Optimization,
Vol. 54, No. 1, February 2005, pp. 27–57

An MPCC approach for stochastic Stackelberg–Nash–Cournot equilibrium

HUIFU XU*

School of Mathematics, University of Southampton, Highfield Southampton, SO17 1BJ, UK

(*Received 1 February 2004; in final form 24 June 2004*)

In this article, we investigate a Stochastic Stackelberg–Nash–Cournot Equilibrium problem by reformulating it as a Mathematical Program with Complementarity Constraints (MPCC). The complementarity constraints are further reformulated as a system of nonsmooth equations. We characterize the followers' Nash–Cournot equilibria by studying the implicit solution of a system of equations. We outline numerical methods for the solution of a stochastic Stackelberg–Nash–Cournot Equilibrium problem with finite distribution of market demand scenarios and propose a discretization approach based on implicit numerical integration to deal with stochastic Stackelberg–Nash–Cournot Equilibrium problem with continuous distribution of demand scenarios. Finally, we discuss the two-leader Stochastic Stackelberg–Nash–Cournot Equilibrium problem.

Keywords: Stackelberg–Nash–Cournot equilibrium; Stochastic mathematical program

Mathematics Subject Classifications 2000: 91A40; 91B50

1. Introduction

Consider a supply side oligopoly market where $N+1$ firms compete to supply a homogeneous product in a non-competitive manner. One of them, called *leader* hereafter, knows the reaction of the others. The leader is to choose his optimal supply so that his profit is maximized. The other firms, called *followers*, attempt to maximize their profits by supplying quantities under Cournot conjecture that the remaining firms will hold their supplies. It is well-known that such a game can be described as the Stackelberg–Nash–Cournot game.

A Stackelberg–Nash–Cournot equilibrium is a situation where the leader chooses an optimal supply that maximizes his profit, given his knowledge of the followers' reaction to his choice of supply, the followers reaching a Nash–Cournot equilibrium where each firm cannot improve his profit by unilaterally changing his supply. The Stackelberg–Nash–Cournot equilibrium model has been studied by Sherali *et al*. [26].

*E-mail: h.xu@maths.soton.ac.uk

Optimization
ISSN 0233-1934 print: ISSN 1029-4945 online
http://www.tandf.co.uk/journals
DOI: 10.1080/02331930412331323863

They used a quadratic programming approach to investigate the followers' Nash–Cournot equilibrium and proposed a numerical method to find the equilibrium. The method is based on the calculation of the objective function value at a set of points spread over the feasible region of the leader's decision variable and the linearization of the followers' aggregate reactive quantity over each interval of two adjacent points.

De Wolf and Smeers [5] considered a stochastic version of the Stackelberg–Nash–Cournot equilibrium model. The stochastic factor comes from some uncertainty of market demand at the time when the leader makes his decision on supply. It is assumed that the leader knows the distribution of the uncertain factor. Since the demand is not realized at the time when the leader makes a decision, what the leader can do at the most is to maximize the expected profit based on his knowledge of the distribution of demand and the followers' reaction in each scenario. De Wolf and Smeers used a method proposed by Sherali *et al.* [26] to find the stochastic Stackelberg–Nash–Cournot equilibrium. The model was applied to the European gas market where a particular gas producer Norway had the opportunity to develop a new and important field and its decision on the development needed to be made in 1990 to become effective 10 years later, the producer faced competitions from other gas producers such as the CEI, the Netherlands, Algeria and the United Kingdom as well as uncertain demand in year 2000.

In this article, we investigate a more general stochastic Stackelberg–Nash–Cournot equilibrium framework by considering a similar market to that discussed by De Wolf and Smeers [5] except that the random factor in demand is allowed to have continuous distribution. As we shall show later, De Wolf and Smeers' model can be treated essentially as a deterministic mathematical program. However, the generalized stochastic Stackelberg–Nash–Cournot equilibrium problem seems to have no easy treatment as such and this is one of the primary motivations of this article.

We propose a new approach, which is another main motivation, to investigate stochastic Stackelberg–Nash–Cournot equilibria. The approach is based on the reformulation of the followers' Nash cournot equilibrium problem as a complementarity problem and a further reformulation of the complementarity problem as a system of nonsmooth equations. The stochastic Stackelberg–Nash–Cournot equilibrium is consequently transformed into a stochastic Mathematical Program with Complementarity Constraints (MPCC) and further into a mathematical program with nonsmooth equality constraints.

The idea of reformulating a Stackelberg game as a Mathematical Program with Equilibrium Constraints (MPEC) is not new [18,20], although we have not seen the use of it to deal with stochastic Stackelberg–Nash–Cournot equilibrium. MPEC has been increasingly investigated over the past few years and has become a relatively new area of optimization. See monograph [18] for the recent developments of the subject. Having benefited from the recent breakthrough in the treatment of nonlinear complementarity programs, MPCC, as a special type of MPEC, has received considerable attention and more mature algorithms. See [8,9,11,12,27] and references therein. This is one of the key reasons why we use MPCC approach to deal with stochastic Stackelberg–Nash–Cournot equilibrium.

The rest of this article is organized as follows. In Section 2, we give a mathematical description of the Stochastic Stackelberg–Nash–Cournot equilibrium problem. Under some appropriate assumptions, we show the existence and uniqueness of the followers'

Nash–Cournot equilibrium. In Section 3, we reformulate the followers' Nash–Cournot equilibrium problem as a complementarity problem and further as a system of non-smooth equations. Then we investigate the properties of the followers' Nash–Cournot equilibrium and the existence of the stochastic Stackelberg–Nash–Cournot equilibrium. In Section 4, we propose a smoothing MPCC method for the solution of a discrete stochastic Stackelberg–Nash–Cournot equilibrium problem. An error bound on the difference between the optimum of the smoothed problem and that of the original problem is given. In Section 5, we propose a discretization method for the solution of a continuous stochastic Stackelberg–Nash–Cournot equilibrium problem. Finally, in Section 6, we give a brief discussion of a two leaders' stochastic Stackelberg–Nash–Cournot equilibrium problem and show the existence of an equilibrium under some circumstances.

2. Mathematical descriptions of the problem

We describe the market demand with the inverse demand function $p(q, \epsilon(\omega))$, where q is the total quantity of supply to the market, $\epsilon\colon \Omega \to \mathbb{R}$ is a random shock with known distribution, and $p(q, \epsilon(\omega))$ is the market price. As we discussed in the preceding section, the random shock reflects the uncertainty of demand at the time when the leader makes a decision.

Let x denote the decision variable of the leader, that is, the quantity supplied by the leader to the market. Let q_i, $i = 1, \ldots, N$, denote the decision variable of follower i, that is, the quantity supplied by firm i to the market.

Followers' Decision Problems Consider a particular future demand scenario $p(\cdot, \epsilon(\omega))$. Assume that the leader's supply is x and the aggregate supplies of followers except i is $\sum_{k=1, k\neq i}^{N} q_k$. If firm i's supply is q_i, then the market price in this demand scenario is $p(x + \sum_{i=1}^{N} q_i, \epsilon(\omega))$. The total revenue of firm i is $q_i p(x + \sum_{i=1}^{N} q_i, \epsilon(\omega))$. Suppose the total cost for firm i to produce q_i is $c_i(q_i)$. Then firm i's profit can be formulated as

$$f_i(q_i) = q_i p\left(x + \sum_{i=1}^{N} q_i, \epsilon(\omega)\right) - c_i(q_i).$$

Since the market price depends on q_i (in other words, firm i has market power), firm i would like to choose an optimal q_i in order to maximize his profit $f_i(q_i)$. Therefore follower i's profit maximization problem can be written as

$$\max_{q_i \geq 0} f_i(q_i) \equiv q_i p\left(x + q_i + \sum_{k=1, k\neq i}^{N} q_k, \epsilon(\omega)\right) - c_i(q_i). \tag{1}$$

In choosing an optimal decision, firm i holds the other firms' supplies as constant. A *Nash–Cournot equilibrium* among followers in demand scenario $p(\cdot, \epsilon(\omega))$ is a situation where no firm can improve its profit by unilaterally changing its supply. We denote such an equilibrium by $(q_1(x, \epsilon(\omega)), \ldots, q_N(x, \epsilon(\omega)))$ where $q_i(x, \epsilon(\omega))$ is the global optimal solution of (1). The existence and uniqueness of such an equilibrium will be dealt with in Proposition 2.6.

Leader's Decision Problem We suppose that the leader knows how the followers will choose their outputs (as described in (1)) in each demand scenario but he does not know which demand scenario will occur in future at the time when he makes a decision on output. What the leader can do at best is to maximize the expected profit based on his knowledge on the market demand distribution and the followers' reaction to his supply. Therefore we can formulate the leader's decision problem as follows:

$$\max_{x\geq 0} f_0(x) \equiv E_\omega\left[xp\left(x+\sum_{i=1}^{N} q_i(x,\epsilon(\omega)),\epsilon(\omega)\right)\right] - c_0(x) \tag{2}$$

where E_ϵ denotes the mathematical expectation. Note that unlike followers' objective functions, the leader's objective function is not necessarily concave. Hence from here on, the 'maximum' of the leader's expected profit refers to the global maximum of (2).

Stochastic Stackelberg–Nash–Cournot Equilibrium We investigate a situation where the leader maximizes the expected profit while the followers reach a Nash–Cournot equilibrium in every demand scenario.

Definition 2.1 A Stackelberg–Nash–Cournot Equilibrium is an $N+1$ dimensional vector $(x^*, q_1(x^*,\cdot),\ldots,q_N(x^*,\cdot))$ such that

$$f_0(x^*) = \max_{x\geq 0}\; E_\omega\left[xp\left(x+\sum_{i=1}^{N} q_i(x,\epsilon(\omega)),\epsilon(\omega)\right)\right] - c_0(x) \tag{3}$$

where

$$q_i(x,\epsilon(\omega)) \in \arg\max_{q_i\geq 0}\left(q_i p\left(x+q_i+\sum_{k=1,k\neq i}^{N} q_k(x,\epsilon(\omega)),\epsilon(\omega)\right) - c_i(q_i)\right). \tag{4}$$

For the simplicity of notation in some cases, let

$$Q(x,\epsilon(\omega)) \equiv \sum_{i=1}^{N} q_i(x,\epsilon(\omega))$$

and

$$Q_{-i}(x,\epsilon(\omega)) \equiv \sum_{k=1,k\neq i}^{N} q_i(x,\epsilon(\omega)).$$

In this article, we are interested in two particular cases depending on the distribution of the random shock $\epsilon(\omega)$.

Case 1 $\epsilon(\omega)$ has a finite discrete distribution, that is,

$$\text{prob}(\epsilon(\omega) = t_l) = \pi_l > 0, \qquad l = 1, \ldots, L, \quad \sum_{l=1}^{L} \pi_l = 1.$$

Consequently the Stackelberg–Nash–Cournot Equilibrium problem becomes

$$\text{(DSSNCE)} \quad \begin{array}{ll} \max & f_0(x) \equiv \sum_{l=1}^{L} \pi_l x p(x + Q(x, t_l), t_l) - c_0(x) \\ \text{s.t.} & x \geq 0, \\ & q_i(x, t_l) \in \arg\max_{q_i \geq 0}(q_i p(x + q_i + Q_{-i}(x, t_l), t_l) - c_i(q_i)), \\ & l = 1, \ldots, L. \end{array} \tag{5}$$

De Wolf and Smeers [5] discussed this model and applied it to study the European gas market. In this case the future demand has L scenarios, each of which has a positive probability. The model is an extension of the Stackelberg–Nash–Cournot equilibrium model, considered by Sherali *et al.* [26].

