

REGULARIZED TWO-STAGE STOCHASTIC VARIATIONAL INEQUALITIES FOR COURNOT-NASH EQUILIBRIUM UNDER UNCERTAINTY*

Jie Jiang, Yun Shi, Xiaozhou Wang and Xiaojun Chen

Department of Applied Mathematics, the Hong Kong Polytechnic University, Hong Kong, China

Email: jie.jiang@connect.polyu.hk, yun.shi@connect.polyu.hk, xzhou.wang@connect.polyu.hk, xiaojun.chen@polyu.edu.hk

Abstract

A convex two-stage non-cooperative multi-agent game under uncertainty is formulated as a two-stage stochastic variational inequality (SVI). Under standard assumptions, we provide sufficient conditions for the existence of solutions of the two-stage SVI and propose a regularized sample average approximation method for solving it. We prove the convergence of the method as the regularization parameter tends to zero and the sample size tends to infinity. Moreover, our approach is applied to a two-stage stochastic production and supply planning problem with homogeneous commodity in an oligopolistic market. Numerical results based on historical data in crude oil market are presented to demonstrate the effectiveness of the two-stage SVI in describing the market share of oil producing agents.

Mathematics subject classification: 90C15, 90C33

Key words: Two-stage stochastic variational inequalities, Cournot-Nash equilibrium, Regularized method, Progressive hedging method, Uncertainty, Oil market share

1. Introduction

We consider a two-stage non-cooperative multi-agent game under uncertainty. The model is used to describe a homogenous commodity production and supply planning problem in an oligopolistic market. In particular, we focus on a J -player stochastic Nash equilibrium problem (SNEP) of Cournot competition, whose solution concept is characterized by stochastic Cournot-Nash (CN) equilibria. Conventional non-cooperative game theory has a long history being an effective model to describe market behaviour, yet mostly under deterministic settings. In order to explore characteristics of real markets, one cannot neglect the presence of uncertainty.

Researchers have studied various real markets through SNEPs in recent years. Jofré, Rockafellar and Wets [24] investigated various economic equilibria using stochastic variational inequalities. A scenario-based multi-stage oligopolistic market equilibrium problem under uncertainty was discussed in [16]. A two-settlement oligopolistic equilibrium with uncertainty in the future market was presented in [48]. For practical applications, electricity markets with hydro-electric distribution have been studied by Philpott, Ferris and Wets [30] in which the levels of water reserves were modelled under uncertainty. The contemporary treatment of classical equilibrium problems are investigated through finite-dimensional variational inequalities (VIs) and complementarity problems (CPs) with a wide range of applications under the assumption of deterministic and single-stage decision, see [14, 20] and references therein. Yet, many practical applications are formulated under uncertainty [9, 10, 21, 28, 29, 47], and the corresponding

* Received xxx / Revised version received xxx / Accepted xxx /

multi-stage SVIs and CPs were studied extensively over the last two decades. In [5], Chen and Fukushima considered the stochastic linear CP (LCP) by expected residual minimization (ERM) procedure. The quasi-Monte Carlo method was adopted to generate scenarios of observation and thus to obtain the discrete approximation problem. Chen, Wets and Zhang [10] investigated SVI problems by the ERM procedure, and the sample average approximation (SAA) method was employed to approximate the expected smoothing residual function. More recently, as an extension from single-stage to multi-stage decision processes, Rockafellar and Wets [40] put forward multi-stage SVI problems, which laid a solid foundation for further research. In [6], Chen, Pang and Wets introduced the ERM procedure for two-stage SVI problems and the Douglas-Rachford splitting method was used to present numerical results. Chen, Sun and Xu [9] considered a two-stage stochastic LCP. Structural properties of the problem were studied under the assumption of strong monotonicity, and a discrete scheme was conducted by partition of the support set and the corresponding convergence assertion was established. More generally, Chen, Shapiro and Sun [8] investigated the SAA of two-stage stochastic nonlinear generalized equations, which included two-stage SVI problems as special cases. Exponential rate of convergence was derived by using the technique of perturbed partial linearization. From the perspective of the numerical calculation, the equilibrium problems are usually rewritten as a minimization problem, mostly nonsmooth. For this class of problems, the smoothing techniques (see [3]) can be employed so that differentiable methods, e.g., Newton's method, become applicable in solving the smoothing problem, see for instance [4, 11–13]. For numerical implementation, Rockafellar and Sun extended the well-known progressive hedging method (PHM) for multi-stage stochastic programming problems to multi-stage stochastic VIs in [33].

The decision vectors of production and supply plan problems, so-called *strategies*, are distinguished into two categories: (i) those of “here-and-now” type, which do not depend on outcomes of random events in the future, and (ii) those treated as responses. The goal of this paper is to establish a model that describes the market mechanism for Cournot competition under uncertainties. The solution concepts of the model assemble the real strategies adopted by participants of the market. We will provide methodology to solve the proposed model and demonstrating its numerical implementation.

We summarize the main contributions of this paper as follows.

- A two-stage stochastic Nash equilibrium problem is proposed to model production and supply competition of a homogenous product under uncertainty in an oligopolistic market.
- The model is recasted as two-stage SVIs whose solutions characterize a Nash-Cournot equilibrium.
- A regularized sample average approximation method is proposed to solve the two-stage SVIs with convergence properties under mild assumptions.
- The model is tested numerically for its effectiveness. Moreover, it is used to describe the market share observation in the world market of crude oil. We show that our model is not only able to reproduce historical in-sample market share but also capable of making out-of-sample predictions based on real data sets.

The remaining of the paper is organized as follows. In section 2, the two-stage SNEPs are developed and recasted into two-stage SVIs. Section 3 contains structural analysis of two-stage SVIs and the corresponding regularized problems. The convergence theorems of the proposed

regularized SAA method are presented in section 4. Numerical results based on randomly generated data and simulation on real market data of crude oil are presented and analyzed in section 6. We conclude our study in the final section.

2. Two-stage stochastic Cournot-Nash game

All stochastic models involve inherently “ordered” components over decision horizons. In particular, the decisions in a strategy may respond to the information that is available only at the present stage. In this paper, we consider a *two-stage stochastic Cournot-Nash equilibrium problem*, which extends the classical deterministic CN equilibrium problem in [27]. Let $\xi : \Omega \rightarrow \mathbb{R}^d$ be a random variable with compact support set $\Xi \subseteq \mathbb{R}^d$, and let (Ξ, \mathcal{F}, P) denote the induced probability space. A strategy pair of agent $j \in \mathcal{J} := \{1, \dots, J\}$ is denoted as

$$(x_j \in \mathbb{R}, y_j : \Xi \rightarrow \mathbb{R}, \mathcal{F}\text{-measurable}), \quad (2.1)$$

where x_j is a first-stage decision vector and $y_j \in \mathcal{Y}$ represents a second-stage *response function* with \mathcal{Y} being the space of \mathcal{F} -measurable functions defined on Ξ . Let \mathfrak{L}_n be the Lebesgue space of \mathbb{R}^n -valued functions with \mathfrak{L}_n^∞ denotes the class of measurable essentially bounded functions. Following a similar treatment as in [6], we further require the second-stage response function of random variable to be essentially bounded, i.e., $y_j \in \mathfrak{L}_1^\infty$. Collectively, the vector of strategy pairs of all agents can be written as

$$(x \in \mathbb{R}^J, y : \Xi \rightarrow \mathbb{R}^J, \mathcal{F}\text{-measurable}). \quad (2.2)$$

A *strategy pair* $(x_j^*, y_j^*) \in \mathbb{R} \times \mathfrak{L}_1^\infty$ is said to be an equilibrium of our stochastic model if it solves the following problem for all agents $j \in \mathcal{J}$.

$$\begin{aligned} & \text{maximize}_{(x_j, y_j)} && W_j^1(x_j, x_{-j}^*) + \mathbb{E}[W_j^2(\xi, y_j(\xi), y_{-j}^*(\xi))], && \text{(objective function)} \\ & \text{subject to} && x_j \in X_j, && \text{(first-stage constraints)} \\ & && y_j(\xi) \in_{a.s.} Y_j, \quad g_j(\xi, x_j, y_j(\xi)) \leq_{a.s.} 0 && \text{(second-stage constraints)} \end{aligned} \quad (2.3)$$

where

$$\begin{aligned} x_{-j}^* &= (x_1^*, \dots, x_{j-1}^*, x_{j+1}^*, \dots, x_J^*), \\ y_{-j}^*(\xi) &= (y_1^*(\xi), \dots, y_{j-1}^*(\xi), y_{j+1}^*(\xi), \dots, y_J^*(\xi)), \end{aligned}$$

and $y_j(\xi)$ denotes the value of response y_j to realization $\xi^{(1)}$, with

- $W_j^1 : \mathbb{R} \times \mathbb{R}^{J-1} \rightarrow \mathbb{R}$ a first-stage wealth function of agent j , concave and continuously differentiable with respect to x_j .
- $W_j^2 : \Xi \times \mathbb{R} \times \mathbb{R}^{J-1} \rightarrow \mathbb{R}$ a second-stage wealth function of agent j , concave, well-defined and finite.
- X_j, Y_j nonempty, closed and convex subsets of \mathbb{R} and the second-stage constraints hold almost surely (*a.s.*).
- $g_j : \Xi \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ a continuously differentiable function with respect to $(x_j, y_j(\xi))$ for almost every (a.e.) $\xi \in \Xi$ and \mathcal{F} -measurable.

¹⁾ For ease the exposition, the same notation ξ is used for both a random variable and its realization in \mathbb{R}^d without causing confusion on the context.

In this paper, the model (2.3) is formulated under the assumption that the uncertainty can be described by a random variable ξ with known distribution. From the perspective of the entire system, the market²⁾ “chooses” $x \in \mathbb{R}^J$ before a realization $\xi \in \Xi$ is revealed and later “selects” $y(\xi) \in \mathbb{R}^J$ with known realization.

2.1. Two-stage stochastic commodity production and supply planning

The application of commodity production and supply in an oligopolistic market serves as a motivation as well as the practical problem of interest. Presented as a stochastic game, the strategy of each agent in supply-side of the market can be described as the solution of a stochastic optimization problem (2.3). The decision process follows that agent $j \in \mathcal{J}$ decides an optimal production quantity x_j of the commodity at the production stage. At the second-stage, each agent decides a supply quantity $y_j(\xi)$ after ξ is observed, and a total quantity $T(y(\xi)) := \sum_{j=1}^J y_j(\xi)$ is supplied to the market. Our focus on oligopolistic markets requires that the price is dominantly affected by the total supplied quantity in the market $T(y(\xi))$. Therefore, all the trading occurs at the price $p : \Xi \times \mathbb{R} \rightarrow \mathbb{R}_+$, determined by a stochastic inverse demand curve $p(\xi, T(y(\xi)))$. In practice, production and supply quantities are subject to physical restrictions, e.g., capability of production plant, logistic restriction, etc., i.e., $x_j \in X_j$ and $y_j(\xi) \in_{a.s.} Y_j$. More specifically, we have non-negative requirements for both production and supply, $X_j = \mathbb{R}_+$ and $Y_j = \mathbb{R}_+$. The relations between stage-wise decision variables x_j and $y_j(\xi)$ are captured by constraints $g_j(\xi, x_j, y_j(\xi)) \leq_{a.s.} 0$ in (2.3). Essentially, every agent needs to formulate and solve a *two-stage stochastic programming problem with recourse* in the sense of achieving equilibria of a J -agents non-cooperative game of the market. We further require that agent j 's supply to the market cannot exceed his/her production quantity, i.e., $y_j(\xi) - x_j \leq_{a.s.} 0$. This can be interpreted as the fact that agents may have no stock to start with, or they need to preserve certain reserved quantities prior to each decision process.

The problem can then be viewed from a slight different perspective than that of problem (2.3). As seen from the first-stage, agent $j \in \mathcal{J}$ wants to find a production quantity $x_j \geq 0$ to

$$\text{maximize } W_j^1(x_j, x_{-j}^*) + \mathbb{E}[\Phi_j(x_j, x_{-j}^*, \xi)], \quad (2.4)$$

where,

$$\Phi_j(x_j, x_{-j}^*, \xi) = \sup_{y_j(\xi) \geq 0} \{W_j^2(\xi, y_j(\xi), y_{-j}^*(\xi)) \mid x_j \geq y_j(\xi), \text{ for a.e. } \xi \in \Xi\}. \quad (2.5)$$

Objective function in (2.4) is regarded as the expected profit of agent j 's and problems (2.4)-(2.5) are termed *intrinsic first-stage problem* following that of a related treatment in [35]. In particular, the analysis of intrinsic first-stage problem and the stochastic programming problem with recourse in convex case were carried out in a series of studies by Rockafellar and Wets [34–37] and more recently in [40]. The key feature of intrinsic first-stage problem compares to formulation (2.3) is the requirement on precise orders of decision execution, commonly known as the constraints of *nonanticipativity*. In (2.4)-(2.5), the second-stage decisions are explicitly determined after the first-stage decision, provided for each x_j the second-stage problem is well-defined [43]. However, the study of optimality condition of (2.4)-(2.5), in the case of a general

²⁾ A “social planner” is commonly termed in literature of economics, and can be interpreted as one individual who oversees the system and assigns strategy to each agents.

probability space $(\Xi, \mathcal{F}, P)^{1)}$, is very complicated since one needs to characterize the order of the decision process explicitly. For ease of analysis, we assume that there exists a multiplier $\lambda_j \in \mathfrak{L}_1^1$ corresponds to second-stage constraint and study the saddle-point condition of the Lagrangian formulation of problem (2.3). It is worth mentioning that the Karush-Kuhn-Tucker (KKT) condition of problem (2.3) (see [34]) introduces a second-stage multiplier $\tilde{\lambda}_j \in (\mathfrak{L}_1^\infty)^*$ for every $j \in \mathcal{J}$ which incorporates the two-stage decision making process. This can be seen from the fact that any element of the dual space $(\mathfrak{L}_1^\infty)^*$ can be decomposed into a component of \mathfrak{L}_1^1 and a “singular” component, corresponding to the multiplier of nonanticipativity. The saddle-point condition is shown to be sufficient and “almost” necessary for optimality of problem (2.3), and we refer the interested readers to [34–37, 40] for more details.

