INTEGER PROGRAMMING APPROACHES TO STOCHASTIC GAMES ARISING IN PAIRED KIDNEY EXCHANGE AND INDUSTRIAL ORGANIZATION

by

Amin Dehghanian

B.S., Sharif University of Technology, 2005

M.S., Sharif University of Technology, 2008

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Amin Dehghanian

It was defended on

July 6, 2015

and approved by

Andrew J. Schaefer, Ph.D., John A. Swanson Chair, Professor, Department of Industrial Engineering

Jeffrey P. Kharoufeh, Ph.D., Associate Professor, Department of Industrial Engineering

Jayant Rajgopal, Ph.D., Professor, Department of Industrial Engineering

Oleg A. Prokopyev, Ph.D., Associate Professor, Department of Industrial Engineering

Mustafa Akan, Ph.D., Associate Professor, Tepper School of Business, Carnegie Mellon University

Dissertation Director: Andrew J. Schaefer, Ph.D., John A. Swanson Chair, Professor, Department of Industrial Engineering

INTEGER PROGRAMMING APPROACHES TO STOCHASTIC GAMES ARISING IN PAIRED KIDNEY EXCHANGE AND INDUSTRIAL ORGANIZATION

Amin Dehghanian, PhD

University of Pittsburgh, 2015

We investigate three different problems in this dissertation. The first two problems are related to games arising in paired kidney exchange, and the third is rooted in a computational branch of the industrial organization literature. We provide more details on these problems in the following.

End-stage renal disease (ESRD), the final stage of chronic kidney disease, is the ninth-leading cause of death in the United States, where it afflicts more than a half million patients, and costs more than forty billion dollars indirect expenses annually. Transplantation is the preferred treatment for ESRD; unfortunately, there is a severe shortage of transplantable kidneys. Kidney exchange is a growing approach to alleviate the shortage of kidneys for transplantation, and the United States is considering creating a national kidney exchange program since such a program provides more and better transplants. A major challenge to establish a national kidney exchange program is the lack of incentives for transplant centers to participate in such a program. To overcome this issue, the kidney transplant community has recently proposed a payment strategy framework that incentivizes transplant centers to participate in a national program. Absent from this debate is a careful investigation of how to design these incentives. We develop a principal-agent model to analyze these incentives and find an equilibrium payment strategy. We develop a mixed-integer bilinear bilevel program to compute an equilibrium payment strategy. We show that this bilevel program can be solved as a mixed-integer linear program. We calibrate our model and provide several data-driven insights about advantages of a national kidney exchange program. We shed light on several controversial policy questions about an equilibrium payment strategy. In particular, we demonstrate that there exists a “win-win” payment strategy that could result in saving thousands of lives and billions of dollars annually.

Consensus stopping games are a class of stochastic games that arises in the context of kidney exchange. Specifically, the problem of finding a socially optimal pure stationary equilibrium of a consensus stopping game is adapted to value a given kidney exchange. However, computational difficulties have limited its applicability. We show that a consensus stopping game may have many pure stationary equilibria, which in turn raises the question of equilibrium selection. Given an objective criterion, we study the problem of finding a best pure stationary equilibrium for the game, which we show to be NP-hard. We characterize the pure stationary equilibria, show that they form an independence system, and develop several families of valid inequalities. We then solve the equilibrium selection problem as a mixed-integer linear program (MILP) by a branch-and-cut approach. Our computational results demonstrate the advantages of our approach over a commercial solver.

Industrial organization is an area of economics that studies firms and markets. Currently, a class of stochastic games are adopted to model behaviors of firms in a market. However, inherent challenges in computability of stationary equilibria have restricted its applicability. To overcome this challenge, we develop several characterizations of stationary equilibria for the class of stochastic games.

Keywords: Kidney exchange, standard acquisition charge, pricing, bilevel programming, principal-agent models, stochastic stopping games, equilibrium selection, consensus decision-making, veto players, independence system, branch-and-cut, industrial organization.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

Dedicated to my father and my mother

I would like to express my deepest gratitude to my advisor, Prof. Andrew J. Schaefer, for his guidance, encouragement, professional attitude, patience, support and constructive criticism. It has been a privilege and pleasure to learn many lessons from him.

My special thanks to Prof. Jeffrey P. Kharoufeh for his collaboration and support.

Many thanks to my committee members- Prof. Jayant Rajgopal, Prof. Oleg A. Prokopyev, and Prof. Mustafa Akan- for their feedback.

My warm thanks to Prof. Bopaya Bidanda for always being there with his support.

This work would have never been possible without my wonderful family: My parents, my sisters, my brothers-in-law, and my nephew. Their dedication, endless love, encouragement, and unconditional support have always been an inspiration and strength to me.

1.0 INTRODUCTION

1.1 SUMMARY OF THE THESIS

This thesis primarily investigates topics arising in paired kidney exchange. In the first topic, we develop an equilibrium payment strategy that incentivizes transplant centers to participate in a national kidney exchange program. In the second topic, we develop a novel approach to solve the problem of finding a socially optimal pure stationary equilibrium of a class of stochastic games arising in contexts such as kidney exchange. In the third topic, we provide several characterizations of stationary equilibria of an important class of stochastic games from the industrial organization literature. Such characterizations facilitate computing stationary equilibria of this class of stochastic games. We provide a brief review of contents of this thesis, in the following.

Chapter 2: An Optimal Incentive Alignment for a National Kidney Exchange Program. Living-donor transplantation is the preferred treatment for chronic kidney disease, the ninth-leading cause of death in the United States. A significant drawback of living-donors is that at least one-third of the patients with a willing living-donor are unable to receive the donor's kidney due to biological incompatibilities. To mitigate this barrier, an emerging clinical practice is paired kidney exchange (PKE), in which incompatible patient-donor pairs exchange donors' kidneys with other pairs in a biologically compatible manner. PKE is currently utilized in a decentralized setting in the United States, which leads to inefficiencies. In fact, each transplant center (a consortium of a few transplant centers) manages a pool of its own incompatible patient-donor pairs, and conducts its exchanges internally. To achieve the benefits of resource pooling and enhance PKE efficiency, the United States is moving toward a national PKE program in which all patient-donor pairs are enrolled and arranged for PKE. However, each transplant center will participate in a national PKE program only if such an action is in its own best interest. Hence, there is a need for a payment strategy in which each transplant center receives a monetary reward in return for enrolling each patient-donor. The clinical community believes that a successful national payment strategy can lead to saving thousands of lives and millions of dollars annually. We present a principal-agent model to capture the interaction between a national PKE program and transplant centers through a payment strategy. We next develop a bilevel program to find an equilibrium payment strategy of the interaction between the national PKE program and the transplant centers. We calibrate our model using a data set provided by the United Network for Organ Sharing, which leads to data-driven insights about advantages of a national PKE program. We shed light on several controversial policy questions about an equilibrium payment strategy. In particular, we demonstrate that there exists a "win-win" equilibrium payment strategy that leads to saving thousands of lives and billions of dollars annually.

Chapter 3: Optimizing over Pure Stationary Equilibria in Consensus Stopping Games. Developing an accurate method to value a given PKE highly improves PKE efficacy. A recent approach to value a given PKE is based on finding a socially optimal pure stationary equilibrium for consensus stopping games, a class of stochastic games that require the consent of all players to terminate and arise in many practical decision-making environments with veto players. However, technical difficulties have restricted the applicability of this approach. Motivated by this challenge, we consider the problem of finding a socially optimal pure stationary equilibrium for consensus stopping games. We represent the problem as a mixed-integer linear program (MILP), and establish its computational complexity. We characterize several combinatorial structures of the equilibria, and subsequently develop various families of valid inequalities that are used to efficiently solve the MILP by a branch-and-cut approach. Our extensive computational experiments on a set of real-world instances demonstrate that our approach can solve some instances in minutes whereas CPLEX cannot solve even their linear programming relaxations within several hours.

We are the first to provide combinatorial characterizations of stationary equilibria for a class of stochastic games. This dissertation is the first attempt to develop a novel cutting plane approach for the problem of finding a socially optimal (pure) stationary equilibrium.

Chapter 4: Characterizing Entry and Exit for Stationary Equilibria of a Dynamic Oligopoly Model. We consider a dynamic industry, from industrial organization literature, composed of two classes of firms: (i) incumbent firms, and (ii) potential entrants. In each period, each incumbent needs to decide whether to *remain* in or *exit* from the industry. If it remains, it next decides how much to invest in this period. Otherwise, it exits, and receives a certain salvage value. In each period, each potential entrant needs to decide whether to *enter* or *stay out*. If it enters, it incurs a certain setup cost, and next decides how much to invest. If it stays out, it permanently disappears from the industry. The current status of the industry is represented by a state variable, which evolves according to a Markovian transition conditioned on the strategies of all firms. The described industry is at the heart of a growing literature on industrial organization. We characterize entry and exit behaviors of the firms in stationary equilibria of the industry. Such characterizations can facilitate the process of computing stationary equilibria.

2.0 AN OPTIMAL INCENTIVE ALIGNMENT FOR A NATIONAL KIDNEY EXCHANGE PROGRAM

2.1 INTRODUCTION

End-stage renal disease (ESRD) leads to permanent failure of both kidneys. As of December 31, 2012, at least 636,905 Americans were suffering from ESRD, and its prevalence rate is still growing. ESRD expenditures exceeded $40 billion in 2012, and Medicare's portion was $28.6 billion (United States Renal Data System 2015). Two available treatment modalities for ESRD are dialysis and transplantation. Dialysis is a temporary, expensive, and dangerous modality, and diminishes patients' quality of life. Kidney transplantation leads to better patient outcomes and is less expensive than dialysis (Laupacis et al. 1996); hence, it is the preferred choice of treatment for ESRD. Based on viability status, kidney donors are classified into two groups: deceased-donors and living-donors. Living-donor transplants generally provide better long-term survivals than those from deceased-donors. In the sequel, we restrict our attention to living-donor transplants, and call a living-donor *her* and an ESRD patient *him*.

A difficulty to greater use of living-donors is that at least one-third of ESRD patients with willing living-donors are physiologically incompatible with their intended donors due to blood-type and/or tissue-type incompatibilities (Montgomery et al. 2005a). To alleviate this difficulty, an emerging clinical practice is paired kidney exchange (PKE), in which incompatible patient-donor pairs swap their donors in a cyclical and physiologically compatible manner (see Figure 1). Note that the size of a PKE cycle can conceptually be any integer greater than or equal to two; however, logistical constraints restrict it to no more than three in effect. Currently, each patient-donor pair visits a transplant center to find out whether they are compatible. If they are compatible, a transplant surgery is conducted in which the patient receives a kidney from his intended donor. Otherwise, if they are incompatible, the transplant center seeks to arrange PKE among its incompatible patient-donor pairs. For convenience, we refer to an incompatible patient-donor pair as a patient-donor (pair) throughout this chapter, unless otherwise stated.

To achieve the benefits of resource pooling and enhance PKE efficiency, the United States is moving toward the creation of a national PKE program in which all patient-donor pairs are registered, and PKE can be arranged through this program. Indeed, as the number of patient-donor pairs registered in a PKE program increases, both the number of transplants and their quality increase (Roth et al. 2004, Segev et al. 2005). In particular, a national PKE program could save at least 1000 additional lives and lead to an annual saving of \$200-\$500 million for the United States' healthcare system (Segev et al. 2005, Rees et al. 2012).

A major challenge that has hindered the establishment of a national PKE program is a lack of incentives for transplant centers to participate. One possible approach is to ask patient-donor pairs to directly register in the national PKE program and omit the transplant center as an intermediary in the registration process. Another is to require that the transplant centers enroll all their patient-donor pairs in the national PKE program, by passing a law. Ashlagi and Roth (2014) discussed these suggestions and provided reasons as to why they are unrealistic. In Section 2.6, our numerical results demonstrate that it is unnecessary to enforce transplant centers to participate in the national PKE program. In fact, we show that transplant centers should be willing to pay for enrolling most of their patient-donor pairs in the national PKE program. In the following, we clarify why transplant centers may lack incentives to participate in a national PKE program.

In a PKE program, patient-donor pairs undergo the following stages: (1) Patient-donor pairs are evaluated to acquire the necessary medical information including compatibility information needed for the next stage. (2) A matching algorithm finds an assignment of donors to recipients subject to compatibility of donors with recipients (cf. Awasthi and Sandholm 2009, Ünver 2010, Glorie et al. 2014). (3) Recipients and donors undergo final tests and subsequently transplant surgeries. From a financial standpoint, the insurance company of a patient is willing to pay for his medical expenses and his donor's medical expenses,

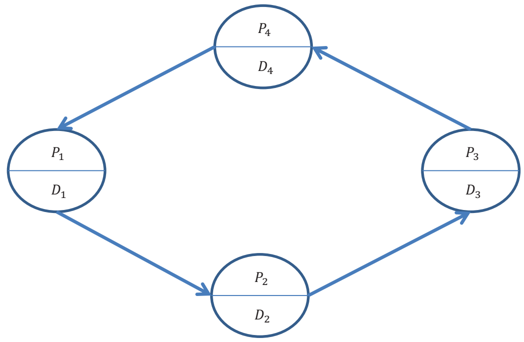

Figure 1: A four-way PKE, where each node represents a patient-donor pair. The donor in each node donates her kidney to the patient in the next node of the cycle.

and the insurance company of a donor does not pay for her medical expenses. Donor's pre-match expenses are incurred before the identity of her recipient is known, so it is unclear who pays for those expenses. This is a pressing question that needs to be addressed before initiating a national PKE program. Deceased donor kidney transplant was suffering from the same challenge in 1970s until the Centers for Medicare and Medicaid Services developed a reimbursement strategy using the concept of a standard acquisition charge (SAC) to pay organ procurement organizations (OPOs) for the overhead costs of evaluating and recovering a deceased donor organ. Analogously, the development of a SAC model for a national PKE program has been recommended so that the issue of pre-match expenses is resolved (Rees et al. 2012, Melcher et al. 2013). This means that a national PKE program needs a payment strategy for the pre-match expenses that is *individually rational* for each transplant center, i.e., each transplant center is better off to harvest patient-donor pairs rather than do nothing at all.

As each transplant center participating in a national PKE program seeks to maximize its benefit, it may decide to only register its hard-to-match pairs in the pool and internally conduct exchanges on its easy-to-match pairs. Such an approach can save the transplant center some overhead cost related to logistics, coordination, and bureaucracy (cf. Ashlagi and Roth 2012). Hence, transplant centers may only send their hard-to-match patient-donor pairs to the national PKE program. Ashlagi and Roth (2014) discussed how transplant centers may interact with a centralized PKE program and how free riding by transplant centers has already been observed as they withhold some of their patient-donor pairs. This implies that the aforementioned payment strategy for a national PKE program needs to be *incentive compatible* for each transplant center, i.e., each transplant center seeks to maximize its benefit, given the payment strategy of the national PKE program.

In this chapter, we develop a payment strategy for a national PKE program that is both individually rational and incentive compatible for transplant centers, and maximizes a social welfare objective.

2.2 LITERATURE REVIEW

Incentives of transplant centers to participate in a national PKE program have been investigated in several papers. Glorie et al. (2014) developed an algorithm to match patient-donor pairs with multiple criteria. Their approach included individual rationality of transplant centers by ensuring that each participating transplant center receives at least as many transplants as those without participation. Ashlagi et al. (2013) studied mechanisms for two-way exchanges under which full participation is a dominant strategy for transplant centers. They also established lower bounds on social welfare loss of pursuing such mechanisms. Ashlagi and Roth (2014) showed that as the number of patient-donor pairs and transplant centers grows, it becomes less costly for society to incorporate individual rationality of transplant centers. In addition, they proposed a bonus mechanism to incentivize transplant centers to enroll all their patient-donor pairs in a large PKE program. Toulis and Parkes (2015) designed a mechanism under which the truthful reporting of patient-donor pairs is efficient and incentive compatible for each transplant center. Their analysis relies on the assumption that the size of each transplant center's pool is at least moderate (greater than 30). Finally, they compared the performance of their mechanism with that of Ashlagi and Roth (2014) by simulation.

What distinguishes our work from the literature is that we are the first to investigate the problem of finding an equilibrium payment strategy that incentivizes transplant centers to participate in a national PKE program. Moreover, we present monetary definitions of individual rationality and incentive compatibility since defining them only in terms of the number of transplants is inaccurate.

2.3 A PRINCIPAL-AGENT MODEL

In this section, we formally define our problem and develop a mathematical model to explore a class of payment strategies for the national PKE program, in which each transplant center receives a monetary reward in return for harvesting, preparing, and enrolling each patient-donor pair in the national pool. Consistent with current practice for deceased donors and Rees et al. (2012), we call such a reward a SAC. We use the term *society* to refer to the national PKE program manager throughout the chapter. We approach the problem as a principal-agent model in which society is the principal and transplant centers are agents.

The parameters of the model are as follows:

- $\mathscr{I} = \{1, 2, \ldots, I\}$: The set of patient-donor and altruistic donor classes.
- $\mathscr{E} := \{(i, j) \in \mathscr{I}^2 |$ donors of class i are compatible with patients of class $j\}$.
- For all $i \in \mathscr{I}$, $\mathscr{E}_{i.} = \{j \in \mathscr{I} | (i, j) \in \mathscr{E}\}$, i.e., set of patient-donor pair classes whose patients are compatible with donors of class i.
- For all $i \in \mathscr{I}$, $\mathscr{E}_{.i} = \{j \in \mathscr{I} | (j, i) \in \mathscr{E}\}$, i.e., set of patient-donor pair classes whose donors are compatible with patients of class i.
- $\mathscr{C} = \{1, 2, \ldots, C\}$: Set of transplant centers.
- λ_{ic}: Arrival rate of patient-donor pairs of class $i \in \mathscr{I}$ to transplant center $c \in \mathscr{C}$.
- α_i: Rate of patient-donor pairs of class $i \in \mathscr{I}$ which are already available in the national pool. These pairs reflect the unmatched pairs left from earlier periods.

  Note that we develop a static model for the problem of assigning the SAC compensations, $\langle s_{ic} \rangle_{i \in \mathscr{I}, c \in \mathscr{C}}$, since when the SAC compensations are assigned in a practical setting, they will remain unchanged for a long time. Hence, it is plausible to investigate patient-donor pairs' arrival process only through its expected value (assuming a stationary distribution). For this reason, we incorporated the rates λ_{ic} and α_i for each $i \in \mathscr{I}$ and $c \in \mathscr{C}$.

- p_{ic}: The probability that a pair of class $i \in \mathscr{I}$ belonging to transplant center $c \in \mathscr{C}$ will be matched if it is enrolled in the national pool. We assume that this parameter is exogenous. This assumption is plausible since the national pool is much larger than that of a transplant center, and in principle it is not influenced by decisions of a transplant center.
- r_{ic}: Monetary reward that transplant center $c \in \mathscr{C}$ receives in return for conducting a transplant operation on the patient of a pair of class $i \in \mathscr{I}$ in its facilities. The amount of money, which is reimbursed by the insurance company of a matched patient, may be embedded in this reward. When we define our objective function subsequently, regardless of whether a patient is internally matched by a transplant center or is externally matched by the national program, we consider that the patient of each pair harvested by a transplant center has his surgery in the same center, and the donor always travels to her intended recipient's transplant center.
- d_{ic}: Monetary cost of harvesting a patient-donor pair and donor's pre-match cost.
- R_i: Societal reward for matching a patient of class $i \in \mathscr{I}$. Similar to r_{ic}, the amount of money, which is reimbursed by the insurance company of a matched patient, may be embedded into this reward.

