Two stage stochastic equilibrium problems with equilibrium constraints: modeling and numerical schemes

August 12, 2010

Dali Zhang [1] and Huifu Xu [2]

Abstract

This paper presents a two stage stochastic equilibrium problem with equilibrium constraints (SEPEC) model. Some source problems which motivate the model are discussed. Monte Carlo sampling method is applied to solve the SEPEC. The convergence analysis on the statistical estimators of Nash equilibria and Nash stationary points are presented.

Key words. Stochastic equilibrium programs with equilibrium constraints, convergence analysis, sample average approximation, Nash stationary point.

1 Introduction

In our earlier work [41], we discussed a one stage stochastic Nash equilibrium model and investigated sample average approximation (SAA) of Nash equilibrium and Nash stationary points. We noted that the model may cover two stage stochastic Nash equilibrium model and included an application of a two stage stochastic equilibrium program with equilibrium constraint (SEPEC) model to study the competition of generators in the electricity wholesale markets with network constraints. However, the work does not explore the unique structure/characteristics of SEPEC model. On the other hand, the SEPEC model, as a natural extension of deterministic EPEC models, has a number of potential applications in a wide domain in engineering design, management and economics. This motivates us to write this note in an attempt to provide an independent discussion of the model and yet not overlap with our earlier work.

Let us start with some literature review. Over the past few years, equilibrium programs with equilibrium constraints (EPEC) and SEPEC have been developed as a new subject in optimization primarily driven by applications in engineering design, management and economics. For instances, Hu and Ralph [15] used EPEC to model bilevel games in a restructured electricity market where each player faces a bilevel optimization problem; Hobbs, Metzler and Pang [14] investigated an oligopolistic market economy consists of several dominant firms in an electric power network, where each generating firm submits bids to an ISO, choosing its bids to maximize profits subject to anticipated reactions by rival firms, and hence the firms' decision problem can be formulated as an EPEC. Yao, Oren and Adler [43] used SEPEC to model generator's competition in a spot with newtwork constraints where the second stage equilibrium constraint is used to characterize independent system operator's optimal decision making problem on generator's dispatch and power flow in the network. Zhang, Xu and Wu [44] developed a two stage SEPEC model to study generator's competition in electricity forward and spot markets and

[1]Department of Industrial and Systems Engineering, National University of Singapore, 117576, Singapore, Email: isezd@nus.edu.sg,

[2]School of Mathematics, University of Southampton, Southampton, UK, Email: h.xu@soton.ac.uk.

their interactions. Henrion and Römisch [13] proposed a two stage SEPEC model for studying competition in electricity markets. The authors investigated *M-stationary points* of the SEPEC model under the condition that the market demand in the spot market has a finite distribution. Recently, Ehrenmann and Neuhoff [7] compared two market designs, the integrated market design and the coordinated transmission auction for electricity trade and transmissions. From the mathematical perspective, the authors showed that the intergrated market design is an instance of an EPEC where generators know that his output decisions will influence the allocation of transmission rights by the SO, and the model can be represented by a Stackelberg model.

Apart from the application in electricity markets, there emerges a trend of using EPEC and SEPEC to characterize the two stage equilibrium problems in some general oligopoly markets where a set of strategic firms or agents (called leaders) non-cooperatively optimize their expected objective function anticipating the reaction of the remaining nonstrategic firms or investors (called followers). Recent work in Su [38] first investigated a deterministic EPEC, and showed the existence of a two stage Nash-Cournot equilibrium. Pang and Fukushima [25] proposed an iterative penality method for solving a generalized Nash equilibrium where each player solves a mathematical program with equilibrium constraints (MPEC), and introduced a class of remedial models for the multi-leader-follower games for the oligopolistic competition models in electric power markets. For a typical two stage competition in the stochastic environment, DeMiguel and Xu [5] developed a stochastic multiple-leader Stackelberg-Nash-Cournot (SMS) model for a homogeneous product (or service) supply market, and derived the existence and uniqueness of the SMS equilibrium. More broadly, for the transportation systems, EPEC models have been applied to analyze the competition behaviors of strategic users in the traffic network by Yang, Xiao and Huang [42]. In a recent work, Koh [17] investigated the potential for implicit collusions between users in the traffic network, and obtained an EPEC model where every players are constrained by a variational inequality. EPEC models have also been used for a game-theoretic analysis of the implications of overlay networks traffic of internet service providers (ISP) in Wang, Chiu and Lui [36].

Along with the increasing interests of modelling issues on EPEC, there synchronically emerges a volume of literature on employing the mathematical programming and game-theoretic analysis to investigate the behaviors of players and the properties of equilibrium in EPEC models. One of the natural questions arising from the EPEC problems is on the existence of Nash equilibrium of these two stage problems, where related results have been well established for some cases with particularly structured objective functions. See [5, 25, 38, 44] for a set of two stage equilibrium problems. However, it is well known that, for some general cases, EPECs may not have any global Nash equilibrium. Instead, several alternatives of *global Nash equilibrium* are introduced for describing players' strategic behaviors within a local feasible set. Hu and Ralph [15] introduced a set of new concepts as *local Nash equilibria* and *Nash stationary points* for EPEC for a bilevel noncooperative game-theoretic model of electricity markets with locational marginal prices. On the other hand, within the framework of Mordukhovich coderivatives, Mordukhovich [21, 22] first investigated the necessary optimality conditions of *M-stationary points* for EPECs. Moreover, Outrara [23] addressed a set of necessary conditions on the stationary points and the local equilibria of EPEC models in the term of coderivatives.

In this paper, we are concerned with the numerical methods for solving the two stage SEPEC models. We analyze the convergence of statistical estimators of Nash equilibria and Nash stationary points obtained from solving the sample average approximation problems. In the case when we obtain a stationary point by solving the sample average approximation, we investigate the convergence to Nash-C-stationary points of the true problem rather than weak ones investigated in [41].

The rest of this article is organized as follows. In the next section, we present a description and some properties of our model followed with the uniqueness of the second stage equilibrium. In Section 3, we present some source problems which motivate the EPEC model. In Section 4, we formulate the sample average approximation of our problem and show the convergence properties of the SAA estimators of Nash equilibria and Nash-C-stationary points, and draw some concluding remarks in Section 5.

Throughout this paper, we will use the following notation. All vectors are thought as column vectors and ^{T} denotes the transpose operation. For $x, y \in \mathbb{R}^s$, $x^T y$ denotes the scalar products of two vectors x and y, $\|x\|$ denotes the Euclidean norm. When $D \subset \mathbb{R}^s$ is a nonempty compact set of vectors, we use the notation $\|\cdot\|$ to denote $\|D\| := \max_{x\in D} \|x\|$. Moreover, $d(x, D) := \inf_{x'\in D} \|x - x'\|$ denotes the distance from point x to set D. For two compact sets D_1 and D_2,

$$\mathbb{D}(D_1, D_2) := \sup_{x\in D_1} d(x, D_2)$$

denotes the deviation from set D_1 to set D_2 (in some references [12] it is also called *excess* of D_1 over D_2), and $\mathbb{H}(D_1, D_2)$ denotes the Hausdorff distance between the two sets, that is,

$$\mathbb{H}(D_1, D_2) := \max\left(\mathbb{D}(D_1, D_2), \mathbb{D}(D_2, D_1)\right).$$

Moreover, we use $D_1 + D_2$ to denote the Minkowski addition of D_1 and D_2, that is, $D_1 + D_2 = \{x + y : x \in D_1, y \in D_2\}$. We also use $B(x, \delta)$ to denote the closed ball with radius δ and center x, that is $B(x, \delta) := \{x' : \|x' - x\| \le \delta\}$. When δ is dropped, $B(x)$ represents a neighborhood of point x. We also use $\mathcal{B}$ to denote the unit ball in a finite dimensional space.

2 The model

Let X_i, $i = 1, 2, \ldots, M$, be a nonempty, closed and convex subset of $\mathbb{R}^{m_i}$ and $X_{-i} := X_1 \times \cdots \times X_{i-1} \times X_{i+1} \times \cdots \times X_M$ denote the Cartesian product of the sets except X_i. Let $X = X_i \times X_{-i}$. We consider the following *stochastic equilibrium program with equilibrium constraints (SEPEC)*: find $x := (x_1, x_2, \cdots, x_M) \in X$ and $y(\cdot)$ such that for $i = 1, 2, \ldots, M$, $(x_i, y(\cdot))$ solves the following problem:

$$\begin{array}{cl} \min\limits_{x_i\in X_i, y(\cdot)} & \mathbb{E}[f_i(x_i, x_{-i}, y(\omega), \xi(\omega))] \\ \text{s.t.} & 0 \in H(x, y(\omega), \xi(\omega)) + \mathcal{N}_Q(y(\omega)), \quad \text{a.e.} \quad \omega \in \Omega, \end{array} \tag{2.1}$$

where $x_{-i} \in X_{-i}$, $f_i : \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_M} \times \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R}$ is a Lipschitz continuous function but not necessarily continuously differentiable, $\xi : \Omega \to \Xi \subset \mathbb{R}^d$ is a random vector defined on probability space $(\Omega, \mathcal{F}, P)$ with support set Ξ, and $\mathbb{E}[\cdot]$ denotes the mathematical expectation with respect to the distribution of ξ. The equilibrium constraint in (2.1) is represented by a parametric variational inequality problem (VIP), where $y(\cdot)$ is the prime variable, and x and $\xi(\omega)$ are treated as parameters, $H : \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R}^k$ is a vector-valued continuous function, Q is a nonempty, convex and closed subset of $\mathbb{R}^k$, and $\mathcal{N}_Q(y)$ is the normal cone to Q at y, which is defined in as follows:

$$\mathcal{N}_Q(y) := \begin{cases} \left\{\eta \in \mathbb{R}^k : \eta^T (\bar{y} - y) \le 0, \forall \bar{y} \in Q\right\}, & \text{if } y \in Q; \\ \emptyset, & \text{if } y \notin Q. \end{cases} \tag{2.2}$$

Problem (2.1) is a *two stage* SEPEC: at the first stage, decision maker/player i, $i \in \{1, \cdots, M\}$, chooses an optimal value $x_i \in X_i$ that maximizes the expected value of f_i under Nash conjecture (for fixed $x_{-i} \in X_{-i}$). At the second stage for the given x and a realization of the random vector ξ, player i finds an optimal $y(x, \xi)$ that solves the following

$$\begin{array}{ll} \min\limits_{y \in Q} & f_i(x_i, x_{-i}, y, \xi) \\ \text{s.t.} & 0 \in H(x, y, \xi) + \mathcal{N}_Q(y). \end{array} \tag{2.3}$$

To ease the notation, in some parts of this paper we will write $\xi(\omega)$ as ξ and the context will make it clear when ξ should be interpreted as a deterministic vector. Let us use $v_i(x_i, x_{-i}, \xi)$ to denote the optimal value function of the second stage problem (2.3). Then (2.1) can be written in an *implicit* form under some moderate conditions as

$$\min_{x_i \in X_i} \mathbb{E}\left[v_i(x_i, x_{-i}, \xi(\omega))\right], \tag{2.4}$$

where "implicit" means that (2.4) does not include the details of the second stage problem.