Case 2 $\epsilon(\omega)$ has a continuous distribution with a density function $\rho(t)$. The support set of ρ is $[0, T]$. Consequently the Stackelberg–Nash–Cournot Equilibrium problem becomes

$$\text{(CSSNCE)} \quad \begin{array}{ll} \max & f_0(x) \equiv \int_0^T x p(x + Q(x, t), t)\rho(t)\, dt - c_0(x) \\ \text{s.t.} & x \geq 0, \\ & q_i(x, t) \in \arg\max_{q_i \geq 0}(q_i p(x + q_i + Q_{-i}(x, t), t) - c_i(q_i)), \\ & t \in [0, T]. \end{array} \tag{6}$$

This is a generalization of the DSSNCE model. It is of both theoretical and practical interests to discuss this model.

We need some assumptions for the study of both the models. Let $\mathcal{T}$ denote the set $[0, T]$ in CSSNCE model and $\{t_1, \ldots, t_L\}$ in DSSNCE model.

Assumption 2.2 For $i = 0, 1, \ldots, N$, $c_i(q)$ is twice continuously differentiable, $c_i'(q) \geq 0$ and $c_i''(q) \geq 0$ for $q \geq 0$.

This assumption is standard, it requires that the cost function of each firm be convex and sufficiently smooth. See [5,26].

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) $p_q'(q, t) + q p_{qq}''(q, t) \leq 0$, for $q \geq 0$ and $t \in \mathcal{T}$.

This assumption is similar to an assumption used by Sherali *et al.* [26] and De Wolf and Smeers [5]. Consider a monopoly market with an extraneous supply $K \geq 0$. If the monopoly's output is q, then its revenue at demand scenario $\epsilon(\omega) = t$ is $q(p(q + K, t)$. The marginal revenue is $p(q + K, t) + q p_q'(q + K, t)$. The rate of change of this marginal revenue with respect to the increase in the extraneous supply K is $p_q'(q + K, t) + q p_{qq}''(q + K, t)$. 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 monopoly's marginal revenue in any demand scenario. See [26] for a similar explanation for a deterministic leader-followers' market.

PROPOSITION 2.4 *Under Assumption* 2.3,

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

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

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

Proof A proof in the deterministic case (without parameter t) was given in [26]. The proof with t is similar. We include it here for completeness.

Part (i) Let $t \in \mathcal{T}$. If $p''_{qq}(q+K,t) \leq 0$, then

$$p'_q(q+K,t) + qp''_{qq}(q+K,t) \leq p'_q(q+K,t) \leq 0.$$

If $p''_{qq}(q+K,t) \geq 0$, then by Assumption 2.3 (ii),

$$p'_q(q+K,t) + qp''_{qq}(q+K,t) \leq p'_q(q+K,t) + (q+K)p''_{qq}(q+K,t) \leq 0.$$

Part (ii) Let $R(q,t) = qp(q+K,t)$. Then

$$R''_{qq}(q,t) = 2p'_q(q+K,t) + qp''_{qq}(q+K,t).$$

The conclusion follows straightforwardly from Part (i). This completes the proof. ■

In the subsequent discussion, we will quote Proposition 2.4 particularly in Propositions 2.6 and 3.3. Note that the strict concavity of $qp(q+x,t)$ does not ensure the boundedness of its maximum. In order to discuss the existence of a Stackelberg–Nash–Cournot Equilibrium, we need an extra assumption as follows.

Assumption 2.5 There exists q^u, such that

$$c'_i(q) \geq p(q,t), \quad \text{for } q \geq q^u, \quad t \in \mathcal{T}, \quad i = 0, 1, \ldots, N.$$

The assumption implies that even firm i was a monopoly, its marginal cost at output level q^u or above would exceed any possible market price. Therefore, none of the firms would wish to supply more than q^u. See a similar assumption in [5,26] and a discussion in [26].

PROPOSITION 2.6 *Under Assumptions* 2.2, 2.3 *and* 2.5,

(i) $f_0(x)$ *is non-negative and is bounded for* $x \geq 0$;
(ii) *for fixed* $x \geq 0$ *and* $t \in \mathcal{T}$, *there exists a unique Nash–Cournot equilibrium among followers*, $(q_1(x,t), \ldots, q_N(x,t))$, *which solves* (4); *moreover*, $q_i(x,t) \in [0, q^u)$, *for* $i = 1, \ldots, N$.

Proof

Part (i) We only prove for $\mathcal{T} = [0, T]$. The discrete distribution case can be analyzed in a similar way. Since $p(q, t)$ is strictly decreasing in q and $Q(x, t) \geq 0$, then

$$\begin{aligned} f_0(x) &= x \int_0^T p(x + Q(x, t), t)\rho(t)\,dt - c_0(x) \\ &\leq x \int_0^T p(x, t)\rho(t)\,dt - c_0(x) \\ &= x \max_{t\in[0,T]} p(x, t) - c_0(x). \end{aligned} \tag{8}$$

Let

$$\tilde{f}(x, t) \equiv xp(x, t) - c_0(x)$$

and

$$\tilde{f}(x) = \max_{t\in[0,T]} \tilde{f}(x, t).$$

It is well-known that $\tilde{f}$ is Lipschitz continuous and

$$\partial\tilde{f}(x) = \bigcup_{t\in\mathcal{T}(x)} f_x{}'(x, t)$$

where

$$\mathcal{T}(x) = \left\{ t \in [0, T]\colon \tilde{f}(x, t) = \tilde{f}(x) \right\}$$

and $\partial\tilde{f}(x)$ is the Clarke subdifferential [4]. See [4, Theorem 2.8.6]. By Assumptions 2.2 and 2.3, we can easily show that $\tilde{f}(x, t)$ is a concave function in x. Moreover, for $x \geq q^u$, by Assumption 2.5

$$\tilde{f}_x{}'(x, t) = p(x, t) + xp'_q(x, t) - c'_0(x) \leq xp'_q(x, t) < 0.$$

This shows that

$$\xi < 0; \quad \text{for } \xi \in \partial\tilde{f}(x).$$

The strict inequality is due to the compactness of $\partial\tilde{f}(x)$. Therefore, $\tilde{f}(x)$ achieves its maximum on $[0, q^u)$. Using (8), we have

$$\sup_{x\geq 0} f_0(x) \leq \max_{x\geq 0} \tilde{f}(x) < +\infty.$$

Part (ii) Let t and x be fixed. For given q_k, $k = 1, \ldots, N$, $k \neq i$, it follows from Proposition 2.4 (i) that firm i's objective function $f_i(q_i)$ is strictly concave. By [25, Theorems 1 and 2], there exists a unique Nash–Cournot equilibrium $(q_1(x,t), \ldots, q_N(x,t))$ which solves (4).

To prove that $q_i(x,t) \in [0, q^u)$, we note that firm i's objective function can be written as

$$f_i(q_i) \equiv q_i p(x + q_i + Q_{-i}, t) - c_i(q_i).$$

Since f_i is a strictly concave and

$$\begin{aligned} f_i{}'(q_i) &= p(x + q_i + Q_{-i}, t) + q_i p_q'(x + q_i + Q_{-i}, t) - c_i'(q_i) \\ &\leq p(q_i, T) + q_i p_q'(x + q_i + Q_{-i}, t) - c_i'(q_i) \\ &\leq q_i p_q'(q_i + Q_{-i}, t) \\ &< 0 \end{aligned}$$

for $q_i \geq q^u$, then f_i achieves maximum on $[0, q^u)$. ■

3. A nonsmooth equation approach for the followers' equilibrium

In this section we discuss the properties of the followers' Nash–Cournot equilibrium and the existence of the Stochastic Stackelberg–Nash–Cournot equilibria.

3.1 *Reformulation of the followers' Nash–Cournot equilibrium problem*

Our first step is to reformulate the followers' Nash–Cournot equilibrium as a nonlinear complementarity problem. Consider a demand scenario where the random shock $\epsilon(\omega) = t$, $t \in \mathcal{T}$ and the leader supply is x. We can characterize the followers' Nash–Cournot equilibrium problem by considering the Karush–Kuhn–Tucker conditions for each follower. For $t \in \mathcal{T}$ and $i = 1, \ldots, N$,

$$\begin{aligned} p(x + Q, t) + q_i p_q'(x + Q, t) - c_i'(q_i) + \mu_i = 0, \\ \mu_i \geq 0, \quad q_i \geq 0, \quad \mu_i q_i = 0, \end{aligned}$$

where $Q = \sum_{i=1}^N q_i$. This is a parameterized N-dimensional nonlinear complementarity problem where both x and t become parameters. Let $\mathbf{q} = (q_1, \ldots, q_N)^T$, $\mathbf{e} = (1, \ldots, 1)^T$, $\mathbf{c}(\mathbf{q}) = (c_1(\mathbf{q}), \ldots, c_N(\mathbf{q}))^T$. Let

$$G(\mathbf{q}, x, t) \equiv -p(x + \mathbf{q}^T\mathbf{e}, t)\mathbf{e} - p_q'(x + \mathbf{q}^T\mathbf{e}, t)\mathbf{q} + \nabla\mathbf{c}(\mathbf{q}).$$

Then the complementarity problem above can be rewritten as

$$0 \leq \mathbf{q} \perp G(\mathbf{q}, x, t) \geq 0. \tag{9}$$

The reformulation of an oligopoly game as a complementarity problem is well-known. See for example [7]. Using (9), we can rewrite the stochastic Stackelberg–Nash–Cournot equilibrium problem as a Stochastic Mathematical Program with Complementarity Constraints (SMPCC):

$$\begin{array}{ll} \max & \mathrm{E}_\omega\big[xp\big(x+\mathbf{e}^T\mathbf{q}(x,\epsilon(\omega)),\epsilon(\omega)\big)\big]-c_0(x) \\ \text{s.t.} & x\geq 0, \\ & \mathbf{q}(x,\epsilon(\omega)) \quad \text{solves } 0\leq \mathbf{q}\perp G(\mathbf{q},x,\epsilon(\omega))\geq 0, \quad \omega\in\Omega. \end{array} \tag{10}$$

There have been a few references in the literature related to SMPCC. Patriksson and Wynter [22] investigated a broad class of stochastic mathematical program with equilibrium constraints. Christiansen *et al*. [3] discussed a stochastic bilevel programming model for a structural optimization problem where the structural equilibrium may be subject to the random properties of materials and randomly varying conditions such as weather and external forces [3].

Recently, Lin *et al*. [16] considered a wait-and-see model for a lunch vendor problem in which a vendor buys lunches from a lunch production company and sells them to customers. Both the company and the vendor face uncertain demands. However the company needs to make a decision on sale price and quantity at once before the realization of the market demand while the vendor can make a decision after the observation on market demand. They modeled the problem as a stochastic mathematical program with equilibrium constraints. Lin and Fukushima [17] also considered a here-and-now model for the same problem in which both the vendor and the company need to make a decision before the realization of market demand and modeled it as a stochastic mathematical program with equilibrium constraints and recourse.