Societal decision variables are as follows:

- y_{ij}: The number of donors of class i matched to patients of class j in the national pool for all $(i, j) \in \mathscr{E}$.
- s_{ic}: The SAC compensation for each pair of class $i \in \mathscr{I}$ enrolled in the national pool by transplant center $c \in \mathscr{C}$. This can be interpreted as the compensation that a transplant center receives in return for the service that the center provides for the national PKE program. This service includes harvesting, preparing, and enrolling a patient-donor of class $i \in \mathscr{I}$ in the national program. Another approach is to assume that the SAC com-

pensation depends both on the class of the enrolled patient-donor pair and the transplant center harvesting the patient-donor pair. We investigate this question in Section 3.8.

Decision variables for each transplant center $c \in \mathscr{C}$ are as follows:

- x_{ijc}: The number of donors of class i matched to patients of class j in the pool of transplant center $c \in \mathscr{C}$ for all $(i,j) \in \mathscr{E}$.
- z_{ic}: Rate of patient-donors of class $i \in \mathscr{I}$ from transplant center $c \in \mathscr{C}$ enrolled in the national pool.

Without loss of generality, we assume that the rewards r_{ic} and R_i are nonnegative for each $i \in \mathscr{I}$ and $c \in \mathscr{C}$, otherwise all associated decision variables will take the optimal value 0 in the subsequent optimization models and may be omitted. We also follow the convention that a term in bold refers to a real-valued vector, e.g., $\mathbf{x}_c := \langle x_{ijc} \rangle_{(i,j)\in\mathscr{E}}$, $\mathbf{y} := \langle y_{ij} \rangle_{(i,j)\in\mathscr{E}}$, $\mathbf{z}_c := \langle z_{ic} \rangle_{i\in\mathscr{I}}$, and $\mathbf{s} := \langle s_{ic} \rangle_{i\in\mathscr{I},c\in\mathscr{C}}$. For each $\mathbf{s} \in \mathbb{R}^{|\mathscr{I}||\mathscr{C}|}$, transplant center $c \in \mathscr{C}$ seeks to maximize its profit by the following problem:

$$\varphi_c(\mathbf{s}) := \max \sum_{(i,j)\in\mathscr{E}} r_{ic} x_{ijc} + \sum_{i\in\mathscr{I}} (p_{ic} r_{ic} + s_{ic}) z_{ic} \tag{2.1a}$$

$$s.t. \quad z_{ic} + \sum_{j\in\mathscr{E}_{i.}} x_{ijc} \le \lambda_{ic} \qquad \forall i \in \mathscr{I}, \tag{2.1b}$$

$$\sum_{j\in\mathscr{E}_{i.}} x_{ijc} = \sum_{j\in\mathscr{E}_{.i}} x_{jic} \qquad \forall i \in \mathscr{I}, \tag{2.1c}$$

$$z_{ic} \ge 0 \qquad \forall i \in \mathscr{I}, \tag{2.1d}$$

$$x_{ijc} \in \mathbb{Z}_+ \qquad \forall (i,j) \in \mathscr{E}. \tag{2.1e}$$

For each $c \in \mathscr{C}$, let $\Psi_c(\mathbf{s}) := \{\langle \mathbf{x}_c, \mathbf{z}_c \rangle \in (2.1b) - (2.1e) : \sum_{i,j\in\mathscr{E}} r_{ic} x_{ijc} + \sum_{i\in\mathscr{I}} (p_{ic} r_{ic} + s_{ic}) z_{ic} \ge \varphi_c(\mathbf{s})\}$, denote the set of optimal solutions for (2.1a) – (2.1e), called a *solution set mapping* in the literature (cf. Robinson 1979, Dempe 2002). Society seeks to maximize social welfare:

$$\max \sum_{(i,j)\in\mathscr{E}} R_i \left(y_{ij} + \sum_{c\in\mathscr{C}} x_{ijc} \right) - \sum_{i\in\mathscr{I},c\in\mathscr{C}} s_{ic} z_{ic} \tag{2.2a}$$

$$s.t. \quad \sum_{j\in\mathscr{E}_{i.}} y_{ij} \leq \alpha_i + \sum_{c\in\mathscr{C}} z_{ic} \qquad \forall i \in \mathscr{I}, \tag{2.2b}$$

$$\sum_{j\in\mathscr{E}_{i.}} y_{ij} = \sum_{j\in\mathscr{E}_{.i}} y_{ji} \qquad \forall i \in \mathscr{I}, \tag{2.2c}$$

$$\sum_{(i,j)\in\mathscr{E}} r_{ic} x_{ijc} + \sum_{i\in\mathscr{I}} (p_{ic} r_{ic} + s_{ic}) z_{ic} - \sum_{i\in\mathscr{I}} \lambda_{ic} d_{ic} \geq 0 \qquad \forall c \in \mathscr{C}, \tag{2.2d}$$

$$\langle \mathbf{x}_c, \mathbf{z}_c \rangle \in \Psi_c(\mathbf{s}) \qquad \forall c \in \mathscr{C}, \tag{2.2e}$$

$$y_{ij} \in \mathbb{Z}_+ \qquad \forall (i,j) \in \mathscr{E}, \tag{2.2f}$$

$$s_{ic} \text{ unrestricted} \qquad \forall i \in \mathscr{I}. \tag{2.2g}$$

The objective function (2.2a) represents the total reward of society due to the number of transplants, subtracted from the amount of money that society spends on the participation of the transplant centers. This objective function reflects the fact that it does not matter whether a pair is matched internally by a transplant center or by the national PKE program, provided that the sizes of both matching are the same. Conditions (2.2b) are *flow conservation* constraints, enforcing that the rate of matched pairs of class $i \in \mathscr{I}$ does not exceed the arrival rate of pairs of the same class to the national PKE program. Conditions (2.2c) are *participation* constraints for pairs, ensuring that each pair of class i donates a kidney only in return for receiving a kidney. Conditions (2.2d) are *individual rationality* constraints for transplant centers, ensuring that the expected benefit of participating and accepting all patient-donor pairs for each transplant center is at least as large as the alternative of rejecting all patient-donor pairs upon admission. Conditions (2.2e) are *incentive compatibility* constraints for transplant centers since they require that given a SAC profile $\mathbf{s}$, each transplant center seeks to maximize its own profit. Indeed, the societal model (2.2a) – (2.2g) is a bilevel program where the transplant center model (2.1a) – (2.1e) for each $c \in \mathscr{C}$ is the lower level problem and reflected by constraints (2.2e). Note that there are multiple ($|\mathscr{C}|$) lower level problems. Whenever there exist multiple optimal solutions for the lower level problem (2.2e) for each $c \in \mathscr{C}$, the one generating the highest profit for society is selected.

In other words, we adopt the notion of optimistic formulation in bilevel programming. For a comprehensive review about different formulations and relations between them in bilevel programming, see Dempe (2002).

Remark 1. *We may truncate our search space to a closed bounded interval in* $\mathbb{R}$ *for variable* s_{ic} *for each* $i \in \mathscr{I}$ *and* $c \in \mathscr{C}$ *in problem* (2.2a) − (2.2g) *because if* $s_{ic} > \sum_{j \in \mathscr{I}} r_{jc}\lambda_{jc}$, *then the optimal solution of problem* (2.1a) − (2.1e) *is* $z^*_{ic} = \lambda_{ic}$, *i.e., the optimal reaction of transplant centers is to enroll all their patient-donor pairs of class* i *in the national PKE program. On the other hand, if* $s_{ic} < -p_{ic}r_{ic}$, *then the optimal solution of problem* (2.1a) − (2.1e) *is* $z^*_{ic} = 0$, *i.e., the optimal reaction of transplant centers is to enroll none of their patient-donor pairs of class* i *in the national PKE program. Let* $\overline{s}_{ic} := \sum_{j \in \mathscr{I}} r_{jc}\lambda_{jc}$ *and* $\underline{s}_{ic} := -p_{ic}r_{ic}$ *be upper and lower bounds of* s_{ic}, *respectively.*

2.4 A SINGLE-LEVEL FORMULATION

A bilevel MILP where the lower level is an LP can be transformed into a single-level MILP that can be solved by state-of-the-art MILP solvers (Fortuny-Amat and McCarl 1981). Conversely, a bilevel MILP where the lower level is an MILP with 40 binary variables and a single budget constraint is unsolvable by state-of-the-art approaches (DeNegre and Ralphs 2009, Tang et al. 2015). Note that the bilevel program (2.2a) − (2.2f) has 300 lower level problems, each of which has at least 32 constraints and 256 integer variables. Our study in this section reveals a rich network structure of the lower level problem (2.1a) − (2.1e) that makes the bilevel program (2.2a) − (2.2f) tractable.

Lemmas 1 and 2 present a reformulation of the lower level problem (2.1a) − (2.1e), which is used to reformulate the bilevel program (2.2a) − (2.2f) into an equivalent MILP. For $a \in \mathbb{R}$, let $a^+ := \max\{a, 0\}$, $\lfloor a \rfloor$ be the floor of a, and $\{a\} := a - \lfloor a \rfloor$.

Lemma 1. *For each $c \in \mathscr{C}$, $(\mathbf{x}_c^*, \mathbf{z}_c^*)$ is an optimal solution of* (2.1a) – (2.1e) *if and only if $\mathbf{x}_c^*$ is an optimal solution of the following IP,*

$$\max \sum_{(i,j)\in\mathscr{E}} (r_{ic} - (p_{ic}r_{ic} + s_{ic})^+)x_{ijc} \tag{2.3a}$$

$$s.t. \quad \sum_{j\in\mathscr{E}_{i.}} x_{ijc} \le \lfloor \lambda_{ic} \rfloor \qquad \forall i \in \mathscr{I}, \tag{2.3b}$$

$$\sum_{j\in\mathscr{E}_{i.}} x_{ijc} = \sum_{j\in\mathscr{E}_{.i}} x_{jic} \qquad \forall i \in \mathscr{I}, \tag{2.3c}$$

$$x_{ijc} \in \mathbb{Z}_+ \qquad \forall (i,j) \in \mathscr{E}, \tag{2.3d}$$

and

$$z_{ic}^* = \lambda_{ic} - \sum_{j\in\mathscr{E}_{i.}} x_{ijc}^* \qquad \forall i \in \mathscr{I} \text{ for which } p_{ic}r_{ic} + s_{ic} \ge 0, \tag{2.4a}$$

$$z_{ic}^* = 0 \qquad \forall i \in \mathscr{I} \text{ for which } p_{ic}r_{ic} + s_{ic} < 0. \tag{2.4b}$$

Proof. $(\Rightarrow)$: In the problem (2.1a) – (2.1e), z_{ic} appears only once in a constraint of (2.1b), for each $c \in \mathscr{C}, i \in \mathscr{I}$. Therefore, if its objective function coefficient is nonnegative, its optimal value should be such that (2.1b) is binding, so that (2.4a) is satisfied. On the other hand, if the objective function coefficient of z_{ic} is strictly negative, then its optimal value should be 0, so that (2.4b) is satisfied. Note that if the objective function coefficient of z_{ic} is 0, its optimal value may be any number between 0 and $\lambda_{ic} - \sum_{j\in\mathscr{E}_{i.}} x_{ijc}^*$, but since we use the optimistic formulation of the bilevel program (2.2a) – (2.2g), the optimal value of z_{ic} should satisfy (2.4a).

Therefore, to solve the problem (2.1a) – (2.1e), we only need to consider $(\mathbf{x}_c, \mathbf{z}_c)$ in which (2.4a) – (2.4b) are satisfied. After substituting z_{ic} by its equivalent from (2.4a) – (2.4b) for each $i \in \mathscr{I}$, projecting it out of the formulation, and considering integrality of $\mathbf{x}_c$, (2.3a) – (2.3d) is obtained as a reformulation of (2.1a) – (2.1e). Note that after substituting z_{ic} by its equivalent from (2.4a) – (2.4b), the objective function (2.1a) is as follows:

$$\sum_{(i,j)\in\mathscr{E}} r_{ic}x_{ijc} + \sum_{i\in\mathscr{I}}(p_{ic}r_{ic}+s_{ic})^+(\lambda_{ic} - \sum_{j\in\mathscr{E}_{i.}} x_{ij}) =$$

$$\sum_{(i,j)\in\mathscr{E}} (r_{ic} - (p_{ic}r_{ic}+s_{ic})^+)x_{ijc} + \sum_{i\in\mathscr{I}}(p_{ic}r_{ic}+s_{ic})^+\lambda_{ic}\ .$$

After removing the last summation, which is a constant, (2.3a) is obtained.

($\Leftarrow$) : The proof follows from the reverse steps of that of the first direction. $\square$

Lemma 2. *For each $c \in \mathscr{C}$, the LP relaxation of* (2.3b) – (2.3d) *is an integral polyhedron.*

Proof. We represent constraints (2.3b) – (2.3c) as the projection of the constraints of a minimum cost network flow problem (Ford and Fulkerson 1962). Note that (2.3c) clearly correspond to the flow conservation constraints in a minimum cost network flow problem, and (2.3b) may be represented as the capacity constraints as follows. We construct a network by splitting each node $i \in \mathscr{I}$ of the underlying network of (2.3b) – (2.3c) into two separate nodes and adding a fictitious arc with the capacity of $\lfloor \lambda_{ic} \rfloor$ between these two nodes. Then, (2.3b) is equal to the capacity constraints after projecting out the variables associated with the fictitious arcs.

Since the family of constraints of a minimum cost network flow problem is an integral polyhedron (Ford and Fulkerson 1962) and the projection of an integral polyhedron is integral (Balas 2005), it follows that the LP relaxation of (2.3b) – (2.3d) is integral.

$\square$

Lemma 3. *For each $c \in \mathscr{C}$,*

$$\left\{\mathbf{x}_c \in \mathbb{R}^{|\mathscr{E}|} : \sum_{j\in\mathscr{E}_{i.}} x_{ijc} = \sum_{j\in\mathscr{E}_{.i}} x_{jic}\ \forall i \in \mathscr{I}\right\} = \left\{\mathbf{x}_c \in \mathbb{R}^{|\mathscr{E}|} : \sum_{j\in\mathscr{E}_{i.}} x_{ijc} \le \sum_{j\in\mathscr{E}_{.i}} x_{jic}\ \forall i \in \mathscr{I}\right\}.$$

Proof. Clearly $\subseteq$ holds. To see the other direction, for each $\mathbf{x}_c$ belonging to the left set, by adding all inequalities, it follows that:

$$\sum_{i\in\mathscr{I}}\sum_{j\in\mathscr{E}_{i.}} x_{ijc} \le \sum_{i\in\mathscr{I}}\sum_{j\in\mathscr{E}_{.i}} x_{jic}.$$

The left-side and the right-side of the last inequality are equal to $\sum_{(i,j)\in\mathscr{E}} x_{ijc}$. This implies that all the inequalities describing the left set may be replaced by the equality form.

$\square$

In the sequel, we seek to linearize problem (2.2a) – (2.2g) by Lemmas 1 – 3. For each $c \in \mathscr{C}$ and $\mathbf{s} \in \mathbb{R}^{|\mathscr{I}||\mathscr{C}|}$, let:

$$\phi_c(\mathbf{s}) := \max \sum_{(i,j)\in\mathscr{E}} (r_{ic} - (p_{ic}r_{ic} + s_{ic})^+)x_{ijc} \tag{2.5a}$$

$$s.t. \quad \sum_{j\in\mathscr{E}_{i.}} x_{ijc} \leq \lfloor \lambda_{ic} \rfloor \qquad \forall i \in \mathscr{I}, \tag{2.5b}$$

$$\sum_{j\in\mathscr{E}_{i.}} x_{ijc} = \sum_{j\in\mathscr{E}_{.i}} x_{jic} \qquad \forall i \in \mathscr{I}, \tag{2.5c}$$

$$x_{ijc} \geq 0 \qquad \forall (i,j) \in \mathscr{E}, \tag{2.5d}$$

and $\Delta_c(\mathbf{s}) := \{\mathbf{x}_c \in (2.5b) - (2.5d) | \sum_{(i,j)\in\mathscr{E}} (r_{ic} - (p_{ic}r_{ic} + s_{ic})^+)x_{ijc} \geq \phi_c(\mathbf{s})\}$. Lemmas 1 and 2 imply that for each $c \in \mathscr{C}$, constraint (2.2e) may be rewritten as follows:

$$\mathbf{x}_c \in \Delta_c(\mathbf{s}), \tag{2.6a}$$

$$z_{ic} = \lambda_{ic} - \sum_{j\in\mathscr{E}_{i.}} x_{ijc} \qquad \forall i \in \mathscr{I} \text{ for which } p_{ic}r_{ic} + s_{ic} \geq 0, \tag{2.6b}$$

$$z_{ic} = 0 \qquad \forall i \in \mathscr{I} \text{ for which } p_{ic}r_{ic} + s_{ic} < 0. \tag{2.6c}$$

Note that for each $c \in \mathscr{C}$ and $\mathbf{s} \in \mathbb{R}^{|\mathscr{I}||\mathscr{C}|}$, the problem (2.5a) – (2.5d) is an LP, and its dual is as follows:

$$\min \sum_{i\in\mathscr{I}} v_{ic,1} \lfloor \lambda_{ic} \rfloor \tag{2.7a}$$

$$s.t. \quad v_{ic,1} + v_{ic,2} - v_{jc,2} \geq r_{ic} - (p_{ic}r_{ic} + s_{ic})^+ \qquad \forall (i,j) \in \mathscr{E}, \tag{2.7b}$$

$$v_{ic,1}, v_{ic,2} \geq 0 \qquad \forall i \in \mathscr{I}, \tag{2.7c}$$

where $v_{ic,1}$ and $v_{ic,2}$ are dual variables for constraints (2.5b) and (2.5c), respectively. Note that variable $v_{ic,2}$ is unrestricted in the dual of (2.5a) – (2.5c), but Lemma 3 implies that it may be considered to be nonnegative as in (2.7c). It can easily be seen that the primal (2.5a) – (2.5c) and its dual (2.7a) – (2.7c) are always feasible and bounded.