2.1 A discussion of the SEPEC model

Let $Y(x, \xi)$ denote the set of solutions to the variational inequality problem

$$0 \in H(x, y, \xi) + \mathcal{N}_Q(y), \tag{2.5}$$

Then we can rewrite the second stage optimization problem (2.3) as

$$\min_{y \in Y(x, \xi)} f_i(x_i, x_{-i}, y, \xi) \tag{2.6}$$

At this point, it might be helpful to give some practical interpretation of problem (2.6): here an optimal equilibrium from $Y(x, \xi)$ is not up to the decision maker i to choose. The mathematical formulation only represents player i's *optimistic* attitude towards a possible equilibrium outcome at scenario ξ, in other words, player i anticipates a best equilibrium outcome which minimizes his objective function $f_i(x_i, x_{-i}, y, \xi)$ at scenario ξ. To properly define $v_i(x_i, x_{-i}, \xi)$ mathematically, we let $v_i(x_i, x_{-i}, \xi) = +\infty$ if the corresponding equilibrium at the second stage does not exist, i.e. the corresponding variational inequality constraint in (2.3) does not have any solution. Then, given that the rivals' decisions are fixed at x_{-i}, decision maker i's expected profit at the first stage can be formulated as

$$\hat{\vartheta}_i(x_i, x_{-i}) := \quad \mathbb{E}\left[v_i(x_i, x_{-i}, \xi(\omega))\right],$$

or equivalently

$$\hat{\vartheta}_i(x_i, x_{-i}) = \quad \mathbb{E}\left[\min_{y(\omega) \in Y(x, \xi(\omega))} f_i(x_i, x_{-i}, y(\omega), \xi(\omega))\right]. \tag{2.7}$$

If all decision makers are optimistic, then the SEPEC problem can be formulated as follows:

$$\min_{x_i \in X_i} \quad \mathbb{E}\left[\min_{y \in Y(x,\xi(\omega))} f_i(x_i, x_{-i}, y, \xi(\omega))\right]. \tag{2.8}$$

Under some moderate conditions, (2.8) coincides with (2.1), see Proposition 5 in [32, Chapter 1].

Let us now consider an opposite case when player i is pessimistic. In such a case, his second stage decision problem can be formulated as

$$\max_{y \in Y(x,\xi)} \quad f_i(x_i, x_{-i}, y, \xi), \tag{2.9}$$

which means that, in making his decision for minimizing $\mathbb{E}\left[f_i(x_i, x_{-i}, y(\omega), \xi)\right]$ at the first stage, decision maker i expects a worst a second stage equilibrium outcome $y(\omega) \in Y(x, \xi(\omega))$ which maximizes $f_i(x_i, x_{-i}, y(\omega), \xi)$. Denote by $\check{v}_i(x_i, x_{-i}, \xi)$ the optimal value function of decision problem (2.9). Then player i's expected objective function can be written as

$$\check{\vartheta}_i(x_i, x_{-i}) := \quad \mathbb{E}\left[\check{v}_i(x_i, x_{-i}, \xi(\omega))\right], \tag{2.10}$$

where $\check{v}_i(x_i, x_{-i}, \xi) = +\infty$ if $Y(x, \xi) = \emptyset$. If all decision makers are pessimistic and try to hedge against a worst possible equilibrium at the second stage, then the SEPEC model becomes: find $x = (x_1, \cdots, x_M)^T$ and $y(\cdot)$ such that for $i = 1, 2, \ldots, M$, $(x_i, y(\cdot))$ solves the following problem:

$$\min_{x_i \in X_i} \quad \mathbb{E}\left[\max_{y \in Y(x,\xi(\omega))} f_i(x_i, x_{-i}, y, \xi(\omega))\right] \tag{2.11}$$

where each decision maker solves a min-max problem.

There could be other cases when some generators are optimistic while others are pessimistic or some decision maker do not really have an extreme view about the future equilibrium. In this paper, we will simplify the discussion by considering the case when the equilibrium problem has a unique solution and subsequently the minimization process in (2.7) can be dropped.

2.2 Uniqueness of the second stage equilibrium

In the remainder of this section, to complete the statement in Section 2.1, we present some assumptions to guarantee we discuss sufficient conditions for the existence and uniqueness of the Nash equilibrium problem (represented by VIP (2.5)) at the second stage and the Lipschitz continuity of the optimal value functions $v_i(x, \xi)$ at the second stage. Based on these assumptions, we can proceed to investigate the first order equilibrium conditions of the SEPEC problem (2.1) and its sample average approximation (4.30) in terms of Clarke generalized gradient in Section 4.

First, let us look at the solution set $Y(x, \xi)$ of VIP (2.5):

$$0 \in H(x, y(\omega), \xi(\omega)) + \mathcal{N}_Q(y(\omega)), \tag{2.12}$$

and decision maker i's objective function $f_i(x_i, x_{-i}, y(x, \xi), \xi)$ at every scenario $\xi \in \Xi$ for $i = 1, 2, \ldots, M$.

Now, what we need to look at is the uniqueness of the solution to variational inequality problem (2.5), which can be guaranteed by the *strict monotonicity* of mapping $H(x, y, \xi)$ for any $x \in X$ and a.e. $\xi \in \Xi$, which is equivalent to the strict concavity of function $R_i(y_i, y_{-i}, x, \xi)$ with respect to y_i for any fixed y_{-i} and $i = 1, 2, \ldots, M$. Here, we present the assumptions on the strict monotonicity of mapping $H(x, y, \xi)$ as follows.

Assumption 2.1 *For $i = 1, 2, \ldots, M$ and almost every $\xi \in \Xi$,*

(a) *$H(x, y, \xi)$ is a Lipschitz continuous function of (x, ξ) on $X \times \Xi$ with a Lipschitz constant independent of y.*

(b) *$H(x, \cdot, \xi)$ is* uniformly strongly monotone *on set Q, that is, for any given x and $\xi \in \Xi$, there exists a constant $c > 0$ such that*

$$(H(x, y', \xi) - H(x, y, \xi))^T (y' - y) \geq c\|y' - y\|^2, \quad \forall y', y \in K. \tag{2.13}$$

(c) *$f_i(\cdot, x_{-i}, \cdot, \xi)$ is Lipschitz continuous on $X_i \times Q$ with modulus $\kappa_i(\xi)$, that is, for all $i = 1, 2, \ldots, M$,*

$$|f_i(x_i', x_{-i}, y', \xi) - f_i(x_i, x_{-i}, y, \xi)| \leq \kappa_i(\xi) \left(\|x_i' - x_i\| + \|y' - y\| \right),$$

where $\mathbb{E}[\kappa_i(\xi)] < \infty$.

Under Assumption 2.1 (b), it follows by virtue of [10, Theorem 2.3.3]) that the VIP (2.5) has a unique solution for every given x and ξ. Moreover, under Assumption 2.1 (a) and (b) we can show the Lipschitzness of $v_i(\cdot, x_{-i}, \xi)$ which will be used in the asymptotic analysis of sample average approximate Nash equilibrium in Proposition 4.1.

Lemma 2.1 *Under Assumption 2.1, for every fixed $x_{-i} \in X_{-i}$ and a.e. $\xi \in \Xi$, $v_i(\cdot, x_{-i}, \xi)$ is Lipschitz continuous on X_i with modulus $\kappa_i'(\xi)$, where $\mathbb{E}[\kappa_i'(\xi)] < \infty$.*

Proof. Under Assumption 2.1 (b), it follows from [10, Theorem 2.3.3 (c)] that the variational inequality

$$0 \in H(x, y, \xi) + \mathcal{N}_Q(y)$$

has a unique solution $y(x, \xi)$ which is Liptschitz continuous on X with a constant modulus. Moreover, under Assumption 2.1 (c), $F_i(x_i, x_{-i}, \cdot, \xi)$ is Lipschitz continuous on Q with modulus $\kappa_i(\xi)$, where $\mathbb{E}[\kappa_i(\xi)] < \infty$. Consequently,for $i = 1, 2, \ldots, M$, there exists a $\kappa_i'(\xi)$ such that $v_i(\cdot, x_{-i}, \xi)$ is Lipschitz continuous on X_i with modulus $\kappa_i'(\xi)$, where $\mathbb{E}[\kappa_i'(\xi)] < \infty$. ∎

3 Source problems

A number of applications of two stage stochastic equilibrium problem with equilibrium constraints arise from a diversity of sources. In this section, we list a few examples.

3.1 Stochastic bilevel games

It is well-known that bilevel programming is closely related to MPEC through KKT conditions at the second stage. It is therefore no surprise that stochastic bilevel games provide rich problem sources for the SEPEC. Consider a two stage stochastic Nash game which consists two sets of players: a set of M players who compete at the first stage and a set of K players who compete at the second stage the decisions of players at the first stage at the first stage are disclosed and exterior uncertainty (such as market demand) is realized. Mathematically, we can formulate this kind of game as follows:

$$\begin{cases} \min\limits_{x_i \in X_i, y(\cdot)} \mathbb{E}[f_i(x_i, x_{-i}, y(x, \xi(\omega)), \xi(\omega))] \\ y_j(x, \xi(\omega)) \text{ solves } \min\limits_{y_j \in Q_j} R_j(y_j, y_{-j}, x, \xi(\omega)), \text{ for } j = 1, \ldots, K, \text{ a.e. } \omega \in \Omega. \end{cases} \tag{3.14}$$

The stochastic multiple leader-followers game investigated by DeMiguel and Xu [5] is a typical example of this type two stage stochastic bilevel game.