It is important to note that although program (10) is a stochastic program, it cannot be simply included in a category of standard well-discussed stochastic programs, such as distribution problem or a two-stage recourse problem. Note also that the distribution of the random shock in demand is assumed known and therefore there is no need to take samples on the uncertainty. However, it will be very interesting and practical to consider the case of the distribution of uncertain factors are not precisely known and consequently the sampling may be needed. See an excellent discussion in this regard by Kleywegt and Shapiro [15].

In what follows, we discuss the followers' Cournot-Nash equilibrium. It is well-known that the complementarity problem (9) can be reformulated as a system of nonsmooth equations

$$F(\mathbf{q},x,t)\equiv \min(G(\mathbf{q},x,t),\mathbf{q})=0. \tag{11}$$

where 'min' is taken componentwise. There are two ways to look at (11). One is to treat x and t as variables and hence (11) is an underdetermined system of equations. The other is to treat x and t as parameters and consequently (11) as a parameterized nonsmooth system of equations. We shall not distinguish them because in both cases we shall solve $\mathbf{q}$ from (11) as a function of x and t.

Our idea is to investigate the dependence of Cournot-Nash equilibrium $\mathbf{q}(x,t)$ on x and t by looking into the implicit solution $\mathbf{q}(x,t)$ of (11). Note that Sherali *et al*. [26]

used a quadratic programming approach to investigate $\mathbf{q}(x, t)$. We shall show that our approach, by taking advantage of the recent development of nonsmooth equations, is better integrated into recently developed powerful numerical methods, which can be readily used to solve the SSNCE problems.

Note that there are many elementary functions known as NCP functions which can be used to reformulate (9) as a system of equations. See [12,28]. Regardless of the choice of NCP functions, the solution of the reformulated system remains the same and it is the Nash–Cournot equilibrium of the followers. However, we shall see later that the derivative of the solution with respect to x and t depends on the reformulation (the choice of NCP functions) and the reformulation (11) seemingly gives us the desired estimation of these derivatives (subdifferentials).

3.2 *Preliminaries in nonsmooth equations*

The discussion on (11) will involve some basic elements in nonsmooth analysis particularly nonsmooth equations. In this subsection, we make some preparations.

Let $H: \mathbb{R}^n \to \mathbb{R}^m$ be a Lipschitz function. The *Clarke generalized Jacobian* [4] of H at $\mathbf{x} \in \mathbb{R}^n$ is defined as

$$\partial H(\mathbf{x}) \equiv \operatorname{conv}\{\overline{\lim_{\substack{\mathbf{y}\in D_H \\ \mathbf{y}\to\mathbf{x}}}} \nabla H(\mathbf{y})\},$$

where 'conv' denotes the convex hull of a set and D_H denotes the set of points in a neighborhood of $\mathbf{x}$ at which H is Frechét differentiable. When $m = 1$ or $n = 1$, ∂H is also called Clarke subdifferential. When $n = m$, the Clarke Jacobian $\partial H(\mathbf{x})$ is said to be *non-singular* if every matrix in $\partial H(\mathbf{x})$ is non-singular.

Let $\Theta = \operatorname{diag}(\theta_1, \ldots, \theta_N) \in \mathbb{R}^{N\times N}$ denote the diagonal matrix with the (i, i)th entry being θ_i, for $i = 1, \ldots, N$. Let I denote the identity matrix in $\mathbb{R}^{N\times N}$. It is easy to verify that the function F defined by (11) is Lipschitz and the Clarke generalized Jacobian of F in $(\mathbf{q}, x, t)$ can be expressed as

$$\partial F(\mathbf{q}, x, t) = \left\{(I - \Theta, \Theta)\begin{pmatrix} I \\ \nabla G(\mathbf{q}, x, t) \end{pmatrix} : \theta_i \in [0, 1], i = 1, \ldots, N\right\}. \tag{12}$$

Moreover $\partial F = \partial F_1 \times \cdots \times F_N$ where ∂F_i denote the Clarke subdifferential of the ith component function F_i.

LEMMA 3.1 *Suppose that $M \in \mathbb{R}^{N\times N}$ is a positive definite matrix. Then there exists a constant $C > 0$ such that*

$$\|((I - \Theta) + \Theta M)^{-1}\| \leq C, \quad \textit{for } \Theta \in \operatorname{diag}([0, 1], \ldots, [0, 1]).$$

Here and later on $\|\cdot\|$ denotes the 2-norm of a matrix and a vector.

Proof The nonsingularity of $(I - \Theta) + \Theta M$ follows from [29, Theorem 9]. The rest follows from the compactness of the set of matrices $\operatorname{diag}([0, 1], \ldots, [0, 1])$. ∎

In order to discuss the solution $\mathbf{q}(x, t)$ of (11), we need the following results which deal with the implicit function of nonsmooth equations.

LEMMA 3.2 *Consider an underdetermined system of nonsmooth equations*

$$H(\mathbf{y}, \mathbf{z}) = 0,$$

where $H: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ *is locally Lipschitz. Let* $(\bar{\mathbf{y}}, \bar{\mathbf{z}}) \in \mathbb{R}^n \times \mathbb{R}^m$ *be such that* $H(\bar{\mathbf{y}}, \bar{\mathbf{z}}) = 0$. *Suppose* $\partial_{\mathbf{y}} H(\bar{\mathbf{y}}, \bar{\mathbf{z}})$ *is nonsingular. Then*

(i) *there exist neighborhoods* Z *of* $\bar{\mathbf{z}}$, Y *of* $\bar{\mathbf{y}}$ *and a locally Lipschitz function* $\mathbf{y}: Z \to Y$, *such that* $\mathbf{y}(\bar{\mathbf{z}}) = \bar{\mathbf{y}}$ *and, for every* $\mathbf{z} \in Z$, $\mathbf{y} = \mathbf{y}(\mathbf{z})$ *is the unique solution of the problem* $H(\mathbf{y}, \mathbf{z}) = 0$, $\mathbf{y} \in Y$;
(ii) *for* $\mathbf{z} \in Z$,

$$\partial \mathbf{y}(\mathbf{z}) \subset \{-R^{-1}U: (R, U) \in \partial H(\mathbf{y}(\mathbf{z}), \mathbf{z}), R \in \mathbb{R}^{n\times n}, U \in \mathbb{R}^{n\times m}\}. \tag{13}$$

Proof Part (i) is a restate of [4, Theorem 7.1.1] and the following corollary. Part (ii) seems known but we are not able to find a reference and hence include a proof for completeness.

Let $\hat{\mathbf{z}} \in Z$. Since $\mathbf{y}(\cdot)$ is Lipschitz on Z, $\mathbf{y}(\cdot)$ is Frechét differentiable almost everywhere in Z. Let $D_{\mathbf{y}}$ denote the set of points at which $\mathbf{y}(\cdot)$ is differentiable. Then by definition

$$\partial \mathbf{y}(\hat{\mathbf{z}}) \equiv \operatorname{conv}\Big\{ \overline{\lim_{\substack{\mathbf{z}\in D_{\mathbf{y}} \\ \mathbf{z}\to\hat{\mathbf{z}}}}} \nabla \mathbf{y}(\mathbf{z}) \Big\}.$$

Let $\mathbf{z} \in D_{\mathbf{y}}$ be a point nearby $\hat{\mathbf{z}}$. If H is differentiable at the point $(\mathbf{y}(\mathbf{z}), \mathbf{z})$, then

$$\nabla \mathbf{y}(\mathbf{z}) = -\nabla_{\mathbf{y}}^{-1} H(\mathbf{y}(\mathbf{z}), \mathbf{z}) \nabla_{\mathbf{z}} H(\mathbf{y}(\mathbf{z}), \mathbf{z}).$$

If H is not differentiable at the point, then since $\mathbf{y}(\cdot)$ is differentiable at $\mathbf{z}$ and

$$H(\mathbf{y}(\mathbf{z}), \mathbf{z}) = \mathbf{0},$$

by [4, Theorem 2.3.10],

$$0 \in \partial H(\mathbf{y}(\mathbf{z}), \mathbf{z}) \begin{pmatrix} \nabla \mathbf{y}(\mathbf{z}) \\ I \end{pmatrix}$$

where $I \in \mathbb{R}^{m\times m}$ is an identity matrix. Driving $\mathbf{z}$ to $\hat{\mathbf{z}}$ and using the upper semicontinuity of the Clarke generalized Jacobian, we obtain (13). ■

3.3 *Properties of the followers' Nash–Cournot equilibrium*

We are now ready to address one of our main results in this section.

PROPOSITION 3.3 *Let* $F(\mathbf{q}, x, t)$ *be defined as in* (11). *Under Assumptions* 2.2, 2.3 *and* 2.5,

(i) $\partial_{\mathbf{q}} F(\mathbf{q}, x, t)$ *is non-singular for any* $\mathbf{q} \geq 0$ *and* $x \geq 0$;
(ii) *for every* $x \geq 0$ *and* $t \in \mathcal{T}$, *there exists a unique* $\mathbf{q}$ *such that* $F(\mathbf{q}, x, t) = 0$;
(iii) *there exists a unique Lipschitz piecewise smooth continuous function* $\mathbf{q}(x, t)$, *such that* $F(\mathbf{q}(x, t), x, t) = 0$.

Proof
Part (i) The Jacobian of G with respect to $\mathbf{q}$ can be worked out explicitly

$$\nabla_{\mathbf{q}} G(\mathbf{q}, x, t) = -p'_q(x + Q, t)\mathbf{e}\mathbf{e}^T - p''_{qq}(x + Q, t)\mathbf{q}\mathbf{e}^T - p'_q(x + Q, t)I + \nabla^2 c(\mathbf{q}) \qquad (14)$$

where $Q = \mathbf{q}^T\mathbf{e}$. Since $p'_q(x + Q, t) < 0$, the eigenvalues of $\nabla_{\mathbf{q}} G(\mathbf{q}, x, t)$ are lower bounded by

$$-(N + 1)p'_q(x + Q, t) - p''_q(x + Q, t)Q + \min_{i=1,\ldots,N} c''_i(q_i).$$

By Assumption 2.2, $c''_i(q_i) \geq 0$. Moreover, by Assumption 2.3 (i), $p(q, t)$ is twice continuously differentiable and $p(q, t) < 0$ for $q \geq 0$, $t \in \mathcal{T}$. Since $\mathcal{T}$ is compact, then there exists a constant $C > 0$, such that

$$\min_{t \in \mathcal{T}} -p'_q(x + Q, t) \geq C, \quad \text{for } x \in [0, q^u], \quad Q \in [0, Nq^u].$$

On the other hand, it follows from Proposition 2.4 (i),

$$-p'_q(x + Q, t) - p''_q(x + Q, t)Q \geq 0.$$

This shows

$$-(N + 1)p'_q(x + Q, t) - p''_q(x + Q, t)Q + \min_{i=1,\ldots,N} c''_i(q_i) \geq NC.$$

and subsequently, $\nabla_{\mathbf{q}} G(\mathbf{q}, x, t)$ is uniformly positive definite. Since

$$\partial_{\mathbf{q}} F(\mathbf{q}, x, t) = \Theta + (I - \Theta)\nabla_{\mathbf{q}} G(\mathbf{q}, x, t)$$

where $\Theta \in \text{diag}([0, 1], \ldots, [0, 1])$, by Lemma 3.1, $\partial_{\mathbf{q}} F(\mathbf{q}, x, t)$ is uniformly nonsingular.