The Lagrangian formulation of problem (2.3) associated with agent j is of the following form

$$L_j(x_j, x_{-j}^*, y_j, y_{-j}^*, \lambda_j) = L_j^1(x_j, x_{-j}^*) + \mathbb{E}[L_j^2(\xi, x_j, y_j(\xi), y_{-j}^*(\xi), \lambda_j(\xi))],$$

where

$$\begin{aligned} L_j^1(x_j, x_{-j}^*) &= W_j^1(x_j, x_{-j}^*), \\ L_j^2(\xi, x_j, y_j(\xi), y_{-j}^*(\xi), \lambda_j(\xi)) &= W_j^2(\xi, y_j(\xi), y_{-j}^*(\xi)) + \lambda_j(\xi)(x_j - y_j(\xi)). \end{aligned}$$

The constraints $y_j(\xi) \leq_{a.s.} x_j$ can be interpreted as the situation under which the profit maximizing supply $y_j^*(\xi)$ of agent j is not necessarily equal to the total production quantity x_j . This feature of our model differs from conventional requirement on production-clearing condition, i.e., all the produced goods are expected to supply to the market.

In order to make further progress in characterizing the CN equilibrium, we need to specify the structures of our wealth functions, W_j^1 and W_j^2 , suitable for our application. We assume that the production cost for j -th agent is quadratic, i.e., for each $j \in \mathcal{J}$ the cost of producing x_j amount of production is $\frac{1}{2}c_j x_j^2 + a_j x_j$, for some $c_j > 0, a_j > 0$. In the second-stage, the cost function of the supply or second-stage is assumed to be linear and of stochastic nature, i.e., for each $j \in \mathcal{J}$ the cost of supplying $y_j(\xi)$ amount of commodity is $h_j(\xi)y_j(\xi)$ for a.e. $\xi \in \Xi$. We adopt a classic stochastic inverse demand curve, see for example [23], that takes the expression $p(\xi, T(y(\xi))) = p_0(\xi) - \gamma(\xi)T(y(\xi))$ for the spot price. In practice, the stochastic *benchmark price* excluding the effect of supply to the market $p_0 : \Xi \mapsto \mathbb{R}_+$ can be estimated via statistical approaches based on real data. The supply discount $\gamma : \Xi \rightarrow \mathbb{R}_+$ acts as a market mechanism to adjust and reflect uncertainty in quantity in the market. In order to respect the market mechanism of supply-demand relation, we make the following assumption through out our study.

Assumption 2.1. *There exists a $\gamma_0 > 0$ such that $\gamma(\xi) \geq \gamma_0$ for a.e. $\xi \in \Xi$.*

Thus, agent j 's stage-wise wealth functions are,

$$W_j^1(x_j, x_{-j}^*) = -\frac{1}{2}c_j x_j^2 - a_j x_j,$$

and

$$W_j^2(\xi, y_j(\xi), y_{-j}^*(\xi)) = (p_j(\xi) - \gamma(\xi)T(y(\xi)))y_j(\xi),$$

where the short-handed notation of the risk-adjusted spot price of agent j 's is denoted by $p_j(\xi) := p_0(\xi) - h_j(\xi)$.

¹⁾ In cases of finitely supported distribution, the equivalence between intrinsic first-stage problem and the original recourse problem can be established, and the optimality condition of the recourse problem can be applied, see for example [40].

We are now ready to consider the specific stochastic programming for every agent $j \in \mathcal{J}$

$$\begin{aligned} & \text{maximize}_{x_j} \quad \mathbb{E}[\Phi_j(\xi, x)] - \frac{1}{2}c_j x_j^2 - a_j x_j \\ & \text{subject to} \quad 0 \leq x_j, \end{aligned} \quad (2.6)$$

where

$$\begin{aligned} \Phi_j(\xi, x) = & \text{maximize}_{y_j(\xi)} \quad \left(p_j(\xi) - \gamma(\xi) \left(\sum_{i \neq j}^J y_i^*(\xi) + y_j(\xi) \right) \right) y_j(\xi) \\ & \text{subject to} \quad 0 \leq y_j(\xi) \leq x_j, \text{ for a.e. } \xi \in \Xi. \end{aligned} \quad (2.7)$$

Note that the requirements in problem (2.7) hold almost surely in accordance with the a.s. constraints of the second-stage in problem (2.3). However, (2.6)-(2.7) is not easy to solve, especially in a SNEP with $J \geq 2$, see [8]. The complication arises since the j -th agent's problem contains that of the other agents' strategy, not yet known at the decision horizon. A commonly used method is to recast problem (2.6)-(2.7) of each agent as a stochastic equilibrium problem. Then, obtaining an equilibrium of the convex J -player game (2.6)-(2.7) is equivalent to finding its solution for all agents. Stochastic equilibrium has been shown to be an effective method to study and to solve two-stage multi-players stochastic game problems, see for instance [9, 28, 33, 40, 46]. We study the saddle-point condition of the problem (2.6)-(2.7), rewritten in the form of problem (2.3). More specifically, for all $j \in \mathcal{J}$, there exists $\bar{\lambda}_j(\xi) \in \mathfrak{L}_1^1$ with $\bar{\lambda}_j(\xi) \geq_{a.s.} 0$ so that a strategy $(\bar{x}_j, \bar{y}_j) \in \mathbb{R}_+ \times \mathfrak{L}_+^\infty$ solves the following system.

$$\begin{aligned} & -c_j \bar{x}_j - a_j + \mathbb{E}[\bar{\lambda}_j(\xi)] \in \mathcal{N}_{[0, \infty)}(\bar{x}_j), \\ & p_j(\xi) - \gamma(\xi) \sum_{i \neq j}^J \bar{y}_i(\xi) - 2\gamma(\xi) \bar{y}_j(\xi) - \bar{\lambda}_j(\xi) \in_{a.s.} \mathcal{N}_{[0, \infty)}(\bar{y}_j(\xi)), \quad (\text{stationarity}) \\ & \bar{x}_j \geq 0, \bar{y}_j(\xi) \geq_{a.s.} 0, \bar{x}_j - \bar{y}_j(\xi) \geq_{a.s.} 0, \quad (\text{feasibility}) \\ & \bar{\lambda}_j(\xi) \geq_{a.s.} 0, \quad (\text{dual feasibility}) \\ & \bar{\lambda}_j(\xi) \perp_{a.s.} (\bar{x}_j - \bar{y}_j(\xi)). \quad (\text{complementarity}) \end{aligned}$$

In particular, *stationarity* comes from the first-order necessary optimality condition under the assertion $\partial \mathbb{E}[\Phi_j(\xi, x)] \subseteq \mathbb{E}[\partial_x \Phi(\xi, x)]$. The assertion is discussed in [2], and the above system can be viewed as a weaker condition for optimality.

Rewritten in a compact form as SVI, the optimal strategy-multiplier pair $(x_j, y_j, \lambda_j) \in \mathbb{R}_+ \times \mathfrak{L}_+^\infty \times \mathfrak{L}_+^1$ must satisfy,

$$\begin{aligned} 0 & \leq x_j \quad \perp \quad c_j x_j + a_j - \mathbb{E}[\lambda_j(\xi)] & \geq 0, \\ 0 & \leq_{a.s.} y_j(\xi) \quad \perp_{a.s.} -p_j(\xi) + \gamma(\xi) \sum_{i \neq j}^J y_i(\xi) + 2\gamma(\xi) y_j(\xi) + \lambda_j(\xi) & \geq_{a.s.} 0, \quad (2.8) \\ 0 & \leq_{a.s.} \lambda_j(\xi) \quad \perp_{a.s.} x_j - y_j(\xi) & \geq_{a.s.} 0. \end{aligned}$$

It follows that since all agents in oligopolistic market act non-cooperatively, we write down the equilibrium interpreted as that of the whole system. More specifically, let $x = (x_1, \dots, x_J)^T$ be the first-stage decision vectors of the system, and for almost every $\xi \in \Xi$, second-stage decision vector $y(\xi) = (y_1(\xi), \dots, y_J(\xi))^T$ and the corresponding multiplier vector $\lambda(\xi) =$

$(\lambda_1(\xi), \dots, \lambda_J(\xi))^T$ are denoted respectively. Similarly, parameter vectors can be written collectively as $a = (a_1, \dots, a_J)^T$, $p(\xi) = (p_1(\xi), \dots, p_J(\xi))^T$. Then, we can treat the SVI for all agents as a two-stage stochastic complementary problem (SCP):

$$\begin{aligned} 0 \leq x & \quad \perp \quad Cx - \mathbb{E}[\lambda(\xi)] + a & \geq 0, \\ 0 \leq \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} & \quad \perp \quad \begin{pmatrix} \Pi(\xi) & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} + \begin{pmatrix} -p(\xi) \\ x \end{pmatrix} & \geq 0, \quad \text{for a.e. } \xi \in \Xi, \end{aligned} \quad (2.9)$$

where

$$C = \text{diag}(c_1, c_2, \dots, c_J), \quad \Pi(\xi) = \gamma(\xi)(ee^T + I).$$

It follows that for the whole system, a J -tuple of strategies

$$(x^*, y^*, \lambda^*) = ((x_1^*, y_1^*, \lambda_1^*), \dots, (x_J^*, y_J^*, \lambda_J^*)) \in \mathbb{R}^J \times \mathfrak{L}_J^\infty \times \mathfrak{L}_J^1$$

is called a solution of the two-stage stochastic complementarity problem (2.9).

3. Structure of two-stage stochastic complementarity problem

In this section, we focus on characterizing solutions of two-stage stochastic linear complementarity problem (2.9). From the derivation of first-order necessary optimality conditions of problem (2.6)-(2.7) and the monotonicity of problem (2.9), we have the following results on existence of solutions.

Proposition 3.1 (Theorem 2, [41]) *For any fixed pair $(x, \xi) \in \mathbb{R}_+^J \times \Xi$, the second-stage problem (2.7) has a unique solution.*

Thus, for the two-stage stochastic linear complementarity problem (2.9), the following proposition holds.

Proposition 3.2. *The two-stage stochastic linear complementarity problem (2.9) has relatively complete recourse, i.e., for any $x \in \mathbb{R}_+^J$ and a.e. $\xi \in \Xi$ the second-stage problem of (2.9) is solvable.*

Proof. The coefficient matrix of the second-stage part of (2.9)

$$M(\xi) = \begin{pmatrix} \Pi(\xi) & I \\ -I & 0 \end{pmatrix}$$

is positive semidefinite for a.e. $\xi \in \Xi$.

For any given $x \in \mathbb{R}_+^J$, it follows that there always exists a pair $(\hat{y}(\xi), \hat{\lambda}(\xi)) \in \mathbb{R}^J \times \mathbb{R}^J$, such that

$$\begin{pmatrix} \hat{y}(\xi) \\ \hat{\lambda}(\xi) \end{pmatrix} \geq 0, \quad \begin{pmatrix} \Pi(\xi) & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \hat{y}(\xi) \\ \hat{\lambda}(\xi) \end{pmatrix} + \begin{pmatrix} -p(\xi) \\ x \end{pmatrix} \geq 0, \quad \text{for a.e. } \xi \in \Xi.$$

In detail, we consider a special choice $\hat{y}(\xi) = 0$ and $\hat{\lambda}(\xi) = \max\{0, p(\xi)\}$, where the max function is taken componentwise. Thus, the corresponding quadratic programming problem of the linear complementarity problem is feasible. It follows from [32, Lemma 3.1.1, Theorem 3.1.2] that there must exist at least a solution which solves the second-stage problem for any given pair (x, ξ) . \square

Although the second-stage problem (2.7) has a unique equilibrium for any given (x, ξ) (see Proposition 3.1), the system (2.9) may admit multiple solutions. To see this, we give an illustrative example.

Example 3.1. Consider a duopoly game, with given $x = (x_1, x_2)^T \geq 0$, and $-p(\xi) \geq_{a.s.} 0$. Then, the corresponding second-stage part of complementarity system (2.9) reads

$$0 \leq \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \lambda_1(\xi) \\ \lambda_2(\xi) \end{pmatrix} \perp \begin{pmatrix} 2\gamma(\xi) & \gamma(\xi) & 1 & 0 \\ \gamma(\xi) & 2\gamma(\xi) & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \lambda_1(\xi) \\ \lambda_2(\xi) \end{pmatrix} + \begin{pmatrix} -p_1(\xi) \\ -p_2(\xi) \\ x_1 \\ x_2 \end{pmatrix} \geq 0, \quad \text{for a.e. } \xi \in \Xi. \quad (3.1)$$

Then, the solution set of (3.1) is of the following form

$$\left\{ (0, 0, \tilde{\lambda}_1(\xi), \tilde{\lambda}_2(\xi)) : \tilde{\lambda}_1(\xi) = \begin{cases} 0, & x_1 > 0 \\ \lambda_1(\xi), & x_1 = 0 \end{cases}, \tilde{\lambda}_2(\xi) = \begin{cases} 0, & x_2 > 0 \\ \lambda_2(\xi), & x_2 = 0 \end{cases}, \quad \text{for a.e. } \xi \in \Xi \right\},$$

where $\lambda_1(\xi) \geq_{a.e.} 0, \lambda_2(\xi) \geq_{a.e.} 0$.

In Example 3.1, the “equilibrium price” λ may admit multiple values when there exist some zero-valued components of x .

Technically, the multiple solutions of the second-stage problem will cause trouble when we handle the two-stage stochastic complementarity system (2.9), both in computation and analysis [43]. The assumption ensuring the uniqueness of second-stage solution is usually made, see for instance [8, 9]. Moreover, interpreted as “equilibrium price” associated with agents’ production clearing, different values of λ would have ambiguous economical interpretations. Motivated by these, we propose a regularized method to seek for one particular choice of “equilibrium price”. Similar approach can be found in for example [7].

For any $\epsilon > 0$, let

$$M^\epsilon(\xi) = \begin{pmatrix} \Pi(\xi) & I \\ -I & \epsilon I \end{pmatrix} \quad \text{and} \quad q(x, \xi) = \begin{pmatrix} -p(\xi) \\ x \end{pmatrix}$$

and thus, we can write the regularized second-stage SCP as

$$0 \leq \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} \perp M^\epsilon(\xi) \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} + q(x, \xi) \geq 0, \quad \text{for a.e. } \xi \in \Xi. \quad (3.2)$$

Then, we have the regularized SCP of (2.9) as follows:

$$\begin{aligned} 0 \leq x & \perp Cx - \mathbb{E}[\lambda(\xi)] + a & \geq 0, \\ 0 \leq \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} & \perp M^\epsilon(\xi) \begin{pmatrix} y(\xi) \\ \lambda(\xi) \end{pmatrix} + q(x, \xi) & \geq 0, \quad \text{for a.e. } \xi \in \Xi. \end{aligned} \quad (3.3)$$

For a given pair $(x, \xi) \in \mathbb{R}_+^J \times \Xi$, the second-stage problem of (2.9) and the regularized second-stage problem (3.2) are denoted by $\text{LCP}(q(x, \xi), M(\xi))$ and $\text{LCP}(q(x, \xi), M^\epsilon(\xi))$ respectively. Their solution functions are chosen from the respective solution sets and expressed by $z(q(x, \xi))$ and $z^\epsilon(q(x, \xi))$. In the sequel, we omit the ξ and x without causing confusion, i.e., $\text{LCP}(q, M) := \text{LCP}(q(x, \xi), M(\xi))$ and $\text{LCP}(q, M^\epsilon) := \text{LCP}(q(x, \xi), M^\epsilon(\xi))$.