Assuming that for $j = 1, 2, \ldots, K$, function $R_j : \mathbb{R}^{k_j} \times \mathbb{R}^{k-k_j} \times \mathbb{R}^m \times \Xi \to \mathbb{R}$ is continuously differentiable w.r.t. y_j on a nonempty convex and closed subset $Q_j \subset \mathbb{R}^{k_j}$, $Q_{-j} := Q_1 \times \cdots \times Q_{j-1} \times Q_{j+1} \times \cdots \times Q_K$, and $k = \sum_{j=1}^{K} k_j$, the we can characterize the optimality condition of each player at the second stage through a generalized equation and combining them gives

$$0 \in \nabla_y R(x, y, \xi) + \mathcal{N}_Q(y), \tag{3.15}$$

where $\nabla_y R(x, y, \xi) := (\nabla_{y_1} R_1(y_1, y_{-1}, x, \xi), \cdots, \nabla_{y_K} R_K(y_K, y_{-K}, x, \xi))^T$. Under some convexity conditions of the objective functions, it is well-known that a solution $y^*(x, \xi)$ is a Nash equilibrium at the second stage if and only if it is a solution to the generalized equation (3.15). Consequently we can reformulate the stochastic bilevel game (3.14) as the following two stage SEPEC:

$$\begin{array}{ll} \min\limits_{x_i \in X_i} & \mathbb{E}[f_i(x_i, x_{-i}, y(x, \xi(\omega)), \xi(\omega))] \\ \text{s.t.} & 0 \in \nabla_y R(x, y(x, \xi(\omega)), \xi(\omega)) + \mathcal{N}_Q(y(x, \xi(\omega))), \quad \text{a.e.} \quad \omega \in \Omega. \end{array}$$

Note that the two stage stochastic bilevel game model (3.14) may cover the capacity expansion model considered by Gürkan and Pang [11] where game at the first stage is viewed as a competition on long term capacity investment at present and second level game as short term competition in future once the capacity expansion is completed and exterior uncertainty is realized. In this case players are two stages may be identical. Note that (3.14) can also be used to model competition in forward-spot electricity markets where generators compete at the first stage in the forward market for long term contracts and then compete for dispatch at spot markets on daily basis, see [44].

3.2 Capital Tax Competition

The enforcement of an effective taxation on savings income has been a long-standing issue both in policy and in academic debates, see [8, 9]. Along with the globalization of the capital market, the tax can be easily evaded if the residence country is unable to monitor the investors' foreign interest incomes, where the countries are linked through perfect capital mobility. These capital links between countries may result in a very complex investment network. Even for a two country economy the flows of real and financial capital might induce a complex system of transactions under the different tax structures of both countries. In this subsection, we consider a capital tax competition between the national tax authorities in two countries, denoted by i and j, respectively.

The analysis of this capital tax competition employs a two stage stochastic equilibrium model. At the first stage, we assume that each tax authority of country, i or j, has three different tax instruments. We give the tax instruments set by tax authority of country i (tax authority i) for example:

(a) The first tax instrument is a *wage tax rate* t_i^w at which it taxes wage income $w_i l_i$. Note that, in most of practical capital markets, wage rate w_i and labor supply l_i may be affected by the different amount of total investment level in country i, and hence we can rewrite them as $w_i(s^i)$ and $l_i(s^i)$, where s^i denotes the amount of capital invested in country i. Here, we set s^i being the sum of s_i^i, the amount of capital invested in country i by residents at home (i.e. country i), and s_j^i invested by the investors from abroad (i.e. country j). Moreover, the wage rate in a country is usually set for a long term and hence independent of the global economic scenario ξ. On the other hand, the labor supply level fluctuates along with the change of economic environment and hence we assume that it is a function of scenario ξ, denoted by $l_i(s^i, \xi)$.

(b) The second tax instrument the authority might choose is on the capital income of residents where we denote its rate by t_i^r. By assuming the perfect information sharing between the tax authorities of the two countries, then the tax base can be formulated as

$$R(\xi)s_i = R(\xi)(s_i^i + s_i^j), \tag{3.16}$$

where $R(\xi)$ is the global return rate to the investment, that is, the return for every unit capital invested in either countries, and is varied by the random shock ξ in the capital market. Since in this subsection, we focus our investigation on the taxation problem, we generally assume that the global return rate for the investment in each country are the same. In (3.16), s_i is the amount of capital invested by residents at country i and is the sum of s_i^i and s_i^j where s_i^j denotes the amount of capital invested by the residents at country i into the market in country j.

(c) Third, a government may tax the capital income generated at home on a source basis, t_i^s, where the tax base can be calculated as

$$R(\xi)s^i = R(\xi)(s_i^i + s_j^i), \tag{3.17}$$

which includes the return of the investment by the residents at country i and from abroad, i.e. country j. Similarly as in country i, we denote the tax instruments set by the tax authority j as t_j^w, t_j^r and t_j^s, respectively.

Differing from the discussion in [8, 9], in the model, we first consider the strategic behaviors of the representative investor in each country. The representative investor (or consumer) in country i maximizes a well-behaved utility function $\vartheta_i(c_{i1}, c_{i2}, l_i, \xi)$, where c_{i1} and c_{i2} are the consumption levels before and after the investment period. Denote the endowment obtained by the investor by e_i. After the realization of uncertainty ξ in the capital market, the investor in country i needs to decide the proportion of e_i to be consumed, c_{i1}, or saved, $s_i = s_i^i + s_i^j$, where $s_i = e_i - c_{i1}$. Because the decision on c_{i1} is made after knowing ξ, consumption level c_{i1} is affected by the uncertainty in the capital market and hence can be taken as a random function of ξ, which is implied by the fact that the consumption level of investors i at country i fluctuates as a response to different economic situations in the capital market. Moreover, c_{i2} denotes the consumption level after the return of the investment and can be formulated as

$$\begin{aligned} c_{i2}(s_i^i, s_i^j, s_j^i, \xi; t_i^r, t_i^s, t_j^s) = \quad & w_i\,(s_i)\, l_i \left(s_i^i + s_j^i, \xi \right) + [1 + R(\xi)\,(1 - t_i^r - t_i^s)]\, s_i^i \\ & + \left[1 + R(\xi) \left(1 - t_i^r - t_j^s \right) \right] s_i^j, \end{aligned} \tag{3.18}$$

where the three terms in the righthand of (3.18) are the wage income, the post-tax income from the home investment, and the post-tax income from the abroad investment, respectively. Note that, in (3.18), $t_i^r + t_i^s$ and $t_i^r + t_j^s$ are the effective tax paid by the presentative investor at country i on its capital income from country i and country j, respectively. By incorporating c_{i1} and c_{i2} into the investor's utility function, for a realized market scenario ξ, the decision problem of the representative investor in country i is to choose the amounts of $s_i^i(\xi)$ and $s_i^j(\xi)$ to maximize their utility function. Assuming that the realization of the uncertainty in the capital market is ξ and the investor in country j rationally fix their optimal investment at $\left(s_j^j(\xi), s_j^i(\xi) \right)$, the representative investor in country i determines its investment policy by solving the following problem,

$$\max_{s_i^i \in S_i^i, s_i^j \in S_i^j} \quad \vartheta_i \left(e_i - s_i^i - s_i^j, c_{i2}(s_i^i, s_i^j, s_j^i, \xi; t_i^r, t_i^s, t_j^s), l_i \left(s_i^i + s_j^i, \xi \right), \xi \right), \tag{3.19}$$

where $S_i^i := [0, \bar{s}_i^i]$ and $S_i^j := [0, \bar{s}_i^j]$, and $\bar{s}_i^i$ and $\bar{s}_i^j$ are the upper bounds of investments s_i^i and s_i^j. By assuming the convexities of functions $l_i(\cdot, \xi)$ and $\vartheta_i(c_{i1}, c_{i2}, l_i, \xi)$, we can show the existence and uniqueness of the investment equilibrium of the investors' competition in the capital market at almost every scenario ξ.

Consequently, the second stage equilibrium problem is: for tax instruments t_i and t_j fixed at the first stage and the realization ξ, find an equilibrium (s_i, s_j) solves the following parametric equilibrium problem,

$$\max_{s_k \in S_k} \quad \vartheta_k \left(e_k - s_k, c_{k2}(s_k^k, s_k^{k'}, s_{k'}^k, \xi; t_k^r, t_k^s, t_{k'}^s), l_k \left(s^k, \xi \right), \xi \right), \tag{3.20}$$

where $s_k \in S_k = S_k^k \times S_k^{k'}$, $k = i, j$ and $k' \neq k \in \{i, j\}$. In (3.20), $t := (t_i, t_j)$ and ξ are treated as parameters. Therefore, the competition between the representative investors at country i and j can be taken as a Cournot-type game, and the equilibrium to problem (3.20) is a function of t and ξ which can be specified as $s_k(t, \xi)$ for $k = i$ and j. Then, we can rewrite (3.19) as the following general equation form,

$$0 \in H(s(t, \xi), t, \xi) + \mathcal{N}_S(s(t, \xi)), \tag{3.21}$$

where $s(t, \xi) = (s_i(t, \xi), s_j(t, \xi))$, and feasible set $S := S_i \times S_j$.