A significant consequence of our arguments so far is that the bilevel model (2.2a) − (2.2g) may be transformed into an MILP, as follows:

$$\max \sum_{(i,j)\in\mathscr{E}} R_i \left(y_{ij} + \sum_{c\in\mathscr{C}} x_{ijc} \right) - \sum_{i\in\mathscr{I}, c\in\mathscr{C}} \left(\pi_{ic} + \sum_{k\in\mathscr{K}_{ic}} 2^k \tau_{ick} \right) \tag{2.8}$$

$s.t.$

$$\sum_{j\in\mathscr{E}_{i.}} y_{ij} \leq \alpha_i + \sum_{c\in\mathscr{C}} z_{ic} \qquad \forall i \in \mathscr{I}, \tag{2.9}$$

$$\sum_{j\in\mathscr{E}_{i.}} y_{ij} = \sum_{j\in\mathscr{E}_{.i}} y_{ji} \qquad \forall i \in \mathscr{I}, \tag{2.10}$$

$$\sum_{(i,j)\in\mathscr{E}} r_{ic} x_{ijc} + \sum_{i\in\mathscr{I}} p_{ic} r_{ic} z_{ic} + \sum_{i\in\mathscr{I}} \left(\pi_{ic} + \sum_{k\in\mathscr{K}_{ic}} 2^k \tau_{ick} \right) - \sum_{i\in\mathscr{I}} \lambda_{ic} d_{ic} \geq 0 \qquad \forall c \in \mathscr{C}, \tag{2.11}$$

$$- M g_{ic} \leq z_{ic} + \sum_{j\in\mathscr{E}_{i.}} x_{ijc} - \lambda_{ic} \leq 0 \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.12}$$

$$0 \leq z_{ic} \leq M(1 - g_{ic}) \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.13}$$

$$- M g_{ic} \leq p_{ic} r_{ic} + s_{ic} \leq M(1 - g_{ic}) \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.14}$$

$$- M w_{ic} \leq \sum_{j\in\mathscr{E}_{i.}} x_{ijc} - \lfloor \lambda_{ic} \rfloor \leq 0 \qquad \forall i \in \mathscr{I}, \tag{2.15}$$

$$\sum_{j\in\mathscr{E}_{i.}} x_{ijc} = \sum_{j\in\mathscr{E}_{.i}} x_{jic} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.16}$$

$$0 \leq x_{ijc} \leq M q_{ijc} \qquad \forall (i,j) \in \mathscr{E}, c \in \mathscr{C}, \tag{2.17}$$

$$0 \leq v_{ic,1} + v_{ic,2} - v_{jc,2} - r_{ic} + h^+_{ic} \leq M(1 - q_{ijc}) \qquad \forall (i,j) \in \mathscr{E}, c \in \mathscr{C}, \tag{2.18}$$

$$h^+_{ic} - h^-_{ic} = p_{ic} r_{ic} + s_{ic} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.19}$$

$$0 \leq v_{ic,1} \leq M(1 - w_{ic}) \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.20}$$

$$z_{ic} = \{\lambda_{ic}\}(1 - g_{ic}) + \sum_{k \in \mathscr{K}_{ic}} 2^k e_{ick} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.21}$$

$$\tau_{ick} \geq s_{ic} - \overline{s}_{ic}(1 - e_{ick}) \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, k \in \mathscr{K}_{ic}, \tag{2.22}$$

$$\tau_{ick} \geq \underline{s}_{ic} e_{ick} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, k \in \mathscr{K}_{ic}, \tag{2.23}$$

$$\tau_{ick} \leq s_{ic} - \underline{s}_{ic}(1 - e_{ick}) \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, k \in \mathscr{K}_{ic}, \tag{2.24}$$

$$\tau_{ick} \leq \overline{s}_{ic} e_{ick} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, k \in \mathscr{K}_{ic}, \tag{2.25}$$

$$\pi_{ic} \geq s_{ic} - \underline{s}_{ic} g_{ic} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.26}$$

$$\pi_{ic} \geq \underline{s}_{ic}(1 - g_{ic}) \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.27}$$

$$\pi_{ic} \leq s_{ic} - \overline{s}_{ic} g_{ic} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.28}$$

$$\pi_{ic} \leq \overline{s}_{ic}(1 - g_{ic}) \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.29}$$

$$e_{ick} \in \{0, 1\}, \tau_{ick} \text{ unrestricted} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, k \in \mathscr{K}_{ic}, \tag{2.30}$$

$$\pi_{ic} \text{ unrestricted} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.31}$$

$$v_{ic,2}, h^+_{ic}, h^-_{ic} \geq 0, w_{ic} \in \{0, 1\} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.32}$$

$$q_{ijc} \in \{0, 1\} \qquad \forall (i,j) \in \mathscr{E}, c \in \mathscr{C}, \tag{2.33}$$

$$g_{ic} \in \{0, 1\} \qquad \forall i \in \mathscr{I}, c \in \mathscr{C}, \tag{2.34}$$

$$y_{ij} \in \mathbb{Z}_+ \qquad \forall (i,j) \in \mathscr{E}, \tag{2.35}$$

$$s_{ic} \text{ unrestricted} \qquad \forall i \in \mathscr{I}, \tag{2.36}$$

where M is a sufficiently large constant. The family of constraints (2.12) – (2.14) and (2.34) is an MILP reformulation of (2.6b) – (2.6c). Also, the family of constraints (2.15) – (2.20) and (2.32) – (2.33) is an MILP variant of (2.6a), and is derived from the complementary slackness theorem of LP duality. For further details on how to reformulate a lower level LP as a family of MILP constraints, see, e.g., Fortuny-Amat and McCarl (1981). The family of constraints (2.21) – (2.31), (2.13), and (2.34) is to linearize the term $s_{ic} z_{ic}$ in (2.2a), and are derived from Lemma (1), replacing the binary decomposition of z_{ic}, and using well known linearization techniques (Glover 1975). Moreover, the base of the logarithm is 2, and $\mathscr{K}_{ic} := \{0, 1, \cdots, \lfloor \log \lambda_{ic} \rfloor\}$ for all $i \in \mathscr{I}$ and $c \in \mathscr{C}$. As the parameter λ_{ic} is small in our practical setting, the binary decomposition of z_{ic} does not explode the model size. The next theorem formalizes our arguments in this section.

Theorem 1. *The optimal solution of the MILP* (2.8) − (2.36) *coincides with that of the bilevel program* (2.2a) − (2.2g).

2.5 CALIBRATION

In this section, we calibrate our model using a data set that will be described in detail in what follows.

Patient-donor pair classes. We partition patient-donor pairs by ABO blood types of the patient and the donor of a pair (cf. Ünver 2010). Specifically, there are four ABO blood types for each patient and each donor: O, A, B, AB. Hence, there are sixteen patient-donor pair classes. We denote a class as X-Y if the patient and the donor of each pair in the class have blood types X and Y, respectively. For instance, Class A-B consists of all pairs whose patients and donors have blood types A and B, respectively.

Compatibility. A patient who is biologically compatible with a donor can receive the donor's kidney. Biological compatibility consists of both blood-type and tissue-type compatibilities. As noted earlier, there are four ABO blood types: O, A, B, AB. Blood-type compatibility possesses a well-defined structure as follow: (1) An O (blood-type) patient is blood-type compatible only with O donors. (2) An A patient is blood-type compatible with both O and A donors. (3) A patient of blood-type B is blood-type compatible with both O and B donors. (4) An AB patient is blood-type compatible with all donors. When a patient and a donor are blood-type compatible, tissue-type compatibility should be investigated. Unlike blood-type compatibility, tissue-type compatibility lacks a well-defined structure. Tissue-type incompatibility occurs due to the presence of antibodies in a patient's blood cells that would damage the kidney from a specific donor. These antibodies rise for several reasons, mostly from exposure to another person's cells, usually due to pregnancy, blood transfusion or prior transplant. To detect the presence of antibodies, the cross-match test is conducted in which the bloods of the patient and the donor are mixed. If the donor's cells die, the cross-match is positive, i.e., the patient cannot receive a kidney from the donor. Otherwise, they are compatible. This implies that before conducting a cross-match test, it is impossible to realize whether a patient is tissue-type compatible with a specific donor. As a consequence of this restriction, tissue-type compatibility has not been modeled in many related papers (e.g., Roth et al. 2007, Ünver 2010). Following this stream of literature, we too skip tissue-type compatibility by the following assumption.

Assumption 1. *(Roth et al. 2007, Ünver 2010) No patient is tissue-type incompatible with the donor of another pair.*

Note that a patient may be tissue-type incompatible with his intended donor, and arrival rates of blood-type compatible patient-donor pairs to the national PKE program is hence allowed to be greater than zero.

Arrival rates of patient-donor pairs. As data regarding the parameter λ_{ic} for all $i \in \mathscr{I}$ and $c \in \mathscr{C}$ are unavailable, we describe a process to estimate this set of parameters. This process consists of two steps: First, we estimate the number of patient-donor pairs in each class in the United States. Second, we estimate how patient-donor pairs are distributed among transplant centers based on the number of kidney transplants performed in each transplant center so far.

By following the approach described in Saidman et al. (2006), the most prevalent approach in the literature, we derive the probability distribution of patient-donor classes in the United States. The necessary data to apply the approach of Saidman et al. (2006) is based on SRTR (2015) and reported in Table 1. Note that the calculated panel reactive antibody (CPRA) in Table 1 is a measure of sensitization level, and estimates the percentage of donors that would be tissue-type incompatible with a patient. Table 2 reports the resulting probability distribution of patient-donor classes in the United States. This table implies that 46.68% of patient-donors are compatible. The number of living-donor transplants in 2012 is 5,346 (SRTR 2015). As of 2010, around 1000 PKEs have been conducted in the United

States (Rees et al. 2012), although over 15000 living-donor transplants have been conducted only in 2010-2012 (SRTR 2015). Hence, with a slight loss of accuracy, we can say that almost all 5,346 living-donor transplants in 2012 have been performed on compatible patient-donor pairs. Therefore, the total number of compatible and incompatible patient-donor pairs in the United States may conservatively be estimated as $11,000 \approx \frac{5{,}346}{0.4668}$. Given that we estimated the patient-donor population size and the distribution of patient-donor classes, the number of patient-donor pairs within each class in the United States may easily be estimated.

The number of kidney transplants performed in each transplant center from the start of living-donor transplant in 1988 to the end of 2014 is available at OPTN Data 2015. Assuming that the number of incompatible patient-donor pairs arriving at each transplant center is proportional to the number of kidney transplants performed in the transplant center, we can estimate how patient-donor pairs are distributed across transplant centers. Knowing this distribution and the number of patient-donor pairs within each class in the United States, we can estimate the parameter λ_{ic} for all $i \in \mathscr{I}$ and $c \in \mathscr{C}$. Finally, we let $\alpha_i = 0$ for all $i \in \mathscr{I}$.

Table 1: Probability characteristics of the patient-donor population.

Patient/Donor blood type	Frequency (percent)
O	51.2
A	30.1
B	15.5
AB	3.2
CPRA	Frequency (percent)
$0- < 1\%$	50.5
$1 - 19\%$	18.4
$20 - 79\%$	16.3
$80 - 100\%$	14.7
Gender	Frequency (percent)
Male	59.3
Female	40.7
Unrelated living donors	Frequency (percent)
Spouse	48.97
Other	51.03

The distributions are based on SRTR (2015).

Probabilities. To estimate the parameter p_{ic} for all $i \in \mathscr{I}$ and $c \in \mathscr{C}$, we consider a hypothetical pool in which all patient-donor pairs are enrolled. Then, we match patient-donor pairs in this pool with the goal of maximizing the number of transplants. Next, we

Table 2: Distribution of patient-donor population.

Incompatible patient-donor population				
	Patient blood type			
Donor blood type	O	A	B	AB
O	0.0716	0.0421	0.0217	0.0045
A	0.1541	0.0247	0.0467	0.0026
B	0.0794	0.0467	0.0066	0.0014
AB	0.0164	0.0096	0.0050	0.0003
Compatible patient-donor population				
	Patient blood type			
Donor blood type	O	A	B	AB
O	0.1906	0.1120	0.0577	0.0119
A	0.0000	0.0659	0.0000	0.0070
B	0.0000	0.0000	0.0175	0.0036
AB	0.0000	0.0000	0.0000	0.0007

This distribution is derived from Table 1.

estimate p_{ic} for all $i \in \mathscr{I}$ and $c \in \mathscr{C}$ by the proportion of patients belonging to class $i \in \mathscr{I}$ who are transplanted in the hypothetical pool. These probabilities are reported in Table 3.

Our approach to estimate the parameter p_{ic} is plausible because: (1) As will be shown, our numerical results demonstrate that a significant portion of patient-donor pairs should be enrolled in the national PKE program under an equilibrium payment strategy. (2) The size of a transplant center's pool is insignificant compared to that of a national pool, and hence the chance of being matched in a national PKE program for an arbitrary patient-donor pair is almost independent of a single transplant center's decisions.

Table 3: Probabilities of being matched in a national PKE program for an incompatible patient-donor pair.

	Patient blood type			
Donor blood type	O	A	B	AB
O	1.0000	1.0000	1.0000	1.0000
A	0.2796	1.0000	1.0000	1.0000
B	0.3165	1.0000	1.0000	1.0000
AB	0.0000	0.3714	0.9630	1.0000

Societal reward. Each kidney transplant saves society one life and thousands of dollars in foregone medical expenses. Matas and Schnitzler (2003) studied the benefits of living-donor transplant and quantified its value by a cost-effectiveness analysis. In their comprehensive study, they considered different factors such as patient survival, cost of dialysis, graft survival, death with function, death after graft loss, cost of organ acquisition, cost of transplant, maintenance cost with graft function, and cost of return to dialysis after living-donor transplant. They found that each additional living-donor transplant is worth at least \$269,319 for society. After incorporating inflation (consumer price index), the updated value of an additional living-donor transplant (R_i for all $i \in \mathscr{I}$) is \$334,648.46 (Bureau of Labor Statistics 2015).

Transplant center costs and rewards. The donor's pre-match cost (d_{ic}) encompasses expenses for a set of procedures such as blood testing, electrocardiogram, chest x-ray, bloodwork, urine tests, pressure check, medical history and physical examination, psychosocial assessment, education session, CT scan, and surgeon visit. Depending on the donor's gender, age, and genetic make-up other medical tests may be required, e.g., mammogram, pap smear, prostate-specific antigen, exercise stress test and a 2-hour glucose tolerance test. In this regard, we contacted the University of Pittsburgh Medical Center staff who estimated d_{ic} to be \$1,554.82 on average for all $i \in \mathscr{I}$ and $c \in \mathscr{C}$ (Tevar 2015).

For conducting each kidney transplant, each transplant center receives an income in return for providing different services such as physical, psychological, and laboratory evaluations, room, board, and ancillary services, and etc. Bentley (2014) investigated the related expenses in the interval of one month pre-transplant to six months post-transplant. The income of a transplant center for each kidney transplant is \$419,200 on average (Bentley 2014). Moreover, hospital operating margin was 5.7% in 2013 (American Hospital Association 2015). As a result, we can estimate r_{ic} to be $\$419,200 \times 0.057 = \$23,894.4$ for all $i \in \mathscr{I}$ and $c \in \mathscr{C}$.

2.6 POLICY INSIGHTS

There are various frameworks on which a national PKE program may be established. Our model can quantify benefits of these frameworks for society and different stakeholders, and provide valuable policy insights.

2.6.1 Benefits of a National PKE Program

Although it is intuitive that a national (centralized) PKE program would perform better than the current (decentralized) PKE program, it is important to provide quantitative insights about benefits of a national PKE program for different stakeholders. Indeed, we must ensure that benefits of establishing a national PKE program outweigh its setup and variable costs. The more the benefits of a national PKE program are realized, the more incentives for society exist to establish such a program as soon as possible. Hence, we compare the performance of a national PKE program with that of a decentralized PKE program with respect to the following criteria:

C.1 The societal objective, which is the value of (2.2a),

C.2 The total number of transplants,

C.3 The total number of internal transplants,

C.4 The total income of transplant centers,

C.5 The total cost of insurance companies.

Table 4 reports these criteria under several different settings. For each setting, **C.1-C.3** are obtained by running the model (2.8) – (2.36) with parameters extracted from the associated data set. Given that each transplant surgery brings $419,200 income to a transplant center, **C.4** is equal to **C.2** multiplied by 419,200. Annual cost of an ESRD patient treated with hemodialysis and transplant are, respectively, $90,026.89 and $32,803.99 in 2015 (Rees et al. 2012, Bureau of Labor Statistics 2015). For each setting, **C.5** is annual total of such costs for all patient-donor pairs.

Table 4 demonstrates that a national PKE program, as compared to a decentralized PKE program, leads to saving 1,340 additional lives, enhancing the societal objective by more than \$0.5 billion, increasing transplant centers' income by more than \$0.5 billion, and decreasing insurance costs by \$75 million annually. Of note, 1,296 patient-donor pairs are internally transplanted in the national PKE program, which is more than half of the number of patient-donor pairs who are all internally transplanted in a decentralized PKE program. Internal transplants are favorable in the sense that the associated patient-donor pairs do not need to move across transplant centers, which is more convenient and suppresses traveling costs for these patient-donor pairs.

2.6.2 SAC Settings

Note that in our model, the SAC compensation s_{ic} depends on patient-donor pair class i and transplant center c. We call this setting *center-and-pair-dependent SAC*. Although this setting seems plausible, other possible settings for SAC compensation are: (1) *Pair-dependent SAC*: Drop the dependency on transplant center c and have a SAC compensation s_i that depends only on patient-donor pair class i. (2) *Center-dependent SAC*: Drop the dependency on patient-donor pair class i and have a SAC compensation s_c that depends only on transplant center c. (3) *Fixed SAC*: Drop the dependency on both transplant center c and patient-donor class i and have a SAC compensation s that is fixed for all transplant center $c \in \mathscr{C}$ and patient-donor class $i \in \mathscr{I}$. Clearly, the societal benefit of a center-and-pair-dependent-SAC setting is higher than that of the other settings. However, it is not clear which setting works better in practice since each one has its own advantages and disadvantages. For instance, as a disadvantage of a center-dependent SAC, Rees et al. (2012) noted that "As the number of KPD [kidney paired donation] transplants performed at each center may be highly variable, such an approach may lead to large variability in center-specific KPD SAC charges and payment of widely disparate costs between transplant centers could be a disincentive to centers participating in such exchanges." To investigate the benefits of these settings, we report their associated **C.1-C.5** in Table 4. As this table shows, all these different settings result in the same number of transplants, which in turn leads to identical transplant centers' income and insurance cost. However, they have different societal objective values. To find out why, recall that the societal objective is $\sum_{(i,j)\in\mathscr{E}} R_i \left(y_{ij} + \sum_{c\in\mathscr{C}} x_{ijc} \right) - \sum_{i\in\mathscr{I},c\in\mathscr{C}} s_{ic}z_{ic}$, whose first term is the reward related to transplants and second term is the amount money spent on (earned by) the participation of transplant centers. The amount of the first term is \$1.269 billion under the optimal SAC compensations for all the settings. However, the amount of the second term is \$88 million, \$65 million, 0, and 0 for the center-and-pair-dependent SAC, the pair-dependent SAC, the center-dependent SAC, and the fixed SAC, respectively, which leads to different societal objectives for these settings. Of note, in the center-and-pair-dependent SAC and the pair-dependent SAC, the national PKE program makes 88, 65 million dollar annual income, respectively, which justifies the cost of establishing and running a national PKE program. Given the space limitation, we are unable to report the optimal SAC compensations for the center-and-pair-dependent SAC and the center-dependent SAC. We report the optimal SAC compensations for the pair-dependent SAC in Table 5. In addition, the optimal SAC compensation for the fixed SAC setting is equal to zero, which means that transplant centers pay nothing for participation in the national PKE program.

2.6.3 Authoritarian Setting

A hypothetical setting of interest is when transplant centers are required by a law to fully participate in a national PKE program. Specifically, transplant centers should enroll all their patient-donor pairs in the national PKE program, and they do not receive (or pay) any money for that. We call this setting *authoritarian*, and report its associated **C.1-C.5** in Table 4. The results of this table show that the authoritarian setting cannot reach more transplants than those of any SAC-based setting, and hence there is no benefit in enforcing transplant centers to participate in a national PKE program. Indeed, as noted earlier, transplant centers are even willing to annually pay \$88, \$65 million for enrolling their patient-donor pairs in a national PKE program with a center-and-pair-dependent SAC or a pair-dependent SAC, respectively.

Table 4: Annual total reward (cost) of different stakeholders under different settings.

