

A Stochastic Multiple-Leader Stackelberg Model: Analysis, Computation, and Application

Victor DeMiguel

Department of Management Science and Operations, London Business School, Regent's Park,
London NW1 4SA, United Kingdom, avmiguel@london.edu

Huifu Xu

School of Mathematics, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom,
h.xu@maths.soton.ac.uk

We study an oligopoly consisting of M leaders and N followers that supply a homogeneous product (or service) noncooperatively. Leaders choose their supply levels first, knowing the demand function only in distribution. Followers make their decisions after observing the leader supply levels and the realized demand function. We term the resulting equilibrium a *stochastic multiple-leader Stackelberg-Nash-Cournot* (SMS) equilibrium. We show the existence and uniqueness of SMS equilibrium under mild assumptions. We also propose a computational approach to find the equilibrium based on the sample average approximation method and analyze its rate of convergence. Finally, we apply this framework to model competition in the telecommunication industry.

Subject classifications: programming; noncooperative games/group decisions; Stackelberg game; equilibrium existence, uniqueness; sample average approximation.

Area of review: Optimization.

History: Received October 2006; revisions received July 2007, February 2008; accepted September 2008.

Published online in *Articles in Advance* June 25, 2009.

1. Introduction

We consider an oligopoly consisting of $M + N$ firms that supply a homogeneous product (or service) noncooperatively. The first M firms (herein leaders) have no capacity installed and thus must make a decision *at the present time* on their *future* supply quantities to allow themselves enough time to build the necessary facilities. As a result, leaders must choose their supply quantities before the demand function is realized and thus know the demand function only in distribution.¹ The remaining N firms (followers) have sufficient capacity installed and thus do not have to make a decision today, but instead they can wait to observe the quantities supplied by the M leaders as well as the realized demand function before making a decision on their supply quantities.

Each of the M leaders maximizes its expected profit (where the expectation is taken with respect to the demand function distribution) by explicitly taking into account the future reaction of the N followers to the supply quantities of the leaders, but assuming that the rest of the leaders will keep their supply levels fixed; that is, leaders behave as Cournot firms with respect to other leaders, but as Stackelberg firms with respect to followers.

Once the demand function is realized, the N followers maximize their profit by assuming all leaders, and the rest of the followers keep their supply levels fixed; that is, followers behave as Cournot firms with respect to both

leaders and other followers. We call the resulting leader follower equilibrium a *stochastic multiple-leader Stackelberg-Nash-Cournot* (SMS) equilibrium.

We apply this framework to model a recurrent situation of competition in the telecommunication industry. Assume that a new service technology is about to enter a telecommunication market (for example, bandwidth). A few companies (leaders) do not have any spare network capacity and thus must decide whether to invest in expanding their network (or installing a new network) and offer the new service to the market. Because the capacity expansion process is time consuming, these companies have to decide the quantity they will supply to the market in advance, and thus know the demand function only in distribution. A few other companies (followers) already have sufficient capacity installed to offer this service, and they need only to decide how much of their capacity to re-allocate from other existing services to the new service technology. Thus, followers have the flexibility to wait until they observe the supply levels of the leaders as well as the realized demand function before deciding how much capacity to allocate to the new service.

Our model is related to a few models in the literature. The classical Stackelberg game (Stackelberg 1934) models competition in a duopoly market composed of one leader and one follower. The leader makes its decision taking into account the reaction of the follower. The follower, on the other hand, makes its decision assuming the leader

will keep its supply quantity fixed. Sherali et al. (1983) extended the classical Stackelberg model by considering the case where there are several followers that reach a Nash-Cournot equilibrium among themselves. Sherali (1984) further extended the model by considering multiple leaders—he gave conditions for the existence of equilibrium as well as for the concavity and piecewise linearity of the objective function of each leader. Sherali also showed the uniqueness of equilibrium for the case where leaders are identical (that is, leaders have identical cost functions) and, in addition, identical leaders make identical decisions. Ehrenmann (2004a, b) demonstrated that the assumption that identical leaders make identical decisions is essential to Sherali’s uniqueness result for the deterministic case. In particular, Ehrenmann (2004a, b) gives a counterexample that shows that identical leaders could reach multiple equilibria (in the deterministic case) if they can make different decisions.

De Wolf and Smeers (1997) made an interesting extension of the single-leader Stackelberg-Nash-Cournot model by incorporating demand uncertainty. Specifically, they assumed that the demand function follows a *discrete* distribution; that is, they consider a finite number of demand scenarios, each characterized by a different demand function. They applied the resulting model to the European gas market. More recently, Xu (2005) extended the stochastic *single-leader* model of De Wolf and Smeers by considering the case where the demand function follows a *continuous* distribution. Xu gives conditions under which there exists an equilibrium for the stochastic single-leader model and proposes a deterministic approach based on numerical integration to approximate the leader’s expected profit.

In this paper, we consider a stochastic multiple-leader Stackelberg-Nash-Cournot (SMS) equilibrium model. Our model extends Sherali’s *deterministic* multiple-leader model (Sherali 1984) because it considers the case where the demand function is stochastic. Our model also extends the work by De Wolf and Smeers (1997) and Xu (2005) because it focuses on the case with multiple leaders. Note that although the majority of the analysis in Xu (2005) focuses on the single-leader model, §6 in Xu (2005) discusses the case with *two leaders* and gives some preliminary analysis under the stringent assumption that every follower supplies a strictly positive quantity at equilibrium. In our analysis, we consider the case with multiple leaders (possibly more than two) and we do not assume that every follower supplies a positive quantity at equilibrium.

Our contribution to the literature is threefold. First, we show the existence and uniqueness of equilibrium for the stochastic model. A crucial result is that we prove, in contrast to the deterministic model, the objective function of each leader is smooth for the stochastic model. This smoothness result allows us to apply the results by Rosen (1965) to show the uniqueness of equilibrium for the stochastic model under conditions that are weaker than those assumed by Sherali (1984) for the deterministic case.

For example, we do not assume that leaders are identical. Also, in the case when leaders are identical, our uniqueness result does not require that identical leaders make identical decisions.

Second, we propose a computational approach to find the SMS equilibrium. Specifically, we use a sample average approximation method (SAA) to approximate the leader expected profits and show that the solutions to the SAA problems converge almost sure to the unique SMS equilibrium and with probability approaching one exponentially fast to an approximate SMS equilibrium. We then reformulate the SAA problems as mixed complementarity problems (using a technique similar to that in Hu 2002 and Leyffer and Munson 2005) and solve the resulting problems with the solver PATH (Ferris and Munson 1998, Ralph 1994). Note that an advantage of using the SAA method is that it allows one to approximate the leader expected profits by sampling directly from historical data. Hence, unlike the numerical integration method proposed in Xu (2005), our method does not require knowledge of the functional form of the demand function distribution.

Our third contribution is to apply this framework to model competition in the telecommunication industry. We use the proposed computational approach to find the resulting telecommunication market equilibrium. We also perform a comparative static analysis to study how economic variables such as the leader and follower supply levels and profits, the market price, the consumer surplus, and the social welfare depend on parameters such as the number of leaders and followers, the leader cost functions, and the variability in the demand function.

The rest of this paper is organized as follows. Section 2 gives a mathematical description of the model and reviews some preliminary results. In §3, we investigate the differentiability of the objective function of a leader. In §4, we show the existence and uniqueness of the SMS equilibrium. In §5, we propose a computational approach to find the SMS equilibrium based on the sample average approximation method and study its convergence. In §6, we apply this framework to model and analyze competition in the telecommunication industry. Section 7 concludes. In addition, the online supplement contains two appendices: Appendix A contains the proofs for some of the results in our paper and Appendix B contains the statements of two frequently used results from the literature. An electronic companion to this paper is available as part of the online version that can be found at <http://or.journal.informs.org/>.

2. The Model and Some Preliminary Results

2.1. The Model

The *market demand* is characterized by a random inverse demand function $p(q, \xi(\omega))$, where $p(q, \xi(\omega))$ is the market price, q is the total supply to the market, and $\xi: \Omega \rightarrow \mathbb{R}$ is a continuous random variable. Specifically, for each realization of the random variable $\xi: \Omega \rightarrow \mathbb{R}$, we obtain a

different inverse demand function $p(q, \xi(\omega))$. The uncertainty in the inverse demand function is then characterized by the distribution of the random variable ξ .

The i th leader's cost function is $C_i(x_i)$, where x_i is the i th leader's supply quantity. Likewise, the j th follower's cost function is $c_j(y_j)$, where y_j is the j th follower's supply quantity. Note that, because we assume that leaders have no capacity installed, leaders' costs will, in practice, be substantially larger than those of the followers.

We will use the following notation. X denotes the aggregate supply of all leaders ($X = \sum_{i=1}^M x_i$), X_{-i} the aggregate supply of all leaders excluding the i th leader ($X_{-i} = \sum_{k \neq i} x_k$), and x_{-i} the vector whose components are the supplies of all leaders excluding the i th leader ($x_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_M)$). Y , Y_{-j} , and y_{-j} denote homologous quantities for followers. Given a function $f(u, v): \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, we denote by $f(\cdot, v): \mathbb{R}^n \rightarrow \mathbb{R}$ the function obtained by fixing the value of v in $f(u, v)$.

The j th follower chooses its supply quantity after observing the aggregate leader supply X and the realized demand function $p(\cdot, \xi(\omega))$. Thus, the total revenue of the j th follower is $y_j p(X + y_j + Y_{-j}, \xi(\omega))$, and its total cost is $c_j(y_j)$. Consequently, the j th follower profit is

$$\psi_j(y_j, Y_{-j}, X, \xi(\omega)) = y_j p(X + y_j + Y_{-j}, \xi(\omega)) - c_j(y_j),$$

and the j th follower's decision problem is to choose the supply quantity y_j that maximizes its profit; that is,

$$\max_{y_j \geq 0} \psi_j(y_j, Y_{-j}, X, \xi(\omega)). \tag{1}$$

Note that when making its decision, the j th follower assumes all other followers and the leaders will keep their supply levels fixed. That is, the j th follower acts as a Cournot firm with respect to the rest of the followers and the leaders. A *Nash-Cournot equilibrium* among the followers (given an aggregate leader supply X and an inverse demand function realization $p(\cdot, \xi(\omega))$) is a situation where no follower can increase its profit by unilaterally changing its supply. We denote this equilibrium by $(y_1(X, \xi(\omega)), \dots, y_N(X, \xi(\omega)))$, where $y_j(X, \xi(\omega))$ is the global minimizer to problem (1). We also denote $Y(X, \xi(\omega)) = \sum_{j=1}^N y_j(X, \xi(\omega))$ and $Y_{-j}(X, \xi(\omega)) = \sum_{k \neq j} y_k(X, \xi(\omega))$. The existence and uniqueness of equilibrium among the followers are addressed in Proposition 2.6.