Assuming the perfect information sharing between the two countries, we have that each national authority determines its tax rates by aiming at the maximization of its expected utility function which consists of two parts: one is the expected return of its representative investor in the capital market, and the other is the production capacity of country i, denoted by g_i, which is seen as a function of the total tax revenue in the country. Consequently, at the first stage, given that tax authority j's optimal tax policy is rationally fixed at t_j^*, the decision problem of tax authority i can be formulated as:

$$\begin{aligned} \max_{t_i} \quad & \mathbb{E}\left[\eta_i(\, t_i, t_j^*, s_i(t_i, t_j^*, \xi), s_j(t_i, t_j^*, \xi), \xi \,) \right], \\ \text{s.t.} \quad & t_i^w \in [0, \bar{t}_i^w],\ t_i^r \in [0, \bar{t}_i^r],\ t_i^s \in [0, \bar{t}_i^s]. \end{aligned} \tag{3.22}$$

where the objective function $\eta_i\left(t_i, t_j^*, s_i, s_j, \xi\right)$ of tax authority i is defined as

$$
\begin{aligned}
\eta_i\left(t_i, t_j, s_i, s_j, \xi\right) \;&:= \vartheta_i\left(e_i - s_i, c_{i2}(s_i^i, s_i^j, s_j^i, \xi; t_i^r, t_i^s, t_j^s), l_i\left(s^i, \xi\right), \xi\right) \\
&+ u\left(g_i\left(t_i, t_j, s_i, s_j, \xi\right)\right),
\end{aligned}
\tag{3.23}
$$

$u(\cdot)$ is the utility function of the production capacity g_i, and the production capacity g_i is a function of the tax revenues as follows

$$
g_i\left(t_i, t_j, s_i, s_j, \xi\right) = t_i^w w_i(\xi) l_i\left(s_i^i + s_j^i, \xi\right) + \left(t_i^s + t_i^r\right) R(\xi) s_i^i + t_i^r R(\xi) s_i^j + t_i^s R(\xi) s_j^i.
$$

Moreover, in decision problem (3.24), the variables $s_i(t_i, t_j, \xi)$ and $s_j(t_i, t_j, \xi)$ are solved from general equation (3.21) for any fixed t_i, t_j and the realization ξ. By inclusively taking the investors' reactions at the second stage into consideration, tax authority i's decision problem can be written as,

$$
\begin{aligned}
\max_{t_i} \quad & \mathbb{E}\left[\eta_i(\, t_i, t_j^*, s_i(t_i, t_j, \xi), s_j(t_i, t_j, \xi), \xi\,)\right], \\
\text{s.t.} \quad & t_i^w \in [0, \bar{t}_i^w],\ t_i^r \in [0, \bar{t}_i^r],\ t_i^s \in [0, \bar{t}_i^s], \\
& 0 \in H(s(t, \xi), t, \xi) + \mathcal{N}_S(s(t, \xi)), \quad \text{a.e. } \xi \in \Xi.
\end{aligned}
\tag{3.24}
$$

which implies that the capital tax competition can be formulated as an SEPEC model.

3.3 Oligopolistic transit market

In this subsection, we look at a two stage stochastic equilibrium problem for a urban transit systems. Over past twenty years, the deregulation of *urban transit systems* has become an appealing alternative to centralized municipal transit policy. In a recent paper [45], a deterministic network equilibrium model with a two stage framework for a deregulated transit system is proposed to describe the fare competition between transit operators where every operator takes into account passengers' responses in making its decision. At the first stage, by assuming that its rivals rationally choose their optimal decisions on their fare structures, each of the transit operators can determine its own fare structure in order to maximize its expected revenue, where in the urban transit system, the transit operators' revenues depends on the number of passengers using their lines. Then, at the second stage, every passenger reacts to the transit operators' fares in the urban transit system.

Stochastic network equilibrium models are widely applied for predicting traffic patterns in the transportation networks at the second stage, in which the traffic flow in the transportation network is characterized by *stochastic user equilibrium* for every possible scenario. In this two stage model, the interaction between the transit operators and the passengers is described in the form of Stackelberg game, that is, at the first stage, in making the decision, every operator takes the passengers' reaction to its fare plan at every traffic scenario into account. On the other hand, the competition between the transit operators can be seen as a Cournot game where each operator makes its decision regarding that its rivals' fares are fixed.

In the problem, the urban transit network is denoted by a directed traffic network as $G = (N, A)$ where N is the set of nodes (or transfer stations) and A is the set of links (or route sections). At the first stage, the transit competition is portrayed as M player (transit operator) noncooperative game of deciding the fares for a set of transfer lines connecting origin-destination

(OD) pairs w for $w \in W$ where W is the set of all OD pairs, and R_w is the set of all routes joining OD pair w. In the first stage equilibrium, transit operator i makes its decision on the fare of the route connecting OD $w \in W$, denoted by $p_w^i = \{p_r^i\}_{r \in R_w}$, so as to maximize its expected profit, where the expectation is taken with respect to the distribution of the traffic uncertainty $\xi \in \Xi$. Note that, in making the decision at the first stage, each operator predictively takes into account the passenger flow in every route of the traffic network at every possible scenario ξ, where the flows are determined by solving a stochastic user equilibrium model at the second stage.

At the second stage problem, with each fixed transit fare structure and a realized traffic uncertainty ξ, the stochastic user equilibrium condition can be mathematically expressed to determine the flow on every route r serving an OD pair $w \in W$ for $r \in R_w$: Denote by $y_r(\xi)$ the traffic flow on path r at traffic scenario ξ. Given that the realization of traffic uncertainty is ξ and operator i's fare for route r is p_r^i for all $r \in R_w$ and $w \in W$, the user (passenger)'s travel cost function on path r can be written as

$$u_r(y, p_w, \xi) = d_r(p_w) + C_r(y, p_w, \xi) + \theta_0 \left(t_r(y, \xi) - \tau_w\right)^2, \tag{3.25}$$

where $y := \{y_r\}_{r \in \mathcal{R}_w}$, $p_w := (p_w^1, \cdots, p_w^M)^T$, $p_w^i = \{p_r^i\}_{r \in R_w}$. In (3.25), $d_r(p_w)$ represents the composite of attributes such as the travel fare for a certain distance which is independent of time/flow, $C_r(y, p_w, \xi)$ denotes the stochastic travel cost on path r which is implicitly determined by the flows on all arc r which may depends on the volume of the passenger flow between OD pair w, the fares of all operators, and changes for different traffic scenario. $t_r(y, \xi)$ denotes the stochastic travel time on path r and τ_w denote the expected travel time between OD pair w. Furthermore, in (3.25), θ_0 is the penalty coefficient when the actual travel time on w deviates the scheduled time τ_w. Then, by assuming the total population traveling between OD pair w is q_w being a deterministic parameter, the feasible set of the traffic flow across the whole network can be expressed as

$$Q = \left\{ y : \sum_{r \in \mathcal{R}_w} y_r = q_w, \, \forall w, y_r \geq 0, \forall r \right\}$$

which is a convex set. Moreover, at the second stage, the stochastic user equilibrium is defined as the state when no passenger believes that he can reduce his perceived travel cost by changing route unilaterally. Hence, we can write this equilibrium condition as following variational inequality problem: find y^* such that

$$\begin{cases} y^* \left(u(y, p, \xi) - u(y^*, p, \xi)\right) = 0, \, \forall y \in Q, \; w \in W \\ y^* \geq 0, \; u(y, p, \xi) - u(y^*, p, \xi) \geq 0, \end{cases} \tag{3.26}$$

for almost every $\xi \in \Xi$ and fixed fare $p = \left\{p_w^i, \; i = 1, 2, \ldots, M, \; w \in W\right\}$, where

$$y^* \left(u(y, p, \xi) - u(y^*, p, \xi)\right) := \sum_{r \in R_w} y_r^* \left(u_r(y, p, \xi) - u_r(y^*, p, \xi)\right),$$

and the fare p and the traffic scenario ξ are treated as parameters. VIP (3.26) can be equivalently rewritten in the following parametric general equation:

$$0 \in H(p, y, \xi(\omega)) + \mathcal{N}_Q(y), \quad \text{a.e. } \omega \in \Omega, \tag{3.27}$$

for a vector-valued function $H(p, y(p, \xi), \xi)$ and feasible set Q of y. Usually, functions $C_r(y, p_j, \xi)$ and $t_r(y, \xi)$ are assumed to be continuously differentiable for every ξ, and hence the cost function $u_r(y, p_j, \xi)$ is a continuously differentiable function of y and $H(p, y, \xi)$ is also a continuous single-valued mapping. Hence, solution y to parametric problem (3.27) can be written as $y(p, \xi)$ which is a function of fare p and traffic scenario ξ.

Let us step back to the operators' problems at the first stage, in which every operator makes its decision on the fare charged at each route to maximize its expected revenue. From the discussion on stochastic user equilibrium, we have that passenger flow $y_r(p, \xi)$ on path r can be implicitly solved by (3.26) or (3.27), and is a continuous function of p for every scenario ξ. Then, we assume that the proportion that passengers chooses operator i's service on path $r \in R_w$ for traveling between OD pair w is a function of $f_w(y, p_w^i, p_w^{-i})$ where p_w^{-i} are the fare structures provided by operator i's rivals between OD pair w. In [45], these proportions are calculated according to the well used logit-type model, and in a more general case, we might assume that $f_w\left(y(p, \xi), p_w^i, p_w^{-i}, \xi\right)$ is a continuously differentiable function of y, p_w^i and p_w^{-i} for all OD pair w, where this proportion only reflects a passenger's attitude towards each operator's fare. Consequently, with this proportion, we can estimate the number of passengers using operator i's service as $v_w^i = f_w\left(y(p, \xi), p_w^i, p_w^{-i}, \xi\right) q_w$. Then operator i's expected revenue function can be written as

$$R_i(p^i, p^{-i}) = \mathbb{E}\left[\sum_{w \in W} f_w\left(y(p_w^i, p_w^{-i}, \xi), p_w^i, p_w^{-i}, \xi\right) q_w p_w^i\right], \tag{3.28}$$

where $p^i = \left\{p_w^i\right\}$ and $p^{-i} = \left\{p_w^{-i}\right\}$ for all $w \in W$, and $\mathbb{E}[\cdot]$ is taken with respect to the distribution of the traffic uncertainty ξ. It should be noted that the operating cost does not appear in the expression due to the assumption of fixed service frequency. Thus, the two stage equilibrium problem is: find $p = \left(p^1, \cdots, p^M\right)$ and $y(\cdot)$ such that for $i = 1, 2, \ldots, M$, $\left(p^i, y(\cdot)\right)$ solves the following problem:

$$\begin{array}{cl} \min\limits_{p_w^i \in [0, \bar{p}_w^i], y(\cdot)} & \mathbb{E}\left[\sum\limits_{w \in W} f_w\left(y(p_w^i, p_w^{-i}, \xi), p_w^i, p_w^{-i}, \xi\right) q_w p_w^i\right] \\ \text{s.t.} & 0 \in H(p, y(p, \xi(\omega)), \xi(\omega)) + \mathcal{N}_Q(y(p, \xi(\omega))), \quad \text{a.e.} \quad \omega \in \Omega, \end{array} \tag{3.29}$$

where $\bar{p}_w^i$ is the ceiling of operator i's fare p_w^i on its route between OD w. From problem (3.29), we have that the competition in the oligopolistic transit market can be modeled by an SEPEC problem. In [45], a deterministic version of EPEC model is proposed for investigating the competition in a deregulated transit network market.