SAC settting	C.1	C.2	C.3	C.4	C.5
Center-and-pair-dependent SAC	\$1.357E+09	3793	1296	\$1.590E+09	\$3.109E+08
Pair-dependent SAC	\$1.335E+09	3793	9	\$1.590E+09	\$3.109E+08
Center-dependent SAC	\$1.269E+09	3793	1586	\$1.590E+09	\$3.109E+08
Fixed SAC	\$1.269E+09	3793	1731	\$1.590E+09	\$3.109E+08
Decentralized	\$8.209E+08	2453	2453	\$1.028E+09	\$3.875E+08
Authoritarian	\$1.269E+09	3793	N/A	\$1.590E+09	\$3.109E+08

2.6.4 Participation of Compatible Patient-donor Pairs

Recall that, based on the physiological compatibility, patient-donor pairs are classified into two groups: (1) compatible patient-donor pairs, and (2) incompatible patient-donor pairs. The patient can receive the donor's kidney for each compatible patient-donor pair, but this is not the case for an incompatible patient-donor pair. PKE is a modality to resolve this issue of incompatible patient-donor pairs, and clearly compatible patient-donor pairs lack incentives to participate in a PKE. Hence, a national PKE program will be restricted only to incompatible patient-donor pairs. However, a national PKE program will reach far more benefits if compatible patient-donor pairs participate in the program by letting it find compatible matches for them. Sonmez and Ünver (2015) proposed a new incentive scheme to encourage compatible patient-donor pairs to participate, and analyzed potentials of the scheme in terms of enhancing efficiency and equity. To illuminate advantages of compatible patient-donor pair participation, we run our model for this case and report its results in Table 6 for settings described in Subsection 2.6.1. Note that the setting "Current" in Table 6 presents the situation where compatible patient-donor pairs do no participate and PKE for incompatible patient-donor pairs is conducted in a decentralized setting.

Table 6 demonstrates that a national PKE program with participation of compatible patient-donor pairs compared to what current practice in the United States PKE program

Table 5: Optimal SAC compensations in US dollars by patient-donor blood types.

	Patient blood type			
Donor blood type	O	A	B	AB
O	\$-23894.4	\$-23894.4	\$-23894.4	\$201724
A	\$-6682.07	\$-23894.4	\$-23894.4	\$0
B	\$-7562.82	\$-23894.4	\$-23225.9	\$0
AB	\$0	\$-8875.1	\$-19359.8	\$0

will save at least 3400 additional lives, enhancing the societal objective by \$1.38 billion, increasing transplant centers' income by \$1.43 billion, and decreasing insurance costs by \$0.195 billion annually. Note that, these benefits are almost three times as large as those if compatible patient-donor pairs do not participate. In particular, a national PKE program with participation of compatible patient-donor pairs creates \$0.847 billion more profit for society than that without participation of compatible patient-donor pairs. This additional profit may be invested in incentivizing compatible patient-donor pairs to participate. Note that we are fully aware of ethical dilemmas on monetary transfer in organ exchange, and we do not recommend its use. We only seek to shed light on advantages of participation of compatible patient-donor pairs.

All the SAC-based settings provide roughly the same number of transplants which in turn leads to the same transplant centers' income and insurance costs. Moreover, center-and-pair dependent SAC, pair-dependent SAC, and center-dependent SAC settings deliver almost the same societal objective. However, the societal objective of the fixed SAC setting is \$0.225 billion smaller than those of the other SAC-based settings. Finally, the authoritarian setting delivers (almost) the same number of transplants as that of any SAC-based setting, and however its societal objective is at least \$0.225 billion smaller than those of center-and-pair dependent SAC, pair-dependent SAC, and center-dependent SAC settings. This \$0.225 billion represents the amount of money that a national PKE program charges transplant centers for enrolling their patient-donor pairs.

Table 6: Annual total reward (cost) of different stakeholders under different settings when compatible patient-donor pairs participate.

SAC setting	**C.1**	**C.2**	**C.3**	**C.4**	**C.5**
Center-and-pair-dependent SAC	\$3.923E+09	10990	2169	\$4.607E+09	\$3.614E+08
Pair-dependent SAC	\$3.903E+09	10992	36	\$4.608E+09	\$3.613E+08
Center-dependent SAC	\$3.923E+09	10990	2169	\$4.607E+09	\$3.614E+08
Fixed SAC	\$3.698E+09	10992	0	\$4.608E+09	\$3.613E+08
Current	\$2.539E+09	7587	2453	\$3.180E+09	\$5.561E+08
Authoritarian	\$3.678E+09	10992	N/A	\$4.608E+09	\$3.613E+08

3.0 OPTIMIZING OVER PURE STATIONARY EQUILIBRIA IN CONSENSUS STOPPING GAMES

3.1 MOTIVATION

Discrete-time stochastic games have modeled dynamic competitive interactions among multiple players since they were introduced by Shapley (1953). A stochastic game consists of periods, states, actions, rewards, players, and transition probabilities. In each period, the stochastic game occupies a state, and all players choose their actions simultaneously and independently. Subsequently, each player receives a reward that depends on the current state and the actions of all players. The game transitions to the next state according to a discrete probability distribution, conditioned on the current state and the chosen actions. Each player seeks to maximize his own reward criterion, e.g., his total expected discounted reward.

A *strategy* for each player specifies a probability distribution over the feasible actions in each period conditioned on the current state and the history of the game up to that period. If this distribution depends only on the current state, then the strategy is called *stationary*. A strategy is called *pure* when all the probabilities are in $\{0, 1\}$. A (stationary) *strategy profile* is a collection of (stationary) strategies of all players that fully specifies all actions in the game, and it must include one and only one (stationary) strategy for each player (Solan 2012). To analyze stochastic games, there are several solution concepts such as Nash equilibrium, subgame-perfect equilibrium, and stationary equilibrium. A *subgame-perfect equilibrium* is a strategy profile that is a Nash equilibrium of every subgame of the original game. A *stationary equilibrium* is a stationary strategy profile that maximizes every

[0] This chapter is based on Dehghanian et al. (2015).

player's reward criterion in each state (among all stationary and non-stationary strategy profiles) given the strategies of the other players. Equivalently, a stationary equilibrium is a subgame-perfect equilibrium that is stationary. There are several reasons to analyze stationary equilibria in stochastic games. First, it prescribes the simplest form of behavior that is consistent with rationality. Second, compared to subgame-perfect equilibrium, a stationary equilibrium represents the notion of being free from the past, i.e., "bygones are bygones", more completely through a state-contingent behavior (Maskin and Tirole 2001). The existence of a stationary equilibrium for a discounted stochastic game with finite state and action spaces has long been established (e.g., Fink 1964, Sobel 1971).

Stochastic games are very challenging since it is often difficult to characterize even stationary equilibria (Doraszelski and Escobar 2010, Hörner et al. 2011). Solan (2012) notes that "Unfortunately, to date there are no efficient algorithms to calculate either the value in zero-sum stochastic games, or equilibria in non-zero-sum games." In many contexts, players may prefer to use pure stationary strategies; however, the characterization of pure stationary equilibria is even more challenging. Another issue that has limited the applicability of stochastic games is the existence of multiple stationary equilibria, and it is difficult to find all such equilibria (Aguirregabiria and Mira 2007). Two major complications arise as consequences of the existence of multiple equilibria (Köppe et al. 2011): First, it becomes more difficult to predict players' behavior, and players may not reach an equilibrium at all. Second, many existing methods find one equilibrium and provide no systematic methodology to find all equilibria. Such complications raise an even more challenging question of finding a best stationary equilibrium with respect to a given criterion in stochastic games.

We consider *consensus stopping games*, a broad class of stochastic stopping games defined over a finite set of players, states, actions, and rewards. Such a game dynamically evolves over an infinite-horizon setting. In each period of the game, each player decides whether to offer to stop or continue the game. If all players offer to stop, the game terminates, and each player receives a lump-sum stopping reward. Conversely, if at least one player decides to continue, each player receives an immediate continuation reward, the game moves into the next state according to a Markovian transition, and the rest of the players must continue, regardless of their decisions. In this chapter, we study the problem of finding a best pure stationary equilibrium for this class of games. Our primary motivations for studying this problem include, but not limited to: (1) Optimizing over (pure stationary) equilibria, rather than identifying such equilibria is inherently interesting, and we are the first who investigate it comprehensively. (2) Consensus stopping games are an important class of stochastic games with various applications, as discussed in the rest of this section.

Consensus stopping games arise in the context of consensus decision-making, which is indeed used by many international organizations in policy making. For instance, the Council of the European Union requires unanimity in some policy areas such as membership, taxation, social security, foreign and common defense policy and operational police cooperation among the Member States (The Treaty of Lisbon 2007). The World Trade Organization, the Association of Southeast Asian Nations, the North Atlantic Treaty Organization, the Conference on Security and Cooperation in Europe, the Executive Committee of the International Monetary Fund, and the Organization for Economic Cooperation and Development all make decisions by consensus (Steinberg 2002). Indeed, in consensus decision-making, every player has a veto in the sense that he can prevent a change from the status quo (Tsebelis 2002).

In principle, consensus stopping games may model many dynamic noncooperative consensus decision-making environments to reach a permanent agreement. We elaborate on the elements of a consensus stopping game, e.g., the players, periods, actions, states, rewards, and transition probabilities, for two specific applications.

- **Organ exchange.** ESRD, as described in Chapter 2, is the final stage of chronic kidney disease in which both kidneys almost fail. The preferred choice of treatment for ESRD is living-donor transplantation, in which a living-donor donates one of his kidneys to the patient. A significant barrier to greater use of living-donors is that at least one-third of the patients with a willing living-donor are unable to receive the donor's kidney due to blood type and/or tissue incompatibilities (Montgomery et al. 2005b). To mitigate this barrier, an emerging clinical practice is PKE in which $N \geq 2$ self-interested patient-donor pairs for whom the only compatible exchange of kidneys is cyclical, swap their donors (see Figure 2). Periodically, each patient-donor pair decides whether to offer to exchange or not. If at least one patient-donor pair decides not to offer to exchange, the transplant

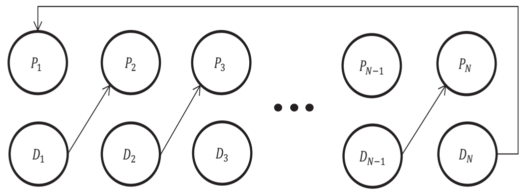

Figure 2: A cycle of N patient-donor pairs for a PKE, where D_i refers to donor of intended recipient (or patient) P_i. Directed arcs in the cycle represent compatibility between patients and donors; that is, no donor is compatible with his intended recipient but only compatible with the intended recipient of the next donor in the cycle.

exchange cycle breaks as each donor is willing to donate a kidney only in return for receiving a kidney for his indended recipient. Therefore, the transplants in the cycle are accomplished if and only if all patient-donor pairs offer to exchange consensually. To value a given PKE cycle when the timing of the transplants is determined by self-interested, autonomous patient-donor pairs, Kurt et al. (2011) propose a consensus stopping game in which the players are the patient-donor pairs, the decision epochs are biweekly, and the players' actions are whether to offer to exchange or not. The glomerular filtration rates (GFRs), a measure of kidney functionality, of all patients are considered as the state of the game. Roughly speaking, the immediate continuation and stopping rewards are estimated by the expected number of days until the next decision epoch (fourteen days) and the expected number of post-transplant survival days for each patient, respectively. A transition probability matrix for the case of not offering to exchange describes the Markovian progression of the GFRs. Finally, Kurt et al. (2011) represent the problem of finding a best pure stationary equilibrium of the game as an MILP which they solve using a commercial MILP solver. Consensus stopping games may be applied to model the timing of transplants for other organ exchange problems, e.g., lung exchange (Ergin et al. 2015).

- **War termination.** Two countries fight a war against each other until reaching peace or one side's definite victory, whichever occurs first. Periodically, each country chooses between continuing the war and offering peace. The war ends if and only if both countries offer peace simultaneously. This situation may be modeled by a consensus stopping game in which the players are countries, the decision epochs are, for example, daily, and the actions are whether to continue the war or offer peace. Political scientists have utilized the Composite Index of National Capabilities (CINC) score to measure power of a country and explain war outcomes. In calculating CINC scores, geopolitical factors such as a country's relative military, economic, and demographic capabilities are considered (e.g., Singer et al. 1972, Singer 1988). Therefore, we may consider the CINC scores of both countries as the state of the game. The expected increase in the area of the occupied territories and the total area of the occupied territories since the beginning of the war may be regarded as the immediate continuation and stopping rewards, respectively. A transition probability matrix for the dynamic evolution of the CINC scores in the case of continuing the war reflects exogenous factors, which are out of the control of both countries, such as natural events, third parties' actions, etc. Other models of war termination may be found in the political science literature. For instance, Filson and Werner (2002) present a two-stage asymmetric information bargaining game by which they provide several insights on the onset and termination of war. As another instance, Cunningham (2006) studies the correlation between the duration of a civil war and the number of veto players. This statistical analysis shows that the more the number of veto players, the longer the conflict.

3.1.1 Summary of Contributions

Motivated by consensus decision-making and its applications, we consider consensus stopping games. Our study reveals a rich structure of these games that allows us to investigate the problem of finding a best pure stationary equilibrium. Such a problem has not been comprehensively investigated before for a class of stochastic games. The specific contributions of this chapter are as follows: First, we establish that the problem of finding a best pure stationary equilibrium of a consensus stopping game is NP-hard. Second, we characterize the pure stationary equilibria, and show that they form an independence system. Accordingly, we derive two families of combinatorial valid inequalities for an MILP developed for the problem of finding a best pure stationary equilibrium. Third, we develop an efficient branch-and-cut algorithm to solve the MILP by applying these valid inequalities. We also develop a family of Pareto-optimal optimality cuts and several algorithmic enhancements. Our extensive computational experiments show that the algorithm significantly outperforms a state-of-the-art commercial MILP solver.

This work is the first to provide combinatorial characterizations of stationary equilibria for a class of stochastic games. It is also the first attempt to develop a novel cutting plane approach for the problem of finding a best pure stationary equilibrium for a class of stochastic games.

3.1.2 Outline of the Chapter

The remainder of this chapter is organized as follows. In Section 3.2 we review the literature. In Section 3.3 we define consensus stopping games and the problem of finding a best pure stationary equilibrium, which we show to be NP-hard in Section 3.4. We characterize the pure stationary equilibria and develop two families of valid inequalities in Section 3.5. In Section 3.6 we develop an MILP formulation that optimizes over the set of pure stationary equilibria. We develop a branch-and-cut approach in Section 3.7, and describe our computational experiments in Section 3.8.

3.2 LITERATURE REVIEW

The stochastic game literature is vast (see the recent survey by Solan 2012). In the economics literature, stochastic games typically model dynamic interactions among firms (Ericson and Pakes 1995, Doraszelski and Satterthwaite 2010). To find a stationary equilibrium of a stochastic game, a common approach is to apply the homotopy method (Herings and Peeters 2004, Borkovsky et al. 2010). Weintraub et al. (2008) introduce the oblivious equilibrium notion for Ericson and Pakes (1995)-style models, and show that it can approximate stationary equilibria under some conditions. Weintraub et al. (2010) develop an algorithm for computing an oblivious equilibrium. In the operations research literature, mathematical programming has been used to compute a stationary equilibrium of stochastic games. For instance, Filar and Vrieze (1997) and Raghavan and Syed (2003) compute a stationary equilibrium for certain classes of stochastic games. Filar et al. (1991) present a nonlinear programming formulation whose global optima are the stationary equilibria of a finite discounted stochastic game; however, there was no computational study. Note that these approaches attempt to identify a single stationary equilibrium, and because of the multiplicity of stationary equilibria in stochastic games (Herings and Peeters 2004, Aguirregabiria and Mira 2007, Doraszelski and Satterthwaite 2010), they may be viewed as heuristics for finding a best stationary equilibrium.

Our work is also related to the literature of stopping games (for a survey, cf. Nowak and Szajowski 2005, Part III). The literature includes different variants for the definition of stopping games (Heller 2012). This stream of research usually addresses the existence of an equilibrium (cf. Solan and Vieille 2001, Shmaya and Solan 2004, Heller 2012).

What distinguishes this work from the literature of stochastic games and stopping games is that we focus on the more challenging question of finding a best pure stationary equilibrium, compared to the question of finding a stationary equilibrium or the question of establishing the existence of an equilibrium because for consensus stopping games, (1) finding a stationary equilibrium is trivial; (2) the game may possess many pure stationary equilibria, many of which may be Pareto-inefficient with respect to the players' payoff profile; and (3) we are able to optimize over pure stationary equilibria by providing effective algorithmic approaches to choose among those equilibria.

3.3 CONSENSUS STOPPING GAMES

We provide a detailed description of consensus stopping games, and present necessary and sufficient conditions for a strategy profile to be a pure stationary equilibrium. We define a consensus stopping game, $\mathcal{G}$, as follows: Let $\mathscr{N} = \{1, 2, \ldots, N\}$ be a set of players, and $\mathscr{S}$

represent the finite state space of $\mathcal{G}$. In each period, each player decides whether to offer to stop or continue based on the current state $s \in \mathscr{S}$. All players make their decisions independently and simultaneously, and $a_i(s) \in \mathbb{B} := \{0, 1\}$ denotes player i's action in state s, where $a_i(s)$ is 1 if he offers to stop, and 0 otherwise. Because we only focus on pure stationary strategies, $a_i(s)$ is sufficient to characterize the action of player i in state s in each period. If all players offer to stop in state s (i.e., $\prod_{i \in \mathscr{N}} a_i(s) = 1$), $\mathcal{G}$ terminates and each player $i \in \mathscr{N}$ receives a lump-sum stopping reward $u_i(s, 1)$. If at least one player decides to continue (i.e., $\prod_{i \in \mathscr{N}} a_i(s) = 0$), $\mathcal{G}$ moves into the next state s' under a Markovian transition probability $\mathscr{P}(s'|s)$ while each player $i \in \mathscr{N}$ receives an immediate continuation reward $u_i(s, 0)$. Each player $i \in \mathscr{N}$ has a periodic discount factor $\lambda_i \in [0, 1)$, accounting for the time value of his future rewards, and he seeks to maximize his total expected discounted reward. $\mathcal{G}$ is an almost perfect information game, i.e., at the beginning of each period, all players are perfectly informed of the current state along with all the actions and states that have already been realized. Note that relative to Kurt et al. (2011), $\mathcal{G}$ permits more general state space and reward structures. Specifically, the reward $u_i(s, 0)$ or $u_i(s, 1)$ may be negative, and the state space $\mathscr{S}$ need not be the Cartesian product of the individual player's state spaces.