Leaders have no capacity installed and thus must choose their supply quantities well in advance to allow themselves enough time to install the necessary facilities. As a result, they choose their supply quantities knowing the demand function only in distribution. Also, because they make their decisions earlier than the followers, leaders explicitly take into account the reaction of the followers to the aggregate leader supply for each future demand function realization. But because all leaders make their decisions simultaneously, they choose their supply levels assuming that all other leaders will keep their supply quantities fixed.

In other words, leaders act as Stackelberg firms with respect to the followers, but as Cournot firms with respect to other leaders. Thus, the i th leader decision problem is

$$\begin{aligned} \max_{x_i \geq 0} \phi_i(x_i, X_{-i}) \\ := \mathbb{E}[x_i p(x_i + X_{-i} + Y(x_i + X_{-i}, \xi(\omega)), \xi(\omega))] - C_i(x_i), \end{aligned} \tag{2}$$

where \mathbb{E} denotes expectation with respect to the random variable ξ and $Y(X, \xi(\omega))$ is the aggregate follower supply given an aggregate leader supply of X and a realization of the random variable $\xi(\omega)$.

A *stochastic multiple-leader Stackelberg-Nash-Cournot equilibrium* is a situation where none of the leaders can improve its expected profit by changing its supply level while assuming that all other leaders keep their supply levels fixed, and for any given leader aggregate supply, the followers will reach a Nash-Cournot equilibrium in every demand scenario. We now give a rigorous definition of the equilibrium.

DEFINITION 2.1. A *stochastic multiple-leader Stackelberg-Nash-Cournot equilibrium* (SMS equilibrium) is an $M + N$ tuple $(x_1^*, \dots, x_M^*; y_1(X^*, \cdot), \dots, y_N(X^*, \cdot))$ such that

$$\phi_i(x_i^*, X_{-i}^*) = \max_{x_i \geq 0} \phi_i(x_i, X_{-i}^*), \quad i = 1, \dots, M, \tag{3}$$

$$\begin{aligned} y_j(X^*, \xi(\omega)) \in \arg \max_{y_j \geq 0} \psi_j(y_j, Y_{-j}(X^*, \xi(\omega)), X^*, \xi(\omega)), \\ j = 1, \dots, N, \end{aligned} \tag{4}$$

and $(y_1(X^*, \xi(\omega)), \dots, y_N(X^*, \xi(\omega)))$ is the Nash-Cournot equilibrium among the followers for aggregate leader supply X^* and realized inverse demand $p(q, \xi(\omega))$.

Herein, we assume that $\xi(\omega)$ is a continuous random variable with density function $\rho(t)$ and supporting set \mathcal{T} . Hence, the objective functions of leaders and followers can be rewritten as

$$\begin{aligned} \phi_i(x_i, X_{-i}) = \int_{\mathcal{T}} x_i p(x_i + X_{-i} + Y(x_i + X_{-i}, t), t) \rho(t) dt \\ - C_i(x_i), \quad i = 1, \dots, M, \end{aligned} \tag{5}$$

and

$$\begin{aligned} \psi_j(y_j, Y_{-j}, X, t) = y_j p(X + y_j + Y_{-j}, t) - c_j(y_j), \\ t \in \mathcal{T}, j = 1, \dots, N. \end{aligned} \tag{6}$$

2.2. Assumptions and Preliminary Results

In this section, we first state a number of standard assumptions from the literature, and then in Proposition 2.6 we show that, under these assumptions, there exists a unique follower Nash-Cournot equilibrium.

ASSUMPTION 2.2. The cost function of each follower ($c_j(q)$ for $j = 1, \dots, N$) is twice continuously differentiable and its derivatives satisfy $c'_j(q) \geq 0$ and $c''_j(q) \geq 0$ for $q \geq 0$.

This assumption requires that the cost function of each follower be nondecreasing, convex, and sufficiently smooth. This is also assumed in Sherali et al. (1983) and De Wolf and Smeers (1997).

ASSUMPTION 2.3. *The inverse demand function $p(q, t)$ satisfies the following:*

- (i) $p(q, t)$ is twice continuously differentiable in q and $p'_q(q, t) < 0$ for $q \geq 0$ and $t \in \mathcal{T}$;
- (ii) $p'_q(q, t) + qp''_{qq}(q, t) \leq 0$ for $q \geq 0$ and $t \in \mathcal{T}$.

Note that Assumption 2.3 holds, for example, for a random linear inverse demand function $p(q, t) = \alpha(t) - \beta(t)q$ with $\alpha(t): \mathcal{T} \rightarrow \mathbb{R}$ and $\beta(t): \mathcal{T} \rightarrow \mathbb{R}$ such that $\alpha(t), \beta(t) > 0$ for all $t \in \mathcal{T}$, where the condition $\alpha(t), \beta(t) > 0$ implies that when the total supply is zero, the price is positive, and that the inverse demand function is strictly decreasing in the total supply. It is easy to show, on the other hand, that Assumption 2.3 (ii) does not hold in general for the isoelastic demand function. The linear inverse demand function also satisfies all other assumptions required in our analysis. This, together with its tractability, makes the linear inverse demand function appropriate for our analysis. Note that a worrying feature of the linear inverse demand function is that, for each realization t of the random variable $\xi(\omega)$, the associated inverse demand function $p(q, t)$ might take negative values for large aggregate supply q . This, however, is not a serious concern in practice because leaders and followers are unlikely to choose supply quantities that may lead to negative prices under certain scenarios. In fact, in our numerical results the price is positive under every scenario for all equilibria we have computed. Finally, note that all previous works in the literature assume linear inverse demand functions with only additive noise in their analysis (Sherali 1984, De Wolf and Smeers 1997, Xu 2005); that is, they all assume that β is deterministic. Thus, our assumptions on the linear inverse demand function are weaker than those made previously in the literature.

Assumption 2.3 is similar to the assumptions used by Sherali et al. (1983) and De Wolf and Smeers (1997), with the main differences derived from our using a continuous distribution for the demand uncertainty. To better understand this assumption, consider a monopoly market with an extraneous supply $K \geq 0$. If the monopolist supply is q , then its revenue at demand realization $\xi(\omega) = t$ is $qp(q + K, t)$ and its marginal revenue is $p(q + K, t) + qp'_q(q + K, t)$. The rate of change of this marginal revenue with respect to an increase in the extraneous supply K is $p'_q(q + K, t) + qp''_{qq}(q + K, t)$. Assumption 2.3 (ii) implies that this rate is not positive when $K = 0$ for any $t \in \mathcal{T}$. In other words, any extraneous supply will potentially reduce the monopolist marginal revenue at any demand realization. See Sherali et al. (1983) for a similar explanation for the deterministic model. This assumption is used in Xu (2005) to prove the strict concavity of the revenue function. For convenience, we review this result here.

LEMMA 2.4 (XU 2005, PROPOSITION 2.4). *Under Assumption 2.3,*

- (i) for fixed $K \geq 0$,

$$p'_q(q + K, t) + qp''_{qq}(q + K, t) \leq 0 \quad \text{for } q \geq 0, t \in \mathcal{T}; \quad (7)$$

- (ii) $qp(q + K, t)$ is strictly concave in q for $q \geq 0, t \in \mathcal{T}$.

Note that the strict concavity of $qp(q + K, t)$ is not sufficient to ensure the boundedness of its maximizer for $K > 0$, which is needed to prove the existence of a follower Nash-Cournot equilibrium. To address this issue, we make one more assumption.

ASSUMPTION 2.5. *For each follower ($j = 1, \dots, N$), there exists a capacity limit y_j^u such that $c'_j(y_j) \geq p(y_j, t)$ for $y_j > y_j^u, t \in \mathcal{T}$.*

The assumption implies that even if the j th follower was a monopolist, its marginal cost above the capacity limit y_j^u would exceed any possible market price. Thus, none of the followers would optimally choose to supply a quantity beyond its capacity limit y_j^u . See a similar assumption in Sherali et al. (1983) and a discussion therein, and De Wolf and Smeers (1997). We are now ready to establish the existence and uniqueness of the follower Nash-Cournot equilibrium.

PROPOSITION 2.6. *Under Assumptions 2.2, 2.3, and 2.5, the following hold:*

- (i) For $j = 1, \dots, N$, $\psi_j(y_j, Y_{-j}, X, t)$ is concave in y_j for y_j in $[0, y_j^u]$, and it has a unique global maximizer with respect to y_j .
- (ii) The follower Nash-Cournot equilibrium problem (4) can be reformulated as a complementarity problem

$$0 \leq y \perp G(X, y, t) \geq 0, \quad t \in \mathcal{T},$$

where

$$G(X, y, t) := -p(X + Y, t)\mathbf{e} - p'_q(X + Y, t)y + \nabla c(y),$$

$\mathbf{e} = (1, \dots, 1)^T$, and $\nabla c(y) := (c'_1(y_1), \dots, c'_N(y_N))^T$. Moreover, for fixed $X \geq 0$ and $t \in \mathcal{T}$, the complementarity problem has a unique solution $(y_1(X, t), \dots, y_N(X, t))$, which is a follower Nash-Cournot equilibrium. Furthermore, $y_j(X, t) \in [0, y_j^u]$ for $j = 1, \dots, N$.

- (iii) For $t \in \mathcal{T}$ and $X \geq 0$, $(y_1(X, t), \dots, y_N(X, t))$ is piecewise smooth with respect to X .

PROOF. Observe that the decision problems of the followers depend only on the aggregate supply of the leaders. From the followers' perspective, X is a constant regardless of whether it comes from the decision of a single leader or from those of several leaders. Consequently, the result follows straightforwardly from Propositions 2.4, 2.6, and 3.3 in Xu (2005), where a stochastic single-leader Stackelberg-Nash-Cournot equilibrium model is considered. \square

A couple of remarks on Proposition 2.6 are in order. First of all, note that the result follows in a straightforward manner from the analysis in Xu (2005) regarding the model with a single leader. In Xu (2005), the follower Nash-Cournot equilibrium problem was reformulated as a nonlinear complementarity problem, and the properties of the equilibrium were investigated by studying the properties of the solution to the complementarity problem. This approach to analyze the problem differs from the approaches used in the literature (Sherali 1984, De Wolf and Smeers 1997), although reformulating an oligopoly game as a complementarity problem is a well-known technique (Ferris and Pang 1997). Second, from part (iii) of the proposition, it is clear that $Y(\cdot, t)$ is also piecewise smooth. At a point X where $Y(\cdot, t)$ is not differentiable, we consider the Clarke subdifferential (Clarke 1983). Recall that for a locally Lipschitz continuous function $h: \mathbb{R}^n \rightarrow \mathbb{R}$, the Clarke subdifferential of h at $z \in \mathbb{R}^n$ is

$$\partial h(z) := \text{conv} \left\{ \lim_{z_k \in D_h, z_k \rightarrow z} \nabla h(z_k) \right\},$$

where conv denotes the convex hull and D_h denotes the set of points near z at which h is differentiable. A simple example is $h(z) = |z|$, which is differentiable everywhere except at $z = 0$. Using the definition of the Clarke subdifferential, we obtain $\partial h(0) = \text{conv}\{-1, 1\} = [-1, 1]$. In our context, because $Y(\cdot, t)$ is piecewise smooth, then $\partial_X Y(X, t)$ is no more than the convex hull of its left and right derivatives at point X , which is a closed interval.

