RECENT APPLICATIONS OF NASH EQUILIBRIA

By

Jinye Zhao

A Thesis Submitted to the Graduate
Faculty of Rensselaer Polytechnic Institute
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
Major Subject: Mathematical Sciences

Approved by the
Examining Committee:

____
John Mitchell, Thesis Adviser

____
Jong-Shi Pang, Thesis Adviser

____
Wai Kin (Victor) Chan, Member

____
Ananth Krishnamurthy, Member

____
Satish Ukkusuri, Member

Rensselaer Polytechnic Institute
Troy, New York

November 2007
(For Graduation December 2007)

RECENT APPLICATIONS OF NASH EQUILIBRIA

By

Jinye Zhao

An Abstract of a Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Mathematical Sciences

The original of the complete thesis is on file
in the Rensselaer Polytechnic Institute Library

Examining Committee:

John Mitchell, Thesis Adviser
Jong-Shi Pang, Thesis Adviser
Wai Kin (Victor) Chan, Member
Ananth Krishnamurthy, Member
Satish Ukkusuri, Member

Rensselaer Polytechnic Institute
Troy, New York

November 2007
(For Graduation December 2007)

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

The completion of this thesis owes to the help of many people. It is my great pleasure to take this opportunity to express my gratitude toward them.

Firstly, I would like to thank my advisors Prof. Jong-Shi Pang and Prof. John Mitchell for their valuable advice toward this thesis. Especially, I would like to express my gratitude to Prof. Jong-Shi Pang for his supervision and guidance from the very beginning of this research. His valuable insights and suggestions on game theory helped to greatly improve my research skills. His constant encouragement and support have made him a great mentor. He provided me with many opportunities to experience multiple research areas. Due to technical issues related to his departure from RPI, he is not able to be served as the chair of my thesis committee.

I gratefully thank Prof. John Mitchell for teaching me fundamental and advanced optimization courses and for providing me with a solid foundation for my research.

I am deeply grateful to Prof. Ananth Krishnamurthy, who enriched my knowledge on supply chain and queuing theory and always has time to answer all my questions. His insights on the supply-assembly system greatly contributed to the second chapter of this thesis.

I gratefully acknowledge Prof. Benjamin F. Hobbs in Johns Hopkins University for his distance collaboration on emissions trading systems in electric power markets. I thank him for providing excellent models, giving the constructive comments from an economics point of view, sending lots of updated literature related to this topic. The third chapter of this thesis could not be completed without his help.

I would like to thank Prof. Satish Ukkusuri, who directed me in the dynamic single bottleneck traffic problem. He always encouraged me to explore the problem further and provided me with detailed comments on my work. His broad knowledge and innovative thinking was very helpful in establishing the results in the forth chapter of this thesis.

I am grateful to my committee member Prof. Wai Kin Chan for his time and effort in reviewing this thesis. His valuable feedback helped me improve this thesis in many ways.

I would like to thank NSF. Part of my research is supported by NSF under research grant from Prof. Pang.

I would also like to acknowledge Dr. Ram Ramakrishnan and Gitakrishnan Ramadurai. Many thanks go to Dr. Ram Ramakrishnan for helping me with computational problems in assembly supply chain models and with my presentation slides. I am grateful for Gitakrishnan Ramadurai for sharing his bright thoughts about the traffic single bottleneck problems with me, which triggers me studying the Nash equilibrium application on this area.

I give my special thanks my husband, Brian Schulkin for proofreading my thesis drafts and offering a lot of comments. I really appreciate his thoughtful support and persistent confidence in me. Moreover, I am indebted to my parents in law, Robort Schulkin and Elaine Schulkin for always being interested in my research and for giving me guidance in various ways.

Last but not least, I would like to express my sincere gratitude to my beloved parents, Zuming Zhao and Lei Li for always being there for me, and for supporting me during my entire life.

ABSTRACT

The problem of how to make decisions in a competitive environment is very common and important in many fields. The Nash equilibrium makes a fundamental contribution to formulating and analyzing the rational decision-making in such a problem. This thesis concentrates on the following recent applications of the Nash equilibrium.

The first application considers a simple synchronized supply-assembly system. This part of the thesis analyzes generalized Nash equilibrium decisions in a simple assembly supply system with common joint performance constraints. While they lead to a much more realistic model of a stochastic supply chain system, the performances constraints derived from queuing analysis are highly nonlinear. As a result, it significantly complicates the game models and imposes challenges for numerical analysis. Quasi-variational inequalities are used to serve as the main theoretical framework for constructive equilibrium analysis. This work makes considerable contributions to supply chain and game equilibria analysis.

The second application pertains to emissions allowance allocation systems. This part of the thesis addresses the pressing issues of how emissions allowances should be initially allocated in electric power markets, via the development of game-theoretic nonlinear complementarity models. A significant difficulty in the existence of an equilibrium in the models is that the resulting variational inequality (VI) is not monotone. Therefore, the classical theories and numerical methods of VI do not apply to such models. This challenge is overcome by applying a degree-theoretic result to the VI formulation. The game-theoretical models explicitly represent the interaction between the emissions and electricity markets, which provides us with valuable insights toward the inefficiency of different allocation rules.

The last application deals with the dynamic traffic flow assignment model of single bottleneck on a link and with heterogeneous commuters classes. A linear complementarity problem is developed to find the dynamic user equilibrium of the discrete-time single bottleneck model. The complementarity formulation provides

a formal framework for rigorous mathematical analysis of the dynamic equilibrium problem, and offers a provably convergent algorithm for computing a solution. The uniqueness of equilibrium is established in the homogeneous case. The departure patterns are investigated in the heterogeneous case.

CHAPTER 1
INTRODUCTION

About sixty years ago, John Nash [42] introduced a solution concept in a non-cooperative game, which is now known as the Nash equilibrium. Since then, Nash equilibrium has had fundamental and wide impact on numerous fields such as economics, political science, communication technology and computer science. The concept of Nash equilibrium is used to analyze the strategic interactions among players with conflicting objectives. As more and more individuals or companies compete for limited resources, this idea has drawn a lot of attention because it is a powerful tool for modeling and a reasonable predictor of players' behavior in a competitive environment.

This thesis focuses on some contemporary applications of Nash equilibrium in such areas as: stochastic assembly systems in a decentralized supply chain, emissions trading systems in electric power markets and dynamic single bottleneck problems in traffic flow assignment models. The existence of an equilibrium is established in each application. Moreover, the uniqueness and the computation of an equilibrium are studied in dynamic single bottleneck problems. The Nash equilibrium provides significant insights on the market design, individual incentives, and so on in these models.

The rest of this chapter is organized as follows. In the next section, some basic concepts of Nash equilibrium are introduced. Section 1.2 includes a brief literature review of game theory in supply chain and related queuing theory research. Section 1.3 addresses the emissions allowance allocation problems arising from electric power markets. Section 1.4 provides the background for dynamic traffic single bottleneck problems. Finally, the outline of the thesis is given in section 1.5.

1.1 Nash Equilibrium Problems

In a game involving two or more players, a Nash equilibrium is a set of strategies such that no player has an incentive to change his or her action unilaterally. The mathematical definition of a non-cooperative Nash equilibrium is given as follows.

Let $N = \{1, ..., n\}$ be a finite set of players. Each player has a set of strategies denoted by $S_i \subset \Re^{n_i}$, which is assumed to be independent of the other players' strategies. Let $\mathbf{s} \equiv (s_1, ..., s_n)$ be the strategy profile, which assigns a strategy $s_i \in S_i$ to each player $i \in N$. Moreover, each player has a payoff function $\theta_i(\mathbf{s})$, that identifies the payoff received by player i given all the players' strategies. We write $\mathbf{s}_{-i} \equiv (s_j : j \neq i)$ to denote other players' strategies, thus $\theta_i(\mathbf{s}) = \theta_i(s_i, \mathbf{s}_{-i})$. Player i's problem is as follows: for each fixed but arbitrary $\mathbf{s}_{-i}$, find s_i such that player i's payoff is maximized:

$$\underset{s_i \in S_i}{\text{maximize}} \ \theta_i(s_i, \mathbf{s}_{-i}).$$

A Nash equilibrium is a tuple $\mathbf{s}^* \equiv (s_i^* : i \in N)$ such that for every $i \in N$, s_i^* is an optimal solution to player i's maximization problem given $\mathbf{s}_{-i}^*$, i.e.,

$$s_i^* \in \underset{s_i \in S_i}{\text{argmax}} \ \theta_i(s_i, \mathbf{s}_{-i}^*), \quad i \in N.$$

A highly effective approach for the computation of Nash equilibria is via the methods of finite-dimensional variational inequalities (VI). The finite-dimensional variational inequality problem VI(K, F), is to determine a vector $x^* \in K \subset \Re^n$, such that

$$(x - x^*)^T F(x^*) \geq 0, \quad \forall x \in K,$$

where $F : K \to \Re^n$ is a given continuous function. Let SOL (K, F) be the solution set of VI (K, F). The VI is a powerful tool for analyzing the Nash equilibrium. The relationship between them is presented as follows.

Proposition 1. *[21, Proposition 1.4.2] Let each S_i be a closed convex subset of $\Re^{n_i}$. Suppose that for each fixed tuple $\boldsymbol{s}_{-i}$, the function $\theta_i(s_i, \boldsymbol{s}_{-i})$ is concave and continuously differentiable in s_i. Then a tuple $\boldsymbol{s}^* \equiv (s_i^* : i \in N)$ is a Nash equilibrium if and only if $\boldsymbol{s}^* \in$SOL $(\boldsymbol{K}, \boldsymbol{F})$, where*

$$\boldsymbol{K} \equiv \prod_{i=1}^{n} S_i \quad \text{and} \quad \boldsymbol{F}(x) \equiv (-\nabla_{s_i} \theta_i(\boldsymbol{s}))_{i=1}^{n}.$$

Many results have been established to guarantee the existence and the uniqueness of solutions to a VI. A couple of them are presented in the propositions below. The existence of solutions to VI (K, F) follows from the boundness of the union of the solutions to its perturbed VIs. Indeed, we have the following:

Proposition 2. *[21] Let K be a closed convex cone in $\Re^n$ and $F : K \to \Re^n$ be continuous. If the union $\bigcup_{\tau>0} SOL\ (K, F + \tau I)$ is bounded, then $VI(K, F)$ has a solution.*

This proposition is an important tool to establish the existence of the solutions in the models studied in the subsequent chapters. The uniqueness of a solution to VI (K, F) can be achieved under strict monotonicity.

Proposition 3. *[39] Let $K \subseteq \Re^n$ be closed convex and $F : K \to \Re^n$ be continuous. If F is strictly monotone on K i.e.,*

$$(F(x) - F(y))^T (x - y) > 0, \quad \forall x, y \in K, x \neq y,$$

then the VI (K, F) has at most one solution.

The VI has been extensively studied in the literature. The 2-volume monograph Facchinei and Pang [21] is an excellent reference on the theoretical analysis and applications of the VI. Moreover, when $K = \Re^n_+$, the VI is equivalent to the nonlinear complementarity problem (NCP), whose definition is given as follows. Given a mapping $F : \Re^n_+ \to \Re^n$, the NCP (F) is to find a vector $x \in \Re^n$ such that

$$0 \leq x \perp F(x) \geq 0, \tag{1.1}$$

where $\perp$ is the complementarity sign; i.e. $0 \leq x \perp y \geq 0$ indicates that $x \geq 0$, $y \geq 0$ and $x^T y = 0$. As a special case of the VI, the NCP also plays a very important role in this thesis. Furthermore, if $F(\bullet)$ in NCP (F) is an affine function, then this special instance of NCP is called a linear complementarity problem; and NCP (F) is denoted as LCP (q, M):

$$0 \leq x \perp Mx + q \geq 0,$$

for some matrix $M \in \Re^{n \times n}$ and vector $q \in \Re^n$. The solution set of the above LCP is written as SOL (q, M). The monograph by Cottle, Pang and Stone [16] is a comprehensive source for all aspects of linear complementarity problems. Before we state an existence result for the LCPs, let us introduce a couple of matrix classes as follows.

Definition 4. *Let $M \in \Re^{n \times n}$. M is called an $\boldsymbol{R}_0$-matrix if SOL $(0, M) = 0$. The class of such matrices is denoted $\boldsymbol{R}_0$*

Definition 5. *$M \in \Re^{n \times n}$ is said to be copositive if $x^T M x \geq 0$, for all $x \in \Re^n_+$.*

The following proposition addresses that if the matrix M belongs to both of the above two matrix classes, then LCP (q, M) is solvable for any vector q.

Proposition 6. *[16] LCP (q, M) has a solution if M is a copositive $\boldsymbol{R}_0$ matrix.*

A proof of Proposition 6 can be found in [16, Corollary 3.8.8]. The argument is based on Proposition 2.

An extension of the Nash equilibrium is the generalized Nash equilibrium where each player i's strategy set S_i depends on other players' strategies. More specifically, each player's set of strategies is redefined by $S_i(\mathbf{s}_{-i})$. With $\mathbf{s}_{-i}$ taken as exogenous, player i solves the payoff maximization problem in the variable s_i:

$$\underset{s_i \in S_i(\mathbf{s}_{-i})}{\text{maximize}} \quad \theta_i(s_i, \mathbf{s}_{-i}).$$

The tuple $\mathbf{s}^*$ is a generalized Nash equilibrium, if for each player $i \in N$, s_i^* is an optimal solution to player i's maximization problem, i.e.,

$$s_i^* \in \underset{s_i \in S_i(\mathbf{s}^*_{-i})}{\text{argmax}} \quad \theta_i(s_i, \mathbf{s}^*_{-i}), \quad \forall i \in N.$$

Typically, player i's set of strategies $S_i(\mathbf{s}_{-i})$ is assumed to be

$$S_i(\mathbf{s}_{-i}) \equiv \{s_i \in \Re^{n_i} : g_i(\mathbf{s}) \leq 0, h_i(s_i) \leq 0\},$$

where $g_i : \Re^n \to \Re^{m_i}$ and $h_i : \Re^{n_i} \to \Re^{l_i}$ are given. Rosen [54] considers an important case of generalized Nash equilibrium problems, where the constraints

$g_i(\mathbf{s})$ in set of strategies $S_i(\mathbf{s}_{-i})$, $i \in N$ are shared by all the players (i.e. $g_i(\mathbf{s}) = g_j(\mathbf{s})$, $i \neq j$) and convex with respect to $\mathbf{s}$. This setting is known as the "coupled constraints" case. Under a boundedness assumption, the existence of a generalized Nash equilibrium is established in [54] by using fixed point theory. As it is shown in Proposition 1, a Nash equilibrium problem can be converted into a VI under some convexity and differentiability assumptions, correspondingly, it is natural to formulate a generalized Nash equilibrium as a quasi-variational inequality (QVI) [15, 27, 49].

The QVI is an extension of the VI. Let K be a point-to-set mapping from $E \subseteq \Re^n$ to $\Re^n$. Given a point-to-point mapping $F : \Re^n \to \Re^n$, the finite-dimensional quasi-variational inequality problem QVI (F, K) is to find a vector $x^* \in K(x^*)$ so that

$$(x - x^*)^T F(x^*) \geq 0, \quad \forall x \in K(x^*).$$

Suppose that each function h_i is continuously differentiable and convex on $\Re^{n_i}$; and for each fixed tuple $\mathbf{s}_{-i}$, the function $\theta_i(s_i, \mathbf{s}_{-i})$ and $g_i(s_i, \mathbf{s}_{-i})$ are continuously differentiable and convex in s_i. Define

$$\mathbf{K}(\mathbf{s}) \equiv \prod_{i=1}^{n} S_i(\mathbf{s}_{-i}) \quad \text{and} \quad \mathbf{F}(x) \equiv (-\nabla_{s_i}\theta_i(\mathbf{s}))_{i=1}^{n}.$$

Similar to Proposition 2, a tuple $\mathbf{s}^* \equiv (s_i^* : i \in N)$ is a generalized Nash equilibrium if and only if $\mathbf{s}^* \in \mathbf{K}(\mathbf{s}^*)$ is a solution to QVI $(\mathbf{F}, \mathbf{K})$. The following is an existence result for QVI.

Proposition 7. *[15] Let F be a continuous point-to-point mapping from $\Re^n$ into itself and let K be a point-to-set mapping from $\Re^n$ into subsets of $\Re^n$. If there exists a compact convex set $T \subset \Re^n$ so that*

(a) for every $x \in T$, $K(x)$ is a nonempty, closed, convex subset of T;

(b) K is continuous at every point in T,

then the QVI (K, F) has a solution.

Unfortunately, the uniqueness of the solutions to QVI (K, F) usually requires very strong assumptions. In many applications, the uniqueness property does not hold.

1.2 Supply Chain and Assembly Systems

In recent years, assemble-to-order production systems have become popular among many manufacturing companies, because these systems require less inventory than conventional make-to-stock systems and they respond to customers' orders faster. Hopp and Spearman [30] is a great reference for fundamental concepts in manufacturing management. Song and Zipkin [59] reviews the research on assemble-to-order operation models. In the traditional make-to-stock operation model, the manufacturer stocks the finished goods before demand occurs. The major disadvantage of the make-to-stock operation model is that the manufacturer may over-produce and end up with a large inventory; or it may under-produce and incur lost sales. On the other hand, in the assemble-to-order mode, the finished product is divided into several components which are provided by suppliers. The manufacturer (or the assembler) assembles the components together right after it receives a confirmed order from a customer. A good example of an assemble-to-order operation model is Dell, who holds a substantial lead in sales of PCs. As a matter of fact, the manufacturer shifts its responsibility to suppliers in assemble-to-order production systems, which enables the manufacturer to lower its component inventories, reduce the production capacity and provide a larger product variety. Nevertheless, when outsourcing the assembly work to suppliers, the manufacturer faces the risk of losing the control in the decentralized supply chain. For example, the delivery of final products would be delayed if one of the components was out of stock in the system. Hence, how to select the inventory level and the production capacity for one component really should depend on the status of the other components. Moreover, each member in a supply chain system usually behaves in a decentralized fashion in reality, instead of cooperating to optimize the overall system objective. Therefore, it is very practical to investigate the non-cooperative game in supply chain, with the presence of the stochastic nature in assembly systems. However, it is already quite challenging to rigorously analyze the system performance measures such as throughput, queue length, and waiting time distribution in a synchronized assembly system. This partly explains why much of the decentralized supply chain literature ignores the stochastic feature of assembly systems, and uses some simple but inaccurate approximations for the system performance measures. On the other hand, much of queuing analysis has only focused on deriving performance evaluation of stochastic supply assembly chains under certain system configuration; but does not investigate the optimal design of decentralized supply chain systems in a competitive environment. With competition among the members in the chain, queueing models alone do not provide solutions to optimize supply chain performance.

One of the main contributions of this research lies in the incorporation of queueing in noncooperative game models of stochastic supply chains. The uncertain nature of supply chain operation is taken into account via the performance measures derived from queuing analysis. Specifically, we consider a simple decentralized supply-assembly system consisting of a single assembler and two suppliers. Two complementary components from the suppliers constitute a final product. In terms of manufacturing techniques, all the members operate in a kanban system, where production is triggered by a demand. (See [30].) Given a confirmed customer order, the suppliers and the assembler act independently, and they make their own investment decisions on production capacities or inventory levels to optimize their individual revenue less cost. This results a three-player game. Moreover, all of the members in the system are subject to a common constraint that guarantees that the throughput of final units meets the basic customer's demand. This means that the three-player game is a generalized Nash equilibrium problem, which is very challenging to analyze. Alternatively, the case where suppliers collude together yields a two-player game. A brief review of the related research on decentralized supply chain and queuing models of assembly systems is presented as follows.

Performance measures, such as throughput, queue length and lead time are very important in evaluating the efficiency of the assembly system with a particular design. A lot of published queuing articles related to kanban systems have successfully derived or approximated the expressions for the performance measures under certain system configurations. Som et al. [58] and Takahashi et al.[60, 61] study the departure process from a fork/join station with finite buffers and exponentially distributed inter-arrival times to these buffers. They derived expressions for the marginal distribution of the inter-departure times. Krishnamurthy et al.[32, 33] present an exact analysis of a fork/join station in a closed queuing network with inputs from servers with two-phase Coxian service distributions; throughput, queue length are determined in their paper. Although queuing analysis yields very good estimations for the performance measures in probabilistic assembly systems, the complex expressions of the resulting approximations, which do not necessarily satisfy the convexity and monotonicity properties, impose a challenge for the optimization analysis.

Game theory is a powerful tool to study decentralized supply chains which is characterized by multiple players with conflicting objects. Cachon [10] and Cachon and Netessine [12] review and extend the supply chain literature on the management of incentive conflicts with contracts. Cachon and Lariviere [11] study a game setting where a single manufacturer offers a contract to induce a single supplier to construct production capacity, and they also explore forecast sharing by the manufacturer and the capacity decision by the supplier under two different compliance regimes. Lariviere and Porteus [34] consider a similar game with a single supplier and single retailer, but they focus on the decentralized decision making under different demand conditions. Wang and Gerchak [65] analyze the decentralized equilibrium capacities under two pricing schemes: the assembler-as-leader and the suppliers-as-leaders; and the firms decide on their capacities before observing the actual demand. Gurnani and Gerchak [26] consider a decentralized assembly system where production yield is uncertain for each of the components but the end-product demand is deterministic. Bernsetin and DeCroix [7] study assembler-as-leader games in multi-tier assembly systems where each supplier of a submodule has its own suppliers for components. Carr and Karmarkar [14] analyze competition in multi-echelon assembly supply chains using a coordinated successive Cournot model. However, the equilibrium analysis in these papers overlooks the stochastic interactions of assembly supply chain. Although recently Caldentey and Wein [13] and Jemai and Karaesmen [31] explicitly implement M/M/1 queuing formulations to a competitive supply chain with a single retailer and single supplier, their concerns are about the production inventory control policies, and are quite different from our assembly systems here.

In this part of the thesis, queuing analysis and game theory are integrated together to improve the reliability of decentralized supply chain analysis. While queuing analysis leads to a much more realistic model of a stochastic supply chain system, the performances constraints derived from queuing models are highly nonlinear. As a result, it significantly complicates the game models and imposes challenges for numerical analysis. Moreover, since the players optimization problems do not satisfy the classic sufficient conditions required for the existence of generalized Nash equilibrium, such as the boundedness and joint convexity of the players strategies, the challenge becomes more pronounced. To resolve these issues, the powerful degree-theoretical existence result Proposition 2 is applied to establish the existence of an equilibrium. It will be shown in the subsequent chapter that the traditional simplified approaches could mislead the manufacturer's and the suppliers' investment decisions while incorporating queuing formulas in the models would result in more satisfactory solutions. Under mild conditions, the existence of an equilibrium to the three-player game, as well that of the two-player game is established. Several sets of numerical results reveal some interesting insights about the price of anarchy and the sensitivity of input parameters.

1.3 Emissions Allowance Allocation Systems

As a result of human activities, greenhouse gases have caused enormous global climate changes. A lot of effort has been made to reduce emissions of greenhouse gases. One of the approaches that is widely accepted by many countries or regions is emissions cap-and-trade systems. In emissions cap-and-trade systems, the regulator, usually the government agency, imposes a cap on the total amount of pollutant emissions. The cap is typically set to a level lower than current emissions. At the beginning of each phase, individual firms, or polluters are granted a certain amount of emissions allowances, which is the rights to emit pollutants. If a firm's emissions are above his allowances quotas, he has to purchase emissions permits from the other firms who are willing to sell their allowances. The cap on the emissions is likely to be harmful to "highly-emitted" firms, while at the same time is helpful to those "environmentally friendly" ones. Since the emissions allowances prices are completely determined by the markets, firms have incentive to avoid buying additional emissions credits by reducing their pollutants. In theory, under certain assumptions, such systems result in a least-cost way to achieve the pollution reduction [62].

The first large scale emission cap-and-trade system was the SO_2 trading system under the framework of the Acid Rain Program of the 1990 Clean Air Act in the U.S. (See Ellerman et al. [20].) The European Union Emission Trading Scheme (or EU ETS) is the largest greenhouse gas CO_2 emissions trading scheme in the world. It commenced operation in January 2005 with all 27 member states of the European Union participating in it. While anticipating a strong US federal policy to reduce greenhouse gas emissions, multiple states are taking the lead and have launched their own emissions trading schemes. The most important state-level emissions reduction program is the Regional Greenhouse Gas Initiative (RGGI). (See Burtraw et al. [8].) Carbon emissions trading is specifically for CO_2 and currently makes up a significant amount of trading [25]. In 2005, the total value of the CO_2 allowances covered by the EU ETS was nearly 7 billion euro, which was much greater than the total of all other emissions trading schemes. The large scale of the EU ETS implies that it could have significant affect on the costs of power companies and electricity prices. EU ETS has been criticized for causing the increase in electricity prices. Sijm et al. [56] analyze the impact of EU CO_2 emissions trading on the price of electricity in Germany and the Netherlands.

A lot of factors are involved in designing an emissions cap-and-trade system to achieve environmental effectiveness and economic efficiency. Especially, setting the initial allocation has been one of the most controversial issues in developing a workable cap-and-trade scheme. About 30 billion euro emissions allowances are traded on the market under EU ETS every year. How these allowances are distributed initially is likely to be a matter of considerable concern to the affected firms. In principle, the emissions allowances could either be auctioned, with polluters purchasing them from their regulators; be free of charge grandfathered to polluters on the basis of their emissions history; or be allocated based on the firms' present or recent performances (contingent rules). Sijm [55] evaluates these different allocation methods in terms of economic efficiency, potential investment distortion, social cost of reducing emissions. In most of EU ETS member states, the regulators give out CO_2 emissions allowances to firms by grandfathering method. This approach has been criticized as being less efficient than auctioning and providing little motivation to invest in low-carbon energy [3, 17, 40]. For example, a generator may emit more pollutants on purpose in order to gain more free allowances in the next phase. Moreover, a plant would rather keep non-economic facilities running than closing it and losing the emissions allowances.

Besides the concern over allocating the emissions allowances to existing generators, the regulator has to think about how to treat new entrants when he determines the allocation plan for the following phase or for the long run. The papers [2, 6, 9, 45] consider different types of strategies to deal with the entry of new investments. Their studies show that the management of new investments has a great influence on the shape of future energy. Inappropriate treatment of the new investment may discourage the use of new technologies and degrade the efficiency of cap-and-trade systems. For example, the benefit of auctioning for new entrants is that it entices new firms to install clean energy. A drawback is that new entrants are at a disadvantage compared to the incumbents who are compensated with free allowances. Another approach is to provide the allowances to new entrants for free. Although this method helps new firms enter the market and lower the energy price, it simultaneously benefits dirtier new sources since they receive more allowances than cleaner new sources [44], which is happening in at least eight EU countries.

This research focuses on the effects of alternative emissions allocation policies on an electric power market, which is assumed to be subject to the same rules throughout entire time periods. In particular, we are interested in two types of contingent allocation rules: the potential emission allocation rule (initial allocation of allowances is determined by the firms' installed capacities), and the actual emission allocation rule (initial allocation of allowances is determined by the firms' actual sales). A nonlinear complementarity model is proposed to evaluate the long-run implications of these two rules regarding economic efficiency, consumer costs, technology distortion. A significant difficulty in both the existence of an equilibrium and the computation of it for this application, is that the resulting VI is not monotone. Therefore, the classical theories and numerical methods of VI do not apply to such a model. To overcome this challenge, Proposition 2 is applied to establish the existence of an equilibrium. To illustrate the results that can be obtained from the proposed NCP, an example with a competitive power market at a single node is studied. Under each allocation rule, the equilibrium solutions are used to quantify the distortions measured by the consumer costs, social costs, investment distribution, etc. The results suggest that the distortion could be sever under mild CO_2 limits. Moreover, as the fraction of allowances grandfathered to new investments increases, the loss of firms and consumers also increases. The effect of the presence of capacity market is studies as well.

Much of research has been done to analyze emissions allocation schemes in electric power markets. The papers [6, 9, 48] implement linear programming or *Haiku* models to simulate the short-term energy market. The works closest to that being discussed in this thesis are the studies by [45, 46]. Assuming that coal plants are always operated in preference to gas plants, Neuhoff et al. [45] explore the distortion on coal and gas investments, prices, and generator revenues under an exogenous emissions allowance price and no demand elasticity in. However, the approach proposed in Neuhoff et al. [45] is only limited to analyze two types of generators: coal-fired and gas fired. Moreover, their model is not analyzed with game theoretic basis. On the other hand, the models presented in this thesis are able to analyze any number of generator types; and the number of allowances allocated to new investment and the allowances price are endogenous. Unlike other papers where only energy and emission markets are considered, this part of the thesis expands to capacity markets, and provides a more complete model than previously reported publications. Moreover, minimum output constraints can be embedded in the model. This adds more realism since some capacity (coal in coal plants) cannot be cycled on and off on a daily basis. Therefore, the models of this thesis are more sophisticated than previously reported emissions allowance allocations analysis.

1.4 Dynamic Traffic Single Bottleneck Models

In the past several decades, much effort has been devoted to the study of traffic flow assignment models for congested traffic networks. The previous studies can be divided into two groups: static traffic equilibrium problems and dynamic traffic equilibrium problems. The underlying guiding principle of the first group is the Wardrop user equilibrium [66], which postulates that users in a transportation network select routes of minimal delay. The Wardrop user equilibrium is closely related to the idea of the Nash equilibrium. Extensive research using complementarity formulations or variational inequality formulations, has been carried out in this area. (See Aashtiani and Magnanti [1], Friesz et al. [22] and Patriksson [50].) However, static traffic equilibrium problems have been challenged by recent work that acknowledges the dynamic nature of departure time decisions of traffic network users.

Unlike static traffic assignment models that only deal with endogenous route decisions, dynamic traffic assignment models consider the traffic network over time domain. As an extension of the Wardrop user equilibrium, the dynamic user equilibrium (DUE) stipulates that users in a traffic network choose the routes and the departure times to minimize the cost, including early and late arrival penalties. Due to its analytical flexibility and more general properties, the VI approach has been widely applied to analyze dynamic traffic assignment models. The use of variational inequalities for dynamic traffic equilibrium problems can be traced back to Friesz et al. [23]. They propose a continuous time variational inequality formulation of DUE with route and departure time decisions. Friesz and Mookherjee [24] represent DUE as an infinite dimensional differential variational inequality and come up a fixed point algorithm to find the solution. Ran and Boyce [53] show that a descretized dynamic route-departure choice equilibrium problem is equivalent to a variational inequality problem. Wie et al. [67] formulate the DUE problem as a variational inequality problem in discrete time. Under certain regularity conditions, a discrete time dynamic network user equilibrium is proved to exist.

Instead of dealing with a large complex traffic network, this part of the thesis focuses on a discrete-time model of single bottleneck on a link and with heterogeneous commuters classes. Vickrey [64] first presents a single bottleneck model with homogeneous commuters classes. In this simple version of dynamic single bottleneck model, a fixed number of commuters travel from an original to a destination along a link with a bottleneck during morning rush-hours. Each user is no difference from the others in terms of preferred arrival time, travel cost, early and late arrival cost penalty. Since there is no need to select route, individual users make decisions on their own departure times to achieve DUE. Although a single bottleneck model is not typical of the general congestion picture, the relatively simple model does provide some valuable insights into the nature of the overall congested traffic network. The subsequent papers [5, 18, 29, 57] enrich Vickrey's model with homogeneous commuters classes in various ways. Arnott, de Palma and Lindsey [5] consider elastic demand (i.e. the number of total commuters is price sensitive) and optimal capacity under a variety of pricing regimes. Hendrickson and Kocur [29] add more realistic elements into the model. Smith [57] gives sufficient conditions for the existence of an equilibrium in the single bottleneck model with homogeneous commuters. However, such existence result is obtained under some differentiability assumptions, which are not satisfied in the model discussed in this thesis. Based on the analysis and assumptions in Smith [57], Daganzo [18] shows that there is a unique equilibrium order of arrivals under certain conditions.

In this research, the major focus is the single bottleneck model with heterogeneous commuters classes, where commuters from distinct classes differ in their cost of travel times, early and late arrival penalties and preferred arrival times. The heterogeneity reflects reality better. For example, self-employed workers have higher value of time, but they have more flexible arrival time; blue collar workers cannot afford arriving at work late; white collar workers have relatively lower late arrival cost, so they usually appear at the end of rush hours. In [38], Lindsey extends Daganzo [18] and Smith [57], and shows the existence and uniqueness of an equilibrium of continuous-time single bottleneck models with heterogeneous commuters classes. Newell [47] illustrates the equilibrium solutions by graphical techniques. VanDerZijpp and Koolstra [63] provide a numerical approach to solve single bottleneck models with heterogeneous commuters.

However, none of the single bottleneck model formulations in the literature listed above is based on rigorous game theoretic approaches. In this part of the thesis, a linear complementarity formulation is developed for the discrete-time single bottleneck model with heterogeneous users classes. Contrast to the traditional analysis in the above citations, the game theoretic based complementarity formulation presents rigorous justification of the existence of DUE in bottleneck models. Such representation also enables computational tractability. Moreover, it is assume that the traffic flows in the bottleneck preserve the first-in-first-out (FIFO) queuing discipline. Including a queuing delay component can partly mitigate the traffic realism issues arising in the analysis of traffic flow assignment models.

In this part of the thesis, given linear cost functions, the existence of DUE in single bottleneck models with heterogeneous commuters is established by using the fundamental LCP existence result Proposition 6. Additionally, the LCP formulation allows us to implement Lemke's method to solve for equilibria. Lemke's method is first introduced by Lemke and Howson [36] on bimatrix games and Lemke [35] for more general LCPs. Regarding the uniqueness of an equilibrium, the nonmonotonicity of the resulting LCP imposes challenges for showing such property. A specialized method is developed to verify the uniqueness of the cost at equilibrium in the homogeneous commuters class model. The proof of the uniqueness also provides a more efficient algorithm than Lemke's method to find the solution in this case. The uniqueness of the equilibrium cost for the case of heterogeneous commuters classes remains an open issue. The order of departures and arrivals from different groups are studied. The departure patterns in the heterogeneous case provide valuable insights on the traffic flow tendency and the incentive of the users from different classes. Understanding the commuters departure behavior is important for determining the effects of various transportation policies, such as road capacity expansion, road tolls pricing, etc. Henderson [28] and Newell [47] study the departure pattern in bottleneck models by different methods. The work closest to this research is Arnott et al. [4], in which the order of departures are analyzed for a few specific cases.

