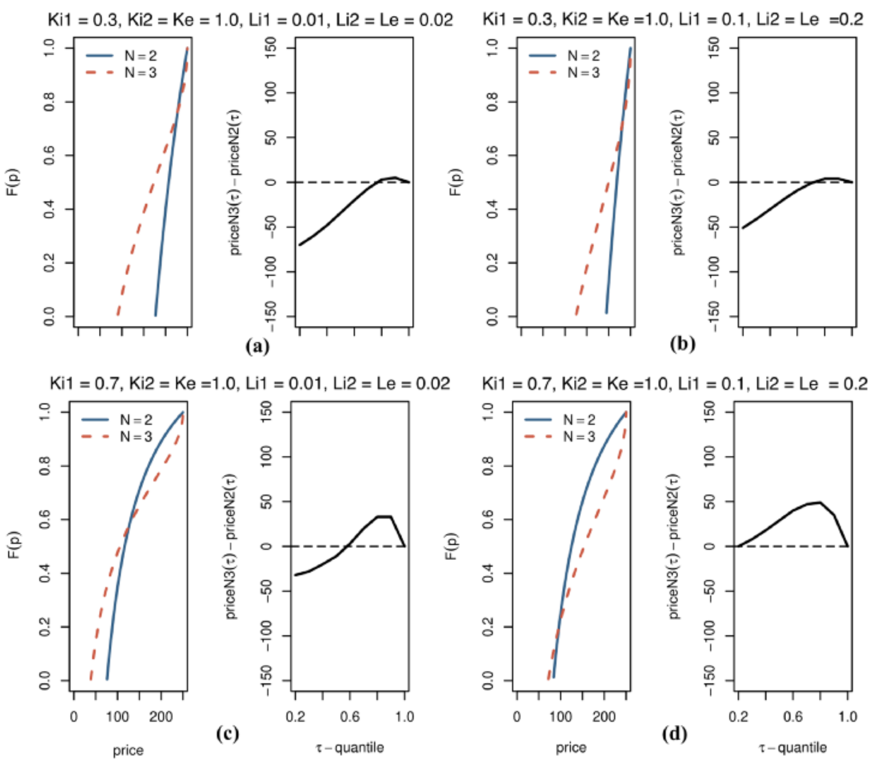

Fig. 3. Effect of competition and loyalty: response of the capacity-unconstrained firm *i*2. Panels (a)–(d): cumulative distribution function (left) and its corresponding quantile function (right).

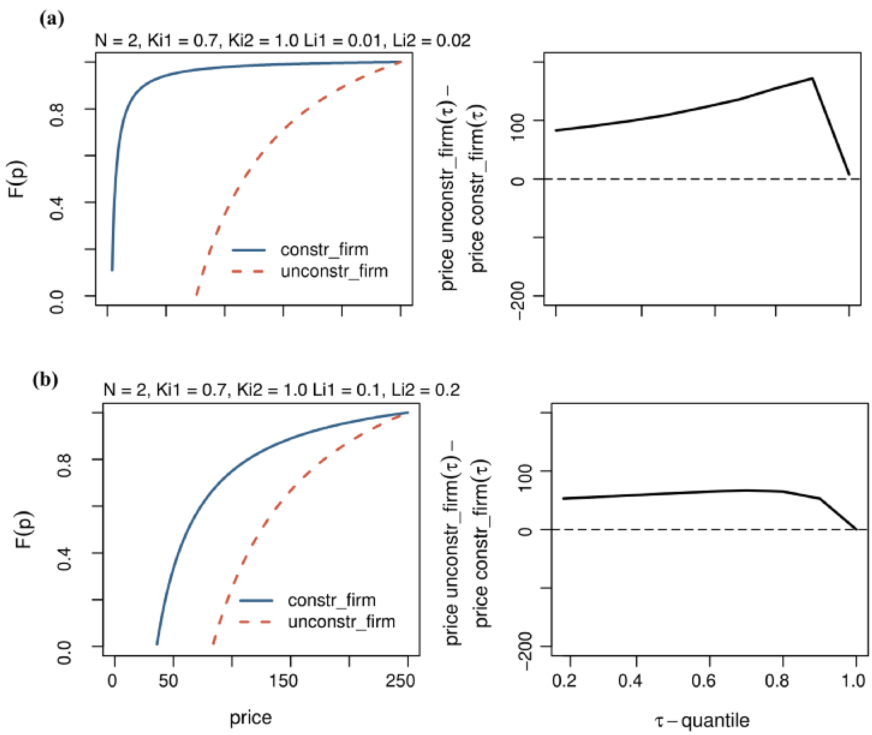

Fig. 4. Relationship between product availability and pricing. Panels (a)–(b): cumulative distribution function (left) and its corresponding quantile function (right).

Result 3. There is a positive relationship between product availability and pricing: capacity-unconstrained firms charge a higher price at each quantile.

Result 3 is presented in Figure 4. We observe that the distribution function of the capacity-unconstrained firm $i2$ stochastically dominates the distribution of the capacity-constrained firm $i1$. Thus, consumers pay higher prices when they buy from a capacity-unconstrained firm. This result is robust to varying sizes of the loyal customer base for each company (panel a and panel b), although the magnitude of the price difference will be smaller when the loyal consumer base in the market is larger.

3. The U.S. corn seed industry

Our analysis in Section 2 posits several hypotheses that can be empirically tested. Empirical studies documenting the effect of competition on price dispersion or on the relationship between product availability and pricing are rare, likely due to lack of data. In this study, we use ample data from the U.S. corn seed industry to test our hypotheses. The seed industry is an excellent case study for three reasons. First, seed companies engage in both temporal and spatial price discrimination; that is, they charge different prices for the same seed product over time and across regions. Hybrid corn seeds, while differing by some observable characteristics that may affect their pricing (for example, embedded biotechnology traits and source of purchase), are generally considered a homogeneous input in essence: they are used for reproduction of a homogeneous output (corn). Therefore, our theoretical model of homogeneous goods is applicable here, after controlling for the observable heterogeneity in pricing relevant features.

Second, seed firms face capacity constraints in production due to the time lag in growing seeds for sale in the following season. We define the capacity constraint in the seed industry as seed companies being unable to produce additional seeds when they are in the current selling season. It takes at least one production season for the seed companies to contract out seed production to the seed farmers (e.g. deciding on the number of seed farms to sign the contract, choosing seed varieties to produce, and the corresponding area of contracted land). There could be a time lag of 3–5 months if the seed farmers are located in the opposite hemisphere, or up to a full year if the seed farmers are local. Therefore, once the seed companies are in the stage of price competition in the marketplace for the current planting season, their seed production quantity is already 'fixed'. Depending on whether their capacities meet market demand for their products, firms can be either capacity constrained (capacity is less than demand) or capacity unconstrained (capacity is more than demand).

Third, the industry is characterised by a mix of brand loyal farmers and price-sensitive farmers.[8] U.S. corn growers plan year $t+1$ production in year t, usually from August$_t$ to April$_{t+1}$, and start planting seed around May$_{t+1}$. From June$_{t+1}$ to September$_{t+1}$ they address in-season challenges and opportunities and harvest fields between October$_{t+1}$ and November$_{t+1}$. The planning process is complicated, as corn growers must choose the right corn seed for their land, with this decision affected by farm location, seed pricing, seed performance, growers' relations to seed companies, and expected seed availability at different times (due to capacity constraints, seeds may not be available throughout the planning period).

In this study, we rely on data collected by GfK Kynetec (hereafter dmrk), St. Louis, MO, which provide information on farm- and firm-level seed purchase for the U.S. corn seed industry between 2004 and 2009. Data are collected annually in June using computer-assisted telephone interviews. Data include seed company identity, seed type (conventional or various genetically modification technology), special seed features, intended end use, net price, seed quantity per transaction, time of order and time of payment, and purchase source.