We follow the convention that a term in bold refers to a real-valued vector; i.e., $\mathbf{v}$ refers to the vector $\langle v(s) \rangle_{s \in \mathscr{S}}$. Given a vector $\mathbf{v} \in \mathbb{R}^{|\mathscr{S}|}$, define $F_i(s, \mathbf{v}) := u_i(s, 0) + \lambda_i \sum_{s' \in \mathscr{S}} \mathscr{P}(s'|s) v(s')$ for any $s \in \mathscr{S}, i \in \mathscr{N}$, which we interpret as player i's expected continuation payoff starting from state s where $\mathbf{v}$ represents the payoffs of all possible states at the next period. In the sequel of this chapter, unless otherwise stated, we use the terms *strategy* and *equilibrium* to refer to pure stationary strategy and pure stationary equilibrium, respectively. Let $\mathbf{a}_i := \langle a_i(s) \rangle_{s \in \mathscr{S}}$ denote a strategy of player i for each $i \in \mathscr{N}$, $\mathbf{a} := \langle \mathbf{a}_i \rangle_{i \in \mathscr{N}}$ denote the resulting strategy profile. Moreover, let $w_i^{\mathbf{a}}(s)$ denote the total expected discounted reward of player i in state s under strategy profile $\mathbf{a}$. To formalize this notion, let s_t denote the state of the game in period t, and $r_i(s_t, \mathbf{a}(s_t))$ denote the reward of player $i \in \mathscr{N}$ under strategy profile $\mathbf{a}$ when state s_t is realized. Then,

$$w_i^{\mathbf{a}}(s) = \mathbb{E} \left\{ \lim_{n \to +\infty} \sum_{t=0}^{n} \lambda_i^t r_i(s_t, \mathbf{a}(s_t)) \middle| s \right\},$$

where $\mathbb{E}(\cdot)$ represents the expectation operator under the probability distribution induced by strategy profile $\mathbf{a}$ over the evolution of the states when the game is initialized in state s. Accordingly, $\mathbf{w}_i^{\mathbf{a}} := \langle w_i^{\mathbf{a}}(s) \rangle_{s \in \mathscr{S}}$ represents the payoff vector for player i under strategy profile $\mathbf{a}$. In a slight abuse of notation, let $\mathbf{w}^{\mathbf{a}} := \langle \mathbf{w}_i^{\mathbf{a}} \rangle_{i \in \mathscr{N}}$ represent the payoff profile under strategy profile $\mathbf{a}$. The outcome of the players' decisions (continuation or termination of the game) at each state s is uniquely characterized by a binary variable $x(s) := \prod_{i \in \mathscr{N}} a_i(s)$, where $\mathbf{x} := \langle x(s) \rangle_{s \in \mathscr{S}} \in \mathbb{B}^{|\mathscr{S}|}$ is the corresponding vector. Note that if all players offer to stop, then $x(s) = 1$; otherwise, $x(s) = 0$. For each $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$, let $A_{\mathbf{x}} := \left\{ \mathbf{a} \in \mathbb{B}^{|\mathscr{S}| \times |N|} | \prod_{i \in \mathscr{N}} a_i(s) = x(s) \, \forall s \in \mathscr{S} \right\}$ which represents the set of strategy profiles inducing the same outcome for all $s \in \mathscr{S}$.

Proposition 1. *Given a strategy profile* $\mathbf{a}$ *with associated payoff profile* $\mathbf{w}^{\mathbf{a}}$*:*
(i) For all $s \in \mathscr{S}, i \in \mathscr{N}$*,*

$$w_i^{\mathbf{a}}(s) = \left(\prod_{i \in \mathscr{N}} a_i(s) \right) u_i(s, 1) + \left(1 - \prod_{i \in \mathscr{N}} a_i(s) \right) F_i(s, \mathbf{w}_i^{\mathbf{a}}). \tag{3.1}$$

Furthermore, $\mathbf{w}^{\mathbf{a}}$ *is the unique solution for this set of equations.*
(ii) $\mathbf{a}$ *is an equilibrium if and only if for all* $s \in \mathscr{S}, i \in \mathscr{N}$*:*

$$w_i^{\mathbf{a}}(s) = \max \left\{ \left(\prod_{\substack{j \in \mathscr{N}, \\ j \neq i}} a_j(s) \right) u_i(s, 1) + \left(1 - \prod_{\substack{j \in \mathscr{N}, \\ j \neq i}} a_j(s) \right) F_i(s, \mathbf{w}_i^{\mathbf{a}}), F_i(s, \mathbf{w}_i^{\mathbf{a}}) \right\}. \tag{3.2}$$

(iii) For each $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$*, there exists an equilibrium in* $A_{\mathbf{x}}$ *if and only if strategy profile* $\bar{\mathbf{a}}$*, defined by* $\bar{\mathbf{a}}_i := \mathbf{x}$ *for all* $i \in \mathscr{N}$*, is an equilibrium.*

Proof. Parts (i) and (ii) are standard results in the literature of discounted stochastic games (cf. Fink 1964). We provide the proof for part (iii).

($\Rightarrow$) Suppose there exists an equilibrium in $A_{\mathbf{x}}$, denoted by $\tilde{\mathbf{a}}$. We show that strategy profile $\bar{\mathbf{a}}$ is an equilibrium as well. Since $\bar{\mathbf{a}}, \tilde{\mathbf{a}} \in A_{\mathbf{x}}$, $\mathbf{w}^{\tilde{\mathbf{a}}} = \mathbf{w}^{\bar{\mathbf{a}}}$ by part (i). There are two cases: If $x(s) = 1$, then for all $i \in \mathscr{N}$, $\tilde{a}_i(s) = \bar{a}_i(s) = 1$, and

$$\begin{aligned} w_i^{\bar{\mathbf{a}}}(s) = w_i^{\tilde{\mathbf{a}}}(s) &= \max\left\{ \left(\prod_{\substack{j\in\mathscr{N},\\ j\neq i}} \tilde{a}_j(s) \right) u_i(s,1) + \left(1 - \prod_{\substack{j\in\mathscr{N},\\ j\neq i}} \tilde{a}_j(s) \right) F_i(s, \mathbf{w}_i^{\tilde{\mathbf{a}}}), F_i(s, \mathbf{w}_i^{\tilde{\mathbf{a}}}) \right\} \\ &= \max\left\{ \left(\prod_{\substack{j\in\mathscr{N},\\ j\neq i}} \bar{a}_j(s) \right) u_i(s,1) + \left(1 - \prod_{\substack{j\in\mathscr{N},\\ j\neq i}} \bar{a}_j(s) \right) F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}}), F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}}) \right\}, \end{aligned}$$

where the second equality follows from the fact that $\tilde{\mathbf{a}}$ is an equilibrium and part (ii).

If $x(s) = 0$, then for all $i \in \mathscr{N}$, $\bar{a}_i(s) = 0$, and

$$w_i^{\bar{\mathbf{a}}}(s) = F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}}) = \max\left\{ \left(\prod_{\substack{j\in\mathscr{N},\\ j\neq i}} \bar{a}_j(s) \right) u_i(s,1) + \left(1 - \prod_{\substack{j\in\mathscr{N},\\ j\neq i}} \bar{a}_j(s) \right) F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}}), F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}}) \right\},$$

where the first equality follows from part (i).

In summary, the equation (3.2) is satisfied for all $s \in \mathscr{S}, i \in \mathscr{N}$. Therefore, $\bar{\mathbf{a}}$ is an equilibrium by part (ii).

($\Leftarrow$) Follows directly from the definitions of $\bar{\mathbf{a}}$ and $A_{\mathbf{x}}$. $\square$

For each $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$, Proposition 1 (i) implies that all strategy profiles in $A_{\mathbf{x}}$ have the same payoff profile. In other words, $\mathbf{x}$ is necessary and sufficient information for characterizing the payoff profile of a strategy profile $\mathbf{a}$. As we are interested in studying the set of equilibrium payoff profiles, by Proposition 1 (iii), it is sufficient to only focus on the set of strategy profiles in which $\mathbf{a}_i = \mathbf{x}$ for all $i \in \mathscr{N}$. We define such a set of strategy profiles as the set of *unanimous* strategy profiles, since all players take the same action at each state $s \in \mathscr{S}$. Hereafter, we restrict our attention to the set of unanimous strategy profiles, and with a slight abuse of notation, $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$ represents a unanimous strategy profile. Accordingly, we define $w_i^{\mathbf{x}}(s)$, $\mathbf{w}_i^{\mathbf{x}} := \langle w_i^{\mathbf{x}}(s) \rangle_{s\in\mathscr{S}}$, and $\mathbf{w}^{\mathbf{x}} := \langle \mathbf{w}_i^{\mathbf{x}} \rangle_{i\in\mathscr{N}}$ for unanimous strategy profile $\mathbf{x}$. We may restate Proposition 1 (i) and (ii) for unanimous strategy profiles as follows. (For convenience we drop the word unanimous hereafter.)

Proposition 2. *Given a strategy profile* $\mathbf{x}$ *with associated payoff profile* $\mathbf{w}^{\mathbf{x}}$*:*
(i) *For all* $s \in \mathscr{S}, i \in \mathscr{N}$,

$$w_i^{\mathbf{x}}(s) = x(s)u_i(s,1) + (1 - x(s))F_i(s, \mathbf{w}_i^{\mathbf{x}}). \tag{3.3}$$

Furthermore, $\mathbf{w}^{\mathbf{x}}$ *is the unique solution for this set of equations.*
(ii) $\mathbf{x}$ *is an equilibrium if and only if for all* $s \in \mathscr{S}, i \in \mathscr{N}$*:*

$$w_i^{\mathbf{x}}(s) = \max\left\{x(s)u_i(s,1) + (1 - x(s))F_i(s, \mathbf{w}_i^{\mathbf{x}}), F_i(s, \mathbf{w}_i^{\mathbf{x}})\right\}. \tag{3.4}$$

Proposition 2 follows immediately from Proposition 1, and the fact that strategy profile $\mathbf{x}$ is unanimous. Part (i) describes how to calculate the payoff profile of a strategy profile, and Part (ii) describes the Bellman-Shapley equations for $\mathcal{G}$. Let $\mathbf{0}$ be the strategy profile in which $x(s) = 0$ for all $s \in \mathscr{S}$.

Remark 2. *An attractive property of Proposition 2* (ii) *is that when we assess equilibrium conditions for a strategy profile* $\mathbf{x}$*, we only need to check if the Bellman-Shapley equation* (3.4) *is satisfied for each* $s \in \mathscr{S}, i \in \mathscr{N}$ *in which* $x(s) = 1$ *since* (3.4) *is trivially satisfied for each* $s \in \mathscr{S}, i \in \mathscr{N}$ *in which* $x(s) = 0$*. In particular, Kurt (2012) formally shows that the strategy profile* $\mathbf{0}$ *is an equilibrium. We prove a more general result, namely, that the set of equilibria forms an independence system, in Section 3.5.*

In this chapter, we study the problem of finding a best equilibrium, with respect to a given linear objective function of payoffs, for $\mathcal{G}$. Such an objective function is well accepted in the literature of group decision analysis (cf. Eliashberg and Winkler 1981 and the references therein). Let $\Psi := \{\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|} | \, \mathbf{x} \text{ is an equilibrium of } \mathcal{G}\}$, $c_i(s) \in \mathbb{R}$ be an objective function coefficient for all $s \in \mathscr{S}, i \in \mathscr{N}$, and $\mathbf{c} := \langle c_i(s) \rangle_{s \in \mathscr{S}, i \in \mathscr{N}}$. Therefore, the problem of finding a best equilibrium is:

$$(\mathbf{P}): \ \max\{\mathbf{c}^\top \mathbf{w}^{\mathbf{x}} | \, \mathbf{x} \in \Psi\}. \tag{3.5}$$

Note that by repeated application of equation (3.3), $\mathbf{w}^{\mathbf{x}}$ can be represented as a (highly) nonlinear function of $\mathbf{x}$. Therefore, $\mathbf{c}^\top \mathbf{w}^{\mathbf{x}}$ is a nonlinear function of $\mathbf{x}$ as well. We present an MILP formulation for $(\mathbf{P})$ in Section 3.6.

3.4 COMPUTATIONAL COMPLEXITY

In one-stage games, the computational complexity of finding a best pure Nash equilibrium is NP-hard (cf. Gairing et al. 2004, Sperber 2010). In any stationary repeated game, each pure stationary equilibrium is identical to a pure Nash equilibrium of a one-stage game, and vice versa. Hence, the problem of finding a best pure stationary equilibrium is NP-hard in repeated games, which in turn implies that the problem is NP-hard in stochastic games. In this section, we establish the computational complexity of the problem for consensus stopping games. For further discussion on computational complexity issues related to stochastic games, see Conitzer and Sandholm (2008).

To establish the computational complexity of the problem of finding a best equilibrium of $\mathcal{G}$, it is sufficient to establish the computational complexity of its associated decision problem, which is as follows: Given an instance of $\mathcal{G}$, $h \in \mathbb{R}$, and $\mathbf{c} \in \mathbb{R}^{|\mathscr{S}| \times N}$, does there exist an equilibrium $\mathbf{x}$ with associated payoff profile $\mathbf{w}^{\mathbf{x}}$ such that $\mathbf{c}^\top \mathbf{w}^{\mathbf{x}} \geq h$?

Proposition 3. *For any fixed number of players ($N \geq 2$), the decision version of finding a best equilibrium of $\mathcal{G}$ is NP-complete.*

Proof. The problem is in NP since given, a strategy profile $\mathbf{x}$, the question of whether $\mathbf{x}$ is an equilibrium with an objective function value of at least h can be verified in polynomial time by Proposition 2.

We provide a proof by a reduction from Knapsack, a well-known NP-complete problem (Karp 1972). Consider strictly positive numbers $a_1, a_2, \ldots, a_n$, b, $c_1, c_2, \ldots, c_n$, k, we need to check whether there exist binary values $y_1, y_2, \ldots, y_n$ satisfying:

$$\sum_{s=1}^{n} c_s y_s \geq k$$
$$\sum_{s=1}^{n} a_s y_s \leq b.$$

Let $N \geq 2$ be an arbitrary integer. We construct (in polynomial time) an instance of $\mathcal{G}$, as follows:

- $\mathscr{N} = \{1, 2, \ldots, N\}$,

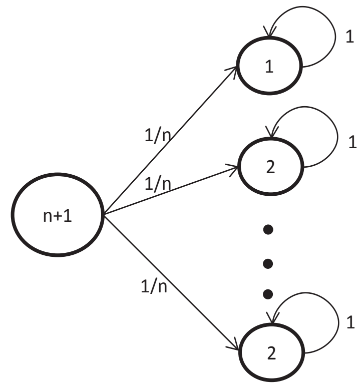

Figure 3: The transition probabilities structure for the instance of $\mathcal{G}$ constructed in the proof of Proposition 3.

- $\mathscr{S} = \{1, 2, \ldots, n, n+1\}$,
- $\lambda_i = 0.5$ for all $i \in \mathscr{N}$,
- $u_i(s, 0) = 0$ for all $s \in \mathscr{S}, i \in \mathscr{N}$,
- $u_1(s, 1) = 2na_s, u_2(s, 1) = 0$ for all $s \in \{1, 2, \ldots, n\}$,
- $u_1(n+1, 1) = b, u_2(n+1, 1) = 1$,
- $u_i(s, 1) = 0$ for all $s \in \mathscr{S}, i \in \{3, \ldots, N\}$,
- $\mathscr{P}(s|n+1) = \frac{1}{n}$ for all $s \in \{1, 2, \ldots, n\}$,
- $\mathscr{P}(s|s) = 1$ for all $s \in \{1, 2, \ldots, n\}$,
- $c_1(s) = \frac{c_s}{2na_s}$ for all $s \in \{1, 2, \ldots, n\}$, and $c_1(n+1) = 0$,
- $c_2(s) = 0$ for all $s \in \{1, 2, \ldots, n\}$, and $c_2(n+1) = 1 + \sum_{s=1}^{n} c_s$,
- $c_i(s) = 0$ for all $i \in \{3, \ldots, N\}, s \in \mathscr{S}$.

We show that there exists an equilibrium with the objective function value of at least $h := 1 + \sum_{s=1}^{n} c_s + k$ for this instance of $\mathcal{G}$ if and only if the instance of Knapsack has a solution with the total value of at least k.

Under any strategy profile $\mathbf{x}$, for each state $s \in \{1, 2, \ldots, n\}$, if $x(s) = 1$, then $w_1(s) = u_1(s, 1) = 2na_s$ and $w_2(s) = u_2(s, 1) = 0$ by (3.3). Otherwise, if $x(s) = 0$, then

$$w_1^{\mathbf{x}}(s) = F_1(s, w_1^{\mathbf{x}}) = u_1(s, 0) + \lambda_1 w_1^{\mathbf{x}}(s) = 0 + \frac{1}{2} w_1^{\mathbf{x}}(s) \Rightarrow w_1^{\mathbf{x}}(s) = 0,$$
$$w_2^{\mathbf{x}}(s) = F_2(s, w_2^{\mathbf{x}}) = u_2(s, 0) + \lambda_2 w_2^{\mathbf{x}}(s) = 0 + \frac{1}{2} w_2^{\mathbf{x}}(s) \Rightarrow w_2^{\mathbf{x}}(s) = 0,$$

where we have made use of (3.3) and the definition of $F_i(s, \cdot)$ to write the above equalities. Consequently, $w_1^{\mathbf{x}}(s) = 2na_s x(s), w_2^{\mathbf{x}}(s) = 0$, and it can easily be verified that the Bellman-Shapley equation (3.4) is satisfied for each $i \in \{1, 2\}, s \in \{1, 2, \ldots, n\}$ under an arbitrary strategy profile $\mathbf{x}$. For state $n+1$ and player 1, $u_1(n+1, 1) = b$, $F_1(n+1, \mathbf{w}_1^{\mathbf{x}}) = 0.5 \sum_{s=1}^{n} \frac{1}{n} w_1^{\mathbf{x}}(s) = \sum_{s=1}^{n} a_s x(s)$ by the definition of $F_i(s, \cdot)$ and our previous arguments regarding computing $w_1^{\mathbf{x}}(s)$ for all $s \in \{1, \cdots, n\}$, so the Bellman-Shapley equation (3.4) is satisfied if and only if either $x(n+1) = 0$, or $x(n+1) = 1$ jointly with $\sum_{s=1}^{n} a_s x(s) \leq b$. For state $n+1$ and player 2, $u_2(n+1, 1) = 1$, $F_2(n+1, \mathbf{w}_2^{\mathbf{x}}) = 0.5 \sum_{s=1}^{n} \frac{1}{n} w_2^{\mathbf{x}}(s) = 0$ by the definition of $F_i(s, \cdot)$ and our previous arguments regarding computing $w_2^{\mathbf{x}}(s)$ for all $s \in \{1, \cdots, n\}$, so that the Bellman-Shapley equation (3.4) is always satisfied, and $w_2^{\mathbf{x}}(n+1) = x(n+1)$. For each $i \in \{3, \ldots, N\}, s \in \mathscr{S}$, it can similarly be verified that $w_i^{\mathbf{x}}(s) = u_i(s, 1) = F_i(s, \mathbf{w}_i^{\mathbf{x}}) = 0$, and hence the Bellman-Shapley equation (3.4) is satisfied. In summary, a strategy profile $\mathbf{x}$ is an equilibrium if and only if either $x(n+1) = 0$, or $x(n+1) = 1$ jointly with $\sum_{s=1}^{n} a_s x(s) \leq b$.