1.5 Thesis Outline

The remaining chapters of the thesis are organized as follows.

In Chapter 2, a simple assembly system is investigated. Under the framework of noncooperative games, two different decentralized structures are considered: selfish suppliers and collusive suppliers. Nonlinear complementarity problems are proposed for studying the models. Queuing approximations for the system performance measures are embedded in the analysis. The existence of an equilibrium is established under mild conditions. Sets of numerical results show the difference between two system structures and the advantage of incorporating queuing formulas.

In Chapter 3, the electric power markets under two alternative emissions allowance allocation rules are studies. Nonlinear complementarity problems are proposed for analyzing the impact of the allocation rules for the new entrants. The existence of the equilibria is shown under mild conditions. An example illustrates the efficiency of the two distinct allocation rules for new investments in a power market.

In Chapter 4, the dynamic user equilibrium problem is solved in the single bottleneck model with heterogeneous commuters classes. The model is formulated as a linear complementarity problem. The existence of an equilibrium is established. The uniqueness of the equilibrium cost for the homogeneous commuters classes model is confirmed. A specified algorithm is proposed. Moreover, departures patterns are studied for the single bottleneck model with heterogeneous commuters classes.

CHAPTER 2
SUPPLY CHAIN AND ASSEMBLY SYSTEMS

In this chapter, an assembly supply system consisting of one assembler and two suppliers is discussed. This chapter is organized as follows. Section 2.1 presents two generalized Nash equilibrium models: the three-player game and the two-player game, which corresponding to the case with selfish suppliers and the case with co-operative suppliers, respectively. Based on different approaches, three alternative performance measure approximations are introduced in Section 2.2. Each generalized equilibrium model is formulated as a nonlinear complementarity problems in Section 2.3. The existence equilibria is established under mild conditions in Section 2.4. In Section 2.5, a sequence of numerical experiment results are studied. The superiority of incorporating queueing formulas in a game-theoretic analysis is shown in the results. The difference between the two-player game and the three-player game is discussed as well. Final remarks are provided in the last section.

2.1 Model Definition

Consider an assembly system that involves of a single assembly firm and two supply firms, namely assembler, supplier 1 and supplier 2 respectively in the model. A final product is composed of two complementary components, and each is produced by a supplier. Once receiving a confirmed order from a customer, the assembler merely needs to kit together the components from each supplier to complete a product (see Figure 2.1). Without loss of generality, it is assumed that only one unit of each component is required to make up a final product. Moreover, each supplier only serves one single assembler. That a single supplier connects to multiple assemblers is out of the scope of this chapter. The time it takes to ship and to assemble the components is relatively insignificant, thus it is neglected in the model.

The table below summarizes all the notations used in this chapter. All the parameters listed are positive. The physical units for the variables and input functions are noted within parentheses.

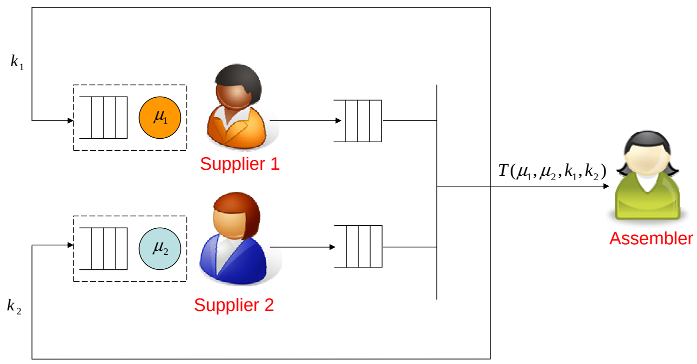

Figure 2.1: A system with two suppliers and a single assembler

Parameters:

α	Confirmed orders received by the assembler from customers or throughput threshold (unit)
$R(\alpha)$	Guaranteed revenue that the assembler promises to the suppliers (\$)
$k_{\min}$	Minimum inventory limit for components inventory (unit)
C_i	Supplier i's capacity investment costs, $i = 1, 2$ (\$/unit)
H_i	Holding cost of the average inventory for supplier i's components, $i = 1, 2$ (\$/unit)
X_i	Supplier i's decision variable μ_i's domain $= [\alpha, \infty)$, $i = 1, 2$
Y_i	Assembler's decision variable k_i' domain $= [k_{\min}, \infty)$, $i = 1, 2$

Variables:

μ_i	Supplier i's production capacity, $i = 1, 2$ (unit)
k_i	Maximum inventory limit for supplier i's components, $i = 1, 2$ (unit)

Functions: all the functions are differentiable.

$T(\mu_1, \mu_2, k_1, k_2)$	Throughput function (unit);
$L_i(\mu_1, \mu_2, k_1, k_2)$	Average inventories from supplier i, $i = 1, 2$ (unit);
$p_i(\bullet)$	Transfer price the assembler pays supplier i for each excess unit sold, (\$/unit), $i = 1, 2$
$\widehat{p}(\bullet)$	Unit sale price of end product (\$/unit)

2.1.1 The assembler's optimization problem

A customer places an order of α units of the end products and ensures the assembler a revenue of $P(\alpha)$. In order to fill the demand within a certain period, the assembler requires each suppliers to guarantee a throughput T of at least equal to α. The assembler promises both suppliers a fraction of his revenue, say $R(\alpha)$ at the beginning of the period. In this way, compensated by a portion of the guaranteed $R(\alpha)$, each supplier is able to start investing on his production capacity to ensure a throughput of α. There are several ways to allocate the guaranteed revenue $R(\alpha)$ between the two suppliers. We introduce the following two rules:

(I) The investment cost allocation rule: each supplier's resulting guaranteed revenue is proportional to their relative capacity investment cost. For example, supplier 1 gets $R(\alpha)C(\mu_1)/[\,C(\mu_1)+C(\mu_2)\,]$.

(II) The capacity allocation rule: each supplier's resulting guaranteed revenue is proportional to their relative capacity investments. For example, supplier 1 gets $R(\alpha)\mu_1/(\mu_1+\mu_2)$ revenue for a capacity investment μ_1 contributing towards the guaranteed throughput α.

Needless to say, other alternative allocation rules are possible, such as the guaranteed revenue $R(\alpha)$ is distributed proportional to the price of the individual components. In the rest of this chapter, we focus on the first rule in the analysis of equilibria.

Beside the revenue of $P(\alpha)$ from the confirmed order α, the assembler might have incentive to sell more units to earn extra profit if there is any demand. It is assumed that the unit sale price of each excess unit $\widehat{p}(T-\alpha)$ is a monotonically decreasing function of the additional final products. Therefore, the revenue by selling the additional $T-\alpha$ units of final products is $(T-\alpha)\widehat{p}(T-\alpha)$. Meanwhile, the assembler has to pay each supplier i the transfer price $p_i(T-\alpha)$ for component i, for each product sold at the price $\widehat{p}(T-\alpha)$.

In additional to the revenue, the costs incurred by the assembler include the holding costs on the average inventories of components from both suppliers. Due to the stochastic nature of an assembly system, the throughput rate might be influenced by the random interruption of the suppliers' operations, such as power outages, machine breakdowns or labor unavailability. Thus, the assembler usually stocks certain amount of components on hand to protect himself against stockout caused by the uncertain fluctuations. Taking into account the physical size of warehouses and the inventory expense, the assembler keeps component i inventory level under k_i units, $i = 1, 2$. It can be shown by queueing analysis that the average inventory of each component is in general subject to the capacity and inventory investment decisions (μ_1, μ_2, k_1, and k_2), i.e. $L_1(\mu_1, \mu_2, k_1, k_2)$ and $L_2(\mu_1, \mu_2, k_1, k_2)$. The detail of alternative queueing approximations of the average inventory functions is presented in the next section. The holding costs include $H_1(L_1(\mu_1, \mu_2, k_1, k_2))$ and $H_2(L_2(\mu_1, \mu_2, k_1, k_2))$ for the two components during the period.

With the suppliers' production capacities μ_1 and μ_2 taken as exogenous in the assembler's optimization problem, the assembler seeks for inventory limits k_1 and k_2 to maximize his revenue less cost:

$$\begin{aligned}
\underset{k_1,k_2}{\text{maximize}} \quad & (\,T(\mu_1, \mu_2, k_1, k_2) - \alpha\,)\,[\,\widehat{p}(T(\mu_1, \mu_2, k_1, k_2) - \alpha\,) \\
& -p_1(\,T(\mu_1, \mu_2, k_1, k_2) - \alpha\,) - p_2(\,T(\mu_1, \mu_2, k_1, k_2) - \alpha\,)\,] \\
& -H_1\,(\,L_1(\mu_1, \mu_2, k_1, k_2)\,) - H_2\,(\,L_2(\mu_1, \mu_2, k_1, k_2)\,) \\
\text{subject to} \quad & T(\mu_1, \mu_2, k_1, k_2) \geq \alpha \quad \text{and} \quad k_1, k_2 \geq k_{\min}.
\end{aligned} \tag{2.1}$$

Since the fixed revenue $P(\alpha)$ and the total guaranteed revenue $R(\alpha)$ are constants in assembler's maximization problem (2.1), they are dropped in the objective function. Under some concavity/convexity assumptions on $T(\mu_1, \mu_2, \bullet, \bullet)$, $L_i(\mu_1, \mu_2, \bullet, \bullet)$ and $H_i(\bullet)$, $i = 1, 2$, the assembler's problem (2.1) can be a concave maximization problem in (k_1, k_2). If all the functions are differentiable, then the Karush-Kuhn-Tucker

(KKT) conditions for (2.1) can be written as follows:

$$
\begin{aligned}
0 \leq\; & H_1'(L_1(\mu_1,\mu_2,k_1,k_2))\frac{\partial L_1(\mu_1,\mu_2,k_1,k_2)}{\partial k_1} - \lambda_a \frac{\partial T(\mu_1,\mu_2,k_1,k_2)}{\partial k_1} \\
& + H_2'(L_2(\mu_1,\mu_2,k_1,k_2))\frac{\partial L_2(\mu_1,\mu_2,k_1,k_2)}{\partial k_1} - \frac{\partial R_a(\mu_1,\mu_2,k_1,k_2)}{\partial k_1} \perp k_1 \geq 0 \\
0 \leq\; & H_1'(L_1(\mu_1,\mu_2,k_1,k_2))\frac{\partial L_1(\mu_1,\mu_2,k_1,k_2)}{\partial k_2} - \lambda_a \frac{\partial T(\mu_1,\mu_2,k_1,k_2)}{\partial k_2} \\
& + H_2'(L_2(\mu_1,\mu_2,k_1,k_2))\frac{\partial L_2(\mu_1,\mu_2,k_1,k_2)}{\partial k_2} - \frac{\partial R_a(\mu_1,\mu_2,k_1,k_2)}{\partial k_2} \perp k_2 \geq 0 \\
0 \leq\; & T(\mu_1,\mu_2,k_1,k_2) - \alpha \perp \lambda_a \geq 0,
\end{aligned}
$$

$$
\begin{aligned}
\text{where } R_a(\mu_1,\mu_2,k_1,k_2) \equiv\; & (T(\mu_1,\mu_2,k_1,k_2)-\alpha)\,[\,\widehat{p}(T(\mu_1,\mu_2,k_1,k_2)-\alpha) \\
& -p_1(T(\mu_1,\mu_2,k_1,k_2)-\alpha) - p_2(T(\mu_1,\mu_2,k_1,k_2)-\alpha)\,],
\end{aligned} \tag{2.2}
$$

and λ_a is the Lagrange multiplier to the throughput constraint. Notice that maximization problem (2.1) satisfies the Slater constraint qualification condition, so the above KKT conditions are both necessary and sufficient for optimality.

2.1.2 The suppliers' optimization problems in the three-player game

In this subsection, a completely decentralized setting is considered between two suppliers, meaning that two suppliers do not cooperate with each other. All the players act independently and try to optimize their own profits. This setting results in a three-player game, which involves the assembler problem (2.1) and two separate supplier's problems described below.

In the three-player game, each supplier needs to make strategic decision on his production capacity μ_i, $i = 1, 2$ ahead of the realization of total demand. The production capacities investments include the amount of equipment to purchase, the amount of labor to hire, etc. The basic requirement for each supplier is to meet the demand of α units of final products. Followed by the assembler's optimization problem in the previous section, suppler i revenue is composed of two parts. The first part is the guaranteed revenue $R(\alpha)C_i(\mu_i)/(\,C_1(\mu_1) + C_2(\mu_2)\,)$ corresponding to the first α units of final products, and the second part is the additional revenue $(T-\alpha)p_i(T-\alpha)$ corresponding to the sale of excess $(T-\alpha)$ units. In terms of costs, the each supplier i incurs convex capacity investment cost, $C_i(\mu_i)$, that is a function of installed capacity, $i=1,2$.

Therefore, for any fixed but arbitrary supplier 2's production capacity μ_2, inventory limits k_1 and k_2, supplier 1 needs to decide on his production capacity μ_1 as to optimize his profit:

$$\begin{aligned} \underset{\mu_1}{\text{maximize}} \quad & (\,T(\mu_1,\mu_2,k_1,k_2)-\alpha\,)\,p_1(T(\mu_1,\mu_2,k_1,k_2)-\alpha) \\ & + R(\alpha)\,\frac{C_1(\mu_1)}{C_1(\mu_1)\,+\,C_2(\mu_2)}\; - \; C_1(\mu_1) \\ \text{subject to} \quad & T(\mu_1,\mu_2,k_1,k_2) \,\geq\, \alpha \quad \text{and} \quad \mu_1 \,\geq\, 0. \end{aligned} \tag{2.3}$$

Similarly, with μ_1, k_1, and k_2 taken as exogenous, supplier 2's problem is to solve for μ_2 in

$$\begin{aligned} \underset{\mu_2}{\text{maximize}} \quad & (\,T(\mu_1,\mu_2,k_1,k_2)-\alpha\,)\,p_2(T(\mu_1,\mu_2,k_1,k_2)-\alpha) \\ & + R(\alpha)\,\frac{C_2(\mu_2)}{C_1(\mu_1)\,+\,C_2(\mu_2)}\; - \; C_2(\mu_2) \\ \text{subject to} \quad & T(\mu_1,\mu_2,k_1,k_2) \,\geq\, \alpha \quad \text{and} \quad \mu_2 \,\geq\, 0. \end{aligned} \tag{2.4}$$

A tuple $(\mu_1^*,\mu_2^*,k_1^*,k_2^*)$ is a three-player Nash equilibrium if

$$\mu_1^* \in \text{argmax of (2.3)}; \qquad \mu_2^* \in \text{argmax of (2.4)};$$

$$\text{and} \qquad (k_1^*,k_2^*) \in \text{argmax of (2.1)}.$$

There are several analytical challenges that need to be addressed. First of all, notice that the optimization problems (2.1), (2.3) and (2.4) have a common constraint $T(\mu_1,\mu_2,k_1,k_2) \geq \alpha$, which contains all the players decision variables. Therefore, the three-player game is a generalized Nash equilibrium problem by the definition given in Chapter 1. If $T(\mu_1,\mu_2,k_1,k_2)$ is not jointly concave in all the players' variables, the uniqueness of an equilibrium does not hold in general. Furthermore, the throughput $T(\bullet)$ and average inventory $L_i(\bullet)$, $i=1,2$, are probabilistic functions of (μ_1, μ_2, k_1, k_2) that reflect the stochastic characteristics of the assembly system. In the next section, some different approaches to determine the throughput and average inventory functions will be presented. As will be demonstrated in subsequent sections, the particular choice of these functions significantly impacts the resulting analysis complexity and equilibrium solutions. Last but not least, in order to analyze the equilibrium problem, each of the three optimization problems (2.1), (2.3) and (2.4) has to be convex program. This means that the throughput function and the average inventory functions need to have certain convexity/concavity properties, which are usually luxuries for most of queueing approximations. Fortunately, all the candidate approximations for the throughput function and the average inventory functions provided in the next section have the desired convexity/concavity properties.

For now, we assume that $T(\bullet, \mu_2, k_1, k_2)$ is concave and $C_1(\bullet)$ is convex and that the revenue function $T \mapsto Tp_1(T)$ is concave and nondecreasing. Consequently, (2.3) is a concave maximization problem in μ_1. Taking these to be differentiable functions, noticing that a Slater point to the constraints readily exists, and letting λ_1 be the Lagrange multiplier to the throughput constraint, we can write down the KKT conditions for (2.3), which are both necessary and sufficient for optimality:

$$
\begin{aligned}
0 \le\ & C_1'(\mu_1) - R(\alpha)\,\frac{C_2(\mu_2)C_1'(\mu_1)}{(\,C_1(\mu_1) + C_2(\mu_2)\,)^2} - \lambda_1\,\frac{\partial T(\mu_1, \mu_2, k_1, k_2)}{\partial \mu_1} \\
& \qquad\qquad - \frac{\partial R_1(\mu_1, \mu_2, k_1, k_2)}{\partial \mu_1}\ \perp\ \mu_1 \ge 0 \qquad (2.5) \\
0 \le\ & T(\mu_1, \mu_2, k_1, k_2) - \alpha\ \perp\ \lambda_1 \ge 0,
\end{aligned}
$$

where $R_1(\mu_1, \mu_2, k_1, k_2) \equiv (T(\mu_1, \mu_2, k_1, k_2) - \alpha)p_1(T(\mu_1, \mu_2, k_1, k_2) - \alpha)$ is supplier 1's revenue when the throughput exceeds the threshold α.

Similarly, under the same properties for supplier 2's objective function, (2.4) is equivalent to its KKT conditions:

$$0 \leq C_2'(\mu_2) - R(\alpha)\,\frac{C_1(\mu_1)C_2'(\mu_2)}{(\,C_1(\mu_1)+C_2(\mu_2)\,)^2} - \lambda_2\,\frac{\partial T(\mu_1,\mu_2,k_1,k_2)}{\partial \mu_2} - \frac{\partial R_2(\mu_1,\mu_2,k_1,k_2)}{\partial \mu_2} \perp \mu_2 \geq 0 \tag{2.6}$$

$$0 \leq T(\mu_1,\mu_2,k_1,k_2) - \alpha \perp \lambda_2 \geq 0,$$

where $R_2(\mu_1,\mu_2,k_1,k_2) \equiv (T(\mu_1,\mu_2,k_1,k_2)-\alpha)p_2(T(\mu_1,\mu_2,k_1,k_2)-\alpha)$ is supplier 2's revenue when the throughput exceeds the threshold α. In the optimization conditions (2.5) and (2.6), if the cost functions $C_i(\mu_i) \equiv C_i\mu_i$, $i = 1, 2$, then the partial derivative of $C_i(\mu_i)/(\,C_1(\mu_1)+C_2(\mu_2)\,)$ becomes $C_1C_2\mu_i/(C_1\mu_1+C_2\mu_2)^2$, whose denominator is equal to zero when $\mu_1 = \mu_2 = 0$ (even though this will not happen at an equilibrium). This technical challenge will be resolved in the analysis later.

2.1.3 The suppliers' optimization problem in the two-player game

Unlike in the three-player game, two suppliers collude together and compete with the assembler in the two-player game. Thus, the assembler's problem remains the same; and the supplier 1's and supplier 2's problems are combined into one single problem stated as follows.

Suppliers' problem: with k_1, and k_2 taken as exogenous, solve for μ_1 and μ_2 in

$$\begin{aligned} \underset{\mu_1,\mu_2}{\text{maximize}} \quad & (\,T(\mu_1,\mu_2,k_1,k_2)-\alpha\,)\,[p_1(T(\mu_1,\mu_2,k_1,k_2)-\alpha) \\ & +p_2(T(\mu_1,\mu_2,k_1,k_2)-\alpha)] - C_1(\mu_1) - C_2(\mu_2) \\ \text{subject to} \quad & T(\mu_1,\mu_2,k_1,k_2) \geq \alpha \quad \text{and} \quad \mu_1,\mu_2 \geq 0. \end{aligned} \tag{2.7}$$

Since two supplies' objective functions in (2.3) and (2.4) are added together in this model, the guaranteed revenue $R(\alpha)$ becomes a constant, which is dropped in the objective function in (2.7).

A tuple $(\mu_1^*,\mu_2^*,k_1^*,k_2^*)$ is a two-player Nash equilibrium if

$$(\mu_1^*,\mu_2^*) \in \text{argmax of (2.7)} \quad \text{and} \quad (k_1^*,k_2^*) \in \text{argmax of (2.1)}.$$

Under that assumption that that $T(\bullet, \bullet, k_1, k_2)$ is jointly concave in (μ_1, μ_2), (2.7) is a concave maximization problem, which is equivalent to its KKT conditions:

$$
\begin{aligned}
&0 \leq C_1'(\mu_1) - \lambda_s \frac{\partial T(\mu_1, \mu_2, k_1, k_2)}{\partial \mu_1} - \frac{\partial R_1(\mu_1, \mu_2, k_1, k_2)}{\partial \mu_1} \perp \mu_1 \geq 0 \\
&0 \leq C_2'(\mu_2) - \lambda_s \frac{\partial T(\mu_1, \mu_2, k_1, k_2)}{\partial \mu_2} - \frac{\partial R_2(\mu_1, \mu_2, k_1, k_2)}{\partial \mu_2} \perp \mu_2 \geq 0 \\
&0 \leq T(\mu_1, \mu_2, k_1, k_2) - \alpha \perp \lambda_s \geq 0,
\end{aligned}
\tag{2.8}
$$

where λ_s is the Lagrange multiplier to the throughput constraint.

2.1.4 The one-player game

In a completely centralized supply system, all the suppliers and the assembler collude together. It results a single optimization problem, which is called one-player game. In particular, the one-player game is to find a tuple (μ_1, μ_2, k_1, k_2) such that the summation of all the players' payoff is maximized:

$$
\begin{aligned}
\underset{\mu_1,\mu_2,k_1,k_2}{\text{maximize}} \quad & (\,T(\mu_1, \mu_2, k_1, k_2) - \alpha\,)\widehat{p}(T(\mu_1, \mu_2, k_1, k_2) - \alpha) - C_1(\mu_1) - C_2(\mu_2) \\
& -H_1(L_1(\mu_1, \mu_2, k_1, k_2)) - H_2(L_2(\mu_1, \mu_2, k_1, k_2)) \\
\text{subject to} \quad & T(\mu_1, \mu_2, k_1, k_2) \geq \alpha, \quad k_1, k_2 \geq k_{\min}, \quad \text{and} \quad \mu_1, \mu_2 \geq 0.
\end{aligned}
$$

Note that the decision variables in the one-player game are the tuple (μ_1, μ_2, k_1, k_2). Thus, to apply the KKT optimality conditions to find the maximal solution of the above problem, it requires the objective function and the constraints to be jointly concave with respect to (μ_1, μ_2, k_1, k_2). However, this condition is a very strong; at least all the throughput functions and average inventory functions provided in the next section do not meet this requirement. Therefore, the one-player game's KKT solution is not guaranteed to be global optimum. Because the one-player game is lack of the necessary concavity property, we focus on the three-player game and the two-player game in the analysis of existence of an equilibrium.

2.2 Approximations for Performance Measures

This section proposes several alternative approximations for the throughput $T(\bullet)$ and average inventory $L_i(\bullet)$, $i = 1, 2$. Although each approximation for these functions is determined by a specific stochastic approach, it is assumed that the throughput function $T(\mu_1, \mu_2, k_1, k_2)$ derived from any these methods will satisfy properties stated in the following subsection.

2.2.1 Assumptions on performance measures

Assumption 8. *There exists* $0 < \kappa < 1$ *such that for any* $(\mu_1, \mu_2, k_1, k_2) \in X_1 \times X_2 \times Y_1 \times Y_2$, *we have*

$$\kappa \min(\mu_1, \mu_2) \leq T(\mu_1, \mu_2, k_1, k_2) \leq \min(\mu_1, \mu_2).$$

Assumption 9. *There exists* $\beta > 0$ *such that for any* $(\mu_1, \mu_2, k_1, k_2) \in X_1 \times X_2 \times Y_1 \times Y_2$, *we have*

$$\frac{\partial T(\mu_1, \mu_2, k_1, k_2)}{\partial \mu_1} \geq \beta, \quad \text{if } \mu_1 \leq \mu_2,$$
$$\frac{\partial T(\mu_1, \mu_2, k_1, k_2)}{\partial \mu_2} \geq \beta, \quad \text{if } \mu_2 \leq \mu_1.$$

Assumption 10. *For any* $(\mu_j, k_1, k_2) \in X_j \times Y_1 \times Y_2$,

$$\liminf_{\mu_i \to \infty} \frac{\partial T(\mu_1, \mu_2, k_1, k_2)}{\partial \mu_i} = 0, \quad i \in \{1, 2\} \text{ and } i \neq j.$$

Assumption 11. *For any* $(\mu_1, \mu_2, k_j) \in X_1 \times X_2 \times Y_j$,

$$\liminf_{k_i \to \infty} \frac{\partial T(\mu_1, \mu_2, k_1, k_2)}{\partial k_i} = 0, \quad i \in \{1, 2\} \text{ and } i \neq j.$$

The right hand side of Assumption 8 emphasizes a fundamental characteristic of throughput in assembly systems. Since, both components are required to form a kit, the throughput rate for kits cannot be greater than the production rates of the individual components. The other three assumptions are technical conditions that are used in the proof of existence of an equilibrium.

2.2.2 Approximations based on upper bounds

Inspired by Assumption 8, $T(\mu_1, \mu_2, k_1, k_2)$ can be estimated by $\min(\mu_1, \mu_2)$. This approximation characterizes the major feature of an assembly supply chain. In particular, the throughput of an assembly system is always bounded by the smaller of the two production capacities. In analogous fashion, the average inventory functions $L_i(\mu_1, \mu_2, k_1, k_2)$, $i = 1, 2$ can be approximated by the upper bounds i.e. $k_i/2$ for $i = 1, 2$, respectively. Thus, we have

$$T^u(\mu_1, \mu_2, k_1, k_2) = \min(\mu_1, \mu_2) \tag{2.9}$$

$$L_i^u(\mu_1, \mu_2, k_1, k_2) = k_i/2 \qquad i = 1, 2, \tag{2.10}$$

where the superscript emphasizes that the approximations are based on the upper bounds. However, the non-differentiable throughput function defined by (2.9) does not satisfy Assumptions 2 and 3, which require the partial derivatives of T with respect to μ_1 and μ_2. In spite of this fact, the Nash equilibrium problem resulting from the two formulas (2.9) and (2.10) is trivially solvable.

The approximations based on the upper bounds are widely employed in a lot of decentralized supply chain literature. Nevertheless, they are too simple to capture the stochastic interactions. For instance, the expression for $T^u(\mu_1, \mu_2, k_1, k_2)$ ignores the effect of the assembler's decision variables k_i, $i = 1, 2$ on throughput. In the next subsection, more realistic approximations are introduced.

2.2.3 Approximations based on queuing models

In general, it is challenging to determine exact closed-form expressions for the steady state probabilities of a synchronized assembly system, unless specific assumptions are imposed on the stochastic behavior of the system. For a relatively simple system like our case, the expressions of the $T(\mu_1, \mu_2, k_1, k_2)$ and $L_i(\mu_1, \mu_2, k_1, k_2)$, $i = 1, 2$ can be deduced from rigorous queueing analysis under Markovian assumptions. The detail of the queueing proof is out of the scope of this chapter. The following are the formulas for the throughput function and the average inventory functions.

$$T^M(\mu_1,\mu_2,k_1,k_2) = \begin{cases} \mu_1\,\mu_2 \left(\dfrac{\mu_1^{k_1+k_2} - \mu_2^{k_1+k_2}}{\mu_1^{k_1+k_2+1} - \mu_2^{k_1+k_2+1}} \right) & \text{if} \quad \mu_1 \neq \mu_2, \\ \mu \left(\dfrac{k_1+k_2}{k_1+k_2+1} \right) & \text{if} \quad \mu_1 = \mu_2 = \mu, \end{cases} \tag{2.11}$$

and

$$L_2^M(\mu_1,\mu_2,k_1,k_2) = \begin{cases} \left(\dfrac{k_2}{1-\rho^{k_1+k_2+1}} \right) - \left(\dfrac{\rho}{1-\rho} \right) \left(\dfrac{1-\rho^{k_2}}{1-\rho^{k_1+k_2+1}} \right) & \\ & \text{if } \rho \equiv \dfrac{\mu_1}{\mu_2} \neq 1 \\ \dfrac{k_2\,(\,k_2+1\,)}{2\,(\,k_1+k_2+1\,)} & \text{if } \rho = 1. \end{cases}$$

Unlike the approximations (2.9) and (2.10), the above formulas clearly show that the dependence of the throughput and average queue lengths on (μ_1, μ_2, k_1, k_2) is rather complex; indeed the dependence is nonlinear. Ramakrishnan and Krishnamurthy [52] propose alternative simpler approximations for the average inventory as follows.

$$\bar{L}_i^M(\mu_1,\mu_2,k_1,k_2) = \frac{k_i^2\,[\,\mu_i - T(\mu_1,\mu_2,k_1,k_2)\,]}{\mu_i k_i - (k_i-1)T(\mu_1,\mu_2,k_1,k_2)} \tag{2.12}$$

or its relaxation

$$\bar{L}_i^M(\mu_1,\mu_2,k_1,k_2) = \frac{k_i^2(\mu_i - \alpha)}{\mu_i k_i - (k_i-1)\alpha}. \tag{2.13}$$

Note the latter formula for $\bar{L}_i^M(\mu_1,\mu_2,k_1,k_2)$ shows its dependence on the i-pair (μ_i, k_i) only.

2.2.4 Approximations based on simple probabilistic models

The expressions derived from the queueing model are accurate, but very complex. The formulas based on the upper bounds are fairly simple but not realistic. The approximations based on simple probabilistic models compromise both approaches discussed previously. T^M shows that throughput function is jointly determined by (μ_1, μ_2, k_1, k_2) while T^u extracts the fact that throughput function is dominated by the smaller of the two capacities. Combining theses two features together, it suggests that $T(\mu_1, \mu_2, k_1, k_2) = g(k_1, k_2) \min(\mu_1, \mu_2)$. If letting the function $g(k_1, k_2)$ be $[1 - 1/(k_1 + k_2 + 1)]$, then it yields that

$$T^S(\mu_1, \mu_2, k_1, k_2) = \min(\mu_1, \mu_2) \left(1 - \frac{1}{k_1 + k_2 + 1}\right) \tag{2.14}$$

$$L_i^S(\mu_1, \mu_2, k_1, k_2) = k_i/2 \qquad i = 1, 2. \tag{2.15}$$

The simple approximation (2.14) assumes that the throughput depend on $\mu_i, i = 1, 2$ and $k_i, i = 1, 2$; this approximation reduces to (2.11) when $\mu_1 = \mu_2$ derived under the Markovian assumptions.

Notice that $T^S(\mu_1, \mu_2, k_1, k_2)$ in (2.14) is not differentiable because of the term $\min(\mu_1, \mu_2)$. This problem can be resolved by considering two cases: $\mu_1 \le \mu_2$ or $\mu_2 \le \mu_1$. Correspondingly, the resulting three-player game includes two Nash equilibrium problems. Specifically, if $\mu_1 \le \mu_2$, then it follows that $T^S(\mu_1, \mu_2, k_1, k_2) = \mu_1\ (1 - 1/(k_1 + k_2 + 1))$. Consequently, the three players' optimization problems are, respectively:

Supplier 1's problem: given μ_2, k_1, and k_2 with $\mu_2 \ge \alpha$ and $\min(k_1, k_2) \ge 1$, solve for μ_1 in

$$\begin{aligned}
\underset{\mu_1}{\text{maximize}} \quad & R(\alpha)\frac{C_1(\mu_1)}{C_1(\mu_1) + C_2(\mu_2)} - C_1(\mu_1) \\
& + \left[\mu_1\left(1 - \frac{1}{k_1 + k_2 + 1}\right) - \alpha\right] p_1\left(\mu_1\left(1 - \frac{1}{k_1 + k_2 + 1}\right) - \alpha\right) \\
\text{subject to} \quad & \mu_1\left(1 - \frac{1}{k_1 + k_2 + 1}\right) \ge \alpha \quad \text{and} \quad \mu_1 \le \mu_2;
\end{aligned}$$

Supplier 2's problem: given μ_1, k_1, and k_2 such that $\mu_1\left(1 - \frac{1}{k_1 + k_2 + 1}\right) \ge \alpha$ and $\min(k_1, k_2) \ge 1$, solve for μ_2 in

$$\begin{aligned}
\underset{\mu_2}{\text{maximize}} \quad & R(\alpha)\frac{C_2(\mu_2)}{C_1(\mu_1) + C_2(\mu_2)} - C_2(\mu_2) \\
\text{subject to} \quad & \mu_2 \ge \mu_1;
\end{aligned}$$

Assembler's problem: given μ_1 and μ_2 such that $\mu_2 \geq \mu_1 \geq \alpha$, solve for k_1 and k_2 in

$$\begin{aligned}\underset{k_1,k_2}{\text{maximize}} \quad & \left[\mu_1\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right]\left[\widehat{p}\left(\mu_1\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right)-\right.\\ & \left. p_1\left(\mu_1\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right)-p_2\left(\mu_1\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right)\right]\\ & -H_1(L_1^S(k_1))-H_2(L_2^S(k_2))\\ \text{subject to} \quad & \mu_1\left(1-\frac{1}{k_1+k_2+1}\right)\geq\alpha \quad \text{and} \quad k_1,k_2\geq 1.\end{aligned}$$

If $\mu_1 \leq \mu_2$, then $T^S(\mu_1,\mu_2,k_1,k_2) = \mu_2\left(1-1/(k_1+k_2+1)\right)$. Thus, the three players' optimization problems are, respectively:

Supplier 1's problem: given μ_2, k_1, and k_2 such that $\mu_2\left(1-\frac{1}{k_1+k_2+1}\right)\geq\alpha$ and $\min(k_1,k_2)\geq 1$, solve for μ_1 in

$$\begin{aligned}\underset{\mu_1}{\text{maximize}} \quad & R(\alpha)\frac{C_1(\mu_1)}{C_1(\mu_1)+C_2(\mu_2)}-C_1(\mu_1)\\ \text{subject to} \quad & \mu_1 \geq \mu_2;\end{aligned}$$

Supplier 2's problem: given μ_1, k_1, and k_2 with $\mu_1 \geq \alpha$ and $\min(k_1,k_2)\geq 1$, solve for μ_2 in

$$\begin{aligned}\underset{\mu_2}{\text{maximize}} \quad & R(\alpha)\frac{C_2(\mu_2)}{C_1(\mu_1)+C_2(\mu_2)}-C_2(\mu_2)\\ & +\left[\mu_2\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right]p_2\left(\mu_2\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right)\\ \text{subject to} \quad & \mu_2\left(1-\frac{1}{k_1+k_2+1}\right)\geq\alpha \quad \text{and} \quad \mu_2\leq\mu_1;\end{aligned}$$

Assembler's problem: μ_1 and μ_2 such that $\mu_1 \geq \mu_2 \geq \alpha$, solve for k_1 and k_2 in

$$\begin{aligned}
\underset{k_1,k_2}{\text{maximize}} \quad & \left[\mu_2\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right]\left[\widehat{p}\left(\mu_2\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right)-\right. \\
& \left. p_1\left(\mu_2\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right)-p_2\left(\mu_2^*\left(1-\frac{1}{k_1+k_2+1}\right)-\alpha\right)\right] \\
& -H_1(L_1^S(k_1))-H_2(L_2^S(k_2)) \\
\text{subject to} \quad & \mu_2\left(1-\frac{1}{k_1+k_2+1}\right) \geq \alpha \quad \text{and} \quad k_1, k_2 \geq 1.
\end{aligned}$$

2.3 Nonlinear Complementarity Formulation

In this section, the three-player game and the two-player game defined in Section 2.1 are formulated as nonlinear complementarity problems under some convexity/concavity assumptions. To simplify the analysis, we take the capacity investment cost functions and the holding cost functions to be linear, i.e. $C_i(\mu_i) = C_i\mu_i$ and $H_i(L_i) = H_iL_i$, $i = 1, 2$ for some positive constants C_i and H_i. The price functions are also taken to be linear: $p_i(T-\alpha) = a_i - b_i(T-\alpha)$ where $b_i > 0$, $i = 1, 2$ and $\widehat{p}(T-\alpha) = \widehat{a} - \widehat{b}(T-\alpha)$, for some constants a_i, b_i, $\widehat{a}$, and $\hat{b} > b_1 + b_2$. Furthermore, to simplify the notation somewhat, we take $L_i(\bullet)$ to be a function of (μ_i, k_i) only (cf. (2.13)) and write $L_i(\mu_i, k_i)$ for $L_i(\mu_1, \mu_2, k_1, k_2)$.