Apart from the applications of the two stage SEPEC models discussed in this section, there are some potential applications in transportation and economics [15, 21, 13, 28], internet service problems [36] and airline revenue management problems [16].

4 Sample average approximation

In this section, we discuss a numerical method for solving the SEPEC problem (2.1). If the random vector ξ has a finite discrete distribution and the distribution is known, then the problem can be easily formulated as a deterministic EPEC for which existing numerical methods may be readily applied to solve it [18]. To cover a broader spectrum of practical applications, here we assume that ξ satisfies a general distribution which could be continuous, and it is impossible to obtain a closed form of $\mathbb{E}[f_i(x_i, x_{-i}, y, \xi)]$ either because it is computationally too expensive or the distribution function is unknown. However, it might be possible to obtain samples of ξ from past data or computer simulation, and a particular numerical scheme we are looking at here is the sample average approximation (SAA). Let $\xi^1, \cdots, \xi^N$ be an independent and identically distributed (iid) sampling of the random vector $\xi(\omega)$. We consider the following sample average approximation of problem: find $x^N := (x_1^N, x_2^N, \cdots, x_M^N)^T \in X_1 \times X_2 \times \cdots \times X_M$ such that

$$\min_{x_i \in X_i} \quad \frac{1}{N} \sum_{n=1}^{N} f_i(x_i, x_{-i}^N, y^n, \xi^n) \tag{4.30}$$

where y^n, $n = 1, 2, \ldots, N$, is a solution of the following variational inequality problem:

$$0 \in H\left(\left(x, x_{-i}^N\right), y^n, \xi^n\right) + \mathcal{N}_Q(y^n) \tag{4.31}$$

$$H(x, y, \xi) := (H_1(x, y, \xi), \cdots, H_K(x, y, \xi))^T$$

and

$$\mathcal{N}_Q(y) := \mathcal{N}_{Q_1}(y_1) \times \cdots \times \mathcal{N}_{Q_K}(y_K).$$

We refer to (2.1) as the *true* (SEPEC) problem and (4.30) as the *sample average approximation* problem. Since (4.31) has a unique solution, we may use the notation of the optimal value function in (2.4) to reformulate (4.30)-(4.31) as follows: find $\left(x_1^N, \cdots, x_M^N\right)^T$ such that x_i solves

$$\min_{x_i \in X_i} \quad \hat{\vartheta}_i^N(x_i, x_{-i}) := \frac{1}{N} \sum_{n=1}^{N} v_i(x_i, x_{-i}^N, \xi^n). \tag{4.32}$$

We call (4.32) the sample average approximation of the implicit Nash SEPEC (2.4).

Sample average approximation is a very popular method in stochastic programming, it is known under various names such as Monte Carlo sampling, sample path optimization and stochastic counterpart, see [27, 30, 34] for SAA in general stochastic programming and [5, 19, 41] for recent applications of the method to stochastic equilibrium problems. Our focus in this section is on the convergence of SAA problems described above to their true counterparts. Specifically, if we obtain a Nash equilibrium or a Nash stationary point (to be defined shortly), denoted by x^N, from solving (4.30), we investigate the convergence of x^N as sample size N increases.

Proposition 4.1 (Convergence of Nash equilibrium estimators) *Let $\{x^N\}$ be a sequence of Nash equilibria obtained from solving (4.30) and Assumption 2.1 holds. Then with probability one an accumulation point of $\{x^N\}$ is a Nash equilibrium of the true problem (2.1).*

The results depend on the Lipschitz continuity of v_i (established in Lemma 2.1) rather than the details of the second stage equilibrium. Therefore the proposition follows straightforwardly from [41, Theorem 4.2 (b)]. We omit the details.

4.1 Nash stationary points

It is well known that the optimal value function of a parametric mathematical program with equilibrium constraints (MPEC) is often nonconvex. In our context, this means that $v_i(x_i, x_{-i}^N, \xi^n)$

may be nonconvex in x_i for fixed x_{-i}^N and ξ^n, and consequently we may obtain a local Nash equilibrium or a Nash stationary point from solving the SAA problem (4.30). The concept of stationary points are important in optimization as it provides some information of optimality. This is particularly so in MPECs where obtaining a global optimal solution is often difficult and consequently various of stationary points are investigated [13, 44]. The concept of Nash stationary point is relatively new: it was introduced by Hu and Ralph [15].

We start with the definition. Based on Assumption 2.1, we have that the optimal value function v_i is usually not continuously differentiable, and the concept of the generalized gradient is needed to characterize the first order optimality conditions. Here we use the *Clarke generalized gradient* for the analysis which is popular and mathematically easy to handle. The Clarke generalized gradient of the optimal value function $v_i(x, y, \xi)$ w.r.t. x coincides with the usual gradient at the points where $v_i(\cdot, y, \xi)$ is strictly differentiable.

Let $v : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz continuous function. Recall that *Clarke generalized derivative* of v at point x in direction d is defined as

$$v^o(x, d) := \limsup_{y \to x, t \downarrow 0} \frac{v(y + td) - v(y)}{t}.$$

v is said to be Clarke *regular* at x if the usual one sided directional derivative, denoted by $v'(x, d)$, exists for all $d \in \mathbb{R}^n$ and $v^o(x, d) = v'(x, d)$. The *Clarke generalized gradient* (also known as Clarke subdifferential) is defined as

$$\partial v(x) := \{\zeta : \zeta^T d \leq v^o(x, d)\}.$$

See [4, Chapter 2].

In Lemma 2.1, we have shown that under some appropriate conditions $v_i(x_i, x_{-i}, \xi)$, $i = 1, \cdots, M$, is Lipschitz continuous w.r.t. x_i with integrable Lipschitz modulus. This implies that $\mathbb{E}[v_i(x_i, x_{-i}, \xi)]$ is also Lipschitz continuous w.r.t. x_i and hence $\partial_{x_i} \mathbb{E}[v_i(x_i, x_{-i}, \xi)]$ is well defined, see [32]. We characterize the first order equilibrium condition of (2.1) at a Nash equilibrium in terms of the Clarke generalized gradients as follows:

$$0 \in \partial_{x_i} \mathbb{E}[v_i(x_i, x_{-i}, \xi)] + \mathcal{N}_{X_i}(x_i), \ i = 1, \cdots, M. \tag{4.33}$$

Here and later on, the addition of the sets is in the sense of Minkowski. We call a point x^* satisfying (4.33) a *stochastic Nash-C-stationary point*. Under some standard constraint qualifications, a Nash equilibrium is a Nash-C-stationary point. Conversely if v_i is convex, then a Nash-C-stationary point is also a Nash equilibrium.

Let us now consider the first order necessary equilibrium condition for the SAA problem (4.30) in terms of Clarke generalized gradient:

$$0 \in \partial_{x_i} \vartheta_i^N(x_i, x_{-i}) := \partial_{x_i} \left(\frac{1}{N} \sum_{n=1}^{N} v_i(x_i, x_{-i}, \xi^n) \right) + \mathcal{N}_{X_i}(x_i), \ i = 1, \cdots, M. \tag{4.34}$$

We call a point $\bar{x}^N$ satisfying (4.34) *SAA Nash-C-stationary point*. Our objective here is to investigate the convergence SAA Nash-C-stationary point to its true counterpart.

For the simplicity of notation, we denote throughout this section the following.

$$\mathcal{A}\vartheta(x) := \partial_{x_1} \mathbb{E}[v_1(x, \xi)] \times \cdots \times \partial_{x_M} \mathbb{E}[v_M(x, \xi)] \tag{4.35}$$

and

$$G_X(x) := \mathcal{N}_{X_1}(x_1) \times \cdots \times \mathcal{N}_{X_M}(x_M). \tag{4.36}$$

The first order equilibrium condition (4.33) can be written as

$$0 \in \mathcal{A}\vartheta(x) + G_X(x). \tag{4.37}$$

Likewise, the first order equilibrium condition (4.34) can be written as

$$0 \in \mathcal{A}\vartheta^N(x) + G_X(x), \tag{4.38}$$

where

$$\mathcal{A}\vartheta^N(x) := \partial_{x_1}\left(\frac{1}{N}\sum_{n=1}^N v_1(x,\xi^n)\right) \times \cdots \times \partial_{x_M}\left(\frac{1}{N}\sum_{n=1}^N v_M(x,\xi^n)\right). \tag{4.39}$$

Let $\{\bar{x}^N\}$ be a sequence of stationary points satisfying optimality condition of the SAA problem (4.34) with sample size N. In what follows, we investigate the convergence of the sequence as sample size N increases. We need the following technical results.