3. Differentiability of the Leader Expected Profit

In this section, we show that the objective function of each leader (that is, its expected profit) is twice continuously differentiable with respect to its own supply level. In the deterministic model, the objective function of each leader (that is, its profit) is not continuously differentiable everywhere (Sherali 1984). The nondifferentiability at a certain supply level x_i means that the marginal profit changes discontinuously when the leader's supply reaches the level x_i . This discontinuity occurs at aggregate leader supply levels for which the marginal profit of one of the followers turns to zero. For aggregate leader supply levels above this point, the aforementioned follower optimally decides to exit the market, and this generates a discontinuity in the leader profit. Needless to say, a leader must take this nondifferentiability into account when making decisions. In this section, we show that for the stochastic model, follower exits at some realizations of the demand function may not be significant enough to cause the nondifferentiability of the leader expected profit.

We focus on the differentiability with respect to x_i of the i th leader objective function

$$\phi_i(x_i, X_{-i}) := \int_{\mathcal{T}} [x_i p(x_i + X_{-i} + Y(x_i + X_{-i}, t), t)] \cdot \rho(t) dt - C_i(x_i),$$

where $Y(x_i + X_{-i}, t) = \sum_{j=1}^N y_j(x_i + X_{-i}, t)$ and $y(x_i + X_{-i}, t)$ is the followers' equilibrium; that is, the solution to the complementarity problem $0 \leq y \perp G(x_i + X_{-i}, y, t) \geq 0$ with $t \in \mathcal{T}$. From the definition of $\phi_i(x_i, X_{-i})$ and our assumptions on $C_i(x_i)$ and $p(q, t)$, we observe that the only possible source of nondifferentiability is the function $Y(x_i + X_{-i}, t)$, which is defined implicitly as the solution to a complementarity problem. Note that at a point X where $Y(X, t)$ is differentiable, we have by the chain rule that $Y'_x(x_i + X_{-i}, t) = Y'_X(x_i + X_{-i}, t)$. Also, at a point X where $Y(X, t)$ is not differentiable, we have that $\partial Y_{x_i}(x_i + X_{-i}, t) = \partial Y_X(x_i + X_{-i}, t)$. Thus, to characterize the differentiability of $Y(x_i + X_{-i}, t)$ with respect to x_i , it is enough to study the differentiability of the function $Y(X, t)$ with respect to X . In this section, we show that $Y(X, t)$ is twice piecewise continuously differentiable with respect to X , and that for every value of X there are only a finite number of values of t for which $Y(X, t)$ is nondifferentiable. Because $Y(X, t) = \sum_{j=1}^N y_j(X, t)$, it is enough to characterize the properties of $y_j(X, t)$ for $j = 1, \dots, N$.

We show in Lemma 3.3 that $y_j(X, t)$ is piecewise twice continuously differentiable and nonincreasing in X as well as globally Lipschitz continuous under certain conditions. First, we state two assumptions that we use to prove the result.

ASSUMPTION 3.1. *The inverse demand function and the leader and follower cost functions satisfy the following:*

(i) $c_j(y_j)$, $j = 1, \dots, N$, is twice continuously differentiable on $[0, y_j^u]$ and there exists a positive constant $\sigma > 0$ such that for $j = 1, \dots, N$,

$$-p'_q(q, t) + c''_j(y_j) \geq \sigma \quad \forall q \geq 0, t \in \mathcal{T}, y_j \in [0, y_j^u].$$

(ii) *There exists a function $L_1(t) > 0$ with $\sup_{t \in \mathcal{T}} L_1(t) < \infty$ such that*

$$\max(|p'_q(q, t)|, |p''_{qq}(q, t)|) < L_1(t) \quad \forall q \geq 0, t \in \mathcal{T}.$$

Note that Assumption 3.1 (i) and (ii) hold, for example, for convex follower cost functions and random linear inverse demand function $p(q, t) = \alpha(t) - \beta(t)q$, where $\alpha(t): \mathcal{T} \rightarrow \mathbb{R}$ and $\beta(t): \mathcal{T} \rightarrow \mathbb{R}$ are such that $\alpha(t), \beta(t) > 0$ for all $t \in \mathcal{T}$ and $\sup_{t \in \mathcal{T}} \beta(t) < \infty$. The condition that $\sup_{t \in \mathcal{T}} \beta(t) < \infty$ holds whenever β is a continuous function and t has bounded support set \mathcal{T} .

ASSUMPTION 3.2. *For each leader ($i = 1, \dots, M$), there exists a capacity limit x_i^u such that $C'_i(x_i) \geq p(x_i, t)$ for $x_i > x_i^u$, $t \in \mathcal{T}$.*

This assumption implies that there is an implicit capacity limit for the leaders. The assumption is reasonable from a practical perspective because every real-world firm has a limited capacity as a result of either the economic limitations of its production technology or of regulatory restrictions. We will use this assumption also in our discussion on existence and uniqueness of SMS equilibrium in the next section.

LEMMA 3.3. *Under Assumptions 2.2, 2.3, 2.5, 3.1, and 3.2, the following hold:*

- (i) $\partial_X Y(X, t) \subset (-1, 0]$ for $X \geq 0$.
- (ii) $y_j(X, t)$, $j = 1, \dots, N$, is piecewise twice continuously differentiable and nonincreasing in X for $X \geq 0$.
- (iii) For almost every $t \in \mathcal{T}$, $y_j(X, t)$ is globally Lipschitz continuous with respect to X with an integrable module; that is, there exists a function $L_2(t)$ such that

$$|y_j(X^{(1)}, t) - y_j(X^{(2)}, t)| \leq L_2(t) |X^{(1)} - X^{(2)}|$$

$$\forall X^{(1)}, X^{(2)} \geq 0,$$

where $\int_{\mathcal{T}} L_2(t) \rho(t) dt < \infty$.

The proof of Lemma 3.3 is given in Appendix A of the online supplement. A few comments are in order. Note that at a point X where $Y(X, t)$ is differentiable with respect to X , $\partial_X Y(X, t) = Y'_X(X, t)$. Thus, for convenience herein we will use $\partial_X Y(X, t)$ to denote both the derivative and Clarke's subdifferential. Also, note that from part (ii) of Lemma 3.3 we know that $y_j(X, t)$ is nonincreasing in X , and thus there exists a unique point, denoted by $X_j(t)$, at which $y_j(X_j(t), t)$ turns from strictly positive to zero as X increases. In economic terms, this is the aggregate leader supply for which the j th follower marginal profit becomes zero. Mathematically, $X_j(t)$ is a *degenerate point* of the complementarity problem (20) because at this point, both $G_j(X_j(t))$ and $y_j(X_j(t), t)$ are zero. Note that $y_j(X, t)$ will not be differentiable with respect to X at such a point if its left derivative is different from zero. The fact that there exists a unique point $X_j(t)$ at which $y_j(X_j(t), t)$ turns from positive to zero will be instrumental to prove that the leader expected profit is twice continuously differentiable.

We are now ready to prove the differentiability of the leader expected profit. We need the following assumption.

ASSUMPTION 3.4. *The following conditions hold:*

- (i) *There exists $L_3(t) \geq 0$ such that $\int_{t \in \mathcal{T}} L_3(t) \rho(t) dt < \infty$ and*

$$\max(|p(q, t)|, |p'''_{qq}(q, t)|, |Y''_{XX}(X, t)|) \leq L_3(t)$$

for all q such that $0 \leq q \leq \sum_{i=1}^M x_i^u + \sum_{j=1}^N y_j^u$, $t \in \mathcal{T}$, and X such that $Y''_{XX}(X, t)$ exists.

- (ii) $p(q, t)$ is twice continuously differentiable with respect to q and t , and there exists a $\delta > 0$ such that $p'_q(q, t) > \delta$ and $p''_{qt}(q, t) \geq 0$ for $t \in \mathcal{T}$ and q such that $0 \leq q \leq \sum_{i=1}^M x_i^u + \sum_{j=1}^N y_j^u$.

Note that Assumption 3.4 (i) regarding the inverse demand function is satisfied, for example, by a random linear inverse demand function $p(q, t) = \alpha(t) - \beta(t)q$, where $\alpha(t): \mathcal{T} \rightarrow \mathbb{R}$ and $\beta(t): \mathcal{T} \rightarrow \mathbb{R}$ are such that $\alpha(t), \beta(t) > 0$ for all $t \in \mathcal{T}$ and $\sup_{t \in \mathcal{T}} \beta(t) < \infty$. Assumption 3.4 (ii) implies that the linear inverse demand is strictly increasing in t and that its cross-partial derivative is nonnegative. This is satisfied, for example, by a random linear inverse

demand function with $\alpha'(t) > \delta$ and $\beta'(t) \leq 0$; that is, by a random linear inverse demand function whose intercept is strictly increasing in t , with first derivative bounded below, and whose slope is decreasing in absolute value in t .

THEOREM 3.5. *Suppose that Assumptions 2.2, 2.3, 2.5, 3.1, 3.2, and 3.4 hold. Then, $\phi_i(x_i, X_{-i})$ is twice continuously differentiable with respect to x_i for $x_i \in [0, x_i^u]$.*

PROOF. We divide the proof in two steps: First, we show that $\phi_i(x_i, X_{-i})$ is once continuously differentiable with respect to x_i , and second, we show that its second derivative is also continuous.

Step 1. $\phi_i(x_i, X_{-i})$ is once continuously differentiable. To show this result, we use Ruszczyński and Shapiro (2003), Proposition 2, which for the reader's convenience we have restated in Appendix B of the online supplement as Lemma B.1. From Lemma B.1, it follows that it suffices to show that the function

$$v(x_i, t) := x_i p(x_i + X_{-i} + Y(x_i + X_{-i}, t), t)$$

satisfies the following two conditions. First, $v(\cdot, t)$ is differentiable at x_i for almost every t , and second, $v(\cdot, t)$ is globally Lipschitz continuous with an integrable module; that is, there exists a function $L_4(t) > 0$ such that

$$\int_{t \in \mathcal{T}} L_4(t) \rho(t) dt < \infty \tag{8}$$

and

$$|v(x_i^{(1)}, t) - v(x_i^{(2)}, t)| \leq L_4(t) |x_i^{(1)} - x_i^{(2)}|$$

$$\forall x_i^{(1)}, x_i^{(2)} \in [0, x_i^u]. \tag{9}$$

We prove these conditions in two separate steps.

Step 1.1. *The function $v(\cdot, t)$ is differentiable at x_i for almost every t .* Note that, for given X_{-i} and x_i , the only values of t at which $v(\cdot, t)$ might not be differentiable with respect to x_i are points t at which the supply of one of the followers $y_j(X, t)$ turns from being strictly positive to being zero. These are the only points at which $Y(X, t)$ might not be differentiable with respect to X and thus $v(\cdot, t)$ might not be differentiable with respect to x_i . We now prove that for given values of X_{-i} and x_i , there exists only a finite number of points t at which the supply of one of the followers turns from being strictly positive to zero.