Assumption 12. The following concavity/convexity conditions are in place:
- For any fixed $(\mu_2, k_1, k_2) \in X_2 \times Y_1 \times Y_2$, $T(\bullet, \mu_2, k_1, k_2)$ and $R_1(\bullet, \mu_2, k_1, k_2)$ are concave on X_1.
- For any fixed $(\mu_1, k_1, k_2) \in X_1 \times Y_1 \times Y_2$, $T(\mu_1, \bullet, k_1, k_2)$ and $R_2(\mu_1, \bullet, k_1, k_2)$ are concave on X_2.
- For any fixed $(\mu_1, \mu_2) \in X_1 \times X_2$, $T(\mu_1, \mu_2, \bullet, \bullet)$ and $R_a(\mu_1, \mu_2, \bullet, \bullet)$ are concave on $Y_1 \times Y_2$.
- For any fixed $\mu_i \in X_i$, $L_i(\mu_i, \bullet)$ is convex and monotonically increasing on Y_i.

Under Assumption 8, it does not hurt to replace the constraint $\mu_i \geq 0$ by $\mu_i \geq \alpha$, $i = 1, 2$ in the optimization problems (2.3) and (2.4). Let $\bar{\mu}_i \equiv \mu_i - \alpha \geq 0$ and $\bar{k}_i \equiv k_i - k_{\min}$.

For the three-player game, the KKT conditions (2.2), (2.5) and (2.6) are equivalent to the corresponding three players' optimization problems under the Assumption 12. Putting (2.2), (2.5) and (2.6) together, the three-player game can be written as a nonlinear complementarity problem (NCP) of the form: $0 \leq x \perp F(x) \geq 0$, where $x = (\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2, \lambda_1, \lambda_2, \lambda_a)$ and $F(x)$ is the vector function given by

$$F(x) = \begin{pmatrix} C_1 - R(\alpha)\,\dfrac{C_1\,C_2\,(\bar{\mu}_2+\alpha)}{(\,C_1(\bar{\mu}_1+\alpha)\;+\;C_2(\bar{\mu}_2+\alpha)\,)^2} \\ \qquad -\big(\,a_1 - 2b_1(T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha)+\lambda_1\,\big)\,\dfrac{\partial T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)}{\partial\bar{\mu}_1} \\ C_2 - R(\alpha)\,\dfrac{C_1\,C_2\,(\bar{\mu}_1+\alpha)}{(\,C_1(\bar{\mu}_1+\alpha)\;+\;C_2(\bar{\mu}_2+\alpha)\,)^2} \\ \qquad -\big(\,a_2 - 2b_2(T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha)+\lambda_2\,\big)\,\dfrac{\partial T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)}{\partial\bar{\mu}_2} \\ H_1\,\dfrac{\partial L_1(\bar{\mu}_1,\bar{k}_1)}{\partial\bar{k}_1} - \big(\,a - 2b(T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha)+\lambda_a\,\big)\,\dfrac{\partial T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)}{\partial\bar{k}_1} \\ H_2\,\dfrac{\partial L_2(\bar{\mu}_2,\bar{k}_2)}{\partial\bar{k}_2} - \big(\,a - 2b(T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha)+\lambda_a\,\big)\,\dfrac{\partial T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)}{\partial\bar{k}_2} \\ T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha \\ T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha \\ T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha \end{pmatrix}.$$

where $a \equiv \widehat{a}-a_1-a_2$ and $b \equiv \widehat{b}-b_1-b_2$. Note that the fraction $C_1C_2\mu_i/(C_1\mu_1+C_2\mu_2)^2$ becomes $C_1C_2(\bar{\mu}_i+\alpha)/(\,C_1(\bar{\mu}_1+\alpha)\;+\;C_2(\bar{\mu}_2+\alpha)\,)^2$, whose denominator is always positive.

In the same way, the two-player game is equivalent to a similar NCP: $0 \leq \widehat{x} \perp$

$\widehat{F}(\widehat{x}) \geq 0$, where $\widehat{x} = (\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2, \lambda_s, \lambda_a)$ and

$$\widehat{F}(\widehat{x}) = \begin{pmatrix} C_1 - \big(a_1 + a_2 - 2(b_1+b_2)(T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2) - \alpha) + \lambda \big) \dfrac{\partial T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2)}{\partial \bar{\mu}_1} \\ C_2 - \big(a_1 + a_2 - 2(b_1+b_2)(T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2) - \alpha) + \lambda \big) \dfrac{\partial T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2)}{\partial \bar{\mu}_2} \\ H_1 \dfrac{\partial L_1(\bar{\mu}_1, \bar{k}_1)}{\partial \bar{k}_1} - \big(a - 2b(T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2) - \alpha) + \lambda_a \big) \dfrac{\partial T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2)}{\partial \bar{k}_1} \\ H_2 \dfrac{\partial L_2(\bar{\mu}_2, \bar{k}_2)}{\partial \bar{k}_2} - \big(a - 2b(T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2) - \alpha) + \lambda_a \big) \dfrac{\partial T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2)}{\partial \bar{k}_2} \\ T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2) - \alpha \\ T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2) - \alpha \end{pmatrix}.$$

Note that the two multipliers λ_1 and λ_2 in the three-player NCP reduce to a single multiplier λ_s in the above two-player NCP and that there is one less term $T(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2) - \alpha$ in the function $\widehat{F}(\widehat{x})$. These two NCPs will be the basis for analyzing the connections between the solutions to the three-player and the two-player models.

2.4 Existence of Equilibria

The following two results assert the existence of equilibria to the three-player and the two-player games with common multipliers.

Theorem 13. Under Assumptions 8–12, the three-player NCP $0 \leq x \perp F(x) \geq 0$ has a solution with a common multiplier to the throughput constraint, i.e., with $\lambda_1 = \lambda_2 = \lambda_a$.

Theorem 14. Under Assumptions 8-12, the two-player NCP $0 \leq \widehat{x} \perp \widehat{F}(\widehat{x}) \geq 0$ has a solution with a common multiplier to the throughput constraint, i.e., with $\lambda = \lambda_a$.

The existence proof is based on the application of Proposition 2, the existence result of the VI introduced in Chapter 1. We rewrite this fundamental degree-theoretic result for a general NCP, which is summarized in the lemma below. For a proof of the lemma, see [21, Theorem 2.6.1].

Lemma 15. Let $\Phi : \Re^n \to \Re^n$ be a continuous function. If there exists a constant $c > 0$ such that all solutions of the NCP: $0 \leq x \perp \Phi(x) + \tau x \geq 0$ for $\tau > 0$ satisfy $\|x\| \leq c$, then the NCP: $0 \leq x \perp \Phi(x) \geq 0$ has a solution. □

2.4.1 Proof for the three-player model

Since the existence proofs for the three-player NCP and the two-player NCP are very similar, we only present the proof for the former. For this purpose, we consider a perturbation of the function $F(x)$: for each $\varepsilon > 0$, let

$$F^\varepsilon(\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2, \bar{\lambda}) = \begin{pmatrix} C_1 - R(\alpha)\dfrac{C_1\,C_2\,(\bar{\mu}_2+\alpha)}{\left(C_1(\bar{\mu}_1+\alpha)\,+\,C_2(\bar{\mu}_2+\alpha)\right)^2} \\ \qquad -\left(a_1 - 2b_1(T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha)+\bar{\lambda}\right)\dfrac{\partial T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)}{\partial \bar{\mu}_1} \\ C_2 - R(\alpha)\dfrac{C_1\,C_2\,(\bar{\mu}_1+\alpha)}{\left(C_1(\bar{\mu}_1+\alpha)\,+\,C_2(\bar{\mu}_2+\alpha)\right)^2} \\ \qquad -\left(a_2 - 2b_2(T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha)+\bar{\lambda}\right)\dfrac{\partial T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)}{\partial \bar{\mu}_2} \\ H_1\left(\dfrac{\partial L_1(\bar{\mu}_1,\bar{k}_1)}{\partial \bar{k}_1}+\varepsilon\right) \\ \qquad -\left(a - 2b(T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha)+\bar{\lambda}\right)\dfrac{\partial T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)}{\partial \bar{k}_1} \\ H_2\left(\dfrac{\partial L_2(\bar{\mu}_2,\bar{k}_2)}{\partial \bar{k}_2}+\varepsilon\right) \\ \qquad -\left(a - 2b(T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha)+\lambda\right)\dfrac{\partial T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)}{\partial \bar{k}_2} \\ T(\bar{\mu}_1,\bar{\mu}_2,\bar{k}_1,\bar{k}_2)-\alpha \end{pmatrix}.$$

First of all, we apply Lemma 15 to the NCP defined by $\Phi \equiv F^\varepsilon$, thus showing that the NCP: $0 \leq \bar{x} \perp F^\varepsilon(\bar{x}) \geq 0$ has a solution for each fixed but arbitrary $\varepsilon > 0$, where $\bar{x} \equiv (\bar{\mu}_1, \bar{\mu}_2, \bar{k}_1, \bar{k}_2, \bar{\lambda})$. We take an arbitrary sequence of positive scalars $\{\tau_\nu\}$;

for each τ_ν, let $(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu, \bar\lambda^\nu)$ be a tuple satisfying

$$\begin{aligned}
0 \le \bar\mu_1^\nu \quad \perp \quad & C_1 - R(\alpha)\,\frac{C_1\,C_2\,(\bar\mu_2^\nu + \alpha)}{(C_1(\bar\mu_1^\nu+\alpha)\,+\,C_2(\bar\mu_2^\nu+\alpha))^2} \\
& - \left(a_1 - 2b_1(T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu) - \alpha) + \bar\lambda^\nu\right) \frac{\partial T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu)}{\partial \bar\mu_1^\nu} + \tau_\nu \bar\mu_1^\nu \ge 0 \\
0 \le \bar\mu_2^\nu \quad \perp \quad & C_2 - R(\alpha)\frac{C_1\,C_2\,(\bar\mu_1^\nu + \alpha)}{(C_1(\bar\mu_1^\nu+\alpha)\,+\,C_2(\bar\mu_2^\nu+\alpha))^2} \\
& - \left(a_2 - 2b_2(T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu) - \alpha) + \bar\lambda^\nu\right) \frac{\partial T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu)}{\partial \bar\mu_2^\nu} + \tau_\nu \bar\mu_2^\nu \ge 0 \\
0 \le \bar k_1^\nu \quad \perp \quad & H_1\left(\frac{\partial L_1(\bar\mu_1^\nu, \bar k_1^\nu)}{\partial \bar k_1^\nu} + \varepsilon\right) \\
& - \left(\,a - 2b(T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu) - \alpha) + \bar\lambda^\nu\,\right) \frac{\partial T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu)}{\partial \bar k_1^\nu} + \tau_\nu \bar k_1^\nu \ge 0 \\
0 \le \bar k_2^\nu \quad \perp \quad & H_2\left(\frac{\partial L_2(\bar\mu_2^\nu, \bar k_2^\nu)}{\partial \bar k_2^\nu} + \varepsilon\right) \\
& - \left(\,a - 2b(T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu) - \alpha) + \bar\lambda^\nu\,\right) \frac{\partial T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu)}{\partial \bar k_2^\nu} + \tau_\nu \bar k_2^\nu \ge 0 \\
0 \le \bar\lambda^\nu \quad \perp \quad & T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu) - \alpha + \tau_\nu \bar\lambda^\nu \ge 0
\end{aligned}$$

We claim that under Assumptions 8–11, the sequence $\{(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu, \bar\lambda^\nu)\}$ is bounded. We prove this in the following order: first the sequence $\{\min(\bar\mu_1^\nu, \bar\mu_2^\nu)\}$; next the sequence $\{\bar\lambda^\nu\}$; then the sequence $\{\bar k_i^\nu\}$, $i = 1, 2$; and finally the sequence $\{\max(\bar\mu_1^\nu, \bar\mu_2^\nu)\}$.

Boundedness of $\{\min(\bar\mu_1^\nu, \bar\mu_2^\nu)\}$. Assume for the sake of contradiction that $\{\min(\bar\mu_1^\nu, \bar\mu_2^\nu)\}$ is unbounded. For an infinite index set $\mathcal{I} \subset \{1, 2, ..., \infty\}$, we have

$$\lim_{\nu(\in\mathcal{I})\to\infty} \min(\bar\mu_1^\nu, \bar\mu_2^\nu) = \infty. \tag{2.16}$$

Without loss of generality, we assume that $\bar\mu_1^\nu = \min(\bar\mu_1^\nu, \bar\mu_2^\nu) > 0$, for all $\mu \in \mathcal{I}$. It immediately follows by Assumption 8 that

$$\lim_{\nu(\in\mathcal{I})\to\infty} T(\bar\mu_1^\nu, \bar\mu_2^\nu, \bar k_1^\nu, \bar k_2^\nu) = \infty. \tag{2.17}$$

Thus, we may assume that $\bar{\lambda}^\nu = 0$ for all $\nu \in \mathcal{I}$ by the throughput constraint. It follows by complementarity that

$$C_1 - R(\alpha) \frac{C_1 C_2 (\bar{\mu}_2^\nu + \alpha)}{(C_1(\bar{\mu}_1^\nu + \alpha) + C_2(\bar{\mu}_2^\nu + \alpha))^2} - \left(a_1 - 2b_1(T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu) - \alpha)\right) \frac{\partial T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)}{\partial \bar{\mu}_1^\nu} + \tau_\nu \bar{\mu}_1^\nu = 0$$

which yields, by Assumption 9,

$$\beta \left(2b_1(T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu) - \alpha) - a_1 \right) \leq R(\alpha) \frac{C_1 C_2 (\bar{\mu}_2^\nu + \alpha)}{(C_1(\bar{\mu}_1^\nu + \alpha) + C_2(\bar{\mu}_2^\nu + \alpha))^2}$$

The right-hand side of above inequality tends to zero as $\nu(\in \mathcal{I}) \to \infty$, since the limits of $\bar{\mu}_1^\nu$ and $\bar{\mu}_2^\nu$ goes to infinity as $\nu(\in \mathcal{I}) \to \infty$. Hence we obtain a contradiction to (2.17). Therefore, $\{\min(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu)\}$ is bounded. As a result, $\{T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)\}$ is bounded as well, under Assumption 8.

Boundedness of $\{\bar{\lambda}^\nu\}$. Assume for the sake of contradiction that $\{\bar{\lambda}^\nu\}$ is unbounded and for an infinite set $\mathcal{I} \subset \{1, 2, ..., \}$,

$$\lim_{\nu(\in \mathcal{I}) \to \infty} \bar{\lambda}^\nu = \infty. \tag{2.18}$$

Without loss of generality, we may assume that $\bar{\lambda}^\nu > 0$ for all $\nu \in \mathcal{I}$. It follows by complementarity that

$$\tau_\nu \bar{\lambda}^\nu = \alpha - T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)$$

Since the right-hand side is bounded, by (2.18), it follows that

$$\lim_{\nu(\in \mathcal{I}) \to \infty} \tau_\nu = 0. \tag{2.19}$$

Without loss of generality, we may assume $\bar{\mu}_1^\nu = \min(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu)$, for all $\nu \in \mathcal{I}$. We have

$$C_1 - R(\alpha) \frac{C_1 C_2 (\bar{\mu}_2^\nu + \alpha)}{(C_1(\bar{\mu}_1^\nu + \alpha) + C_2(\bar{\mu}_2^\nu + \alpha))^2} - \left(a_1 - 2b_1(T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu) - \alpha) + \bar{\lambda}^\nu\right) \frac{\partial T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)}{\partial \bar{\mu}_1^\nu} + \tau_\nu \bar{\mu}_1^\nu \geq 0$$

which yields, by Assumption 9,

$$\beta\left(a_1 - 2b_1\left(T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu) - \alpha\right) + \bar{\lambda}^\nu\right) \leq C_1 + \tau_\nu \bar{\mu}_1^\nu$$

Because of (2.18) and the boundedness of $\{T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)\}$, the left-hand side of the above inequality tends to infinity as $\nu(\in \mathcal{I}) \to \infty$. On the other hand, the right-hand side is bounded by (2.19) and the boundedness of $\{\min(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu)\}$. We obtain a contradiction. Therefore, $\{\bar{\lambda}^\nu\}$ is bounded.

Boundedness of $\{\bar{k}_i^\nu\}$**,** $i = 1, 2$. Assume for the sake of contradiction that $\{\bar{k}_1^\nu\}$ is unbounded and for an infinite set $\mathcal{I} \subset \{1, 2, ...,\}$,

$$\lim_{\nu(\in\mathcal{I})\to\infty} \bar{k}_1^\nu = \infty. \tag{2.20}$$

Without loss of generality, we may assume that $\bar{k}_1^\nu > 0$ for all $\nu \in \mathcal{I}$. It follows by complementarity that

$$H_1\left(\frac{\partial L_1(\bar{\mu}_1^\nu, \bar{k}_1^\nu)}{\partial \bar{k}_1^\nu} + \varepsilon\right) - \left(a - 2b(T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu) - \alpha) + \bar{\lambda}^\nu\right)\frac{\partial T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)}{\partial \bar{k}_1^\nu} + \tau_\nu \bar{k}_1^\nu = 0.$$

Since $L_1(\bar{\mu}_1^\nu, \bullet)$ is monotonically increasing i.e. $\dfrac{\partial L_1(\bar{\mu}_1^\nu, \bar{k}_1^\nu)}{\partial \bar{k}_1^\nu} \geq 0$, the above equality yields

$$H_1\varepsilon \leq (a - 2b(T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu) - \alpha) + \bar{\lambda}^\nu)\frac{\partial T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)}{\partial \bar{k}_1^\nu}.$$

Because of the boundedness of $\{T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)\}$ and $\{\bar{\lambda}^\nu\}$, the right-hand side of above inequality tends to zeros as $\nu(\in \mathcal{I}) \to \infty$ by Assumption 11, which is a contradiction. Therefore, $\{\bar{k}_1^\nu\}$ is bounded. Similarly, it can be shown that $\{\bar{k}_2^\nu\}$ is bounded as well.

Boundedness of $\{\max(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu)\}$. Assume for the sake of contradiction that $\{\max(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu)\}$ is unbounded and for an infinite set $\mathcal{I} \subset \{1, 2, ...,\}$,

$$\lim_{\nu(\in\mathcal{I})\to\infty} \max(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu) = \infty. \tag{2.21}$$

Without loss of generality, we assume $\bar{\mu}_2^\nu = \max(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu) > 0$ for all $\nu \in \mathcal{I}$. It follows by complementarity that

$$C_2 - R(\alpha)\frac{C_1\,C_2\,(\bar{\mu}_1^\nu + \alpha)}{(C_1(\bar{\mu}_1^\nu + \alpha)\; + \; C_2(\bar{\mu}_2^\nu + \alpha))^2} \\ - \left(a_2 - 2b_2(T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu) - \alpha) + \bar{\lambda}^\nu\right)\frac{\partial T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)}{\partial \bar{\mu}_2^\nu} + \tau_\nu \bar{\mu}_2^\nu = 0$$

which yields

$$\begin{aligned} C_2 \;\; \leq \;\; & R(\alpha)\frac{C_1\,C_2\,(\bar{\mu}_1^\nu + \alpha)}{(C_1(\bar{\mu}_1^\nu + \alpha)\; + \; C_2(\bar{\mu}_2^\nu + \alpha))^2} \\ & + \left(a_2 - 2b_2(T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu) - \alpha) + \bar{\lambda}^\nu\right)\frac{\partial T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)}{\partial \bar{\mu}_2^\nu} \end{aligned}$$

Because $\{\min(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu)\}$, $\{T(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu)\}$ and $\{\bar{\lambda}^\nu\}$ are bounded, together with the limit in (2.21), the right-hand side tends to zero as $\nu(\in \mathcal{I}) \to \infty$ by Assumption 10. We obtain a contradiction. Therefore, $\{\max(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu)\}$ is bounded.

Summarizing, we deduce that the sequence $\{(\bar{\mu}_1^\nu, \bar{\mu}_2^\nu, \bar{k}_1^\nu, \bar{k}_2^\nu, \bar{\lambda}^\nu)\}$ is bounded, under Assumptions 8–12. By Lemma 15, there exists a solution to the NCP $0 \leq$

$\bar{x} \perp F^{\varepsilon}(\bar{x}) \geq 0$ for all $\varepsilon > 0$, namely $(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon}, \bar{\lambda}^{\varepsilon})$. For each $\varepsilon > 0$, we have

$$
\begin{aligned}
0 \leq \bar{\mu}_1^{\varepsilon} \quad \perp \quad & C_1 - R(\alpha) \frac{C_1\, C_2\, (\bar{\mu}_2^{\varepsilon} + \alpha)}{\left(C_1(\bar{\mu}_1^{\varepsilon} + \alpha) \, + \, C_2(\bar{\mu}_2^{\varepsilon} + \alpha)\right)^2} \\
& - \left(a_1 - 2b_1(T(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon}) - \alpha) + \bar{\lambda}^{\varepsilon}\right) \frac{\partial T(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon})}{\partial \bar{\mu}_1^{\varepsilon}} \geq 0 \\
0 \leq \bar{\mu}_2^{\varepsilon} \quad \perp \quad & C_2 - R(\alpha) \frac{C_1\, C_2\, (\bar{\mu}_1^{\varepsilon} + \alpha)}{\left(C_1(\bar{\mu}_1^{\varepsilon} + \alpha) \, + \, C_2(\bar{\mu}_2^{\varepsilon} + \alpha)\right)^2} \\
& - \left(a_2 - 2b_2(T(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon}) - \alpha) + \bar{\lambda}^{\varepsilon}\right) \frac{\partial T(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon})}{\partial \bar{\mu}_2^{\varepsilon}} \geq 0 \\
0 \leq \bar{k}_1^{\varepsilon} \quad \perp \quad & H_1(\frac{\partial L_1(\bar{\mu}_1^{\varepsilon}, \bar{k}_1^{\varepsilon})}{\partial \bar{k}_1^{\varepsilon}} + \varepsilon) - (a - 2b(T(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon}) - \alpha) + \bar{\lambda}^{\varepsilon}) \\
& \qquad \frac{\partial T(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon})}{\partial \bar{k}_1^{\varepsilon}} \geq 0 \\
0 \leq \bar{k}_2^{\varepsilon} \quad \perp \quad & H_2(\frac{\partial L_2(\bar{\mu}_2^{\varepsilon}, \bar{k}_2^{\varepsilon})}{\partial \bar{k}_2^{\varepsilon}} + \varepsilon) - (a - 2b(T(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon}) - \alpha) + \bar{\lambda}^{\varepsilon}) \\
& \qquad \frac{\partial T(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon})}{\partial \bar{k}_2^{\varepsilon}} \geq 0 \\
0 \leq \bar{\lambda}^{\varepsilon} \quad \perp \quad & T(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon}) - \alpha \geq 0
\end{aligned}
$$

By the same proof sequence as before, we can show that $\limsup_{\varepsilon \downarrow 0} \|(\bar{\mu}_1^{\varepsilon}, \bar{\mu}_2^{\varepsilon}, \bar{k}_1^{\varepsilon}, \bar{k}_2^{\varepsilon}, \bar{\lambda}^{\varepsilon})\| < \infty$. Thus the sequence $\{(\bar{\mu}_1^{\varepsilon_\nu}, \bar{\mu}_2^{\varepsilon_\nu}, \bar{k}_1^{\varepsilon_\nu}, \bar{k}_2^{\varepsilon_\nu}, \bar{\lambda}^{\varepsilon_\nu})\}$ must accumulate to a limit, say $(\bar{\mu}_1^*, \bar{\mu}_2^*, \bar{k}_1^*, \bar{k}_2^*, \bar{\lambda}^*)$. It is obvious that $(\bar{\mu}_1^*, \bar{\mu}_2^*, \bar{k}_1^*, \bar{k}_2^*, \bar{\lambda}^*)$ is a solution to the three-player NCP $0 \leq x \perp F(x) \geq 0$ with a common multiplier $\lambda_1 = \lambda_2 = \lambda_a = \bar{\lambda}^*$ to the same throughout constraint $T(\mu_1, \mu_2, k_1, k_2) \geq \alpha$.

2.5 Numerical Analysis

In this section, several sets of numerical experiments are provided. The outline for this section is as follows. In subsection 2.5.1, a sequence of experiments are designed to get insights regarding the equilibria's sensitivity to input parameters. In subsection 2.5.2, the three-player game and the two-player game are compared under a symmetric and an asymmetric parameters setting. Finally, subsection 2.5.3 presents a couple of examples to illustrate the advantage of incorporating queueing analysis and game theoretical analysis in the assembly supply chain. The NCPs are solved by using the PATH solver available on the NEOS server (http://neos.mcs.anl.gov/neos/solvers/index.html).

2.5.1 The effects of confirmed order α

In the three-player game, we are particularly interested in the influence of the confirmed order α on the system's performance. An assemble-supply system with the following characteristics is considered.

• Define $\tilde{p}_i(\bullet)$, $i = 1, 2, a$, as supplier 1 and 2's components price functions and assembler's final product's price functions. Let $\tilde{p}_1(T) = 1.20 - 0.012T$, $\tilde{p}_2(T) = 4.80 - 0.048T$ and $\tilde{p}_a(T) = 10.00 - 0.100T$. Thus, the relations between $\tilde{p}(T)$, $i = 1, 2, a$ and the price functions for the excess units $p_i(T - \alpha) = a_i - b_i(T - \alpha)$, $i = 1, 2$, $\widehat{p}(T - \alpha) = \widehat{a} - \widehat{b}(T - \alpha)$ are the following:

$$\begin{aligned} p_1(T-\alpha) &= \tilde{p}_1(\alpha) - 0.012(T-\alpha), \\ p_2(T-\alpha) &= \tilde{p}_2(\alpha) - 0.048(T-\alpha), \\ \widehat{p}\,(T-\alpha) &= \tilde{p}_a(\alpha) - 0.100(T-\alpha). \end{aligned} \tag{2.22}$$

The price functions' parameters imply that $\tilde{p}_1(0) = 1.20$ \$, $\tilde{p}_2(0) = 4.80$ \$ and $\tilde{p}_a = 10.00$ \$ are the highest price for supplier 1' components, supplier 2's components and final products, respectively; and the demand intercepts of the price functions are all equal to 100 units, which is the maximal demand.

• $C_1 = 0.30$ \$/unit and $C_2 = 1.20$ \$/unit, which are a quarter of the suppliers' components highest price 1.20 \$ and 4.80 \$, accordingly. $H_1 = 0.20$ \$/unit, $H_2 = 0.60$ \$/unit.

• Since the given price functions satisfy $\tilde{p}_1(\bullet) + \tilde{p}_2(\bullet) = 0.6\tilde{p}_a(\bullet)$, it is fair for the assembler to allocate 60% of the revenue he gains from selling the first α units order to the suppliers, i.e.

$$R(\alpha) = 60\% \int_0^{\alpha} \tilde{p}_a(t)dt. \tag{2.23}$$

• To investigate the effect of the confirmed order α, we solve a set of the three-player NCPs with the value of α varying from 1 to 90 in increments of 10 in general, assuming a fixed maximal demand 100 units. In turn, it follows by (2.22) and (2.23) that $p_i(\bullet)$, $i = 1, 2$, $\widehat{p}(\bullet)$ and $R(\alpha)$ also change accordingly. The detail is summarized in Table 2.1.

α	1.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00
$R(\alpha)$	5.97	57.00	108.00	153.00	192.00	225.00	252.00	273.00	288.00	297.00
a_1	1.19	1.08	0.96	0.84	0.72	0.60	0.48	0.36	0.24	0.12
a_2	4.75	4.32	3.84	3.36	2.88	2.40	1.92	1.44	0.96	0.48
$\widehat{a}$	9.90	9.00	8.00	7.00	6.00	5.00	4.00	3.00	2.00	1.00

Table 2.1: Parameter setting for the experiments (asymmetric input)

For each of the above ten cases, equilibrium solutions are found by the PATH solver using the throughput and mean queue length expressions (2.11) and (2.13) obtained from Markovian models. The results including the players' decision variables, profits and the critical performance measures are reported in the Table 2.2. In the table, S_i, $i = 1, 2$ stand for two suppliers' profits respectively, which are the objective values at the equilibrium in optimality problems (2.3) and (2.4). The symbol A represents the assembler's profit, which is the objective value in (2.1) plus the constant $\int_0^\alpha \tilde{p}_a(t)dt - R(\alpha)$.

α	μ_1	μ_2	k_1	k_2	T	L_1	L_2	S_1	S_2	A
1	27.60	27.38	14.46	1.00	25.82	14.42	0.96	15.01	60.28	74.16
10	45.79	41.69	12.80	1.00	40.33	12.53	0.76	20.26	81.56	107.43
20	73.35	51.46	7.67	1.00	50.72	7.31	0.61	24.54	90.54	130.73
30	113.61	58.63	5.01	1.00	58.09	4.67	0.49	29.97	89.22	147.87
40	149.40	64.93	4.10	1.00	64.40	3.76	0.38	35.72	85.66	161.76
50	178.84	70.72	3.70	1.00	70.17	3.35	0.29	40.72	81.86	173.22
60	202.21	76.17	3.53	1.00	75.60	3.15	0.21	44.43	78.34	182.47
70	219.73	81.34	3.16	1.36	80.76	2.75	0.24	46.61	75.28	189.59
80	231.46	86.28	1.94	2.69	85.71	1.53	0.47	47.16	72.77	194.68
90	236.84	90.74	1.00	5.22	90.60	0.62	0.22	46.29	71.11	197.97

Table 2.2: Results for the three-player game using Markovian formulas

First, it is noted that in all the cases in Table 2.2, the equilibrium value of throughput is strictly greater than α, implying that at equilibrium, the common throughput constraint is not binding in each players problem. Moreover, the equilibrium solutions are all asymmetric, i.e. $\mu_1 \neq \mu_2$ and $k_1 \neq k_2$, which is to be expected when supplier costs are asymmetric. However, the individual decisions are not simply in the proportion of the corresponding costs.

When $\alpha = 1$, two suppliers' production investments μ_i, $i = 1, 2$ are very close. Because the guaranteed revenue $R(\alpha)$ is only 5.97, both suppliers just keep up with each other so that they can make enough profit for themselves. On the other hand, the assembler hopes to have more throughput, so he sets the total inventory limit $k_1 + k_2$ very large in order to encourage the suppliers to produce more components. In particular, $k_1 = 14.46$ is much higher than $k_2 = 1.00$, which only reaches its minimal inventory limit. This is because given $H_1 = 0.2$ and $H_2 = 0.6$, the assembler would rather have a higher inventory limit for the components with cheaper inventory cost.

As α increases, supplier 1's production investment increases much faster than supplier 2's. For example, when $\alpha = 30$, μ_2 is almost twice as much as μ_1. What happens is that $R(\alpha)$ becomes a big bulk of revenue of the suppliers as α grows. Recall that supplier 1's investment cost is only one quarter of supplier 2's and he can receive $R(\alpha)C_1\mu_1/(C_1\mu_1 + C_2\mu_2)$ amount of guaranteed revenue from the assembler. The lower capacity investment costs of supplier 1 permits him to make significant investments in capacity potentially increase revenues. Since supplier 2's investment cost is much higher, it is more expensive for him to increase μ_2 to receive more guaranteed revenue. Therefore, the high capacity investment cost for supplier 2 limits the capacity investments to levels that allow the suppliers to meet the constraint on throughput. For example, in spite of supplier 1's large production capacity, the throughput T at $\alpha = 30$ is equal to 58.09, which is close to $\mu_2 = 58.63$. Meanwhile, the assembler still keeps a relatively high inventory limit for supplier 1's components and remains the inventory limit for supplier 2's components at its lowest; however, the volume of k_1 decreases. This is because the assembler notices that supplier 1's large production capacity does not really help to improve the final throughput. Moreover, the final products' price $\widehat{p}(\bullet)$ declines as throughput increases, the assembler has to reduce his inventory cost to maximize his total profit.

This situation lasts until $\alpha = 70$, where $p_1(T - \alpha) = 0.23$ falls below $C_1 = 0.30$

and $p_2(T-\alpha) = 0.92$ is less than $C_2 = 1.20$. Because suppliers start losing money when they sell the components additional to α, supplier 1 slows down his investment on production capacity. For the cases with $\alpha \geq 70$, supplier 2 merely produces enough components to meet the minimal demand α. From assembler's point of view, he finds out that supplier 1 is not willing to increase his capacity too much, so he cuts down k_1. On the other hand, the assembler begins raising the inventory limit for the supplier 2's components and hopes supplier 2 to increase μ_2, so that the total throughput will grow up.