This study focuses on the most agriculturally productive corn area of the United States, the Corn Belt region. Note that seed is a 'local' product subject

8 We consulted with UW-Extension specialists, regional sales managers and farmers in the state of Wisconsin to confirm the existence of capacity constraints and brand loyalty in the seed industry.

to local agro-climatic conditions. A seed variety developed and marketed for Wisconsin farmers is very different from that for California farmers. Therefore, we define the local market (competition region) at the Crop Reporting District level (CRD). By United States Department of Agriculture (USDA) definition, regions in a given CRD share similar agro-climatic conditions and are likely targeted by similar seed varieties. To account for regional differences, we divide data into *Corn Belt Fringe* (where farmers are likely to alternate between different crops) and *Corn Belt Core* (where corn dominates and crop substitution is less likely). This distinction allows us to assess spatial differences relating to effects of competition on price dispersion and to examine the relationship between product availability and pricing.

The samples are screened to include transactions in CRDs with more than 10 farmers sampled every year to ensure statistical representativity. We further dropped observations with zero or missing prices, unknown seed hybrid numbers and unknown order and payment time. Also excluded are transactions with intended use as 'corn for seed'[9] and purchase source as 'seed left over from last year'. The screening processes reduced our sample size by 5 per cent and 2 per cent, respectively. Our final data set for the *Fringe Corn Belt* includes 53,413 farm-firm-level sales observations for 55 CRDs out of 13 states. For the *Core Corn Belt*, it includes 61,835 farm-firm-level sales observations from 26 CRDs in six states.[10]

Currently, at least 100 seed companies operate in the U.S. market. However, due to the locality of seed, number of firms operating at the local (county or CRD) market is much lower than the number of firms operating at the state or national level. Table 1a shows the average number of seed firms in the United States from 2004 to 2009 at the state, CRD and county level. On average, there are around 4–6 seed firms selling at the county level in the Fringe and Core region. The corresponding numbers are 3–4 times higher at the CRD level and 5–6 times or more higher at the state level for the two regions, respectively. Another trend in the U.S. seed market is the increasing consolidation over time. Table 1a suggests that this consolidation is more evident at the CRD and state level. From 2004 to 2009, the number of firms decreased by 22 per cent and 36 per cent (from 16.27 and 25.54 to 12.63 and 16.5) at the CRD level, and by 27 per cent and 36 per cent (from 31.46 and 44 to 22.85 and 28) at the state level, in the Fringe region and Core region, respectively.

We classify a company as being capacity constrained if all of its sales occur between August$_t$ and December$_t$, and capacity unconstrained otherwise. Table 1b suggests that throughout the study period, over a quarter of total firms are capacity constrained. However, they supply only between 1.5 per cent and 2.7 per cent of corn area in the Core region, and between

9 Seed companies contract farmers to grow crops for seeds that will be sold in the following season.

10 The states with CRDs in the Fringe region are Colorado, Illinois, Indiana, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota and Wisconsin. The states with CRDs in the Core region are Illinois, Indiana, Iowa, Missouri, Nebraska and South Dakota.

Table 1a. Average number of firms at the county, CRD and state level (by region and year)

Year	Fringe			Core		
	Average # of firms (county)	*Average # of firms (CRD)*	*Average # of firms (state)*	*Average # of firms (county)*	*Average # of firms (CRD)*	*Average # of firms (state)*
2004	4.62	16.27	31.46	6.73	25.54	44.00
2005	4.43	14.63	29.38	6.36	23.15	42.67
2006	4.71	15.43	29.38	6.66	24.19	43.67
2007	4.48	13.82	26.85	6.24	20.88	37.00
2008	4.32	13.72	25.46	5.35	19.00	33.17
2009	4.20	12.63	22.85	5.07	16.5	28.0

Table 1b. Share of capacity-constrained firms (by region and year)

Year	Fringe			Core		
	# of firms	*# of capacity-constrained firms*	*Area share of capacity-constrained firms (%)*	*# of firms*	*# of capacity-constrained firms*	*Area share of capacity-constrained firms (%)*
2004	164	41	2.0	138	42	1.8
2005	171	49	2.3	140	43	2.7
2006	171	48	2.2	145	46	1.7
2007	178	45	2.3	136	34	1.5
2008	155	40	2.1	122	38	1.8
2009	146	38	1.7	111	36	2.3

1.7 per cent and 2.3 per cent of corn acreage in the Fringe region. Capacity-constrained firms seem small in size as compared to capacity-unconstrained firms.

By our definition, *loyal consumers* always purchase from one company, while price-sensitive consumers often shop across brands. Our data show that on average farmers placed five orders from two firms throughout the planning and planting seasons. Moreover, 42.2 per cent of farmers in the Fringe region and 38.0 per cent of farmers in the Core region buy their seed inputs from only one company while the rest of the farmers purchase from 2 to 14 different companies.[11] These statistics suggest that there might

11 To further check the evidence for brand loyalty among farmers, we selected a 3-year panel sampled between 2004 and 2006 out of our data to see how many farmers stay with the same single company throughout the 3-year period. Of the 736 panel farmers, 32.3 per cent of farmers in the Fringe region and 32.6 per cent of farmers in the Core region purchase seeds from the same single company. We checked other panel subsamples and found robust evidence that there are farmers who purchase only from one company and farmers who purchase from multiple companies from year to year.

exist both loyal farmers and price-sensitive farmers in the U.S. corn seed market.

Finally, it must be recognised that heterogeneity among seed products might account for price dispersion. Seeds sold in the market differ primarily by whether they incorporate Genetic modification (GM) technology. First introduced to the U.S. corn field in 1995, GM seeds soon gained farmer acceptance, especially after 2004 GM corn seeds fall into two major categories: (i) genes/traits for insect resistance and (ii) genes/traits for herbicide tolerance. Insect resistance (IR) traits are designed to control for damages caused by the European corn borer (ECB) and rootworms (RW). Herbicide-tolerant (HT) traits are designed for tolerance to non-selective post-emergence herbicides such as glyphosate, glufosinate and other herbicides. GM traits were initially introduced as single traits but then were stacked together. To account for product heterogeneity, our analysis groups seeds into four types: (i) conventional (non-GM); (ii) GM IR single-stacked (ECB, RW); (iii) GM IR double-stacked (ECB + RW) and (iv) GM IR and HT multi-stacked.[12] IR-traited GM seeds are treated differently from IR and HT stacked seeds to further capture price differences inherently implied by GM technology.

3.1. Scenario samples

The theoretical model in Section 2 is tested using seed market data presented above. Given the distribution of capacity-constrained and capacity-unconstrained seed firms in the two regions, we divide data in two sample scenarios. In **Scenario 1** competition is among symmetrically capacity-unconstrained firms and in **Scenario 2** competition is among firms with asymmetric capacities; that is, among both capacity-constrained and capacity-unconstrained firms.

We define a CRD as symmetrically unconstrained if market share of unconstrained seed companies in a year equals or exceeds 98 per cent and as asymmetric otherwise.[13] This yields an unbalanced panel sample of symmetrically unconstrained firms with 36,385 and 40,178 observations in Fringe and Core regions, respectively. We denote this sample as **unconstr_sample**. The asymmetrically constrained sample is also an unbalanced panel sample, with 17,028 and 21,657 observations in Fringe and Core regions, respectively. We denote this sample as **asymm_sample**.