Observe that the objective function of the game is:

$$\begin{aligned}
\sum_{i=1}^{N} \sum_{s=1}^{n+1} c_i(s) w_i^{\mathbf{x}}(s) &= \sum_{i=1}^{2} \sum_{s=1}^{n+1} c_i(s) w_i^{\mathbf{x}}(s) \\
&= c_2(n+1) w_2^{\mathbf{x}}(n+1) + \sum_{s=1}^{n} c_1(s) w_1^{\mathbf{x}}(s) \\
&= c_2(n+1) x(n+1) + \sum_{s=1}^{n} c_1(s)(2na_s x(s)) \\
&= (1 + \sum_{s=1}^{n} c_s) x(n+1) + \sum_{s=1}^{n} c_s x(s).
\end{aligned}$$

For the nontrivial case $k > 0$, any equilibrium with the objective function value of at least h must have $x(n+1) = 1$ since h is larger than the objective function portion for all $s \in \{1, \ldots, n\}$. Let us couple each strategy profile $\langle x(s) \rangle_{s \in \{1,2,\ldots,n\}}, x(n+1) = 1$ with a binary vector $\langle y_s \rangle_{s\in\{1,2,\ldots,n\}}$ by the one-to-one mapping $x(s) = y_s$ for all $s \in \{1,\ldots,n\}$, and $x(n+1) = 1$. As noted earlier, a strategy profile $\langle x(s) \rangle_{s\in\{1,2,\ldots,n\}}, x(n+1) = 1$ is an equilibrium if and only if $\sum_{s=1}^{n} a_s x(s) \leq b$. Hence, a strategy profile $\langle x(s) \rangle_{s\in\{1,2,\ldots,n\}}, x(n+1) = 1$ is an equilibrium if and only if its associated binary vector $\langle y_s \rangle_{s\in\{1,2,\ldots,n\}}$ is a feasible solution for the instance of KNAPSACK. In addition, the objective function of the instance of $\mathcal{G}$ for a strategy profile $\langle x(s) \rangle_{s\in\{1,2,\ldots,n\}}, x(n+1) = 1$ is at least h if and only if the total value of the associated binary vector $\langle y_s \rangle_{s\in\{1,2,\ldots,n\}}$ for the instance of KNAPSACK is at least k. □

3.5 CHARACTERIZING EQUILIBRIA AND COMBINATORIAL VALID INEQUALITIES

In this section, we characterize the equilibria of $\mathcal{G}$ and develop two families of combinatorial valid inequalities for (**P**) to improve its representation. We first need to define several functions from the space of strategy profiles to the collection of all subsets of $\mathscr{S}$ as follows. Given a strategy profile $\mathbf{x}$, let:

$$\mathscr{S}^k(\mathbf{x}) = \{s \in \mathscr{S} \mid x(s) = k\} \qquad \forall k \in \{0,1\}, \tag{3.6a}$$

$$\mathscr{S}^k_{i,nv}(\mathbf{x}) = \{s \in \mathscr{S}^k(\mathbf{x}) \mid u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})\} \qquad \forall k \in \{0,1\}, i \in \mathscr{N}, \tag{3.6b}$$

$$\mathscr{S}^k_{i,v}(\mathbf{x}) = \{s \in \mathscr{S}^k(\mathbf{x}) \mid u_i(s,1) < F_i(s, \mathbf{w}_i^{\mathbf{x}})\} \qquad \forall k \in \{0,1\}, i \in \mathscr{N}. \tag{3.6c}$$

It can easily be seen that $\{\mathscr{S}^k_{i,v}(\mathbf{x}), \mathscr{S}^k_{i,nv}(\mathbf{x})\}$ is a partition of $\mathscr{S}^k(\mathbf{x})$ for all $i \in \mathscr{N}, k \in \{0,1\}$, and $\{\mathscr{S}^1(\mathbf{x}), \mathscr{S}^0(\mathbf{x})\}$ is a partition of $\mathscr{S}$ for each strategy profile $\mathbf{x}$. $\mathscr{S}^1(\mathbf{x}), \mathscr{S}^0(\mathbf{x})$ are the sets of *stopping* and *continuing* states under $\mathbf{x}$, respectively. $\mathscr{S}^1_{i,v}(\mathbf{x})$ represents the set of stopping states in which the Bellman-Shapley equation (3.4) is violated for player i under $\mathbf{x}$; conversely, $\mathscr{S}^1_{i,nv}(\mathbf{x})$ represents the set of stopping states in which the Bellman-Shapley equation (3.4) is satisfied for player i under $\mathbf{x}$. Although the Bellman-Shapley equation (3.4) is satisfied for states of $\mathscr{S}^0_{i,v}(\mathbf{x})$ for player i, it could have been violated if $x(s)$ had been 1 in those states. Hence, it may be thought of as the set of potentially equilibrium-violating states for player i under $\mathbf{x}$. The set $\mathscr{S}^0_{i,nv}(\mathbf{x})$ can similarly be interpreted for player i under $\mathbf{x}$.

Specifically, it represents the set of continuing states in which the Bellman-Shapley equation (3.4) is satisfied even if $x(s)$ had been 1, so that they do not violate the Bellman-Shapley equation (3.4) even with small perturbations in the strategy profile.

Proposition 4. *Given a strategy profile $\bar{\mathbf{x}}$ with associated payoff profile $\mathbf{w}^{\bar{\mathbf{x}}}$:*

(i) For some strategy profile $\mathbf{x}$ with associated payoff profile $\mathbf{w}^{\mathbf{x}}$, suppose $\exists i \in \mathscr{N}$ such that $\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$. Then $w_i^{\mathbf{x}}(s) \geq w_i^{\bar{\mathbf{x}}}(s)$ for all $s \in \mathscr{S}$.

(ii) For some strategy profile $\mathbf{x}$ with associated payoff profile $\mathbf{w}^{\mathbf{x}}$, suppose $\exists i \in \mathscr{N}$ such that $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$. Then $w_i^{\mathbf{x}}(s) \leq w_i^{\bar{\mathbf{x}}}(s)$ for all $s \in \mathscr{S}$.

Proof. Recall that under a fixed strategy profile $\mathbf{x}$, the payoffs of each player represent a stationary Markov reward process, and hence can be calculated by value iteration (Denardo 1967). Let $[w_i^{\mathbf{x}}(s)]^n$ denote the value associated with state $s \in \mathscr{S}$ at the n^{th} iteration of the value iteration algorithm under strategy profile $\mathbf{x}$ for player i. Furthermore, we initialize our value iteration with payoffs of player i under $\bar{\mathbf{x}}$, i.e., $[w_i^{\mathbf{x}}(s)]^0 = w_i^{\bar{\mathbf{x}}}(s)$ for all $s \in \mathscr{S}$.

(i) From Proposition 2 (i) and the hypothesis about the relation between $\mathbf{x}$ and $\bar{\mathbf{x}}$, there are four cases:

If $s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$, then $x(s) = 1$ and $[w_i^{\mathbf{x}}(s)]^1 = u_i(s,1) = w_i^{\bar{\mathbf{x}}}(s) = [w_i^{\mathbf{x}}(s)]^0$.

If $s \in \mathscr{S}^0(\bar{\mathbf{x}})$, then $x(s) = 0$ and $[w_i^{\mathbf{x}}(s)]^1 = F_i(s, [\mathbf{w}_i^{\mathbf{x}}]^0) = F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) = w_i^{\bar{\mathbf{x}}}(s) = [w_i^{\mathbf{x}}(s)]^0$.

If $s \in \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$ and $x(s) = 1$, then $[w_i^{\mathbf{x}}(s)]^1 = u_i(s,1) = w_i^{\bar{\mathbf{x}}}(s) = [w_i^{\mathbf{x}}(s)]^0$.

If $s \in \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$ and $x(s) = 0$, then $[w_i^{\mathbf{x}}(s)]^1 = F_i(s, [\mathbf{w}_i^{\mathbf{x}}]^0) = F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) \geq u_i(s,1) = w_i^{\bar{\mathbf{x}}}(s) = [w_i^{\mathbf{x}}(s)]^0$ where the inequality follows from $s \in \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$.

From all four cases, it follows that $[w_i^{\mathbf{x}}(s)]^1 \geq [w_i^{\mathbf{x}}(s)]^0$ for all $s \in \mathscr{S}$. By the monotonicity of the dynamic programming operator induced by strategy profile $\mathbf{x}$ for player i (Blackwell 1965), it follows that for any n, $[w_i^{\mathbf{x}}(s)]^{n+1} \geq [w_i^{\mathbf{x}}(s)]^n$ for all $s \in \mathscr{S}$. As a result, $w_i^{\mathbf{x}}(s) = \lim_{n\to\infty} [w_i^{\mathbf{x}}(s)]^n \geq [w_i^{\mathbf{x}}(s)]^0 = w_i^{\bar{\mathbf{x}}}(s)$.

(ii) The proof is similar to that of part (i).

□

Remark 3. *In part (i) of Proposition 4, if $\mathscr{S}^1(\mathbf{x}) \subset \mathscr{S}^1(\bar{\mathbf{x}})$, then there exists some $\hat{s} \in \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$ for which $w_i^{\mathbf{x}}(\hat{s}) > w_i^{\bar{\mathbf{x}}}(\hat{s})$. In part (ii) of Proposition 4, if there exists some $\hat{s} \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$ such that $u_i(\hat{s},1) > F_i(\hat{s}, \mathbf{w}_i^{\bar{\mathbf{x}}})$ and $x(\hat{s}) = 0$, then $w_i^{\mathbf{x}}(\hat{s}) < w_i^{\bar{\mathbf{x}}}(\hat{s})$.*

Lemma 4. *A strategy profile* $\mathbf{x}$ *is an equilibrium if and only if* $\mathscr{S}^1_{i,v}(\mathbf{x}) = \emptyset$ *for all* $i \in \mathscr{N}$.

Proof. ($\Rightarrow$) If $s \in \mathscr{S}^1(\mathbf{x})$, then $w_i^{\mathbf{x}}(s) = u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ by Proposition 2. Therefore, $u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ for all $s \in \mathscr{S}^1(\mathbf{x}), i \in \mathscr{N}$. It follows that $\mathscr{S}^1_{i,v}(\mathbf{x}) = \emptyset$ for all $i \in \mathscr{N}$.

($\Leftarrow$) $\mathscr{S}^1(\mathbf{x}) = \mathscr{S}^1_{i,nv}(\mathbf{x}) \cup \mathscr{S}^1_{i,v}(\mathbf{x})$ and $\mathscr{S}^1_{i,v}(\mathbf{x}) = \emptyset$ for all $i \in \mathscr{N}$. It follows that $\mathscr{S}^1(\mathbf{x}) = \mathscr{S}^1_{i,nv}(\mathbf{x})$ for all $i \in \mathscr{N}$. Therefore, if $s \in \mathscr{S}^1(\mathbf{x})$, then for all $i \in \mathscr{N}$, $u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ and Proposition 2 (i) implies that $w_i^{\mathbf{x}}(s) = u_i(s,1)$. Hence, the Bellman-Shapley equation (3.4) is satisfied for all $s \in \mathscr{S}^1(\mathbf{x}), i \in \mathscr{N}$. As noted in Remark 2, the Bellman-Shapley equation (3.4) is trivially satisfied for all $s \in \mathscr{S}^0(\mathbf{x}), i \in \mathscr{N}$. Consequently, $\mathbf{x}$ is an equilibrium by Proposition 2 (ii).

$\square$

Proposition 5. *Suppose a strategy profile* $\bar{\mathbf{x}}$ *is an equilibrium. For any strategy profile* $\mathbf{x}$ *where* $\mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$*, the following hold:*

(i) *For all* $s \in \mathscr{S}, i \in \mathscr{N}$, $w_i^{\mathbf{x}}(s) \leq w_i^{\bar{\mathbf{x}}}(s)$.

(ii) $\mathbf{x}$ *is an equilibrium.*

Proof. (i) As $\bar{\mathbf{x}}$ is an equilibrium, $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) = \emptyset$ for all $i \in \mathscr{N}$ by Lemma 4, and the hypothesis states that $\mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$. Therefore, $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$ for all $i \in \mathscr{N}$. The result follows from Proposition 4 (ii).

(ii) If $s \in \mathscr{S}^1(\mathbf{x})$, then $s \in \mathscr{S}^1(\bar{\mathbf{x}}) = \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) = \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$ for all $i \in \mathscr{N}$ since $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) = \emptyset$ for all $i \in \mathscr{N}$ by Lemma 4. This implies that $u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}})$ for all $s \in \mathscr{S}^1(\mathbf{x}), i \in \mathscr{N}$. Moreover, $w_i^{\bar{\mathbf{x}}}(s) \geq w_i^{\mathbf{x}}(s)$ for all $s \in \mathscr{S}, i \in \mathscr{N}$ by part (i), and hence $F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ for all $s \in \mathscr{S}, i \in \mathscr{N}$. Therefore, $u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ for all $s \in \mathscr{S}^1(\mathbf{x}), i \in \mathscr{N}$. So, $\mathscr{S}^1_{i,v}(\mathbf{x}) = \emptyset$ for all $i \in \mathscr{N}$. It follows from Lemma 4 that $\mathbf{x}$ is an equilibrium.

$\square$

The next corollary is an immediate consequence of Proposition 5 (ii).

Corollary 1. *If a strategy profile* $\bar{\mathbf{x}}$ *is not an equilibrium, then any strategy profile* $\mathbf{x}$ *where* $\mathscr{S}^1(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$ *is not an equilibrium.*

A well-known combinatorial structure is an *independence system*. Let I be a set, and $\mathscr{J}$ be a set of subsets of I. The pair $(I, \mathscr{J})$ is an independence system if it satisfies two conditions: First, the empty set is in $\mathscr{J}$. Second, if $I_1 \subset I_2$ and $I_2 \in \mathscr{J}$, then $I_1 \in \mathscr{J}$ (cf. Nemhauser and Trotter Jr. 1974).

The next corollary immediately follows from Remark 2 and Proposition 5 (ii).

Corollary 2. *Let $\mathcal{J}$ be the collection of all $\mathscr{S}^1(\mathbf{x})$ such that $\mathbf{x}$ is an equilibrium. The pair $(\mathscr{S}, \mathcal{J})$ is an independence system.*

Definition 1. *An equilibrium $\bar{\mathbf{x}}$ is* maximal *if there does not exist any equilibrium $\mathbf{x}$ such that $\mathscr{S}^1(\bar{\mathbf{x}}) \subset \mathscr{S}^1(\mathbf{x})$.*

Proposition 5 demonstrates that an equilibrium $\bar{\mathbf{x}}$ yields $2^{\sum_{s\in\mathscr{S}} \bar{x}(s)} - 1$ additional equilibria, and this property describes why $\mathcal{G}$ may possess many equilibria in general. However, by Proposition 5, all such equilibria are payoff-wise dominated, and if $c_i(s) \geq 0$ for all $s \in \mathscr{S}, i \in \mathscr{N}$, they can be eliminated from consideration when we search for a best equilibrium. Based on this dominance, we may restrict our search to maximal equilibria and develop the following optimality valid inequality for $(\mathbf{P})$ if the objective function coefficients are non-negative.

Proposition 6. (i) *If $c_i(s) \geq 0$ for all $s \in \mathscr{S}, i \in \mathscr{N}$, then there exists an optimal equilibrium that is maximal.*

(ii) *If a strategy profile $\bar{\mathbf{x}}$ is an equilibrium, then the inequality*

$$\sum_{s\in\mathscr{S}^1(\bar{\mathbf{x}})} [1 - x(s)] \leq |\mathscr{S}^1(\bar{\mathbf{x}})| \sum_{s\in\mathscr{S}^0(\bar{\mathbf{x}})} x(s) \tag{3.7}$$

is satisfied by every maximal equilibrium $\mathbf{x}$.

Proof. (i) Immediate from Remark 2, Proposition 5 (i), and $c_i(s) \geq 0$ for all $s \in \mathscr{S}, i \in \mathscr{N}$.

(ii) If $\sum_{s\in\mathscr{S}^0(\bar{\mathbf{x}})} x(s) > 0$, (3.7) is redundant. Otherwise, $\sum_{s\in\mathscr{S}^0(\bar{\mathbf{x}})} x(s) = 0$, implying $\mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$, so (3.7) cuts off an equilibrium $\mathbf{x}$ if and only if $\mathscr{S}^1(\mathbf{x}) \subset \mathscr{S}^1(\bar{\mathbf{x}})$. Moreover, if $\mathscr{S}^1(\mathbf{x}) \subset \mathscr{S}^1(\bar{\mathbf{x}})$, then $\mathbf{x}$ is a non-maximal equilibrium since $\bar{\mathbf{x}}$ is an equilibrium. □

Proposition 7. *Given a strategy profile $\bar{\mathbf{x}}$, any strategy profile $\mathbf{x}$ where $\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$ and $\mathscr{S}^1(\mathbf{x}) \cap (\mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})) \neq \emptyset$ for some $i \in \mathscr{N}$, is not an equilibrium.*

Proof. Let $\Xi_{\bar{\mathbf{x}}}$ be the set of strategy profiles $\tilde{\mathbf{x}}$ for which there exists an state $\hat{s} \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$ such that:

- $\tilde{x}(s) = 1$ for all $s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$,
- $\tilde{x}(s) = 0$ for all $s \in \mathscr{S}^0_{i,nv}(\bar{\mathbf{x}})$,
- $\tilde{x}(\hat{s}) = 1$ for some $\hat{s} \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$,
- $\tilde{x}(s) = 0$ for all $s \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})/\{\hat{s}\}$.

We show that each $\tilde{\mathbf{x}} \in \Xi_{\bar{\mathbf{x}}}$ is not an equilibrium. Since $\hat{s} \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$, there are two non-overlapping cases on where $\hat{s}$ belongs to:

Case 1) If $\hat{s} \in \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$, then $w_i^{\tilde{\mathbf{x}}}(s) \geq w_i^{\bar{\mathbf{x}}}(s)$ for all $s \in \mathscr{S}$ by Proposition 4 (i). As a result, $F_i(s, \mathbf{w}_i^{\tilde{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}})$ for all $s \in \mathscr{S}$. In particular, $F_i(\hat{s}, \mathbf{w}_i^{\tilde{\mathbf{x}}}) \geq F_i(\hat{s}, \mathbf{w}_i^{\bar{\mathbf{x}}}) > u_i(\hat{s}, 1)$. Since $F_i(\hat{\mathbf{s}}, \mathbf{w}_i^{\tilde{\mathbf{x}}}) > u_i(\hat{s}, 1)$ and $\tilde{x}(\hat{s}) = 1$, the Bellman-Shapley equation (3.4) is violated in state $\hat{s}$ under $\tilde{\mathbf{x}}$, so it is not an equilibrium.

Case 2) If $\hat{s} \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}})$, then consider strategy profile $\breve{\mathbf{x}}$ defined as follows:

- $\breve{x}(s) = 1$ for all $s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$,
- $\breve{x}(s) = 0$ for all $s \in \mathscr{S}/\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$.

Therefore, $\tilde{\mathbf{x}}$ and $\breve{\mathbf{x}}$ take the same value for all $s \in \mathscr{S}/\{\hat{s}\}$. By Corollary 1, if the strategy profile $\breve{\mathbf{x}}$ is not an equilibrium, $\tilde{\mathbf{x}}$ cannot be an equilibrium since $\mathscr{S}^1(\breve{\mathbf{x}}) \subseteq \mathscr{S}^1(\tilde{\mathbf{x}})$.

Suppose $\breve{\mathbf{x}}$ is an equilibrium. By Proposition 4 (i), $w_i^{\breve{\mathbf{x}}}(s) \geq w_i^{\bar{\mathbf{x}}}(s)$ for all $s \in \mathscr{S}$. Therefore, $F_i(s, \mathbf{w}_i^{\breve{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) > u_i(s, 1)$ for all $s \in \mathscr{S}^0_{i,v}(\bar{x})$. In particular, $F_i(\hat{s}, \mathbf{w}_i^{\breve{\mathbf{x}}}) > u_i(\hat{s}, 1)$. Moreover, suppose that $\tilde{\mathbf{x}}$ is an equilibrium; thus $w_i^{\tilde{\mathbf{x}}}(s) \geq w_i^{\breve{\mathbf{x}}}(s)$ for all $s \in \mathscr{S}$ by Proposition 5 (i). As a result, $F_i(s, \mathbf{w}_i^{\tilde{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\breve{\mathbf{x}}})$ for all $s \in \mathscr{S}$. In particular, $F_i(\hat{s}, \mathbf{w}_i^{\tilde{\mathbf{x}}}) \geq F_i(\hat{s}, \mathbf{w}_i^{\breve{\mathbf{x}}}) > u_i(\hat{s}, 1)$, and this means that the Bellman-Shapley equation (3.4) is violated in state $\hat{s}$ under $\tilde{\mathbf{x}}$. Therefore, $\tilde{\mathbf{x}}$ is not an equilibrium, which is a contradiction.