Part (ii) The conclusion follows straightforwardly from Proposition 2.6 (ii) and the definition of F.

Part (iii) From part (i), $\partial_{\mathbf{q}} F(\mathbf{q}, x, t)$ is non-singular. By part (ii) and Lemma 3.2 (i), there exists a unique function $\mathbf{q}(x, t)$, such that

$$F(\mathbf{q}(x, t), x, t) = 0$$

in a neighborhood of (x, t). The implicit function can be easily extended. The piecewise smoothness follows from [24, Lemma 4.11] which was essentially due to [18,23]. ■

The proposition shows that the system of nonsmooth equations (11) has a unique solution for every t. Since (11) is equivalent to (9), this means (9) has a unique solution. On the other hand, a followers' Nash–Cournot equilibrium must satisfy the complementarity system (9), which means this solution must be the followers' Nash–Cournot equilibrium. This shows that the unique solution of (11) is the followers' Nash–Cournot equilibrium and vice versa.

Note that the proposition also shows the Lipschitz continuity of $\mathbf{q}(x,t)$, hence of $Q(x,t)$ and of $f_0(x)$. This paves the way for the investigation of derivatives of $Q(x,t)$ and f_0 in the next proposition.

PROPOSITION 3.4 *Let $F(\mathbf{q},x,t)$ be defined as in* (11). *Suppose that Assumptions* 2.2, 2.3 *and* 2.5, *are satisfied. Suppose also that*

$$\min_{i=1,\ldots,N} c_i'(0) < p(x+Q(x,t),t), \quad t \in \mathcal{T}. \tag{15}$$

Then

(i) *the Clarke subdifferential of* $\mathbf{q}$ *with respect to x and t can be estimated as follows:*

$$\partial_x \mathbf{q}(x,t) \subset \left\{-R^{-1}U : (R,U,V) \in \partial F(\mathbf{q}(x,t),x,t), R \in \mathbb{R}^{N\times N}, U \in \mathbb{R}^N, V \in \mathbb{R}^N\right\}$$

and

$$\partial_t \mathbf{q}(x,t) \subset \{-R^{-1}V : (R,U,V) \in \partial F(\mathbf{q}(x,t),x,t), R \in \mathbb{R}^{N\times N}, U \in \mathbb{R}^N, V \in \mathbb{R}^N\},$$

where ∂F is given in (12);

(ii) *the Clarke subdifferential of $Q(x,t)$ with respect to x can be estimated as*

$$\partial_x Q(x,t) \subset (-1,0);$$

(iii) *for problem* (6), if $p_{qt}''(q,t)=0$, *then $q_i(x,t)$ is increasing in t*; *moreover*, *if there exists a constant $C>0$ such that*

$$p_q'(q,t) + qp_{qq}''(q,t) < -C, \quad \text{for } q \geq 0, \text{ and } t \in [0,T], \tag{16}$$

then the Clarke subdifferential of $Q(x,t)$ with respect to t can be estimated as

$$\partial_t Q(x,t) \subset \left(0, \frac{1}{C}p_t'(x+Q(x,t),t)\right],$$

where C is a constant;

(iv) *$f_0(x)$ is locally Lipschitz continuous and*

$$\partial_x f_0(x) \subset \int_0^T (p(x+Q(x,t),t) + xp'(x+Q(x,t),t)\partial_x Q(x,t))\rho(t)\,dt - c_0'(x).$$

Proof

Part (i) The first inclusion follows straightforwardly from Lemma 3.2 (ii).

Part (ii) Since

$$\partial_x Q(x,t) \subset \mathbf{e}^T \partial_x \mathbf{q}(x,t), \tag{17}$$

we have from Part (i)

$$\partial_x Q(x,t) \subset \mathbf{e}^T\{-R^{-1}U: (R,U,V) \in \partial F(\mathbf{q},x,t)\}.$$

Since F_i is a piecewise smooth function, by the definition of Clarke subdifferential,

$$\partial F_i(\mathbf{q},x,t) = \{\theta_i \nabla G_i(\mathbf{q},x,t) + (1-\theta_i)\mathbf{e}_i\}, \quad \text{for } i = 1,\dots,N,$$

where $\theta_i \in [0,1]$, and $\mathbf{e}_i$ is an $N+2$ dimensional vector with the ith component being 1 and the rest being zero, for $i = 1,\dots,N$. Note that $\theta_i = 0$ only when $F_i(\mathbf{q},x,t) = q_i$. Let $\Theta = \mathrm{diag}(\theta_1,\dots,\theta_N)$ denote the $N \times N$ diagonal matrix. We first show that, under (15), $\mathbf{q}(x,t) \neq 0$.

By (15), there exists $i_0 \in \{1,\dots,N\}$ such that

$$c'_{i_0}(0) < p(x + Q(x,t),t), \quad t \in [0,T]. \tag{18}$$

By definition, $q_{i_0}(x,t)$ solves the following

$$\max_{q_{i_0} \geq 0} f_{i_0}(q_{i_0}) \equiv q_{i_0} p\left(x + q_{i_0} + \sum_{k=1,\, k \neq i_0}^{N} q_k(x,t), t\right) - c_{i_0}(q_{i_0}).$$

The first-order necessary condition gives

$$\begin{aligned} q_{i_0}(x,t) f_{i_0}'(q_{i_0}(x,t)) = q_{i_0}(x,t)[p(x + \mathbf{e}^T\mathbf{q}(x,t),t) \\ + p'(x + \mathbf{e}^T\mathbf{q}(x,t),t) q_{i_0}(x,t) - \nabla \mathbf{c}_{i_0}(q_{i_0}(x,t))] = 0, \\ q_{i_0}(x,t) \geq 0, \\ -p(x + \mathbf{e}^T\mathbf{q}(x,t),t) - p'(x + \mathbf{e}^T\mathbf{q}(x,t),t) q_{i_0}(x,t) + \nabla \mathbf{c}_{i_0}(q_{i_0}(x,t)) > 0. \end{aligned}$$

If $q_{i_0}(x,t) = 0$, then by (18)

$$f_{i_0}'(0) = p(x + Q(x,t),t) - c'_{i_0}(0)) > 0,$$

which shows the strict complementarity condition. This shows $q_{i_0}(x,t) \neq 0$. Subsequently

$$F_{i_0}(\mathbf{q}(x,t),x,t) = G_{i_0}(\mathbf{q}(x,t),x,t) = 0,$$

hence $\theta_{i_0} = 1$ at $(\mathbf{q}(x,t),x,t)$. This shows Θ is not a zero matrix under (15).

Consider an arbitrary matrix (R,U,V) of $\partial F(\mathbf{q}(x,t),x,t)$. We want to prove that

$$\mathbf{e}^T(-R^{-1}U) \in (0,1). \tag{19}$$

If we can do so, then the conclusion follows from (17) and (19). By definition

$$\begin{aligned} R &= \Theta \nabla_{\mathbf{q}} G(\mathbf{q}, x, t) + (I - \Theta) \\ &= \Theta(-p'_q \mathbf{e} - p''_{qq}\mathbf{q})\mathbf{e}^T + \Theta(-p'_q I + \nabla^2 c(\mathbf{q})) + (I - \Theta). \end{aligned}$$

where $\nabla_{\mathbf{q}} G(\mathbf{q}, x, t)$ is given in (14) and

$$U = \Theta G'_x(\mathbf{q}, x, t) = \Theta(-p'_q \mathbf{e} - p''_q \mathbf{q})$$

where

$$G'_x(\mathbf{q}, x, t) = -p'_q(x + \mathbf{q}^T\mathbf{e}, t)\mathbf{e} - p''_{qq}(x + \mathbf{q}^T\mathbf{e}, t)\mathbf{q}.$$

Note that (R, U, V) is an arbitrary element of the $\partial_C F(\mathbf{q}(x, t), x, t)$, it depends on point $(\mathbf{q}(x, t), x, t)$. However for the simplicity of notation, instead of writing $\mathbf{q}(x, t)$, we write $\mathbf{q}$. Similarly we write p'_q and p''_{qq} for $p'_q(x + \mathbf{q}^T\mathbf{e})$ and $p''_{qq}(x + \mathbf{q}^T\mathbf{e})$.

Let $D = \Theta(-p'_q I + \nabla^2 c(\mathbf{q})) + (I - \Theta)$. D is an $N \times N$ diagonal matrix. It is easy to verify that D is non-singular. Using Sherman-Morrison formula, we obtain

$$R^{-1} = D^{-1} - \frac{1}{\eta} D^{-1}\Theta(-p'_q \mathbf{e} - p''_{qq}\mathbf{q})\mathbf{e}^T D^{-1}$$

where

$$D^{-1} = \text{diag}\left(\frac{1}{\theta_1(-p'_q + c''_1(q_1)) + (1 - \theta_1)}, \ldots, \frac{1}{\theta_N(-p'_q + c''_N(q_N)) + (1 - \theta_N)}\right)$$

and

$$\eta = 1 + \gamma,$$

and

$$\gamma = \mathbf{e}^T D^{-1}\Theta(-p'_q \mathbf{e} - p''_{qq}\mathbf{q}) = \sum_{i=1}^{N} \frac{\theta_i(-p'_q - p''_{qq} q_i)}{\theta_i(-p'_q + c''_i(q_i)) + (1 - \theta_i)}.$$

Note that $\gamma > 0$. This is obvious when $p''_{qq} < 0$. When $p''_{qq} \geq 0$,

$$\gamma = \sum_{i=1}^{N} \frac{\theta_i(-p'_q - p''_{qq} q_i)}{\theta_i(-p'_q + c''_i(q_i)) + (1 - \theta_i)} \geq \sum_{i=1}^{N} \frac{\theta_i(-p'_q - p''_{qq} Q)}{\theta_i(-p'_q + c''_i(q_i)) + (1 - \theta_i)} > 0.$$

The last inequality is due to Proposition 2.4 and the fact that Θ is not a zero matrix. On the other hand, after a few calculations, we obtain

$$-\mathbf{e}^T R^{-1} U = -\gamma + \frac{1}{\eta}\gamma^2 = -\frac{\gamma}{1 + \gamma} \in (-1, 0).$$

The conclusion follows.

Part (iii) The proof is similar to that of Part (ii). Since $p''_{q,t} = 0$,

$$V = \Theta \nabla_t G(\mathbf{q}, x, t) = \Theta(-p'_t \mathbf{e})$$

and

$$-\mathbf{e}^T R^{-1} V = \frac{1}{1+\gamma} p'_t \mathbf{e}^T D^{-1} \Theta \mathbf{e}.$$

By (16),

$$\begin{aligned}\gamma &= \sum_{i=1}^{N} \frac{\theta;_i (-p'_q - p''_{qq} q_i)}{\theta_i(-p'_q + c''_i(q_i)) + (1-\theta_i)} \\ &\geq C \sum_{i=1}^{N} \frac{\theta_i}{\theta_i(-p'_q + c''_i(q_i)) + (1-\theta_i)} \\ &= C\mathbf{e}^T D^{-1} \Theta \mathbf{e}.\end{aligned}$$

Therefore

$$-\mathbf{e}^T R^{-1} U \leq \frac{p'_t \mathbf{e}^T D^{-1} \Theta \mathbf{e}}{1 + C\mathbf{e}^T D^{-1} \Theta \mathbf{e}} \leq \frac{p'_t}{C}.$$

The conclusion follows.