3.2. Evidence of price dispersion

Table 2 presents detailed summary statistics of seed prices by seed types (conventional vs. various GM), regions (Fringe vs. Core of Corn Belt) and sample scenarios (unconstr_sample vs. asymm_sample). We report the

12 Includes GM IR and HT double-, triple- and quadruple-stacked.

13 The 98 per cent threshold was chosen arbitrarily to represent that firms competing in a region are very likely to be symmetrically capacity unconstrained. The empirical results are robust to different threshold levels.

Table 2. Summary statistics of seed prices in $ per 50 lb./bag (by region and type of seed)

Seed type	Fringe					Core				
	N	*Gini*	*0.1-q*	*0.5-q*	*0.9-q*	*N*	*Gini*	*0.1-q*	*0.5-q*	*0.9-q*
Conventional:										
unconstr_ sample	8,951	0.13	68	93	120	9,447	0.12	70	95	120
asymm_ sample	5,767	0.13	67	90	119	6,666	0.11	71	92	116
GM IR single-stacked:										
unconstr_ sample	3,024	0.1	89	112	135	5,052	0.09	90	113	135
asymm_ sample	1,267	0.1	85	110	133	2,723	0.1	86	110	133
GM HT single-stacked:										
unconstr_ sample	7,983	0.15	91	125	181	5,803	0.16	92	129	187
asymm_ sample	3,351	0.16	87	127	189	3,229	0.17	85	125	192
GM IR double-stacked:										
unconstr_ sample	198	0.11	110	139	173	585	0.09	110	134	156
asymm_ sample	128	0.11	104	130	166	340	0.10	106	131	159
GM IR and HT double-stacked:										
unconstr_ sample	7,196	0.11	102	126	162	8,066	0.10	104	127	161
asymm_ sample	2,219	0.12	100	126	170	3,285	0.10	100	124	154
GM IR and HT triple-stacked:										
unconstr_ sample	7,547	0.13	128	180	240	9,265	0.14	130	180	250
asymm_ sample	3,602	0.13	130	185	249	4,582	0.15	130	186	261
GM IR & HT quad- stacked:										
unconstr_ sample	1,486	0.11	143	193	238	1,960	0.12	140	189	246
asymm_ sample	693	0.11	143	193	235	832	0.12	143	200	248

0.1-, 0.5- and 0.9-quantiles of the net price[14] and the median Gini coefficient[15]

14 The k^{th}-quantile of the net price is a value p such that the probability that the price will be less than p is at most k and the probability that the net price will be greater than p is at most $1-k$.

15 The Gini coefficient G is computed as follows: for price values $p_i, i = 1, \ldots, N$, and mean price $\bar{p}$, $G = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \left| p_i - p_j \right|}{2N^2 \bar{p}}$.

for conventional seeds and various GM seeds. The law of one price is easily rejected by statistics in Table 2. Seed price varies greatly for any given seed type, by region and by scenario samples; thus, price dispersion is generally present in the market. The Gini coefficient is in the 0.1–0.2 range, corresponding to an expected price difference between 20 per cent and 40 per cent of the mean net price for any two randomly selected seed products in a given sample. GM seeds are on average priced at a premium over conventional seeds, and GM seeds with multiple-trait-stacking systems are generally priced more than single-trait GM seeds.

4. Econometric specification

Our theoretical model in Section 2 suggests distributional effects to competition, which may interact with product availability and loyal customer base. The empirical estimation builds on the equilibrium distribution function derived in equation (3). We use quantile regression analysis, which allows for variables of interest (firm-specific and market-level factors) to have heterogeneous effects at different points in the distribution function of price.

We consider two linear parameters' econometric specifications corresponding to Scenarios 1 and 2 as follows:

$$\begin{aligned}\ln\left(P\left(k\right)\right)_{ict} \;=\; & \beta_{0(k)} - \beta_{1(k)} HHI_{ct} \times 100 + \Upsilon^{1}{}_{(k)}\boldsymbol{M} + \Upsilon^{2}{}_{(k)}\boldsymbol{F} + \Upsilon^{3}{}_{(k)}\boldsymbol{C} \\ & + \Upsilon^{4}{}_{(k)}\boldsymbol{S} + \Upsilon^{5}{}_{(k)}\boldsymbol{X} + \alpha_{c(k)} + \epsilon_{ict} \end{aligned} \tag{4}$$

where $P(k)_{ict}$ is the k^{th} price quantile for seed supplied by firm i in CRD c at time t; HHI_{ct} is the Herfindahl–Hirschman Index (HHI) in CRD c at time t; **M, F, C, S, X** are vectors of covariates related to discriminatory pricing, grouped into market attributes (**M**), farmer attributes (**F**), company attributes (**C**), seed attributes (**S**) and other variables (**X**); Alpha's $\left(\alpha_{c(k)}\right)$ capture the CRD fixed effect accounting for regional heterogeneity, beta and gammas are the parameter coefficients and ϵ_{ijt} is an unobservable error term.[16] We further account for spatial price discrimination by estimating separate regressions for Fringe and Core regions. Summary statistics of selective variables are presented in Table 3.

4.1. Market attributes

We use HHI to measure market concentration computed at the CRD level: $\text{HHI} = \sum_{i=1}^{N} s_i^2$, where N is number of firms in a CRD and s_i is market share of firm i computed as firm-specific seed area divided by total seed area in a CRD. On average, the market in the Core region is less concentrated than that in the Fringe region, with HHI at 0.16 vs. 0.21 for the unconstrained sample and at

16 Note that we are interested in the distribution of the price variable $\ln\left(\mathrm{P}_{\mathrm{ict(K)}}\right)$, relative to CRD fixed effects, $\alpha_{\mathrm{c(k)}}$; thus our parameters will reflect the estimation of the conditional distribution of $\left[\ln\left(P_{\mathrm{ict(K)}}\right) - \alpha_{\mathrm{c(K)}}\right]$.

Table 3. Descriptive statistics of selective control variables

Variable	Mean		Standard Deviation		Min		Max	
	Fringe	Core	Fringe	Core	Fringe	Core	Fringe	Core
Market attributes								
HHI								
unconstr_sample	0.21	0.16	0.09	0.04	0.07	0.08	0.81	0.32
asymm_sample	0.17	0.14	0.07	0.04	0.06	0.07	0.50	0.35
Product density (%)								
unconstr_sample	6	11	2.64	2.86	1	2	12	16
asymm_sample	5	10	1.68	3.03	1	2	8	15
Farmer attributes								
Loyal farmer								
unconstr_sample	0.78	0.83	0.41	0.38	0	0	1	1
asymm_sample	0.79	0.84	0.40	0.36	0	0	1	1
Share of loyal farmers in a CRD (%)								
unconstr_sample	78	83	9.66	6.11	27	61	97	97
asymm_sample	79	84	9.52	7.73	20	57	97	97
Number of loyal farmers in a CRD at the firm level								
unconstr_sample	40	64	46.68	68.27	0	0	202	266
asymm_sample	27	50	42.14	51.37	0	0	257	218
Difference between order time and payment time (in months)								
unconstr_sample	1.78	1.65	2.95	2.73	0	0	11	11
asymm_sample	1.91	1.66	2.98	2.56	0	0	11	11
Company attributes								
Biotechnology company								
unconstr_sample	0.80	0.80	0.40	0.40	0	0	1	1
asymm_sample	0.74	0.73	0.44	0.44	0	0	1	1
Seed attributes								
Seed is conventional								
unconstr_sample	0.25	0.24	0.43	0.42	0	0	1	1
asymm_sample	0.34	0.31	0.47	0.46	0	0	1	1
HLFS and/or HLES feature								
unconstr_sample	0.25	0.24	0.43	0.43	0	0	1	1
asymm_sample	0.19	0.17	0.39	0.37	0	0	1	1
Seed output is intended for ethanol production								
unconstr_sample	0.11	0.10	0.31	0.30	0	0	1	1
asymm_sample	0.07	0.08	0.25	0.27	0	0	1	1

(continued)

Table 3. (Continued)

Variable	Mean		Standard Deviation		Min		Max	
	Fringe	Core	Fringe	Core	Fringe	Core	Fringe	Core
New seed								
unconstr_sample	0.49	0.48	0.50	0.50	0	0	1	1
asymm_sample	0.55	0.49	0.5	0.5	0	0	1	1
New seed is conventional								
unconstr_sample	0.09	0.07	0.28	0.26	0	0	1	1
asymm_sample	0.12	0.11	0.33	0.31	0	0	1	1
Observations								
unconstr_sample	36,385	40,178						
asymm_sample	17,028	21,657						

0.14 vs. 0.17 for the asymm_sample. Variation in individual CRDs is larger, with HHI ranging from 0.06 to 0.81 across samples and regions.