The profit figures indicate that the relative increase in profit is different for each player, and appears to strongly depend on the cost structure. The assembler's profit keeps going up as α increases. Supplier 1's profit develops dramatically before $\alpha = 70$, due to the benefit he obtains from the guaranteed revenue. After $\alpha = 70$, his profit stays around 46, since μ_1 does not vary a lot. Supplier 2's profit improves before $\alpha = 20$. After that, his profit gradually decreases as α increases.

2.5.2 Compare the two-player game with the three-player game

In this subsection, numerical results for the two-player game where the two suppliers collude are examined. The NCP for the two player game is discussed in Section 2.3. We did two sets of experiments under asymmetric and symmetric parameters input, respectively. The results for the two player game provide insights on the effect of collusion among the suppliers.

For all the asymmetric input experiments, the input setting and experiment design are remained the same as the ones in subsection 2.5.1. Instead of solving a set of three-player NCPs, the PATH solver works on the two-player NCP using the throughput function and average inventory functions approximated by Markovian model. The results are listed in Table 2.3. To distinct from notations used in the three-player game, we use S_1', S_2' and A' to indicate the suppliers' profits and assembler's profits in the two-player game, and they are still calculated in the same way as in the three-player game. It is still assumed that the guaranteed revenue $R(\alpha)$ is divided between two suppliers by the investment cost allocation rule in the centralized suppliers game.

α	μ_1	μ_2	k_1	k_2	T	L_1	L_2	S'_1	S'_2	A'
1	44.47	38.30	10.78	1.00	37.23	10.76	0.97	15.29	67.82	92.21
10	49.46	42.90	11.47	1.00	41.77	11.22	0.77	20.11	81.56	109.29
20	54.96	48.00	12.20	1.00	46.81	11.65	0.58	24.66	94.81	126.36
30	60.43	53.09	12.90	1.00	51.85	11.98	0.43	28.39	105.89	141.42
40	65.86	58.18	13.59	1.00	56.88	12.20	0.31	31.33	114.77	154.48
50	71.24	63.25	14.28	1.00	61.91	12.26	0.21	33.51	121.43	165.56
60	76.56	68.31	14.51	1.53	66.93	11.61	0.27	34.94	125.88	174.69
70	81.66	73.35	13.07	4.08	71.97	8.96	0.67	35.60	128.19	182.02
80	89.08	81.27	10.15	10.26	80.00	5.43	1.44	35.22	128.53	190.05
90	98.92	91.24	11.57	11.85	90.00	6.18	1.66	33.66	124.18	195.77

Table 2.3: Results for the two-player game using Markovian formulas (asymmetric input)

Unlike in the three-player game, supplier 1 and supplier 2's production capacities are close in the two-player game. Since the guaranteed revenue $R(\alpha)$ appears as an constant in the suppliers' objective function in the two-player game, supplier 1 does not have incentive to overinvest in his capacity to receive more guaranteed revenue. When the suppliers collude together, the throughput in all the cases with $\alpha \geq 20$ is mildly less than the one in the game where they act selfishly. This is because supplier 1 is not as aggressive as he is in the three-player game and supplier 2's production capacity is a little smaller, it results lower throughput. Unlike in the three-player game, the assembler keeps adding the inventory limit for supplier 1's components until α goes up to 60. However, the larger k_1 does not seem to be good enough to push up the final throughput, so the assembler starts increasing k_2 dramatically at $\alpha = 70$ and reducing k_1.

To compare the two-player game and the three-player game, we subtract each player's profit (S_i, $i = 1, 2$ and A) and the system's profit ($S_1 + S_2 + A$) in the three-player game from their counterparts in the two-player game. The resulting differences are denoted by $\Delta Profit(S_i)$, $i = 1, 2$, $\Delta Profit(A)$ and $\Delta Profit(S)$, respectively; and they are plotted in Figure 2.2. Except $\alpha = 1$, in most of the cases, supplier 1 and the assembler's profits decline in the two-player game, as shown in Figure 2.2 that $\Delta Profit(S_1)$ and $\Delta Profit(A)$ are less than 0. This variation is exaggerated as $R(\alpha)$ becomes larger. Due to the higher inventory cost caused by larger $k_1 + k_2$, the assembler's profit is worse off in the two-player game. Since

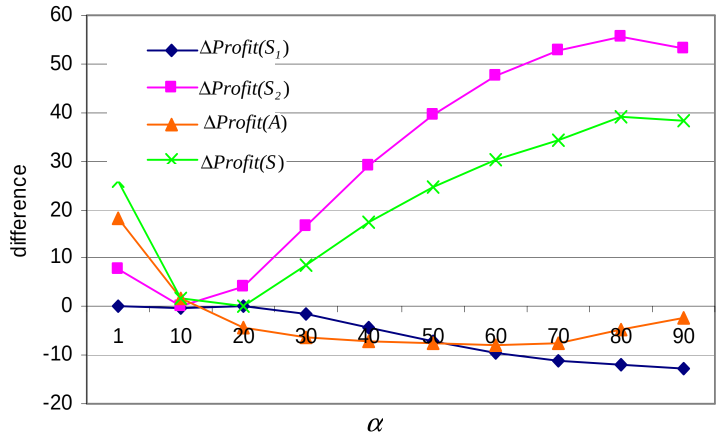

Figure 2.2: Compare the two-player game and the three-player game (asymmetric input)

supplier 1 loses his advantage from the guaranteed revenue in the two-player game, his profit also reduces. Supplier 2 is the only one that does better in the two-player game. This is because supplier 1's production capacity is deflated by almost an order of magnitude in collusive game, consequently the guaranteed revenue $R(\alpha)$ allocated to supplier 2 increases. Regarding the system profit difference, the overall profit is higher in the two-player game than the corresponding one in the three-player game, although the throughput is slightly less in this collusive configuration than the one in completely decentralized system. This is because μ_1 is much smaller in the two-player game than the one in the three-player game, it results lower total system cost.

In the set of experiments with symmetric input setting, all the parameters regarding supplier 1 and supplier 2 are the same. In particular, $\tilde{p}_1(T) = \tilde{p}_2(T) = 3.00 - 0.03T$, $C_1 = C_2 = 0.75$ \$/unit and $H_1 = H_2 = 0.40$ \$/unit. The final product price $\tilde{p}_a(T) = 10.00 - 0.1T$ is remained the same, so the maximal demand is still 100 units and $\tilde{p}_1(\bullet) + \tilde{p}_2(\bullet)$ is equal to 60% of $\tilde{p}_a(\bullet)$. As α varies from 1 to 90, the parameters for the price functions $p_i(T - \alpha)$, $i = 1, 2$, $\widehat{p}(T - \alpha)$ and the guaranteed revenue $R(\alpha)$, which are determined by (2.22) and (2.23), respectively, are summarized in Table 2.4.

Under the symmetric input setting, the equilibrium solutions are symmetric as well. The differences of the profits between the two-player game and three-

α	1	10	20	30	40	50	60	70	80	90
$R(\alpha)$	1.791	17.1	32.4	45.9	57.6	67.5	75.6	81.9	86.4	89.1
a_1	2.97	2.7	2.4	2.1	1.8	1.5	1.2	0.9	0.6	0.3
a_2	2.97	2.7	2.4	2.1	1.8	1.5	1.2	0.9	0.6	0.3
$\widehat{a}$	9.9	9	8	7	6	5	4	3	2	1

Table 2.4: Parameter setting for the experiments (symmetric input)

player game are plotted in Figure 2.3. Unlike in Figure 2.2, all the players' profits improve when two suppliers collaborate. Therefore, depending on the inputs, the collusion of two suppliers might hurt some players' profit, while it might help all the players to gain more benefits. Furthermore, the differences of the profits diminish as α increases; nevertheless, the distortion caused by suppliers' cooperation expands under the asymmetric inputs. Thus, increasing α does not necessarily amplify the difference between centralized-suppliers game and completely decentralized game.

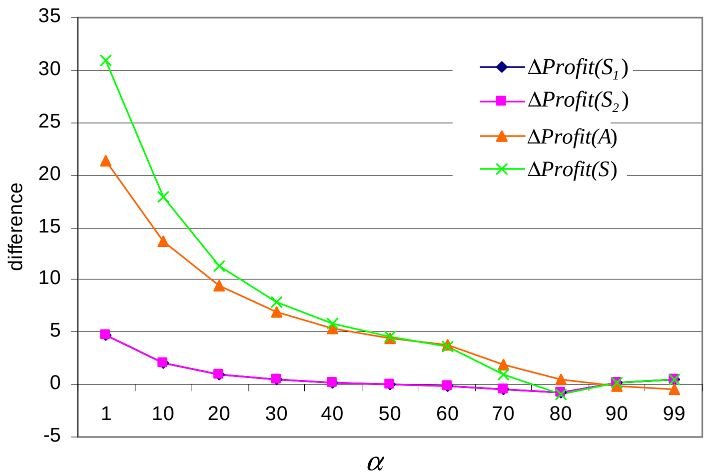

Figure 2.3: Compare the two-player game and the three-player game (symmetric input)

2.5.3 Compare non-queuing and queueing approximations

In this subsection, a couple of examples are provided to show that the non-queuing approximations for the throughput and the average inventory may mislead the players' decisions, result in under-producing and incur lost sales. Table 2.5 lists the parameters setting for two examples.

Under these two sets of inputs, the three-player NCPs are solved using the

Case	a_1	a_2	$\widehat{a}$	b_1	b_2	$\widehat{b}$	α	$R(\alpha)$	C_1	C_2	H_1	H_2
1	0.24	0.96	2.00	0.012	0.048	0.100	80.00	288.00	0.30	1.20	0.20	0.60
2	0.12	0.48	1.00	0.012	0.048	0.100	90.00	297.00	0.30	1.20	0.20	0.60

Table 2.5: Parameter setting for the experiments using non-queuing approximation

throughput function (2.9) and the average inventory function (2.10), which are derived from upper bounds in Section 2.2. The results for each input setting are labelled as case N_1 and case N_2 respectively in Table 2.6. Similarly, using the simple probabilistic approximations (2.14) and (2.15), the results are marked by S_1 and S_2, while M_1 and M_2 imply the results obtain by Markovian expressions (2.11) and (2.13). Given each set of the decision variables (μ_1, μ_2, k_1, k_2), the corresponding values of the exact throughput under Markovian assumptions can be evaluated by the expression (2.11). The actual throughput values are listed at T^* column in Table 2.6. S_1^*, S_2^* and A^* represent the players' actual profits which are evaluated by T^* instead of T.

Case	μ_1	μ_2	k_1	k_2	T	T^*	L_1	L_2	S_1^*	S_2^*	A^*
N1	230.74	85.89	1.00	1.00	85.89	78.02	0.65	0.07	45.96	67.13	190.08
N2	236.54	90.65	1.00	1.00	90.65	81.95	0.62	0.01	44.55	63.99	192.06
S1	230.88	85.67	15.23	1.00	85.67	85.67	14.71	0.07	47.64	74.16	192.27
S2	236.70	90.43	16.19	1.00	90.43	90.43	15.59	0.004	46.52	71.21	195.04
M1	231.46	86.28	1.94	2.69	85.71	/	1.53	0.47	47.16	72.77	194.68
M2	236.84	90.74	1.00	5.22	90.60	/	0.62	0.22	46.29	71.11	197.97

Table 2.6: Compare the effect of the queuing approximations and the non-queuing approximations

As with the Markovian model, it is observed from the equilibrium solutions that the throughput constraint is not binding in the players problems. In case N_1 and N_2, it is noticeable that all the k_i's are equal to 1 at the equilibrium, which does not happen in the other cases. This is because that the throughput defined by (2.9) is only determined by μ_1 and μ_2 and ignores the effect of the assembler's decision variables k_i, $i = 1, 2$ on throughput, which is clearly not very realistic. In addition, the value of T is larger than the accurate throughput T^* in both cases, which causes the failure to meet the confirmed demand. For instance, in case N_1

where $\alpha = 80$, if the players employed the formulas based on upper bounds, then they would expect to have have enough throughput $T = 85.89$ to meet the minimal order $\alpha = 80$. Unfortunately, the exact throughput T^* is only 78.02, less than the required final products. Furthermore, all the players' actual profits are smaller than the corresponding ones obtain in Markovian model. This means that using the formulas based on upper bounds in model investigation might mislead the players' judgement of their equilibrium decisions, which is the drawback of the conventional decentralized supply chain analysis. In case S_1 and S_2, the deviation is not as sever as the one in case N_1 and N_2. This suggests that if exact Markovian analysis leads to expressions that are complex, there is scope for developing a class of simple probabilistic approximations that provide good performance estimates as well as could lead to realistic equilibrium decisions.

2.6 Conclusion

In this chapter, queuing models and noncooperative game theory are integrated to analyze stochastic assembly systems. This combination helps to correct the error caused by overlooking the stochastic elements in the assembly supply chain systems. To take a detailed look at the effect of incorporating queuing models, a system containing one assembler and two supplier is studied. A three-player game and a two-player game are formulated, given that two suppliers operate in a decentralized fashion or a centralized fashion, respectively. Under some concavity/convexity assumptions, both games are equivalent to generalized Nash equilibrium problems. The existence of solutions was established under mild conditions. Numerical experiments illustrates the advantage of embedding queuing theory in assembly supply chain analysis. A sequence of numerical results also provide marginal insight on the difference between the three-player game and the two-player game and the effect of the confirmed order.

In the future, the current one assembler and two suppliers model could be extended to more general supply chain network models, for instance, a system involves more multiple products, components, suppliers and assemblers. Additionally, some realistic features such as lead time constraints and uncertain demands, can be added to the models. From the game theory perspective, it would be interesting to study leader-follower games.

CHAPTER 3
EMISSIONS ALLOWANCE ALLOCATION SYSTEMS

In this chapter, we focus on the effects of emissions allocation rules on electric power markets. A nonlinear complementarity model is proposed to evaluate the long-run implications of different allocation schemes for economic efficiency and consumer costs of the electric power sector. This chapter is organized as follows. The model is stated and two different emission allowance allocation rules are presented in Section 4.1. Nonlinear complementarity models are set up in Section 3.2. The existence of equilibria is established under general conditions in Section 3.3. Several sets of numerical results are obtain by using PATH solver, and the efficiency of two different rules are discussed in Section 3.4. The conclusion and future work are provided in Section 3.5.

3.1 Model Definition

The notations used in the model formulation are summarized in the following table. The physical units are noted within parentheses. Figure 3.1 illustrates the interaction among the involved markets.

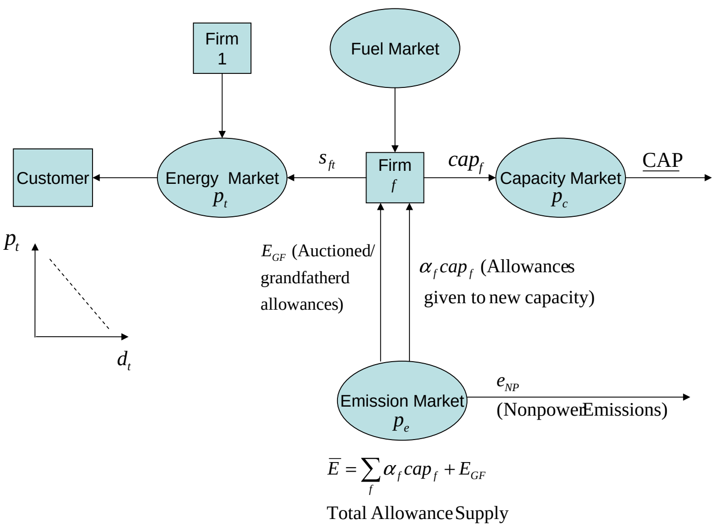

Figure 3.1: Market structure

Parameters: all positive except possibly CAP_f and $\underline{\text{CAP}}$ which can be zero,

$\mathcal{F}$	Set of firms
$\mathcal{T}$	Set of time periods $\equiv \{1, \cdots, T\}$
CAP_f	Minimal amount of energy that firm f has to generate (MW)
MC_f	Marginal cost for firm f, excluding cost of emission allowances (EURO/MWh)
E_f	Emission rate for firm f (tons/MWh)
F_f	Annualized investment cost of firm f's capacity (EURO/MWyr)
R_f	Ratio of allowances allocated to firm f per unit of capacity relative to firm 1
$\widehat{R}_f$	Ratio of allowances allocated to firm f per weighted unit of sales relative to firm 1
$\overline{E}$	Total emission allowances supply (tons/yr): $\overline{E} > E_{\text{GF}}$
E_{GF}	Amount of emission allowances that are grandfathered or auctioned (tons/yr)
H_t	Hours in period t (hr/yr), here assumed to be $8760/T$
χ	Unit converter = 1 MW^2 yr/EURO
$\underline{\text{CAP}}$	Total capacity requirement (MW)

The restriction $\overline{E} > E_{\text{GF}}$ means that a certain volume of free allowances is guaranteed to new entrants. For example, In EU ETS phase I, a fraction of the allowances have been reserved for eligible new entrants.

Functions:

$d_t(\bullet)$	Demand function for energy, strictly decreasing (MW)
$\pi_t(\bullet)$	The inverse of $d_t(\bullet)$; (EURO/MWh)
$e_{\text{NP}}(\bullet)$	The sales from nonpower sector to power sector, nonincreasing (tons/yr)

Variables:

p_t	$= \pi_t\left(\sum_{g\in\mathcal{F}} s_{gt}\right)$: Energy price during period t (EURO/MWh)
p_e	Emission allowance price (EURO/ton)
p_c	Capacity price (EURO/MW/yr)
α_f	Emission allowance for firm f (tons/MWyr)
s_{ft}	Energy sold by firm f in period t (MW)
$\bar{s}_{ft}$	$= s_{ft} - \mathrm{CAP}_f$ (MW)
cap_f	Capacity for firm f (MW)
μ_{ft}	Dual variable associated with firm f's capacity constraint in period t (EURO/MWyr)

The model of this chapter is characterized by the following three key factors: (a) firms' profit maximization problems, (b) market clearing conditions for the emission, capacity, and energy markets, and (c) allowances allocation rules. Each of these factors is elaborated in detail below.

3.1.1 Firms' optimization problems

Without loss of generality, it is assumed that each firm invests in only one type of generating capacity. The proportion of allowances that firm f receives per MW of its new investment is reflected by α_f. Indeed, the number of allowances allocated to firm f is related to his generating capacity, according to the emission allocation rules described later. However, as an individual, firm f does not realize that his behavior would affect his granted allowances, so he treats α_f as exogenous. Moreover, firm f is also assumed to be a price-taker, meaning that the prices (for energy p_t for $t \in \mathcal{T}$, emission allowances p_e, capacity p_c) are exogenous, too. Thus, firm f's optimization problem is to make decision on his capacity cap_f and sales s_{ft} so that the revenue less cost is maximized:

$$\begin{aligned} \underset{\mathrm{cap}_f,\,(s_{ft})_{t\in\mathcal{T}}}{\text{maximize}} \quad & \sum_{t\in\mathcal{T}} H_t\,(\,p_t - \mathrm{MC}_f - p_e\,E_f\,)\,s_{ft} + (\,p_c + p_e\alpha_f - F_f\,)\,\mathrm{cap}_f \\ \text{subject to} \quad & \mathrm{CAP}_f \,\le\, s_{ft} \,\le\, \mathrm{cap}_f, \quad \forall\, t \,\in\, \mathcal{T} \end{aligned} \tag{3.1}$$

The problem (3.1) is a linear program whose optimality conditions are straightforward to write down:

$$
\begin{array}{lll}
0 \le \bar{s}_{ft} & \perp \; H_t\,(\,-p_t + \mathrm{MC}_f + p_e\,E_f\,) + \mu_{ft} \;\ge\; 0, & \forall\, t \,\in\, \mathcal{T} \\
0 \le \mu_{ft} & \perp \; \mathrm{cap}_f - \bar{s}_{ft} - \mathrm{CAP}_f \;\ge\; 0, & \forall\, t \,\in\, \mathcal{T} \\
0 \le \mathrm{cap}_f & \perp \; -p_c - p_e\,\alpha_f + F_f - \sum_{t \in \mathcal{T}} \mu_{ft} \;\ge\; 0. &
\end{array}
\tag{3.2}
$$

If firms are able to alter the energy price (market power case), their revenues from energy sales change from linear to nonlinear functions of the sales variables:

$$
\sum_{t \in \mathcal{T}} H_t\, s_{ft}\, p_t \;\longrightarrow\; \sum_{t \in \mathcal{T}} H_t\, s_{ft}\, p_t \left(\sum_{g \in \mathcal{F}} s_{gt} \right).
$$

Thus, each firm's optimization problem becomes a nonlinear convex program. The corresponding optimality conditions are similar to (3.2), except that there is an extra term corresponding to the derivative of the price function $p_t(\bullet)$ with respect to s_{ft} in the first complementarity condition. Moreover, the firms' problem can also be enriched by introducing linear constraints, such as a "min-run capacity constraint" that is of the form $\mathrm{s}_f \ge \gamma_f \mathrm{cap}_f$ for a firm-dependent constant $\gamma_f > 0$. The above refinements can be taken as further research direction. The rest of this chapter focuses on the price-taker case (3.1) and its equivalent optimality conditions (3.2).

3.1.2 Market clearing conditions

The model includes three market clearing conditions. The first is a complementarity condition for the emission allowance price p_e. Allowance price is positive only when demand for allowances equals the available supply:

$$
0 \;\le\; p_e \;\perp\; \overline{E} - e_{\mathrm{NP}}(p_e) - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t\, E_g\, s_{gt} \;\ge\; 0, \tag{3.3}
$$

where the function $e_{\mathrm{NP}}(p_e)$ represents the effective demand for allowances from sectors of the economy other than electric power sector [37]. Condition (3.3) is under the assumption that allowances can be traded among other sectors, which is the realism in EU ETS.

The second market clearing condition is a complementarity condition for the capacity price p_c, which is similar to the first one. Capacity price is positive only when demand for capacity equals the supply:

$$0 \leq p_c \perp \sum_{g \in \mathcal{F}} \text{cap}_g - \underline{\text{CAP}} \geq 0. \tag{3.4}$$

The last is the balance condition stating that energy supplies equal the quantity demanded:

$$\sum_{f \in \mathcal{F}} s_{ft} = d_t(p_t), \text{ for all } t \in \mathcal{T}, \text{ or equivalently, } p_t = \pi_t \left(\sum_{f \in \mathcal{F}} s_{ft} \right), \tag{3.5}$$

which indicates that the energy price p_t is not necessarily nonnegative.

3.1.3 Emissions allocation rules

Two emissions allocation rules are considered: the potential emission rule and the actual emission rule. The former depends on the weighted average of CO_2 emission in terms of firms' capacity (or input). The latter is subject to the weighted average of CO_2 emission in terms of firms' sale (or output). The mathematical formulations of each emission allocation rule are given as follows:

(I) The potential emission rule:

$$\alpha_f \,\text{cap}_f = \frac{R_f \,\text{cap}_f}{\sum_{g \in \mathcal{F}} R_g \,\text{cap}_g} (\overline{E} - E_{\text{GF}}), \quad \forall f \in \mathcal{F}, \tag{3.6}$$

provided that the denominator is positive. In particular, $\overline{E} - E_{\text{GF}}$ equals the total amount of allowances available for allocation. $\alpha_f \,\text{cap}_f$ the emissions allowances allocated to firm f's new capacity is proportional to its potential emissions $R_f \,\text{cap}_f$, which is the maximal amount of CO_2 firm f can emit.

(II) The actual emission rule:

$$\alpha_f \, \text{cap}_f = \frac{\widehat{R}_f \sum_{t \in \mathcal{T}} H_t \, s_{ft}}{\sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_g \, \widehat{R}_g \, s_{gt}} \, (\, \overline{E} - E_{\text{GF}} \,), \quad \forall \, f \, \in \, \mathcal{F}; \tag{3.7}$$

provided that the denominator is positive. Under this rule, the emissions allowances distributed to firm f's new capacity is proportion to its actual emissions $\widehat{R}_f \sum_{t \in \mathcal{T}} H_t \, s_{ft}$, which is the amount of CO_2 emitted from its actual operation. As an extreme case of the actual emission rule, if firm f has no sale, then he will not get any allowance. However, under the potential emission rule, as long as firm f installs the new capacity of a particular type, he will receive a certain amount of allowances before the realization of the actual sale.

Additionally, an emission balance condition must be maintained for both emission allocation rules, so that the amount of allowances available for allocation equals the amount allocated to capacity:

$$\overline{E} - E_{\text{GF}} \; = \sum_{f \in \mathcal{T}} \alpha_f \, \text{cap}_f. \tag{3.8}$$

There are simple conditions ensuring that the denominators in (3.6) and (3.7) are positive, such as $\text{CAP}_f \, > \, 0$ for some $f \, \in \, \mathcal{F}$. Another way to avoid the zero denominators is to rewrite two rules in the following fashion:

$$\alpha_f \, \text{cap}_f \; = \begin{cases} \alpha \, R_f \, \text{cap}_f & \text{for (3.6)} \\ \widehat{\alpha} \, \widehat{R}_f \sum_{t \in \mathcal{T}} H_t \, s_{ft} & \text{for (3.7)} \end{cases} \tag{3.9}$$

for some nonnegative common endogenous variables $\widehat{\alpha}$ to be determined, due to the allowance allocation balance (3.8).

Put together firms' optimality conditions (3.2), market clearing conditions (3.3), (3.4), (3.5) and emission allocation rule (3.9), (3.8), then the following mixed nonlinear complementarity problem is obtained:

$$
\begin{array}{lll}
0 \leq \bar{s}_{ft} & \perp & H_t \left[-\pi_t \left(\sum_{g \in \mathcal{F}} (\bar{s}_{gt} + \text{CAP}_g) \right) + \text{MC}_f + p_e E_f \right] + \mu_{ft} \geq 0 \\
& & \forall \, (f,t) \in \mathcal{F} \times \mathcal{T} \\
0 \leq \mu_{ft} & \perp & \text{cap}_f - \bar{s}_{ft} - \text{CAP}_f \, \geq \, 0 \quad \forall \, (f,t) \in \mathcal{F} \times \mathcal{T} \\
0 \leq \text{cap}_f & \perp & -p_c - p_e \, \alpha_f + F_f - \sum_{t \in \mathcal{T}} \mu_{ft} \, \geq \, 0 \quad \forall \, f \in \mathcal{F} \\
0 \leq p_e & \perp & \overline{E} - e_{\text{NP}}(p_e) - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t \, E_g \, (\, \bar{s}_{gt} + \text{CAP}_g \,) \, \geq \, 0 \\
0 \leq p_c & \perp & \sum_{g \in \mathcal{F}} \text{cap}_g - \underline{\text{CAP}} \, \geq \, 0 \\
& (3.9) \ \text{and} & \sum_{g \in \mathcal{F}} \alpha_g \, \text{cap}_g - (\, \overline{E} - E_{\text{GF}} \,) \, = \, 0.
\end{array}
\tag{3.10}
$$

The model is to find $(\mathbf{x}, \alpha)$ under the potential emission rule (or $(\mathbf{x}, \widehat{\alpha})$ under the actual emission rule), where

$$
\mathbf{x} \, \equiv \, \left\{ \, (s_{ft})_{(f,t) \in \mathcal{F} \times \mathcal{T}} \, , (\mu_{ft})_{(f,t) \in \mathcal{F} \times \mathcal{T}} \, , \left(\text{cap}_f\right)_{f \in \mathcal{F}} \, , (\alpha_f)_{f \in \mathcal{F}} \, , p_e, p_c \right\},
$$

satisfying the mixed NCP (3.10).

3.2 Nonlinear Complementarity Formulation

In order to apply game-theoretic analysis to establish the existence of a solution to the model with the emission rules (I) and (II), we rewrite the mixed NCP (3.10) in the form of a standard NCP.

3.2.1 The potential emission rule

The expression of (3.9) in (3.10) is $\alpha_f \text{cap}_f = \alpha R_f \text{cap}_f$ for all $f \in \mathcal{F}$, under the potential emission rule. Multiply the emission allowance balancing constraint $\sum_{f \in \mathcal{F}} \alpha_f \text{cap}_f = \overline{E} - E_{\text{GF}}$ by the variable p_e and introduce the nonnegative variable $\sigma \equiv \alpha p_e$. Together with (3.9), we get $\sigma \sum_{f \in \mathcal{F}} R_f \text{cap}_f = \overline{E} - E_{\text{GF}}$, which is equivalent to imposing complementarity between σ and the modified emission inequality. Thus, the mixed NCP (3.10) can be reformulated as follows.

$$\begin{array}{lll}
0 \le \bar{s}_{ft} & \perp & H_t \left[-\pi_t \left(\sum_{g\in\mathcal{F}} (\bar{s}_{gt} + \mathrm{CAP}_g) \right) + \mathrm{MC}_f + p_e E_f \right] + \mu_{ft} \ge 0 \\
& & \forall\, (f,t) \in \mathcal{F}\times\mathcal{T} \\
0 \le \mu_{ft} & \perp & \mathrm{cap}_f - \bar{s}_{ft} - \mathrm{CAP}_f \ \ge\ 0 \quad \forall\, (f,t) \in \mathcal{F}\times\mathcal{T} \\
0 \le \mathrm{cap}_f & \perp & -p_c - \sigma\, R_f + F_f - \sum_{t\in\mathcal{T}} \mu_{ft} \ \ge\ 0 \quad \forall\, f \in \mathcal{F} \\
0 \le p_e & \perp & \overline{E} - e_{\mathrm{NP}}(p_e) - \sum_{(g,t)\in\mathcal{F}\times\mathcal{T}} H_t\, E_g\ (\,\bar{s}_{gt} + \mathrm{CAP}_g\,) \ \ge\ 0 \\
0 \le p_c & \perp & \sum_{g\in\mathcal{F}} \mathrm{cap}_g - \underline{\mathrm{CAP}} \ \ge\ 0 \\
0 \le \sigma & \perp & \sigma \sum_{g\in\mathcal{F}} R_g\, \mathrm{cap}_g - (\,\overline{E} - E_{\mathrm{GF}}\,)\, p_e \ \ge\ 0.
\end{array} \tag{3.11}$$

(3.11) is a standard NCP with variable $\mathbf{x}^{\mathrm{I}}$, where

$$\mathbf{x}^{\mathrm{I}} \equiv \left\{ (s_{ft})_{(f,t)\in\mathcal{F}\times\mathcal{T}}\,, (\mu_{ft})_{(f,t)\in\mathcal{F}\times\mathcal{T}}\,, \left(\mathrm{cap}_f\right)_{f\in\mathcal{F}}\,, p_e, p_c, \sigma \right\}.$$

To establish the relation between (3.11) and (3.10), we introduce the following mild condition on the supply of allowances:

$$\overline{E} \ >\ e_{\mathrm{NP}}(0) + \sum_{(g,t)\in\mathcal{F}\times\mathcal{T}} H_t\, E_g\, \mathrm{CAP}_g. \tag{3.12}$$

This condition merely says that the total number of allowances is sufficient to cover the emissions resulting from the sum across firms of the lower bounds upon generation, net of the supply of allowances from other sectors at an allowance price of zero. This is not restrictive, because at a price of zero, it is likely that the supply from other sectors is nonpositive (as other sectors would likely be willing to buy allowances at such a low price), and because the minimum required generation is likely to be a small fraction of total generation, thereby requiring few allowances.

We also impose a condition on firms' investment.

$$\max\left(\underline{\mathrm{CAP}}+\sum_{f\in\mathcal{F}}\mathrm{CAP}_f,\ \chi\max_{f\in\mathcal{F}}\left\{\sum_{t\in\mathcal{T}}H_t\left[\,\pi_t(0)-\mathrm{MC}_f\,\right]-F_f\right\}\right)>0. \quad (3.13)$$

Condition (3.13) implies that if $\underline{\mathrm{CAP}}+\sum_{f\in\mathcal{F}}\mathrm{CAP}_f=0$, then

$$\max_{f\in\mathcal{F}}\left\{\sum_{t\in\mathcal{T}}H_t\left[\,\pi_t(0)-\mathrm{MC}_f\,\right]-F_f\right\}>0,$$

which allows each firm to produce zero power. If no firm sells any power, then the power price will be expected to very high at each time interval. Therefore, it is reasonable to assume that there is at least one firm which will find it profitable to invest; i.e., whose total short-run and investment cost is less than their revenue, on a per MW of investment basis.

The following result summarizes the connection between (3.11) and (3.10) with the emission rule (I) under the following two mild conditions (3.12) and (3.13).

Proposition 16. Under (3.12) and (3.13), if $(\mathbf{x},\alpha)$ is a solution of (3.10) under the potential emission rule, then

$$\sigma\equiv\frac{(\,\overline{E}-E_{\mathrm{GF}}\,)\,p_e}{\sum_{g\in\mathcal{F}}R_g\,\mathrm{cap}_g} \quad (3.14)$$

is well defined and $\mathbf{x}^{\mathrm{I}}$ is a solution of (3.11). Conversely, if $\mathbf{x}^{\mathrm{I}}$ is a solution of (3.11), then

$$\alpha\equiv\frac{\overline{E}-E_{\mathrm{GF}}}{\sum_{g\in\mathcal{F}}R_g\,\mathrm{cap}_g} \quad (3.15)$$

is well defined, and with $\alpha_f\equiv\alpha R_f$ for all $f\in\mathcal{F}$, $(\mathbf{x},\alpha)$ is a solution of (3.10) under the potential emission rule.