Let $X_j(t)$ be the value of X at which the supply of the j th follower turns from being strictly positive to zero. Then, it suffices to show that $X_j(t)$ is strictly increasing in t . To see this, note that if $X_j(t)$ is strictly increasing in t , we then have that for each value of X there are at most N different values of t for which the supply of one of the followers turns from being strictly positive to zero, where N is the number of followers.

We now show that $X_j(t)$ is strictly increasing in t . To see this, first note that by the definition of $X_j(t)$, we have that $y_j(X_j(t) - \epsilon, t) > 0$ for small positive ϵ . Then, because

by Assumption 3.4 (ii) we have that $p'_i > 0$ and $p''_{qt} \geq 0$, it follows from Xu (2005), Proposition 3.4, part (iii) that $y_j(X, t)$ is strictly increasing in t for $X = X_j(t) - \epsilon$ provided that $\epsilon > 0$ is small enough.² Moreover, because $p'_i > \delta > 0$ and $p''_{qt} \geq 0$, it follows from Xu (2005), Proposition 3.4, part (iii) that the subdifferential of $y_j(X, t)$ with respect to t is bounded below by a constant positive number. Also, we know that $y_j(X, t)$ is continuous because, by Assumptions 2.3 and 3.4, the random inverse demand function is continuously differentiable in both X and t , and hence $y_j(X, t)$ (the solution to the parametric complementarity problem $0 \leq y \perp G(x_i + X_{-i}, y, t) \geq 0$) is continuous in both X and t . Hence, we have by continuity that $X_j(t)$ must be strictly increasing in t .

Step 1.2. $v(\cdot, t)$ is globally Lipschitz continuous with an integrable module. We have shown in Step 1.1 that $v(\cdot, t)$ is differentiable at x_i for almost every t . At a point where Y is differentiable with respect to X , by the chain rule and the fact that $X'_i = 1$, we have that

$$v'_{x_i}(x_i, t) = p(x_i + X_{-i} + Y(x_i + X_{-i}, t), t) + x_i p'_q(x_i + X_{-i} + Y(x_i + X_{-i}, t), t) \cdot (1 + Y'_X(x_i + X_{-i}, t)).$$

By Assumption 3.4, $p(q, t)$ is bounded by $L_3(t)$; and by Assumption 3.2, x_i is bounded by x_i^u . By Assumption 3.1 (ii), p'_q is bounded by $L_1(t)$; and by Lemma 3.3 (i), $1 + Y'_X$ is bounded by 1. Therefore, $v'_{x_i}(x_i, t)$ is bounded by $L_3(t) + x_i^u L_1(t)$. Let $L_4(t) := L_3(t) + x_i^u L_1(t)$. Then, $L_4(t)$ satisfies (8). By the mean value theorem,

$$|v(x_i^{(1)}, t) - v(x_i^{(2)}, t)| \leq \int_0^1 |v'_{x_i}(x_i^{(2)} + \theta(x_i^{(1)} - x_i^{(2)}, t))| |x_i^{(1)} - x_i^{(2)}| d\theta \leq L_4(t) |x_i^{(1)} - x_i^{(2)}|.$$

This shows (9). From Lemma B.1, it follows that

$$(\phi_i)'_{x_i}(x_i, X_{-i}) = \int_{t \in \mathcal{T}} v'_{x_i}(x_i, t) \rho(t) dt - C'_i(x_i).$$

Step 2. The second derivative of $\phi_i(x_i, X_{-i})$ is also continuous. We first show that the second derivative exists, and then we show that it is continuous.

Step 2.1. The second derivative of $\phi_i(x_i, X_{-i})$ exists. To show this, we again apply Ruszczyński and Shapiro (2003), Proposition 2, which for the reader's convenience we have restated in Appendix B of the online supplement as Lemma B.1. It suffices to show that $v'_{x_i}(\cdot, t)$ is differentiable with respect to x_i for almost every $t \in \mathcal{T}$, and there exists $L_5(t) > 0$ such that

$$\int_{t \in \mathcal{T}} L_5(t) \rho(t) dt < \infty$$

and

$$|v'_{x_i}(x_i^{(1)}, t) - v'_{x_i}(x_i^{(2)}, t)| \leq L_5(t) |x_i^{(1)} - x_i^{(2)}| \quad \forall x_i^{(1)}, x_i^{(2)} \in [0, x_i^u].$$

At a point where Y'_X is differentiable with respect to X , we have

$$v''_{x_i x_i}(x_i, t) = 2p'_q(X + Y(X, t), t)(1 + Y'_X(X, t), t) + x_i p''_{qq}(X + Y(X, t), t)(1 + Y'_X(X, t), t) + x_i p'_q(X + Y(X, t), t) Y''_{XX}(X, t).$$

Under the assumptions of this theorem, it is easy to derive that

$$|v''_{x_i x_i}(x_i, t)| \leq L_5(t) := 2L_1(t) + x_i^u L_1(t) + x_i^u L_1(t) L_2(t).$$

Because $L_1(t)$ is bounded and $L_2(t)$ is integrable, then $L_5(t)$ is integrable. The rest follows from the mean value theorem and Lemma B.1. This shows that $\phi_i(x_i, X_{-i})$ is twice differentiable and

$$(\phi_i)''_{x_i x_i}(x_i, X_{-i}) = \int_{\mathcal{T}} [2p'_q(X + Y(X, t), t)(1 + Y'_X(X, t), t) + x_i p''_{qq}(X + Y(X, t), t)(1 + Y'_X(X, t), t) + x_i p'_q(X + Y(X, t), t) Y''_{XX}(X, t)] \rho(t) dt - C''_i(x_i).$$

Step 2.2. The second derivative of $\phi_i(x_i, X_{-i})$ is continuous. To show the continuity of $(\phi_i)''_{x_i x_i}(x_i, X_{-i})$, we note that $v'_{x_i}(\cdot, t)$ is continuous on $[0, x_i^u]$ for almost every $t \in \mathcal{T}$ and $v''_{x_i x_i}(x_i, t)$ is dominated by $L_5(t)$, which is integrable. By the Lebesgue dominated convergence theorem,

$$\begin{aligned} \lim_{z \rightarrow x_i} (\phi_i)''_{x_i x_i}(z, X_{-i}) &= \int_{\mathcal{T}} \lim_{z \rightarrow x_i} v''_{x_i x_i}(z, t) \rho(t) dt - C''_i(x_i) \\ &= \int_{\mathcal{T}} v''_{x_i x_i}(x_i, t) \rho(t) dt - C''_i(x_i) \\ &= (\phi_i)''_{x_i x_i}(x_i, X_{-i}). \end{aligned}$$

This completes the proof. \square

4. Existence and Uniqueness of Equilibrium

It is well known that a multiple-leader-follower Nash game might not have an equilibrium; see, for example, Pang and Fukushima (2005). Sherali (1984), however, showed the existence and uniqueness of equilibrium for the deterministic multiple-leader Stackelberg-Nash-Cournot model under certain conditions. In this section, we prove the existence and uniqueness of SMS equilibrium under conditions that are either similar to or weaker than those used by Sherali (1984).

4.1. Existence

We start by showing that the expected profit function of a leader is concave. Under certain conditions, Sherali (1984) demonstrated the strict concavity of the profit of each leader for the deterministic game. Here we show that this property also holds for the stochastic case under similar or weaker conditions. We then move on to investigate the existence of SMS equilibrium.

THEOREM 4.1. *Assume that $Y(X, t)$ is convex with respect to X for all $t \in \mathcal{T}$. Assume also that Assumptions 2.2, 2.3, 2.5, and 3.2 hold and that $C_i(x_i)$ is convex on $[0, x_i^u]$ for $i = 1, \dots, M$. Then, $\phi_i(x_i, X_{-i})$ is concave in x_i for $x_i \in [0, x_i^u]$ and $i = 1, \dots, N$.*

PROOF. Under Assumptions 2.2, 2.3, 2.5, and 3.2, it follows from Sherali’s work for the deterministic model (Sherali 1984, Lemma 1) that the function $x_i p(x_i + X_{-i} + Y(x_i + X_{-i}, t), t)$ is concave with respect to x_i for every $t \in \mathcal{T}$. Integrating this function with respect to t preserves this concavity. Thus, $\int_{\mathcal{T}} [x_i p(x_i + X_{-i} + Y(x_i + X_{-i}, t))] \rho(t) dt$ is concave in x_i . Moreover, because $C_i(x_i)$ is assumed to be convex, then we have that $\phi_i(x_i, X_{-i})$ is concave in x_i . \square

Note that to show the concavity of the expected profit of a leader, we have assumed that the follower aggregate reaction function $Y(X, t)$ is convex in X . This assumption does not hold in general, but in the following proposition we show that it holds under certain conditions such as linearity of the inverse demand function and either convex quadratic or symmetric follower cost functions. Sherali showed similar results for the deterministic model (see Sherali 1984, Theorem 3, and its corollary), but our results are stronger in two senses. First, we allow for a general (multiplicative) random linear inverse demand function $p(q, t) = \alpha(t) - \beta(t)q$, while Sherali’s results hold only for the additive random linear inverse demand function $p(q, t) = \alpha - \beta q + t$. Second, we allow for the follower cost functions c_j to be quadratic and possibly different for each of the followers, while Sherali’s results hold only for either linear or symmetric cost functions.

PROPOSITION 4.2. *Let $p(q, t)$ be a random linear inverse demand function $p(q, t) = \alpha(t) - \beta(t)q$, where $\alpha(t): \mathcal{T} \rightarrow \mathbb{R}$ and $\beta(t): \mathcal{T} \rightarrow \mathbb{R}$ are such that $\alpha(t), \beta(t) > 0$ for all $t \in \mathcal{T}$, let $c_j(y_j)$, $j = 1, \dots, N$, be convex functions, and let Assumptions 2.2, 2.3, 2.5, 3.1, and 3.2 hold. Then, $Y(X, t)$ is convex for $X \geq 0$ if one of the following conditions holds:*

(i) *The follower cost functions ($c_j(y_j)$, $j = 1, \dots, N$) are quadratic and convex ($c_j''(y_j) \geq 0$) and possibly different for each follower.*

(ii) *The follower cost functions are convex and identical (that is, $c_j(y_j) = c(y_j)$ for $j = 1, \dots, N$ with $c''(y) \geq 0$) and $c'''(y) \leq 0$.*

The proof of Proposition 4.2 is given in Appendix A in the online supplement. Now that we have established the concavity of the expected profit of each leader, we are ready to show the existence of SMS equilibrium.