Another market attribute is crowdedness of seed-product varieties. When farmers can choose among many seed products with similar technologies, there is high substitutability in demand for seed products. This affects firms' seed pricing. We construct a product density variable to measure such market crowdedness, calculated as the per cent of number of seed hybrid varieties in a given CRD market relative to total number of seed hybrids present in all CRD markets in Fringe or Core regions. A higher product density measure suggests more intensive product competition in a market. Similar to the HHI, the market in the Core region is more competitive than that in the Fringe region in terms of crowdedness, with an almost doubling crowded product space: 11 per cent vs. 6 per cent for the unconstr_sample and 10 per cent vs. 5 per cent for the asymm_sample.

4.2. Farmer attributes

We distinguish between price-sensitive and loyal farmers. Anecdotal evidence suggests that loyal farmers are more likely to order early in the season, often before the end of the year, whereas price-sensitive farmers shop around for the entire season (from August_t to April_{t+1}) and tend to order later to obtain seeds on sale. Such a claim is supported by our data: 74 per cent of farmers in the Fringe region and 87 per cent of farmers in the Core region, who do business with only one company, place their seed orders before the end of the year. Therefore, we construct a dummy variable to proxy the loyal farmer type: $loyal = 1$ if a farmer ordered his/her seed inputs between August and December. Otherwise, the farmer is labeled as price sensitive.[17]

17 We test the validity of the proxy variable for brand loyalty using a 3-year (2004–2006) farmer panel built from our data. By computing the Jaccard–Needham similarity index between a dummy for 'purchasing before December 31' and a dummy for sticking with the same seed company throughout the 3 years, we found that the two variables share 75.7 per cent and

We construct two more variables as proxies for size of the loyal customer base in any local market. One relates to the share of loyal farmers in a CRD defined as per cent of transactions in each CRD occurring between August$_t$ and December$_t$ over those during the entire season; the other one relates to the number of firm-specific loyal farmers in a CRD, defined as total number of transactions per company occurring between August$_t$ and December$_t$ at the CRD level. On average, loyal farmers in the unconstrained sample place more firm-specific orders than those in the asymm_sample: 40 loyal farmers vs. 27 price-sensitive farmers in the Fringe region and 64 loyal farmers vs. 50 price-sensitive in the Core region.

Also, some seed companies, as price discrimination schemes, offer cash discounts for farmers paying in advance.[18] We capture this effect by measuring the difference between order date and payment date. Table 3 suggests that, on average, it takes farmers longer to pay in the Fringe region (about 2 months) than farmers in the Core region (about one and a half months). However, at the individual farmer level, time variance is greater, ranging from on-site payment (no time difference) to 11 months.

4.3. Company attributes

Seed companies are differentiated by vertical structure affiliation. A firm is either independent (e.g. Beck's Hybrids and Unity Seeds), or belongs to a vertically integrated biotechnology seed firm (e.g. Asgrow is part of Monsanto's corn seed division). In both scenario samples, biotechnology companies dominate independent companies in the market, having an average market share of 80 per cent in the unconstr_sample and 73–74 per cent in the asymm_sample. However, at local market levels, their share ranges from 0 per cent to 100 per cent, indicating a large variance within the sample. For Scenario 2 only, we also distinguish between capacity-constrained and -unconstrained firms. Further, we allow for the possibility that unconstrained vertically integrated biotechnology seed companies may use a different pricing strategy by introducing the corresponding interaction terms.

4.4. Seed attributes

Seeds differ by technology (conventional vs. various GM) and whether seed hybrid has higher levels of fermentable starch (HLFS) and/or extractable starch (HLES)—special features widely recognised as suitable for ethanol production. Farmers also reported intended output use being ethanol or not.

The average share of conventional seed in the Core region is 24 per cent in the unconstr_sample and 31 per cent in the asymm_sample. Similar patterns are observed in the Fringe region. Conventional seed share decreases quickly over time, reflecting general adoption of GM seeds by U.S. corn farmers. Seeds

71.5 per cent of the members in the Fringe region and Core region, respectively. Note that we do not have a direct measurement of brand loyalty. For testing purpose, we think that farmers sticking with the same seed company over time might be most likely loyal and thus use that dummy as the true loyal customer indicator.

18 An example of cash discount reported on a seed company website is as follows: 8 per cent for payments no later than November 10 and 6 per cent for payments no later than January 5.

with special features (HLFS and/or HLES) have a larger share in the unconstr_sample than that in the asymm_sample for both Core and Fringe regions: 25 per cent vs. 19 per cent and 24 per cent vs. 17 per cent, respectively. However, farmers have indicated intended output use in ethanol production for less than half of these seeds.

We control for whether seed is new to the market or not (i.e. if it had never presented in the dmrk data since 1994). Seed company pricing strategy and farmer willingness to pay likely differ for seeds new to the market. Around 50 per cent of seeds are new in the market across samples and regions, indicating that seed companies continuously develop new seed varieties. We also interact with the variables representing ‘new seed’ and ‘conventional seed’ to assess if farmers are more (or less) risk averse to conventional new seeds as opposed to GM seeds.

4.5. Other variables

We control for price dispersion attributed to other factors such as purchase sources, time of order and year trends. Purchasing sources include ‘directly from the seed company or its representative’, ‘Myself being a dealer for the seed company’, ‘other farmer who is a dealer or agent’ and ‘others’. At least 60 per cent of transactions were classified into one of the first three categories.

Prices may fluctuate throughout the season, which may depend on the probability at time of order that demand will exceed capacity. We construct two time-trend variables to control for such an effect. Using January$_t$ as the benchmark, orders placed prior to or after this month will take the corresponding values that equal to the month’s distance to January$_t$. For example, if an order is placed in August$_{t-1}$, then $priorJAN = 5$ and $postJAN = 0$. If an order is placed in February$_t$, then $priorJAN = 0$ and $postJAN = 1$. Our data suggest that most orders are placed between November$_{t-1}$ and January$_t$. Finally, we include a year trend variable to capture the time-trend effect that might be associated with technology advances, inflation and other potential time-consistent structural changes.

5. Econometric estimation

Our theoretical model allows us to examine the whole distribution function of pricing using a quantile regression approach to accomplish research objectives. One challenge to the identification strategy in the econometric specification (see Section 4) is related to the assumption that the error term and explanatory variables are contemporaneously uncorrelated. However, both market concentration, as measured by HHI, and seed prices are likely endogenous as jointly determined in the model. A firm’s decision to enter a market is affected by its marketing strategies. Econometricians can hardly observe all determinants of these strategies. Therefore, the market concentration measurement HHI may be correlated with unobserved factors affecting the response variable. For example, a seed company may understand demand, with knowledge of customer base correlated with the decision to enter the market and affecting pricing strategies; yet, such private information cannot be observed in our data set. If so, the standard quantile regression (QR) proposed by Koenker and Bassett (1978) will provide biased and inconsistent estimates for behavioural parameters.