Lemma 4.1 *Let $F(x,\xi) : \mathbb{R}^m \times \Xi \to \mathbb{R}$ be a continuous function, and $\mathcal{X}$ be a compact subset. Assume that $F(x,\xi)$ is locally Lipschitz continuous with respect to x for almost every ξ with modulus $L(x,\xi)$ which is bounded by a positive constant C. Then*

$$\lim_{N\to\infty}\sup_{x\in\mathcal{X}} \mathbb{H}\left(\partial\left(\frac{1}{N}\sum_{n=1}^N F(x,\xi^n)\right), \partial\mathbb{E}[F(x,\xi)]\right) = 0 \tag{4.40}$$

Proof. The assertion is a special case of a recently established result [20, Lemma 5.1]. We include a proof for the completeness. For the simplicity of the notation, let

$$P_N := \frac{1}{N}\sum_{n=1}^N \mathbb{1}_{\xi^n}(\omega)$$

where

$$\mathbb{1}_{\xi^n}(\omega) := \begin{cases} 1, & \text{if } \xi(\omega) = \xi^n, \\ 0, & \text{if } \xi(\omega) \neq \xi^n. \end{cases}$$

Then $\mathbb{E}_{P_N}[F(x,\xi)] = \frac{1}{N}\sum_{n=1}^N F(x,\xi^n)$ and hence

$$\partial\mathbb{E}_{P_N}[F(x,\xi)] = \partial\left(\frac{1}{N}\sum_{n=1}^N F(x,\xi^n)\right).$$

Let $f_{P_N}(x) = \mathbb{E}_{P_N}[F(x,\xi)]$ and $f_P(x) = \mathbb{E}[F(x,\xi)]$. Under condition (a), both $f_{P_N}(x)$ and $f_P(x)$ are globally Lipschitz continuous, therefore the Clarke's generalized derivatives of $f_{P_N}(x)$ and $f_P(x)$, denoted by $f^o_{P_N}(x;h)$ and $f^o_P(x;h)$ respectively, are well-defined for any fixed nonzero vector $h \in \mathbb{R}^m$, where

$$f^o_{P_N}(x;h) = \limsup_{x'\to x, \tau\downarrow 0} \frac{1}{\tau}\left(f_{P_N}(x'+\tau h) - f_{P_N}(x')\right)$$

and

$$f_P^o(x;h) = \limsup_{x'\to x, \tau\downarrow 0} \frac{1}{\tau}\left(f_P(x'+\tau h) - f_P(x')\right).$$

Our idea is to study the Hausdorff distance $\mathbb{H}(\partial f_{P_N}(x), \partial f_P(x))$ through certain "distance" of the Clarke generalized derivatives $f_{P_N}^o(x;h)$ and $f_P^o(x;h)$. Let D_1, D_2 be two convex and compact subsets of $\mathbb{R}^m$. Let $\sigma(D_1,u)$ and $\sigma(D_2,u)$ denote the support functions of D_1 and D_2 respectively, where $\sigma(D_i,u)$ is defined as for $i=1$ and 2

$$\sigma(D_i,u) = \sup_{x\in D_i} u^T x$$

for every $u \in \mathbb{R}^m$. Then

$$\mathbb{D}(D_1,D_2) = \max_{\|u\|\le 1}(\sigma(D_1,u) - \sigma(D_2,u))$$

and

$$\mathbb{H}(D_1,D_2) = \max_{\|u\|\le 1}|\sigma(D_1,u) - \sigma(D_2,u)|.$$

The above relationships are known as Hömander's formulae, see [3, Theorem II-18]. Applying the second formula to our setting, we have

$$\mathbb{H}(\partial f_{P_N}(x), \partial f_P(x)) \quad = \quad \sup_{\|h\|\le 1}\left|\sigma(\partial f_{P_N}(x),h) - \sigma(\partial f_{P_N}(x),h)\right|.$$

Using the relationship between Clarke's subdifferential and Clarke's generalized derivative, we have that $f_{P_N}^o(x;h) = \sigma(\partial f_{P_N}(x),h)$ and $f_P^o(x;h) = \sigma(\partial f_P(x),h)$. Consequently,

$$\begin{aligned}
&\mathbb{H}\left(\partial f_{P_N}(x), \partial f_P(x)\right) = \sup_{\|h\|\le 1}\left|f_P^o(x;h) - f_{P_N}^o(x;h)\right| \\
&= \sup_{\|h\|\le 1}\left|\limsup_{x'\to x,\tau\downarrow 0}\frac{1}{\tau}\left(f_P(x'+\tau h) - f_P(x')\right) - \limsup_{x'\to x,\tau\downarrow 0}\frac{1}{\tau}\left(f_{P_N}(x'+\tau h) - f_{P_N}(x')\right)\right|.
\end{aligned}$$

Note that for any bounded sequence $\{a_n\}$ and $\{b_n\}$, we have

$$\left|\limsup_{n\to\infty} a_n - \limsup_{n\to\infty} b_n\right| \le \limsup_{n\to\infty}|a_n - b_n|. \tag{4.41}$$

To see this, let $\{a_{n_j}\}$ be a subsequence such that $\limsup_{n\to\infty} a_n = \lim_{n_j\to\infty} a_{n_j}$. Then

$$\begin{aligned}
\limsup_{n\to\infty}|a_n - b_n| &\ge \limsup_{n_j\to\infty}|a_{n_j} - b_{n_j}| \\
&\ge \limsup_{n_j\to\infty}(a_{n_j} - b_{n_j}) \\
&= \limsup_{n\to\infty} a_n + \limsup_{n_j\to\infty}(-b_{n_j}) \\
&\ge \limsup_{n\to\infty} a_n + \liminf_{n_j\to\infty}(-b_{n_j}) \\
&\ge \limsup_{n\to\infty} a_n + \liminf_{n\to\infty}(-b_n) \\
&= \limsup_{n\to\infty} a_n - \limsup_{n\to\infty} b_n.
\end{aligned}$$

Since a_n and b_n are in a symmetric position, we have that

$$\limsup_{n\to\infty}|a_n - b_n| \ge \limsup_{n\to\infty} b_n - \limsup_{n\to\infty} a_n.$$

This verifies (4.41). Using (4.41), we have

$$\begin{aligned}\mathbb{H}(\partial f_{P_N}(x), \partial f_P(x)) &\leq \sup_{\|h\|\leq 1} \limsup_{x'\to x, \tau\downarrow 0} \left| \frac{1}{\tau}\left(f_P(x'+\tau h) - f_P(x')\right) - \frac{1}{\tau}\left(f_{P_N}(x'+\tau h) - f_{P_N}(x')\right) \right| \\ &= \sup_{\|h\|\leq 1} \limsup_{x'\to x, \tau\downarrow 0} \left| \int_\Xi \frac{1}{\tau}\left(F(x'+\tau h,\xi) - F(x',\xi)\right) d(P-P_N)(\xi) \right|.\end{aligned}$$

Since P_N converges to P in distribution, and the integrand $\frac{1}{\tau}(F(x'+\tau h,\xi) - F(x',\xi))$ is continuous w.r.t ξ and it is bounded by L, by virtue of [2, Theorem 2.1]

$$\lim_{N\to\infty} \sup_{x\in\mathcal{X}} \sup_{\|h\|\leq 1} \limsup_{x'\to x, \tau\downarrow 0} \left| \int_\Xi \frac{1}{\tau}(F(x'+\tau h,\xi) - F(x',\xi)) d(P-P_N)(\xi) \right| = 0.$$

This completes the proof. ■

Let $\phi : \mathbb{R}^n \times \Xi \to \mathbb{R}$ be a real valued function and $\xi : \Omega \to \Xi \subset \mathbb{R}^k$ a random vector defined on probability space $(\Omega, \mathcal{F}, P)$, let $\mathcal{X}$ be a subset of $\mathbb{R}^n$ and $x \in \mathcal{X}$. Recall that ϕ is said to be *H-calm at x from above* with modulus $\kappa(\xi)$ and order γ if $\phi(x,\xi)$ is finite and there exists a (measurable) function $\kappa : \Xi \to \mathbb{R}_+$, positive numbers γ and δ such that

$$\phi(x',\xi) - \phi(x,\xi) \leq \kappa(\xi)\|x'-x\|^\gamma \tag{4.42}$$

for all $x' \in \mathcal{X}$ with $\|x'-x\| \leq \delta$ and almost every $\xi \in \Xi$, see [41].

Using Lemma 4.1 and the concept of H-calmness, we are ready to present one of the main results in this section which state the uniform and exponential convergence of the subdifferentials of underlying functions in defining the Nash equilibrium conditions.

Theorem 4.1 *Let $\{x^N\}$ be a sequence of SAA Nash-C-stationary points. Assume: (a) w.p.1 $\{x^N\}$ is contained in a compact subset $\mathcal{X}$ of X, (b) Assumption 2.1 holds, (c) the Lipschitz modulus of $v_i(x_i, x_{-i}, \xi)$ w.r.t. x_i, $i = 1, \cdots, M$, is bounded by a positive constant C. Then*

(i) w.p.1

$$\lim_{N\to\infty} \sup_{x\in\mathcal{X}} \mathbb{H}(\mathcal{A}\vartheta^N(x), \mathcal{A}\vartheta(x)) = 0. \tag{4.43}$$

(ii) Assume in addition that for every $\xi \in \Xi$ and $x_{-i} \in X_{-i}$: (d) $v_i(x_i, x_{-i}, \xi)$ is piecewise twice continuously differentiable w.r.t. x_i, (e) $v_i(x_i, x_{-i}, \xi)$ is Clarke regular w.r.t. x_i for a.e. $\xi \in \Xi$, (f) the support set of ξ is bounded. Then for every small positive number $\epsilon > 0$, there exist $\hat{c}(\epsilon) > 0$ and $\hat{\beta}(\epsilon) > 0$, independent of N, such that

$$\text{Prob}\left\{ \sup_{x\in\mathcal{X}} \mathbb{D}(\mathcal{A}\vartheta^N(x), \mathcal{A}\vartheta(x)) \geq \epsilon \right\} \;\leq\; \hat{c}(\epsilon) e^{-\hat{\beta}(\epsilon)N} \tag{4.44}$$

for N sufficiently large.