Part (vi) From Part (iii), we know that $\partial_x Q(x,t)$ is uniformly bounded with respect to x and t. Therefore $Q(x,t)$ is uniformly Lipschitz continuous in x and f_0 is locally Lipschitz. The rest follows from [4, Theorem 2.7.2]. ■

Note that Part (ii) of the proposition indicates that a unit increase of the leader's supply will result in a less than one unit decrease by the aggregate supply of followers' in each demand scenario. See a similar observation in [5,26]. Note also that assumption (16) can be replaced by

$$p'_q(q,t) + qp''_{qq}(q,t) < 0, \quad \text{for } q \geq 0, \quad \text{and} \quad t \in [0,T], \tag{20}$$

in that p is twice continuously differentiable and $x + Q(x,t) \in [0,(N+1)q^u]$ and $t \in [0,T]$, (20) implies (16). Finally, we note that since $\mathbf{q}(x,t)$ is piecewise smooth, a tighter estimation of the subdifferential (or derivative) can be possibly obtained with similar analysis to [24]. In extreme case, if for $t \in [0,T]$,

$$\max_{i=1,\cdots,N} c'_i(0) < p(x + Q(x,t), t), \tag{21}$$

then we can prove that (9) satisfies strict complementarity condition and subsequently $F(\mathbf{q}, x, t)$, $\mathbf{q}(x,t)$, $Q(x,t)$ and $f_0(x)$ are continuously differentiable at $(\mathbf{q}(x,t), x, t)$.

3.4 *Existence of the stochastic Stackelberg–Nash–Cournot equilibria*

With Propositions 3.3 and 3.4, we are able to show the existence of the Stochastic Stackelberg–Nash–Cournot Equilibria.

THEOREM 3.5 *A stochastic Stackelberg–Nash–Cournot equilibrium* $(x^*, q_1(x^*, \cdot), \ldots, q_N(x^*, \cdot))$ *exists under Assumptions* 2.2, 2.3 *and* 2.5 *where* $x^* \in [0, q^u)$.

Proof We only consider (6). The discrete case can be dealt with similarly. By Proposition 3.3 (iii), $Q(x, t)$ Lipschitz and hence $f_0(x)$ is also Lipschitz. By Proposition 3.4 (vi),

$$\partial_x f_0(x) \subset \int_0^T (p(x + Q(x, t), t) + xp'(x + Q(x, t))(1 + \partial_x Q(x, t)))\rho(t)\, dt - c_0'(x). \quad (22)$$

For $x \geq q^u$,

$$\begin{aligned}\int_0^T p(x + Q(x, t), t)\rho(t)\, dt - c_0'(x) &\leq \int_0^T p(x, t)\rho(t)\, dt - c_0'(x) \\ &\leq \max_{t \in [0, T]} (p(x, t) - c_0'(x)) \\ &< 0. \end{aligned} \quad (23)$$

The last inequality is due to Assumption 2.5. Note also that by Proposition 3.4 (ii), $\partial_x Q(x, t) \subset (-1, 0)$. Combining this with (22) and (23), we have

$$\partial_x f_0(x) \subset (-\infty, 0)$$

for $x \geq q^u$. This shows $f_0(x)$ strictly decreasing for $x \geq q^u$. Thus there exists $x^* \in [0, q^u)$ such that f_0 achieves maximum at x^*.

Given x^*, it follows from Proposition 2.4 that $f_i(q_i)$ is strictly concave for $i = 1, \ldots, N$. By [19, Theorems 1 and 2], there is a unique Cournot equilibrium $q^*(x^*, t)$ among followers. ∎

Note that although, as we commented following Assumption 2.5, the leader would not wish to supply more than q^u under the assumption, it seems technically difficult to prove $x^* < q^u$ due to the possible concavity of $f_0(x)$. The nonconcavity problem may also contribute to possible multiplicity of stochastic Stackelberg–Nash–Cournot equilibria.

4. A smoothing MPCC approach for discrete SSNCE problem

In this section, we discuss the numerical methods for the solution of DSSNCE model (5). De Wolf and Smeers [5] considered a two-stage method, originally proposed by Sherali *et al.* [26] for solving (5). The basic idea is to calculate $f_0(x)$ at a set of points spread over the interval $[0, q^u]$ and to linearize $Q(x, t_l)$ over each interval of two adjacent points. The leader's objective function is then concave over the interval and a maximizer is found and added to the set of grids. The approximate maximum is chosen from the calculated values of f_0 on the points. Error bounds are estimated at each step.

In this section, we propose a new approach for solving DSSNCE problem (5). Our idea is to transform the DSSNCE problem into an MPCC and then solve the latter with available numerical methods. Following the discussion in the preceding section, (5) can be transformed into the following program

$$\begin{array}{ll} \max & \sum_{l=1}^{L} \pi_l x p\big(x + \mathbf{e}^T \mathbf{q}(x, t_l), t_l\big) - c_0(x) \\ \text{s.t.} & x \geq 0, \\ & \mathbf{q}(x, t_l) \quad \text{solves } 0 \leq \mathbf{q} \perp G(\mathbf{q}, x, t_l) \geq 0, \quad l = 1, \ldots, L, \end{array}$$

where $\pi_l > 0, l = 1, \ldots, L$, $\sum_{l=1}^{L} \pi_l = 1$. It is not difficult to see that the program above is essentially an implicit deterministic mathematical program with complementarity constraints. The implicity is in the sense that the followers' decision variables $\mathbf{q}(x, t_l)$ are expressed as an implicit function of the leader's decision variable x. Unless $\mathbf{q}(x, t_l)$ can be explicitly obtained in terms of x, this form disadvantages the use of existing numerical methods for MPCC. Therefore, for the sake of numerical solution of the program, we reformulate it by changing variables as follows:

$$\begin{array}{ll} \max & \sum_{l=1}^{L} \pi_l x p\big(x + \mathbf{e}^T \mathbf{q}_l, t_l\big) - c_0(x) \\ \text{s.t.} & x \geq 0, \\ & 0 \leq \mathbf{q}_l \perp G(\mathbf{q}_l, x, t_l) \geq 0, \quad l = 1, \ldots, L. \end{array} \tag{24}$$

In this form, the leader's decision variable and the followers' decision variables are treated equally in the constraints. The problem (24) is a special MPCC and can be solved existing numerical methods for general MPCC, such as the smoothing method of Facchinei *et al.* [6] and the smooth SQP method of Jiang and Ralph [12]. In what follows, we will use the smoothing method by Facchinei *et al.* [6] to deal with (24) with a slightly different treatment of the complementarity problem. Instead of discussing algorithmic details, our focus here is to obtain a bound on the difference between the optimum of the original problem and that of a smoothing program.

Using the 'min' function as we discussed in the preceding section, we can rewrite (24) as the following nonsmooth equality constrained program

$$\begin{array}{ll} \max & \sum_{l=1}^{L} \pi_l x p\big(x + \mathbf{e}^T \mathbf{q}_l, t_l\big) - c_0(x) \\ \text{s.t.} & x \geq 0, \\ & \min(\mathbf{q}_l, G(\mathbf{q}_l, x, t_l)) = 0, \quad l = 1, \ldots, L. \end{array} \tag{25}$$

To tackle the nonsmoothness in the equality constraints, we consider the following smoothing scheme

$$\begin{array}{ll} \max & \sum_{l=1}^{L} \pi_l x p\big(x + \mathbf{e}^T \mathbf{q}_l, t_l\big) - c_0(x) \\ \text{s.t.} & x \geq 0, \\ & \Psi(\mathbf{q}_l, x, t_l, \mu) = 0, \quad l = 1, \ldots, L, \end{array} \tag{26}$$

where

$$\Psi(\mathbf{q}_l, x, t_l, \mu) \equiv \begin{pmatrix} \psi(\mathbf{q}_{l1}, G_1(\mathbf{q}_l, x, t_l), \mu) \\ \vdots \\ \psi(\mathbf{q}_{lN}, G_N(\mathbf{q}_l, x, t_l), \mu) \end{pmatrix} \tag{27}$$

and

$$\psi(a, b, c) = -\frac{1}{2}\left(\sqrt{(a-b)^2 + c^2} - (a+b)\right), \tag{28}$$

and $\mathbf{q}_{li}$ denotes the ith component of vector $\mathbf{q}_l$. Our idea is to solve (25) by solving the smoothed program (26). Note that the smoothing methods for solving general MPCC problems have been well-discussed by Facchinei *et al.* [6] and Jiang and Ralph [12]. The ψ function (strictly speaking -2ψ) was first considered by Kanzow and used in [6,12] for dealing with general MPCCs. ψ is a smoothing of the 'min' function $\min(a, b)$ in that $\psi \in C^\infty$ for $c \neq 0$ and

$$\psi(a, b, 0) = \min(a, b).$$

It is well-known that ψ is globally Lipschitz and at $(a, b, 0)$, it satisfies the *strong Jacobian consistency* [2] in the sense of the following relationship between the Clarke generalized Jacobian of ψ at $(a, b, 0)$ and the Clarke generalized Jacobian of ϕ at (a, b)

$$\Pi_{a,b}\partial\psi(a, b, 0) = \partial \min(a, b).$$

Therefore

$$\Psi(\mathbf{q}_l, x, t_l, 0) = F(\mathbf{q}_l, x, t_l)$$

and by the strong Jacobian consistency of ψ, it is easy to verify

$$\Pi_{\mathbf{q}_l}\partial\Psi(\mathbf{q}_l, x, t_l, 0) = \partial_{\mathbf{q}_l}F(\mathbf{q}_l, x, t_l) \tag{29}$$

where

$$\Pi_{\mathbf{q}_l}\partial\Psi(\mathbf{q}_l, x, t_l, 0) = \left\{R\colon (R, U, V) \in \partial F(\mathbf{q}_l, x, t), R \in \mathbb{R}^{N\times N}, U \in \mathbb{R}^N, V \in \mathbb{R}^N\right\}$$

and

$$\partial_{\mathbf{q}_l}F(\mathbf{q}_l, x, t_l) = \left\{I - \Theta + \Theta\nabla_{\mathbf{q}_l}G(\mathbf{q}_l(x, t_l), x, t_l)\colon \Theta \in \mathrm{diag}([0, 1], \ldots, [0, 1])\right\}.$$

Lemma 4.1 *Under Assumptions* 2.2, 2.3 *and* 2.5, *there exists* $\mu_0 > 0$ *and a unique implicit function* $\mathbf{q}_l(x, \mu)$ *such that* $\mathbf{q}_l(x, 0) = \mathbf{q}_l(x)$,

$$\Psi(\mathbf{q}_l(x, \mu), x, t_l, \mu) = 0, \quad \textit{for } \mu \in [0, \mu_0],$$

and $\mathbf{q}_l(x,\mu)$ *is continuously differentiable for* $\mu \neq 0$. *Moreover there exists a constant* $C > 0$, *such that*

$$\|\mathbf{q}_l(x,\mu) - \mathbf{q}_l(x)\| \leq \sqrt{N} C\mu, \quad \textit{for } \mu \in [0, \mu_0]. \tag{30}$$