THEOREM 4.3. *Suppose that the assumptions in Theorem 4.1 hold. Then, there exists an SMS equilibrium.*

PROOF. The existence follows from Rosen (1965), Theorem 1 because the expected profit of leader i ($\phi_i(x_i, X_{-i})$), $i = 1, \dots, M$, is concave in x_i for $x_i \in [0, x_i^u]$. \square

4.2. Uniqueness

To show uniqueness, we apply Theorem 2 in the paper by Rosen (1965). Let the i th leader expected profit as a function of its own supply x_i and the aggregate supply of the rest of the leaders X_{-i} be $\phi_i(x_i, X_{-i})$. Define the function

$$\sigma(x, r) := \sum_{i=1}^M r_i \phi_i(x_i, X_{-i}),$$

where $x = (x_1, x_2, \dots, x_M)$ and $r = (r_1, r_2, \dots, r_M)$. According to Theorem 2 in Rosen (1965), if $\sigma(x, r)$ is diagonally strictly concave for some $r > 0$, then the leaders equilibrium $x^* = (x_1^*, x_2^*, \dots, x_M^*)$ is unique. Moreover, Theorem 6 in Rosen (1965), gives a sufficient condition for $\sigma(x, r)$ to be diagonally strictly concave. Concretely, Rosen shows that if the Jacobian $\nabla_x g(x, r)$ of the function

$$g(x, r) := \begin{pmatrix} r_1 (\phi_1)_{x_1}'(x_1, X_{-1}) \\ \vdots \\ r_M (\phi_M)_{x_M}'(x_M, X_{-M}) \end{pmatrix}$$

is negative definite for all $x \in [0, x_1^u] \times [0, x_2^u] \times \dots \times [0, x_M^u]$, then $\sigma(x, r)$ is diagonally strictly concave, and thus the leader’s equilibrium is unique.

In the following theorem, we give conditions under which the Jacobian matrix $\nabla_x g(x, r)$ corresponding to our model is indeed negative definite, and thus there exists a unique leader equilibrium.

We note that applying the uniqueness analysis in §3 of Rosen (1965) to our model requires the functions $\phi_i(x_i, X_{-i})$ to be continuously differentiable with respect to x_i —this assumption is stated in Rosen (1965, p. 523). Hence, our smoothness analysis in §3 is what allows us to apply Rosen’s results.

THEOREM 4.4. *Let the assumptions of Theorems 3.5 and 4.1 hold. Define*

$$\zeta_i := \int_{t \in \mathcal{T}} [p_q'(1 + Y_X') + x_i p_{qq}''(1 + Y_X')^2 + x_i p_q' Y_X''] \rho(t) dt$$

and

$$\delta_i = \left(\int_{t \in \mathcal{T}} p_q'(1 + Y_X') \rho(t) dt - C_i''(x_i) \right).$$

Suppose that there exists a constant vector $r = (r_1, r_2, \dots, r_M)$ such that $r > 0$,

$$\sum_{i=1}^M r_i \zeta_i < 0, \tag{10}$$

and

$$\sum_{i=1}^M r_i \zeta_i + \sqrt{M \sum_{i=1}^M r_i^2 \zeta_i^2} + 2 \max_{i=1, \dots, M} r_i \delta_i < 0. \tag{11}$$

Then, there exists a unique SMS equilibrium.

The proof of Theorem 4.4 is given in Appendix A in the online supplement. Note that Sherali (1984) established his uniqueness result for the deterministic game under three conditions (among others): (a) the demand function is linear with a deterministic parameter β , (b) leaders are identical (that is, they have identical cost functions), and (c) identical leaders make identical decisions. Our uniqueness result requires weaker conditions because we allow β in the linear demand to depend on the random shock $\xi(\omega)$, and we do not require conditions (b) or (c).

To see that our conditions are weaker than Sherali’s, we first show that if conditions (a), (b), and (c) hold, then conditions (10) and (11) hold. To see this, note that condition (10) holds (with any $r > 0$) for the random linear inverse demand function $p(q, t) = \alpha(t) - \beta(t)q$ with $\alpha(t), \beta(t) > 0$ for all $t \in \mathcal{T}$. In particular, observe that $p'_q(1 + Y'_X) + x_i p''_{qq}(1 + Y'_X)^2 + x_i p'_q Y''_X < 0$ for all x_i, X_{-i} , and t because $p'_q < 0, p''_{qq} = 0, 1 + Y'_X > 0$ by Lemma 3.3, and $Y''_X > 0$ by Proposition 4.2. This gives $\zeta_i < 0$ for all i , and hence (10) holds. Note that this also implies that condition (10) holds for the linear inverse demand function with deterministic β . Also, condition (11) holds when identical leaders make identical decisions because $x_1 = x_2 = \dots = x_M$, and hence $\zeta_1 = \zeta_2 = \dots = \zeta_M < 0$; consequently $\sum_{i=1}^M \zeta_i + \sqrt{M \sum_{i=1}^M \zeta_i^2} = 0$. Moreover, for the linear demand function, we have that $\delta_i < 0$ because $p'_q < 0, 1 + Y'_X > 0$ by Lemma 3.3, and $C'' \geq 0$ by the convexity of the leader cost functions. This shows that Sherali’s conditions (a), (b), and (c) imply (10) and (11). Finally, note that conditions (10) and (11) might also hold in general for the cases in which the leader cost functions are not identical, or for the case where the leaders have identical cost functions, but they might make different decisions. To see this, simply note that conditions (10) and (11) do not require in general δ_i or ζ_i to be identical for $i = 1, \dots, M$.

5. Computing Equilibria

To compute the SMS equilibrium, our first step is to deal with the expected value in the leader objective functions. To do this, we propose a sample average approximation (SAA) method to approximate the leader expected profits and analyze the convergence of the solutions to the SAA problems as the sample size increases.

For a fixed sample, the SAA problem is a deterministic equilibrium program with equilibrium constraints (EPEC). Our second step is to solve this deterministic EPEC. We use a popular method initially proposed by Hu (2002) and used by many others (e.g., Hu and Ralph 2007, Ehrenmann 2004b, Leyffer and Munson 2005), which reformulates the deterministic EPEC as a mixed nonlinear

complementarity problem, and then we use the commercial solver PATH (Ferris and Munson 1998, Ralph 1994) to solve this mixed complementarity problem.

5.1. Sample Average Approximations

SAA is a popular method in stochastic programming; see Robinson (1996), Gürkan et al. (1999), and Shapiro (2003) and the references therein. The basic idea behind the SAA method is to approximate the expected value function by the sample average of the random function. An advantage of the SAA approach is that it requires only a sample of $\xi(\omega)$ rather than the distribution of $\xi(\omega)$ as in other discretization methods. Samples of $\xi(\omega)$ could be obtained, for example, from historical data or computer-based simulations without the need to know the exact distribution of the random variable $\xi(\omega)$.

Let ξ^1, \dots, ξ^k be an independent identically distributed (i.i.d) sample of $\xi(\omega)$. We approximate the i th leader’s decision problem by the following SAA problem:

$$\begin{aligned} \min \quad & \phi_i^k(x_i, X_{-i}) := \frac{1}{k} \sum_{l=1}^k x_i p(X + Y(X, \xi^l), \xi^l) - C_i(x_i) \\ \text{s.t.} \quad & x_i \in [0, x_i^u]. \end{aligned} \tag{12}$$

If (x_1^k, \dots, x_M^k) satisfies $\phi_i^k(x_i^k, X_{-i}^k) = \max_{x_i \in [0, x_i^u]} \phi_i^k(x_i, X_{-i}^k)$ for $i = 1, \dots, M$, then we say (x_1^k, \dots, x_M^k) is a *stochastic multiple-leader Stackelberg-Nash-Cournot equilibrium of the SAA problem* (SAA-SMS equilibrium, for brevity). It is not difficult to prove that under the conditions of Theorem 4.3, an SAA-SMS equilibrium exists.

Our focus here is to investigate the convergence of a sequence of SAA-SMS equilibria as the sample size k increases. To simplify notation, let

$$\begin{aligned} f_i(x_1, \dots, z_i, \dots, x_M, \xi(\omega)) & := z_i p(z_i + X_{-i} + Y(z_i + X_{-i}, \xi(\omega)), \xi(\omega)) - C_i(z_i), \\ F(x, z, \xi(\omega)) & := \sum_{i=1}^M f_i(x_1, \dots, z_i, \dots, x_M, \xi(\omega)), \end{aligned}$$

and

$$\Phi(x, z) := \mathbb{E}[F(x, z, \xi(\omega))].$$

We first note that $x^* = (x_1^*, \dots, x_M^*)$ is an SMS equilibrium if and only if

$$\Phi(x^*, x^*) = \max_{z \in \mathcal{W}} \Phi(x^*, z),$$

where $\mathcal{W} := [0, x_1^u] \times \dots \times [0, x_M^u]$. If the conditions of Theorem 4.1 hold, then the function $f_i(x_1, \dots, z_i, \dots, x_M, \xi(\omega))$ is concave with respect to z_i . Moreover, if the conditions

of Theorem 3.5 hold, then $\Phi(x^*, z)$ is differentiable at x^* . Therefore, x^* is an equilibrium if and only if the first-order optimality conditions hold:

$$0 \in \nabla_z \Phi(x^*, x^*) + \mathcal{N}_{\mathcal{W}}(x^*), \quad (13)$$

where $\mathcal{N}_{\mathcal{W}}(x^*)$ denotes the normal cone of \mathcal{W} at x^* ; that is, $\mathcal{N}_{\mathcal{W}}(x^*) := \{u \in \mathbb{R}^M : u^T(x' - x^*) \leq 0 \ \forall x' \in \mathcal{W}\}$.

From a numerical point of view, it might be difficult and/or unnecessary to obtain an exact SMS equilibrium. Hence, we are also interested in an approximate equilibrium defined as follows:

$$0 \in \nabla_z \Phi(x^k, x^k) + \mathcal{N}_{\mathcal{W}}(x^k) + \eta \mathcal{B}, \quad (14)$$

where η is a small positive number and \mathcal{B} denotes the unit ball in \mathbb{R}^M ; that is, $\mathcal{B} = \{x \in \mathbb{R}^M \text{ such that } \|x\| = 1\}$.

We now turn our attention to the SAA problem (12). Let

$$\Phi_k(x, z) := \frac{1}{k} \sum_{l=1}^k F(x, z, \xi^l).$$

Then, $x^k = (x_1^k, \dots, x_M^k)$ is an SAA-SMS equilibrium if and only if

$$\Phi_k(x^k, x^k) = \max_{z \in \mathcal{W}} \Phi_k(x^k, z). \quad (15)$$

If the conditions of Theorem 4.1 hold, then $f_i(x_1, \dots, z_i, \dots, x_M, \xi(\omega))$ is concave with respect to z_i . Moreover, if the conditions of Lemma 3.3 hold, then f_i is piecewise twice continuously differentiable with respect to z_i . Then, x^k is an SAA-SMS equilibrium if and only if

$$0 \in \partial_z \Phi_k(x^k, x^k) + \mathcal{N}_{\mathcal{W}}(x^k), \quad (16)$$

where \mathcal{N} denotes the normal cone of \mathcal{W} at x^k and $\partial_z F_k$ denotes the subdifferential of the concave function F_k with respect to z ; that is, if $h(x)$ is a concave function, then $\partial h(y)$ is the set of vectors u such that

$$u^T(z - y) \geq h(z) - h(y) \quad \forall z.$$

Note that the reason we have to consider subdifferential is that $\Phi_k(x^k, z)$ might not be differentiable at x^k .