For clearer demonstration, recall our illustrative Example 3.1, and consider its regularization approach. Thus, the second-stage of the regularized problem takes the following form

$$0 \leq \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \lambda_1(\xi) \\ \lambda_2(\xi) \end{pmatrix} \perp \begin{pmatrix} 2\gamma(\xi) & \gamma(\xi) & 1 & 0 \\ \gamma(\xi) & 2\gamma(\xi) & 0 & 1 \\ -1 & 0 & \epsilon & 0 \\ 0 & -1 & 0 & \epsilon \end{pmatrix} \begin{pmatrix} y_1(\xi) \\ y_2(\xi) \\ \lambda_1(\xi) \\ \lambda_2(\xi) \end{pmatrix} + \begin{pmatrix} -p_1(\xi) \\ -p_2(\xi) \\ x_1 \\ x_2 \end{pmatrix} \geq 0, \quad \text{for a.e. } \xi \in \Xi. \quad (3.4)$$

Under the same condition as in Example 3.1, we can obtain the unique solution of (3.4), which $\tilde{y}_1, \tilde{y}_2, \tilde{\lambda}_1, \tilde{\lambda}_2$ equal to 0 for a.e. $\xi \in \Xi$. Due to the positive definiteness of C , it follows that we obtain the unique solution of the first-stage problem is $x_1 = 0, x_2 = 0$. Then, we have obtained one particular solution of the original problem, the trivial solution in this example. The key feature of our regularized method is that it promises the existence and uniqueness of solution due to the strongly monotone of regularized two-stage problem.

In the remaining of this section, we concern ourselves with the solution z^ϵ of $\text{LCP}(q, M^\epsilon)$ and explore the structure of the second-stage solution.

Proposition 3.3. *For any fixed $\epsilon > 0$, the regularized problem (3.3) has a unique solution $(x^\epsilon, y^\epsilon, \lambda^\epsilon) \in \mathbb{R}^J \times \mathcal{Y} \times \mathcal{Y}$.*

Proof. The result can be obtained via a similar procedure as in [9, Proposition 2.1 (i)] and we only need to show that the condition in [9, Assumption 1] holds. Recall that Assumption 2.1 holds, then for a.e. $\xi \in \Xi$

$$\begin{pmatrix} x \\ u(\xi) \\ v(\xi) \end{pmatrix}^T \begin{pmatrix} C & 0 & -I \\ 0 & \Pi(\xi) & I \\ I & -I & \epsilon I \end{pmatrix} \begin{pmatrix} x \\ u(\xi) \\ v(\xi) \end{pmatrix} \geq \tau(\|x\|^2 + \|u(\xi)\|^2 + \|v(\xi)\|^2),$$

where $\tau = 2 \min\{\bar{c}, \gamma_0(J+1), \epsilon\}$ with \bar{c} denoting the minimum diagonal element of C . \square

Theorem 3.1. *For any fixed $\epsilon > 0$, $x \geq 0$ and a.e. $\xi \in \Xi$, the j -th component of the solution of problem (3.2) $((y^\epsilon)_j, (\lambda^\epsilon)_j)$ is either $(0, 0)$, or one of the following two forms:*

$$\begin{aligned} & - \left(\frac{\gamma(\xi)T^\epsilon - p_j(\xi)}{\gamma(\xi)}, \quad 0 \right), \\ & - \left(\frac{\epsilon(\gamma(\xi)T^\epsilon - p_j(\xi)) - x_j}{\epsilon\gamma(\xi) + 1}, \quad \frac{\gamma(\xi)(T^\epsilon + x_j) - p_j(\xi)}{\epsilon\gamma(\xi) + 1} \right) \end{aligned} \quad (3.5)$$

for $j \in \mathcal{J}$, where

$$T^\epsilon := \sum_{i=1}^J (y^\epsilon)_i = \frac{\gamma(\xi) \sum_{i \in \mathcal{I}_3} x_i + \epsilon\gamma(\xi) \sum_{i \in \mathcal{I}_2 \cup \mathcal{I}_3} p_i(\xi) + \sum_{i \in \mathcal{I}_2} p_i(\xi)}{(\epsilon\gamma(\xi)(|\mathcal{I}_2| + |\mathcal{I}_3| + 1) + |\mathcal{I}_2| + 1)\gamma(\xi)} \quad (3.6)$$

with

$$\begin{aligned} \mathcal{I}_2 &= \{j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, \quad (y^\epsilon)_j - x_j \leq 0\}, \\ \mathcal{I}_3 &= \{j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, \quad (y^\epsilon)_j - x_j > 0\}, \end{aligned}$$

where $|\mathcal{I}_2|$ and $|\mathcal{I}_3|$ denote the cardinality of \mathcal{I}_2 and \mathcal{I}_3 respectively.

The detailed proof of the Theorem 3.1 is given in Appendix. Note that the above theorem gives the forms of the unique solution of the second-stage regularized problem (3.2). However, it may be used to assist numerical calculation since the partition of the index set is not known in advance. Nevertheless, it is suffice for our purposes of deriving additional properties of the solutions. Due to the positive definiteness of M^ϵ and special structure of problem (3.2), we first obtain the following Lipschitz continuous property, following [11, Corollary 2.1].

Lemma 3.1. *There exists $L > 0$ such that for any fixed $\epsilon \in (0, 1]$, we have*

$$\|z^\epsilon(q(x_1, \xi)) - z^\epsilon(q(x_2, \xi))\| \leq L\|x_1 - x_2\|, \quad \text{for } x_1, x_2 \in \mathbb{R}_+^J \text{ and } \xi \in \Xi.$$

Lemma 3.2. *For any fixed $\epsilon > 0$ and $(x, \xi) \in \mathbb{R}_+^J \times \Xi$, T^ϵ has the following upper bound:*

$$T^\epsilon \leq \|x\|_1 + \left(\epsilon + \frac{1}{\gamma(\xi)} \right) \|p(\xi)\|_1.$$

Proof. We have the following derivation from (3.6) that

$$\begin{aligned} T^\epsilon &= \frac{\gamma(\xi) \sum_{i \in \mathcal{I}_3} x_i + \epsilon \gamma(\xi) \sum_{i \in \mathcal{I}_2 \cup \mathcal{I}_3} p_i(\xi) + \sum_{i \in \mathcal{I}_2} p_i(\xi)}{(\epsilon \gamma(\xi)(|\mathcal{I}_2| + |\mathcal{I}_3| + 1) + |\mathcal{I}_2| + 1) \gamma(\xi)} \\ &\leq \frac{\gamma(\xi) \sum_{i=1}^J x_i + (\epsilon \gamma(\xi) + 1) \sum_{i=1}^J |p_i(\xi)|}{(\epsilon \gamma(\xi)(|\mathcal{I}_2| + |\mathcal{I}_3| + 1) + |\mathcal{I}_2| + 1) \gamma(\xi)} \\ &\leq \frac{\gamma(\xi) \sum_{i=1}^J x_i + (\epsilon \gamma(\xi) + 1) \sum_{i=1}^J |p_i(\xi)|}{\gamma(\xi)} \\ &= \|x\|_1 + \left(\epsilon + \frac{1}{\gamma(\xi)} \right) \|p_1(\xi)\|_1. \end{aligned}$$

□

We end this section by establishing the convergence result of the second-stage LCP(q, M^ϵ) solutions as $\epsilon \downarrow 0$ for any given pair $(x, \xi) \in \mathbb{R}^J \times \Xi$.

Proposition 3.4. *For any fixed $\epsilon > 0$ and $(x, \xi) \in \mathbb{R}_+^J \times \Xi$, let $z^\epsilon(\xi) = (y^\epsilon(\xi), \lambda^\epsilon(\xi))$ denote the unique solution of the regularized problem LCP(q, M^ϵ). Then*

$$\lim_{\epsilon \downarrow 0} \|z^\epsilon(\xi) - \bar{z}(\xi)\| = 0,$$

where $\bar{z}(\xi) = (\bar{y}(\xi), \bar{\lambda}(\xi))$ denotes the unique least l_2 -norm solution of the LCP(q, M). Moreover, the j -th component of the least l_2 -norm solution of problem (2.9) has one of the following three forms:

$$\left\{ (0, 0), \left(-\frac{\gamma(\xi)\bar{T} - p_j(\xi)}{\gamma(\xi)}, 0 \right), \left(x_j, -\gamma(\xi)(\bar{T} + x_j) + p_j(\xi) \right) \right\} \quad (3.7)$$

for $j \in \mathcal{J}$, where

$$\bar{T} := \lim_{\epsilon \downarrow 0} T^\epsilon = \sum_{i=1}^J \bar{y}_i.$$

Furthermore, for a.e. $\xi \in \Xi$ there exists $\bar{\kappa}(\xi) > 0$, such that

$$\|\lambda^\epsilon(\xi) - \bar{\lambda}(\xi)\| \leq \bar{\kappa}(\xi)\epsilon. \quad (3.8)$$

The proof of Proposition 3.4 is given in Appendix.

4. Convergence analysis

In this section, we first prove the convergence of the unique solution of the regularized problem (3.3) to the solution set of the original problem as the regularized parameter ϵ decreases

to zero. Next, we will study the sample average approximation (SAA) to solve the regularized problem. See [7]. Combined with our regularization approaches, we demonstrate the convergence property of the solution of our regularized SAA model as the number of samples goes to infinity. More specifically, the convergence analysis in this section is divided into two parts: the convergence analysis of the regularized problem as the regularized parameter ϵ tends to zero, and the analysis of regularized SAA. We finally build up the convergence relationship between the regularized SAA approach and the original problem.

4.1. Convergence of the regularized model

In this subsection, we only need to consider the convergence properties of the first-stage decision vector, i.e., $x^\epsilon \in \mathbb{R}_+^J$ that solves problem (3.3), when the regularized parameter ϵ tends to zero. The convergence property of the solution $(x^\epsilon, y^\epsilon, \lambda^\epsilon)$ then follows combining the result of section 3. From Proposition 3.3, we know that for fixed $\epsilon > 0$ problem (3.3) admits a unique first-stage solution x^ϵ . In the following, we concern about the sequence of accumulation points of $\{x^\epsilon\}$ as $\epsilon \downarrow 0$.

For the existence of accumulation points, we have the following result.

Proposition 4.1. *Suppose there exists $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then, with $\epsilon \downarrow 0$, $\{x^\epsilon\}$ is bounded.*

Proof. From the condition on $p_j(\xi)$, there must a sufficiently large $\alpha > 0$ such that for any $j \in \mathcal{J}$

$$\gamma_0 \alpha - p_j(\xi) > 0, \quad \text{for a.e. } \xi \in \Xi.$$

Then we have

$$-\frac{\gamma(\xi)(T^\epsilon + \alpha) - p_j(\xi)}{\epsilon \gamma(\xi) + 1} < 0, \quad \text{for a.e. } \xi \in \Xi.$$

Assume that $\{x^\epsilon\}$ is unbounded for the purpose of arriving at a contradiction. Then, it follows that there exist some indices $j \in \mathcal{J}$, such that $(x^\epsilon)_j \geq \alpha$. Then we consider the j -th component of the first-stage complementarity relation,

$$0 \leq (x^\epsilon)_j \perp c_j(x^\epsilon)_j - \mathbb{E}[(\lambda^\epsilon(\xi))_j] + a_j \geq 0,$$

which can be expressed, from (3.5), as

$$0 \leq (x^\epsilon)_j \perp c_j(x^\epsilon)_j + a_j \geq 0.$$

However, this complementarity relation cannot be obtained because $(x^\epsilon)_j > 0$ and $c_j(x^\epsilon)_j + a_j > 0$. This completes our proof. \square

Note that the conditions $p_j(\xi) \leq_{a.s.} p_0$ can be easily satisfied in many practical applications. For example, with given data sets of $p(\xi)$ we can always find an upper bound $p_0 := \max_j \{p_j(\xi)\}$.

Lemma 4.1. *Suppose there exist a constant $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then, there exists a subsequence $\{\epsilon_k\}_{k=1}^\infty$ with $\epsilon_k \downarrow 0$ as $k \rightarrow \infty$ such that with $\lambda^k = \lambda^{\epsilon_k}$,*

$$\lim_{k \rightarrow \infty} \mathbb{E}[\lambda^k(\xi)] = \mathbb{E}[\bar{\lambda}(\xi)].$$

Proof. From the results of Proposition 4.1, there must exist an accumulation point of $\{x_\epsilon\}$ as $\epsilon \downarrow 0$, denoted by \hat{x} . Taking a subsequence $\{\epsilon_k\}_{k=1}^\infty$ with $\epsilon_k \downarrow 0$ as $k \rightarrow \infty$. We derive from Proposition 4.1 and (3.8) that

$$\begin{aligned} & \left\| \mathbb{E}[\lambda^k(\xi)] - \mathbb{E}[\bar{\lambda}(\xi)] \right\| \\ & \leq \left\| \mathbb{E}[\lambda^k(\xi)] - \mathbb{E}[\hat{\lambda}(\xi)] \right\| + \left\| \mathbb{E}[\hat{\lambda}(\xi)] - \mathbb{E}[\bar{\lambda}(\xi)] \right\| \\ & \leq \mathbb{E}[\bar{\kappa}(\xi)]\epsilon_k + \left\| \mathbb{E}[\hat{\lambda}(\xi)] - \mathbb{E}[\bar{\lambda}(\xi)] \right\|, \end{aligned}$$

where for given $\xi \in \Xi$, $\hat{\lambda}(\xi)$ denotes the multiplier component of the least norm solution of problem (2.9) with \hat{x} . It follows from Lemma 3.1 that for a.e. $\xi \in \Xi$

$$\left\| \bar{\lambda}(\xi) - \hat{\lambda}(\xi) \right\| = 0.$$

Furthermore, we derive that the estimation

$$\begin{aligned} \left\| \bar{\lambda}(\xi) - \hat{\lambda}(\xi) \right\| & \leq \left\| \bar{\lambda}(\xi) \right\| + \left\| \hat{\lambda}(\xi) \right\| \\ & \leq 4\sqrt{J}(\gamma(\xi) \|w\|_1 + \|p(\xi)\|_1), \quad \text{for a.e. } \xi \in \Xi, \end{aligned}$$

where the last term comes from (3.7) with some $w > 0$. It follows from the Lebesgue Dominated Convergence Theorem, we have

$$\left\| \mathbb{E}[\hat{\lambda}(\xi)] - \mathbb{E}[\bar{\lambda}(\xi)] \right\| = 0.$$

□

Theorem 4.1. *Any accumulation point of $\{x^\epsilon, y^\epsilon, \lambda^\epsilon\}$ as $\epsilon \downarrow 0$ is a solution of problem (2.9).*

Proof. We only need to verify that for any $\epsilon_k \downarrow 0$, the accumulation point \hat{x} of subsequence $\{x^{\epsilon_k}\}$ is a first-stage solution of (2.9). Since x^{ϵ_k} is the first-stage solution of problem (3.3) for any $\epsilon_k > 0$, we have with $x^k = x^{\epsilon_k}$

$$0 \leq x^k \perp Cx^k - \mathbb{E}[\lambda^k(\xi)] + a \geq 0,$$

which, by using the ‘min’ NCP function (see, for example, [32]), can be rewritten as

$$\min\{x^k, Cx^k - \mathbb{E}[\lambda^k(\xi)] + a\} = 0.$$

By Lemma 4.1, we have

$$0 = \lim_{k \rightarrow \infty} \min\{x^k, Cx^k - \mathbb{E}[\lambda^k(\xi)] + a\} = \min\{\hat{x}, C\hat{x} - \mathbb{E}[\bar{\lambda}(\xi)] + a\}$$

as $k \rightarrow \infty$. Thus we obtain that

$$\min\{\hat{x}, C\hat{x} - \mathbb{E}[\bar{\lambda}(\xi)] + a\} = 0.$$

The statement then follows from Proposition 3.4. □

4.2. Convergence of the regularized SAA model

In this subsection, we study the SAA scheme for solving the regularized problem (3.3) and focus on the convergence of the regularized SAA approach. More specifically, we focus on the SAA convergence analysis and the solution of the first-stage problem. It is noteworthy that Chen, Sun and Xu considered a discrete approximation scheme in [9], which also leads to an approximation of the response variable in the second-stage problem.