Proof The first part of the lemma is well-known. See [6,12,13] for a similar claim. We only prove (30). It is easy to check that

$$\nabla_{\mathbf{q}_l} \Psi(\mathbf{q}_l, x, t_l, \mu) \in \partial_{\mathbf{q}_l} F(\mathbf{q}_l, x, t_l), \quad \forall \mu > 0.$$

By Proposition 3.3 (i), $\partial_{\mathbf{q}_l} F(\mathbf{q}_l, x, t_l)$ is uniformly nonsingular, that is, there exist a constant $C > 0$ such that

$$\max_{V \in \partial_{\mathbf{q}_l} F(\mathbf{q}_l, x, t_l)} \|V^{-1}\| \leq C.$$

Therefore

$$\|\nabla_{\mathbf{q}_l} \Psi(\mathbf{q}_l, x, t_l, \mu)^{-1}\| \leq \max_{V \in \partial_{\mathbf{q}_l} F(\mathbf{q}_l, x, t_l)} \|V^{-1}\| \leq C.$$

For $\mu \neq 0$,

$$\nabla_\mu \Psi(\mathbf{q}_l(x,\mu), x, t_l, \mu) = \begin{pmatrix} \dfrac{\mu}{\sqrt{(q_{l1}(x,\mu) - G_1(\mathbf{q}(x,\mu), x, t_l))^2 + \mu^2}} \\ \vdots \\ \dfrac{\mu}{\sqrt{(q_{lN}(x,\mu) - G_N(\mathbf{q}(x,\mu), x, t_l))^2 + \mu^2}} \end{pmatrix} \in \begin{pmatrix} (0,1) \\ \vdots \\ (0,1), \end{pmatrix}.$$

For $\mu = 0$,

$$\partial_\mu \Psi(\mathbf{q}_l(x,0), x, t_l, 0) = \begin{pmatrix} [0,1] \\ \vdots \\ [0,1] \end{pmatrix}.$$

Since $\mathbf{q}_l(x,\mu)$ is continuously differentiable for $\mu > 0$, then

$$\mathbf{q}_l(x,\mu) - \mathbf{q}_l(x,0) = \mu \int_0^1 \nabla_\mu \mathbf{q}(x, \mu\nu)\, d\nu.$$

Therefore for $\mu \in [0, \mu_0]$,

$$\begin{aligned} \|\mathbf{q}_l(x,\mu) - \mathbf{q}_l(x,0)\| &\le \mu \int_0^1 \|\nabla_\mu \mathbf{q}(x,\mu\nu)\| d\nu \\ &= \mu \int_0^1 \|-\nabla_{\mathbf{q}_l} \Psi(\mathbf{q}_l(x,\mu\nu), x, t_l, \mu\nu)^{-1} \nabla_\mu \Psi(\mathbf{q}_l(x,\mu\nu), x, t_l, \mu\nu)\| d\nu \\ &\le \sqrt{N} C \mu. \end{aligned}$$

This completes the proof. ■

Lemma 4.1 shows that for μ sufficiently small and $x \ge 0$, the equality constraints of program (26) define a unique solution $(\mathbf{q}_1(x,\mu), \ldots, \mathbf{q}_L(x,\nu))$. It also gives a bound on the difference between $\mathbf{q}_l(x,\mu)$ and $\mathbf{q}_l(x,0)$, which will be used in the next theorem. With a similar analysis to the proof of Theorem 3.5, we can show that the smoothed program (26) is well defined in the sense that its global maximum is achievable.

Let $\{\mu_k\}$ be a sequence such that $\mu_k \downarrow 0$ as $k \to \infty$. Let $\{(x_k, \mathbf{q}_l(x_k, \mu_k)_{l=1,\ldots,L})\}$ be a sequence of solutions of (26) corresponding to $\mu = \mu_k$. Let

$$f_0^{\mu}(x) \equiv \sum_{l=1}^{L} \pi_l x p\big(x + \mathbf{e}^T \mathbf{q}_l(x,\mu), t_l\big) - c_0(x). \tag{31}$$

THEOREM 4.2 *Under Assumptions* 2.2, 2.3 *and* 2.5,

(i) *any accumulation point of* $\{(x_k, \mathbf{q}_l(x_k, \mu_k)_{l=1,\ldots,L})\}$ *is a solution of* (24);
(ii) *there exists a constant* $\hat{C} > 0$ *such that*

$$|f_0^{\mu_k}(x_k) - f_0^*| \le \hat{C}\mu_k,$$

where f_0^* *denote the global maximum of* (24).

Proof
Part (i) First, we prove that $f_0^{\mu_k}(x)$ achieves its maximum on $[0, q^u)$. Let $x \ge 0$, $\mu \in [0, \mu_0]$ where μ_0 is given in Lemma 4.1. Let $\mathbf{q}_l(x,\mu)$ be the solution of $\Psi(\mathbf{q}_l, x, t_l, \mu) = 0$. Then

$$f_0^{\mu}(x) = \sum_{l=1}^{L} \pi_l x p(x + Q_l(x,\mu), t_l) - c_0(x),$$

where $Q_l(x,\mu) = \mathbf{e}^T \mathbf{q}_l(x,\mu)$. Moreover

$$\begin{aligned} (f_0^{\mu})'(x) = &\sum_{l=1}^{L} \pi_l p(x + Q_l(x,\mu), t_l) - c_0'(x) \\ &+ \sum_{l=1}^{L} \pi_l x p_q'(x + Q_l(x,\mu), t_l)\big(1 + (Q_l)_x'(x,\mu)\big). \end{aligned}$$

By the classical implicit function theorem and the upper semicontinuity of the Clarke generalized Jacobian, we obtain

$$(Q_l)'_x(x,\mu) = -\mathbf{e}^T \nabla_{\mathbf{q}_l} \Psi(\mathbf{q}_l, x, t_l, \mu)^{-1} \nabla_x \Psi(\mathbf{q}_l, x, t_l, \mu)$$
$$\in -\mathbf{e}^T \nabla_{\mathbf{q}_l} \Psi(\mathbf{q}_l, x, t_l, 0)^{-1} \nabla_x \Psi(\mathbf{q}_l, x, t_l, 0) + \sigma(-1,1), \quad (32)$$

where σ depends on μ and it can be arbitrarily small so long as μ is sufficiently small. Combining (29) and (32), and using Part (ii) of Proposition 3.4, we obtain

$$(Q_l)'_x(x,\mu) = (-1,0) + \sigma(-1,1).$$

Note that for $x \geq q^u$, by Assumption 2.5,

$$\sum_{l=1}^{L} \pi_l p(x + Q_l(x,\mu), t_l) - c_0'(x) \leq \max_{l=1,\dots,L} p(x, t_l) - c_0'(x) < 0.$$

This shows for μ_0 sufficiently small

$$(f_0^{\mu})'(x) < 0, \quad \text{for } x \geq q^u, \quad \mu \in [0, \mu_0].$$

Therefore $f_0^{\mu}(x)$ achieves its maximum on $[0, q^u)$ for $\mu \in [0, \mu_0]$.

Since $x_k \in [0, q^u]$, the sequence $\{(x_k, \mathbf{q}_l(x_k, \mu_k)_{l=1,\dots,L})\}$ is bounded. Let $(\hat{x}, \mathbf{q}_l(\hat{x}, 0)_{l=1,\dots,L})$ be an accumulation point of the sequence. Let $(x^*, \mathbf{q}_l(x^*, 0)_{l=1,\dots,L})$ be a solution of (24). Then

$$f_0^{\mu_k}(x^*) - f_0(x^*) = \sum_{l=1}^{L} \pi_l x^* [p(x^* + Q_l(x^*, \mu_k), t_l) - p(x^* + Q_l(x^*, 0), t_l)]$$
$$= \sum_{l=1}^{L} \pi_l x^* p_q'(x^* + \theta_k Q_l(x^*, \mu_k) + (1-\theta_k) Q_l(x^*, 0), t_l)(Q_l(x^*, \mu_k) - Q_l(x^*, 0))$$

where $\theta_k \in (0,1)$. By Lemma 4.1,

$$|Q_l(x^*, \mu_k) - Q_l(x^*, 0)| = \|\mathbf{e}\| \|\mathbf{q}_l(x^*, \mu_k) - \mathbf{q}_l(x^*, 0)\| \leq NC\mu_k,$$

where C is the constant given in Lemma 4.1. Moreover since p_q' is bounded, and $x^* \in [0, q^u]$, then there exists $\hat{C} > 0$ such that

$$f_0^{\mu_k}(x_k) \geq f_0^{\mu_k}(x^*) \geq f_0(x^*) - \hat{C}\mu_k. \quad (33)$$

Note that here $\hat{C}$ does not depend on L. By taking a subsequence if necessary and letting $k \to \infty$, we have

$$f_0(\hat{x}) \geq f_0(x^*)$$

On the other hand, by definition

$$f_0(\hat{x}) \le f_0(x^*).$$

This shows

$$f_0(\hat{x}) = f_0(x^*).$$

Part (ii) follows straightforwardly from (33). ■

Note that Lin *et al* . [16] recently discussed a smoothing method for a discrete stochastic program with linear complementarity constraints and presented a similar analysis on error bounds between a smoothing program and the original program. Despite the differences of the original programs and smoothing functions, it seems that our results are stronger in that the error bound given in Theorem 4.2 does not depend on L, and this apparently has to do with the estimation method. See [16] for detail.

5. Discretization of the CSSNCE problem

In this section, we discuss the discretization of CSSNCE problem (6) where the random shock in demand has a continuous distribution. In DSSNCE problem (5), there exists a finite number of followers' equilibria each of which corresponds to a particular demand scenario. The problem is essentially a deterministic MPCC and can be dealt with using available numerical methods for MPCC. In CSSNCE problem, there exists an uncountable number of demand scenarios. For each x, the leader's supply, the equilibrium among followers can be written as $\mathbf{q}(x, \epsilon(\omega))$ at a demand scenario $p(q, \epsilon(\omega))$. Since $\epsilon(\omega)$ takes value on $[0, T]$, the equlibria $\mathbf{q}(x, \cdot) : [0, T] \to \mathbb{R}^N$ forms a one-dimensional manifold in $\mathbb{R}^N$. It is evident that unless we discretize $\mathbf{q}(x, \cdot)$ over $[0, T]$, we are not able to deal with (6) with any available numerical method for MPCC.

Our idea of discretization is as follows: we choose a set of points $\{t_0, t_1, \dots, t_L\}$ on the interval $[0, T]$ and calculate $\mathbf{q}(x, t_i)$ for each t_l, and then calculate the numerical integration of $xp(x + Q(x, t), t)\rho(t)$ based on its values on these points. The calculation of $\mathbf{q}(x, t_l)$, $i = 0, \dots, L$ and the numerical integration can be accomplished by the solution of a single deterministic MPCC as we discussed in the preceding section for DSSNCE problem (5).