We consider an alternative econometrics approach, the fixed-effect instrumental variable quantile regression (FEIVQR) proposed by Harding and Lamarche (2009). The FEIVQR builds on Chernozhukov and Hansen (2008)'s model. It facilitates estimation of covariate effects at different quantiles while controlling for additive fixed effect as introduced in Koenker (2004) that may affect response and are correlated with independent variables.

Since market concentration is likely to be corelated with unobservables in the model, we propose an instrument for the HHI variable: the lagged value of HHI (HHI_{t-1}) to account for lags in the seed production process that takes on average 8 months. As a result, firm managers may use previous year's information on market concentration to decide seed production quantity for next year's market.

The lagged value HHI_{t-1} is an appropriate instrument if it is correlated with the endogenous regressor HHI_t (not a 'weak instrument') and not with the error term (the orthogonality condition). Since the model is 'exactly identified', the orthogonality condition holds by construction in equation (4). We use the Arellano and Bond (1991) estimator to test for first, second and third serial correlations in the idiosyncratic errors. The null hypothesis of 'no serial correlation' in all samples cannot be rejected in favour of the alternative. The p-value exceeds 0.10 for all levels of correlation tested. To test for 'weak instruments', we examine the reduced-form regression and evaluate the explanatory power of HHI_{t-1}. The values of the F statistics for the null hypothesis in the unconstr_sample and asymm_sample are 259.78 and 68.63 in the Fringe region and 558.34 and 499.47 in the Core region. The p-value is less than 0.0001 in all samples. Using the simple rule-of-thumb of $F = 10$ proposed by Stock and Watson (2003), results suggest our instrument may not be weak.

Another challenge to the identification is the homoscedasticity assumption that the error term is independent and identically distributed with mean zero and constant variance. The Breusch–Pagan test result applied to the econometric specification (see Section 4) rejects null of homoskedasticity assumption in all samples. The chi statistics are in the range of 330 and 1,402 with p-values less than 0.0001. To account for the heteroskedasticity of the error terms, we estimate the FEIVQR model using the 'xy-pair' bootstrap robust standard errors (Efron and Tibshirani, 1994). Besides, we expect that the correlation of the error terms in a CRD is likely driven by a common shock process, thus we add CRD-specific fixed effects to control for the within-cluster correlation of the error (e.g. following Cameron and Miller, 2015).

6. Empirical finding

Tables 4–6 report the econometric results of our analysis of distributional effects of competition on price, as well as the relationship between product availability and pricing. We estimate the model using the FEIVQR method, with bootstrap robust standard errors. Results will be discussed for each sample scenario and for both regions. We also compare results from Two Stage Least Square Instumental Variable (2SLS IV) to illustrate the likely misleading results if only a mean regression is conducted.

6.1. The case of unconstr_sample

Estimation results in the unconstr_sample are presented in Table 4 for Fringe and Core regions. We report bootstrapped standard errors[19] (for the FEIVQR model) and heteroskedastic robust standard errors (for the 2SLS IV model) in parentheses.

6.1.1. Market attributes

The negative of the $HHI \times 100$ is reported to capture the effect of a decrease in market concentration on the log of net price. Statistical evidence shows that firms operating in the unconstrained sample are governed by different pricing strategies in the Fringe vs. the Core regions. The effect of a decrease in market concentration is statistically significant and negative across the distribution function of price in the Fringe region and positive in the Core region. In the Fringe region, estimates suggest that a 1 per cent point increase in $-HHI * 100$ may lead to a decrease in seed price by -0.9 per cent ($-\$0.72$ per bag) at the 0.1-quantile and by -1.9 per cent decrease ($-\$3.03$ per bag) at the 0.9-quantile. This price effect is positive in the Core region with a 10.2 per cent increase (\$8.63 per bag) at the 0.1-quantile and a 7.3 per cent increase (\$11.95 per bag) at the 0.9-quantile. We test for equality of the slope parameters across quantiles using Wald test statistics. Table 6 reports test results for selected quantile pairs. In the Fringe region, the null is rejected in favour of the alternative for all quantile cases, while in the Core region, the slope differs between the upper tail and lower tail but remains the same across quantiles in the middle portion of distribution. The market concentration effect differs along the pricing distribution in these regions.

The 2SLS IV results find that a decrease in market concentration significantly increases mean price in both Fringe and Core regions.[20] Thus, the FEIVQR method provides a much richer description of the distributional effects of market concentration. Our results in the Core region confirm with Result 1 illustrated in the theoretical prediction, as described in Section 2.3. An increase in the number of capacity-unconstrained firms has a positive effect on the distribution function, leading to an increase in prices for all quantiles. The effect is with a greater magnitude in the lower quantiles, implying that

19 Obtained by sampling 1,000 samples with replacement from the original sample.

20 The 2SLS IV estimates are larger than the Ordinary Least Square (OLS) estimates in the Fringe and Core regions, implying that the OLS estimates are downward biased (e.g. if pricing and knowledge of the demand side are negatively correlated).

Table 4. FEIVQR and 2SLS IV regression results for selective variables. Core and fringe regions, unconstr_sample

	Dependent variable: log net price							
	Fringe region				Core region			
	FEIVQR			2SLSIV	FEIVQR			2SLS IV
	0.1-q	0.5-q	0.9-q	Mean	0.1-q	0.5-q	0.9-q	Mean
Market attributes								
(-) HHI * 100	−0.009*	−0.042***	−0.019***	0.027***	0.102***	0.081***	0.073***	0.107***
	(0.005)	(−0.007)	(0.003)	(0.004)	(0.015)	(0.007)	(0.007)	(0.008)
Product Density	0.079**	0.062***	0.015***	−0.030***	−0.053***	−0.122***	−0.192***	−0.067***
	(0.039)	(0.010)	(0.004)	(0.004)	(0.007)	(0.011)	(0.019)	(0.004)
Farmer attributes								
Loyal farmer	−0.005	0.006	0.004	0.0002	0.032***	0.051***	0.054***	0.037***
	(0.009)	(0.007)	(0.006)	(0.005)	(0.008)	(0.011)	(0.016)	(0.007)
Share of loyal farmers in a CRD	0.001*	0.0005	0.00009	0.003***	−0.0001	0.016***	0.023***	0.007***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.002)	(0.003)	(0.003)	(0.001)
Number of loyal farmers in a CRD at the firm level	0.0005*	0.00004	0.0003***	0.001***	0.001	0.001***	0.001***	0.0007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Company attributes								
Biotechnology company	0.065***	0.055***	0.029***	0.035***	0.039***	0.041***	0.039***	0.056***
	(0.008)	(0.006)	(0.005)	(0.004)	(0.007)	(0.008)	(0.011)	(0.005)
Seed attributes								
Seed is conventional	−0.346***	−0.328***	−0.284***	−0.292***	−0.286***	−0.283***	−0.291***	−0.300***
	(0.014)	(0.008)	(0.007)	(0.005)	(0.007)	(0.010)	(0.015)	(0.007)
HLFS and/or HLES features	0.005	−0.0036***	−0.042***	0.014***	−0.013***	−0.052***	−0.073***	−0.031***
	(0.019)	(0.007)	(0.004)	(0.004)	(0.006)	(0.008)	(0.011)	(0.005)

(continued)

Table 4. (Continued)