Proof. Part (i). Observe that

$$\mathbb{H}(\mathcal{A}\vartheta^N(x), \mathcal{A}\vartheta(x)) \leq \sum_{i=1}^M \mathbb{H}(\mathcal{A}\vartheta_i^N(x), \mathcal{A}\vartheta_i(x))$$

where $\mathcal{A}\vartheta_i^N(x) = \partial_{x_i}\left(\frac{1}{N}\sum_{n=1}^N v_i(x,\xi^k)\right)$ and $\mathcal{A}\vartheta_i(x) = \partial_{x_i}\mathbb{E}[v_i(x,\xi)]$. Under conditions (a)-(c), it follows by Lemma 4.1 that

$$\lim_{N\to\infty}\sup_{x\in\mathcal{X}}\mathbb{H}(\mathcal{A}\vartheta_i^N(x),\mathcal{A}\vartheta_i(x)) = 0$$

w.p.1 for $i = 1,\cdots,M$, which immediately yields (4.43).

Part (ii). Following the proof in [24, Proposition 3.4], we have the following equation

$$\mathbb{E}\left[\sigma\left(\mathcal{A}v(x,\xi),u\right)\right] = \sigma\left(\mathbb{E}\left[\mathcal{A}v(x,\xi)\right],u\right). \tag{4.45}$$

On the other hand, we observe that it is easy to verify the following inequality

$$\mathbb{D}(\mathcal{A}\vartheta^N(x),\mathcal{A}\vartheta(x)) \le \sum_{i=1}^M \mathbb{D}(\mathcal{A}\vartheta_i^N(x),\mathcal{A}\vartheta_i(x)).$$

By Hömander's formulae [3, Theorem II-18],

$$\mathbb{D}(\mathcal{A}\vartheta_i^N(x),\mathcal{A}\vartheta_i(x)) = \max_{\|u\|\le 1}\sigma(\mathcal{A}\vartheta_i^N(x),u) - \sigma(\mathcal{A}\vartheta_i(x),u).$$

Since

$$\mathcal{A}\vartheta_i^N(x) \subset \frac{1}{N}\sum_{n=1}^N \partial_{x_i}v_i(x,\xi^n),$$

and $\sigma(\partial_{x_i}v_i(x,\xi^n),u) = (v_i)^o_{x_i}(x,\xi^n)$, then

$$\sigma(\mathcal{A}\vartheta_i^N(x),u) \le \frac{1}{N}\sum_{n=1}^N \sigma(\partial_{x_i}v_i(x,\xi^n),u) \le \frac{1}{N}\sum_{n=1}^N (v_i)^o_{x_i}(x,\xi^n;u).$$

where $f^o_{x_i}(x;u)$ denotes the Clarke generalized directional derivative of f w.r.t. x_i at point x in the direction u. Moreover, under conditions (e), it follows from [4, Theorem 2.7.2]

$$\sigma(\mathcal{A}\vartheta_i(x),u) = (\vartheta_i)^o_{x_i}(x;u) = \mathbb{E}\left[(v_i)^o_{x_i}(x,\xi;u)\right].$$

Consequently, we have

$$\begin{aligned}\sup_{x\in\mathcal{X}}\mathbb{D}(\mathcal{A}\vartheta^N(x),\mathcal{A}\vartheta(x)) &\le \sup_{x\in\mathcal{X}}\sum_{i=1}^M \mathbb{D}(\mathcal{A}\vartheta_i^N(x),\mathcal{A}\vartheta_i(x))\\ &\le \sup_{x\in\mathcal{X}}\max_{\|u\|\le 1}\frac{1}{N}\sum_{n=1}^N\left[(v_i)^o_{x_i}(x,\xi^n;u)\right] - \mathbb{E}\left[(v_i)^o_{x_i}(x,\xi;u)\right].\end{aligned}$$

In what follows, we show the uniform exponential convergence of the right hand side of the above inequality. Observe that

$$\|(v_i)^o_{x_i}(x,\xi;u_i)\| \le \|\partial_{x_i}v_i(x,\xi)\| \le \kappa_i(\xi).$$

Under condition (d), we can verify that

$$(v_i)^o_{x_i}(x',\xi;u_i') - (v_i)^o_{x_i}(x,\xi;u_i) \le a_i(\xi)\|x'-x\|^\gamma + \kappa_i(\xi)\|u_i'-u_i\|.$$

Let $z_i := (x, u_i)$ and $Z_i := \mathcal{X} \times \{u_i \in \mathbb{R}^{m_i} : \|u_i\| \leq 1\}$. The inequalities above shows that $(v_i)^o_{x_i}(\cdot, \xi^n; \cdot)$ is H-calm from above on set Z_i. Moreover, under Condition (f), the moment generating function $M_x(t) := \mathbb{E}\left\{e^{t[a_i(\xi)+\kappa_i(\xi)]}\right\}$ is finite valued for t close to 0. By virtue of [41, Proposition 4.1], we have that for any $\epsilon_i > 0$, there exist positive constants $\hat{c}_i(\epsilon_i)$ and $\hat{\beta}_i(\epsilon_i)$, independent of N such that

$$\text{Prob}\left\{\sup_{(x,u)\in Z_i}\left(\frac{1}{N}\sum_{n=1}^{N}(v_i)^o_{x_i}(x, \xi^n; u) - \mathbb{E}\left[(v_i)^o_{x_i}(x, \xi^n; u)\right] \geq \epsilon_i\right)\right\} \leq \hat{c}_i(\epsilon_i)e^{-N\hat{\beta}_i(\epsilon_i)},$$

for $i = 1, 2, \cdots, M$. For any $\epsilon > 0$, let $\epsilon_i > 0$ be such that $\sum_{i=1}^{M} \epsilon_i < \epsilon$. Then, we can show that for every small positive number $\epsilon > 0$, there exist $\hat{c}(\epsilon) > 0$ and $\hat{\beta}(\epsilon) > 0$, independent of N, such that

$$\text{Prob}\left\{\sup_{x\in\mathcal{X}}\max_{\|u\|\leq 1}\sum_{i=1}^{M}\frac{1}{N}\sum_{n=1}^{N}\left[(v_i)^o_{x_i}(x, \xi^n; u)\right] - \mathbb{E}\left[(v_i)^o_{x_i}(x, \xi; u)\right] \geq \epsilon\right\} \leq \hat{c}(\epsilon)e^{-\hat{\beta}(\epsilon)N}, \quad (4.46)$$

where $\hat{c}(\epsilon) = M \max_{i=1}^{M} \hat{c}_i(\epsilon_i)$ and $\hat{\beta}(\epsilon) = \min_{i=1}^{M} \hat{\beta}_i(\epsilon_i)$. The conclusion follows. ■

In what follows, we translate the uniform convergence of the subdifferential in Theorem 4.1 into the convergence of Nash-C-stationary points. We need a perturbation theorem on generalized equation.

Consider the following generalized equation

$$0 \in G(x) + \mathcal{N}_{\mathcal{C}}(x), \quad (4.47)$$

where $G(x) : \mathcal{C} \to 2^{\mathbb{R}^m}$ is a closed set-valued mapping, $\mathcal{C}$ is a closed convex subset of $\mathbb{R}^m$. Let $\tilde{G}(x)$ be a perturbation of $G(x)$ and we consider the perturbed equation

$$0 \in \tilde{G}(x) + \mathcal{N}_{\mathcal{C}}(x). \quad (4.48)$$

Recall that a set-valued mapping F is said to be *outer semicontinuous* (osc for brevity) at $\bar{x} \in \mathbb{R}^n$ if $\overline{\lim}_{x\to\bar{x}} F(x) \subseteq F(\bar{x})$ or equivalently $\lim_{x\to\bar{x}} \mathbb{D}(F(x), F(\bar{x})) = 0$, where

$$\overline{\lim_{x\to\bar{x}}}\, F(x) := \{v \in R^m : \exists \text{ seqences } x_k \to \bar{x}, v_k \to v \text{ with } v_k \in F(x_k)\}.$$

The following lemma states that when $\mathbb{D}(\tilde{G}(x)), G(x))$ is sufficiently small uniformly w.r.t x, then the solution set of (4.48) is close to the solution set of (4.47).

Lemma 4.2 *([39]) Let $\mathcal{W}$ be a compact subset of $\mathcal{C}$. Let X^* denote the set of solutions to (4.47) in $\mathcal{W}$ and Y^* denote the set of solutions to (4.48) in $\mathcal{W}$. Assume that X^* and Y^* are nonempty. Then*

(i) for any $\epsilon > 0$ there exists a $\delta > 0$ such that if $\sup_{x\in\mathcal{C}} \mathbb{D}(\tilde{G}(x), G(x)) < \delta$ and G is outer semicontinuous in $\mathcal{W}$, then $\mathbb{D}(Y^*, X^*) < \epsilon$;

(ii) for any $\epsilon > 0$ there exists a $\delta > 0$ such that if $\sup_{x\in\mathcal{C}} \mathbb{H}(\tilde{G}(x), G(x)) < \delta$ and $\tilde{G}(x)$ is also outer semicontinuous in $\mathcal{W}$, then $\mathbb{H}(Y^*, X^*) < \epsilon$.

Theorem 4.2 *Let $\{x^N\}$ be a sequence of Nash-C-stationary points satisfying (4.34). If conditions (a)-(c) of Theorem 4.1 hold, then w.p.1, an accumulation point of $\{x^N\}$ is a Nash-C-stationary point of the true problem which satisfies the first order necessary equilibrium condition (4.33). If, in addition, conditions (d)-(f) of Theorem 4.1 are satisfied, then for every small positive number $\epsilon > 0$, there exist $\hat{c}(\epsilon) > 0$ and $\hat{\beta}(\epsilon) > 0$, independent of N, such that*

$$\text{Prob}\left\{d(x^N, X^*) \geq \epsilon\right\} \leq \hat{c}(\epsilon)e^{-\hat{\beta}(\epsilon)N} \tag{4.49}$$

for N sufficiently large, where X^* denotes the set of Nash-C-stationary point of the true problem.