Let $\xi_1, \xi_2, \dots, \xi_\nu$ denote ν independent identically distributed (i.i.d.) samples. Then, with slight abuse of notation, we can obtain the following formulation of problem (3.3) with SAA:

$$\begin{aligned} 0 \leq x \perp Cx - \frac{1}{\nu} \sum_{\ell=1}^{\nu} \lambda(\xi_\ell) + a \geq 0, \\ 0 \leq \begin{pmatrix} y(\xi_\ell) \\ \lambda(\xi_\ell) \end{pmatrix} \perp \begin{pmatrix} \Pi(\xi_\ell) & I \\ -I & \epsilon I \end{pmatrix} \begin{pmatrix} y(\xi_\ell) \\ \lambda(\xi_\ell) \end{pmatrix} + \begin{pmatrix} -p(\xi_\ell) \\ x \end{pmatrix} \geq 0, \ell = 1, \dots, \nu. \end{aligned} \quad (4.1)$$

Or, we can write the problem collectively for all ν samples

$$0 \leq \begin{pmatrix} x \\ v_1 \\ \vdots \\ v_\nu \end{pmatrix} \perp \begin{pmatrix} C & \frac{1}{\nu}B & \dots & \frac{1}{\nu}B \\ -B^T & D_1^\epsilon & & \\ \vdots & & \ddots & \\ -B^T & & & D_\nu^\epsilon \end{pmatrix} \begin{pmatrix} x \\ v_1 \\ \vdots \\ v_\nu \end{pmatrix} + \begin{pmatrix} a \\ q_1 \\ \vdots \\ q_\nu \end{pmatrix} \geq 0, \quad (4.2)$$

where $C \in R^{J \times J}$, $B = \begin{pmatrix} 0 & -I \end{pmatrix} \in R^{J \times 2J}$, and for $\ell = 1, \dots, \nu$, $D_\ell^\epsilon = \begin{pmatrix} \Pi(\xi_\ell) & I \\ -I & \epsilon I \end{pmatrix} \in R^{2J \times 2J}$, $v_\ell = (y(\xi_\ell), \lambda(\xi_\ell))^T$, $q_\ell = (-p(\xi_\ell), 0)^T$. Thus, (4.2) is treated as a large-scale deterministic linear complementarity problem:

$$0 \leq z \perp H^\epsilon z + \bar{q} \geq 0, \quad (4.3)$$

where $z = (x, v_1, \dots, v_\nu)^T$, $\bar{q} = (a, q_1, \dots, q_\nu)^T$, and H^ϵ denotes the coefficient matrix in (4.2).

We have the following assertion of existence and uniqueness of problem (4.1) by [32, Theorem 3.1.6].

Proposition 4.2. *For any fixed $\epsilon > 0$ and positive integer ν , there exists a unique solution of problem (4.1).*

Recall the result of Lemma 3.1 and the following proposition can be shown in a similar way as in [7, Proposition 3.7].

Proposition 4.3. *Let $(y^\epsilon(\xi), \lambda^\epsilon(\xi))$ be the unique solution of the regularized second-stage problem (3.3) for any $(x, \xi) \in \mathbb{R}_+^J \times \Xi$. Then,*

$$\frac{1}{\nu} \sum_{\ell=1}^{\nu} \lambda^\epsilon(\xi_\ell) \rightarrow \mathbb{E}[\lambda^\epsilon(\xi)]$$

with probability (w.p.) 1 as $\nu \rightarrow \infty$ uniformly on $\mathbb{B}(x, \delta) \cap \mathbb{R}_+^J$ for any $\delta > 0$,

Let x_ν^ϵ denote the first J -components of the unique solution of problem of (4.1), and we have the following assertion.

Lemma 4.2. *Suppose there exists $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then, with $\epsilon \downarrow 0$, $\{x_\nu^\epsilon\}$ is bounded.*

We omit the proof since it can be shown analogously as in Proposition 4.1.

Theorem 4.2. *Suppose there exists $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then, for any fixed $\epsilon > 0$, $x_\nu^\epsilon \rightarrow x^\epsilon$ w.p. 1 as $\nu \rightarrow \infty$.*

Proof. From Propositions 3.3, Proposition 4.2, and Lemma 4.2, for any fixed $\epsilon > 0$, both the regularized problem (3.3) and its SAA-regularized problem (4.1) have solutions and contained in some compact subset in \mathbb{R}_+^J . We know from Proposition 4.3 that

$$\frac{1}{\nu} \sum_{\ell=1}^{\nu} \lambda^\epsilon(\xi_\ell) \rightarrow \mathbb{E}[\lambda^\epsilon(\xi)]$$

as $\nu \rightarrow \infty$, uniformly with respect to x on any compact set. Then, we have $x_\nu^\epsilon \rightarrow x^\epsilon$ w.p. 1 as $\nu \rightarrow \infty$ by [44, Proposition 19]. \square

Combining Theorem 4.1 with Theorem 4.2, we have the following convergence result.

Theorem 4.3. *Suppose there exists $p_0 > 0$ such that for all $j \in \mathcal{J}$, $p_j(\xi) \leq_{a.s.} p_0$. Then,*

$$\limsup_{\epsilon \downarrow 0} \lim_{\nu \rightarrow \infty} x_\nu^\epsilon \subseteq S^*$$

w.p. 1, where S^ denotes the optimal solution set of the first-stage problem of (2.9).*

5. Numerical tests and applications on crude oil market

In this section, we firstly carry out numerical experiments using randomly generated data to illustrate the effectiveness of our model and its solution approach. Furthermore, we adopt our two-stage stochastic Nash equilibrium problem to describe market share competition in crude oil market. The results show that the model is capable of reproducing the actual oil market share based on in-sample data. Moreover, it is also shown that the model can make good out-of-sample predictions using historical data. All the tests are run in MATLAB 2016b on a personal computer with 32GB RAM and 8-core processor ($3.6 \times 8GHz$).

5.1. Progressive hedging method and smoothing Newton sub-algorithm

Firstly, randomly generated problems are used for testing our regularized SAA approach to solve the two-stage stochastic Nash equilibrium problem. Recall that the model of interests takes the form of a scenario-based linear complementarity problem (4.1) or its equivalent expression (4.2) with sufficiently small ϵ . The solution process adopts the well-known PHM. The PHM is globally convergent and the convergence rate is linear for problem (4.2) with SAA approach.

The classical PHM to solve (4.2) is as follows

Algorithm 1: Progressive hedging method

Step 0. Given an initial point $x^0 \in \mathbb{R}^J$, let $x_\ell^0 = x^0 \in \mathbb{R}^J$, $v_\ell^0 \in \mathbb{R}^{2J}$ and $w_\ell^0 \in \mathbb{R}^J$, for $\ell = 1, \dots, \nu$, such that $\frac{1}{\nu} \sum_{\ell=1}^{\nu} w_\ell^0 = 0$. Set the initial point $z^0 = (x^0, v_1^0, \dots, v_\nu^0)^T$. Choose a step size $r > 0$. Set $k = 0$.

Step 1. If the point z^k satisfies the condition

$$\|\min(z^k, H^\epsilon z^k + \bar{q})\| \leq 10^{-6},$$

output the solution z^k and terminate the algorithm; otherwise, go to Step 2.

Step 2. For $\ell = 1, \dots, \nu$, find $(\hat{x}_\ell^k, \hat{v}_\ell^k)$ that solves linear complementarity problems

$$\begin{aligned} 0 &\leq x_\ell \perp Cx_\ell + Bv_\ell + a + w_\ell^k + r(x_\ell - x_\ell^k) \geq 0, \\ 0 &\leq v_\ell \perp -B^T x_\ell + D_\ell^\epsilon v_\ell + q_\ell + r(v_\ell - v_\ell^k) \geq 0. \end{aligned} \quad (5.1)$$

Then let $\bar{x}^{k+1} = \frac{1}{\nu} \sum_{\ell=1}^{\nu} \hat{x}_\ell^k$, and for $\ell = 1, \dots, \nu$, update

$$x_\ell^{k+1} = \bar{x}^{k+1}, v_\ell^{k+1} = \hat{v}_\ell^k, w_\ell^{k+1} = w_\ell^k + r(\hat{x}_\ell^k - x_\ell^{k+1}),$$

to get point $z^{k+1} = (\bar{x}^{k+1}, v_1^{k+1}, \dots, v_\nu^{k+1})^T$.

Step 3. Set $k := k + 1$; go back to Step 1.

The PHM involves solving ν independently sample-based LCP (5.1) at each iteration, Problem (5.1) is well-defined, since for $\ell = 1, \dots, \nu$, the coefficient matrix

$$\begin{pmatrix} C + rI & B \\ -B^T & D_\ell^\epsilon + rI \end{pmatrix} \in \mathbb{R}^{3J \times 3J}$$

is positive definite for any $\epsilon > 0$. Thus, it has a unique solution for each $\ell = 1, \dots, \nu$. For simplicity, denote (5.1) as

$$0 \leq z_\ell \perp \tilde{H}_\ell^\epsilon z_\ell + \tilde{q}_\ell \geq 0, \quad (5.2)$$

where $z_\ell = \begin{pmatrix} x_\ell \\ v_\ell \end{pmatrix}$, $\tilde{H}_\ell^\epsilon = \begin{pmatrix} C + rI & B \\ -B^T & D_\ell^\epsilon + rI \end{pmatrix}$, $\tilde{q}_\ell = \begin{pmatrix} a + w_\ell^k - rx_\ell^k \\ q_\ell - rv_\ell^k \end{pmatrix}$ with $\ell = 1, 2, \dots, \nu$.

Then, we can equivalently write a large-scale LCP for all ν samples and solve

$$0 \leq \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_\nu \end{pmatrix} \perp \begin{pmatrix} \tilde{H}_1^\epsilon & & \\ & \tilde{H}_2^\epsilon & \\ & & \ddots \\ & & & \tilde{H}_\nu^\epsilon \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_\nu \end{pmatrix} + \begin{pmatrix} \tilde{q}_1 \\ \tilde{q}_2 \\ \vdots \\ \tilde{q}_\nu \end{pmatrix} \geq 0. \quad (5.3)$$

The structure of problem (5.3) enables us to use block computation to solve it, which can significantly improves the efficiency of the PHM. For example, suppose $\nu = mN$ with m, N being positive integer number. The equivalent m -block reformulation for (5.3) reads

$$0 \leq \begin{pmatrix} z_{i_1} \\ z_{i_2} \\ \vdots \\ z_{i_m} \end{pmatrix} \perp \begin{pmatrix} \tilde{H}_{i_1}^\epsilon & & \\ & \tilde{H}_{i_2}^\epsilon & \\ & & \ddots \\ & & & \tilde{H}_{i_m}^\epsilon \end{pmatrix} \begin{pmatrix} z_{i_1} \\ z_{i_2} \\ \vdots \\ z_{i_m} \end{pmatrix} + \begin{pmatrix} \tilde{q}_{i_1} \\ \tilde{q}_{i_2} \\ \vdots \\ \tilde{q}_{i_m} \end{pmatrix} \geq 0, \quad (5.4)$$

where $i_j = iN + j, j = 1, \dots, m, i = 0, \dots, (m-1)N$.

The main computation cost of the PHM is in **Step 2** due to the large sample size ν despite the relative low cost in solving for one sample. Block implementation speeds up the computation by better exploring and rebalancing the computational load. In practice, the number of blocks are adjusted so that the over all computation time is minimized. To improve the efficiency of the PHM, we also use the warm-start technique suggested in [33] to choose an initial point for subproblem (5.1). More specifically, the solution z^k of subproblem (5.1) at the k -th iteration is used as a starting point for the $(k+1)$ -th iteration.