	Dependent variable: log net price							
	Fringe region				Core region			
	FEIVQR			2SLSIV	FEIVQR			2SLS IV
	0.1-q	0.5-q	0.9-q	Mean	0.1-q	0.5-q	0.9-q	Mean
Seed output is intended for	0.051***	0.036***	0.002	0.002	−0.013*	−0.058***	−0.103***	−0.038***
ethanol production	(0.021)	(0.020)	(0.005)	(0.004)	(0.007)	(0.010)	(0.015)	(0.006)
New seed	0.002	−0.014**	−0.002	0.015***	0.015***	−0.006	−0.024***	−0.006
	(0.011)	(0.008)	(0.005)	(0.003)	(0.005)	(0.006)	(0.010)	(0.004)
New seed is conventional	−0.028	−0.010	−0.03***	−0.059***	−0.054***	−0.054***	−0.056***	−0.046***
	(0.021)	(0.011)	(0.009)	(0.007)	(0.011)	(0.014)	(0.021)	(0.010)
Constant	2.575***	1.980***	4.061***	5.732***	5.982***	7.005***	8.931***	6.077***
	(0.907)	(0.306)	(0.077)	(0.159)	(0.147)	(0.179)	(0.394)	(0.091)
CRD Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Observations	36,385				40,178			

Note: Statistical significance is noted by an asterisk (*) at the 10 per cent level, two asterisks (**) at the 5 per cent level and three asterisks (***) at the 1 per cent level. Results not reported here but discussed in the text include the difference between order time and payment time, different types of GM seeds (IR single-stacked, IR double-stacked, IR and HT multi-stacked), temporal pricing (prior or after January), purchase source effects and a time trend.

Table 5. FEIVQR and 2SLS IV regression results—Fringe and Core region, asymm_sample

	Dependent variable: log net price							
	Fringe region				Core region			
	FEIVQR			2SLS IV	FEIVQR			2SLS IV
	0.1-q	0.5-q	0.9-q	Mean	0.1-q	0.5-q	0.9-q	Mean
Market attributes								
(-) HHI * 100	0.039**	0.034***	0.020*	0.052***	−0.003	0.021***	0.083***	0.031***
	(0.015)	(0.013)	(0.011)	(0.013)	(0.008)	(0.004)	(0.011)	(0.005)
Product density	−0.110*	−0.125	−0.113	−0.067***	−0.020***	−0.029***	−0.096***	−0.034***
	(0.048)	(0.062)	(0.062)	(0.014)	(0.006)	(0.003)	(0.012)	(0.003)
Company attributes								
Biotechnology company	0.118*	0.095	0.025	0.109***	0.126***	0.080***	0.003	0.098***
	(0.064)	(0.066)	(0.050)	(0.023)	(0.037)	(0.028)	(0.028)	(0.018)
Unconstrained company	0.093**	0.046	0.014	0.068***	0.069***	0.042***	0.003	0.044***
	(0.044)	(0.034)	(0.027)	(0.016)	(0.015)	(0.012)	(0.015)	(0.009)
Unconstrained firm is	−0.088	−0.057	0.011	−0.074**	−0.096**	−0.056**	0.033	−0.076***
biotechnology	(0.061)	(0.060)	(0.045)	(0.023)	(0.038)	(0.029)	(0.029)	(0.019)
Farmer attributes								
Loyal farmer	0.029*	0.012	−0.004	0.008	0.0001	0.012**	−0.008	0.005
	(0.016)	(0.016)	(0.015)	(0.007)	(0.010)	(0.006)	(0.007)	(0.005)
Share of loyal farmers in a CRD	0.006	0.003	0.002	0.0005	−0.007***	−0.012***	−0.014***	−0.011***
	(0.005)	(0.005)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of loyal farmers	0.002***	0.001***	0.001**	0.001***	0.001***	0.001***	0.001***	0.0009**
in a CRD at the firm level	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

(continued)

Table 5. (Continued)

	Dependent variable: log net price							
	Fringe region				Core region			
	FEIVQR			2SLS IV	FEIVQR			2SLS IV
	0.1-q	0.5-q	0.9-q	Mean	0.1-q	0.5-q	0.9-q	Mean
Seed attributes								
Seed is conventional	−0.336***	−0.302***	−0.283***	−0.329***	−0.312***	−0.295***	0.287***	−0.318***
	(0.019)	(0.017)	(0.016)	(0.008)	(0.009)	(0.005)	(0.007)	(0.005)
HLFS and/or HLES features	0.062***	0.024***	−0.005	0.03***	0.004	−0.011***	−0.015***	0.013**
	(0.012)	(0.012)	(0.010)	(0.005)	(0.008)	(0.004)	(0.005)	(0.004)
Seed output is intended	0.105***	0.083*	0.063	0.068***	0.024***	0.018***	0.025***	0.024***
for ethanol production	(0.041)	(0.047)	(0.030)	(0.010)	(0.009)	(0.006)	(0.007)	(0.005)
New seed	0.011	−0.006	−0.004	−0.002	0.008	0.011***	0.006	0.004
	(0.015)	(0.015)	(0.011)	(0.005)	(0.006)	(0.004)	(0.005)	(0.004)
New seed is conventional	−0.034	−0.044**	−0.04**	−0.032***	−0.033***	−0.034***	−0.026***	−0.032***
	(0.022)	(0.020)	(0.016)	(0.009)	(0.013)	(0.008)	(0.009)	(0.007)
Constant	9.287***	8.876***	7.796***	6.818***	4.938***	5.887***	7.868***	6.010***
	(3.262)	(2.438)	(3.220)	(515.000)	(0.200)	(0.103)	(0.311)	(0.141)
CRD Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Observations	17,028				21,657			

Note: Statistical significance is noted by an asterisk (*) at the 10 per cent level, two asterisks (**) at the 5 per cent level and three asterisks (***) at the 1 per cent level. Results not reported here but discussed in the text include the difference between order time and payment time, different types of GM seeds (IR single-stacked, IR double-stacked, IR and HT multi-stacked), temporal pricing (prior or after January), purchase source effects and a time trend.

Table 6. Wald test for equality of slope parameters in FEIVQR

	$\beta_{0.1}-\beta_{0.5}$		$\beta_{0.5}-\beta_{0.9}$		$\beta_{0.1}-\beta_{0.9}$	
	Wald_test	p_val	Wald_test	p_val	Wald_test	p_val
unconstr_sample						
Fringe region						
(-) HHI * 100	30.784	0.000	9.468	0.002	3.453	0.063
Core region						
(-) HHI * 100	2.004	0.157	2.269	0.132	3.509	0.061
asymm_sample						
Fringe region						
(-) HHI * 100	0.211	0.646	1.854	0.173	2.129	0.145
Unconstrained firm	1.510	0.219	1.219	0.270	3.463	0.063
Unconstrained firm is biotechnology	0.361	0.548	2.099	0.147	3.154	0.076
Core region						
(-) HHI * 100	10.285	0.001	30.558	0.000	40.957	0.000
Unconstrained firm	2.838	0.092	5.365	0.021	9.645	0.002
Unconstrained firm is biotechnology	1.027	0.311	6.985	0.008	8.071	0.004

seed firms in this region may benefit from a large number of loyal farmers. They focus on extracting surplus from the loyal group and avoid competition in the price-sensitive farmer group when facing new market entries.

The product density effect is statistically significant and positive across the entire distribution function of prices in the Fringe region, but significant and negative in the Core region.