Proof. To prove the first statement, let $(\mathbf{x},\alpha)$ be as given. We first show that $\mathrm{cap}_f>0$ for some $f\in\mathcal{F}$. Suppose not, then $\mathrm{cap}_f=s_{ft}=0$ for all $(f,t)\in\mathcal{F}\times\mathcal{T}$, which implies $\underline{\text{CAP}} + \sum_{f \in \mathcal{F}} \text{CAP}_f = 0$. We claim that $p_e = 0$. Indeed, if $p_e > 0$, then

$$0 \;=\; \overline{E} - e_{\text{NP}}(p_e) - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t \, E_g \, \text{CAP}_g \;\geq\; \overline{E} - e_{\text{NP}}(0) - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t \, E_g \, \text{CAP}_g \;>\; 0,$$

which is a contradiction. From the first and third line in (3.10), we deduce, for all $f \in \mathcal{F}$,

$$\sum_{t \in \mathcal{T}} H_t \, [\, -\pi_t(0) + \text{MC}_f \,] + F_f \;\geq\; 0;$$

or equivalently,

$$\max_{f \in \mathcal{F}} \left\{ \sum_{t \in \mathcal{T}} H_t \, [\, \pi_t(0) - \text{MC}_f \,] - F_f \right\} \;\leq\; 0,$$

which contradicts (3.13). Therefore, the scalar σ in (3.14) is well defined; moreover, $\sigma = p_e \alpha$, yielding $\sigma R_f = p_e \alpha_f$. Consequently, (3.11) follows from (3.10). Conversely, let $\mathbf{x}^{\text{I}}$ be a solution of (3.11). By the same argument as before, we deduce that $\text{cap}_f > 0$ for some $f \in \mathcal{F}$. Therefore, the scalar α in (3.15) is well defined; let $\alpha_f \equiv \alpha R_f$. We then have $\sum_{f \in \mathcal{F}} \alpha_f \text{cap}_f = \overline{E} - E_{\text{GF}}$. Consequently, $(\mathbf{x}, \alpha)$ is a solution (3.10) under the emission rule (I). □

It is worth to notice that the NCP (3.11) can be reformulated as a variational inequality (VI) under some conditions. Specially, if $\underline{\text{CAP}} + \sum_{f \in \mathcal{F}} \text{CAP}_f > 0$, then the NCP (3.11) is identical to the set of KKT conditions of the variational inequality (VI) defined by the pair $(K^{\text{I}}, \Phi^{\text{I}})$, where $K^{\text{I}} \equiv K \times \Re_+$ with

$$K \equiv \left\{ \begin{array}{ll} (\, \bar{\mathbf{s}}, \mathbf{cap} \,) \;\geq\; 0 \;: & \displaystyle\sum_{g \in \mathcal{F}} \text{cap}_g - \underline{\text{CAP}} \;\geq\; 0 \\ & \text{and} \;\; \text{cap}_f - \bar{s}_{ft} - \text{CAP}_f \;\geq\; 0, \forall \, (f,t) \in \mathcal{F} \times \mathcal{T} \end{array} \right\}$$

being an unbounded polyhedron in the variables $\bar{\mathbf{s}} \equiv (\bar{s}_{ft})_{(f,t) \in \mathcal{F} \times \mathcal{T}}$ and $\mathbf{cap} \equiv$

$\left(\text{cap}_f\right)_{f \in \mathcal{F}}$, and

$$\Phi^{\rm I}(\bar{\mathbf{s}}, \mathbf{cap}, p_e) \equiv \left(\begin{array}{c} \left(H_t \left[-\pi_t \left(\sum_{g \in \mathcal{F}} (\, \bar{s}_{gt} + \text{CAP}_g \,) \right) + \text{MC}_f + p_e \, E_f \right] \right)_{(f,t) \in \mathcal{F} \times \mathcal{T}} \\ \left(F_f - \dfrac{(\, \overline{E} - E_{\rm GF} \,) \, p_e}{\displaystyle\sum_{g \in \mathcal{F}} R_g \, \text{cap}_g} \, R_f \right)_{f \in \mathcal{F}} \\ \overline{E} - \displaystyle\sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t \, E_g \, (\, \bar{s}_{gt} + \text{CAP}_g \,) \, - e_{\rm NP}(p_e) \end{array} \right)$$

is a non-monotone map. Indeed, note that $\Phi^{\rm I}$ is well defined on the set $K^{\rm I}$ because for every element $(\bar{\mathbf{s}}, \mathbf{cap}) \in K$, we must have $\text{cap}_f > 0$ for some $f \in \mathcal{F}$. Letting p_c and μ_{ft} be the multipliers of the functional constraints in K, we can readily write down the KKT conditions of the VI $(K^{\rm I}, \Phi^{\rm I})$ and conclude that they are equivalent to the NCP (3.11) under the identification (3.14) for σ. When $\text{CAP} = 0 < \sum_{f \in \mathcal{F}} \text{CAP}_f$, the set

$$K = \prod_{f \in \mathcal{F}} \left\{ \left((\, s_{ft} \,)_{t \in \mathcal{T}}, \text{cap}_f \right) \geq 0 \, : \, \text{cap}_f - \bar{s}_{ft} - \text{CAP}_f \geq 0, \, \forall \, t \in \mathcal{T} \right\}$$

is the Cartesian product of separable sets. While the VI formulation is quite compact, one obvious advantage of the NCP (3.11) is that it applies to the case where $\underline{\text{CAP}} = \text{CAP}_f = 0$ for all $f \in \mathcal{F}$; in the latter case, the set $K^{\rm I}$ contains the origin where the function $\Phi^{\rm I}$ fails to be well defined.

3.2.2 The actual emission rule

Because of the specific expression of (3.9) under actual emission rule, the NCP reformulation for (3.10) is handled differently. In particular, consider the NCP:

$$
\begin{array}{lll}
0 \le \bar{s}_{ft} & \perp & H_t \left[-\pi_t \left(\sum_{g \in \mathcal{F}} (\bar{s}_{gt} + \mathrm{CAP}_g) \right) + \mathrm{MC}_f + p_e E_f \right] + \mu_{ft} \ge 0 \\
 & & \forall\, (f,t) \in \mathcal{F} \times \mathcal{T} \\
0 \le \mu_{ft} & \perp & \mathrm{cap}_f - \bar{s}_{ft} - \mathrm{CAP}_f \ \ge\ 0, \quad \forall\, (f,t) \in \mathcal{F} \times \mathcal{T} \\
0 \le \mathrm{cap}_f & \perp & -p_c - \alpha_f\, p_e + F_f - \sum_{t \in \mathcal{T}} \mu_{ft} \ \ge\ 0, \quad \forall\, f \in \mathcal{F} \\
0 \le \alpha_f & \perp & \alpha_f\, \mathrm{cap}_f - \widehat{\alpha}\, \widehat{R}_f \sum_{t \in \mathcal{T}} H_t\, (\, \bar{s}_{ft} + \mathrm{CAP}_f\,) \ \ge\ 0, \quad \forall\, f \in \mathcal{F} \\
0 \le p_e & \perp & \overline{E} - e_{\mathrm{NP}}(p_e) - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t\, E_g\, (\, \bar{s}_{gt} + \mathrm{CAP}_g\,) \ \ge\ 0 \\
0 \le p_c & \perp & \sum_{g \in \mathcal{F}} \mathrm{cap}_g - \underline{\mathrm{CAP}} \ \ge\ 0 \\
0 \le \widehat{\alpha} & \perp & \sum_{g \in \mathcal{F}} \alpha_g\, \mathrm{cap}_g - (\, \overline{E} - E_{\mathrm{GF}}\,) \ \ge\ 0
\end{array}
\tag{3.16}
$$

in the variable $(\mathbf{x}, \widehat{\alpha})$. The following result verifies that the above NCP is equivalent to (3.10) under the actual emission rule.

Proposition 17. A pair $(\mathbf{x}, \widehat{\alpha})$ is a solution of (3.10) under the actual emission rule if and only if it is a solution of (3.16).

Proof. It is easy to prove the "only if" statement. Conversely, let $(\mathbf{x}, \alpha)$ be a solution of (3.16). It suffices to show that (3.9) under the actual emission rule and the emission allowances balance equation hold, i.e.

$$
\sum_{g \in \mathcal{F}} \alpha_g\, \mathrm{cap}_g - (\, \overline{E} - E_{\mathrm{GF}}\,) \ =\ 0
$$

and

$$
\alpha_f\, \mathrm{cap}_f - \widehat{\alpha}\, \widehat{R}_f \sum_{t \in \mathcal{T}} H_t\, E_f\, (\, \bar{s}_{ft} + \mathrm{CAP}_f\,) \ =\ 0,\ \forall\, f\ \in\ \mathcal{F}.
$$

We prove these two equalities by contradiction. Assume that the first equality does not hold, then $\widehat{\alpha} = 0$ by complementarity, which follows $0 \leq \alpha_f \perp \alpha_f \text{cap}_f \geq 0$. In turn, this implies $\alpha_f \text{cap}_f = 0$ for all $f \in \mathcal{F}$; thus $\overline{E} - E_{\text{GF}} = 0$, which contradicts the assumption that $\overline{E} > E_{\text{GF}}$. Similarly, if $\alpha_f \text{cap}_f - \widehat{\alpha}\widehat{R}_f \sum_{t \in \mathcal{T}} H_t E_f(\bar{s}_{ft} + \text{CAP}_f) > 0$ for some $f \in \mathcal{F}$, then $\alpha_f = 0$ by complementarity, which contradicts the inequality itself. □

Similar to the VI $(K^{\text{I}}, \Phi^{\text{I}})$, if $\text{CAP}_f > 0$ for all $f \in \mathcal{F}$, the NCP (3.16) is equivalent to the KKT conditions of the VI $(K^{\text{II}}, \Phi^{\text{II}})$, where $K^{\text{II}} \equiv K^{\text{I}} = K \times \Re_+$ and

$$\Phi^{\text{II}}(\bar{\mathbf{s}}, \mathbf{cap}, p_e) \equiv \begin{pmatrix} \left(H_t \left[-\pi_t \left(\sum_{g \in \mathcal{F}} (\bar{s}_{gt} + \text{CAP}_g) \right) + \text{MC}_f + p_e\, E_f \right] \right)_{(f,t) \in \mathcal{F} \times \mathcal{T}} \\ \left(F_f - \dfrac{(\overline{E} - E_{\text{GF}})\, p_e\, \widehat{R}_f \sum_{t \in \mathcal{T}} H_t\, (\bar{s}_{ft} + \text{CAP}_f)}{\text{cap}_f \sum_{g \in \mathcal{F}} \widehat{R}_g \sum_{t \in \mathcal{T}} H_t\, (\bar{s}_{gt} + \text{CAP}_g)} \right)_{f \in \mathcal{F}} \\ \overline{E} - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t\, E_g\, (\bar{s}_{gt} + \text{CAP}_g) - e_{\text{NP}}(p_e) \end{pmatrix}.$$

3.3 Existence of Solutions

For the analysis in this section, we impose condition (3.12) and the following condition that is a slight strengthening of (3.13):

$$\max_{f \in \mathcal{F}} \left\{ \sum_{t \in \mathcal{T}} H_t \left[\pi_t \left(\sum_{g \in \mathcal{F}} \text{CAP}_g \right) - \text{MC}_f \right] - F_f \right\} > 0. \tag{3.17}$$

This condition states that at least one firm would find investment profitable (fixed and variable cost less than revenue) if every generator is producing just its individual lower bound. This is a mild restriction, as the power price would likely be very high when everyone is producing at their lowest possible level.

The following proposition shows that under these two conditions, any solution to the model (3.10) is nontrivial.

Proposition 18. Under (3.12) and (3.17), any solution of the NCP (3.11) and (3.16) must have $\bar{s}_{ft} > 0$ for some $(f,t) \in \mathcal{F} \times \mathcal{T}$.

Proof. We prove the proposition only for (3.16). Assume for the sake of contradiction that some solution of this NCP has $\bar{s}_{ft} = 0$ for all $(f,t) \in \mathcal{F} \times \mathcal{T}$. We claim that $p_e = 0$. Indeed, if $p_e > 0$, then

$$0 \;=\; \overline{E} - e_{\mathrm{NP}}(p_e) - \sum_{(g,t)\in\mathcal{F}\times\mathcal{T}} H_t\, E_g\, \mathrm{CAP}_g \;\geq\; \overline{E} - e_{\mathrm{NP}}(0) - \sum_{(g,t)\in\mathcal{F}\times\mathcal{T}} H_t\, E_g\, \mathrm{CAP}_g \;>\; 0,$$

which is a contradiction. Hence, we have, for each $f \in \mathcal{T}$,

$$\begin{aligned} 0 \;&\leq\; \sum_{t\in\mathcal{T}} \left\{ H_t \left[-\pi_t \left(\sum_{g\in\mathcal{F}} \mathrm{CAP}_g \right) + \mathrm{MC}_f \right] + \mu_{ft} \right\} \\ &\leq\; \sum_{t\in\mathcal{T}} H_t \left[-\pi_t \left(\sum_{g\in\mathcal{F}} \mathrm{CAP}_g \right) + \mathrm{MC}_f \right] + F_f, \end{aligned}$$

which contradicts (3.17). □

We recall that the temporal price functions $\pi_t(\bullet)$ are strictly decreasing and the nonpower emission function $e_{NP}(\bullet)$ is nonincreasing. The following is the main existence theorem for the model (3.10) with the emission rules (I) and (II).

Theorem 19. Under conditions (3.12) and (3.17), the model (3.10) has a solution.

We prove the above theorem via the two equivalent NCPs: (3.11) for emission rule (I) and (3.16) for emission rule (II). In turn, the proofs for these two NCPs are quite similar. Both are based on the application of a fundamental existence result for a general NCP, which is summarized in Lemma 12 in the previous chapter.

To avoid repetition, we present the proof for the NCP (3.16) only; see Subsection 3.3.1. We choose this NCP because there is an extra perturbation step that is needed in applying the lemma, whereas one can follow the same argument and directly apply the lemma to the NCP (3.11).

3.3.1 Proof for the NCP under the actual emission rule

Toward the proof of solution existence to the NCP (3.16), we consider a perturbation of the function Φ^{II} in order to deal with the general case where some $\text{CAP}_f = 0$. Specifically, for each $\varepsilon > 0$, let

$$\Phi^{\text{II}}_{\varepsilon}(\bar{\mathbf{s}}, \mathbf{cap}, p_e) \equiv \left(\begin{array}{c} \left(H_t \left[-\pi_t \left(\sum_{g \in \mathcal{F}} (\bar{s}_{gt} + \text{CAP}_g) \right) + \text{MC}_f + p_e \, E_f \right] \right)_{(f,t) \in \mathcal{F} \times \mathcal{T}} \\ \left(F_f - \dfrac{(\overline{E} - E_{\text{GF}}) \, p_e \, \widehat{R}_f \sum_{t \in \mathcal{T}} H_t (\bar{s}_{ft} + \text{CAP}_f)}{(\text{cap}_f + \varepsilon) \left(\varepsilon + \sum_{g \in \mathcal{F}} \widehat{R}_g \sum_{t \in \mathcal{T}} H_t (\bar{s}_{gt} + \text{CAP}_g) \right)} \right)_{f \in \mathcal{F}} \\ \overline{E} - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t \, E_g (\bar{s}_{gt} + \text{CAP}_g) - e_{\text{NP}}(p_e) \end{array} \right),$$

which is well defined on the set K^{II}. We first show that the VI $(K^{\text{II}}, \Phi^{\text{II}}_{\varepsilon})$ has a solution for each fixed but arbitrary $\varepsilon > 0$. For this purpose, we take an arbitrary sequence of positive scalars $\{\tau_k\}$; for each k, let $(\bar{\mathbf{s}}^{\varepsilon,k}, \mathbf{cap}^{\varepsilon,k}, \boldsymbol{\mu}^{\varepsilon,k}, p_e^{\varepsilon,k}, p_c^{\varepsilon,k})$ be a tuple satisfying

$$0 \leq \bar{s}_{ft}^{\varepsilon,k} \;\perp\; H_t \left[-\pi_t \left(\sum_{g \in \mathcal{F}} (\bar{s}_{gt}^{\varepsilon,k} + \mathrm{CAP}_g) \right) + \mathrm{MC}_f + p_e^{\varepsilon,k} E_f \right] + \mu_{ft}^{\varepsilon,k} + \tau_k \bar{s}_{ft}^{\varepsilon,k} \geq 0,$$

$$\forall\, (f,t) \in \mathcal{F} \times \mathcal{T} \tag{3.18}$$

$$0 \leq \mu_{ft}^{\varepsilon,k} \;\perp\; \mathrm{cap}_f^{\varepsilon,k} - \bar{s}_{ft}^{\varepsilon,k} - \mathrm{CAP}_f + \tau_k\, \mu_{ft}^{\varepsilon,k} \;\geq\; 0, \quad \forall\, (f,t) \in \mathcal{F} \times \mathcal{T} \tag{3.19}$$

$$0 \leq \mathrm{cap}_f^{\varepsilon,k} \;\perp\; -p_c^{\varepsilon,k} + F_f - \frac{(\overline{E} - E_{\mathrm{GF}}) p_e^{\varepsilon,k} \widehat{R}_f \sum_{t \in \mathcal{T}} H_t (\bar{s}_{ft}^{\varepsilon,k} + \mathrm{CAP}_f)}{(\mathrm{cap}_f^{\varepsilon,k} + \varepsilon) \left(\varepsilon + \sum_{g \in \mathcal{F}} \widehat{R}_g \sum_{t \in \mathcal{T}} H_t (\bar{s}_{gt}^{\varepsilon,k} + \mathrm{CAP}_g) \right)}$$

$$- \sum_{t \in \mathcal{T}} \mu_{ft}^{\varepsilon,k} + \tau_k \mathrm{cap}_f^{\varepsilon,k} \geq 0, \quad \forall\, f \in \mathcal{F} \tag{3.20}$$

$$0 \leq p_e^{\varepsilon,k} \;\perp\; \overline{E} - e_{\mathrm{NP}}(p_e^{\varepsilon,k}) - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t\, E_g \left(\bar{s}_{gt}^{\varepsilon,k} + \mathrm{CAP}_g \right) + \tau_k\, p_e^{\varepsilon,k} \;\geq\; 0 \tag{3.21}$$

$$0 \leq p_c^{\varepsilon,k} \;\perp\; \sum_{g \in \mathcal{F}} \mathrm{cap}_g^{\varepsilon,k} - \underline{\mathrm{CAP}} + \tau_k\, p_c^{\varepsilon,k} \;\geq\; 0. \tag{3.22}$$

We claim that under condition (3.12), the sequence $\{(\bar{\mathbf{s}}^{\varepsilon,k}, \mathbf{cap}^{\varepsilon,k}, \boldsymbol{\mu}^{\varepsilon,k}, p_e^{\varepsilon,k}, p_c^{\varepsilon,k})\}$ is bounded. We show this in several steps: first the sequence $\{p_e^{\varepsilon,k}\}$; next the sequence $\{\bar{s}_{ft}^{\varepsilon,k}\}$ for all $(f,t) \in \mathcal{F} \times \mathcal{T}$; then the sequence $\{\mathrm{cap}_f^{\varepsilon,k}\}$ for all $f \in \mathcal{F}$.

Boundedness of $\{p_e^{\varepsilon,k}\}$. Assume for the sake of contradiction that $\{p_e^{\varepsilon,k}\}$ is unbounded. Then for an infinite index set $\kappa \subset \{1, 2, \cdots, \infty\}$, we have

$$\lim_{k(\in\kappa)\to\infty} p_e^{\varepsilon,k} \;=\; \infty. \tag{3.23}$$

Without loss of generality, we may assume that $p_e^{\varepsilon,k} > 0$ for all $k \in \kappa$. It follows by complementarity condition (3.21) that

$$\overline{E} - e_{\rm NP}(p_e^{\varepsilon,k}) - \sum_{(g,t)\in\mathcal{F}\times\mathcal{T}} H_t\, E_g \left(\bar{s}_{gt}^{\varepsilon,k} + \mathrm{CAP}_g \right) + \tau_k\, p_e^{\varepsilon,k} \;=\; 0, \quad \forall\, k \;\in\; \kappa.$$

For any $k \in \kappa$ such that $\bar{s}_{f_0t_0}^{\varepsilon,k} > 0$ for some pair (f_0, t_0), we have by complementarity condition (3.18)

$$H_{t_0} \left[-\pi_{t_0} \left(\sum_{g\in\mathcal{F}} (\bar{s}_{gt_0}^{\varepsilon,k} + \mathrm{CAP}_g) \right) + \mathrm{MC}_{f_0} + p_e^{\varepsilon,k}\, E_{f_0} \right] + \mu_{f_0t_0}^{\varepsilon,k} + \tau_k \bar{s}_{f_0t_0}^{\varepsilon,k} \;=\; 0,$$

which yields

$$p_e^{\varepsilon,k} \;\leq\; E_{f_0}^{-1} \left[\pi_{t_0} \left(\sum_{g\in\mathcal{F}} \mathrm{CAP}_g \right) - \mathrm{MC}_{f_0} \right]. \tag{3.24}$$

On the other hand, if $k \in \kappa$ is such that $\bar{s}_{ft}^{\varepsilon,k} = 0$ for all (f, t), then we have

$$\begin{aligned} 0 \;&=\; \overline{E} - e_{\rm NP}(p_e^{\varepsilon,k}) - \sum_{(g,t)\in\mathcal{F}\times\mathcal{T}} H_t\, E_g\, \mathrm{CAP}_g + \tau_k\, p_e^{\varepsilon,k} \\ &\geq\; \overline{E} - e_{\rm NP}(0) - \sum_{(g,t)\in\mathcal{F}\times\mathcal{T}} H_t\, E_g\, \mathrm{CAP}_g \;>\; 0, \end{aligned}$$

which contradicts (3.12). Consequently, the bound (3.24) holds for all $k \in \kappa$; contradicting the limit (3.23).

Boundedness of $\{\bar{s}_{ft}^{\varepsilon,k}\}$ for every $(f, t) \in \mathcal{F} \times \mathcal{T}$. Assume for the sake of contradiction that for some pair (f_0, t_0) and an infinite set $\kappa \subset \{1, 2, \cdots, \}$,

$$\lim_{k(\in\kappa)\to\infty} \bar{s}_{f_0t_0}^{\varepsilon,k} \;=\; \infty. \tag{3.25}$$

Without loss of generality, we may assume that $\bar{s}_{f_0t_0}^{\varepsilon,k} > 0$ for all $k \in \kappa$. It follows by complementarity condition (3.18) that

$$\begin{aligned} 0 &= H_{t_0} \left[-\pi_{t_0} \left(\sum_{g \in \mathcal{F}} (\bar{s}^{\varepsilon,k}_{gt_0} + \text{CAP}_g) \right) + \text{MC}_{f_0} + p^{\varepsilon,k}_e \, E_{f_0} \right] + \mu^{\varepsilon,k}_{f_0 t_0} + \tau_k \, \bar{s}^{\varepsilon,k}_{f_0 t_0} \\ &\geq H_{t_0} \left[-\pi_{t_0} \left(\sum_{g \in \mathcal{F}} \text{CAP}_g \right) + \text{MC}_f + p^{\varepsilon,k}_e \, E_{f_0} \right] + \max(\mu^{\varepsilon,k}_{f_0 t_0}, \, \tau_k \, \bar{s}^{\varepsilon,k}_{f_0 t_0}) \end{aligned}$$

which implies

$$\max(\mu^{\varepsilon,k}_{f_0 t_0}, \, \tau_k \, \bar{s}^{\varepsilon,k}_{f_0 t_0}) \, \leq \, H_{t_0} \left[\pi_{t_0} \left(\sum_{g \in \mathcal{F}} \text{CAP}_g \right) - \text{MC}_f - p^{\varepsilon,k}_e \, E_{f_0} \right].$$

Since the right-hand side is bounded, by (3.25), it follows that

$$\lim_{k (\in \kappa) \to \infty} \tau_k \, = \, 0. \tag{3.26}$$

But this contradicts the inequality in (3.21):

$$\overline{E} - e_{\text{NP}}(p^{\varepsilon,k}_e) - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t E_g \left(\bar{s}^{\varepsilon,k}_{gt} + \text{CAP}_g \right) + \tau_k p^{\varepsilon,k}_e \geq 0.$$

Therefore, $\{\bar{s}^{\varepsilon,k}_{ft}\}$ is bounded for all $(f,t) \in \mathcal{F} \times \mathcal{T}$.

Boundedness of $\{\text{cap}^{\varepsilon,k}_f\}$ for every $f \in \mathcal{F}$. Assume for the sake of contradiction that for some f_0 and an infinite set $\kappa \subset \{1, 2, \cdots, \}$,

$$\lim_{k (\in \kappa) \to \infty} \text{cap}^{\varepsilon,k}_{f_0} \, = \, \infty. \tag{3.27}$$

Thus, by complementarity, we must have $p^{\varepsilon,k}_c = 0$ for all $k \in \kappa$ sufficiently large. Since $\{\bar{s}^{\varepsilon,k}_{f_0 t}\}$ is bounded, we deduce $\mu^{\varepsilon,k}_{f_0 t} = 0$ for all $t \in \mathcal{T}$ and all $k \in \kappa$ sufficiently large. Without loss of generality, we may assume that $\text{cap}^{\varepsilon,k}_{f_0} > 0$ for all $k \in \kappa$. It follows by complementarity condition (3.20) that for all $k \in \kappa$ sufficiently large,

$$F_{f_0} - \frac{(\overline{E} - E_{\text{GF}})\, p_e^{\varepsilon,k}\, \widehat{R}_{f_0} \sum_{t \in \mathcal{T}} H_t\,(\,\bar{s}_{f_0 t}^{\varepsilon,k} + \text{CAP}_{f_0}\,)}{(\,\text{cap}_{f_0}^{\varepsilon,k} + \varepsilon\,)\left(\varepsilon + \sum_{g \in \mathcal{F}} \widehat{R}_g \sum_{t \in \mathcal{T}} H_t(\,\bar{s}_{gt}^{\varepsilon,k} + \text{CAP}_g)\right)} + \tau_k\, \text{cap}_{f_0}^{\varepsilon,k} \;=\; 0,$$

which implies that

$$F_{f_0} \;\le\; \frac{(\overline{E} - E_{\text{GF}})\, p_e^{\varepsilon,k}\, \widehat{R}_{f_0} \sum_{t \in \mathcal{T}} H_t\,(\,\bar{s}_{f_0 t}^{\varepsilon,k} + \text{CAP}_{f_0}\,)}{(\,\text{cap}_{f_0}^{\varepsilon,k} + \varepsilon\,)\left(\varepsilon + \sum_{g \in \mathcal{F}} \widehat{R}_g \sum_{t \in \mathcal{T}} H_t(\,\bar{s}_{gt}^{\varepsilon,k} + \text{CAP}_g)\right)}.$$

The limit (3.27) implies that the right-hand side tends to zero as $k(\in \kappa) \to \infty$, which is a contradiction. Hence $\{\text{cap}_f^{\varepsilon,k}\}$ is bounded for all $f \in \mathcal{F}$.

Boundedness of $\{\mu_{ft}^{\varepsilon,k}\}$ for all $(f,t) \in \mathcal{F} \times \mathcal{T}$. Assume for the sake of contradiction that for some $(f_0, t_0) \in \mathcal{F} \times \mathcal{T}$ and an infinite set $\kappa \subset \{1, 2, \cdots, \}$,

$$\lim_{k(\in\kappa)\to\infty} \mu_{f_0 t_0}^{\varepsilon,k} \;=\; \infty. \tag{3.28}$$

Without loss of generality we may assume that $\mu_{f_0 t_0}^{\varepsilon,k} > 0$ for all $k \in \kappa$. By complementarity condition (3.19), we deduce $\text{cap}_{f_0}^{\varepsilon,k} - \bar{s}_{f_0 t_0}^{\varepsilon,k} + \text{CAP}_{f_0} + \tau_k \mu_{f_0 t_0}^{\varepsilon,k} = 0$. Since $\{(\text{cap}_{f_0}^{\varepsilon,k}, \bar{s}_{f_0 t_0}^{\varepsilon,k})\}$ is bounded, (3.28) implies that

$$\lim_{k(\in\kappa)\to\infty} \tau_k \;=\; 0.$$

Since

$$\mu^{\varepsilon,k}_{f_0 t_0} \leq \sum_{t \in \mathcal{T}} \mu^{\varepsilon,k}_{f_0 t} \leq -p^{\varepsilon,k}_c + F_{f_0} - \frac{(\overline{E} - E_{\mathrm{GF}}) p^{\varepsilon,k}_e \widehat{R}_{f_0} \sum_{t \in \mathcal{T}} H_t(\bar{s}^{\varepsilon,k}_{f_0 t} + \mathrm{CAP}_{f_0})}{(\mathrm{cap}^{\varepsilon,k}_{f_0} + \varepsilon)\left(\varepsilon + \sum_{g \in \mathcal{F}} \widehat{R}_g \sum_{t \in \mathcal{T}} H_t(\bar{s}^{\varepsilon,k}_{gt} + \mathrm{CAP}_g)\right)} + \tau_k \mathrm{cap}^{\varepsilon,k}_{f_0}$$

$$\leq F_{f_0} + \tau_k \, \mathrm{cap}^{\varepsilon,k}_{f_0}$$

and $\tau_k \mathrm{cap}^{\varepsilon,k}_{f_0} \to 0$ as $k(\in \kappa) \to \infty$, we obtain a contradiction to (3.28).

Boundedness of $\{p^{\varepsilon,k}_c\}$. This is similar to the above proof of the μ-sequence.

We have therefore completed the proof of the boundedness of the sequence $\{(\bar{\mathbf{s}}^{\varepsilon,k}, \mathbf{cap}^{\varepsilon,k}, \boldsymbol{\mu}^{\varepsilon,k}, p^{\varepsilon,k}_e, p^{\varepsilon,k}_c)\}$ under the condition (3.12). This is enough to apply Lemma 15 to deduce the existence of a solution to the VI $(K^{\mathrm{II}}, \Phi^{\mathrm{II}}_\varepsilon)$ for all $\varepsilon > 0$ via its KKT formulation. Let $(\bar{\mathbf{s}}^\varepsilon, \mathbf{cap}^\varepsilon, p^\varepsilon_e)$ be one such solution. For each $\varepsilon > 0$, there exists $(\mu^\varepsilon_{ft}, p^\varepsilon_c)$ such that

$$0 \leq \bar{s}^\varepsilon_{ft} \quad \perp \quad H_t \left[-\pi_t \left(\sum_{g \in \mathcal{F}} (\bar{s}^\varepsilon_{gt} + \mathrm{CAP}_g) \right) + \mathrm{MC}_f + p^\varepsilon_e E_f \right] + \mu^\varepsilon_{ft} \geq 0, \quad \forall (f,t) \in \mathcal{F} \times \mathcal{T}$$

$$0 \leq \mu^\varepsilon_{ft} \quad \perp \quad \mathrm{cap}^\varepsilon_f - \bar{s}^\varepsilon_{ft} - \mathrm{CAP}_f \geq 0, \quad \forall (f,t) \in \mathcal{F} \times \mathcal{T}$$

$$0 \leq \mathrm{cap}^\varepsilon_f \quad \perp \quad -p^\varepsilon_c + F_f - \frac{(\overline{E} - E_{\mathrm{GF}})\, p^\varepsilon_e \widehat{R}_f \sum_{t \in \mathcal{T}} H_t\,(\bar{s}^\varepsilon_{ft} + \mathrm{CAP}_f)}{(\mathrm{cap}^\varepsilon_f + \varepsilon)\left(\varepsilon + \sum_{g \in \mathcal{F}} \widehat{R}_g \sum_{t \in \mathcal{T}} H_t\,(\bar{s}^\varepsilon_{gt} + \mathrm{CAP}_g)\right)} - \sum_{t \in \mathcal{T}} \mu^\varepsilon_{ft} \geq 0, \quad \forall f \in \mathcal{F}$$

$$0 \leq p^\varepsilon_e \quad \perp \quad \overline{E} - e_{NP}(p^\varepsilon_e) - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t\, E_g \left(\bar{s}^\varepsilon_{gt} + \mathrm{CAP}_g\right) \geq 0$$

$$0 \leq p^\varepsilon_c \quad \perp \quad \sum_{g \in \mathcal{F}} \mathrm{cap}^\varepsilon_g - \underline{\mathrm{CAP}} \geq 0.$$

By the same proof sequence as before, we can show that

$$\limsup_{\varepsilon \downarrow 0} \| \, (\, \bar{\mathbf{s}}^{\varepsilon}, \mathbf{cap}^{\varepsilon}, \boldsymbol{\mu}^{\varepsilon}, p_e^{\varepsilon}, p_c^{\varepsilon} \,) \, \| \, < \, \infty.$$

Let $(\widehat{\mathbf{s}}, \widehat{\mathbf{cap}}, \widehat{\boldsymbol{\mu}}, \widehat{p}_e, \widehat{p}_c)$ be the limit of a convergence sequence $\{(\bar{\mathbf{s}}^{\varepsilon_k}, \mathbf{cap}^{\varepsilon_k}, \boldsymbol{\mu}^{\varepsilon_k}, p_e^{\varepsilon_k}, p_c^{\varepsilon_k})\}$ corresponding to a sequence of positive scalars $\{\varepsilon_k\} \downarrow 0$. It follows readily that $(\widehat{\mathbf{s}}, \widehat{\mathbf{cap}}, \widehat{\boldsymbol{\mu}}, \widehat{p}_e, \widehat{p}_c)$ satisfies

$$\begin{aligned}
0 \, \leq \, \widehat{s}_{ft} \quad & \perp \quad H_t \left[-\pi_t \left(\sum_{g \in \mathcal{F}} (\, \widehat{s}_{gt} + \text{CAP}_g \,) \right) + \text{MC}_f + \widehat{p}_e \, E_f \right] + \widehat{\mu}_{ft} \, \geq \, 0, \\
& \quad \forall \, (f, t) \in \mathcal{F} \times \mathcal{T} \\
0 \, \leq \, \widehat{\mu}_{ft} \quad & \perp \quad \widehat{\text{cap}}_f - \widehat{s}_{ft} - \text{CAP}_f \, \geq \, 0, \quad \forall \, (f, t) \in \mathcal{F} \times \mathcal{T} \\
0 \, \leq \, \widehat{p}_e \quad & \perp \quad \overline{E} - e_{\text{NP}}(\widehat{p}_e) - \sum_{(g,t) \in \mathcal{F} \times \mathcal{T}} H_t \, E_g \, \left(\widehat{s}_{gt} + \text{CAP}_g \right) \, \geq \, 0 \\
0 \, \leq \, \widehat{p}_c \quad & \perp \quad \sum_{g \in \mathcal{F}} \widehat{\text{cap}}_g - \underline{\text{CAP}} \, \geq \, 0.
\end{aligned}$$

By the same proof as that of Proposition 18, we can show that $\widehat{\mathbf{s}} \neq 0$. Since

$$\frac{\displaystyle\sum_{t \in \mathcal{T}} H_t \, (\, \bar{s}_{ft}^{\varepsilon_k} + \text{CAP}_f \,)}{(\, \text{cap}_f^{\varepsilon_k} + \varepsilon_k \,)} \, \leq \, \frac{\displaystyle\sum_{t \in \mathcal{T}} H_t \, \text{cap}_f^{\varepsilon_k}}{(\, \text{cap}_f^{\varepsilon_k} + \varepsilon_k \,)} \, < \, \sum_{t \in \mathcal{T}} H_t$$

it follows that the sequence $\left\{ \frac{\displaystyle\sum_{t \in \mathcal{T}} H_t \, (\, \bar{s}_{ft}^{\varepsilon_k} + \text{CAP}_f \,)}{(\, \text{cap}_f^{\varepsilon_k} + \varepsilon_k \,)} \right\}$ must have at least one accumulation point. Without loss of generality, we may assume that this sequence converges to a limit, say $\gamma_f \geq 0$. Note that

$$\gamma_f \, = \, \frac{\displaystyle\sum_{t \in \mathcal{T}} H_t \, (\, \widehat{s}_{ft} + \text{CAP}_f \,)}{\widehat{\text{cap}}_f}, \quad \text{if the denominator is positive.}$$

Generator type	coal steam	natural gas-fired combined cycle	natural gas-fired combustion turbine
$\text{CAP}_f(MW)$	0	0	0
MC_f(\$/MWh)	20	40	80
F_f(\$/MWh)	120,000	75,000	50,000
E_f(ton/MWh)	1	0.35	0.6
$\underline{\text{CAP}}$ (MW)	11,000 if CAP > 0		
Demands	$d_t(p_t) = a_t - b_t p_t$, with $a_t = 500t$ and $b_t = t/2$		
Nonpower emission	$e_{\text{NP}}(p_e) = 0$		
Time	$T = 20$ periods per year, $H_t = 438$ hr		

Table 3.1: Parameter setting for the simple system

Define, for each $f \in \mathcal{F}$,

$$\alpha_f \;\equiv\; \frac{(\overline{E} - E_{\text{GF}})\,\widehat{R}_f\,\gamma_f}{\displaystyle\sum_{g \in \mathcal{F}} \widehat{R}_g \sum_{t \in \mathcal{T}} H_t\,(\,\widehat{s}_{gt} + \text{CAP}_g\,)}\,\widehat{p}_e.$$

It is easy to show that the following complementarity holds:

$$0 \;\leq\; \widehat{\text{cap}}_f \;\perp\; -\widehat{p}_c + F_f - \alpha_f\,\widehat{p}_e + \sum_{t \in \mathcal{T}} \widehat{\mu}_{ft} \;\geq\; 0, \quad \forall\, f \;\in\; \mathcal{F}.$$

With

$$\widehat{\alpha} \;\equiv\; \frac{(\overline{E} - E_{\text{GF}})\,\widehat{p}_e}{\displaystyle\sum_{g \in \mathcal{F}} \widehat{R}_g \sum_{t \in \mathcal{T}} H_t\,(\,\widehat{s}_{gt} + \text{CAP}_g\,)},$$

it is easy to see that all conditions in (3.16) are satisfied.