Theorem 5.1 gives two results regarding the convergence of $\{x^k\}$ as the sample size k tends to ∞ . First, $\{x^k\}$ converges to x^* w.p.1. Second, $\{x^k\}$ converges at an exponential rate to an approximate equilibrium satisfying (14). The proof is given in Appendix A in the online supplement.

THEOREM 5.1. *Suppose that the conditions in Theorem 4.4 hold and $\xi(\omega)$ has a bounded support set. If $x^k = (x_1^k, \dots, x_M^k)$ is an SMS equilibrium of the SAA problem (12), then:*

(i) *with probability one the sequence $\{x^k\}$ converges to the unique SMS equilibrium x^* ,*

(ii) *with probability approaching one exponentially fast with the increase of the sample size, the sequence $\{x^k\}$ converges to an approximate SMS equilibrium satisfying (14).*

5.2. Solving the Sample Average Approximation Problem

The sample average approximation problem for the i th leader is

$$\max_{x_i} \frac{1}{k} \sum_{l=1}^k x_i p(X + Y(X, \xi^l), \xi^l) - C_i(x_i) \quad (17)$$

s.t. $x_i \geq 0$,

where $Y(X, \xi^l)$ is defined implicitly as the equilibrium for the follower Nash-Cournot game given the aggregate supply of all leaders X and demand scenario ξ^l . Note that under Assumption 3.2 the maximizer of problem (17) lies in the interval $[0, x_i^u]$. Thus, to simplify notation, we do not explicitly include an upper bound x_i^u for the variable x_i in problem (17). Moreover, from Proposition 2.6 we know that $Y(X, \xi^l) = \mathbf{e}^T y^l(X, \xi^l)$, where \mathbf{e} is the N -dimensional vector of ones and $y^l(X, \xi^l)$ is the solution to the nonlinear complementarity problem $0 \leq y^l \perp G(X, y^l, \xi^l) \geq 0$, and $G(X, y, \xi^l) = -p(X + Y, t)\mathbf{e} - p'_q(X + Y, t)y + \nabla c(y)$. Thus, the i th leader problem can be written as the following mathematical program with equilibrium constraints:

$$\max_{x_i, y^l} \frac{1}{k} \sum_{l=1}^k x_i p(X + Y^l, \xi^l) - C_i(x_i) \quad (18)$$

s.t. $0 \leq y^l \perp G(X, y^l, \xi^l) \geq 0 \quad \forall l$,

$x_i \geq 0$.

Moreover, this mathematical program with equilibrium constraints can be reformulated as the following standard nonlinear program:

$$\max_{x_i, y^l} \frac{1}{k} \sum_{l=1}^k x_i p(X + Y^l, \xi^l) - C_i(x_i) \quad (19)$$

s.t. $y^l \geq 0 \quad \forall l$,

$G(X, y^l, \xi^l) \geq 0 \quad \forall l$,

$-y^l \circ G(X, y^l, \xi^l) \geq 0 \quad \forall l$,

$x_i \geq 0$,

where \circ represents the Hadamard (componentwise) product. The KKT optimality conditions for this problem are:

$$\frac{1}{k} \sum_{l=1}^k [p(X + Y^l, \xi^l) + x_i p'_q(X + Y^l, \xi^l)] - C'_i(x_i)$$

$$+ \sum_{l=1}^k \lambda_{il}^T (G'_X)_l - \sum_{l=1}^k \nu_{il}^T y^l \circ (G'_X)_l + \pi_i = 0,$$

$$\frac{1}{k} x_i p'_q(X + Y^l, \xi^l) \mathbf{e} + (G'_y)^T \lambda_{il} + \mu_i$$

$$- \nu_{il} [G + y^l \circ (G'_y)]_l = 0 \quad \forall l,$$

$$0 \leq G(X, y^l, \xi^l) \perp \lambda_{il} \geq 0 \quad \forall l,$$

$$0 \leq y^l \perp \mu_{il} \geq 0 \quad \forall l,$$

$$0 \leq -y^l \circ G(X, y^l, \xi^l) \perp \nu_{il} \geq 0 \quad \forall l,$$

$$0 \leq x_i \perp \pi_i \geq 0 \quad \forall i,$$

where $(G'_x)_l = G'_x(X, y^l, \xi^l)$ and $(G'_y)_l = G'_y(X, y^l, \xi^l)$. To find an SMS equilibrium for the SAA problem, it is enough then to compile these KKT conditions for all M leaders and then solve the resulting mixed nonlinear complementarity problem using a solver such as PATH (Ralph 1994, Ferris and Munson 1998). This methodology is similar to that proposed by Hu (2002) and used by others for solving deterministic equilibrium problems with equilibrium constraints (Hu and Ralph 2007, Ehrenmann 2004b, Leyffer and Munson 2005).

6. Modeling Competition in the Telecommunication Industry

We apply the proposed framework to model a recurrent situation of competition in telecommunication markets. We use the proposed computational approach to find equilibria for this market. Finally, we perform a comparative static analysis to study how economic variables such as the leader and follower supply levels and profits, the market price, the consumer surplus, and the social welfare depend on parameters such as the number of leaders and followers, the leader cost functions, and the uncertainty in the demand function.

This section is divided into three parts. In the first part, we describe the competition model. In the second part, we discuss the parameters used for our computations. Finally, in the third part, we discuss the numerical results.

6.1. Model

Assume that a new service technology is about to enter a telecommunication market (for example, bandwidth). A few companies (leaders) have no spare network capacity and thus must decide whether to invest on expanding their network capacity (or install a new network) and offer the new service to the market. Because the network expansion process is time consuming, these companies have to make a decision in advance and thus know the demand function for the new technology only in distribution. A few other companies (followers) already have sufficient network capacity installed to offer this technology and they just need to decide how much of their capacity to re-allocate from other existing services to the new technology. Thus, followers have the flexibility to wait until they observe the leaders supply decision as well as the realized demand function before deciding how much of their network capacity to allocate to the new service. This situation can be seen as one of a few companies that are established in the telecommunication market (incumbents), and a few other companies that are considering entering this market (entrants) to offer the new service technology but need to install their network capacity.³

This situation can be modeled as the stochastic multiple-leader Stackelberg-Nash-Cournot model that we have analyzed in this paper. We assume a random linear inverse demand function $p(q, \xi) = \alpha(\xi) - \beta(\xi)q$, where ξ is a random variable. The follower cost functions are assumed

to be quadratic, $c_j(y_j) = \delta_j y_j + (\gamma_j/2)y_j^2$, where y_j is the network capacity allocated by the j th follower (incumbent) to the new technology. This cost represents the opportunity cost of allocating capacity to the new service technology as opposed to previously existing technologies. The leader cost functions are also quadratic, $C_i(x_i) = \delta_i x_i + (\gamma_i/2)x_i^2$, where x_i is the capacity of the network installed by the i th leader (entrant). The leaders' costs represent the installation cost associated with the new capacity and thus are much larger, in general, than follower costs.

6.2. Parameterization

For simplicity of exposition, we assume a symmetric game. We consider a *base case* with the following values of the different model parameters. We assume that there are two leaders (entrants) and two followers (incumbents). The coefficients for the cost functions of the leaders are $\delta_i = 0.4$ and $\gamma_i = 0.4$, and for the cost functions of the followers are $\delta_j = 0.25$ and $\gamma_j = 0.25$. This implies the total cost of the leaders, which includes installation cost, is substantially larger than that of the followers. For the linear inverse demand function, we assume that the intercept is $\alpha(\xi) = 1.5 - \xi$ and the negative slope is $\beta = 1.0$. That is, for the base case we assume that all the uncertainty in the demand function arises from the intercept α . We assume that the random variable ξ follows a truncated normal distribution with zero mean, standard deviation of 0.25, and is truncated at two standard deviations above and below the mean. For the base case, we solve a sample average approximation problem with 200 samples.

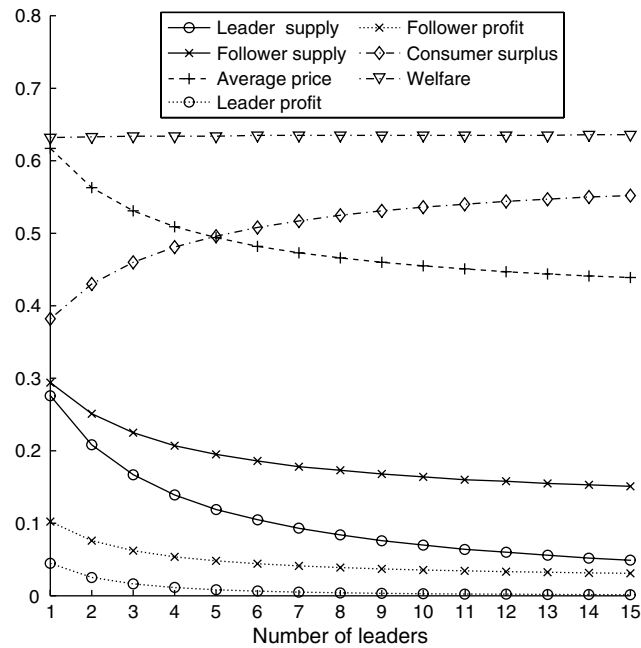
We then perform comparative static analysis with respect to the number of leaders (M ranging from 1 to 15), the number of followers (N ranging from 1 to 15), the leader cost coefficients (δ_i equal to γ_i and ranging from 0.4 to 0.8), and the uncertainty in the demand function. Finally, to illustrate the convergence properties of the SAA methodology, we solve problems with a number of samples ranging from 5 to 500.

We must note that the quadratic cost functions and the linear inverse demand function we use for the base case and for all of our comparative static analysis satisfy the assumptions made for our analysis in previous sections. Thus, our theoretical results guarantee that there exists a unique SMS equilibrium and that the solutions to the SAA problems converge to the true unique equilibrium.

6.3. Numerical Results

We first solve the case corresponding to the base case parameters given in §6.2. For the base case, the supply by each leader is 0.208 and the average supply of each follower is 0.251. Also, each leader obtains an average profit of 0.0253, whereas each follower obtains an average profit of 0.0760.⁴ That is, for the base case, on average, the leaders supply less and make less profit than the followers. The reason for this is twofold: Leaders lack perfect information

Figure 1. Comparative static analysis with respect to the number of leaders.



Notes. The horizontal axis gives the number of leaders ranging from 1 to 15. The vertical axis gives the leader production (circles and solid line), average follower production (crosses and solid line), average price (pluses and dashed line), average leader profit (circles and dotted line), average follower profit (crosses and dotted line), average consumer surplus (diamonds and dashed dotted line), and average welfare (triangles and dashed dotted line).