6.1.2. Farmer attributes

Being a loyal farmer has a positive and significant effect on the mean and across the distribution of prices in the Core region, yet it is not significant in the Fringe region. Additionally, the share of loyal farmers in a CRD is positive and statistically significant in both regions. The economic magnitude is relatively small and significant only in the lower tail in the Fringe region, yet larger and statistically different across quantiles in the Core region. The total number of transactions per firm occurred in August$_t$ to December$_t$ at the CRD level and positively affects prices as well; the corresponding coefficient is statistically significant in the Fringe and Core regions but has a small magnitude. These results indicate that an increase in a firm's loyal customer base may raise the price of seeds for farmers in the two regions. Thus, they confirm our theoretical predictions, described in Section 2.3, relating to the effect of an increase in the value of the loyal customer base parameter (L). We also find that a 1-month increase between payment time and order time may result in a price increase between 0.3 per cent and 0.6 per cent in the Core region.

6.1.3. Company attributes

We find evidence of distributional price effects associated with the type of seed company. Biotechnology companies charge a price premium at the mean and across the distribution of prices in both regions, with a higher magnitude in the lower tail of the distribution.

6.1.4. Seed attributes

The variable corresponding to conventional seed and variables controlling for various GM technologies show that GM seeds are sold at a price premium over conventional seeds. Additionally, GM seeds with single-stacking IR systems or with multiple technologies are generally more expensive than GM seeds with single-stacking HT systems. We also find that farmers in the Fringe region might be willing to pay a price premium to plant seeds with HLFS and/or HLES features or for seeds intended for ethanol production. Negative coefficients on 'new seed' and 'new seed is conventional' indicate that farmers in the two regions associate conventional new seeds with higher performance uncertainty as opposed to GM.

6.1.5. Other variables

Variables of top three purchase sources are positive or negative, or change signs across the price distribution in the Fringe region, and positive in the Core region. These results may reflect the presence of price discrimination across regions and across purchase sources.

Variables capturing transaction timing (prior or after January) indicate that seed companies may engage in time-based price discrimination. Both loyal and price-sensitive farmers in the Core region can save on input costs if they order seeds early in the season; loyal farmers in the Fringe save if they order earlier and price-sensitive farmers save if they order later.

Finally, we find the time trend effect is positive and significant across regions, with a greater magnitude in the Core. This indicates that technology advancement, inflation and other potential time-consistent structural changes contribute to part of the observed price increase.

6.2. The case of asymm_sample

Estimation results in the asymm_sample are reported in Table 5 for both Fringe and Core regions. Again, bootstrapped standard errors used in the FEIVQR model and the heteroskedastic robust standard errors used in the 2SLS IV model are reported in the corresponding parentheses.

6.2.1. Market attributes

The effect of a one-point increase in market concentration as measured by $-HHI \times 100$ is positive and statistically significant at the mean and across the distribution function of prices in both Fringe and Core regions. Estimates range between 3.9 per cent ($2.95 per bag) at the 0.1-quantile and 2.0 per cent ($3.10 per bag) at the 0.9-quantile in the Fringe region, and between 2.1 per cent ($2.40 per bag) at the 0.1-quantile and 8.3 per cent ($13.48 per bag) at the 0.9-quantile in the Core region. Wald test statistics reported in Table 6 suggest that the null hypotheses of equality between the HHI coefficient for 0.1- vs. 0.5-quantile, 0.5- vs. 0.9-quantile and 0.1- vs. 0.9-quantile are rejected in favour of the alternative in the Core region. It is not rejected in the Fringe region. Thus, market concentration effects in the lower and the upper tail of the price distribution differ only in the Core region. The 2SLS IV results find a positive and significant market concentration effect in both regions.[21] These results confirm that the FEIVQR method is a richer description of the distributional effects of competition.

Our results here are consistent with Result 2 developed in the theoretical predictions in Section 2.3. If the size of incumbent firms and number of loyal farmers in the market are sufficiently large, an increase in competition among firms with asymmetric capacity levels increases prices paid by corn growers. In the Fringe region, we find the effect is greater at the lower price quantiles, suggesting the presence of a larger number of loyal farmers. This result contrasts with the Core region, where the effect is greater at the upper quantiles.

Similar to the unconstrained sample, the coefficient corresponding to the product density variable is negative and significant in the Core regions. However, the coefficient is either statistically insignificant or marginally significant but negative in the Fringe regions. Price will decrease (weakly) when the market becomes crowded.

6.2.2. Farmer attributes

An increase in firms' loyal customer base may increase the price of seeds paid by price-sensitive farmers in the Fringe region: the coefficient is significant and positive at 0.1-quantiles. The coefficient for company-specific total number of transactions purchased by loyal customers (those occurred in August$_t$ to December$_t$) at the CRD level is also positive and significant across the price distribution in both regions but the magnitude is small.

The coefficient estimates corresponding to the share of loyal customers in a CRD suggest that it does not affect prices in the Fringe region but may have a negative impact on the Core region. The findings in the Fringe region confirm our theoretical prediction. Figures 3 and 4, described in Section 2.3, relate to the positive effect of an increase in the value of the loyal farmer parameter (L). The coefficient measuring the difference between payment time and order time is positive and significant in both regions, ranging between 0.02 per cent and 0.4 per cent.

21 The 2SLS IV estimates are larger than the OLS estimates in both regions, implying that the OLS estimates are downward biased (e.g. if pricing and knowledge of the demand side are negatively correlated).

6.2.3. Company attributes

Biotechnology companies charge a price premium at the mean and lower price quantiles in both regions. Evidence also strongly suggests that product availability and pricing are positively correlated. In the Fringe region, the term related to unconstrained firm is positive and significant in the lower tail of the price distribution, ranging from 9.3 per cent ($7.03 per bag) at the 0.1-quantile to 5.0 per cent ($5.32 per bag) at the 0.3-quantile. The mean effect is closer to the 0.3-quantile. Additionally, capacity-unconstrained firms belonging to vertically integrated biotechnology seed companies seem to charge less for product availability on average. The value is −7.4 per cent (−$9.71 per bag) and is closer to the lower tail of the price distribution. Wald test statistics reported in Table 6 suggest that the null hypothesis of equality across quantiles for parameters related to the unconstrained firm and its interaction with the biotechnology firm variable is rejected in favour of the alternative only at the 0.1- vs. 0.9-quantile.

In the Core region, the coefficient related to unconstrained firm is positive and statistically different across quantiles, ranging from 6.9 per cent ($5.72 per bag) at the 0.1-quantile to 3.6 per cent ($5.22 per bag) at the 0.7-quantile. The mean coefficient is closer to the median. However, vertically integrated capacity-unconstrained biotechnology firms charge less for product availability across the distribution function of prices. The coefficient ranges from −9.6 per cent (−$7.95 per bag) at the 0.1-quantile to −6.4 per cent (−$9.27 per bag) at the 0.7-quantile.

Table 6 suggests that the null hypothesis of equality between the unconstrained firm slope parameters 0.1- vs. 0.5-quantile, 0.5- vs. 0.9-quantile and 0.1- vs. 0.9-quantile is rejected in favour of the alternative with a p-value between 1 per cent and 10 per cent. The null hypothesis for the parameter corresponding to the interaction between unconstrained firm and biotechnology variables is rejected only for the 0.5- vs. 0.9-quantile and 0.1- vs. 0.9-quantile with p-value between 5 per cent and 10 per cent.

These findings confirm Result 3 in Section 2.3 of our theoretical model relating to the relationship between product availability and pricing. Farmers in the two regions are willing to pay higher prices for increased product availability. Capacity-unconstrained firms other than vertically integrated biotechnology seed giants are able to extract more surplus from price-sensitive farmers in the Fringe region and from all farmers in the Core region. Additionally, the biotechnology capacity-unconstrained firms do not extract additional profits from farmers located in the Fringe region and may offer a price reduction for farmers located in the Core region.

6.2.4. Seed attributes

Similar to the unconstr_sample, the estimates corresponding to different types of seeds suggest price premiums associated with the GM technology.