3.4 Application

In this section, a simple system is considered to illustrate the efficiency of the potential emission rule and the actual emission rule under various market settings. All the complementarity problems are solved by the PATH solver. The parameters that characterized the system are summarized in Table 3.1.

The demand function $d_t(p_t)$ satisfies the following properties: given the same energy price at each time period, the demand grows up as t increases; in turn, the peak demand occurs at the last period; at each period, price can be reached up to \$1000/MWh, so the demand is relatively inelastic at the equilibrium price, which is usually much lower than \$1000/MWh. Moreover, if the capacity market is taken into account, it is assume to be 10% more than the peak demand of 10,000 MW.

In order to quantify the investment, efficiency and pricing outcomes of the two allocation rules, we introduce the following system performance measures:

- *Generation cost* (M\$/yr), total generation investment and fuel costs:

$$\sum_{f\in\mathcal{F}}\left(F_f \text{cap}_f + \sum_{t\in\mathcal{T}} H_t MC_f s_f\right);$$

- *Social cost* (M\$/yr), the cost of generation plus the cost of price-induced changes in energy consumption: PS + CS + GS, where PS is the equilibrium producer surplus, equal to the sum over all firms of the objectives in (3.1); CS is the equilibrium consumer surplus: $\sum_{t\in\mathcal{T}} H_t \left(\int_0^{d_t(p_t)} \pi_t(x)dx - d_t(p_t)p_t\right)$; and GS is the auction or grandfathering surplus, equal to the economic rent accruing to the original owners of grandfathered allowances (or, equivalently, the revenue received by the government if it instead auctioned those allowances). However, we exclude environmental costs from this performance measure;
- *Consumer payments* (M\$/yr): $p_c \sum_{f\in\mathcal{F}} \text{cap}_f + \sum_{t\in\mathcal{T}} d_t(p_t)p_t$; and
- *Capacity factor*, the ratio of annual generation to potential generation:

$$\frac{\sum_{t\in\mathcal{T}} H_t s_{ft}}{\sum_{t\in\mathcal{T}} H_t \text{cap}_f}.$$

3.4.1 The base run

The base run works on the scenario without a CO_2 cap i.e. $\overline{E} = \infty$; consequently, $p_e = 0$. It provides a reference to compare the two emission rules. Due to the absence of the CO_2 limit, the original model results in a the following linear complementarity problems:

$$
\begin{aligned}
0 \le \bar{s}_{ft} \quad &\perp \quad H_t\left[-\pi_t\left(\sum_{g\in\mathcal{F}}(\bar{s}_{gt}+\mathrm{CAP}_g)\right)+\mathrm{MC}_f\right]+\mu_{ft}\ge 0, && \forall\,(f,t)\in\mathcal{F}\times\mathcal{T}\\
0 \le \mu_{ft} \quad &\perp \quad \mathrm{cap}_f-\bar{s}_{ft}-\mathrm{CAP}_f\ \ge\ 0, && \forall\,(f,t)\in\mathcal{F}\times\mathcal{T}\\
0 \le \mathrm{cap}_f \quad &\perp \quad -p_c+F_f-\sum_{t\in\mathcal{T}}\mu_{ft}\ \ge\ 0, && \forall\,f\in\mathcal{F}\\
0 \le p_c \quad &\perp \quad \sum_{g\in\mathcal{F}}\mathrm{cap}_g-\underline{\mathrm{CAP}}\ \ge\ 0.
\end{aligned}
\tag{3.29}
$$

The results of the base run are shown in the table in the appendix. As an example, when the total capacity requirement $\underline{\mathrm{CAP}}$ is 11,000 MW, coal plants make up 66.6% of the energy, followed by 14.8% from combined cycle plants. The investment in combustion turbines capacities is merely enough to meet the capacity market requirement. Because of the absence of emissions limit and the zero profit free-entry assumption, the generation cost 2049 M\$/yr is equals to the consumer payment. The capacity price p_c agrees with the annual cost of combustion turbine capacity, 50,000 \$/MW/yr. The sum of consumer surplus and producer surplus is 20911 M\$/yr.

3.4.2 Comparative results for the allowance allocation rules

In this subsection, the potential emissions allocation rule and the actual emissions allocation rule are compared under different system settings. The table in the appendix lists the main system performance measures for a series of experiments representing alternative assumptions regarding: the type of the emission allocation as well as the base run where $\overline{E}=\infty$; the presence or absence of a capacity market; and the presence or absence of minimum output levels (in the form of a min-run capacity constraint) for coal plants only. For each set of assumptions, results are presented for CO_2 limits of 20 (a severe restriction) and 40 (a mild restriction) Mtons/yr, and for three cases of 0%, 50% or 100% allowances grandfathered, i.e. $E_{GF}/\overline{E}=0, 0.5, 1$, respectively. R_f in (3.6) and $\widehat{R}_f$ in (3.7) are set to be both equal to E_f/E_1.

For the purpose of the comparison, we focus on the percentage of increase in generation cost, social cost (equal to the loss of social surplus), and consumer payments relative to the base run solution obtained in Subsection 3.4.1. The three percentage increases are calculated, respectively, as:

• relative generation cost increase:
100% (generation cost - 2049 M$/yr)/2049 M$/yr;

• relative social cost increase:
100% [20911 M$/yr - (PS + CS + GS)]/2049 M$/yr;

• relative consumer payments increase:
100% (consumer payments - 2049 M$/yr)/2049 M$/yr.

With the presence of a capacity market, runs 1–6 show the results for the potential emission rule and runs 7–12 show the results for the actual emission rule. In the 100% grandfathered case, the results under the potential rule are the same as the ones under the actual rule. This is because no allowances are allocated to new investment in this case, consequently how the allowances are distributed does not matter. In the 0% case, where all allowances are instead allocated to new entries by the actual emission rule (runs 7,10), we observe that the investment is greatly distorted and costs are much higher than if allowances are completely grandfathered (or auctioned). The distortion is worse under the potential emission rule, because it is possible for new generators to be compensated by free allowances without generating any power under the potential rule. For instance, in the tight CO_2 restriction case ($\overline{E}$ = 20 Mton/yr), more than 99% of of the combustion turbine capacity is never used, and it is just built to receive free allowances. This is reflected by the combustion turbine's capacity factor in runs 1–3. Unlike the actual allocation rule, the cost distortion is much worse in the mild CO_2 limit ($\overline{E}$ = 40 Mton/yr) under the potential allocation rule. The social cost or generation cost (in run 1) is as high as the one in the 20 Mton/yr case (run 4) when $E_{GF} = 0\%\overline{E}$. This means that it should not be assumed that the risk of distortion is less if CO_2 limits are less severe.

The results listed in runs 13–24 address the effect of the absence of capacity market. When there is not separate market for capacity, the energy prices measured by customer payments are much higher than the corresponding ones with the presence of capacity market. What happens is that the energy prices grow so hight that consumer demand does not exceed available capacity. Moreover, the capacity investment decreases compared with the one in runs 1–12. The decreased investment is most dramatic for combustion turbines plant, which does not have any capacity investment under the actual emission rule. Furthermore, under the potential emission rule, the costs and generation mixes are unchanged at the presence of a capacity market under in 0% grandfathering case where all the allowances are granted to new investment. This is because the large amount of free allowances encourage the firm to overinvest in the capacity under the potential allocation rule. Consequently, the capacity constraint is not binding.

Runs 25–27 show the results obtained from the model with imposing a min-run constraint on coal plants under potential emission rule. The minimum output assumption captures the reality that the coal plants cannot be cycled on and off (shift between low and high coal capacity factors). The output of coal plants is assumed to be no less than 35% of their capacity. Imposing the min-run restriction causes higher generation cost and consumer payment in the base run case than the other base runs. As a result, the percentages of relative increase in cost are reduced by over half. This can be seen by comparing runs 25-27 with runs 1-3. The only assumption difference between them is the the min-run constraint.

3.4.3 Results for the actual emissions rule

In this subsection, we are interested in the effect of grandfathering under the actual emission allocation rule. We solve a set of NCP (3.16) with the percentage of allowances grandfathered changing from 0% to 100% by increments of 10%, with the presence of capacity market ($\underline{\text{CAP}} = 11,000$ MW). We show the results for the actual emission rule under two CO_2 limit levels: the mild CO_2 restriction ($\overline{E} = 40$ Mton/yr) and the tight CO_2 restriction ($\overline{E} = 20$ Mton/yr).

The effect of the different systems upon the costs and the allowance price p_e is plotted in Figure 3.2 and 3.3. The the distortion of investment and operation of coal and nature gas-fired combined cycle plants is illustrated in Figure 3.4 and 3.5. In the horizontal axis of each figure, the percentage of allowances grandfathered varies from 0% to 100%, which results several different systems. For example, at the left-hand extreme i.e. $E_{GF} = 0\%\overline{E}$, it represents the system where all the emissions allowances are all reserved for new investments, so new entries can receive the allowances for free. At the right-hand extreme i.e. $E_{GF} = 100\%\overline{E}$, it corresponds to the complete grandfathered system where all the allowances have already been grandfatherd or auctioned. Hence, each new participant has to purchase all the emissions rights he needs to operate. Any percentage less than 100% presents the system mixing grandfathering and partial allocation to new investment.

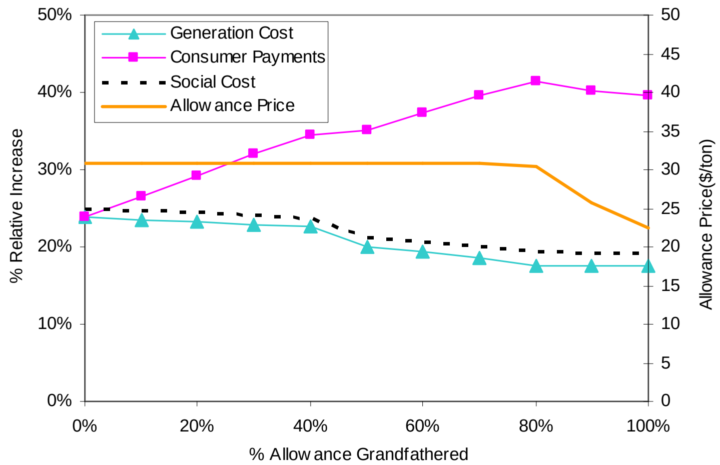

Figure 3.2: Cost and price comparison (20 Mton limit)

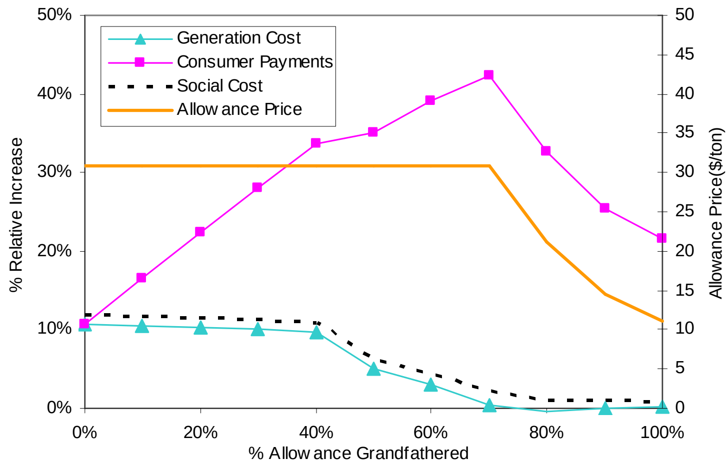

Figure 3.3: Cost and price comparison (40 Mton limit)

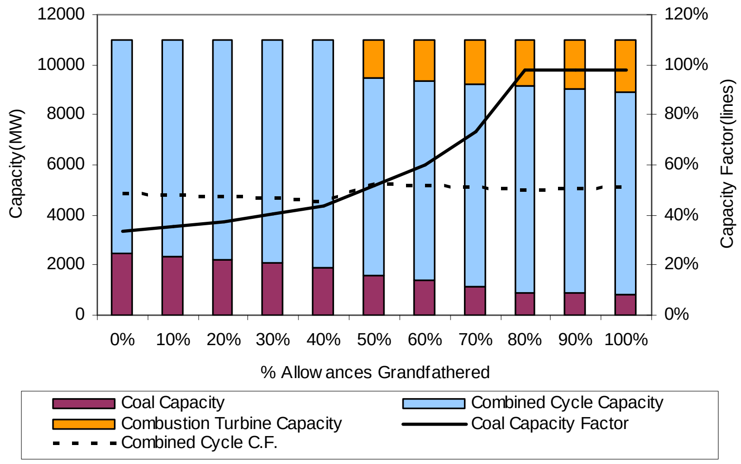

Figure 3.4: Capacity and capacity factor comparison (20 Mton limit)

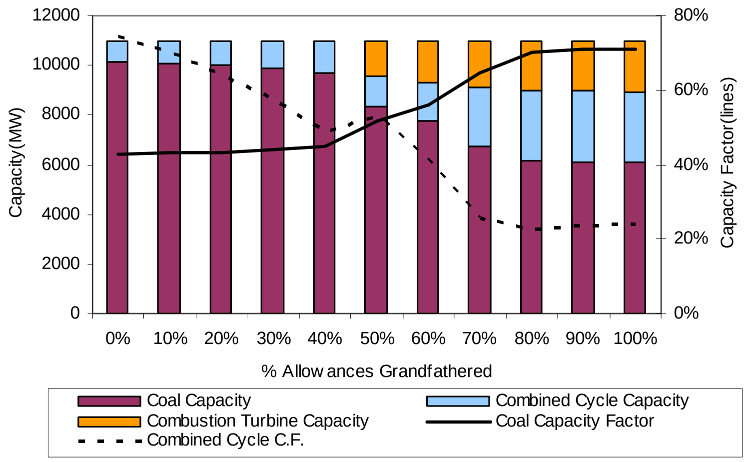

Figure 3.5: Capacity and capacity factor comparison (40 Mton limit)

In both Figure 3.2 and Figure 3.3, the social cost follows the tendency of the generation cost. In fact, from the generators' point of view, the expense associated with emission allowances is passed from the government to consumers by charging higher energy price. Therefore, in the terms of the overall social cost, it is partially consistent with the generation cost. Moreover, as more and more allowances are allocated freely to new investments, the generation cost or social cost increases in both cases. More specifically, the generation cost changes slightly when the fraction is between 100% and 80%. As the fraction falls below 80%, the distortion measured by the generation cost or social cost exaggerates a lot. Furthermore, Figure 3.3 shows that in complete grandfathering system ($E_{GF} = 100\%\overline{E}$), the 40 Mton/yr limit has a very small social cost which is only 1% of the unconstrained social cost. However, under the stricter 20 Mton/yr limit, the cost is much more expensive; it is almost 20% of the unconstrained social cost (see Figure 3.2.) In the system where $E_{\mathrm{G}F} = 0\%\overline{E}$, the social cost raises to almost 12% and 25% under mild and strict CO_2 restrictions respectively. This means that by applying the actual emission allocation rule, the social cost or generation cost under the tight CO_2 limit is almost twice as much as the one under the loose CO_2 limit. This can be explained by Figure 3.4 and 3.5. In order to meet the stricter CO_2 restriction, there are much more investments in the gas-fired capacities than the ones in the coal-fired capacity under $\overline{E} = 20$ tons/yr. It is vice versa under the mild restriction. Since the marginal cost of gas-fired capacity is higher than the one of the other two types of capacities, it results greater generation cost or social cost.

Another observation is that the investment in coal-fired facility is unexpectedly large in Figure 3.5, but its capacity factor is lower than combined cycle capacity factor. The reason is that when allowances price reaches up to \$ 31/ton when $E_{\mathrm{G}F} = 0\%\overline{E}$, natural gas-fired plants are cheaper to run, on the margin, than coal-fired plants. As the allowances price goes down, the coal-fired plants start generating more power. Two capacity factors shift at the percentage 50%, as shown in the figures. In the 100% grandfathered case, combined cycle plants is base loaded.

Last but not least, as the fraction of allowances grandfathered increases, meaning that the new entries have higher allowances expenses, consumer payments also increase. This is partially consistent with the conjecture of the Netherlands Bureau for Economic Policy Analysis [43] that much of the value of emissions allowances is passed back to consumers in the long run.

3.5 Conclusion

In this chapter, a nonlinear complementarity problem is formulated for the analysis of alternative emissions allowance allocation systems in electric power markets. Under mild conditions, the existence of solutions are established by applying a fundamental degree-theoretic result for the NCP. An example illustrates the distortions cause by different allowance allocation schemes on consumer costs, social cost, investment, etc. Moreover, the inefficiency of the allocation rules in various systems are studied as well.

Future work could address formulation of more realistic models including, for instance, transmission or carbon sequestration alternatives; parameterization based on actual markets; extension to other allowance allocation systems, such as output-based allocation [19]; and representation of interlinked markets in which different markets are subject to different rules, as is presently the case in the European Union.

CHAPTER 4
DYNAMIC SINGLE BOTTLENECK MODELS

In this chapter, a dynamic single bottleneck with heterogeneous commuters classes model is studied. The detail description of the model is given in the subsequent section. This chapter is organized as follows. First, the model is formulated as a linear complementarity problem in Section 4.1. Next, the existence of an equilibrium is established in Section 4.2. Section 4.3 studies departure patterns in the heterogeneous commuters classes case. Section 4.4 investigates the homogeneous commuters class case. Especially, the uniqueness of an equilibrium cost is confirmed and a specified algorithm is proposed. Some concluding remarks and suggestions for future work are provided in Section 4.5.

4.1 Model Definition

Consider a link with a single bottleneck (Figure 4.1). The commuters from distinct classes travel from the original to the destination along the link during a certain time period. The total time period is equally partitioned into several discrete time intervals. Each commuter needs to make decision on his or her departure time such that DUE is achieved. The notations used in the model formulation are summarized below; first the parameters, then followed by the models' variables.

Parameters: all the number parameters are positive

$\mathcal{F}$	Set of classes
$\mathcal{T}$	Set of time intervals $\equiv \{1, \cdots, T\}$
N_g	Total number of commuters of class g (unit)
s	Bottleneck capacity (unit)
α_g	Class g's unit cost of travel time (\$/time unit)
β_g	Class g's unit cost of arriving early to work (\$/time unit)
γ_g	Class g's unit cost of arriving late to work (\$/time unit)
$t_{*,g}$	Preferred arrival time of class g (time unit)

Moreover, it is reasonable assumed that $\beta_g < \alpha_g < \gamma_g$, meaning that a commuter would rather arrive at work early than waiting in the middle of a traffic congestion. Of course, it is the worst to arrive at work late. In the homogeneous commuter case, all the g subscripts are removed.

Variables:

$r_{t,g}$	Number of departures of class g at time interval t (unit)
TT_t	Travel time of individual who departs at time interval t (time unit)
$e_{t,g}$	Duration between arriving early and starting working if individuals of class g depart at time interval t (time unit)
$l_{t,g}$	Duration of lost work time if individuals of class g departing at time interval t (time unit)
$C_{t,g}$	Cost of commuters of class g departing at time interval t (\$)
$C_{*,g}$	Equilibrium cost of class g (\$).

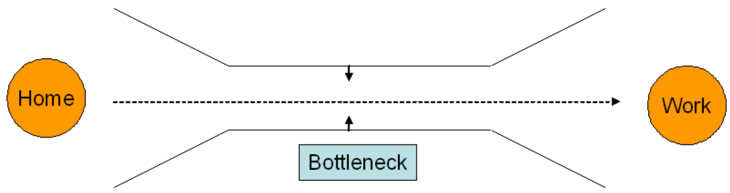

Figure 4.1: Single bottleneck

Travel time TT_t is the summation of the fixed free-flow travel time and the delay time caused by the bottleneck. Without loss of generality, the free-flow travel time can be ignored. If there is a delay occurred at the bottleneck at time period t, the delay time is approximated by $TT_{t-1} + \left(\sum_{g \in \mathcal{F}} r_{t,g} - s\right) / s$. This approximation is widely used in the literature on the single bottleneck models. Basically, it means that the current travel time TT_t is effected by the previous travel time TT_{t-1} and the congestion condition at the current time $\left(\sum_{g\in\mathcal{F}} r_{t,g} - s\right)/s$. Thus,

$$TT_t = \max\left(0, TT_{t-1} + \frac{\sum_{g\in\mathcal{F}} r_{t,g} - s}{s}\right), \tag{4.1}$$

where TT_0 is equal to 0. For the convenience, $e_{t,g}$ is called the "early arrival time", which is given by

$$e_{t,g} = \max(0, t_{*,g} - t - TT_t). \tag{4.2}$$

If individual arrives at work late, then $e_{t,g} = 0$. Similarly, $l_{t,g}$ is the "late arrival time", which is given by

$$l_{t,g} = \max(0, TT_t + t - t_{*,g}). \tag{4.3}$$

If individual arrives at work early, then $l_{t,g} = 0$. Ideally, the commuters from class g want to arrive at work right at time $t_{*,g}$. However, if a commuter from class g arrives early, then he will get an early arrival penalty $\beta_g e_{t,g}$. Similarly, if a user from class g arrives late, then a late arrival penalty $\gamma_g l_{t,g}$ will apply to him. The cost of an individual of class g departing at time interval t is defined as follows:

$$C_{t,g} \equiv \alpha_g TT_t + \beta_g e_{t,g} + \gamma_g l_{t,g}. \tag{4.4}$$

There are N_g identical commuters of class g who are assumed to drive from home to work along the link, each in his or her own car. The link passes through a single bottleneck with fixed capacity s. For all the demands to be met, it is required that $sT \geq N_g$, for every $g \in \mathcal{F}$. It is also assumed that $s \leq N_g$, for every $g \in \mathcal{F}$, otherwise, no queue would occur at the bottleneck. The users of class g have a desired arriving time at work $t_{*,g}$. Each individual can schedule their own departure time, regarding the travel time and traffic delay. $(r_{t,g})_{(t,g)\in\mathcal{T}\times\mathcal{F}}$ is the dynamic user equilibrium if for each class g, the cost of the time period t with departure $(r_{t,g} > 0)$

is equal to the equilibrium cost $C_{*,g}$, i.e.

$$0 \leq r_{t,g} \perp C_{t,g} - C_{*,g} \geq 0, \quad \text{for all } t \in \mathcal{T}$$

where $C_{t,g}$ is given by (4.4). Together with (4.1), (4.2) , (4.3) and demand constraint below

$$\sum_{t \in \mathcal{T}} r_{t,g} = N_g, \quad \text{for all } g \in \mathcal{F},$$

the dynamic single bottleneck with heterogeneous users model can be formulated as the following mixed LCP:

$$\begin{aligned}
0 \leq r_{t,g} \quad &\perp \quad \alpha_g TT_t + \beta_g e_{t,g} + \gamma_g \left(e_{t,g} - (t_{*,g} - t - TT_t)\right) - C_{*,g} \geq 0, \\
&\quad\ \forall\, t \in \mathcal{T}, \forall\, g \in \mathcal{F} \\
0 \leq TT_t \quad &\perp \quad TT_t - \left(TT_{t-1} + \frac{\sum_{g \in \mathcal{F}} r_{t,g} - s}{s} \right) \geq 0, \quad \forall\, t \in \mathcal{T} \\
0 \leq e_{t,g} \quad &\perp \quad e_{t,g} - (t_{*,g} - t - TT_t) \geq 0, \quad \forall\, t \in \mathcal{T}, \forall\, g \in \mathcal{F} \\
&\quad\ \sum_{t \in \mathcal{T}} r_{t,g} - N_g = 0, \quad \forall\, g \in \mathcal{F}
\end{aligned} \tag{4.5}$$

Ramadurai et al. [51] show that the mixed LCP (4.5) is equivalent to the following

LCP:

$$
\begin{aligned}
0 \le r_{t,g} \quad &\perp \quad \alpha_g TT_t + \beta_g e_{t,g} + \gamma_g \left(e_{t,g} - (t_{*,g} - t - TT_t)\right) - C_{*,g} \ge 0, \\
&\qquad \forall\, t \in \mathcal{T}, \forall\, g \in \mathcal{F} \\
0 \le TT_t \quad &\perp \quad TT_t - \left(TT_{t-1} + \frac{\displaystyle\sum_{g\in\mathcal{F}} r_{t,g} - s}{s} \right) \ge 0, \quad \forall\, t \in \mathcal{T} \\
0 \le e_{t,g} \quad &\perp \quad e_{t,g} - (t_{*,g} - t - TT_t) \ge 0, \quad \forall\, t \in \mathcal{T}, \forall\, g \in \mathcal{F} \\
0 \le C_{*,g} \quad &\perp \quad \sum_{t\in\mathcal{T}} r_{t,g} - N_g \ge 0, \quad \forall\, g \in \mathcal{F}.
\end{aligned}
\tag{4.6}
$$

The last equation in mixed LCP (4.5) is replaced by a complementarity condition in LCP (4.6). If $\mathcal{F}$ is only a singleton, the single bottleneck model with heterogeneous commuters classes is reduced to a special case: the single bottleneck model with homogeneous commuters classes.

4.2 Existence of Equilibria

In this section, the existence of DUE in the single bottleneck model is established by applying the fundamental existence result for LCP. Without changing the complementarity problem itself, we normalize LCP (4.6), and rewrite it in a standard LCP fashion for the convenience of analysis. By multiplying the right-hand side of the first condition in LCP (4.6) by $1/(\alpha_g + \gamma_g)$, the right-hand side of the second condition by s and the right-hand side of the last condition by $1/(\alpha_g + \gamma_g)$, then LCP (4.6) is equivalent to the following LCP($\mathbf{q}, \mathbf{M}$) of the standard form:

$$0 \le \mathbf{x} \perp \mathbf{M}\mathbf{x} + \mathbf{q} \ge 0,$$

where the vector of variables $\mathbf{x}$ is of the following form:

$$\mathbf{x} \equiv \begin{pmatrix} \mathbf{r} \\ \mathbf{TT} \\ \mathbf{e} \\ \boldsymbol{C_*} \end{pmatrix}$$

where $\mathbf{r} \equiv (r_{t,g})_{(t,g)\in\mathcal{T}\times\mathcal{F}}$, $\mathbf{TT} \equiv (TT_t)_{t\in\mathcal{T}}$, $\mathbf{e} \equiv (e_{t,g})_{(t,g)\in\mathcal{T}\times\mathcal{F}}$ and $\boldsymbol{C_*} \equiv (C_{*,g})_{g\in\mathcal{F}}$.

$\mathbf{q}$ is the constant vector:

$$\mathbf{q} \equiv \begin{pmatrix} -\frac{\gamma_1}{\alpha_1+\gamma_1}(t_{*,1}-t)_{t\in\mathcal{T}} \\ \vdots \\ -\frac{\gamma_g}{\alpha_g+\gamma_g}(t_{*,g}-t)_{t\in\mathcal{T}} \\ s\mathbf{1} \\ -(t_{*,1}-t)_{t\in\mathcal{T}} \\ \vdots \\ -(t_{*,g}-t)_{t\in\mathcal{T}} \\ -\frac{N_1}{\alpha_1+\gamma_1} \\ \vdots \\ -\frac{N_g}{\alpha_g+\gamma_g} \end{pmatrix},$$

and the matrix $\mathbf{M}$, partitioned in accordance with the vectors $\mathbf{x}$ and $\mathbf{q}$ is given by

$$\mathbf{M} \equiv \begin{pmatrix} \mathbf{0} & \mathbf{M}_1 & \mathbf{M}_2 & -\mathbf{M}_3^T \\ -\mathbf{M}_1^T & \mathbf{S} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_1 & \mathbf{I} & \mathbf{0} \\ \mathbf{M}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4.7}$$

where

$$\mathbf{M}_1 \equiv \begin{pmatrix} \mathbf{I} \\ \vdots \\ \mathbf{I} \end{pmatrix} \in \Re^{(|\mathcal{T}|\times|\mathcal{F}|)} \times \Re^{|\mathcal{T}|},$$

$$\mathbf{M}_2 \equiv \begin{pmatrix} \frac{\beta_1+\gamma_1}{\alpha_1+\gamma_1}\mathbf{I} & & \\ & \ddots & \\ & & \frac{\beta_g+\gamma_g}{\alpha_g+\gamma_g}\mathbf{I} \end{pmatrix} \in \Re^{(|\mathcal{T}|\times|\mathcal{F}|)} \times \Re^{(|\mathcal{T}|\times|\mathcal{F}|)}$$

$$\mathbf{M}_3 \equiv \begin{pmatrix} \frac{1}{\alpha_1+\gamma_1}\mathbf{1}^T & & \\ & \ddots & \\ & & \frac{1}{\alpha_g+\gamma_g}\mathbf{1}^T \end{pmatrix} \in \Re^{|\mathcal{F}|} \times \Re^{(|\mathcal{T}|\times|\mathcal{F}|)}$$

and

$$\mathbf{S} \equiv \begin{pmatrix} s & & & \\ -s & \ddots & & \\ & \ddots & \ddots & \\ & & -s & s \end{pmatrix} \in \Re^{|\mathcal{T}|} \times \Re^{|\mathcal{T}|};$$

$\mathbf{I}$ is the identical matrix with appropriate dimension, $\mathbf{1}$ is all ones $\Re^{|\mathcal{T}|}$-vector.

Because $\mathbf{S}$ is a positive definite matrix, the matrix $\mathbf{M}$ can be decomposed as the sum of a positive semi-definite matrix $\mathbf{P}$ and a nonnegative matrix $\mathbf{Q}$, i.e $\mathbf{M} = \mathbf{P} + \mathbf{Q}$, where

$$\mathbf{P} \equiv \begin{pmatrix} \mathbf{0} & \mathbf{M}_1 & \mathbf{0} & -\mathbf{M}_3^T \\ -\mathbf{M}_1^T & \mathbf{S} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{M}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \text{ and } \mathbf{Q} \equiv \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{M}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

In Chapter 1, two matrix classes are introduced: $\mathbf{R}_0$-matrix class and copositive matrix class, whose definitions are given by (4) and (5), respectively. In the following lemmas, we show that $\mathbf{M}$ given by (4.7) belongs to these two matrix classes.

Lemma 20. *$\boldsymbol{M}$ is a copositive matrix.*

Proof. For any $\mathbf{v} \equiv (\mathbf{r}, \mathbf{TT}, \mathbf{e}, \boldsymbol{C}_*) \in \Re_+^{3(|\mathcal{T}|+|\mathcal{F}|)}$,

$$\mathbf{v}^T\mathbf{M}\mathbf{v} = \mathbf{v}^T\mathbf{P}\mathbf{v} + \mathbf{v}^T\mathbf{Q}\mathbf{v} \geq 0,$$

since matrix $\mathbf{P}$ is a positive semi-definite matrix and $\mathbf{Q}$ is a nonnegative matrix. By definition, $\mathbf{M}$ is a copositive matrix. □

Lemma 21. *$\boldsymbol{M}$ is an $\boldsymbol{R}_0$-matrix.*

Proof. To prove the R_0-matrix property, we need to show that zero is the only solution to LCP $(\mathbf{0}, \mathbf{M})$. In other words: $\mathbf{M}\mathbf{v} \geq 0, \mathbf{v} \geq 0$ and $\mathbf{v}^T\mathbf{M}\mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

If $\mathbf{v} = \mathbf{0}$, it is obvious that $\mathbf{M}\mathbf{v} \geq 0, \mathbf{v} \geq 0$ and $\mathbf{v}^T\mathbf{M}\mathbf{v} = 0$ holds.