In the remaining of this subsection, we focusing on problem (5.2) for a given sample and omit the subscription ℓ for ease of expression. In order to take advantage of the sparse structure of the subproblem (5.2), we apply the smoothing Newton method proposed by Chen and Ye [12] to solve it. In what follows, we give a brief description of the smoothing Newton method. It is well-known that solving (5.2) for a given sample is equivalent to solving the nonsmooth equation

$$F(z) = \min(z, \tilde{H}^\epsilon z + \tilde{q}) = 0. \quad (5.5)$$

The main idea of the smoothing Newton method is to use a smooth approximation function to approximate the nonsmooth function F and then solve the corresponding linear system. We use the smooth Gariel-Moré approximation function $f : \mathbb{R}^{3J} \times \mathbb{R}_{++} \rightarrow \mathbb{R}^{3J}$ to approximate the nonsmooth function F . In our numerical tests, the j -th component of f reads

$$f_j(z, \delta) = (z)_j - \frac{1}{2} \left(\sqrt{(\tilde{H}^\epsilon z + \tilde{q} - z)_j^2 + 4\delta^2} + (z - \tilde{H}^\epsilon z - \tilde{q})_j \right), \quad j = 1, \dots, 3J,$$

where the corresponding j -th diagonal element of the Jacobian $\bar{D}(z_i)$ is

$$\bar{D}_{jj} = \frac{1}{2} \left(\frac{(z - \tilde{H}^\epsilon z - \tilde{q})_j}{\sqrt{(z - \tilde{H}^\epsilon z - \tilde{q})_j^2 + 4\delta^2}} + 1 \right), \quad j = 1, \dots, 3J.$$

Then, the smoothing Newton method for solving subproblem (5.5) requires to solve a linear equation to determine d^k at each iteration, namely

$$\nabla_z f(z^k, \delta_k) d^k + F(z^k) = 0, \quad (5.6)$$

where δ_k decreases to 0 according to the criterion in article [12]. To guarantee the well-defineness of the (5.6), we make use of the following result.

Theorem 5.1. [15] *For any diagonal matrix $\tilde{D} = \text{diag}(\tilde{D}_{jj}) \in \mathbb{R}^{J \times J}$ with $0 \leq \tilde{D}_{jj} \leq 1, j = 1, 2, \dots, J$, the matrix $I - \tilde{D}(I - A)$ is nonsingular if and only if A is a P-matrix.* \square

For any given sample, it is known that H^ϵ is positive definite and hence a P-matrix. Moreover, the $\bar{D}(z)$ is a diagonal matrix with its element on the interval $[0, 1]$ for any $(z, \delta) \in \mathbb{R}^{3J} \times \mathbb{R}_{++}$. Therefore, using Theorem 5.1, the Jacobian $\nabla_z f(z, \delta)$ is nonsingular for any $(z, \delta) \in \mathbb{R}^{3J} \times \mathbb{R}_{++}$. Thus, the linear equation (5.6) is well-defined.

Denoting the matrix $\bar{D}(z) = \text{diag}(\bar{D}_1(z), \bar{D}_2(z), \bar{D}_3(z))$, the Jacobian $\nabla_z f(z, \delta)$ at the point z is of the following structure

$$\begin{aligned} \nabla_z f(z, \delta) &= (I - \bar{D}(z)) + \bar{D}(z) \tilde{H}^\epsilon \\ &\triangleq \begin{pmatrix} \Lambda_1(z) & 0 & -\bar{D}_1(z) \\ 0 & u_1(z)e^T + \Lambda_2(z) & \bar{D}_2(z) \\ \bar{D}_3(z) & -\bar{D}_3(z) & \Lambda_3(z) \end{pmatrix}, \end{aligned} \quad (5.7)$$

where $\Lambda_1(z) = \bar{D}_1(z)(C + (r-1)I) + I$, $\Lambda_2(z) = (\gamma(\xi) + (r-1))\bar{D}_2(z) + I$, $\Lambda_3(z) = (\epsilon + (r-1))\bar{D}_3(z) + I$ are all diagonal matrices, and $u_1(z) = \gamma(\xi)\bar{D}_2(z)e$. We can take advantage of the sparse structure of (5.7) in its inverse computation. Specifically, $\nabla_z f(z, \delta)$ consists of only diagonal sub-matrix and the matrix $u_1(z)e^T + \Lambda_2(z)$, where the later is a sum of a diagonal sub-matrix and a rank-one matrix.

Noticing from (5.7), linear equation (5.6) is of the following form

$$\begin{pmatrix} \Lambda_1 & 0 & \Lambda_2 \\ 0 & u_1 u_2^T + \Lambda_3 & \Lambda_4 \\ \Lambda_5 & \Lambda_6 & \Lambda_7 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \quad (5.8)$$

where $\Lambda_i \in \mathbb{R}^J, i = 1, 2, \dots, 7$ are diagonal matrices with Λ_1 and Λ_3 being nonsingular, $s_i \in \mathbb{R}^J, b_i \in \mathbb{R}^J, i = 1, 2, 3, u_1, u_2 \in \mathbb{R}^J$. For ease of notation, we use $\hat{\Lambda} = \text{diag}(1/\Lambda_{11}, \dots, 1/\Lambda_{JJ})$ to represent the inverse of any invertible diagonal matrix $\Lambda = \text{diag}(\Lambda_{11}, \dots, \Lambda_{JJ})$. Given the sparse structure of the coefficient matrix, we can solve (5.8) efficiently. More exactly, by the first two equations of (5.8), we can get

$$s_1 = \hat{\Lambda}_1(b_1 - \Lambda_2 s_3), \quad (5.9)$$

$$s_2 = (u_1 u_2^T + \Lambda_3)^{-1}(b_2 - \Lambda_4 s_3). \quad (5.10)$$

Directly substituting (5.9) and (5.10) into the third equation of (5.8), then we have

$$(\Lambda_7 - \Lambda_5 \hat{\Lambda}_1 \Lambda_2 - \Lambda_6(u_1 u_2^T + \Lambda_3)^{-1} \Lambda_4) s_3 = b_3 + \text{const}, \quad (5.11)$$

where $\text{const} = -(\Lambda_5 \hat{\Lambda}_1 b_1 + \Lambda_6(u_1 u_2^T + \Lambda_3)^{-1} b_2)$. For computing the inverse matrix of $(u_1 u_2^T + \Lambda_3)$, the Sherman-Morrison formula is useful. We have by the Sherman-Morrison formula

$$(u_1 u_2^T + \Lambda_3)^{-1} = \hat{\Lambda}_3 - \frac{\hat{\Lambda}_3 u_1 u_2^T \hat{\Lambda}_3}{1 + u_2^T \hat{\Lambda}_3 u_1}. \quad (5.12)$$

Substituting (5.12) into (5.11), we get

$$(\Lambda_0 + \alpha \tilde{u}_1 \tilde{u}_2^T) s_3 = b_3 + \text{const},$$

where $\alpha = 1/(1 + u_2^T \hat{\Lambda}_3 u_1)$, $\Lambda_0 = \Lambda_7 - \Lambda_5 \hat{\Lambda}_1 \Lambda_2 - \Lambda_6 \hat{\Lambda}_3 \Lambda_4$, $\tilde{u}_1 = \Lambda_6 \hat{\Lambda}_3 u_1$, $\tilde{u}_2 = \Lambda_4 \hat{\Lambda}_3 u_2$. Then, if Λ_0 is nonsingular, using the Sherman-Morrison formula again, we can immediately get the solution of s_3

$$s_3 = \left(\hat{\Lambda}_0 - \frac{\alpha \hat{\Lambda}_0 \tilde{u}_1 \tilde{u}_2^T \hat{\Lambda}_0}{1 + \alpha \tilde{u}_2^T \hat{\Lambda}_0 \tilde{u}_1} \right) (b_3 + \text{const}). \quad (5.13)$$

Then, substituting the s_3 into (5.9) and (5.10), we get the solution of s_1 and s_2 , respectively.

From (5.13), one can know that the computation cost of s_3 is trivial, since we only need to compute inverse of several diagonal matrix, namely, Λ_0, Λ_1 , and Λ_3 . Once s_3 is obtained, the calculation of s_1 and s_2 just needs to perform matrix-vector production. Therefore, the linear equation (5.8) can be solve efficiently.

5.2. Randomly generated problems

For the first part of numerical test, we randomly generated the problem of the form (4.2). More specifically, we generate a set of i.i.d. samples $\{\xi_\ell\}_{\ell=1}^\nu$ from a uniformly distribution over the interval $[0, 1]$. For $\ell = 1, \dots, \nu$, set

$$p(\xi_\ell) = ((\xi_\ell)_1, (\xi_\ell)_2, \dots, (\xi_\ell)_J)^T, \quad \Pi(\xi_\ell) = \gamma(\xi_\ell)(ee^T + I) \triangleq (\xi_\ell)_1(ee^T + I).$$

Diagonal matrix $C \in \mathbb{R}^{J \times J}$ and $a \in \mathbb{R}^J$ are generated with its elements uniformly distributed over the interval $[1, 2]$. All the numerical results are based on the average of 10 independent runs.

To show the feasibility of the solution of the regularized problem compared to that of the original problem, we compute the following residual value

$$\mathbf{Res} = \|\min(z, Hz + \bar{q})\|,$$

where H denotes the coefficient matrix (4.2) with $\epsilon = 0$.

Selected numerical results for $J = 10$ were listed in the Table 5.1 and 5.2. The first table is the solution obtained by PHM without block computation, while table 5.2 illustrates results with block implementation. For our generated experiments, we found $N = 50$ is the best block choice (measured by CPU time) for $J = 10$. The average number of iterations, the average cpu time, and the average value of \mathbf{Res} were recorded in both table. It is easy to see that the block implementation greatly reduces the cpu time. For the same value of ϵ , the number of iterations increases slightly when the sample size ν increases. In cases where the sample size ν is kept constant and the values of regularization parameter ϵ are chosen from $\epsilon = 10^{-3}$ to $\epsilon = 10^{-12}$, the iteration numbers are barely influenced as well as the cpu time. Furthermore, we observe the convergence of our regularization approach with decreasing values of ϵ , as have been proved in previous sections. Also notice that, the value of \mathbf{Res} decreases when the ϵ diminishes from 10^{-3} to 10^{-12} . Numerically, it shows that the solution of the regularized problem is also that of the original problem when $\epsilon = 10^{-12}$.

Figure 5.1 illustrates the performance of the PHM measured by the number of iterations and cpu time for 10 players. It is also worth mentioning that although one might expect the problem to be more difficult to solve for a small ϵ , the numerical performance in our experiments remain roughly unaffected with decreasing values of ϵ .

Figure 5.2 demonstrates the convergence property of the first-stage solution x when the sample size gets large for the case $J = 10$. The convergence trend can be seen component-wisely as the solution x converges when the sample size ν gets large.

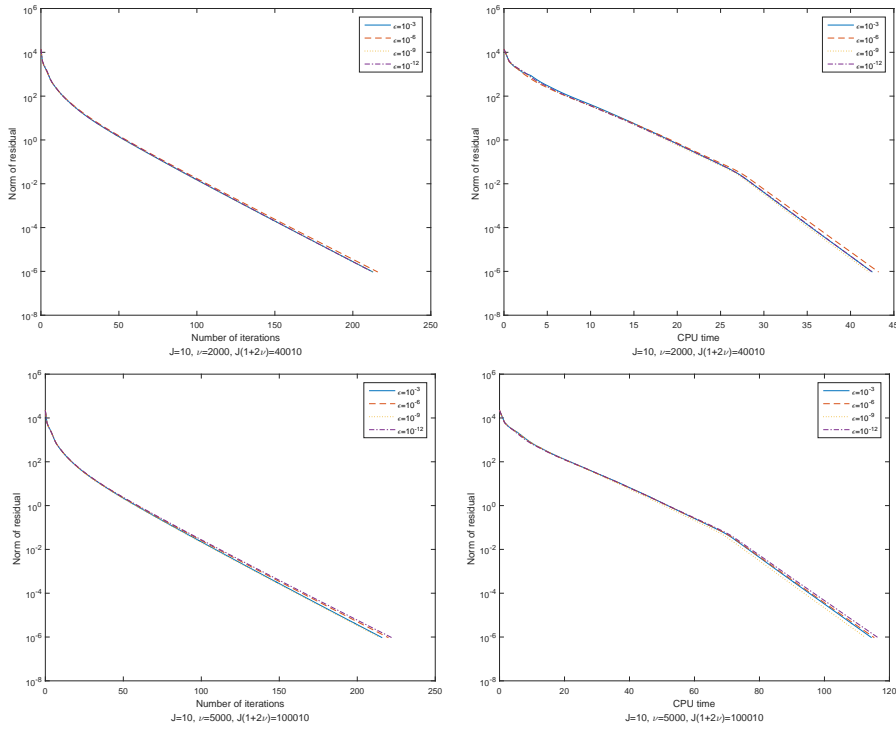
ν	$J(1 + 2\nu)$	Iter	CPU time/s	\mathbf{Res}	Iter	CPU time/s	\mathbf{Res}
				$\epsilon = 10^{-3}$			
10	210	146.30	0.26	4.42e-01	176.20	0.32	3.88e-04
50	1010	194.70	1.81	9.35e-01	197.40	1.83	9.32e-04
500	10010	208.70	26.72	3.00e+00	212.20	27.21	2.99e-03
2000	40010	222.60	154.97	5.93e+00	220.50	153.54	6.00e-03
5000	100010	224.70	623.53	9.49e+00	226.40	627.53	9.48e-03
				$\epsilon = 10^{-9}$			
10	210	152.70	0.27	1.08e-06	169.40	0.30	9.49e-07
50	1010	197.20	1.83	1.41e-06	194.40	1.80	9.75e-07
500	10010	212.70	27.21	3.21e-06	209.70	26.85	9.59e-07
2000	40010	220.30	153.34	6.16e-06	220.70	153.73	9.51e-07
5000	100010	226.70	628.89	9.60e-06	226.20	627.58	9.60e-07

Table 5.1: Numerical results for different ϵ and sample size ν , $J = 10$ with individual sample.

5.3. Crude oil market

In this subsection, we demonstrate one application of our two-stage stochastic Nash equilibrium model in describing the crude oil market share. Namely, we investigate the strategies of crude oil exporting agents via solutions of our reformulated SVI problem so as to recreate

ν	$J(1+2\nu)$	Iter	CPU time/s	Res	Iter	CPU time/s	Res
$\epsilon = 10^{-3}$							
10	210	162.60	0.21	4.04e-01	155.50	0.20	4.14e-04
50	1010	180.60	0.94	9.42e-01	184.00	0.93	9.31e-04
500	10010	203.30	10.67	2.97e+00	204.10	10.61	2.98e-03
2000	40010	213.80	42.38	5.89e+00	213.70	42.36	5.91e-03
5000	100010	218.20	115.15	9.36e+00	219.60	115.63	9.36e-03
$\epsilon = 10^{-9}$							
10	210	143.20	0.19	1.08e-06	149.50	0.19	9.59e-07
50	1010	197.40	1.02	1.38e-06	191.10	0.97	9.58e-07
500	10010	205.10	10.58	3.19e-06	202.40	10.67	9.52e-07
2000	40010	214.10	42.46	6.07e-06	213.40	42.62	9.73e-07
5000	100010	219.00	115.35	9.50e-06	219.70	115.58	9.74e-07
$\epsilon = 10^{-12}$							
10	210	143.20	0.19	1.08e-06	149.50	0.19	9.59e-07
50	1010	197.40	1.02	1.38e-06	191.10	0.97	9.58e-07
500	10010	205.10	10.58	3.19e-06	202.40	10.67	9.52e-07
2000	40010	214.10	42.46	6.07e-06	213.40	42.62	9.73e-07
5000	100010	219.00	115.35	9.50e-06	219.70	115.58	9.74e-07

Table 5.2: Numerical results for different ϵ and sample size ν , $J = 10$ with block implementation.Fig. 5.1. Numerical comparisons among different ϵ , $J = 10$.

actual market shares. Based on historical data on crude oil market, we make in-sample backtracking test to establish the effectiveness and validity of our model while explaining the market behaviour. Furthermore, the out-of-sample prediction capability of our model is demonstrated when training data is used to specify model parameters. From the results of our numerical tests, we conclude that our model is suitable to reproduce, predict and potentially capable to explain stable market shares of crude oil.