Let $\mathcal{T}_L = \{t_0, t_1, \dots, t_L\}$ and $\mathcal{A}$ be a numerical integration method which only require the calculation of function values. The discretized CSSNCE program of (6) can be written as

$$\begin{aligned} &\max f_0^K(x) \equiv \mathcal{A}\big(p\big(x + \mathbf{e}^T\mathbf{q}(x, \cdot), \cdot\big)\rho(\cdot), \mathcal{T}_L\big) - c_0(x) \\ \text{(DCSSNCE)}\quad &\text{s.t.}\quad x \ge 0, \\ &\mathbf{q}(x, t_l) \quad \text{solves } 0 \le \mathbf{q} \perp G(\mathbf{q}, x, t_l) \ge 0, \quad l \in \mathcal{T}_L. \end{aligned} \tag{34}$$

Many numerical integration methods can be used. In this section we consider the Trapezoidal method, where

$$\mathcal{T}_L = \{t_l\colon t_0 = 0, t_l = t_{l-1} + \frac{T}{L}, \text{ for } i = 1, \dots, L\}$$

and

$$f_0{}^L(x) = \frac{T}{L}\sum_{l=0}^{L} xp(x + Q(x, t_l), t_l)\rho(t_l) - c_0(x).$$

The discretized DCSSNCE problem (34) becomes

$$\begin{aligned} &\max f_0{}^L(x) \equiv \frac{T}{L}\sum_{l=0}^{L} xp(x + Q(x, t_l), t_l)\rho(t_l) - c_0(x) \\ &\text{s.t.} \quad x \geq 0, \\ &\qquad \mathbf{q}(x, t_l) \quad \text{solves } 0 \leq \mathbf{q} \perp G(\mathbf{q}, x, t_l) \geq 0, \quad l = 1, \ldots, L. \end{aligned} \tag{35}$$

In what follows, we shall focus on discussing the approximation of CSSNCE problem (6) with the program defined in (35). The fundamental issue is the uniform approximation of the objective function. Since $Q(x, t)$ is implicitly defined, the numerical integration approximation requires the uniform approximation of $Q(x, \cdot)$ and $Q_x(x, \cdot)$.

Throughout this section, we consider a special class of demand function which satisfies the following:

$$p'_t(q, t) > 0 \text{ and } p''_{q,t}(q, t) = 0. \tag{36}$$

This means the demand curves are parallel to each other under random shock. This type of demand function was considered by Klemperer and Meyer [14] in the study of supply function equilibria in an oligopoly market and it has been used by many others particularly in modelling competitions in modern electricity markets [10]. Anderson and Xu [1] observed that a demand function satisfying (36) can be reformulated as

$$p(q, \epsilon(\omega)) = p(q) + \epsilon(\omega) \tag{37}$$

after a change of random variable. Throughout this section, we assume that the market demand function takes a form of (37) and the density function of $\epsilon(\cdot)$ is $\rho(t)$ with support set $[0, T]$. We need to make some more specific assumptions on $p(q)$.

Assumption 5.1 $p(q)$ is twice continuously differentiable and $qp(q)$ is strictly concave for $q \geq 0$. Moreover there exists q^u such that

$$c'_i(q) \geq p(q) + T, \quad \text{for } q \geq q^u, \quad i = 0, 1, \ldots, N.$$

Under Assumption 5.1, Assumptions 2.3 and 2.5 are satisfied. Also since $qp(q)$ is twice continuously differentiable, the strict concavity of $qp(q)$ in assumption implies the strong concavity of $qp(q)$ over $[0, q^u]$.

THEOREM 5.2 *Let $f_0{}^L$ be defined as in* (35). *Suppose $\rho(t)$ is continuously differentiable and*

$$\int_0^T |\rho'(t)|\, dt < +\infty.$$

Then under Assumptions 2.2 *and* 5.1, *there exists a constant* $\tilde{C} > 0$ *such that*

$$\left|f_0{}^L(x) - f_0(x)\right| \leq \frac{\tilde{C}T}{L}. \tag{38}$$

Proof By definition

$$f_0(x) - f_0{}^L(x) = \sum_{l=1}^{L} \int_{t_{l-1}}^{t_l} x[(p(x + Q(x,t)) + t)\rho(t) - (p(x + Q(x,t_l)) + t_l)\rho(t_l)]\, dt.$$

Let

$$\Delta(x,t) = (p(x + Q(x,t)) + t)\rho(t) - (p(x + Q(x,t_l)) + t_l)\rho(t_l), \quad \text{for } t \in (t_{l-1}, t_l).$$

We need to estimate $\Delta(x,t)$. Note first that the demand function $p(q,t) = p(q) + t$ satisfies Assumption 2.3. Moreover, $p''_{q,t}(q,t) = 0$, $p'_t(q,t) = 1$ and $qp(q,t)$ is uniformly strongly concave over $[0, q^u]$. By Part (iii) of Proposition 3.4, there exists a constant $C > 0$, such that

$$\partial_t Q(x,t) \subset \left(0, \frac{1}{C}\right].$$

Note also that both $p(q)$ and $p'(q)$ are strictly decreasing. Thus for $t \in [t_{l-1}, t_l]$,

$$\begin{aligned} |\Delta(x,t)| &\leq \frac{T}{L} \max_{t\in[t_{l-1},t_l],\, \xi\in\partial_t Q(x,t)} \left|(p'(x + Q(x,t))\xi + 1)\rho(t) + (p(x + Q(x,t)) + t)\rho'(t)\right| \\ &< \frac{T}{L}\left(\frac{1}{C}(|p'((N+1)q^u)| + 1) \max_{t\in[t_{l-1},t_l]} \rho(t) + (p(x) + T) \max_{t\in[t_{l-1},t_l]} |\rho'(t)|\right). \end{aligned}$$

Since both $\rho(t)$ and $\rho'(t)$ are continuously differentiable and integrable over $[0, T]$, there exists $\delta > 0$ such that

$$\frac{T}{L}\sum_{l=1}^{L} \max_{t\in[t_{l-1},t_l]} |\rho(t)| < 1 + \delta$$

and

$$\frac{T}{L}\sum_{l=1}^{L} \max_{t\in[t_{l-1},t_l]} |\rho'(t)| < \int_0^T |\rho'(t)|\, dt + \delta.$$

Let

$$\tilde{C} = q^u\left[\frac{1}{C}(|p'((N+1)q^u)| + 1)(1 + \delta) + (p(0) + T)\left(\int_0^T |\rho'(t)|\, dt + \delta\right)\right].$$

The conclusion follows. ■

The theorem shows that if the derivative of the density function $\rho(t)$ is integrable over its support, then we can obtain a uniform approximation to the program (6) with (35). This together with the smoothing method for the discrete program discussed in the previous section should be adequate to solve CSSNCE problem (6). We summarize this in the following corollary.

COROLLARY 5.3 *Suppose that CSSNCE problem in* (6) *is discretized through program* (34) *and program* (34) *is solved through a smoothing program* (26). *Let* $\mu = 1/L$. *Suppose the assumptions in Theorem* 5.2 *are satisfied. Then*

$$\left| {f_0}^{1/L}(x_L) - {f_0}^* \right| \leq \frac{1}{L}(\hat{C} + \tilde{C}T)$$

where x_L *solves* (26) *and* ${f_0}^*$ *denotes the global maximum of* (6).

A potential disadvantage of the proposed discretization approach is that the program (35) is of high dimension when L is large. It is worthwhile to consider some numerical integration method which requires a small set of $\mathcal{T}_L$. This is beyond the focus of this article.

Note also that Theorem 5.2 can be generalized to the case when $\rho(t)$ is finitely piecewise smooth and/or the support set is unbounded (although the unboundedness of the support set is not sensible in this context). A more challenging case is that $\epsilon(\omega)$ is a vector of random variables and consequently $f_0(x)$ involves a multi-dimensional integration. It would be interesting to discuss a discretization method for such a case, although again there is no need in this context.

6. Multi-leader stochastic Stackelberg–Nash–Cournot equilibrium

In the preceding sections, we have discussed stochastic Stackelberg–Nash–Cournot equilibria in an oligopoly market and numerical methods for the solution of a stochastic Stackelberg–Nash–Cournot equilibrium. In this section, we present a brief discussion on a related but slightly different problem. We assume that there are two leaders and N followers in the market. The competition between the leader and followers remains the same as described previously while competition between two leaders are assumed to be of Nash–Cournot, that is, each choosing its supply by holding the others offer as constant and the followers' reaction as known. This discussion is inspired by a recent paper of Pang and Fukushima [21] which presented a general description of multi-leader–follower model. Their model is mathematically linked to quasi-variational inequality problems and is practically linked to the competition in modern electricity markets. See [21] for details.

Throughout this section, we use different notation for leaders. We use A and B to denote the two leaders and x, y denote their decision variables, that is, the supply to the market. Assume that leader B's supply is y. Then leader A's decision problem is

$$\begin{aligned}
\max \quad & f_A(x) \equiv x \int_0^T p\left(x + y + \sum_{i=1}^N q_i, t\right)\rho(t)\, dt - c_A(x) \\
\text{s.t.} \quad & x \geq 0, \\
& q_i^A(x+y,t) \in \arg\max_{q_i \geq 0}\left(q_i p\left(x + y + q_i + \sum_{k=1,\, k\neq i}^N q_k, t\right) - c_i(q)\right), \\
& t \in \mathcal{T}, i = 1, \ldots, N,
\end{aligned} \tag{39}$$

where c_A denotes leader A's cost function. In a similar manner, by assuming leader A's offer of x, leader B's decision problem is

$$\begin{aligned}
\max \quad & f_B(y) \equiv y \int_0^T p\left(x + y + \sum_{i=1}^N q_i, t\right)\rho(t)\, dt - c_B(y) \\
\text{s.t.} \quad & y \geq 0, \\
& q_i^B(x+y,t) \in \arg\max_{q_i \geq 0}\left(q_i p\left(x + y + q_i + \sum_{k=1,\, k\neq i}^N q_k, t\right) - c_i(q)\right), \\
& t \in \mathcal{T}, \quad i = 1, \ldots, N,
\end{aligned} \tag{40}$$

where c_B denotes leader B's cost function.

Note that $(q_1^A(x+y,t), \ldots, q_N^A(x+y,t))$ is a Nash–Cournot equilibrium of the followers which, leader A believes, will be reached if he supplies x and leader B's supply y is held constant. Similarly $(q_1^B(x+y,t), \ldots, q_N^B(x+y,t))$ is a Nash–Cournot equilibrium of the followers which, leader B believes, will be reached if he supplies y and leader A's supply x is held constant. In general, when there are multiple followers' equilibria, $(q_1^A(x+y,t), \ldots, q_N^A(x+y,t))$ may differ from $(q_1^B(x+y,t), \ldots, q_N^B(x+y,t))$. See [21] for a general discussion. However, under our assumption in the previous discussion on followers, it is obvious that $(q_1^A(x+y,t), \ldots, q_N^A(x+y,t))$ equals $(q_1^B(x+y,t), \ldots, q_N^B(x+y,t))$ and we denote it by $(q_1(x+y,t), \ldots, q_N(x+y,t))$. We are interested in an equilibrium $(x^*, y^*, \mathbf{q}(x^*+y^*, \cdot))$ such that x^* solves (39) and y^* solves (40).