The coefficient of seeds with HLFS and/or HLES features is positive and significant only in the lower tail of the price distribution in the Fringe region, and negative and significant only in the upper tail in the Core region. These results indicate that price-sensitive farmers located in the Fringe region may find it profitable to pay a price premium for seeds featuring HLFS and/or HLES; loyal farmers in the Core region may demand a price reduction for these seeds. Our estimates also indicate seed companies may charge a premium to all types of farmers in both regions if the stated use of seed is for ethanol production. Note that the result in the Core region contrasts with the unconstr_sample findings, which predict significant negative price effects across the whole distribution. It may reflect differences in farmer connection with ethanol plants across regions. According to estimates from the U.S. Energy Information Administration, six states (Iowa, Nebraska, Illinois, Minnesota, Indiana and South Dakota) accounted for 72 per cent of ethanol production in the United States in 2016. Five of these six states are included in our Core region. Therefore, farmers in the Core region may find the production for ethanol plants more profitable than those in the fringe region.

Finally, the coefficient of new conventional seeds is negative along the price distribution, indicating that farmers in the two regions are in general averse to experimenting with new conventional seeds but not necessarily with the genetically modified varieties.

6.2.5. Other variables

Some statistically significant price differences arise across purchase sources as well. Compared to other purchasing sources, farmers may pay lower prices for seeds purchased directly from seed companies in both the Fringe and Core regions, or if they are also seed company dealers in the Fringe region. Additionally, our estimates indicate that seed companies competing in a market with firms of asymmetric capacity levels also engage in time-based price discrimination. Fringe and Core farmers may pay lower prices for seed if orders are placed earlier or later than January. Finally, we find the time trend effect *Year* positive across the entire distribution of prices and the sign and magnitude of the coefficients comparable with those in the unconstr_sample.

7. Conclusions

This study examines how firms price differently for given products and contributes to an understanding of the relationship between product availability and pricing. We propose a clearinghouse model of price dispersion to explain the roles that constraints in firm capacity and differences in consumer preferences play in forming temporal price dispersion for a homogeneous product. Simulation results are investigated empirically for the U.S. corn seed industry. The data provide farm-firm-level purchase information for conventional and genetically modified corn seeds sold by different firms between 2004 and 2009 in the Fringe and Core regions of the U.S. Corn Belt. The empirical model is estimated using the FEIVQR.

Our research findings yield several major conclusions. First, our model predicts a positive relationship between competition and pricing. Our empirical results in Core and Fringe regions confirm these predictions. Second, our model indicates that product availability and pricing move in the same direction. We find evidence that capacity-unconstrained firms charge a price premium in both regions. These results suggest that farmers in our samples are willing to pay a price premium to buy from a seed company that can guarantee seed availability later in the season. This finding is similar to those in the airline company or hoteling industry, where capacity constraint exists. Reservations with flexibility of changing time of travel will be more expensive than those with fixed dates of travel.

Third, we investigate whether an increase in the firms' loyal customer base leads to an increase in price along the distribution function. We found this pattern is highly significant when competition is among symmetrically capacity-unconstrained firms located in the Core region. Fourth, we find that seed companies engage in time-based price discrimination. Farmers in the Core region may pay lower prices if they order seeds early in the season, and farmers in the Fringe region benefit by placing orders earlier or later in the season.

Such effects on the distribution function of prices may be of concern to policymakers interested in developing antitrust and consumer protection laws or policy. For example, current antitrust laws are concerned that some mergers and acquisitions change the functioning of markets in ways that can lead to higher prices and other inefficiencies. However, for some industries, a new entrant may not be beneficial to consumers if firm capacity constraints and consumer brand loyalty play significant roles. The entry may actually induce implicit collusion with incumbent firms and harm all consumers if incumbent firms charge prices specially designed to attract surplus from the loyal customer base. Policies designed to prevent anticompetitive mergers and acquisitions may have unintended consequences if failing to account for these particularities. In recent years, a 'tsunami' of merger and acquisitions proposals between and within the already consolidated global biotechnology seed and agrochemical industries underwent regulatory scrutiny. In 2015, Monsanto launched a takeover bid to Syngenta with no success. In 2016, a new wave of mergers and acquisitions emerged: DuPont/Dow, ChemChina/Syngenta and Bayer/Monsanto. The U.S. Department of Justice, the U.S. Federal Trade Commission, the European Union and other countries with interest in the matter investigated these deals due to anticompetitive concerns. Our analysis may provide useful guidance in evaluating the potential impacts of such mergers on pricing dispersion and on how farmer groups are being affected.

Supplementary data

Supplementary data is available at *ERAE* online.

Funding

This material is based upon work supported by the United States Department of Agriculture - Agriculture and Food Research Initiative (USDA-AFRI) under Grant No. 2011-67023-30070, and the United States Department of Agriculture - National Institute of Food and Agriculture (USDA-NIFA) under Hatch Grant No. 142-PRJ37WQ.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies* 58: 277–297.

Baye, M. R. and Morgan, J. (2001). Information gatekeepers on the internet and the competitiveness of homogeneous product markets. *American Economic Review* 91: 454–474.

Baye, M. R., Morgan, J. and Scholten, P. (2006). Information, search and price dispersion. *Handbook on Economics and Information Systems* 1: 323–375.

Beckmann, M. (1965). Efgeworth-Bertrand duopoly revisited. *Operations Research-Verfahren* 3: Misenheim am Glan Hain 55–68.

Colin, C. A. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources* 50: 317–372.

Chernozhukov, V. and Hansen, C. (2008). Instrumental variable quantile regression: a robust inference approach. *Journal of Econometrics* 142: 379–398.

Davidson, C. and Deneckere, R. (1986). Long-run competition in capacity, short-run competition in price, and the Cournot model. *The Rand Journal of Economics* 17: 404–415.

Efron, B. and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. CRC press, Cleveland, OH.

Harding, M. and Lamarche, C. (2009). A quantile regression approach for estimating panel data models using instrumental variables. *Economics Letters* 104: 133–135.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis* 91: 74–89.

Koenker, R. and Bassett, G. Jr (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society* 46: 33–50.

Kreps, D. M. and Scheinkman, J. A. (1983). Quantity precommitment and Bertrand competition yield Cournot outcomes. *The Bell Journal of Economics* 14: 326–337.

MacMinn, R. D. (1980). Search and market equilibrium. *The Journal of Political Economy* 88: 308–327.

Narasimhan, C. (1988). Competitive promotional strategies. *Journal of Business* 61: 427–449.

Reinganum, J. F. (1979). A simple model of equilibrium price dispersion. *The Journal of Political Economy* 87: 851–858.

Rosenthal, R. W. (1980). A model in which an increase in the number of sellers leads to a higher price. *Econometrica: Journal of the Econometric Society* 48: 1575–1579.

Rothschild, M. (1973). Models of market organization with imperfect information: a survey. *The Journal of Political Economy* 81: 1283–1308.
Salop, S. and Stiglitz, J. (1977). Bargains and ripoffs: a model of monopolistically competitive price dispersion. *The Review of Economic Studies* 44: 493–510.
Shilony, Y. (1977). Mixed pricing in oligopoly. *Journal of Economic Theory* 14: 373–388.
Spulber, D. F. (1995). Bertrand competition when rivals' costs are unknown. *The Journal of Industrial Economics* 43: 1–11.
Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy* 69: 213–225.
Stock, J. H. and Watson, M. W. (2003). *Introduction to Econometrics*. Vol. 104, Boston: Addison Wesley.
Varian, H. R. (1980). A model of sales. *The American Economic Review* 70: 651–659.