Next, we prove "only if" direction. Assume that $\mathbf{v}^T\mathbf{M}\mathbf{v} = \mathbf{v}^T\mathbf{P}\mathbf{v} + \mathbf{v}^T\mathbf{Q}\mathbf{v} = 0$. Since $\mathbf{Q}$ is a nonnegative matrix and $\mathbf{v} \geq 0$, it yields $\mathbf{v}^T\mathbf{P}\mathbf{v} = (\mathbf{TT}^T)\mathbf{S}(\mathbf{TT}) + \mathbf{e}^T\mathbf{I}\mathbf{e} = 0$. Since $\mathbf{S}$ and $\mathbf{I}$ are positive definite matrices, it follows that $\mathbf{TT} = 0$ and $\mathbf{e} = 0$. As a result, we have

$$\mathbf{M}\mathbf{v} = \begin{pmatrix} -\mathbf{M}_3^T\mathbf{C}_* \\ -\mathbf{M}_1^T\mathbf{r} \\ \mathbf{0} \\ \mathbf{M}_3\mathbf{r} \end{pmatrix} \geq 0.$$

which implies that $\mathbf{r} = 0$ and $\mathbf{C}_* = 0$. Thus, we obtain that $\mathbf{v} = 0$. □

Recall that Proposition 6 in Chapter 1 indicates that if the matrix $\mathbf{M}$ is a copositive $\mathbf{R}_0$-matrix, then LCP $(\mathbf{q}, \mathbf{M})$ is solvable for any vector $\mathbf{q}$. Therefore, together with Lemma 20 and Lemma 21, the existence of a solution to LCP $(\mathbf{q}, \mathbf{M})$ is established in the following theorem. Furthermore, since the model is formulated in a LCP fashion, the equilibria can be found by Lemke's method.

Theorem 22. *LCP* $(\boldsymbol{q}, \boldsymbol{M})$ *has a solution and the solution can be computed by Lemke's method.*

4.3 The Heterogeneous Commuters Classes

In this section, we study the departure patterns in the single bottleneck model with heterogeneous commuters classes. In particular, we consider a special case where all the commuters classes have the same preferred arrival times, i.e. $t_{*,g}$ for all $g \in \mathcal{F}$ are identical, and are equal to t_*.

If a user from class g departs at time interval t and arrives at destination early, then we say that class g has an "early departure" at time interval t. Similarly, class g has a "late departure" at time t, if a user from class g leaves at time interval t and arrives at work late. Recall that all the classes have the same preferred arrival time. As a result, if an early departure happens at time interval t, then any departures before time t must be an early departure as well. For the same reason, once a late departure occurs, then any departures afterward must be a late departure. Thus, there exists a certain time interval t, such that any departures before time t are early departures, while any departures after time t are late departures. Therefore, departures patterns can be divided into two types in the case with identical preferred arrival times: the order of early departures and the order of late departures.

For each $g \in \mathcal{F}$, define

$$\widehat{TT}_{t,g} \equiv \min\left(\frac{C_{*,g} - \beta_g(t_* - t)}{\alpha_g - \beta_g}, \frac{C_{*,g} + \gamma_g(t_* - t)}{\alpha_g + \gamma_g}\right), \qquad \text{for every } t \in \mathcal{T}. \tag{4.8}$$

The following lemma states that the class who has a departure at a certain time interval t_o must has the greatest $\widehat{TT}_{t_o,g}$ among all the classes.

Lemma 23. *If class* g *has an early departure at some time interval* $t_o \in \mathcal{T}$*, then*

$$TT_{t_o} = \max_{f \in \mathcal{F}}\left(\widehat{TT}_{t_o,f}\right) = \widehat{TT}_{t_o,g} = \frac{C_{*,g} - \beta_g(t_* - t_o)}{\alpha_g - \beta_g};$$

if class g has a late departure at some time interval $t_o \in \mathcal{T}$, then

$$TT_{t_o} = \max_{f \in \mathcal{F}} \left(\widehat{TT}_{t_o,f} \right) = \widehat{TT}_{t_o,g} = \frac{C_{*,g} + \gamma_g (t_* - t_o)}{\alpha_g + \gamma_g}.$$

The proof of Lemma 23 can be found in Appendix A.

By using the above lemma, let us investigate the order of early departures. We sort the relative early arrival penalties $\dfrac{\beta_g}{\alpha_g}$ for all $g \in \mathcal{F}$ in ascending order. The following proposition reveals the corresponding early departure pattern.

Proposition 24. *Suppose that $\dfrac{\beta_1}{\alpha_1} < \dfrac{\beta_2}{\alpha_2} < \cdots < \dfrac{\beta_{|\mathcal{F}|}}{\alpha_{|\mathcal{F}|}}$. The early departures must follow the order: class 1, class 2, ..., class n, but some classes might be skipped.*

Proof. Suppose that $\dfrac{\beta_g}{\alpha_g} < \dfrac{\beta_f}{\alpha_f}$ for some $g, f \in \mathcal{F}$. We claim that class g cannot have any early departures after class f has an early departure. Assume for the sake of contradiction that class f has an early departure at time interval $\bar{t}$ and class g has an early departures at some time interval $\tilde{t}$, where $\bar{t} < \tilde{t}$. By Lemma 23, it follows that

$$\frac{C_{*,g} - \beta_g (t_* - \bar{t})}{\alpha_g - \beta_g} \leq \frac{C_{*,f} - \beta_f (t_* - \bar{t})}{\alpha_f - \beta_f} \tag{4.9}$$

and

$$\frac{C_{*,g} - \beta_g (t_* - \tilde{t})}{\alpha_g - \beta_g} \geq \frac{C_{*,f} - \beta_f (t_* - \tilde{t})}{\alpha_f - \beta_f}. \tag{4.10}$$

Inequality (4.9) implies

$$\bar{t} \geq t_* - \frac{(1 - \dfrac{\beta_g}{\alpha_g}) \dfrac{C_{*,f}}{\alpha_f} - (1 - \dfrac{\beta_f}{\alpha_f}) \dfrac{C_{*,g}}{\alpha_g}}{\dfrac{\beta_f}{\alpha_f} - \dfrac{\beta_g}{\alpha_g}},$$

and inequality (4.10) implies

$$\tilde{t} \leq t_* - \frac{(1 - \dfrac{\beta_g}{\alpha_g}) \dfrac{C_{*,f}}{\alpha_f} - (1 - \dfrac{\beta_f}{\alpha_f}) \dfrac{C_{*,g}}{\alpha_g}}{\dfrac{\beta_f}{\alpha_f} - \dfrac{\beta_g}{\alpha_g}},$$

so it implies $\tilde{t} \leq \bar{t}$. However, it contradicts the assumption $\tilde{t} > \bar{t}$. □

Proposition 24 is consistent with the intuition that commuters with relatively high early arrival cost do not have incentive to get to work too early. Analogously, the following proposition investigates the late departures order.

Proposition 25. *Suppose that* $\frac{\gamma_1}{\alpha_1} > \frac{\gamma_2}{\alpha_2} > \cdots > \frac{\gamma_{|\mathcal{F}|}}{\alpha_{|\mathcal{F}|}}$. *The late departures must follow the order: class 1, class 2, ..., class* n*, but some classes might be skipped.*

Proof. Suppose that $\frac{\gamma_f}{\alpha_f} > \frac{\gamma_g}{\alpha_g}$. We claim that class g cannot have any late departures before class f has a late departure. Assume for the sake of contradiction that class g has late departures at time interval $\tilde{t}$ and class f has late departures at time interval $\bar{t}$, where $\tilde{t} < \bar{t}$. By Lemma 23, it follows that

$$\frac{C_{*,g} + \gamma_g(t_* - \bar{t})}{\alpha_g + \gamma_g} \leq \frac{C_{*,f} + \gamma_f(t_* - \bar{t})}{\alpha_f + \gamma_f} \tag{4.11}$$

and

$$\frac{C_{*,g} + \gamma_g(t_* - \tilde{t})}{\alpha_g + \gamma_g} \geq \frac{C_{*,f} + \gamma_f(t_* - \tilde{t})}{\alpha_f + \gamma_f}. \tag{4.12}$$

Inequality (4.11) implies

$$\bar{t} \leq t_* - \frac{(1 + \frac{\gamma_g}{\alpha_g})\frac{C_{*,f}}{\alpha_f} - (1 + \frac{\gamma_f}{\alpha_f})\frac{C_{*,g}}{\alpha_g}}{\frac{\gamma_g}{\alpha_g} + \frac{\gamma_f}{\alpha_f}},$$

and inequality (4.12) implies

$$\tilde{t} \geq t_* - \frac{(1 + \frac{\gamma_g}{\alpha_g})\frac{C_{*,f}}{\alpha_f} - (1 + \frac{\gamma_f}{\alpha_f})\frac{C_{*,g}}{\alpha_g}}{\frac{\gamma_g}{\alpha_g} + \frac{\gamma_f}{\alpha_f}}.$$

As a result, we have $\bar{t} \leq \tilde{t}$, which contradicts the assumption $\tilde{t} < \bar{t}$. □

Proposition 25 implies that the class with relatively small late arrival penalty would arrive at work later than the ones with larger late arrival penalty.

The results obtained in Arnott et al. [4] partially agree with Proposition 24 and Proposition 25. For example, Arnott et al. [4] obtain a similar early departure pattern result as Proposition 24 under a more restrict assumption that users have the same relative cost of late arrival $\frac{\gamma_g}{\beta_g}$ and preferred arrival time. Similarly, Proposition 25 is also more general than the corresponding one in Arnott et al. [4].

4.4 The Homogeneous Commuters Classes

If the set of class is a singleton, the heterogeneous commuters classes model is recognized as the homogeneous commuters class model. In this section, the uniqueness of the equilibrium cost of the homogeneous user case is established. Moreover, an algorithm other than the Lemke's method is developed to solve this special case. Since there is only one commuters class in the homogeneous case, the subscript g can be drop in the LCP (4.6) for the heterogeneous case.

4.4.1 The uniqueness of the equilibrium cost C_*

In this subsection, the cost equilibrium C_* is proved to be unique in the single bottleneck with homogeneous users class. To this end, first we drop the demand satisfaction constraint given by $\sum_{t \in \mathcal{T}} r_t = N$ in LCP (4.6). As a result, the remaining complementarity conditions in LCP (4.6) is equivalent to the following LCP, which is parameterized by $C_* > 0$:

$$\begin{aligned} 0 \leq r_t \quad & \perp \quad TT_t + \frac{\beta + \gamma}{\alpha + \gamma} e_t - \frac{C_* + \gamma(t_* - t)}{\alpha + \gamma} \geq 0, \quad \forall\, t \in \mathcal{T} \\ 0 \leq TT_t \quad & \perp \quad -r_t + s(TT_t - TT_{t-1}) + s \geq 0, \quad \forall\, t \in \mathcal{T} \\ 0 \leq e_t \quad & \perp \quad TT_t + e_t - (t_* - t) \geq 0, \quad \forall\, t \in \mathcal{T}, \end{aligned} \tag{4.13}$$

For any given C_*, let $\mathbf{x}(C_*) \equiv (\mathbf{r}(C_*), \mathbf{TT}(C_*), \mathbf{e}(C_*))$ be an arbitrary solution to the above parameterized LCP. To show there there is a unique C_* at the equilibrium of LCP (4.6), we prove that there is only one C_* satisfying the demand satisfaction constraint $\sum_{t \in \mathcal{T}} r_t(C_*) = N$, even though there may be multiple subvectors $\mathbf{r}(C_*)$. First of all, the following lemma provides the expressions of $\mathbf{r}(C_*)$ and $\mathbf{TT}(C_*)$.

Lemma 26. The subvector $\mathbf{TT}(C_*)$ is unique and given by

$$\begin{aligned}
TT_t(C_*) &= 0, && \text{if } t \le \max\left(0, t_* - \frac{C_*}{\beta}\right)\\
TT_t(C_*) &= \frac{C_* - \beta(t_* - t)}{\alpha - \beta}, && \text{if } \max\left(0, t_* - \frac{C_*}{\beta}\right) < t < \max\left(0, t_* - \frac{C_*}{\alpha}\right)\\
TT_t(C_*) &= \frac{C_* + \gamma(t_* - t)}{\alpha + \gamma}, && \text{if } \max\left(0, t_* - \frac{C_*}{\alpha}\right) \le t < t_* + \frac{C_*}{\gamma}\\
TT_t(C_*) &= \max(TT_{t-1}(C_*) - 1, 0), && \text{if } t \ge t_* + \frac{C_*}{\gamma};
\end{aligned} \tag{4.14}$$

moreover, $0 \le \mathbf{TT}(C_*) \le C_*/\alpha$. The subvector $\mathbf{r}(C_*)$ is multi-valued and satisfies

$$\begin{aligned}
&r_t(C_*) = 0 && \text{if } t < \max\left(0, t_* - \frac{C_*}{\beta}\right)\\
&0 \le r_t(C_*) \le s && \text{if } t = \max\left(0, t_* - \frac{C_*}{\beta}\right)\\
&r_t(C_*) = s(TT_t - TT_{t-1} + 1) && \text{if } \max\left(0, t_* - \frac{C_*}{\beta}\right) < t < t_* + \frac{C_*}{\gamma}\\
&0 \le r_t(C_*) \le s(1 - TT_{t-1}) && \text{if } t = t_* + \frac{C_*}{\gamma}\\
&r_t(C_*) = 0 && \text{if } t > t_* + \frac{C_*}{\gamma}.
\end{aligned} \tag{4.15}$$

The proof for Lemma 26 can be found in Appendix B. Recalling the ceiling $\lceil x \rceil$ = smallest integer $\ge x$ and floor $\lfloor x \rfloor$ = largest integer $\le x$ of a real number x, we let

$$t_1 \equiv \left\lfloor t_* - \frac{C_*}{\beta} \right\rfloor, \quad t_2 \equiv \left\lceil t_* - \frac{C_*}{\alpha} \right\rceil, \quad \text{and} \quad t_3 \equiv \left\lceil t_* + \frac{C_*}{\gamma} \right\rceil.$$

It follows that $TT_{t_3-1}(C_*) = \dfrac{\gamma}{\alpha+\gamma}\left(\dfrac{C_*}{\gamma} + 1 - \left\lceil \dfrac{C_*}{\gamma} \right\rceil\right)$.

To show that there is only one C_* satisfying $\sum_{t \in \mathcal{T}} r_t(C_*) = N$, we demonstrate that the trajectory of $\sum_{t \in \mathcal{T}} r_t(C_*)$ is monotonically increasing in C_* as follows. Let

$\Theta \equiv \{\theta_1, \theta_2, ...\}$ be the sequence such that

$$\Theta = \{m\beta, \text{ where } m \in \mathbb{Z}_+\} \bigcup \{n\gamma, \text{ where } n \in \mathbb{Z}_+\} \quad \text{and} \quad \theta_i < \theta_j \text{ if } i < j.$$

We claim that the trajectory of $\sum_{t\in\mathcal{T}} r_t(C_*)$ is strictly increasing (see Figure 4.4.1). More specifically, $\sum_{t\in\mathcal{T}} r_t(C_*)$ is monotonically increasing in any open interval (θ_i, θ_{i+1}). At the each breakpoint θ_i, $\sum_{t\in\mathcal{T}} r_t(\theta_i)$ may have multiple values, and they are greater than the summation at its left and less than the summation at its right. The details are described as follows.

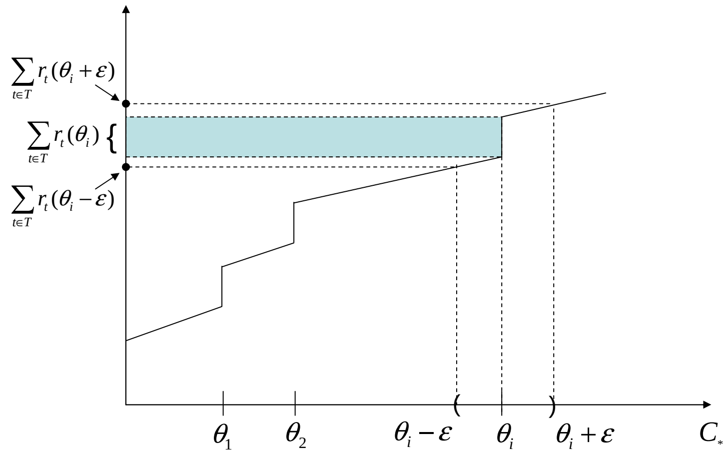

Figure 4.2: Trajectory of $\sum_{t\in\mathcal{T}} r_t(C_*)$

If $\theta_i < C_* < \theta_{i+1}$, for any $\theta_i \in \Theta$, then by Lemma 26, we have

$$\begin{aligned}
\sum_{t\in\mathcal{T}} r_t(C_*) &= \sum_{t=t_1+1}^{t_3-1} r_t(C_*) \\
&= \sum_{t=t_1+1}^{t_3-1} (TT_t(C_*) - TT_{t-1}(C_*) + 1) \\
&= s\,(TT_{t_3-1}(C_*) - TT_{t_1}(C_*)) + s(t_3 - t_1 - 1),
\end{aligned}$$

where $TT_{t_3-1}(C_*) = \dfrac{C_* - \gamma\lceil\frac{C_*}{\gamma}\rceil + \gamma}{\alpha+\gamma}$ and $TT_{t_1}(C_*) = 0$. It results in

$$\sum_{t\in\mathcal{T}} r_t(C_*) = s\frac{C_* + \alpha\lceil\frac{C_*}{\gamma}\rceil}{\alpha+\gamma} + s\left\lceil\frac{C_*}{\beta}\right\rceil - s\frac{\alpha}{\alpha+\gamma}, \tag{4.16}$$

which verifies that $\sum_{t\in\mathcal{T}} r_t(C_*)$ is increasing in the open interval (θ_i, θ_i+1). Next, let us examine $\sum_{t\in\mathcal{T}} r_t(C_*)$ at any arbitrary breakpoint θ_i.

If $C_* = \theta_i$, for some $\theta_i \in \Theta$, then

$$\sum_{t\in\mathcal{T}} r_t(C_*) = \begin{cases} s\dfrac{\gamma}{\alpha+\gamma} + s\left(\dfrac{C_*}{\gamma} + \left\lceil\dfrac{C_*}{\beta}\right\rceil - 1\right) + r_{t_3}(C_*) & \text{if } \dfrac{C_*}{\beta} \notin \mathbb{Z} \\ s\dfrac{C_* - \gamma\lceil\frac{C_*}{\gamma}\rceil + \gamma}{\alpha+\gamma} + s\left(\left\lceil\dfrac{C_*}{\gamma}\right\rceil + \dfrac{C_*}{\beta} - 1\right) + r_{t_1}(C_*) & \text{if } \dfrac{C_*}{\gamma} \notin \mathbb{Z} \\ s\dfrac{\gamma}{\alpha+\gamma} + s\left(\dfrac{C_*}{\gamma} + \dfrac{C_*}{\beta} - 1\right) + r_{t_1}(C_*) + r_{t_3}(C_*) & \text{if } \dfrac{C_*}{\gamma}, \dfrac{C_*}{\beta} \in \mathbb{Z} \end{cases} \tag{4.17}$$

where $0 \le r_{t_1}(C_*) \le s$ and $0 \le r_{t_3}(C_*) \le s(1 - TT_{t_3-1}(C_*))$. Notice that the value of $\sum_{t\in\mathcal{T}} r_t(C_*)$ is not unique at $C_* = \theta_i$. We claim that for any $\varepsilon > 0$ sufficiently small, we have the following inequalities:

$$\sum_{t\in\mathcal{T}} r_t(C_* - \varepsilon) < \sum_{t\in\mathcal{T}} r_t(C_*) < \sum_{t\in\mathcal{T}} r_t(C_* + \varepsilon) \tag{4.18}$$

for any values of $r_{t_1}(C_*)$ and $r_{t_3}(C_*)$ in the intervals $[0, s]$ and $[0, s(1 - TT_{t_3-1}(C_*)]$, respectively. Figure 4.4.1 illustrates the rough configuration around a breakpoint θ_i.

First, let us consider $\sum_{t\in\mathcal{T}} r_t(C_* + \varepsilon)$. For any sufficiently small $\varepsilon > 0$, deriving from (4.16), we have

$$\sum_{t\in\mathcal{T}} r_t(C_*+\varepsilon) = \begin{cases} s\dfrac{\varepsilon}{\alpha+\gamma} + s\left(\dfrac{C_*}{\gamma} + \left\lceil\dfrac{C_*}{\beta}\right\rceil - 1\right) + s & \text{if } \dfrac{C_*}{\beta} \notin \mathbb{Z} \\ s\dfrac{C_* + \varepsilon - \gamma\lceil\frac{C_*}{\gamma}\rceil + \gamma}{\alpha+\gamma} + s\left(\left\lceil\dfrac{C_*}{\gamma}\right\rceil + \dfrac{C_*}{\beta} - 1\right) + s & \text{if } \dfrac{C_*}{\gamma} \notin \mathbb{Z} \\ s\dfrac{\varepsilon}{\alpha+\gamma} + s\left(\dfrac{C_*}{\gamma} + \dfrac{C_*}{\beta} - 1\right) + 2s & \text{if } \dfrac{C_*}{\gamma}, \dfrac{C_*}{\beta} \in \mathbb{Z} \end{cases} \tag{4.19}$$

By equalities (4.17), we can derive the following inequalities:

$$\sum_{t\in\mathcal{T}} r_t(C_*) \leq \begin{cases} s\left(\dfrac{C_*}{\gamma} + \left\lceil\dfrac{C_*}{\beta}\right\rceil - 1\right) + s & \text{if } \dfrac{C_*}{\beta} \notin \mathbb{Z} \\ s\dfrac{C_* - \gamma\lceil\frac{C_*}{\gamma}\rceil + \gamma}{\alpha+\gamma} + s\left(\left\lceil\dfrac{C_*}{\gamma}\right\rceil + \dfrac{C_*}{\beta} - 1\right) + s & \text{if } \dfrac{C_*}{\gamma} \notin \mathbb{Z} \\ s\left(\dfrac{C_*}{\gamma} + \dfrac{C_*}{\beta} - 1\right) + 2s & \text{if } \dfrac{C_*}{\gamma}, \dfrac{C_*}{\beta} \in \mathbb{Z}. \end{cases} \tag{4.20}$$

If we compare equalities (4.19) and inequalities (4.20), it yields that for any sufficiently small $\varepsilon > 0$,

$$\sum_{t\in\mathcal{T}} r_t(C_*) < \sum_{t\in\mathcal{T}} r_t(C_* + \varepsilon).$$

Next, let us consider $\sum_{t\in\mathcal{T}} r_t(C_*-\varepsilon)$. For any sufficiently small $\varepsilon > 0$, by Lemma

26 we have

$$\sum_{t\in\mathcal{T}} r_t(C_* - \varepsilon) = \begin{cases} s\dfrac{\gamma-\varepsilon}{\alpha+\gamma} + s\left(\dfrac{C_*}{\gamma} + \left\lceil\dfrac{C_*}{\beta}\right\rceil - 1\right) & \text{if } \dfrac{C_*}{\beta} \notin \mathbb{Z} \\ s\dfrac{C_* - \varepsilon - \gamma\lceil\frac{C_*}{\gamma}\rceil + \gamma}{\alpha+\gamma} + s\left(\left\lceil\dfrac{C_*}{\gamma}\right\rceil + \dfrac{C_*}{\beta} - 1\right) & \text{if } \dfrac{C_*}{\gamma} \notin \mathbb{Z} \\ s\dfrac{\gamma-\varepsilon}{\alpha+\gamma} + s\left(\dfrac{C_*}{\gamma} + \dfrac{C_*}{\beta} - 1\right) & \text{if } \dfrac{C_*}{\gamma}, \dfrac{C_*}{\beta} \in \mathbb{Z} \end{cases} \tag{4.21}$$

If we compare equalities (4.17) and (4.21), it shows that for ε sufficiently small,

$$\sum_{t\in\mathcal{T}} r_t(C_* - \varepsilon) < \sum_{t\in\mathcal{T}} r_t(C_*).$$

By the above argument, the trajectory of $\sum_{t\in\mathcal{T}} r_t(C_*)$ is increasing in C_*. Thus, for a given $N > 0$, there is only one C_* such that $\sum_{t\in\mathcal{T}} r_t(C_*) = N$.

4.4.2 Algorithm for the homogeneous commuters classes case

In the previous subsection, we derive the explicit expression for $\sum_{t\in\mathcal{T}} r_t(C_*)$ in the homogeneous users case, and also show that the curve of $\sum_{t\in\mathcal{T}} r_t(C_*)$ is increasing in C_*. Following the argument for the uniqueness of equilibrium cost C_*, we are able to construct an algorithm other than Lemke's method to solve the homogeneous users case.

Since $r_t(C_*) = 0$ for all $t > t_* + \frac{C_*}{\gamma}$ do not contribute to $\sum_{t\in\mathcal{T}} r_t(C_*)$, we are only interested in those time intervals $t \in \mathcal{T}$, such that $t \le t_* + \frac{C_*}{\gamma}$. In other words, C_* is only required to be large enough to satisfy $t_* + \frac{C_*}{\gamma} \le |\mathcal{T}|$. Equivalently, C_* must be in the interval $(0, \gamma(|\mathcal{T}| - t_*)]$. The idea of the algorithm is that we search for C_* along the interval $(0, \gamma(|\mathcal{T}| - t_*)]$ until we find the one satisfying $\sum_{t\in\mathcal{T}} r_t(C_*) = N$. The details of the algorithm is elaborated as follows:

Step 1: Sort the breakpoints.

Let $\Theta \equiv \{\theta_1, \theta_2, ..., \theta_k\}$ be the breakpoint set such that

$$\Theta = \{m\beta \leq \gamma\left(|\mathcal{T}| - t_*\right), m \in \mathbb{Z}_+\} \bigcup \{n\gamma \leq \gamma\left(|\mathcal{T}| - t_*\right), n \in \mathbb{Z}_+\} \text{ and } \theta_i < \theta_j \text{ if } i < j.$$

Step 2: Check each subinterval.

$i = 1 : k - 1$

If there exists a solution $C_* \in (\theta_i, \theta_{i+1})$ such that the following equation holds:

$$s\frac{C_* + \alpha\lceil\frac{C_*}{\gamma}\rceil}{\alpha + \gamma} + s\left\lceil\frac{C_*}{\beta}\right\rceil - s\frac{\alpha}{\alpha + \gamma} = N$$

where

$$\left\lceil\frac{C_*}{\gamma}\right\rceil = \frac{\theta_i}{\gamma} + 1, \left\lceil\frac{C_*}{\beta}\right\rceil = \frac{\theta_{i+1}}{\beta} \quad \text{if } \frac{\theta_i}{\gamma} \in \mathbb{Z}$$

$$\left\lceil\frac{C_*}{\gamma}\right\rceil = \left\lceil\frac{\theta_i}{\gamma}\right\rceil, \left\lceil\frac{C_*}{\beta}\right\rceil = \frac{\theta_i}{\beta} + 1 \quad \text{if } \frac{\theta_i}{\gamma} \notin \mathbb{Z}$$

then return the equilibrium cost C_*. Otherwise, go to the next subinterval.

Step 3: Check each breakpoint.

$i = 1 : k$

If θ_i satisfies $\sum_{t\in\mathcal{T}} r_t(\theta_i) = N$, where $r_t(\bullet)$ is given by equation (4.17), the equilibrium cost is equal to θ_i. Otherwise, go to the next breakpoint.

Step 4: Solve the remaining variables.

Once C_* is determined, we can immediately obtain **TT** and **r** by using Lemma 26. In addition, **e** can be derived by

$$e_t = \max(0, t_* - t - TT_t), \quad \text{for all } t \in \mathcal{T}$$

by the last complementarity condition in LCP (4.13). □

Theorem 22 implies that the equilibria in the homogeneous commuters class case can be found by Lemke's method. However, its computational complexity can be exponential in the worst case. Murty [41] gives an example to show exponential growth for Lemke's algorithm for the LCPs. On the other hand, the following proposition shows that the above specified algorithm is expected to have the complexity of $O(\kappa^2)$, where $\kappa \equiv \lceil \frac{\gamma(|\mathcal{T}| - t_*)}{\beta} \rceil$. By the definition, κ is actually the maximal number of the elements in the breakpoint set $\{m\beta \leq \gamma\left(|\mathcal{T}| - t_*\right), m \in \mathbb{Z}\}$. In turn, the number of the elements in the breakpoint set $\{n\gamma \leq \gamma\left(|\mathcal{T}| - t_*\right), n \in \mathbb{Z}\}$ cannot be larger than κ, since $\gamma > \beta$.

Proposition 27. *The complexity of the specialized algorithm for the single bottleneck model with homogeneous commuters is $O(\kappa^2)$ in the worse case.*

Proof. In the first step, since the components in the sets $\{m\beta \leq \gamma\left(|\mathcal{T}| - t_*\right), m \in \mathbb{Z}\}$ and $\{n\gamma \leq \gamma\left(|\mathcal{T}| - t_*\right), n \in \mathbb{Z}\}$ are already in order, respectively. The the worst-case complexity of the first step is $O\left(\frac{\kappa(\kappa+1)}{2}\right)$. It takes $O(\kappa)$ time to complete each step afterward. Therefore, the complexity of the whole algorithm is $O(\kappa^2)$. $\square$

Proposition 27 implies that it is more efficient to apply the specialized algorithm to solve for DUE in the homogeneous users case than Lemek's method.

4.5 Conclusion

In this chapter, we study dynamic single bottleneck problems. Depending on the preferred arrival times, travel cost, early and late arrival penalty, the commuters are classified into different groups, i.e. heterogeneous commuters classes. Each commuter choose his or her departure time to avoid the traffic congestion incurred at the bottleneck, and at the same time to minimize his or her cost. This part of thesis is distinguished from the other related work by its linear complementarity formulation. Such expression provides rigorous mathematical analysis of the problem and also allows us to use Lemke's method to find equilibria. For the homogeneous commuters class case, the uniqueness of equilibrium cost is verified and a specified algorithm is developed. For the homogeneous commuters classes case, the departure patterns are investigated.

The formulation in this chapter assumes a linear cost function. A future extension of the current bottleneck model is to implement more general cost functions, such as smooth convex function used in Daganzo [18] and Smith [57]. Moreover, it is important to confirm the uniqueness of equilibrium cost in the heterogeneous commuters classes case.

BIBLIOGRAPHY

[1] H.Z. Aashtiani and T.L. Magnanti (1981). Equilibria on a congested transportation network. *SIAM Journal on Agebraic and Discrete Methods* **2** 213-226.

[2] M. Ahman, D. Burtraw, J.A. Kruger and L. Zetterberg (2005). The Ten-Year Rule: Allocation of Emission Allowances in the EU Emissions Trading System, *Resources for the Future Discussion Paper* 05-30.

[3] M. Ahman and K. Holmgren (2006). New entrant allocation in the Nordic energy sectors: Incentives and options in the EU ETS. *Climate Policy*, in press.

[4] R. Arnott, A. dePalma and R. Lindsey (1988). Schedule delay and departure time decisions with heterogeneous commuters. *Transportation Research Record* **1197** 56–67.

[5] R. Arnott, A. dePalma and R. Lindsey (1993). A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *American Economic Review* **83**(1) 161–179.

[6] M. Bartels, and F. Musgens (2006). Do technology specific CO_2-allocations distort investments? Institute of Energy Economics, University of Cologne, Germany.

[7] F. Bernstein and G.A. DeCroix (2004). Decentralized pricing and capacity decisions in a multitier system with modular assembly, *Management Science* **50** 1293–308.

[8] D. Burtraw, D. Kahn and K. Palmer (2005). CO_2 allowance allocation in the regional greenhouse gas initiative and the effect on electricity investors. Discussion papers dp-05-55, *Resources For the Future*.

[9] D. Burtraw, K. Palmer, R. Bharvirkar, and A. Paul (2001). The effect of allowance allocation on the cost of carbon emission trading. RFF DP 01-30, Resources for the Future, Washington, DC.

[10] G. P. Cachon (2003). Supply chain coordination with contracts, in A. G. de Kok and S. C. Graves editors, *Handbooks in operations research and management science*, **11**, North Holland (Amsterdam 2003).

[11] G. P. Cachon and M. A. Lariviere (2001). Contracting to assure supply: how to share demand forecasts in a supply chain, *Management Science*, **47** 629–646.

[12] G. P. Cachon and S. Netessine (2004). Game theory in supply chain analysis, in *Supply Chain Analysis in the eBusiness Era* (Kluwer 2004).

[13] R. Caldentey and L.M. Wein (2003). Analysis of a decentralized production-inventory system, *Manufacturing and Service Operations Management* **5** 1–17.

[14] M.S. Carr and U.S. Karmarkar (2005). Competition in multiechelon assembly supply chains, *Management Science* **51** 45–59.

[15] D. Chan and J.S. Pang (1982). The generalized quasi-variational inequality problem. *Mathematics of Operations Research* **7**(2) 211–222.

[16] R.W. Cottle, J.S. Pang and R.E. Stone(1992). *The Linear Complementarity Problem*, Academic Press, Boston.

[17] P. Cramton and S. Kerr (2002). Tradable carbon permit auctions, how and why to auction not grandfather, *Energy Policy* **30** 333-345.

[18] C. Daganzo (1985). The uniqueness of a time-dependent equilibrium distribution of arrivals at a single bottleneck. *Transportation Science* **19**(1) 29–37.

[19] D. Demailly and P. Quirion (2006). CO_2 abatement, competitiveness and leakage in the European cement industry under the EU ETS: grandfathering versus output-based allocation. *Climate Policy* **6** 93–113.

[20] D. Ellerman, P.L. Joskow, R. Schmalensee, J.P. Montero and E.M. Bailey (2000). *Markets for Clean Air: The U.S. Acid Rain Program*, Cambridge University Press.

[21] F. Facchinei and J.S. Pang (2003). *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volumes I and II, Springer-Verlag, New York.

[22] T.L. Friesz, R.L. Tobin, T. Smith and P.T. Harker (1983). A nonlinear complementarity formulation and solution procedure for the general derived demand network equilibrium problem. *Journal of Regional Science* **23** 337-359.

[23] T.L. Friesz, D. Bernstein, T.E. Smith, R.L. Tobin and B.W. Wie (1993). A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research* **41**(1) 179–191.

[24] T. L. Friesz and R. Mookherjee (2006). Solving the dynamic network user equilibrium problem with state-Dependent time Shifts. *Transportation Research Part B* **40**(3) 207–229.

[25] M. Grubb and K. Neuhoff (2006). Allocation and competitiveness in the EU emissions trading scheme: Policy overview. *Climate Policy* **6**(1) 7–30.

[26] H. Gurnani and Y. Gerchak (2007). Coordination in decentralized assembly systems with uncertain component yields, *European Journal of Operations Research* **176** (3) 1559–1576.