Let

$$Q(x+y,t) \equiv \sum_{i=1}^N q_i(x+y,t)$$

In general, $Q(x+y,t)$ is not necessarily convex in x or y and henec $f_A(x)$ and $f_B(y)$ may not be concave. As a result, two-leaders' stochastic Stackelberg–Nash–Cournot equilibrium may not exist.

We consider a particular case when market demand function is affine, that is,

$$p(q,t) = a - bq + t,$$

the cost function of follower i is quadratic, that is,

$$c_i(q_i) = \alpha_i + \beta_i q_i + \frac{1}{2}\gamma_i q_i^2, \quad i = 1, \ldots, N.$$

The followers' equilibrium problem can be simplified as

$$0 \leq \mathbf{q} \perp \left[\beta - (a + t - b(x + y))\mathbf{e} + \left(b\mathbf{e}\mathbf{e}^T + bI + \Gamma\right)\mathbf{q}\right] \geq 0 \tag{41}$$

where, $I \in \mathbb{R}^{N \times N}$ is the identity matrix, $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_N)$ and $\beta = (\beta_1, \ldots, \beta_N)^T$. Our anticipation is to find an equilibrium in this relatively simple case.

Proposition 6.1 *If*

$$\max_{i \in \{1,\ldots,N\}} \beta_i < a + t - b(x + y), \tag{42}$$

then the followers have a unique equilibrium

$$\mathbf{q}(x + y, t) = R^{-1}[(a + t - b(x + y))\mathbf{e} - \beta]. \tag{43}$$

where

$$R = bI + b\mathbf{e}\mathbf{e}^T + \Gamma.$$

Proof The condition (42) implies the strict complementarity condition of the followers' Nash–Cournot equilibrium problem (41). It is easy to verify that the $\mathbf{q}(x, t)$ given in (43) is the unique equilibrium solving (41). ■

In order to discuss two leaders Stackelberg–Nash–Cournot equilibria, we also need to consider specific forms of leaders' cost functions.

Suppose that leader A's cost function is

$$C_A(x) = \alpha_A + \beta_A x + \frac{1}{2}\gamma_A x^2$$

and leader B's cost function is

$$C_B(y) = \alpha_B + \beta_B y + \frac{1}{2}\gamma_B y^2.$$

Then the leader A's profit maximization problem is

$$\begin{aligned}\max_{x \geq 0} f_A(x) \equiv & x\left\{\int_0^T \left[a + t - b\left(x + y + (a + t - bx)\mathbf{e}^T R^{-1}\mathbf{e} - \mathbf{e}^T R^{-1}\beta\right)\right]\rho(t)\,dt\right\} \\ & - \alpha_A - \beta_A x - \frac{1}{2}\gamma_A x^2\end{aligned}$$

and leader B's profit maximization problem is

$$\max_{y\geq 0} f_B(y) \equiv y\left\{\int_0^T \left[a+t-b\left(y+x+(a+t-by)\mathbf{e}^T R^{-1}\mathbf{e}-\mathbf{e}^T R^{-1}\beta\right)\right]\rho(t)\,dt\right\}$$
$$-\alpha_B-\beta_B y-\frac{1}{2}\gamma_B y^2.$$

Note that

$$f_A'(x)=(\lambda-\gamma_A)x-by+a+\eta-\beta_A=0$$

and

$$f_B'(y)=-bx+(\lambda-\gamma_B)y+a+\eta-\beta_B=0$$

where

$$\lambda=-2b+2b^2\mathbf{e}^T R^{-1}\mathbf{e}$$

and

$$\eta=(1+b\mathbf{e}^T R^{-1}\mathbf{e})\int_0^T t\rho(t)\,dt.$$

It is easy to check that

$$\mathbf{e}^T R^{-1}\mathbf{e}<\frac{1}{b}.$$

Thus $\lambda<0$ and

$$(\lambda-\gamma_A)(\lambda-\gamma_B)-b^2>3b^2.$$

Proposition 6.2 *Let*

$$\hat{x}=\frac{(\beta_A-a-\eta)(\lambda-\gamma_B)+b(\beta_B-a-\eta)}{(\lambda-\gamma_A)(\lambda-\gamma_B)-b^2}$$

and

$$\hat{y}=\frac{(\beta_B-a-\eta)(\lambda-\gamma_A)+b(\beta_A-a-\eta)}{(\lambda-\gamma_B)(\lambda-\gamma_A)-b^2}.$$

There exists a unique Stackelberg–Nash–Cournot equilibrium $(x^*, y^*, q(x^* + y^*, \cdot))$ *where*

$$(x^*, y^*) = \begin{cases} (\hat{x}, \hat{y}) & \text{if } \hat{x} > 0, \text{ and } \hat{y} > 0 \\ \left(0, \dfrac{a + \eta - \beta_B}{\gamma_B - \lambda}\right) & \text{if } \hat{x} \le 0, \text{ and } \hat{y} > 0 \\ \left(\dfrac{a + \eta - \beta_A}{\gamma_A - \lambda}, 0\right) & \text{if } \hat{x} > 0, \text{ and } \hat{y} \le 0 \end{cases}$$

and

$$\mathbf{q}(x^* + y^*, \cdot) = R^{-1}[(a + \cdot - b(x^* + y^*))\mathbf{e} - \beta].$$

Proof There exist three possibilities: $\hat{x} > 0, \hat{y} > 0$; $\hat{x} \le 0, \hat{y} > 0$; $\hat{x} > 0, \hat{y} \le 0$. It is easy to verify that each case leads to a unique equilibrium as described. ■

The equilibrium we obtained are based on assumptions that the demand function is affine and the cost functions of firms are quadratic. These assumptions guarantee the existence, uniqueness and linearity of the followers' Nash–Cournot equilibrium as well as the concavity of the leaders' objective functions. It is unclear if there exist such equilibria under general circumstances and further research in this regard will be very interesting.

Acknowledgements

I would like to thank Dr. Daniel Ralph for his helpful comments and suggestions on an earlier version of this article. I would also like to thank two anonymous referees for their valuable comments. The work has been partially supported by the United Kingdom Engineering and Physical Sciences Research Council grant GR/S90850/01.

References

[1] Anderson, E.J. and Xu, H., 2001, *Contracts and Supply Functions in Electricity Markets*. Working paper 01-006, Australian Graduate School of Management, University of New South Wales, Sydney, Australia. To appear in *Journal of Optimization Theory and Applications*.

[2] Chen, X., Qi, L. and Sun, D., 1998, Global and superlinear convergence of the smoothing Newton's method and its application to general box constrained variational inequalities. *Mathematics of Computation*, **67**, 519–540.

[3] Christiansen, S., Patriksson, M. and Wynter, L., 2001, Stochastic bilevel programming in structral optimization. *Structural Optimization*, **21**, 361–371.

[4] Clarke, F.H., 1983, *Optimization and Nonsmooth Analysis* (New York: Wiley).

[5] De Wolf, D. and Smeers, Y., 1997, A stochastic version of a Stackelverg–Nash–Cournot equilibrium model. *Management Sciences*, **43**, 190–197.

[6] Facchinei, F., Jiang, H. and Qi, L., 1999, A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming*, **85**, 81–106.

[7] Ferris, M. and Pang, J.S., 1997, Engineering and economic applications of complementarity problems. *SIAM Review*, **39**, 669–713.

[8] Fukushima, M. and Pang, J.S., 1998, Convergence of a smoothing continuation method for mathematical programs with complementarity constraints. Ill-posed variational problems and regularization techniques (Trier). *Lecture Notes in Economics and Mathematical Systems*, Vol. 477 (Berlin: Springer), pp. 99–110.

[9] Fukushima, M. and Tseng, P., 2002, An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complementarity constraints. *SIAM Journal on Optimization*, **12**, 724–739.
[10] Green, R.J. and Newbery, D.M., 1992, Competition in the British electricity spot market. *Journal of Political Economy*, **100**, 929–953.
[11] Hu, X. and Ralph, D., 2002, A note on sensitivity of value functions of mathematical programs with complementarity constraints. *Mathematical Programming*, **93**, 265–279.
[12] Jiang, H. and Ralph, D., 2000, Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. *SIAM Journal on Optimization*, **10**, 779–808.
[13] Kanzow, C. and Jiang, H., 1998, A continuation method for (strongly) monotone variational inequalities. *Mathematical Programming*, **81**, 103–125.
[14] Klemperer, P.D. and Meyer, M.A., 1989, Supply function equilibria in oligopoly under uncertainty. *Econometrica*, **57**, 1243–1277.
[15] Kleywegt, A.J. and Shapiro, A., 2001, Stochastic optimization. In: G. Salvendy (Ed.) *Handbook of Industrial Engineering*, 3rd edition, (New York, NY: John Wiley & Sons), pp. 2625–2649.
[16] Lin, G.-H., Chen, X. and Fukushima, M., 2003, Smoothing implicit programming approaches for stochastic mathematical programs with linear complementarity constraints, http://www.amp.i.kyoto-u.ac.jp/tecrep/index-e.html.
[17] Lin, G.-H. and Fukushima, M., 2003, A class of stochastic mathematical programs with complementarity constraints, http://www.amp.i.kyoto-u.ac.jp/tecrep/index-e.html.
[18] Luo, Z.-Q., Pang, J.-S. and Ralph, D., 1996, *Mathematical Programs with Equilibrium Constraints* (Cambridge University Press).
[19] Murphy, F.H., Sherali, H.D. and Soyster, A.L., 1982, A mathematical programming approach for determining oligopolistic market equilibrium. *Mathematical Programming*, **24**, 92–106.
[20] Outrata, J., Kocvara, M. and Zowe, J., 1998, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints–Theory, Applications and Numerical Constraints* (Kluwer Publishers).
[21] Pang, J.S. and Fukushima, M., 2002, Quasi-variational inequalities, generalized Nash equilibrium and multi-leader-follower games.
[22] Patriksson, M. and Wynter, L., 1999, Stochastic mathematical programs with equilibrium constraints. *Operations Research Letters*, **25**, 159–167.
[23] Ralph, D. and Scholtes, S., 1997, Sensitivity analysis of composite piecewise smooth equations. *Mathematical Programming*, **76**, 593–612.
[24] Ralph, D. and Xu, H., 1999, *Implicit Smoothing and its Application to Optimization with Piecewise Smooth Equality Constraints* (preprint). Department of Mathematics and Statistics, The University of Melbourne, Parkville, Australia. To appear in *Journal of Optimization Theory and Applications*.
[25] Rosen, J.B., 1965, Existence and uniqueness of equilibrium points for concave *N*-person games. *Econometrica*, **33**, 520–534.
[26] Sherali, H.D., Soyster, A.L. and Murphy, F.H., 1983, Stackelberg–Nash–Cournot equilibria: characterizations and computations. *Operations Research*, **31**, 253–276.
[27] Scholtes, S., 2001, Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, **11**, 918–936.
[28] Sun, D. and Qi, L., 1999, On NCP functions. *Computational Optimization and Applications*, **13**, 201–220.
[29] Yamashita, N. and Fukushima, M., 1997, Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems. *Mathematical Programming*, **76**, 469–497.