Crude oil market is one of the most widely studied commodity market in the world. The market has long been described as of an ‘intermediate between monopoly and perfect competition [22]’, supported by extensive historical data and market observation. When the market encounters large sudden events, it took “oil shocks” [17, 18, 25]. Conceptually, consequences of supply and demand fluctuation, along with many other factors, are reflected most directly in

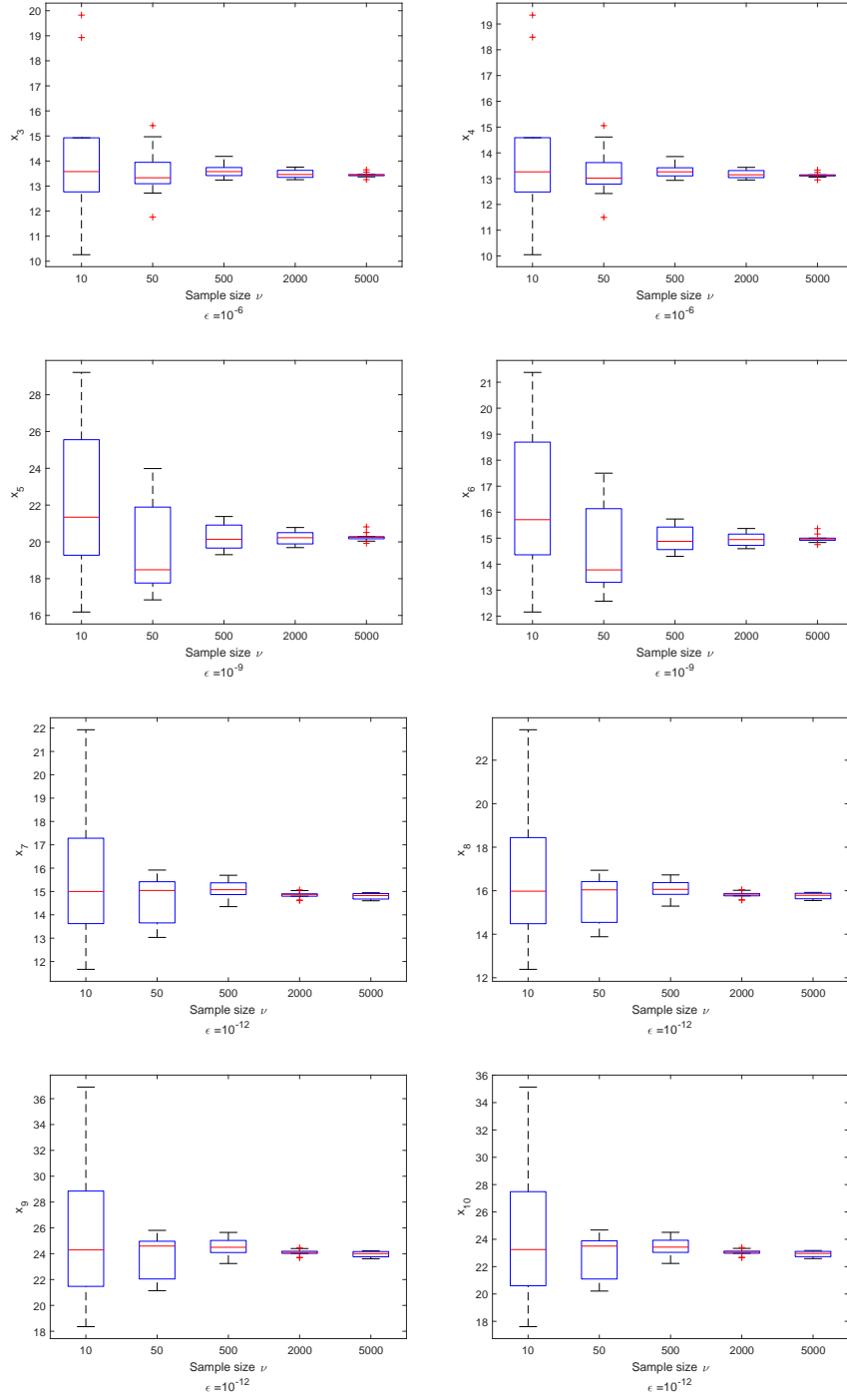


Fig. 5.2. Convergence property of x with increasing ν , $J = 10$.

dramatic changes in oil price. One interesting observation is that the market share behaves rather smoothly even during periods of oil shocks [25]. Majority of the world's crude oil is supplied by a few large oil exporting countries and they are viewed collectively as a finite number of large

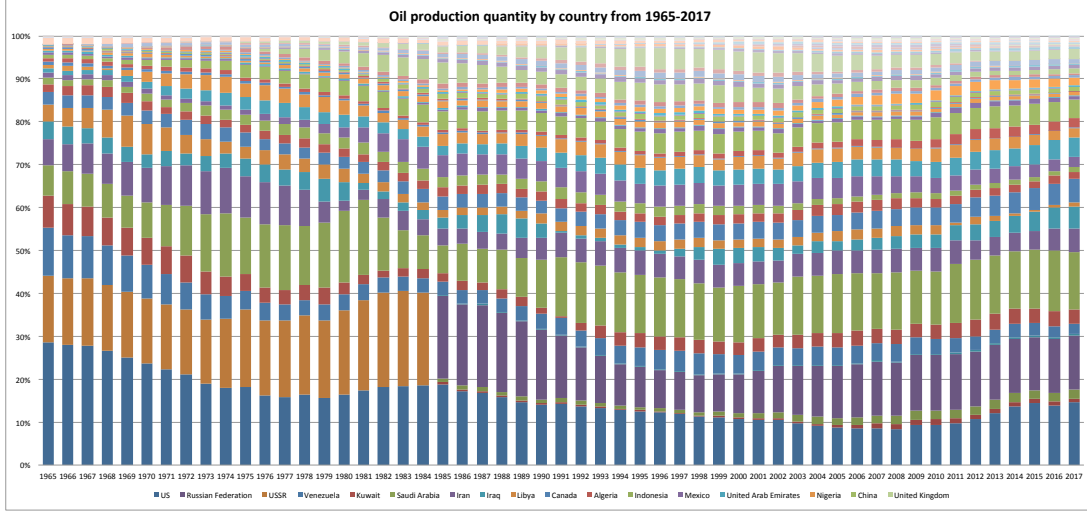


Fig. 5.3. Market shares of different oil-producing countries, 1965-2017.

agents from which price-taking consumers purchase product at the same price [45]. Hence, it is no surprise that Cournot-Nash approaches have been adopted from earlier days of study in this field, see [42]. The study of oil shock gives rise to rich field of economic researches in oil market structure and as well as factors affecting oil price, see [1, 26, 49], etc. Our model does not attempt to make explanation or future predictions on oil shocks but rather aims to make sense of the stable characteristics of supply-side of oil market share: major agents acting non-cooperatively to achieve market equilibrium. We focus on the supply factor of oil, while treat other factors, e.g., demand, world economic situation, population, etc., as known information with uncertainty. Our model is found to be able to reproduce the market shares as well as to forecast future production plans of different oil-producing countries based on historical data. In short, we firstly use historical data to determine model parameters as well as to approximate distribution for uncertainty in observations of the market. Model parameters are tested within a fixed sampling window so that the in-sample results obtained by solving our model matches that of training data set. It then follows that over a short decision horizon in the future, we can adopt the trained model to prediction future production plans. More specifically, we took i.i.d. samples using a fixed-size rolling window sampling method while the prediction is obtained by solving our model with the adjusted set of parameters obtained by in-sample tests within the training window.

5.3.1. Data set, parameter selection and uncertainty description

In order to implement real market data into our model (2.6)-(2.7), we need to determine parameters c_j, a_j correspond to production costs of j -th agent by analyzing training set. In practice, model parameters are adjusted via a brute force learning process. More specifically, if the training set consists only one trading day of information, we adjust the parameters so that the solution, i.e., production quantities of J agents, matches the estimated historical observation. When more data become available, the parameters are adjusted so that the average of solutions is close to the real estimate. It is found from our numerical implementation that the best fit solutions are obtained when values of c_j and a_j are taken to be inversely proportional to known market share within the sampling period. It is firstly learned that the common belief of “lower

unit cost would lead to large sale” does not hold in our model, which also motivates the choice of quadratic production cost in our model. Moreover, it is found in our numerical experiment that the smaller the quadratic cost parameters c_j the greater market share the agent would have. This can be understood from the fact that the large oil producing agents are more flexible in adjusting their production quantities as compare to smaller agents. A more mathematical interpretation is to think of c_j as related to the penalty parameters of the augmented Lagrangian formulation corresponding to production constraints of each agent: production cost is upper bounded with some unit production cost \tilde{a}_j . In this way, one can think of this observation been that the countries has greater reserve of oil suffers less when the production upper bound gets violated. Nevertheless, we found that the quadratic is more than adequate to reproduce historical observations.

The scenarios of stochastic parameters are taken from empirical distribution of the sampling data set, guided by the analysis of [19]. In detail, we draw samples to generate scenarios of market’s supply discounting factor $\gamma(\xi)$ and j -th agent’s adjusted price $p_j(\xi) = p_0(\xi) - h_j(\xi)$. The data set used in this study are from (i) *Statistical Review of World Energy*¹⁾, yearly data from 1984, published every June by bp. Inc; (ii) *Oil Price Dynamics Report*²⁾, weekly by Federal Reserve Bank of New York since 1986; (iii) *U.S. Energy Information Administration*³⁾, weekly and daily spot price of Brent. The empirical distribution are generated from the contents of oil dynamic report, consists of time series of weekly percentage change of Brent spot price and its corresponding components’ contributions, i.e., those from demand, residual, and supply. More specifically, the percentage change in price in $k + 1$ -th⁴⁾ week compared to that of the k -th week is denoted by Δp^k , while the contribution of demand and supply changes are denoted as ΔD^k and ΔS^k respectively. The contribution to the price change that does not correspond to demand and supply is denoted as residual ΔR^k . Therefore, for any week within the report, it holds that the Brent price change is deterministically represented by

$$\Delta p^k = \Delta D^k + \Delta T^k + \Delta R^k. \quad (5.14)$$

Recall that our expression for (scenario-based) inverse demand function of quantity offered to the market for a given scenario ξ^ℓ is in the following form

$$p^k(\xi_\ell) - \gamma^k(\xi_\ell)T^k(y_{\xi_\ell}).$$

In accordance with data structure, $p^k(\xi_\ell)$ corresponds to the adjusted predicted price from contributions other than supply, i.e.,

$$p^k(\xi_\ell) = p_0^k(1 + \Delta d^k(\xi_\ell) + \Delta r^k(\xi_\ell)),$$

where p_0^k denotes the known Brent spot price prior to that of the concerned time. Both $\Delta d^k(\xi_\ell)$ and $\Delta r^k(\xi_\ell)$ are random scenarios taken from empirical distributions of historical demand and residual distributions within the sampling set respectively. To estimate $\gamma^k(\xi)$, corresponding to k -week data, we need the data of spot quantity supplied to the market, T^k . In practice, it is very difficult to obtain reliable data on total supply to the market over short observation window. Rather, we observe that trust-worthy estimate on daily oil supply based on annual

¹⁾ <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html>

²⁾ https://www.newyorkfed.org/research/policy/oil_price_dynamics_report

³⁾ <https://www.eia.gov>

⁴⁾ k is used to denote the order in time, distinguished from that of i.i.d. samples

data is available along with the observation that a steady growth of about 1% per year over the last four decades, according to *Statistical Review of World Energy*. Therefore, we take random instances of daily market supply $T^k(y_\xi)$ which are taken from a uniformly distributed interval between 99% and 101% of yearly based daily estimate. Then, we can generate a set of data of stochastic excessive supply discount factor $\{\gamma^k(\xi_\ell)\}$,

$$\gamma^k(\xi_\ell) = \frac{|p^k(\xi_\ell) - p_0^k|}{T^k(y_\xi)},$$

where p_0^k is known with certainty within the testing data set, and the $|\cdot|$ ensures that increase in quantity has a negative influence on price.

We present the results for our numerical experiments on reproduced oil market share. As shown in Fig 5.4, the in-sample results over the periods of oil shocks 2007, 2009 and 2014 are reproduced using solely from the available information within that year. For example, the in-sample experiment of 2009 uses all the available data obtained within 2009, e.g., daily price from 01/01/2009 to 31/12/2009. For the generation of random instances in-sample, we use weekly price contribution covering the year 2009 to form empirical distributions for supply excluded spot prices $p(\xi_k)$, and discounting factor $\gamma(\xi_k)$, from which we took i.i.d. samples to represent random scenarios. Hence, for in-sample experiments, our goal is to reproduce known market share with the adjustments of cost parameters a_j, c_j corresponding to j -th oil producing agent.

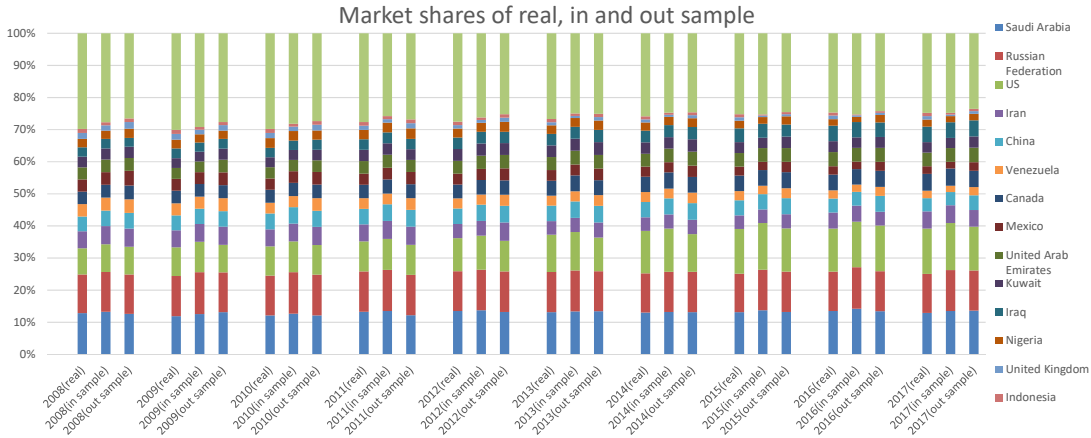


Fig. 5.4. Real, in-sample and out-of-sample market shares results, 2008-2017.