Proof. The conclusion follows straightforwardly from Theorem 4.1 and Lemma 4.2. We omit the details. ■

Remark 4.1 *The convergence results established here are stronger than those in our previous work [41]. To see this, recall that in [41] we considered the so-called weak Nash equilibrium conditions for a one stage stochastic Nash equilibrium problem:*

$$0 \in \mathbb{E}[\partial_{x_i} v_i(x_i, x_{-i}, \xi)] + \mathcal{N}_{X_i}(x_i), \ i = 1, \cdots, M, \tag{4.50}$$

where v_i is player i's objective function and $\mathbb{E}[\partial_{x_i} v_i(x_i, x_{-i}, \xi)]$ denotes Aumann's integral of Clarke subdifferential $\partial_{x_i} v_i(x_i, x_{-i}, \xi)$ [1]. Condition (4.50) is weaker than (4.33) in that

$$\partial_{x_i}\mathbb{E}[v_i(x_i, x_{-i}, \xi)] \subset \mathbb{E}[\partial_{x_i} v_i(x_i, x_{-i}, \xi)],$$

see [4, Theorem 2.7.2]. The corresponding first order optimality conditions for the SAA problem considered there are:

$$0 \in \left(\frac{1}{N}\sum_{n=1}^{N} \partial_{x_i} v_i(x_i, x_{-i}, \xi^n)\right) + \mathcal{N}_{X_i}(x_i), \ i = 1, \cdots, M. \tag{4.51}$$

Since

$$\partial_{x_i}\left(\frac{1}{N}\sum_{n=1}^{N} v_i(x_i, x_{-i}, \xi^n)\right) \subset \left(\frac{1}{N}\sum_{n=1}^{N} \partial_{x_i} v_i(x_i, x_{-i}, \xi^n)\right),$$

condition (4.51) is also weaker than (4.34). Roughly speaking, the convergence results established in [41] are about weak Nash-C-stationary point defined through (4.51) to its true counterpart defined through (4.50), whereas the convergence results in Theorem 4.2 are for normal SAA Nash-C-stationary point defined by (4.51) to its true counterpart which satisfies (4.33).

5 Concluding remarks

In this paper, we discuss a two stage stochastic equilibrium problem with equilibrium constraints model and present a few source problems to motivate the model. The model may be extended by including some terms either in the objective or in the constraints which reflect risks such as variance, conditional value at risk [29], chance constraints [26] or certain dominance constraints [6]. To solve the two stage stochastic equilibrium model, we propose to apply the well known sample average approximation method. The exponential rate of convergence means that the sample size will not be very large to obtain a reasonably reliable solution. In the case when the distribution of ξ is finite and known, SAA is not needed. The true problem may be solved through decomposition method or stochastic approximation method, see [26, 37] for stochastic bilevel programs. Finally, we note that the EPEC model is often nonconvex and therefore it would be practically interesting but challenging to identify a Nash equilibrium from a set of obtained Nash stationary points.

References

[1] R. J. Aumann, Integrals of set-valued functions, *Journal of Mathematical Analysis and Applications*, Vol. 12, pp. 1-12, 1965.

[2] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.

[3] C. Castaing and M. Valadier, convex analysis and measurable multifunctions, Lecture Notes in Mathematics, Vol 580, Springer, Berlin, 1977.

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

[5] V. DeMiguel and H. Xu, A Stochastic Multiple Leader Stackelberg Model: Analysis, Computation, and Application, *Operations Research* Vol. 57, pp. 1220-1235, 2009.

[6] D. Dentcheva and A. Ruszczyński, Optimality and duality theory for stochastic optimization with nonlinear dominance constraints, *Mathematical Programming*, Vol. 99, pp. 329–350, 2004.

[7] A. Ehrenmann and K. Neuhoff, A comparison of electricity market designs in networks, *Operations Research*, Vol. 57, pp. 274-286, 2009.

[8] W. Eggert and A. Haufler, Capital taxation and production efficiency in an open economy, *Economics Letters* , Vol. 62, pp. 85-90, 1999.

[9] W. Eggert and M. Kolmar, The taxation of financial capital under asymmetric information and the tax– competition paradox, *Scandinavian Journal of Economics* , Vol. 106, pp. 83-104, 2004.

[10] F. Facchinei and J-S Pang, *Finite-dimensional Variational Inequalities and Complementarity Problems*, Springer, 2003.

[11] G. Gürkan and J-S. Pang, Approximations of Nash equilibria, *Mathematical Programming*, Vol. 117, pp. 223-253, 2008.

[12] C. Hess, Set-valued integration and set-valued probability theory: an overview. Handbook of measure theory, Vol. I, II, 617–673, North-Holland, Amsterdam, 2002.

[13] R. Henerion and W. Römisch, On M-stationary point for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling, *Applications of Mathematics*, Vol. 52, pp. 473-494, 2007.

[14] B. F. Hobbs, C. B. Metzler and J.-S. Pang. Strategic gaming analysis for electric power systems: An MPEC approach. *IEEE Transaction on Power Systems*, Vol. 15, pp. 637-645, 2000.

[15] X. Hu and D. Ralph, Using EPECs to model bilevel games in restructured electricity markets with locational prices, *Operations Research*, Vol. 55, pp. 809-827, 2007.

[16] H. Jiang, D. Ralph and S. Scholtes, Equilibrium problems with equilibrium constraints: A new modelling paradigm for revenue management, report, The 4th Annual INFORMS Revenue Management and Pricing Section Conference, 2004.

[17] A. Koh, Nash Equilibrium, Collusion in Games and the Coevolutionary Particle Swarm Algorithm, White Rose research papers, URL: http://eprints.whiterose.ac.uk/7784/.

[18] S. Leyffer and T. Munson, Solving multi-leader-follower games, *Optimization Methods and Software*, Vol. 25, pp. 601-623, 2010.

[19] G. Lin, H. Xu and M. Fukushima, Monte Carlo and quasi-Monte Carlo sampling methods for a class of stochastic mathematical programs with equilibrium constraints, *Mathematical Methods of Operations Research*, Vol. 67, pp. 423-441, 2008.

[20] Y. Liu, H. Xu and G.H Lin, Stability analysis of two stage stochastic mathematical programs with complementarity constraints via NLP-regularization, Optimization online, February, 2010.

[21] B.S. Mordukhovich, Equilibrium problems with equilibrium constraints via multiobjective optimization, *Optimization Methods and Software*, Vol. 19, pp. 479-492, 2004.

[22] B.S. Mordukhovich, Optimization and equilibrium problems with equilibrium constraints, *Omega*, Vol. 33, pp. 379-384, 2005.

[23] J. Outrata, A note on a class of equilibrium problems with equilibrium constraints, *Kybernetika*, Vol. 40, pp. 585-594, 2004.

[24] N. Papageorgiou, On the theory of Banach space valued multifunctions 1. Integration and conditional expectation, *Journal of Multivariate Analysis*, Vol. 17, pp. 185-206, 1985.

[25] J. S. Pang and M. Fukushima, Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games, *Computational Management Science*, Vol. 2, pp. 21-56, 2005.

[26] G. C. Pflug and N. Broussev, Electricity swing options: behavioral models and pricing, *European Journal of Operational Research*, Vol. 197, pp. 1041–1050, 2009.

[27] E. L. Plambeck, B. R. Fu, S. M. Robinson and R. Suri, Sample-path optimization of convex stochastic performances functions, *Mathematical Programming*, Vol. 75, pp. 137–176, 1996.

[28] D. Ralph. Mathematical programs with complementarity constraints in traffic and telecommunications networks. Royal Society of London. Philosophical Transactions. Mathematical, Physical and Engineering Sciences, Vol. 366, pp. 1973-1987, 2008.

[29] U. Ravat and U. V. Shanbhag, On the characterization of solution sets of smooth and nonsmooth stochastic Nash games, *Proceedings of the American Control Conference (ACC)*, Baltimore, 2010.

[30] S. M. Robinson, Analysis of sample-path optimization, *Mathematics of Operations Research*, Vol. 21, pp. 513–528, 1996.

[31] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.

[32] A. Rusczyński and A.Shapiro, Stochastic Programming Models, in *Stochastic Programming*, A. Rusczyński and A.Shapiro Eds., Handbooks in OR & MS, Vol. 10, North-Holland Publishing Company, Amsterdam, 2003.

[33] A. Shapiro, Stochastic programming with equilibrium constraints, *Journal of Optimization Theory and Application* , Vol. 128, pp. 223-243, 2006.

[34] A. Shapiro, Monte Carlo sampling methods, in A. Rusczyński and A.Shapiro, editors, *Stochastic Programming*, Handbooks in OR & MS, Vol. 10, North-Holland Publishing Company, Amsterdam, 2003.

[35] A. Shapiro and H. Xu, Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation, *Optimization*, Vol. 57, pp. 395-418, 2008.

[36] J. Wang, D. Chiu and C. Lui, A game theoretic analysis of the implications of overlay network traffic on ISP peering, *Computer Networks*, Vol. 52, pp. 2961-2974, 2008.

[37] A. S. Werner, Bilevel Stochastic Programming Problems: Analysis and Application to Telecommunications, PhD dissertation, Norwegian University of Science and Technology, 2004.

[38] C-L. Su, Analysis on the forward market equilibrium model, *Operations Research Letter*, Vol. 35, pp. 74-82, 2007.

[39] H. Xu, Uniform exponential convergence of sample average random functions under general sampling with applications in stochastic programming, to appear in *Journal of Mathematical Analysis and Applications*, 2010.

[40] H. Xu and D. Zhang, Smooth sample average approximation of nonsmooth stochastic problems and applications, *Mathematical Programming Series A*, 2008.

[41] H. Xu and D. Zhang, Stochastic Nash equilibrium problems: sample average approximation and applications, submitted, www.optimization-online.org, 2009.

[42] H. Yang, F. Xiao and H. Huang , Private road competition and equilibrium with traffic equilibrium constraints, *Journal of Advanced Transportation,* Vol. 43, pp. 21-45, 2009.

[43] J. Yao, S. S. Oren and I. Adler, Two-settlement electricity markets with price caps and Cournot generation firms, *European Journal of Operational Research,* Vol. 181, pp. 1279-1296, 2007.

[44] D. Zhang, H. Xu and Y. Wu, A stochastic two stage equilibrium model for electricity markets with two way contracts, *Mathematical Methods of Operations Research*, Vol. 71, pp. 1-45, 2010.

[45] J. Zhou, W. H. K. Lam and B. G. Heydecker, The generalized Nash equilibrium model for oligopolistic transit market with elastic demand, *Transportation Research Part B*, Vol. 39, pp. 519-544,