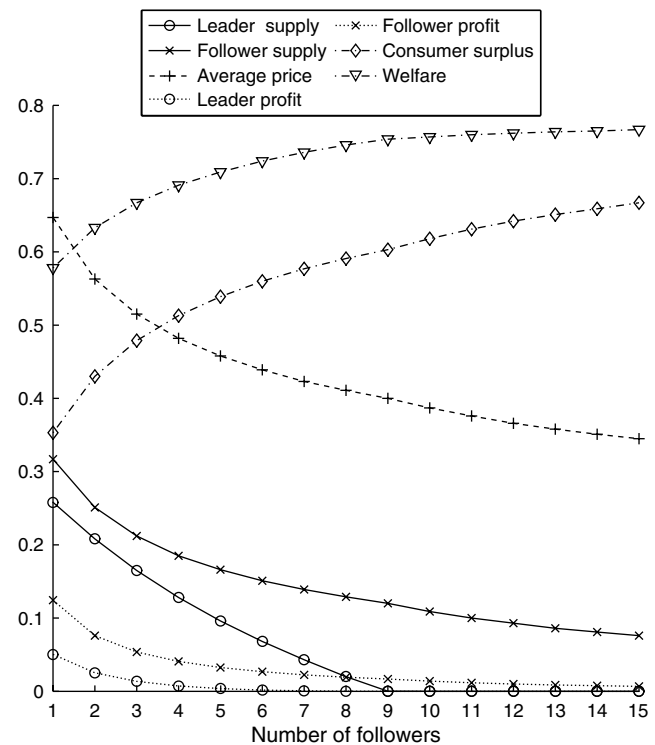
on demand at the time when they make their decisions, and they incur higher costs than the followers (because they have to install their network capacity).

We perform comparative static analysis with respect to the number of leaders, the number of followers, the leader cost coefficients, and the uncertainty in the demand function. For each of these parameters, we set all other parameters to their base case values and then study how the equilibrium changes when we vary the parameter under study.

The results are given in Figures 1–3. Each figure depicts the different values of the parameter under study on the horizontal axis. The vertical axis depicts six different equilibrium variables: the supply of each leader (circles and solid line), the average supply of each follower (crosses and solid line), the average price (pluses and dashed line), the average profit of each leader (circles and dotted line), the average profit of each follower (crosses and dotted line), the average consumer surplus⁵ (diamonds and dashed dotted line), and the average welfare⁶ (triangles and dashed dotted line).

Figure 1 depicts how the market equilibrium changes when the number of firms potentially interested in entering the market to offer the new service technology (i.e., the number of leaders) increases from 1 to 15. Note that the case with one leader corresponds to the stochastic single-leader Stackelberg-Nash-Cournot model proposed by

Figure 2. Comparative static analysis with respect to the number of followers.

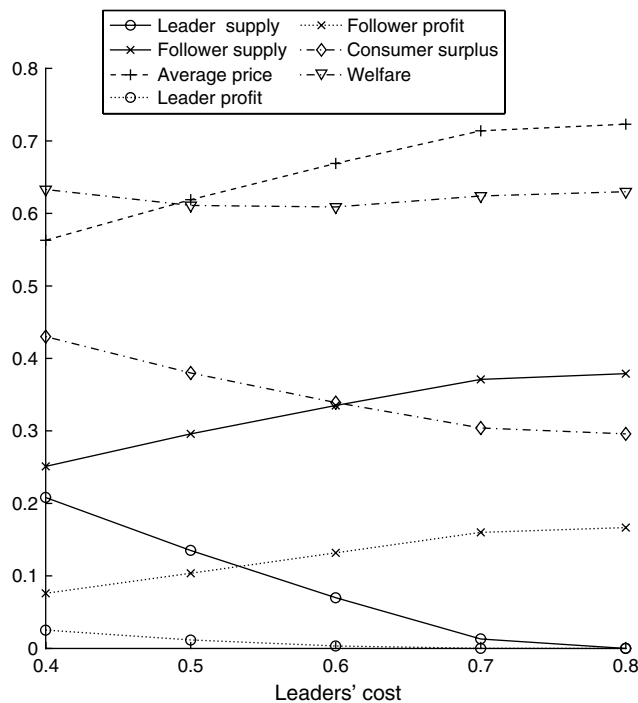


Notes. The horizontal axis gives the number of followers ranging from 1 to 15. The vertical axis gives the leader production (circles and solid line), average follower production (crosses and solid line), average price (pluses and dashed line), average leader profit (circles and dotted line), average follower profit (crosses and dotted line), average consumer surplus (diamonds and dashed dotted line), and average welfare (triangles and dashed dotted line).

De Wolf and Smeers (1997) and studied by Xu (2005). From the figure, we observe that when additional leaders enter the market, the average supply of each of the two followers decreases. Also, the supply of each of the leaders decreases, but their total supply increases (because there are more leaders in the market).⁷ As a result, the average total supply of all firms (leaders plus followers) increases, and consequently, the average market price decreases. Also, the profit of each firm (leader or follower) decreases with the presence of additional leaders, and the average consumer surplus increases. Summarizing, the presence of additional leaders (potential entrants) in the market increases competition and as a result the price for the new service technology decreases, leader and follower profits decrease, and consumer surplus increases.

Figure 2 gives the results for the case where the number of followers increases from 1 to 15. Note that the average supply of each follower and leader decreases with the presence of additional followers. Moreover, the two leaders decide not to supply the new service technology whenever there are nine or more followers. Nonetheless, the average total supply of all firms (leaders plus followers) increases when there are more followers in the market, and thus the average

Figure 3. Comparative static analysis with respect to the leaders' cost.



Notes. The horizontal axis gives the leaders' unit variable cost (δ_i equal to γ_i and ranging from 0.4 to 0.8). The vertical axis gives the leader production (circles and solid line), average follower production (crosses and solid line), average price (pluses and dashed line), average leader profit (circles and dotted line), average follower profit (crosses and dotted line), average consumer surplus (diamonds and dashed dotted line), and average welfare (triangles and dashed dotted line).

price decreases. As a result, consumer surplus and social welfare increase sharply with the number of followers.

It is also interesting to compare the results for the case where the number of leaders increases (Figure 1) to those for the case where the number of followers increases (Figure 2). In both cases, the presence of additional firms in the market (leaders or followers) increases competition and thus leads to lower prices, higher consumer surpluses, and lower profits for each individual firm. The figures, however, demonstrate that increasing the number of followers has a larger impact in competition than increasing the number of leaders. To see this, note that when the number of followers increases the price cuts are more substantial than when the number of leaders increases.

There are two reasons why followers are more competitive than leaders. First, followers have their capacity already installed and thus face lower costs than the leaders. Second, followers observe the realized demand function before they decide their supply quantities. Leaders, on the other hand, only know the demand function distribution when they make their decisions. A consequence of the higher competitiveness of followers is that when there are nine or more followers in the market, leaders do not find it profitable to enter the market at all, and thus they decide

to stay out of the market. When the number of leaders increases up to 15, on the other hand, followers decide to stay in the market.

Note that our model is the first in the literature that allows a sensible comparison of the impact that increasing the number of leaders and followers has on competition. Concretely, unlike previous models in the literature (Sherali et al. 1983, Sherali 1984, De Wolf and Smeers 1997), our model considers both multiple leaders and stochastic demand. These two aspects are necessary to carry out a fair comparison of the impact on competition of increasing the number of leaders and followers.

Figure 3 depicts how the equilibrium changes when the leader cost coefficients increase (δ_i equal to γ_i and ranging from 0.4 to 0.8). This represents an increase in the capacity installation costs to supply the new service technology. As these costs increase, the supply of each leader decreases, and leaders decide not to enter the market for δ_i and γ_i above 0.7. The average supply of each follower, on the other hand, increases. But the average total supply decreases and, as a result, the average price increases substantially. Thus, while the profit of each leader decreases, the profit of each follower increases substantially. The consumer surplus decreases dramatically. Clearly, when the leaders' costs are very high, the leaders do not find it profitable to supply high quantities (or even to enter the market), and thus the followers increase their profits by exercising their market power and driving prices up.

We now perform comparative static analysis with respect to the uncertainty in the demand function. We first consider the case where, as in the base case, all the uncertainty in the demand function originates from the intercept ($\alpha(\xi) = 1.5 - \xi$ and $\beta = 1.0$), but we study the effect of changing the standard deviation of the random variable σ_ξ from 0 to 0.5.⁸ Note that the case with zero standard deviation corresponds to the deterministic multiple-leader Stackelberg-Nash-Cournot model of Sherali (1984).⁹ We study how the value of the stochastic solution (that is, the percentage profit loss incurred by a leader who ignores uncertainty) changes when the standard deviation of ξ ranges from 0 to 0.5.¹⁰ We observe that, as expected, the value of the stochastic solution is zero for the case where demand is in fact deterministic, $\sigma_\xi = 0$. But the value of the stochastic solution grows with σ_ξ and is 4.93% for $\sigma_\xi = 0.5$. Thus, for the range of standard deviations considered, a leader might lose up to 5% of its profits by ignoring demand.

From our comparative static analysis, we also observe that market prices vary substantially for different realizations of the demand function. For example, the price for the deterministic case $\sigma_\xi = 0$ is 0.567, but for the case with $\sigma_\xi = 0.25$, market prices range between 0.386 and 0.762, and for the case with $\sigma_\xi = 0.5$, market prices range between 0.142 and 0.959. Likewise, the market share of the two followers is 70% for the deterministic case $\sigma_\xi = 0$, whereas it ranges between 51% and 80% for the case with $\sigma_\xi = 0.25$, and between 0% and 85% for the case with $\sigma_\xi = 0.5$. Our

results confirm that ignoring uncertainty can lead to large losses for the leaders and that the market structure might change dramatically for different realizations of the demand function.

We have also performed comparative static analysis to check whether the insights discussed above are also robust with respect to origin (intercept or slope) of the uncertainty in the demand function. Specifically, we have considered situations where the uncertainty in the demand function arises from both the intercept and the slope. To do this, we have considered the cases where $\alpha(\xi) = 1.5 - \tau\xi$ and $\beta = 1.0 - (1.01 - \tau)\xi$, with $\sigma_\xi = 0.25$ and for values of $\tau = 0.01, 0.26, 0.51, 0.76, 1.01$.¹¹ Note that the case with $\tau = 1.01$ is almost identical to our base case; that is, in this case, almost all the uncertainty is associated with the intercept of the demand function. For the case with $\tau = 0.01$, on the other hand, virtually all the uncertainty is associated with the demand function slope.¹² Other values of τ correspond to situations where part of the uncertainty originates from the intercept and part from the slope. Overall, we have kept a similar “amount” of uncertainty by using the parameter τ to split the effect of the random variable ξ between the intercept and the slope.