The out-of-sample tests cover the period from 2007 to 2017 with the *one year length equivalent of sampling window* in Figure 5.4. Note that one year sampling window is adopted since independent numerical results of selected years demonstrated that the market has a short memory in the sense that longer historical data would not provide extra information, at least in the setup of our model and approaches. Prior to the date of interests, all the historical data are assumed to be available while the rolling window sampling permits the usage of newly acquired data as soon as it becomes available. From the construction and interpretation of our two-stage model, it is reasonable to assume that cost of oil production remain unchanged over a short decision horizon. Furthermore, we tested different lengths of the sampling window, especially during times of oil shocks for better in-sample fitting, and the best fit sampling window length is used in out-of-sample tests. For example, we are interested in obtaining the market share

of 2009, during times of recent financial crisis, using the fixed length sampling windows whose length varies from one year to five years. If we want to predict the production on 01/01/2009 using a sampling window of one year, we took one year sampling window from 01/01/2008 to 31/01/2008 to obtain empirical distributions from which scenarios of $\gamma(\xi_k)$, $p(\xi_k)$ are sampled. Taking ν samples, the solutions method follows that described in previous subsection. For the result on 02/01/2019, the sampling window rolls forward using the data from 02/01/2008 to 01/01/2009 since the last day information becomes known, while the model remains unchanged. Therefore, for the market share prediction of any given year, we obtained the daily results by taking average of 250 samples, while the yearly estimate is formed by taking the average on daily results. From Figure 5.4, it is observed that we can obtain fair good prediction quantities through out the testing period. However, it is worth mentioning that during the numerical experiments across that decision horizon the computational cost is much higher during periods over oil shocks compare to those of relative stable oil prices.

6. Conclusion

A two-stage SVI is used to describe the equilibrium of a convex two-stage non-cooperative multi-agent game under uncertainty. Sufficient conditions for the existence of solutions of the two-stage SVI are provided. The numerical implementation is constructed by proposing a regularized SAA method. Furthermore, we prove the convergence of the regularized SAA method as the regularization parameter tends to zero and the sample size tends to infinity. A real application of crude oil market equilibrium is studied using our approach. More specifically, the oligopolistic market is characterized by a two-stage stochastic production and supply planning game and the market share is achieved at the equilibrium. Numerical results are presented based on historical data over the last fifty years and the effectiveness of our two-stage SVI approach of market share is demonstrated.

References

- [1] G. Bornstein, P. Krusell and S. Rebelo, Lags, costs, and shocks: An equilibrium model of the oil industry, *National Bureau of Economic Research*, 2017.
- [2] J.V. Burke, X. Chen and H. Sun, Subdifferentiation and smoothing of nonsmooth integral functionals, Department of applied mathematics, The Hong Kong Polytechnic University, 2017.
- [3] X. Chen, Smoothing methods for nonsmooth, nonconvex minimization, *Math. Program.*, **134** (2012), 71-99.
- [4] X. Chen and M. Fukushima, A smoothing method for a mathematical program with P-matrix linear complementarity constraints, *Comput. Optim. Appl.*, **27** (2004), 223-246.
- [5] X. Chen and M. Fukushima, Expected residual minimization method for stochastic linear complementarity problems, *Math. Oper. Res.*, **30** (2005), 1022-1038.
- [6] X. Chen, T.K. Pong and R.J.B. Wets, Two-stage stochastic variational inequalities: an ERM-solution procedure, *Math. Program.*, **165** (2017), 71-111.
- [7] X. Chen, H. Sun and R.J.B. Wets, Regularized mathematical programs with stochastic equilibrium constraints: estimating structural demand models, *SIAM J. Optim.*, **25** (2015), 53-75.
- [8] X. Chen, A. Shapiro and H. Sun, Convergence analysis of sample average approximation of two-state stochastic generalized equations, *SIAM J. Optim.*, to appear.
- [9] X. Chen, H. Sun and H. Xu, Discrete approximation of two-stage stochastic and distributionally robust linear complementarity problems, *Math. Program.*, (2018), to appear.

- [10] X. Chen, R.J.B. Wets and Y. Zhang, Stochastic variational inequalities: residual minimization smoothing sample average approximations, *SIAM J. Optim.*, **22** (2012), 649-673.
- [11] X. Chen and S. Xiang, Newton iterations in implicit time-stepping scheme for differential linear complementarity systems, *Math. Program.*, **138** (2013), 579-606.
- [12] X. Chen and Y. Ye, On homotopy-smoothing methods for box-constrained variational inequalities, *SIAM J. Control. Optim.*, **37** (1999), 589-616.
- [13] X. Chen and Y. Ye, On smoothing methods for the P_0 matrix linear complementarity problem, *SIAM J. Optim.*, **11** (2000), 341-363.
- [14] F. Facchinei and J.S. Pang, *Finite-dimensional variational inequalities and complementarity problems*, Springer Science & Business Media, 2007.
- [15] S.A. Gabriel and J.J. More, Smoothing of mixed complementarity problems, in *Complementarity and Variational Problems: State of the Art*, M.C. Ferris and J.S. Pang, eds., SIAM, (1997), 105-116.
- [16] T.S. Genc, S.S. Reynolds and S. Sen, Dynamic oligopolistic games under uncertainty: A stochastic programming approach, *J. Econ. Dyn. Control.*, **31** (2007), 55-80.
- [17] J.D. Hamilton, This is what happened to the oil price-macroeconomy relationship, *J. Monet. Econ.*, **38** (1996), 215-220.
- [18] J.D. Hamilton, What is an oil shock?, *J. Econom.*, **113**:2 (2003), 363-398.
- [19] J.D. Hamilton, Historical oil shocks, *National Bureau of Economic Research*, 2011.
- [20] P.T. Harker and J.S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications, *Math. Program.*, **48** (1990), 161-200.
- [21] B.F. Hobbs and J.S. Pang, Nash-Cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints, *Oper. Res.*, **55** (2007), 113-127.
- [22] H. Hotelling, The economics of exhaustible resources, *J. Polit. Econ.*, **39** (1931), 137-175.
- [23] J. Huang, M. Leng and M. Parlar, Demand functions in decision modeling: A comprehensive survey and research directions, *Decis. Sci.*, **44** (2013), 557-609.
- [24] A. Jofré, R.T. Rockafellar and R.J.B. Wets, Variational inequalities and economic equilibrium, *Math. Oper. Res.*, **32** (2007), 32-50.
- [25] L. Kilian, Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market, *Am. Econ. Rev.*, **99** (2009), 1053-1069.
- [26] L. Kilian and D.P. Murphy, The role of inventories and speculative trading in the global market for crude oil, *J. Appl. Econom.*, **29** (2014), 454-478.
- [27] F.H. Murphy, H.D. Sherali and A.L. Soyster, A mathematical programming approach for determining oligopolistic market equilibrium, *Math. Program.*, **24** (1982), 92-105.
- [28] J.S. Pang, S. Sen and U.V. Shanbhag, Two-stage non-cooperative games with risk-averse players, *Math. Program.*, **165** (2017), 235-290.
- [29] J.S. Pang, C.L. Su, Y.C. Lee, A constructive approach to estimating pure characteristics demand models with pricing, *Oper. Res.*, **63** (2015), 639-659.
- [30] A. Philpott, M.C. Ferris and R.J.B. Wets, Equilibrium, uncertainty and risk in hydro-thermal electricity systems, *Math. Program.*, **157** (2016), 483-513.
- [31] J. Outrata, M.C. Ferris, M. Červinka and M. Outrata, On Cournot-Nash-Walras equilibria and their computation, *Set-Valued Var. Anal.*, **24** (2016), 387-402.
- [32] C.W. Richard, J.S. Pang and S.E. Richard, *The linear complementarity problem*, Academic Press, 1992.
- [33] R.T. Rockafellar and J. Sun, Solving monotone stochastic variational inequalities and complementarity problems by progressive hedging, *Math. Program.*, (2018), to appear.
- [34] R.T. Rockafellar and R.J.B. Wets, Stochastic convex programming: Kuhn-Tucker conditions, *J. Math. Econ.*, **2** (1975), 349-370.
- [35] R.T. Rockafellar and R.J.B. Wets, Stochastic convex programming: Basic duality, *Pac. J. Math.*,

- 62** (1976), 173-195.
- [36] R.T. Rockafellar and R.J.B. Wets, Stochastic convex programming: Singular multipliers and extended duality singular multipliers and duality, *Pac. J. Math.*, **62** (1976), 173-195.
 - [37] R.T. Rockafellar and R.J.B. Wets, Stochastic convex programming: Relatively complete recourse and induced feasibility, *SIAM. J. Control. Optim.*, **14** (1976), 574-589.
 - [38] R.T. Rockafellar and R.J.B. Wets, Scenarios and policy aggregation in optimization under uncertainty, *Math. Oper. Res.*, **16** (1991), 119-147.
 - [39] R.T. Rockafellar and R.J.B. Wets, *Variational analysis*, Springer Science & Business Media, 2009.
 - [40] R.T. Rockafellar and R.J.B. Wets, Stochastic variational inequalities: single-stage to multistage, *Math. Program.*, **165**, (2017), 331-360.
 - [41] B.J. Rosen, Existence and uniqueness of equilibrium points for concave n-person games, *Econometrica*, (1965), 520-534.
 - [42] S.W. Salant, Exhaustible resources and industrial structure: A Nash-Cournot approach to the world oil market, *J. Polit. Econ.*, **84** (1976), 1079-1093.
 - [43] A. Shapiro, D. Dentcheva and A. Ruszczyński, *Lectures on stochastic programming: modeling and theory*, SIAM, 2009.
 - [44] A. Shapiro, Monte Carlo sampling methods, *Handbooks in operations research and management science*, **10** (2003), 353-425.
 - [45] J.L. Smith, World oil: market or mayhem?, *J. Econ. Perspect.*, **23** (2009), 145-164.
 - [46] H. Sun, C.L. Su and X. Chen, SAA-regularized methods for multiproduct price optimization under the pure characteristics demand model, *Math. Program.*, **165** (2017), 361-389.
 - [47] H. Xu, J.S. Pang, F. Ordóñez and M. Dessouky, Complementarity models for traffic equilibrium with ridesharing, *Transport. Res. B-meth.*, **81** (2015), 161-182.
 - [48] J. Yao, I. Adler and S.S. Oren, Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network, *Oper. Res.*, **56** (2008), 34-47.
 - [49] L. Juvenal and I. Petrella, Speculation in the oil market, *J. Appl. Econom.*, **30** (2015), 621-649.

A. Appendix

Proof of Theorem 3.1.

Proof. By direct computation, we have that

$$\begin{aligned} & \left(M^\epsilon(\xi) \begin{pmatrix} y^\epsilon \\ \lambda^\epsilon \end{pmatrix} + q(x, \xi) \right)_j \\ &= \begin{cases} \gamma(\xi)(y^\epsilon)_j + \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi), & j = 1, \dots, J; \\ x_{j-J} - (y^\epsilon)_{j-J} + \epsilon(\lambda^\epsilon)_{j-J}, & j = J+1, \dots, 2J. \end{cases} \end{aligned}$$

Then, we can rewrite problem (3.2) as below:

$$\begin{cases} 0 \leq (y^\epsilon)_j \perp \gamma(\xi)(y^\epsilon)_j + \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) \geq 0, \\ 0 \leq (\lambda^\epsilon)_j \perp x_j - (y^\epsilon)_j + \epsilon(\lambda^\epsilon)_j \geq 0, \end{cases} \quad (\text{A.1})$$

for $j \in \mathcal{J}$. From the first complementarity condition in (A.1), we have $(y^\epsilon)_j$ as follows:

$$(y^\epsilon)_j = \begin{cases} -\frac{\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi)}{\gamma(\xi)}, & \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0; \\ 0, & \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) \geq 0 \end{cases} \quad (\text{A.2})$$

for $j \in \mathcal{J}$. Similarly, we can derive that

$$(\lambda^\epsilon)_j = \begin{cases} \frac{(y^\epsilon)_j - x_j}{\epsilon}, & (y^\epsilon)_j - x_j > 0; \\ 0, & (y^\epsilon)_j - x_j \leq 0 \end{cases} \quad (\text{A.3})$$

for $j \in \mathcal{J}$. Note that $(y^\epsilon)_j = 0$ implies $(y^\epsilon)_j = 0 \leq x_j$, and we have $(\lambda^\epsilon)_j = 0$. Then, based on (A.2) and (A.3), we have for all three cases:

$$\begin{cases} (y^\epsilon)_j = 0, (\lambda^\epsilon)_j = 0 & \text{for } j \in \mathcal{I}_1; \\ (y^\epsilon)_j = -\frac{\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi)}{\gamma(\xi)}, (\lambda^\epsilon)_j = 0 & \text{for } j \in \mathcal{I}_2; \\ (y^\epsilon)_j = -\frac{\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi)}{\gamma(\xi)}, (\lambda^\epsilon)_j = \frac{(y^\epsilon)_j - x_j}{\epsilon} & \text{for } j \in \mathcal{I}_3, \end{cases}$$

where

$$\begin{aligned} \mathcal{I}_1 &:= \{j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) \geq 0, (y^\epsilon)_j - x_j \leq 0\}, \\ \mathcal{I}_2 &:= \{j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, (y^\epsilon)_j - x_j \leq 0\}, \\ \mathcal{I}_3 &:= \{j \in \mathcal{J} : \gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, (y^\epsilon)_j - x_j > 0\}. \end{aligned}$$