So far, we have shown that each $\tilde{\mathbf{x}} \in \Xi_{\bar{\mathbf{x}}}$ is not an equilibrium. For any strategy profile $\mathbf{x}$, satisfying the conditions of Proposition 7, there exists a strategy profile $\tilde{\mathbf{x}} \in \Xi_{\bar{\mathbf{x}}}$ such that $\mathscr{S}^1(\tilde{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$, so $\mathbf{x}$ cannot be an equilibrium by Corollary 1. $\square$

An interesting feature of Proposition 7 is that it can provide insights about equilibria irrespective of whether strategy profile $\bar{\mathbf{x}}$ is an equilibrium or not.

Proposition 8. *Given a strategy profile $\bar{\mathbf{x}}$, the following inequalities are valid for Ψ.*

$$\sum_{s \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})} x(s) \leq \left(|\mathscr{S}^0_{i,v}(\bar{\mathbf{x}})| + |\mathscr{S}^1_{i,v}(\bar{\mathbf{x}})|\right) \sum_{s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})} [1 - x(s)] \qquad \forall i \in \mathscr{N}. \tag{3.8}$$

Proof. If $\sum_{s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})} [1 - x(s)] > 0$, then (3.8) is redundant. Now, consider the other case in which $\sum_{s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})} [1 - x(s)] = 0$. This implies $\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$. If $\mathbf{x}$ is an equilibrium and $\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$, then $\mathscr{S}^1(\mathbf{x}) \cap (\mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})) = \emptyset$ by Proposition 7. This is equivalent to saying that if $\mathbf{x}$ is an equilibrium and $\sum_{s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})} [1 - x(s)] = 0$, then $x(s) = 0$ for all $s \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$. $\square$

The valid inequalities given in (3.8) remove the set of strategy profiles that are not equilibria by Proposition 7. In fact, Proposition 7 implies the following logical disjunction: either $x(s) = 1$ for all $s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$, $x(s) = 0$ for all $s \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cap \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$, or $x(s) = 0$ for some $s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$. Although there are other valid inequalities that may ensure the logical disjunction, by following arguments similar to those of Balas (1979), it can be shown that (3.8) is a best disjunctive valid inequality in the sense that the amount by which $\bar{\mathbf{x}}$ violates the valid inequality, is maximized by (3.8) among all disjunctive valid inequalities. Another advantage of Proposition 8 is that for each player $i \in \mathscr{N}$ where $\mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \neq \emptyset$, we can derive a nontrivial valid inequality.

We conclude this section by describing how our approach may be extended to establish structures of stationary equilibria in other classes of stochastic games. In Markov decision processes (MDPs), value iteration is often adopted to establish structural properties of the optimal value function and the optimal stationary policy. Apart from insights behind the derived properties, they are sometimes applied to develop a method to compute the optimal value function, especially in approximate dynamic programming (cf. Powell 2007). Although stochastic games may be viewed as a multi-player generalization of MDPs, there is no analogue of value iteration for discounted stochastic games. Consequently, there have been only a few characterizations of stationary equilibria for stochastic games aside from the well-known Bellman-Shapley equations. An important consequence of Propositions 4, 5, and 7 is that in the cost of providing the payoff profile of a given strategy profile, we can characterize the payoff profile and equilibrium behavior of a set of strategy profiles. Our results are the first combinatorial characterizations for a class of stochastic games, and it is anticipated that this approach opens the door to analyze stationary equilibria of more complicated stochastic games.

3.6 EQUILIBRIUM SELECTION FORMULATION

In this section, we present an MILP formulation for $(\mathbf{P})$. Let coefficient $V_i(s)$ be an upper bound for equilibria payoffs of player i in state s. Kurt et al. (2011) suggest that for each player $i \in \mathscr{N}$, $\mathbf{V}_i := \langle V_i(s) \rangle_{s \in \mathscr{S}}$ may be calculated as a solution of the MDP equations $V_i(s) = \max\{u_i(s,1), F_i(s, \mathbf{V}_i)\}$ for all $s \in \mathscr{S}$, and show it is a valid upper bound as it represents optimal value function of player i in an MDP where he is maximizing his own payoffs when the autonomy of the other players is suppressed. By using this set of parameters, they propose an MILP to represent the set of equilibria of $\mathcal{G}$, relying on the assumption that $u_i(s,0), u_i(s,1) \geq 0$ for all $s \in \mathscr{S}, i \in \mathscr{N}$. Let $\mathbf{d} := \langle d_i(s) \rangle_{s \in \mathscr{S}, i \in \mathscr{N}}$ be the payoff profile of the strategy profile $\mathbf{0}$. In words, $d_i(s)$ represents the total expected discounted reward of player i starting from state s when he always decides to continue. We propose a similar formulation to represent the set of equilibria of $\mathcal{G}$ in the following:

$$
\begin{aligned}
& w_i(s) \geq F_i(s, \mathbf{w}_i) && \forall s \in \mathscr{S}, i \in \mathscr{N}, && (3.9a)\\
& w_i(s) \leq F_i(s, \mathbf{w}_i) + [u_i(s,1) - d_i(s)]x(s) && \forall s \in \mathscr{S}, i \in \mathscr{N}, && (3.9b)\\
& w_i(s) \geq [u_i(s,1) - d_i(s)]x(s) + d_i(s) && \forall s \in \mathscr{S}, i \in \mathscr{N}, && (3.9c)\\
& w_i(s) \leq u_i(s,1)x(s) + F_i(s, \mathbf{V}_i)[1 - x(s)] && \forall s \in \mathscr{S}, i \in \mathscr{N}, && (3.9d)\\
& w_i(s) \text{ unrestricted} && \forall s \in \mathscr{S}, i \in \mathscr{N}, && (3.9e)\\
& x(s) \in \{0,1\} && \forall s \in \mathscr{S}. && (3.9f)
\end{aligned}
$$

Let $\Delta := \{\langle \mathbf{x}, \mathbf{w} \rangle \,|(3.9a) - (3.9f)\}$.

Lemma 5. *(i)* $d_i(s) = F_i(s, \mathbf{d}_i)$ *for all* $s \in \mathscr{S}, i \in \mathscr{N}$.

(ii) If a strategy profile $\mathbf{x}$ *is an equilibrium, then* $w_i^{\mathbf{x}}(s) \geq d_i(s)$ *for all* $s \in \mathscr{S}, i \in \mathscr{N}$.

(iii) If $u_i(\bar{s}, 1) < d_i(\bar{s})$ *for some* $\bar{s} \in \mathscr{S}, i \in \mathscr{N}$*, then* $x(\bar{s}) = 0$ *for each equilibrium* $\mathbf{x}$.

Proof. (i) Immediate from Proposition 2 (i).

(ii) Note that $\mathscr{S}^1(\mathbf{0}) = \emptyset \subseteq \mathscr{S}^1(\mathbf{x})$ for each equilibrium $\mathbf{x}$. The result follows from Proposition 5 (i).

(iii) Substituting $\mathbf{0}$ for $\bar{\mathbf{x}}$ in Proposition 8 implies that the following set of inequalities are valid for Ψ.

$$\sum_{s \in \mathscr{S}^0_{i,v}(\mathbf{0}) \cup \mathscr{S}^1_{i,v}(\mathbf{0})} x(s) \leq (|\mathscr{S}^0_{i,v}(\mathbf{0})| + |\mathscr{S}^1_{i,v}(\mathbf{0})|) \sum_{s \in \mathscr{S}^1_{i,nv}(\mathbf{0})} [1 - x(s)] = 0 \qquad \forall i \in \mathscr{N},$$

where we have made use of the fact that $\mathscr{S}^1_{i,nv}(\mathbf{0}) \subseteq \mathscr{S}^1(\mathbf{0}) = \emptyset$, to write the equality. Note that $\bar{s} \in \mathscr{S}^0_{i,v}(\mathbf{0})$ since $u_i(\bar{s}, 1) < d_i(\bar{s})$. By the above set of inequalities, $x(\bar{s}) = 0$ is valid for Ψ, and hence $x(\bar{s}) = 0$ for each equilibrium $\mathbf{x}$. □

Proposition 9. *(i)* $\langle \mathbf{x}, \mathbf{w} \rangle \in \Delta$ *if and only if strategy profile* $\mathbf{x}$ *is an equilibrium with associated payoff profile* $\mathbf{w}$.

(ii) If $u_i(s, 0), u_i(s, 1) \geq 0$ *for all* $s \in \mathscr{S}, i \in \mathscr{N}$*, then* Δ *is at least as strong as the formulation of Kurt et al. (2011).*

Proof. (i) $(\Rightarrow)$ $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$ by (3.9f). For each $s \in \mathscr{S}$, there are two cases:
If $x(s) = 0$, $w_i(s) = F_i(s, \mathbf{w}_i)$ for all $i \in \mathscr{N}$ by (3.9a) – (3.9b).
If $x(s) = 1$, $w_i(s) = u_i(s, 1)$ for all $i \in \mathscr{N}$ by (3.9c) – (3.9d).
As a result, $\mathbf{w}$ is associated payoff profile of $\mathbf{x}$ by Proposition 2 (i). It also follows from (3.9a) that the Bellman-Shapley equation (3.4) is satisfied for all $s \in \mathscr{S}^1(\mathbf{x})$. Therefore, $\mathbf{x}$ is an equilibrium by Proposition 2 (ii).

$(\Leftarrow)$ Suppose $\mathbf{x}$ is an equilibrium with associated payoff profile $\mathbf{w}$. Constraint (3.9f) is satisfied since $\mathbf{x}$ is a (pure) strategy profile. Constraint (3.9a) is satisfied by Proposition 2 (ii). For each $s \in \mathscr{S}$, there are two cases:

If $x(s) = 0$, then $w_i(s) = F_i(s, \mathbf{w}_i)$ by Proposition 2 (i). Constraint (3.9b) is obviously satisfied, and (3.9c) is satisfied by Lemma 5 (ii). Constraint (3.9d) is satisfied since:

$$w_i(s) = F_i(s, \mathbf{w}_i) \leq F_i(s, \mathbf{V}_i),$$

where the inequality follows from the definition of $\mathbf{V}_i$.

If $x(s) = 1$, then $w_i(s) = u_i(s, 1)$ by Proposition 2 (i). Constraint (3.9b) is satisfied since:

$$w_i(s) = u_i(s, 1) \leq u_i(s, 1) + F_i(s, \mathbf{w}_i) - F_i(s, \mathbf{d}_i) = u_i(s, 1) + F_i(s, \mathbf{w}_i) - d_i(s),$$

where the first inequality follows from Lemma 5 (ii), and the second equality follows from Lemma 5 (i). Constraints (3.9c) $-$ (3.9d) are obviously satisfied.

(ii) If $u_i(s, 0), u_i(s, 1) \geq 0$ for all $s \in \mathscr{S}, i \in \mathscr{N}$, it can easily be shown by value iteration that $d_i(s) \geq 0$ for all $s \in \mathscr{S}, i \in \mathscr{N}$. The rest of the proof is straightforward.

$\square$

Proposition 9 (i) implies that Ψ is the projection of Δ onto the $\mathbf{x}$-space and $(\mathbf{P})$ may be reformulated as the following MILP:

$$\max \sum_{s \in \mathscr{S}, i \in \mathscr{N}} c_i(s) w_i(s) \tag{3.10a}$$

$$s.t. \quad \langle \mathbf{x}, \mathbf{w} \rangle \in \Delta. \tag{3.10b}$$

3.7 ALGORITHMIC APPROACH

In this section we develop an algorithm to solve ($\mathbf{P}$) efficiently. Problem ($\mathbf{P}$) has a so-called "block-ladder" structure where $\mathbf{x}$ are the linking variables, and is amenable to Benders' decomposition (Benders 1962). However, our computational experiments in Tables 7 – 9 show that even a state-of-the-art implementation of Benders' decomposition (Fischetti et al. 2010) is ineffective. In this section, our goal is to develop a cutting plane approach for solving ($\mathbf{P}$). Let θ_i be an artificial variable that approximates $\sum_{s\in\mathscr{S}} c_i(s)w_i(s)$ for any $i \in \mathscr{N}$, and let $\boldsymbol{\theta} := \langle\theta_i\rangle_{i\in\mathscr{N}}$. Furthermore, $RMP, LB, UB, \mathbf{x}_{incum}$, and ϵ are the restricted master problem, lower bound, upper bound, the incumbent solution, and a termination tolerance, respectively.

Decomposition Algorithm

0. Initialization: Let $LB := -\infty, UB := +\infty, \mathbf{x}_{incum} := \emptyset$ and RMP be as follows:

$$\max \sum_{i\in\mathscr{N}} \theta_i \tag{3.11a}$$

$$s.t. \quad \theta_i \le \sum_{s\in\mathscr{S}} c_i(s)V_i(s) \qquad \forall i \in \mathscr{N}, \tag{3.11b}$$

$$\theta_i \text{ unrestricted} \qquad \forall i \in \mathscr{N}, \tag{3.11c}$$

$$x(s) \in \{0,1\} \qquad \forall s \in \mathscr{S}. \tag{3.11d}$$

1. Restricted Master: Solve RMP and obtain an optimal solution $(\bar{\mathbf{x}}, \bar{\boldsymbol{\theta}})$, and let $UB := \sum_{i\in\mathscr{N}} \bar{\theta}_i$.

2. Separation:

a. Calculate payoff profile $\mathbf{w}^{\bar{\mathbf{x}}}$ associated with $\bar{\mathbf{x}}$ using Proposition 2 (i), and characterize the sets $\mathscr{S}^0_{i,v}(\bar{\mathbf{x}}), \mathscr{S}^0_{i,nv}(\bar{\mathbf{x}}), \mathscr{S}^1_{i,v}(\bar{\mathbf{x}}), \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$ for all $i \in \mathscr{N}$.

b. For any i for which $\mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \neq \emptyset$, add the valid inequality (3.8) to RMP.

c. If $\bar{\mathbf{x}}$ is an equilibrium (i.e., $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) = \emptyset$ for all $i \in \mathscr{N}$) and $\bar{\theta}_i > \sum_{s\in\mathscr{S}} c_i(s)w_i^{\bar{\mathbf{x}}}(s)$ for some $i \in \mathscr{N}$, then 1. Add a Benders' optimality cut (Benders 1962) to RMP for $i \in \mathscr{N}$ for which $\bar{\theta}_i > \sum_{s\in\mathscr{S}} c_i(s)w_i^{\bar{\mathbf{x}}}(s)$; 2. Add the valid inequality (3.7) to RMP if $c_i(s) \ge 0$ for all $s \in \mathscr{S}, i \in \mathscr{N}$; 3. If $\mathbf{c}^\top\mathbf{w}^{\bar{\mathbf{x}}} \ge LB$, let $\mathbf{x}_{incum} := \bar{\mathbf{x}}, LB := \mathbf{c}^\top\mathbf{w}^{\bar{\mathbf{x}}}$.

3. Termination: If $UB - LB \leq \epsilon$, then terminate the algorithm. Otherwise, go to Step 1.

If $\bar{\mathbf{x}}$ is not an equilibrium, there exists some $i \in \mathscr{N}$ such that $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \neq \emptyset$, and hence (3.8) cuts off $\bar{\mathbf{x}}$ in Step (b). Hence, the Decomposition Algorithm finitely converges for any $\epsilon \in \mathbb{R}_+$. There are several differences between our proposed algorithm and Benders' decomposition: First, we solve the set of equations (3.3) as the subproblem. Second, we use a combinatorial feasibility cut. Third, we use a combinatorial optimality cut in addition to a Benders' optimality cut.

In order to calculate $\mathbf{w}_i^{\bar{\mathbf{x}}}$ in Step (a) for each $i \in \mathscr{N}$, by Proposition 2 (i) we solve the following linear program (LP):

$$
\begin{aligned}
\mathcal{R}_i(\bar{\mathbf{x}}) : \max \quad & \sum_{s \in \mathscr{S}} c_i(s) w_i(s) \\
s.t. \quad & w_i(s) = F_i(s, \mathbf{w}_i) && \forall s \in \mathscr{S}^0(\bar{\mathbf{x}}), \\
& w_i(s) = u_i(s, 1) && \forall s \in \mathscr{S}^1(\bar{\mathbf{x}}), \\
& w_i(s) \text{ unrestricted} && \forall s \in \mathscr{S}.
\end{aligned}
$$

For each $i \in \mathscr{N}$, let $\gamma_i(s)$ be optimal dual multipliers of $\mathcal{R}_i(\bar{\mathbf{x}})$ for all $s \in \mathscr{S}$ and $y^+ := \max\{0, y\}$ for any $y \in \mathbb{R}$. To enhance the Decomposition Algorithm, we seek a Pareto-optimal Benders' optimality cut (Magnanti and Wong 1981).