Our results (not reported in any figure, for brevity) show that the main insights from our analysis in this section are robust to whether the uncertainty originates from the demand function intercept or slope. Concretely, we have observed that most average equilibrium quantities (average leader and follower supply levels and profits, average market price, average consumer surplus, and average social welfare) are relatively similar in equilibrium regardless of whether the uncertainty arises from the demand function intercept or slope. We must note, however, that the source of the demand function uncertainty (intercept or slope) does make a difference to the equilibrium quantities for different demand function realizations. For example, we have observed that for the case with $\tau = 1.01$ (i.e., for our base case), market prices range between 0.386 and 0.762, whereas for the case with $\tau = 0.01$, market prices range only between 0.480 and 0.694. This is because intercept variability affects prices equally at all supply levels, whereas slope variability affects prices more substantially for larger aggregate supply levels. Hence, although most of the insights discussed in this section are robust to the source of the uncertainty (intercept or slope), the equilibrium quantities and, more specifically, the equilibrium quantities across different demand function realizations, differ substantially depending on the source of the uncertainty (intercept or slope).

Finally, we also study how the SAA solutions change when the number of samples ranges from 5 to 500. We observe that the SAA solutions converge to the true equilibrium at a fast rate. In particular, for the number of samples below 150, there is some degree of variability in the solutions to the SAA problems, but the solutions to the SAA problems with the number of samples on or above 200 are practically identical.

7. Conclusion

We have studied the stochastic multiple-leader Stackelberg-Nash-Cournot model, which is an extension of Sherali's deterministic multiple-leader model (Sherali 1984) and De Wolf and Smeers' stochastic single-leader model (De Wolf and Smeers 1997). We have shown the existence and uniqueness of equilibrium under conditions that are either similar or weaker than those in the literature (Sherali 1984). We have also shown the twice continuous differentiability of the leader objective function—a result that contrasts with the existing results for the deterministic model.

In addition, we have proposed a computational approach to find equilibria based on the SAA method, and we have shown that the solution to the SAA problems converges to an approximate equilibrium to the true problem at an exponential rate. The exponential convergence of the SAA method for solving stochastic optimization problems has been recently shown by Shapiro and Homem-de-Mello (2000) and Shapiro (2003). But to the best of our knowledge, our result is the first extension of these convergence results to the case where the SAA method is applied to a class of stochastic equilibrium programs with equilibrium constraints.

Finally, we have applied the proposed model to investigate competition in the telecommunication industry. One interesting economic insight from the comparative static analysis is that the followers have an important competitive advantage over the leaders because they incur substantially lower costs and have perfect information on the demand function. As a result, a sufficient number of followers in the market might prevent leaders from entering the market, whereas followers stay in the market even in situations where there is a very large number of leaders.

8. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at <http://or.journal.informs.org/>.

The following references are cited only in the electronic companion: Artstein (1974), Artstein and Vitale (1975), Murphy et al. (1982), Papageorgiou (1985), Patriksson and Wynter (1999), Ralph and Xu (2005), Rockafellar and Wets (1998), Rubinstein and Shapiro (1993), and Shapiro and Xu (2008).

Endnotes

1. For computational purposes, we do not require that leaders know the demand function distribution; we require only that leaders can *sample* the demand function distribution.
2. Note that the analysis in Xu (2005), Proposition 3.4, part (iii) is made under the assumption $p''_{qt} = 0$, but it is easy to see that $y_j(X, t)$ is increasing in t even for the case where $p''_{qt} \geq 0$. Also, Xu (2005), Proposition 3.4, part (iii) is proven under the assumption that at least one of the

followers supplies a strictly positive quantity, but this is definitely the case at points $X = X_j(t) - \epsilon$ with $\epsilon > 0$ small enough.

3. We apply the term “leaders” to the entrants because they *lead* the decision-making process. Concretely, the entrants have to make their supply quantity decisions *today*. Incumbents, on the other hand, can wait to observe the entrants’ decisions and the realized demand function before making their decisions, and thus they make their decisions at a later stage than the entrants. In this sense, the entrants lead the decision-making process.

4. Averages are taken with respect to the samples used in the SAA problem, with 200 samples in the base case.

5. The consumer surplus per unit of product is the difference between the consumer willingness to pay and the price paid.

6. The social welfare is the total firm profit plus the consumer surplus.

7. Although total leader (follower) supply is not shown in the figure, it can be inferred from the figure by multiplying the supply of a single leader (follower) by the number of leaders (followers), because our numerical example is a symmetric game.

8. That is, ξ follows a truncated normal distribution with zero mean, standard deviations of $\sigma_\xi = 0.00$ – 0.50 , and truncated at two standard deviations above and below the mean.

9. Sherali (1984) gives an analytical solution to the symmetric deterministic multiple-leader Stackelberg-Nash-Cournot model. For our numerical experiments, we consider the symmetric *stochastic* multiple-leader Stackelberg-Nash-Cournot model, which, to the best of our knowledge, can not be solved analytically.

10. To compute the value of the stochastic solution, we first solve the problem with deterministic demand $\sigma_\xi = 0$ and thus determine the supply level of a leader for the deterministic case x_{det} . We then compute the followers’ Nash-Cournot equilibrium for an aggregate leader supply of $X = Mx_{det}$ but with $\sigma_\xi > 0$. From this follower equilibrium, we compute the corresponding market prices and the resulting average profit of a leader who ignores uncertainty and thus supplies x_{det} . This profit is then compared to the optimal average leader profit obtained by the leaders when they take uncertainty into account. For a discussion of value of the stochastic solution, see Birge and Louveaux (1997).

11. Note that because $\tau > 0$, $\alpha'(\xi) > 0$ for all cases considered. All other assumptions in our analysis in the previous sections are also satisfied by the demand functions considered.

12. This case can also be interpreted as having uncertainty in the demand elasticity because it is easy to see that the smaller the value of β in the *inverse* demand function, the larger the demand elasticity.

Acknowledgments

The authors thank the area and associate editors and two referees for many helpful comments and suggestions that helped improve this manuscript substantially. They also gratefully acknowledge comments from Elodie Adida, Houyuan Jiang, Danny Ralph, Stefan Scholtes, and seminar participants at the Judge Business School at the University of Cambridge, the Johnson School at Cornell University, the Stern School of Business at New York University, the University of Chicago Graduate School of Business, the 7th Annual INFORMS Revenue Management and Pricing Section Conference, and the 2007 and 2008 INFORMS annual meetings. This research was partially supported by the United Kingdom Engineering and Physical Sciences Research Council grant GR/S90850/01.

References

- Artstein, Z. 1974. On the calculus of closed set-valued functions. *Indiana Univ. Math. J.* **24** 433–441.
- Artstein, Z., R. A. Vitale. 1975. A strong law of large numbers for random compact sets. *Ann. Probab.* **3** 879–882.
- Birge, J. R., F. Louveaux. 1997. *Introduction to Stochastic Programming*. Springer-Verlag, New York.
- Clarke, F. H. 1983. *Optimization and Nonsmooth Analysis*. Wiley, New York.
- De Wolf, D., Y. Smeers. 1997. A stochastic version of a Stackelberg-Nash-Cournot equilibrium model. *Management Sci.* **43** 190–197.
- Ehrenmann, A. 2004a. Manifolds of multi-leader Cournot equilibria. *Oper. Res. Lett.* **32** 121–125.
- Ehrenmann, A. 2004b. Equilibrium problems with equilibrium constraints and their applications in electricity markets. Ph.D. dissertation, Judge Institute of Management, Cambridge University, Cambridge, UK.
- Ferris, M. C., T. Munson. 1998. GAMS solver manual for the PATH solver. <http://www.gams.com/dd/docs/solvers/path.pdf>.
- Ferris, M., J. S. Pang. 1997. Engineering and economic applications of complementarity problems. *SIAM Rev.* **39** 669–713.
- Gürkan, G., A. Y. Özge, S. M. Robinson. 1999. Sample-path solution of stochastic variational inequalities. *Math. Programming* **84** 313–333.
- Hiriart-Urruty, J. B., C. Lemaréchal. 1993. *Convex Analysis and Minimization Algorithms: I*. Springer-Verlag, Berlin.
- Hu, X. 2002. Mathematical programs with complementarity constraints and game theory in electricity markets. Ph.D. dissertation, Department of Mathematics and Statistics, University of Melbourne, Melbourne, Australia.
- Hu, X., D. Ralph. 2007. Using EPECs to model bilevel games in restructured electricity markets with locational prices. *Oper. Res.* **55**(5) 809–827.
- Leyffer, S., T. Munson. 2005. Solving multi-leader-follower games. Argonne National Laboratory, Argonne, IL.
- Murphy, F. H., H. D. Sherali, A. L. Soyster. 1982. A mathematical programming approach for determining oligopolistic market equilibrium. *Math. Programming* **24** 92–106.
- Pang, J. S., M. Fukushima. 2005. Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. *Comput. Management Sci.* **2** 21–56.
- Papageorgiou, N. 1985. On the theory of Banach space valued multifunctions 1. Integration and conditional expectation. *J. Multivariate Anal.* **17** 185–206.
- Patriksson, M., L. Wynter. 1999. Stochastic mathematical programs with equilibrium constraints. *Oper. Res. Lett.* **25** 159–167.

- Ralph, D. 1994. Global convergence of damped Newton's method for non-smooth equations via the path search. *Math. Oper. Res.* **19** 352–389.
- Ralph, D., H. Xu. 2005. Implicit smoothing and its application to optimization with piecewise smooth equality constraints. *J. Optim. Theory Appl.* **124** 673–699.
- Robinson, S. M. 1996. Analysis of sample-path optimization. *Math. Oper. Res.* **21** 513–528.
- Rockafellar, R. T., R. J.-B. Wets. 1998. *Variational Analysis*. Springer-Verlag, Berlin.
- Rosen, J. B. 1965. Existence and uniqueness of equilibrium points for concave N -person games. *Econometrica* **33** 520–534.
- Rubinstein, R. Y., A. Shapiro. 1993. *Discrete Events Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Methods*. John Wiley and Sons, New York.
- Ruszczynski, A., A. Shapiro. 2003. Optimality conditions and duality in stochastic programming. A. Ruszczyński, A. Shapiro, eds. *Stochastic Programming, Handbooks in OR & MS*, Vol. 10. North-Holland Publishing Company, Amsterdam, 141–211.
- Shapiro, A. 2003. Monte Carlo sampling methods. A. Ruszczyński, A. Shapiro, eds. *Stochastic Programming, Handbooks in OR & MS*, Vol. 10. North-Holland Publishing Company, Amsterdam, 353–425.
- Shapiro, A., T. Homem-de-Mello. 2000. On rate of convergence of Monte Carlo approximations of stochastic programs. *SIAM J. Optim.* **11** 70–86.
- Shapiro, A., H. Xu. 2008. Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation. *Optimization* **57** 395–418.
- Sherali, H. D. 1984. A multiple leader Stackelberg model and analysis. *Oper. Res.* **32** 390–405.
- Sherali, H. D., A. L. Soyster, F. H. Murphy. 1983. Stackelberg-Nash-Cournot equilibria: Characterizations and computations. *Oper. Res.* **31** 253–276.
- Stackelberg, H. 1934. *Marktform und Gleichgewicht*. Julius Springer, Vienna.
- Xu, H. 2005. An MPCC approach for stochastic Stackelberg-Nash-Cournot equilibrium. *Optimization* **54** 27–57.