It follows that,

$$((y^\epsilon)_j, (\lambda^\epsilon)_j) = \begin{cases} (0, 0), & j \in \mathcal{I}_1; \\ \left(-\frac{\gamma(\xi)T^\epsilon - p_j(\xi)}{\gamma(\xi)}, 0\right), & j \in \mathcal{I}_2; \\ \left(-\frac{\epsilon\gamma(\xi)T^\epsilon - x_j - \epsilon p_j(\xi)}{\epsilon\gamma(\xi) + 1}, -\frac{\gamma(\xi)(T^\epsilon + x_j) - p_j(\xi)}{\epsilon\gamma(\xi) + 1}\right), & j \in \mathcal{I}_3, \end{cases}$$

which verifies (3.5). For the remaining of the proof, let $j \in \mathcal{I}_2$, we have

$$-\gamma(\xi)(y^\epsilon)_j = \gamma(\xi)T^\epsilon - p_j(\xi)$$

and thus

$$-\gamma(\xi) \sum_{i \in \mathcal{I}_2} (y^\epsilon)_i = |\mathcal{I}_2| \gamma(\xi) T^\epsilon - \sum_{i \in \mathcal{I}_2} p_i(\xi). \quad (\text{A.4})$$

Analogously, we can derive from

$$(y^\epsilon)_j = -\frac{\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi)}{\gamma(\xi)} \text{ and } (\lambda^\epsilon)_j = \frac{(y^\epsilon)_j - x_j}{\epsilon} \text{ for } j \in \mathcal{I}_3$$

that

$$-\gamma(\xi) \sum_{i \in \mathcal{I}_3} (y^\epsilon)_i = |\mathcal{I}_3| \gamma(\xi) T^\epsilon + \frac{1}{\epsilon} \sum_{i \in \mathcal{I}_3} (y^\epsilon)_i - \frac{1}{\epsilon} \sum_{i \in \mathcal{I}_3} x_i - \sum_{i \in \mathcal{I}_3} p_i(\xi). \quad (\text{A.5})$$

Combining that of (A.4) and (A.5), we obtain

$$-\gamma(\xi)T^\epsilon = (|\mathcal{I}_2| + |\mathcal{I}_3|)\gamma(\xi)T^\epsilon + \frac{1}{\epsilon} \sum_{i \in \mathcal{I}_3} (y^\epsilon)_i - \frac{1}{\epsilon} \sum_{i \in \mathcal{I}_3} x_i - \sum_{i \in \mathcal{I}_2 \cup \mathcal{I}_3} p_i(\xi).$$

Therefore, we have

$$\frac{1}{\epsilon} \sum_{i \in \mathcal{I}_3} (y^\epsilon)_i = -(|\mathcal{I}_2| + |\mathcal{I}_3| + 1) \gamma(\xi) T^\epsilon + \frac{1}{\epsilon} \sum_{i \in \mathcal{I}_3} x_i + \sum_{i \in \mathcal{I}_2 \cup \mathcal{I}_3} p_i(\xi). \quad (\text{A.6})$$

Substituting (A.6) into (A.5), we have derivation

$$\begin{aligned} |\mathcal{I}_3| \gamma(\xi) T^\epsilon &= - \left(\gamma(\xi) + \frac{1}{\epsilon} \right) \sum_{i \in \mathcal{I}_3} (y^\epsilon)_i + \frac{1}{\epsilon} \sum_{i \in \mathcal{I}_3} x_i + \sum_{i \in \mathcal{I}_3} p_i(\xi) \\ &= -(\epsilon \gamma(\xi) + 1) \left(-(|\mathcal{I}_2| + |\mathcal{I}_3| + 1) \gamma(\xi) T^\epsilon + \frac{1}{\epsilon} \sum_{i \in \mathcal{I}_3} x_i + \sum_{i \in \mathcal{I}_2 \cup \mathcal{I}_3} p_i(\xi) \right) \\ &\quad + \frac{1}{\epsilon} \sum_{i \in \mathcal{I}_3} x_i + \sum_{i \in \mathcal{I}_3} p_i(\xi) \\ &= (\epsilon \gamma(\xi) + 1) (|\mathcal{I}_2| + |\mathcal{I}_3| + 1) \gamma(\xi) T^\epsilon - \gamma(\xi) \sum_{i \in \mathcal{I}_3} x_i - \epsilon \gamma(\xi) \sum_{i \in \mathcal{I}_2 \cup \mathcal{I}_3} p_i(\xi) - \sum_{i \in \mathcal{I}_2} p_i(\xi). \end{aligned}$$

Then, we get

$$(\epsilon \gamma(\xi) (|\mathcal{I}_2| + |\mathcal{I}_3| + 1) + |\mathcal{I}_2| + 1) \gamma(\xi) T^\epsilon = \gamma(\xi) \sum_{i \in \mathcal{I}_3} x_i + \epsilon \gamma(\xi) \sum_{i \in \mathcal{I}_2 \cup \mathcal{I}_3} p_i(\xi) + \sum_{i \in \mathcal{I}_2} p_i(\xi),$$

that is,

$$T^\epsilon = \frac{\gamma(\xi) \sum_{i \in \mathcal{I}_3} x_i + \epsilon \gamma(\xi) \sum_{i \in \mathcal{I}_2 \cup \mathcal{I}_3} p_i(\xi) + \sum_{i \in \mathcal{I}_2} p_i(\xi)}{(\epsilon \gamma(\xi) (|\mathcal{I}_2| + |\mathcal{I}_3| + 1) + |\mathcal{I}_2| + 1) \gamma(\xi)}.$$

This completes the proof. \square

Proof of Proposition 3.4.

Proof. Let $\hat{z} = (\hat{y}, \hat{\lambda})$ be any solution of $\text{LCP}(q, M)$ and we have the derivation:

$$\begin{aligned} 0 &\geq (z^\epsilon - \hat{z})^T (M^\epsilon z^\epsilon + q - (M \hat{z} + q)) \\ &= (z^\epsilon - \hat{z})^T (M^\epsilon z^\epsilon - M \hat{z}) \\ &= (z^\epsilon - \hat{z})^T M (z^\epsilon - \hat{z}) + (z^\epsilon - \hat{z})^T \begin{pmatrix} 0 \\ \epsilon \lambda^\epsilon \end{pmatrix} \\ &\geq (z^\epsilon - \hat{z})^T \begin{pmatrix} 0 \\ \epsilon \lambda^\epsilon \end{pmatrix} \\ &= \epsilon (\lambda^\epsilon - \hat{\lambda})^T \lambda^\epsilon, \end{aligned}$$

where the second inequality follows from the positive semidefiniteness of M . Then, we have

$$\|\lambda^\epsilon\|^2 \leq \hat{\lambda}^T \lambda^\epsilon \leq \|\hat{\lambda}\| \|\lambda^\epsilon\|,$$

which implies the boundedness of λ^ϵ ,

$$\|\lambda^\epsilon\| \leq \|\hat{\lambda}\|. \quad (\text{A.7})$$

It follows from (A.7) that any accumulation point of $\{\lambda^\epsilon\}$ as $\epsilon \downarrow 0$ is the least l_2 -norm solution. Since M is positive semidefinite, we know from [32, Theorem 5.6.2] that there is a unique least

l_2 -norm solution. On the other hand, we know from Proposition 3.1, for any fixed (x, ξ) , \hat{y} is unique. Therefore, the limit of z^ϵ exists as $\epsilon \downarrow 0$ and converges to the least l_2 -norm solution of LCP(q, M).

Due to the existence of limit for z^ϵ as $\epsilon \downarrow 0$, (3.7) can be derived directly from (3.5). In what follows, we focus on deriving the expression (3.8). To this end, for each $j \in \mathcal{J}$, three cases are discussed:

$$\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) \geq 0, \quad (y^\epsilon)_j - x_j \leq 0, \quad (\text{A.8})$$

$$\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, \quad (y^\epsilon)_j - x_j \leq 0, \quad (\text{A.9})$$

$$\gamma(\xi)T^\epsilon + (\lambda^\epsilon)_j - p_j(\xi) < 0, \quad (y^\epsilon)_j - x_j > 0. \quad (\text{A.10})$$

Case 1: If there exists a sequence $\{\epsilon_k\}_{k=1}^\infty$ converging to 0 such that (A.8) holds, we have

$$\lim_{k \rightarrow \infty} ((y^{\epsilon_k})_j, (\lambda^{\epsilon_k})_j) = (0, 0).$$

Thus, $|(\lambda^{\epsilon_k})_j - \bar{\lambda}_j| = 0$.

Case 2: If there exists a sequence $\{\epsilon_k\}_{k=1}^\infty$ converging to 0 such that (A.9) holds, we have an estimation

$$\begin{aligned} \lim_{k \rightarrow \infty} ((y^{\epsilon_k})_j, (\lambda^{\epsilon_k})_j) &= \lim_{k \rightarrow \infty} \left(-\frac{\gamma(\xi)T^{\epsilon_k} - p_j(\xi)}{\gamma(\xi)}, 0 \right) \\ &= \left(-\frac{\gamma(\xi) \lim_{k \rightarrow \infty} T^{\epsilon_k} - p_j(\xi)}{\gamma(\xi)}, 0 \right) \\ &= \left(-\frac{\gamma(\xi)\bar{T} - p_j(\xi)}{\gamma(\xi)}, 0 \right). \end{aligned}$$

Thus, $|(\lambda^{\epsilon_k})_j - \bar{\lambda}_j| = 0$.

Case 3: If there exists a sequence $\{\epsilon_k\}_{k=1}^\infty$ converging to 0 such that (A.10) holds, we have

$$\begin{aligned} \lim_{k \rightarrow \infty} ((y^{\epsilon_k})_j, (\lambda^{\epsilon_k})_j) &= \lim_{k \rightarrow \infty} \left(-\frac{\epsilon_k \gamma(\xi) T^{\epsilon_k} - x_j - \epsilon_k p_j(\xi)}{\epsilon_k \gamma(\xi) + 1}, -\frac{\gamma(\xi)(T^{\epsilon_k} + x_j) - p_j(\xi)}{\epsilon_k \gamma(\xi) + 1} \right) \\ &= (x_j, -\gamma(\xi)(\bar{T} + x_j) + p_j(\xi)). \end{aligned}$$

Thus, we have

$$\begin{aligned} &|(\lambda^{\epsilon_k})_j - \bar{\lambda}_j| \\ &= |(\lambda^{\epsilon_k})_j + \gamma(\xi)(\bar{T} + x_j) - p_j(\xi)| \\ &= \left| -\frac{\gamma(\xi)(T^{\epsilon_k} + x_j) - p_j(\xi)}{\epsilon_k \gamma(\xi) + 1} + \gamma(\xi)(\bar{T} + x_j) - p_j(\xi) \right| \\ &= \frac{|-\gamma(\xi)(T^{\epsilon_k} + x_j) + p_j(\xi) + \gamma(\xi)(\bar{T} + x_j) - p_j(\xi) + \epsilon_k \gamma(\xi)(\gamma(\xi)(\bar{T} + x_j) - p_j(\xi))|}{\epsilon_k \gamma(\xi) + 1} \\ &\leq \frac{\gamma(\xi) |T^{\epsilon_k} - \bar{T}| + |\gamma(\xi)(\gamma(\xi)(\bar{T} + x_j) - p_j(\xi))| \epsilon_k}{\epsilon_k \gamma(\xi) + 1}. \end{aligned}$$

Collectively, we know from **Case 1**, **Case 2** and **Case 3** that

$$(y^{\epsilon_k})_j - \bar{y}_j = 0, \quad (\text{A.11})$$

$$(y^{\epsilon_k})_j - \bar{y}_j = -(T^{\epsilon_k} - \bar{T}), \quad (\text{A.12})$$

$$(y^{\epsilon_k})_j - \bar{y}_j = \frac{-\gamma(\xi)T^{\epsilon_k} + p_j(\xi) - \gamma(\xi)x_j}{\epsilon_k \gamma(\xi) + 1} \cdot \epsilon_k. \quad (\text{A.13})$$

Furthermore, we have that $T^{\epsilon_k} - \bar{T} \geq 0$ always holds. For the purpose of arriving at a contradiction, we assume $T^{\epsilon_k} - \bar{T} < 0$. Then (A.12) implies that

$$y^{\epsilon_k})_j - \bar{y}_j > 0.$$

Moreover, (A.11) and (A.13) induce

$$\begin{aligned} (y^{\epsilon_k})_j - \bar{y}_j &= 0, \\ (y^{\epsilon_k})_j - \bar{y}_j &\geq x_j - \bar{y}_j \geq 0, \end{aligned}$$

respectively. Clearly, we have $T^{\epsilon_k} - \bar{T} \geq 0$, which contradicts our assumption. In addition, we have

$$T^{\epsilon_k} - \bar{T} \leq \frac{-\gamma(\xi)T^{\epsilon_k} + \|p(\xi)\|_1 + \gamma(\xi)\|x\|_1}{\epsilon_k\gamma(\xi) + 1} \cdot \epsilon_k \leq (\|p(\xi)\|_1 + \gamma(\xi)\|x\|_1) \epsilon_k.$$

Then, it follows that

$$\begin{aligned} & |(\lambda^{\epsilon_k})_j - \bar{\lambda}_j| \\ & \leq \frac{\gamma(\xi)|T^{\epsilon_k} - \bar{T}| + |\gamma(\xi)(\gamma(\xi)(\bar{T} + x_j) - p_j(\xi))| \epsilon_k}{\epsilon_k\gamma(\xi) + 1} \\ & \leq \frac{\gamma(\xi)(\|p(\xi)\|_1 + \gamma(\xi)\|x\|_1) + |\gamma(\xi)(\gamma(\xi)(\bar{T} + x_j) - p_j(\xi))|}{\epsilon_k\gamma(\xi) + 1} \cdot \epsilon_k \\ & \leq \left(\gamma(\xi)(\|p(\xi)\|_1 + \gamma(\xi)\|x\|_1) + \gamma(\xi)^2 \left(\|x\|_1 + \frac{\|p(\xi)\|_1}{\gamma(\xi)} + \|x\|_1 \right) + \gamma(\xi)\|p(\xi)\|_1 \right) \epsilon_k \\ & \leq 3(\gamma(\xi)^2\|x\|_1 + \gamma(\xi)\|p(\xi)\|_1) \epsilon_k, \end{aligned}$$

where the third inequality follows Lemma 3.2 and the continuity of T^ϵ that

$$\bar{T} \leq \|x\|_1 + \frac{\|p(\xi)\|_1}{\gamma(\xi)}.$$

To summarize, for each $j \in \mathcal{J}$, we always have

$$|(\lambda^\epsilon)_j - \bar{\lambda}_j| \leq 3(\gamma(\xi)^2\|x\|_1 + \gamma(\xi)\|p(\xi)\|_1) \epsilon.$$

Then, according to the definition of l_2 -norm, for any given $x \in \mathbb{R}_+^J$ one can compute

$$\bar{\kappa}(\xi) := 3\sqrt{J}(\gamma(\xi)^2\|x\|_1 + \gamma(\xi)\|p(\xi)\|_1).$$

□