[27] P.T. Harker (1991). Generalized Nash games and quasivariational inequalities. *European Journal of Operations Research* **54** 81–94.

[28] J.V. Henderson (1981). The economics of staggered work hours. *Journal of Urban Economics* **9** 349–364.

[29] C. Hendrickson and G. Kocur (1981). Schedule delay and departure time decisions in a deterministic model. *Transportation Science* **15**(1) 62–77.

[30] W.J. Hopp and M.L. Spearman (2000). *Factory Physics: The Foundations of Manufacturing Management*, 2nd edition, McGraw-Hill, Boston.

[31] Z. Jemai and F. Karaesmen (2007). Decentralized inventory control in a two-stage capacitated supply chain, *IIE Transactions* **39** 501–512.

[32] A. Krishnamurthy, R. Suri and M. Vernon. Two-moment approximations for throughput and mean queue length of a fork/join station with general input processes, in J.G. Shanthikumar, D.D. Yao, and W.H.M. Zijm, editors, *Stochastic modeling and optimization of manufacturing systems and supply chains*, Kluwer Academic Publishers (Dordretch 2003) pp. 87–126.

[33] A. Krishnamurthy, R. Suri, and M. Vernon (2004). Analysis of a fork/join synchronization station with inputs from coxian servers in a closed queuing network, *Annals of Operations Research* **125** 69–94.

[34] M.A. Lariviere and E.L. Porteus (2001). Selling to a newsvendor: an analysis of price-only contracts, *Manufacturing and Service Operations Management* **3** 293–305.

[35] C.E. Lemke (1965). Bimatrix equilibrium points and mathematical programming. *Management Science* **11** 681–689.

[36] C.E. Lemke and J.T. Howson, JR. (1964). Equilibrium points of bimatrix games. *SIAM Journal on Control and Optimization* **18** 444–454.

[37] P. Linares, F.J. Santos, M. Ventosa, and L. Lapiedra (2006). Impacts of the European emissions trading scheme directive and permit assignment methods on the Spanish electricity sector. *The Energy Journal* **27**(1) 79–98.

[38] R. Lindsey (2004). Existence, uniqueness, and the trip cost function properties of user equilibrium in the bottleneck model with multiple user classes. *Transportation Science* **38**(3) 293-314.

[39] O. Mancino and C. Stampacchia (1972). Convex programming and variational inequalities. *Journal of Optimization Theory and Application* **9** 3–23.

[40] F. Matthes, V. Graichen, and J. Repenning (2005). The environmental effectiveness and economic efficiency of the European Union Emissions Trading Scheme: Structural aspects of allocation. [Prepared for the World Wildlife Federation, Institute for Applied Ecology, Berlin.] Netherlands Bureau for Economic Policy Analysis, Ministry of Economic Affairs. Explanation of CPB Vision on Relationship Emissions Trading - Power Prices. Unpublished memo, The Hague.

[41] K.G. Murty (1978). Computational complexity of complementary pivot methods. *Mathematical Programming Study* **7** 61–73.

[42] J. Nash (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences* **36** (1) 48–49.

[43] Netherlands Bureau for Economic Policy Analysis (2005). Explanation of CPB vision on relationship emissions trading–power prices. Unpublished memo, The Hague.

[44] K. Neuhoff, M. Ahman, R. Betz, J. Cludius, F. Ferrario, K. Holmgren, G. Pal, M. Grubb, F. Matthes, K. Rogge, M. Sato, J. Schleich, J. Sijm, A. Tuerk, C. Kettner, and N. Walker (2006). Comparison of national allocation plans for the period 2008-2012. EPRG Working Paper, Electric Power Research Group, Cambridge University.

[45] K. Neuhoff, M. Grubb, and K. Keats (2005). Impact of allowance allocation on prices and efficiency. CWPE 0552 and EPRG 08, Electric Power Research Group, Cambridge University.

[46] K. Neuhoff, K.K. Martinez, and M. Sato (2006). Allocation, incentives and distortions: The impact of EU ETS emission allowance allocations to the electricity sector. *Climate Policy* **6**(5) 73–91.

[47] G. Newell (1987). The morning commute for nonidentical travelers. *Transportation Science* **21**(2) 74–88.

[48] K. Palmer, D. Burtraw, and D. Kahn (2006). Simple rules for targeting CO_2 allowance allocations to compensate firms. RFF DP 06-28, Resources for the Future, Washington, DC.

[49] J.S. Pang and M. Fukushima (2005). Quasi-variational inequalities, generalized Nash equilibria and multi-leader-follower games. *Computational Management Science* **1** 21–56.

[50] M. Patriksson (1994). *The Traffic Assignment Problem, Models and Methods*, VSP (Utretch).

[51] G. Ramadurai, S.V. Ukkusuri, J. Zhao and J.S. Pang (2007). Dynamic equilibrium in multi-user class single bottleneck models: a complementarity formulation. *In Proceedings of the Triannual Symposium of Transportation System Analysis (TRISTAN) VI conference* Phuket, Thailand, June 2007.

[52] R. Ramakrishnan, and A. Krishnamurthy (2006). Analysis of kitting operations in manufacturing systems. *Asia-Pacific Journal of Operations Research*, accepted.

[53] B. Ran and D.E. Boyce (1996). A link-based variational inequality formulation of ideal dynamic user-optimal route choice problem. *Transportation Research* **4C** (1) 1-12.

[54] J.B. Rosen (1965). Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica* **33** (3) 520–534.

[55] J. Sijm (2006). EU ETS allocation: Evaluation of present system and options beyond 2012. *Zeitshrift für Energiewirtschaft* **30**(4) 285–292.

[56] J. Sijm, K. Neuhoff, and Y. Chen (2006). CO_2 cost pass-through and windfall profits in the power sector. *Climate Policy* **6**(5) 49–72.

[57] M.J. Smith (1984). The existence of a time-dependent equilibrium distribution of arrivals at a single bottleneck. *Transportation Science* **18**(5) 385–394.

[58] P. Som, W. Wilhelm and R. Disney (1994). Kitting process in a stochastic assembly system. *Queueing Systems* **17** 471–490.

[59] J.S. Song and P. Zipkin (2003). Supply chain operations: assemble-to-order systems, in A. G. de Kok and S. C. Graves editors, *Handbooks in operations research and management science*, **11**, North Holland (Amsterdam 2003).

[60] M. Takahashi, H. Osawa and T. Fujisawa (1996). A stochastic assembly system with resume levels, *Asia-Pacific Journal of Operations Research* **15** 127–146.

[61] M. Takahashi, H. Osawa, and T. Fujisawa (2000). On a synchronization queue with two finite buffers. *Queueing Systems* **36** 107–123.

[62] T. Tietenberg (2006). *Emissions Trading-Principles and Practice*. Resources for the Future Press, Washington, DC.

[63] N. VanDerZijpp and K. Koolstra (2002). Multiclass continuous-time equilibrium model for departure time choice on single-bottleneck network. *Transportation Research Record* **1783** 134–141.

[64] W.S. Vickrey (1969). Congestion theory and transport investment. *American Economic Review* **59** 251-261.

[65] Y. Wang and Y. Gerchak (2003). Capacity games in assembly systems with uncertain demand, *Manufacturing and Service Operations Management* **5** 252–267.

[66] J.G. Wardrop (1952). Some theoretical aspects of road traffic research. *Proceeding of the Institute of Civil Engineers, Part II* 325–378.

[67] B.W. Wie, R.L. Tobin, T.L. Friesz and D. Bernstein (1995). A discrete time, nested cost operator approach to the dynamic network user equilibrium problem. *Transportation Science* **29** (1) 79–92.

APPENDIX A
Proof for Lemma 23

The proof of Lemma 23 consists of several steps. Given that $t_{*,g}$ for all $g \in \mathcal{F}$ are the same, step 1-5 reveal the expressions for travel time TT_t over some critical time intervals. Step 5 verifies that the class who has a departure at a certain time interval $\bar{t}$ must has the greatest $\widehat{TT}_{\bar{t},g}$ among all the classes.

Suppose that $t_{*,g} = t_*$ for all $g \in \mathcal{F}$. LCP (4.6) for the heterogeneous commuters classes case can be rewritten as the following LCP:

$$0 \le r_{t,g} \quad \perp \quad TT_t + \frac{\beta_g + \gamma_g}{\alpha_g + \gamma_g} e_{t,g} - \frac{C_{*,g} + \gamma_g (t_* - t)}{\alpha_g + \gamma_g} \ge 0,$$

$$\forall t \in \mathcal{T}, \forall g \in \mathcal{F} \tag{A.1}$$

$$0 \le TT_t \quad \perp \quad -\sum_{f \in \mathcal{F}} r_{t,f} + s\, TT_t - (TT_{t-1} - 1) \ge 0, \quad \forall t \in \mathcal{T} \tag{A.2}$$

$$0 \le e_{t,g} \quad \perp \quad TT_t + e_{t,g} - (t_* - t) \ge 0, \quad \forall t \in \mathcal{T}, \forall g \in \mathcal{F} \tag{A.3}$$

$$0 \le C_{*,g} \quad \perp \quad \sum_{t \in \mathcal{T}} r_{t,g} - N_g \ge 0, \quad \forall g \in \mathcal{F}. \tag{A.4}$$

Let $\underline{t}_g \equiv \max\left(0, t_* - \dfrac{C_{*,g}}{\beta_g}\right)$, $\tilde{t}_g \equiv \max\left(0, t_* - \dfrac{C_{*,g}}{\alpha_g}\right)$ and $\bar{t}_g \equiv t_* + \dfrac{C_{*,g}}{\gamma_g}$, $\forall g \in \mathcal{F}$. These points are the breakpoints which are used in the following steps.

Step 1. For each $g \in \mathcal{F}$, $r_{t,g} = 0$ for all $t < \underline{t}_g$ or $t > \bar{t}_g$. Moreover, $TT_t = 0$, for $t < \min\limits_{g \in \mathcal{F}} \left(\underline{t}_g\right)$.

Proof. For each class $g \in \mathcal{F}$, assume for the sake of contradiction that $r_{t_o,g} > 0$ for some $t_o < \underline{t}_g$, then it follows by complementarity condition (A.1) that

$$TT_{t_o} = -\frac{\beta_g + \gamma_g}{\alpha_g + \gamma_g} e_{t_o,g} + \frac{C_{*,g} + \gamma_g (t_* - t_o)}{\alpha_g + \gamma_g}. \tag{A.5}$$

The complementarity condition (A.3) becomes

$$0 \leq e_{t_o,g} \perp \frac{\alpha_g - \beta_g}{\alpha_g + \gamma_g} e_{t_o,g} + \frac{C_{*,g} - \alpha_g(t_* - t_o)}{\alpha_g + \gamma_g} \geq 0,$$

which implies $e_{t_o,g} = \dfrac{\alpha_g(t_* - t_o) - C_{*,g}}{\alpha_g - \beta_g}$. Together with (A.5) and $t_o < \underline{t}_g$, it yields

$$TT_{t_o} = \frac{C_{*,g} - \beta_g(t_* - t_o)}{\alpha_g - \beta_g} < 0,$$

which is a contradiction. Thus, for each $g \in \mathcal{F}$, $r_{t,g} = 0$, for $t < \underline{t}_g$.

For any $g \in \mathcal{F}$, if $t > \bar{t}_g$, then by complementarity condition (A.1), it follows $r_{t,g} = 0$.

By the above results and $TT_0 = 0$, it follows immediately that $TT_t = 0$, for $t < \min_{g \in \mathcal{F}} \left(\underline{t}_g\right)$. □

Step 2. $TT_t > 0$, for $\underline{t}_f < t < \bar{t}_f$, for every $f \in \mathcal{F}$.

Proof. Assume for the sake of contradiction that $TT_{t_o} = 0$, for some $\underline{t}_{f_o} < t_o < \bar{t}_{f_o}$, for some $f_o \in \mathcal{F}$, then complementarity condition (A.3) implies $e_{t_o,f_o} = \max(0, t_* - t_o)$. Thus, the right-hand side of complementarity condition (A.1) becomes

$$\begin{aligned} & \frac{\beta_{f_o} + \gamma_{f_o}}{\alpha_{f_o} + \gamma_{f_o}} \max(0, t_* - t_o) - \frac{C_{*,f_o} + \gamma_{f_o}(t_* - t_o)}{\alpha_{f_o} + \gamma_{f_o}} \\ = \quad & \max\left(-\frac{C_{*,f_o} + \gamma_{f_o}(t_* - t_o)}{\alpha_{f_o} + \gamma_{f_o}}, \frac{\beta_{f_o}(t_* - t_o) - C_{*,f_o}}{\alpha_{f_o} + \gamma_{f_o}}\right), \end{aligned}$$

which is a contradiction. □

As a result step 2, for each class $g \in \mathcal{F}$, it follows by complementarity

$$\sum_{f \in \mathcal{F}} r_{t,f} = s(TT_t - TT_{t-1} + 1), \qquad \text{for } \underline{t}_g < t < \bar{t}_g. \tag{A.6}$$

In turn, we can eliminate complementarity condition (A.2) in LCP (A.1) - (A.4) by equality (A.6), then we obtain the following equivalent LCP for t satisfying

$\underline{t}_g < t < \bar{t}_g$:

$$0 \le r_{t,g} \quad \perp \quad \sum_{f \in \mathcal{F}} r_{t,f} \, + \, s\frac{\beta_g + \gamma_g}{\alpha_g + \gamma_g} e_{t,g} - s\frac{C_{*,g} + \gamma_g(t_* - t)}{\alpha_g + \gamma_g} + s\,(TT_{t-1} - 1) \, \ge 0, \quad \forall g \in \mathcal{F}$$

$$0 \le e_{t,g} \quad \perp \quad \sum_{f \in \mathcal{F}} r_{t,f} \, + \, s\, e_{t,g} \, - \, s\,(t_* - t) + s\,(TT_{t-1} - 1) \ge 0, \quad \forall g \in \mathcal{F}. \tag{A.7}$$

LCP (A.7) is used in Step 3 and 4 to obtain the expressions for TT_t.

Step 3. For each class $g \in \mathcal{F}$, if $r_{t,g} > 0$, for $\underline{t}_g < t < \tilde{t}_g$, then $e_{t,g} > 0$ and $TT_t = \dfrac{C_{*,g} - \beta_g(t_* - t)}{\alpha_g - \beta_g}$.

Proof. For each $g \in \mathcal{F}$, assume for the sake of contradiction that $e_{t_o,g} = 0$ given that $r_{t_o,g} > 0$, for some $\underline{t}_g < t_o < \tilde{t}_g$, then by the first complementarity condition in LCP (A.7), it follows that

$$\sum_{f \in \mathcal{F}} r_{t_o,f} = s\frac{C_{*,g} + \gamma(t_* - t_o)}{\alpha_g + \gamma_g} - s(TT_{t_o-1} - 1).$$

Plug $\sum_{f \in \mathcal{F}} r_{t_o,f}$ into the right-hand side of the seconde complementarity condition in LCP (A.7), then it yields that $s\dfrac{C_{*,g} - \alpha_g(t_* - t_o)}{\alpha_g + \gamma_g} < 0$, which is a contradiction. Thus, if $r_{t,g} > 0$, for $\underline{t}_g < t < \tilde{t}_g$, then $e_{t,g} > 0$. Moreover, LCP (A.7) can be reduced to the following complementarity condition for class g:

$$0 \le r_{t,g} \;\perp\; \frac{\alpha_g - \beta_g}{\alpha_g + \gamma_g} \sum_{f \in \mathcal{F}} r_{t,f} + s\frac{\beta_g(t_* - t) - C_{*,g}}{\alpha_g + \gamma_g} + s\frac{\alpha_g - \beta_g}{\alpha_g + \gamma_g}(TT_{t-1} - 1) \ge 0.$$

Since $r_{t,g}$ is assumed to be positive, it follows by complementarity that

$$\sum_{f \in \mathcal{F}} r_{t,f} + s\frac{\beta_g(t_* - t) - C_{*,g}}{\alpha_g - \beta_g} + s(TT_{t-1} - 1) = 0.$$

By (A.6), it yields that

$$TT_t = \frac{1}{s}\sum_{f\in\mathcal{F}} r_{t,f} + (TT_{t-1} - 1) = \frac{C_{*,g} - \beta_g(t_* - t)}{\alpha_g - \beta_g}.$$

□

Step 4. For each class $g \in \mathcal{F}$, $e_{t,g} = 0$ for $t \geq \tilde{t}_g$. Moreover, if $r_{t,g} > 0$, for $\tilde{t}_g \leq t < \bar{t}_g$, then $TT_t = \dfrac{C_{*,g} + \gamma_g(t_* - t)}{\alpha_g + \gamma_g}$.

Proof. For each class $g \in \mathcal{F}$, assume for the sake of contradiction that $e_{t_o,g} > 0$, for some t_o such that $t_o \geq \tilde{t}_g$, then the complementarity condition (A.3), we have

$$TT_{t_o} = -e_{t_o,g} + (t_* - t_o).$$

Plug into the right-hand side of complementarity condition (A.1), we have

$$-\frac{\alpha_g - \beta_g}{\alpha_g + \gamma_g} e_{\bar{t},g} + \frac{\alpha_g(t_* - t_o) - C_{*,g}}{\alpha_g + \gamma_g} < 0,$$

which is a contradiction. Therefore, $e_{t,g} = 0$, for $t \geq \tilde{t}_g$. Moreover, If $r_{t,g} > 0$, for $\tilde{t}_g \leq t < \bar{t}_g$, then by the first complementarity condition in LCP (A.7), we have

$$\sum_{f\in\mathcal{F}} r_{t,f} - s\frac{C_{t,g} + \gamma_g(t_* - t)}{\alpha_g + \gamma_g} + s(TT_{t-1} - 1) = 0.$$

By (A.6), it follows

$$TT_t = \frac{C_{*,g} + \gamma_g(t_* - t)}{\alpha_g + \gamma_g}.$$

□

Step 5. For each class $g \in \mathcal{F}$, if $t = \underline{t}_g$ or $t = \bar{t}_g$, then $r_{t,g} > 0$ implies $TT_t = 0$.

Proof. If $t = \underline{t}_g$ and $r_{t,g} > 0$, then the right-hand side of complementarity condition (A.1) is binding, meaning

$$TT_t = -\frac{\beta_g + \gamma_g}{\alpha_g + \beta_g} e_{t,g} + \frac{C_{*,g} + \gamma_g(t_* - t)}{\alpha_g + \gamma_g}. \quad \text{(A.8)}$$

Plug into complementarity condition (A.3), then we obtain $e_{t,g} = \dfrac{-C_{*,g} + \alpha_g(t_* - t)}{\alpha_g - \beta_g}$, which results in $TT_t = \dfrac{C_{*,g} - \beta_g(t_* - t)}{\alpha_g - \beta_g} = 0$ by (A.8).

If $t = \bar{t}_g$, then complementarity condition (A.1) becomes

$$0 \le r_{t,g} \perp TT_t + \frac{\beta_g + \gamma_g}{\alpha_g + \gamma_g} e_{t,g} \ge 0.$$

If $r_{t,g} > 0$, then $TT_t = 0$. □

Step 6. If class g has an early departure at some time interval $t_o \in \mathcal{T}$, then

$$TT_{t_o} = \max_{f \in \mathcal{F}} \left(\widehat{TT}_{t_o,f} \right) = \widehat{TT}_{t_o,g} = \frac{C_{*,g} - \beta_g(t_* - t_o)}{\alpha_g - \beta_g};$$

if class g has a late departure at some time interval $t_o \in \mathcal{T}$, then

$$TT_{t_o} = \max_{f \in \mathcal{F}} \left(\widehat{TT}_{t_o,f} \right) = \widehat{TT}_{t_o,g} = \frac{C_{*,g} + \gamma_g(t_* - t_o)}{\alpha_g + \gamma_g}.$$

Proof. The users from class g only depart during $\underline{t}_g \le t \le \bar{t}_g$ step 1. If the users from class g departs at a time interval t_o such that $t_o = \underline{t}_g$ or $t_o = \bar{t}_g$, then it follows by step 5 that $TT_{t_o} = \widehat{TT}_{t_o,g} = 0$. By complementarity condition (A.3), it implies $e_{t_o,f} = \max(0, t_* - t_o)$, for any $f \in \mathcal{F}$ and $f \neq g$. Plug into the right-hand side of complementarity condition (A.1), we get

$$\max \left(-\frac{C_{*,f} + \gamma_f(t_* - t_o)}{\alpha_f + \gamma_f}, -\frac{C_{*,f} - \beta_f(t_* - t_o)}{\alpha_f - \beta_f} \right) \ge 0, \quad \forall f \in \mathcal{F}, f \neq g,$$

or equivalently

$$\widehat{TT}_{t_o,f} = \min \left(\frac{C_{*,f} + \gamma_f(t_* - t_o)}{\alpha_f + \gamma_f}, \frac{C_{*,f} - \beta_f(t_* - t_o)}{\alpha_f - \beta_f} \right) \le 0, \quad \forall f \in \mathcal{F}, f \neq g.$$

Thus, it implies $\widehat{TT}_{t_o,g} = \max_{f \in \mathcal{F}} \left(\widehat{TT}_{t_o,f} \right)$.

If $\underline{t}_g < t_o < \tilde{t}_g$, it follows by step 3 that $TT_{t_o} = \widehat{TT}_{t_o,g} = \dfrac{C_{*,g} - \beta_g(t_* - t_o)}{\alpha_g - \beta_g}$.

By (A.6), we have

$$\sum_{f\in\mathcal{F}} r_{t_o,f} = s\frac{C_{*,g}-\beta_g(t_*-t_o)}{\alpha_g-\beta_g} - s(TT_{t_o-1}-1).$$

Thus, by the last complementarity condition in LCP (A.7), it yields

$$e_{t_o,f} = \max\left(0, (t_*-t_o) - \frac{C_{*,g}-\alpha_g(t_*-t_o)}{\alpha_g-\beta_g}\right), \quad \forall f\in\mathcal{F}, f\neq g.$$

Plug into the right-hand side of the first complementarity condition in LCP (A.7), then it follows that

$$s\max\left(\frac{C_{*,g}-\beta_g(t_*-t_o)}{\alpha_g-\beta_g} - \frac{C_{*,f}+\gamma_f(t_*-t_o)}{\alpha_f+\gamma_f},\right.$$
$$\left.\frac{\alpha_f-\beta_f}{\alpha_f+\gamma_f}\left(\frac{C_{*,g}-\beta_g(t_*-t_o)}{\alpha_g-\beta_g} - \frac{C_{*,f}-\beta_f(t_*-t_o)}{\alpha_f-\beta_f}\right)\right) \geq 0, \quad \forall f\in\mathcal{F}, f\neq g.$$

It is equivalent

$$\frac{C_{*,g}-\beta_g(t_*-t_o)}{\alpha_g-\beta_g} \geq \frac{C_{*,f}+\gamma_f(t_*-t_o)}{\alpha_f+\gamma_f}, \quad \forall f\in\mathcal{F}$$

or

$$\frac{C_{*,g}-\beta_g(t_*-t_o)}{\alpha_g-\beta_g} \geq \frac{C_{*,f}-\beta_f(t_*-t_o)}{\alpha_f-\beta_f}, \quad \forall f\in\mathcal{F},$$

meaning that $\frac{C_{*,g}-\beta_g(t_*-t_o)}{\alpha_g-\beta_g} \geq \min\left(\frac{C_{*,f}+\gamma_f(t_{*,f}-\bar{t})}{\alpha_f+\gamma_f}, \frac{C_{*,f}-\beta_f(t_*-t_o)}{\alpha_f-\beta_f}\right), \forall f \in \mathcal{F}$, $f\neq g$. Thus, $\widehat{TT}_{t_o,g} = \max_{f\in\mathcal{F}}\left(\widehat{TT}_{t_o,f}\right)$.

If $\tilde{t}_g \leq t_o < \bar{t}_g$, the proof is similar. □

APPENDIX B
Proof for Lemma 26

In the homogeneous users case, we focus on LCP (4.13). All the subscripts g are dropped. By using the results in Appendix A, we are able to derive the solution expressions to LCP (4.13).

Case 1: $t < \max\left(0, t_* - \dfrac{C_*}{\beta}\right)$. Step 1 in Appendix A implies that $r_t = 0$ and $TT_t = 0$.

Case 2: $t = \max\left(0, t_* - \dfrac{C_*}{\beta}\right)$. Because $TT_{t-1} = 0$ by the previous case, the second condition in LCP (4.13) implies that

$$TT_t = \max(0, r_t - s). \tag{B.1}$$

We claim that $r_t - s \leq 0$. Assume for the sake of contradiction that $r_t - s > 0$, then it follows immediately that $TT_t > 0$ by equation (B.1). On the other hand, since $r_t > s$, by complementarity we have

$$TT_t = -\frac{\beta + \gamma}{\alpha + \gamma} e_t + \frac{C_* + \gamma(t_* - t)}{\alpha + \gamma}. \tag{B.2}$$

Hence the last complementary condition in LCP (4.13) becomes

$$0 \leq e_t \perp \frac{\alpha - \beta}{\alpha + \gamma} e_t + \frac{C_* - \alpha(t_* - t)}{\alpha + \gamma} \geq 0,$$

which follows by the complementarity that $e_t = \dfrac{-C_* + \alpha(t_* - t)}{\alpha - \beta}$ given that $C_* - \alpha(t_* - t) < 0$. Together with equation (B.2), we obtain that $TT_t = \dfrac{C_* - \beta(t_* - t)}{\alpha - \beta} = 0$, which is a contradiction. Hence, $TT_t = 0$ results from $r_t - s \leq 0$. Therefore, $0 \leq r_t \leq s$, $TT_t = 0$ if $t = \max\left(0, t_* - \dfrac{C_*}{\beta}\right)$.

Before discussing case 3, we claim that $r_t > 0$, for $\max\left(0, t_* - \dfrac{C_*}{\beta}\right) < t <$

$t_* + \frac{C_*}{\gamma}$.

Proof. Assume for the sake of contradiction that $r_{\bar{t}} = 0$, for some $\bar{t}$ satisfying $\max\left(0, t^* - \frac{C_*}{\beta}\right) < \bar{t} < \max\left(0, t_* - \frac{C_*}{\alpha}\right)$ and $r_{\bar{t}-1} > 0$. Thus, by the last complementarity condition in LCP (A.7) for homogeneous commuters case, it follows that

$$e_{\bar{t}} = \max(0, t_* - \bar{t} - (TT_{\bar{t}-1} - 1))\,.$$

The right-hand side of the first complementarity condition in LCP (A.7) becomes

$$s \max\left(\frac{-C_* - \gamma(t_* - \bar{t})}{\alpha + \gamma} + (TT_{\bar{t}-1} - 1), \frac{\beta(t_* - \bar{t}) - C_*}{\alpha + \gamma} + \frac{\alpha - \beta}{\alpha + \gamma}(TT_{\bar{t}-1} - 1)\right) \geq 0,$$

which is equivalent to

$$TT_{\bar{t}-1} - 1 \geq \frac{C_* + \gamma(t_* - \bar{t})}{\alpha + \gamma} \tag{B.3}$$

or

$$TT_{\bar{t}-1} - 1 \geq \frac{C_* - \beta(t_* - \bar{t})}{\alpha - \beta}. \tag{B.4}$$

Since $r_{\bar{t}-1} > 0$, it implies that $\max\left(0, t_* - \frac{C_*}{\beta}\right) \leq \bar{t} - 1 < \max\left(0, t_* - \frac{C_*}{\alpha}\right)$ by case 1 and case 2. Moreover, it follows by step 6 in Appendix A and case 2 that $TT_{\bar{t}-1} = \frac{C_* - \beta(t_* - \bar{t} + 1)}{\alpha - \beta}$. Plug into inequality (B.3), then it yields

$$\bar{t} - 1 \geq t_* - \frac{C_*}{\alpha} + \frac{\alpha - \beta}{\beta + \gamma},$$

which contradicts to $\bar{t} - 1 < \max\left(0, t_* - \frac{C_*}{\alpha}\right)$. Moreover, Inequality B.4 is also violated by $TT_{\bar{t}-1} = \frac{C_* - \beta(t_* - \bar{t} + 1)}{\alpha - \beta}$. Therefore, $r_t > 0$, for all $\max\left(0, t_* - \frac{C_*}{\beta}\right) < t < \max\left(0, t_* - \frac{C_*}{\alpha}\right)$. □

Case 3: $\max\left(0, t_* - \frac{C_*}{\beta}\right) < t < \max\left(0, t_* - \frac{C_*}{\alpha}\right)$. Equality (A.6) in Appendix A implies that

$$r_{\bar{t}} = s\frac{C_* + \gamma(t_* - \bar{t})}{\alpha + \gamma} - s(TT_{\bar{t}-1} - 1).$$

By step 3 in Appendix A and the above claim, we have $TT_t = \dfrac{C_* - \beta(t_* - t)}{\alpha - \beta}$.

Case 4: $\max\left(0, t_* - \dfrac{C_*}{\alpha}\right) \le t < t_* + \dfrac{C_*}{\gamma}$. Equality (A.6) in Appendix A

$$r_{\bar{t}} = s\frac{C_* + \gamma(t_* - \bar{t})}{\alpha + \gamma} - s(TT_{\bar{t}-1} - 1).$$

By step 4 in Appendix A and the above claim, we have $TT_t = \dfrac{C_* + \gamma(t_* - t)}{\alpha + \gamma}$.

Case 5: $t = t_* + \dfrac{C_*}{\gamma}$. Because $t_* - t = -\dfrac{C_*}{\gamma} < 0$, it follows by the last complementarity condition that $e_t = 0$. Thus, LCP (4.13) is equivalent to

$$\begin{aligned} 0 \le r_t \quad &\perp \quad TT_t(C_*) \ge 0 \\ 0 \le TT_t \quad &\perp \quad -r_t + s(TT_t - TT_{t-1}) + s \ge 0 \end{aligned}$$

If $1 - TT_{t-1} \ge 0$, then it follows by complementarity that

$$TT_t = 0 \text{ and } 0 \le r_t \le s(1 - TT_{t-1}).$$

If $1 - TT_{t-1} < 0$, then it follows by complementarity that

$$r_t = 0 \text{ and } TT_t = TT_{t-1} - 1.$$

Therefore, we have $TT_t = \max(TT_{t-1} - 1, 0)$ and $0 \le r_t \le s(1 - TT_{t-1})$.

Case 6: $t > t_* + \dfrac{C_*}{\gamma}$. Step 1 and step 4 in Appendix A implies that $r_t = 0$ and $e_t = 0$. Hence, the second complementarity condition in LCP (4.13) implies that $TT_t = \max(TT_{t-1} - 1, 0)$.

Table 1. Summary of Model Results

Capacity Market (MW)	Min Output Level	Contingent Allocation System	Run	CO2 Limit(Mton/yr)	% Grandfathered	% Increase Relative to Base Run			Capacity(MW)			% Capacity Factor			p_e ($/ton)	p_c ($/ton)
						Generation Cost	Social Cost	Consumer Payments	Coal	Comb. Cycle	Comb. Turbine	Coal	Comb. Cycle	Comb. Turbine		
11000	None	Potential Emission Rule	1	20	0%	28.7%	31.6%	28.7%	3088	5540	5361	27.6%	73.2%	0.3%	34.35	0
			2		50%	16.9%	19.4%	31.3%	863	7747	2390	97.8%	52.6%	0.9%	29.54	14617
			3		100%	17.6%	19.3%	39.5%	852	8084	2064	97.8%	51.1%	0.3%	22.45	50000
			4	40	0%	26.5%	29.2%	26.5%	8094	559	12155	53.9%	99.2%	0.1%	32.46	0
			5		50%	-0.2%	2.5%	29.8%	6803	1729	2468	64.0%	32.0%	1.0%	30.77	8462
			6		100%	0.1%	0.9%	21.6%	6076	2871	2053	71.1%	23.9%	0.5%	11.01	50000
		Actual Emission Rule	7	20	0%	23.8%	25.0%	23.8%	2459	8541	0	33.7%	48.6%	N/A	30.77	29115
			8		50%	20.1%	21.2%	35.1%	1600	7889	1511	51.8%	52.7%	0.0%	30.77	50000
			9		100%	Same results as Run 3										
			10	40	0%	10.6%	11.8%	10.6%	10140	860	0	42.8%	74.6%	N/A	30.77	4593
			11		50%	5.0%	6.2%	35.0%	8359	1211	1430	51.9%	53.0%	0.0%	30.77	50000
			12		100%	Same results as Run 6										
		Base Run		No Limit		2049	20911	2049	7329	1628	2042	65.10%	17.65%	0.59%	0.00	50000
None	None	Potential Emission Rule	13	20	0%	Same as results as Run 1										N/A
			14		50%	22.8%	23.8%	39.1%	1976	6557	290	43.2%	61.9%	5.6%	30.77	
			15		100%	18.7%	20.6%	42.4%	886	7601	0	97.7%	53.3%	N/A	22.45	
			16	40	0%	Same as results as Run 4										
			17		50%	5.1%	5.8%	38.2%	8038	553	301	54.3%	99.3%	5.0%	31.43	
			18		100%	-0.1%	0.9%	23.2%	6160	2365	0	70.7%	25.7%	N/A	11.01	
		Actual Emission Rule	19	20	0%	25.3%	24.8%	25.3%	2226	6851	0	37.8%	60.2%	N/A	30.77	N/A
			20		50%	21.6%	22.5%	37.9%	1606	7109	0	53.4%	57.3%	N/A	30.77	
			21		100%	Same results as Run 15										
			22	40	0%	17.1%	15.9%	17.1%	9752	807	0	44.5%	79.5%	N/A	30.77	
			23		50%	4.2%	4.9%	36.7%	7859	889	0	55.6%	63.4%	47.3%	30.77	
			24		100%	Same results as Run 18										
		Base Run		No Limit		1893	21020	1893	7329	1232	0	65.10%	19.79%	N/A	0.00	N/A
11000	Coal Only	Potential Emission Rule	25	20	0%	13.4%	13.5%	13.4%	1441	7018	5811	58.5%	57.8%	0.6%	30.77	0
			26		50%	5.1%	5.0%	18.1%	863	7747	2390	97.8%	52.6%	0.9%	29.54	14617
			27		100%	5.8%	4.9%	25.5%	852	8084	2064	97.8%	51.1%	0.3%	22.45	50000
		Base Run		No Limit		2278	20628	2278	2287	6671	2042	0.9%	0.4%	0.0%	0.00	50000