Proposition 10. *The following Benders' optimality cut is Pareto-optimal.*

$$
\begin{aligned}
\theta_i \leq & \sum_{s \in \mathscr{S}^0(\bar{\mathbf{x}})} \gamma_i(s) \left(u_i(s, 0) + [u_i(s, 1) - d_i(s)]^+ x(s) \right) + \\
& \sum_{s \in \mathscr{S}^1(\bar{\mathbf{x}})} \gamma_i(s) \ (u_i(s, 1) x(s) + F_i(s, \mathbf{V}_i)[1 - x(s)]) \qquad \forall i \in \mathscr{N}. \qquad (3.12)
\end{aligned}
$$

Proof. Let $\mathcal{P}_i(\bar{\mathbf{x}})$ denote the subproblem associated with each player i in (3.10a) – (3.10b) when $\mathbf{x}$ is set to $\bar{\mathbf{x}}$:

$$
\begin{aligned}
\mathcal{P}_i(\bar{\mathbf{x}}) : \max \sum_{s\in\mathscr{S}} & c_i(s)w_i(s) \\
s.t. \quad & w_i(s) \geq F_i(s, \mathbf{w}_i) && \forall s \in \mathscr{S}, \\
& w_i(s) \leq F_i(s, \mathbf{w}_i) + [u_i(s,1) - d_i(s)]\bar{x}(s) && \forall s \in \mathscr{S}, \\
& w_i(s) \geq [u_i(s,1) - d_i(s)]\bar{x}(s) + d_i(s) && \forall s \in \mathscr{S}, \\
& w_i(s) \leq u_i(s,1)\bar{x}(s) + F_i(s, \mathbf{V}_i)[1 - \bar{x}(s)] && \forall s \in \mathscr{S}, \\
& w_i(s) \text{ unrestricted} && \forall s \in \mathscr{S}.
\end{aligned}
$$

In order to generate a Benders' optimality cut in Step (c), optimal dual multipliers of $\mathcal{P}_i(\bar{\mathbf{x}})$ are needed while we only know optimal dual multipliers of $\mathcal{R}_i(\bar{\mathbf{x}})$. In fact, optimal dual multipliers of $\mathcal{R}_i(\bar{\mathbf{x}})$ and $\mathcal{P}_i(\bar{\mathbf{x}})$ are closely related. Let $\pi_{i,1}(s), \pi_{i,2}(s), \pi_{i,3}(s), \pi_{i,4}(s)$ be optimal dual multipliers of $\mathcal{P}_i(\bar{\mathbf{x}})$ associated with equations (3.9a) – (3.9d), for all $s \in \mathscr{S}, i \in \mathscr{N}$, respectively. It can easily be seen that for all $i \in \mathscr{N}$:

$$\pi_{i,1}(s) = \pi_{i,2}(s) = 0 \qquad \forall s \in \mathscr{S}^1(\bar{\mathbf{x}}), \tag{3.13a}$$

$$\pi_{i,1}(s) + \pi_{i,2}(s) = \gamma_i(s) \qquad \forall s \in \mathscr{S}^0(\bar{\mathbf{x}}), \tag{3.13b}$$

$$\pi_{i,3}(s) = \pi_{i,4}(s) = 0 \qquad \forall s \in \mathscr{S}^0(\bar{\mathbf{x}}), \tag{3.13c}$$

$$\pi_{i,3}(s) + \pi_{i,4}(s) = \gamma_i(s) \qquad \forall s \in \mathscr{S}^1(\bar{\mathbf{x}}), \tag{3.13d}$$

$$\pi_{i,1}(s) \leq 0, \pi_{i,2}(s) \geq 0, \pi_{i,3}(s) \leq 0, \pi_{i,4}(s) \geq 0 \qquad \forall s \in \mathscr{S}. \tag{3.13e}$$

At each iteration of the Decomposition Algorithm, $\mathcal{R}_i(\bar{\mathbf{x}})$ is solved, and $\gamma_i(s)$ is obtained for all $s \in \mathscr{S}, i \in \mathscr{N}$. There are multiple dual optimal solutions for $\mathcal{P}_i(\bar{\mathbf{x}})$, and we may use any $\pi_{i,1}(s), \pi_{i,2}(s), \pi_{i,3}(s), \pi_{i,4}(s)$, satisfying (3.13a) – (3.13e), to generate a Benders' optimality cut. A Benders' optimality cut is as follows:

$$
\begin{aligned}
\theta_i \leq \sum_{s\in\mathscr{S}} & \pi_{i,1}(s)\, u_i(s,0) + \pi_{i,2}(s) \Big(u_i(s,0) + [u_i(s,1) - d_i(s)]x(s) \Big) + \\
& \pi_{i,3}(s) \Big([u_i(s,1) - d_i(s)]x(s) + d_i(s) \Big) + \pi_{i,4}(s) \Big(u_i(s,1)x(s) + F_i(s, \mathbf{V}_i)[1 - x(s)] \Big) \quad \forall i \in \mathscr{N}.
\end{aligned} \tag{3.14}
$$

Since $\mathcal{R}_i(\bar{\mathbf{x}})$ is not degenerate, $\langle\gamma_i(s)\rangle_{s\in\mathscr{S}}$ is the unique dual solution of $\mathcal{R}_i(\bar{\mathbf{x}})$. In order to generate a Pareto-optimal Benders' optimality cut, we need to find $\pi^*_{i,1}(s), \pi^*_{i,2}(s), \pi^*_{i,3}(s), \pi^*_{i,4}(s)$ for all $s \in \mathscr{S}$ such that they minimize the term on the right-hand side of (3.14) subject to (3.13a) – (3.13e) for each equilibrium $\mathbf{x}$. The right-hand side of (3.14) subject to (3.13a) – (3.13e) may be rewritten as follows:

$$
\begin{aligned}
&\sum_{s\in\mathscr{S}} \pi_{i,1}(s)\, u_i(s,0) + \pi_{i,2}(s)\left(u_i(s,0) + [u_i(s,1) - d_i(s)]x(s)\right) + \\
&\pi_{i,3}(s)\left([u_i(s,1) - d_i(s)]x(s) + d_i(s)\right) + \pi_{i,4}(s)\left(u_i(s,1)x(s) + F_i(s,\mathbf{V}_i)[1 - x(s)]\right) = \\
&\sum_{s\in\mathscr{S}^0(\bar{\mathbf{x}})} [\pi_{i,1}(s) + \pi_{i,2}(s)]\, u_i(s,0) + \pi_{i,2}(s)\, [u_i(s,1) - d_i(s)]x(s) + \\
&\sum_{s\in\mathscr{S}^1(\bar{\mathbf{x}})} [\pi_{i,3}(s) + \pi_{i,4}(s)]\, u_i(s,1)x(s) + [\pi_{i,3}(s)\, d_i(s) + \pi_{i,4}(s)\, F_i(s,\mathbf{V}_i)][1 - x(s)] = \\
&\sum_{s\mathscr{S}^0(\bar{\mathbf{x}})} \gamma_i(s)\, u_i(s,0) + \pi_{i,2}(s)\, [u_i(s,1) - d_i(s)]x(s) + \\
&\sum_{s\in\mathscr{S}^1(\bar{\mathbf{x}})} \gamma_i(s)\, u_i(s,1)x(s) + [\pi_{i,3}(s)\, d_i(s) + \pi_{i,4}(s)\, F_i(s,\mathbf{V}_i)][1 - x(s)],
\end{aligned}
\tag{3.15}
$$

where the first and second equality follow from (3.13a), (3.13c) and (3.13b), (3.13d), respectively. In order to minimize (3.15) subject to (3.13a) – (3.13e), we may seek to minimize (3.15) for each $s \in \mathscr{S}$ separately. There are three cases:

If $s \in \mathscr{S}^0(\bar{\mathbf{x}})$ and $u_i(s,1) \geq d_i(s)$, then $\pi^*_{i,1}(s) := 0, \pi^*_{i,2}(s) := \gamma_i(s), \pi^*_{i,3}(s) := 0, \pi^*_{i,4}(s) := 0$ minimizes (3.15) subject to (3.13a) – (3.13e).

If $s \in \mathscr{S}^0(\bar{\mathbf{x}})$ and $u_i(s,1) < d_i(s)$, then $x(s)$ is equal to 0 for any equilibrium $\mathbf{x}$ by Lemma 5. As a result, $\pi^*_{i,1}(s) := \gamma_i(s), \pi^*_{i,2}(s) := 0, \pi^*_{i,3}(s) := 0, \pi^*_{i,4}(s) := 0$ minimizes (3.15) subject to (3.13a) – (3.13e).

If $s \in \mathscr{S}^1(\bar{\mathbf{x}})$, then $\pi^*_{i,1}(s) := 0, \pi^*_{i,2}(s) := 0, \pi^*_{i,3}(s) := 0, \pi^*_{i,4}(s) := \gamma_i(s)$ minimizes (3.15) subject to (3.13a) – (3.13e) since $F_i(s,\mathbf{V}_i) \geq F_i(s,\mathbf{d}_i) = d_i(s)$.

□

It is worth noting that the cut (3.12) requires only optimal dual multipliers of $\mathcal{R}_i(\bar{\mathbf{x}})$. Moreover, we solve the following separation problem in Step 2 of the algorithm: Given an integral feasible point $(\bar{\mathbf{x}}, \bar{\boldsymbol{\theta}})$, generate a valid inequality, separating the point from the convex hull of the master problem. The computational complexity of this separation problem is clearly polynomial since $\bar{\mathbf{x}}$ is integral.

3.7.1 Branch-and-Cut

In the Decomposition Algorithm, we need to solve the restricted master problem repeatedly. However, solving a mixed-integer restricted master repeatedly may be time-consuming. In order to deal with this difficulty, we use a branch-and-cut framework. At the root node, we start with the MILP in Step 0 of the Decomposition Algorithm. Then, we solve the LP relaxation at each node and generate a violated cut at each integral node by using the separation procedure described in Step 2 of the Decomposition Algorithm. We use the default branching and node selection strategies of the MILP solver, ILOG-CPLEX 12.4 (2012).

Dynamic Variable Fixing and Pruning. At node t of the branch and cut tree, let J_t^k be the set of all states in which $x(s)$ is fixed to k, for any $k = \{0, 1\}$. A strategy profile $\mathbf{x}^t := \langle x^t(s) \rangle_{s \in \mathscr{S}}$ can be assigned to each node t as follows:

$$x^t(s) = \begin{cases} 1 & \text{if } s \in J_t^1, \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathbf{x}^t$ is not an equilibrium, node t is pruned because all possible strategy profiles at this node (admissible with respect to the binary variables being fixed so far) are not equilibria by Corollary 1. On the other hand, if $\mathbf{x}^t$ is an equilibrium, we can apply Proposition 8 to fix some of unfixed binary variables at node t and its offspring. For this reason, $x(s)$ is set to 0 at node t for all $s \in \cup_{i \in \mathscr{N}} \mathscr{S}_{i,v}^0(\mathbf{x}^t)$.

Dynamic Coefficient Strengthening. Recall that $\langle V_i(s) \rangle_{s \in \mathscr{S}}$ is the optimal value function of player i when the autonomy of the other players is suppressed. The coefficient $V_i(s)$ may be replaced with any other upper bound of $w_i(s)$. In general, $V_i(s)$ may be relatively large, which weakens the LP relaxation of Δ. To address this, we dynamically tighten upper bounds for $w_i(s)$ in progress of the branch-and-cut tree rather than using a fixed value as an upper bound for $w_i(s)$.

At leaf node t, we can obtain a tightened upper bound of $w_i(s)$ by using the optimal value function of player i for a certain MDP. In particular, consider an MDP in which the autonomy of all players except for i is suppressed, while player i is restricted to strategies admissible with respect to the binary variables which are fixed at node t. Since the set of fixed binary variables increases as we go down further in the search tree, the tightened upper bound of $w_i(s)$ does not increase. Let $V_i^t(s)$ be the tightened upper bound for $w_i(s)$ at leaf node t. For each player i, $\mathbf{V}_i^t := \langle V_i^t(s)\rangle_{s\in\mathscr{S}}$ can be calculated as the unique solution of the following MDP equations:

$$
\begin{aligned}
&V_i^t(s) = F_i(s, V_i^t) && \forall s \in J_t^0,\\
&V_i^t(s) = u_i(s,1) && \forall s \in J_t^1,\\
&V_i^t(s) = \max\{u_i(s,1), F_i(s, \mathbf{V}_i^t)\} && \forall s \in \mathscr{S}/J_t^0 \cup J_t^1.
\end{aligned}
$$

Similarly, the coefficient $d_i(s)$ is a lower bound for $w_i(s)$, and we may tighten it as we go down in the search tree. At leaf node t, consider payoff profile $\mathbf{w}^{\mathbf{x}^t}$ associated with strategy profile $\mathbf{x}^t$, as defined in the previous part. As noted earlier, if $\mathbf{x}^t$ is not an equilibrium, node t is pruned. Otherwise, $\mathbf{x}^t$ is an equilibrium, and by Proposition 5 (i), $w_i^{\mathbf{x}^t}(s)$ is a lower bound for $w_i(s)$ at node t and its offspring. Compared to $d_i(s)$, $w_i^{\mathbf{x}^t}(s)$ is a tighter lower bound by Proposition 5. The new set of upper and lower bounds is applied to generate a Benders' optimality cut (3.12), which is only locally valid. Since the upper and lower bounds get tighter as we go down further in the search tree, deeper Benders' optimality cuts will be generated. Needless to say, this idea should only be implemented at nodes at which we generate a Benders' optimality cut. The idea of coefficient strengthening has received some recent attention in the optimization community (e.g., Qiu et al. 2014).

Dynamic Player-Aggregated Upper Bounds. In stochastic games, it is natural to assume that the players compete in the same environment, and therefore share the same discount factor (cf. Herings and Peeters 2004, Hörner et al. 2011). Suppose that the discount factors are equal for all players, i.e., $\lambda_i = \lambda$ for all $i \in \mathscr{N}$. We present a family of upper bounds for the objective function at each leaf node t. Let $\alpha_i \in \mathbb{R}$ for all $i \in \mathscr{N}$, and $\alpha := \langle \alpha_i \rangle_{i\in\mathscr{N}}$. We define an aggregated MDP, $\mathcal{G}^{\alpha,t}$, over the state space $\mathscr{S}$ as follows. In each state $s \in \mathscr{S}\backslash(J_t^0 \cup J_t^1)$, we may decide whether to stop or continue. In each state $s \in J_t^1$ (J_t^0), we have to stop (continue). If we decide to stop, then the MDP terminates and we receive a stopping reward $u^{\alpha}(s,1) := \sum_{i\in\mathscr{N}} \alpha_i u_i(s,1)$. Conversely, if we decide to continue, then the MDP moves into a new state s' under the Markovian transition probability $\mathscr{P}(s'|s)$ and we receive an immediate continuation reward $u^{\alpha}(s,0) := \sum_{i\in\mathscr{N}} \alpha_i u_i(s,0)$. All future rewards are discounted at rate λ. Therefore, $\mathcal{G}^{\alpha,t}$ has the same dynamic evolution as $\mathcal{G}$, but the players' rewards are aggregated in $\mathcal{G}^{\alpha,t}$ and its action space is admissible with respect to the binary variables which are fixed at node t. Similar to $\mathcal{G}$, we can define strategy $\mathbf{x}$ and its associated payoffs $\mathbf{w}^{\mathbf{x},\alpha,t}$ for $\mathcal{G}^{\alpha,t}$. For each strategy $\mathbf{x}$ with associated payoffs $\mathbf{w}^{\mathbf{x},\alpha,t}$ for $\mathcal{G}^{\alpha,t}$ and associated payoff profile $\mathbf{w}^{\mathbf{x}}$ for $\mathcal{G}$, it can easily be seen that

$$w^{\mathbf{x},\alpha,t}(s) = \sum_{i\in\mathscr{N}} \alpha_i w_i^{\mathbf{x}}(s) \qquad \forall s \in \mathscr{S}. \tag{3.16}$$

Let $\mathbf{V}^{\alpha,t} := \langle V^{\alpha,t}(s)\rangle_{s\in\mathscr{S}}$ be the optimal value function of $\mathcal{G}^{\alpha,t}$, which is calculated as the unique solution of the following MDP equations:

$$\begin{aligned}
&V^{\alpha,t}(s) = u^{\alpha}(s,0) + \lambda \sum_{s'\in\mathscr{S}} \mathscr{P}(s'|s)V^{\alpha,t}(s') && \forall s \in J_t^0,\\
&V^{\alpha,t}(s) = u^{\alpha}(s,1) && \forall s \in J_t^1,\\
&V^{\alpha,t}(s) = \max\{u^{\alpha}(s,1), u^{\alpha}(s,0) + \lambda \sum_{s'\in\mathscr{S}} \mathscr{P}(s'|s)V^{\alpha,t}(s')\} && \forall s \in \mathscr{S}/J_t^0 \cup J_t^1.
\end{aligned}$$

Consider the following LP:

$$\bar{\theta}^t := \max \sum_{s\in\mathscr{S}, i\in\mathscr{N}} c_i(s) w_i(s) \tag{3.17a}$$

$$s.t. \quad w_i(s) = F_i(s, \mathbf{w}_i) \qquad \forall s \in J_t^0, i \in \mathscr{N}, \tag{3.17b}$$

$$w_i(s) \geq F_i(s, \mathbf{w}_i) \qquad \forall s \in \mathscr{S}/J_t^0, i \in \mathscr{N}, \tag{3.17c}$$

$$w_i(s) = u_i(s,1) \qquad \forall s \in J_t^1, i \in \mathscr{N}, \tag{3.17d}$$

$$\sum_{i\in\mathscr{N}} \alpha_i w_i(s) \leq V^{\alpha,t}(s) \qquad \forall s \in \mathscr{S}, \tag{3.17e}$$

$$w_i(s) \leq V_i^t(s) \qquad \forall s \in \mathscr{S}, i \in \mathscr{N}, \tag{3.17f}$$

$$w_i(s) \text{ unrestricted} \qquad \forall s \in \mathscr{S}, i \in \mathscr{N}. \tag{3.17g}$$

If a strategy profile $\mathbf{x}$ is an equilibrium and admissible at node t, then $\mathbf{w}^{\mathbf{x}}$ satisfies (3.17b) – (3.17d) by Proposition 2, and satisfies (3.17e) – (3.17f) by (3.16) and the definitions of $\mathbf{V}^{\alpha,t}, \mathbf{V}_i^t$. Hence, the following inequality is locally valid at node t and its offspring:

$$\sum_{i \in \mathscr{N}} \theta_i \leq \bar{\theta}^t. \tag{3.18}$$

Of special interest is the case when α is equal to e_i, the i^{th} unit vector in $\mathbb{R}^{|\mathscr{N}|}$. In this case, we do not need the above-mentioned assumption about equality of the discount factors, and we use λ_i as the discount factor of $\mathcal{G}^{e_i,t}$. Consider the following LP for each $i \in \mathscr{N}$:

$$\begin{aligned}
\bar{\theta}_i^t := \max \sum_{s \in \mathscr{S}} c_i(s) w_i(s) & & \quad (3.19\text{a})\\
s.t. \quad w_i(s) = F_i(s, \mathbf{w}_i) & \qquad \forall s \in J_t^0, & \quad (3.19\text{b})\\
w_i(s) \geq F_i(s, \mathbf{w}_i) & \qquad \forall s \in \mathscr{S}/J_t^0, & \quad (3.19\text{c})\\
w_i(s) = u_i(s, 1) & \qquad \forall s \in J_t^1, & \quad (3.19\text{d})\\
w_i(s) \leq V_i^t(s) & \qquad \forall s \in \mathscr{S}, & \quad (3.19\text{e})\\
w_i(s) \text{ unrestricted} & \qquad \forall s \in \mathscr{S}. & \quad (3.19\text{f})
\end{aligned}$$

A reasoning similar to that for validity of the inequality (3.18), shows that following inequality is locally valid at node t and its offspring:

$$\theta_i \leq \bar{\theta}_i^t \qquad \forall i \in \mathscr{N}. \tag{3.20}$$

3.8 COMPUTATIONAL EXPERIMENTS

3.8.1 Implementation and Test Instances

We implemented the branch-and-cut algorithm described in Section 3.7, using the ILOG-CPLEX 12.4 Callable Library embedded in C++ under Microsoft Visual Studio 2010. We conducted our computational experiments on an Intel Xeon PC with 4.0 GHz CPU, 32 GB RAM, and Windows 7 (64-bit) operating system. Each instance of our test bed was processed three times within a 4-hour time limit: First by our implementation of the branch-and-cut algorithm described in Section 3.7 within ILOG-CPLEX 12.4; second by solving the original MILP model (3.10a) – (3.10b) through ILOG-CPLEX 12.4 (with default settings); and third by our implementation of a state-of-the-art Benders' decomposition (Fischetti et al. 2010) within ILOG-CPLEX 12.4. In the implementation of our branch-and-cut algorithm, all cuts are generated at integral nodes and added locally. The valid inequalities (3.18) and (3.20) are added locally in non-integral nodes as well. Dynamic variable fixing and pruning are implemented via a branch callback routine.

In our implementation of Benders' decomposition, the vector $\mathbf{x}$ is considered as the linking variables for the basic MILP (3.10a) – (3.10b). We adopt a branch-and-cut framework as follows. At the root node, we start with the initial problem (3.11a) – (3.11d). At each node, we use the constraint-generation scheme proposed by Fischetti et al. (2010) to identify a violated constraint (if any) for each $i \in \mathscr{N}$. When no constraint is generated at a node, we let the default setting of ILOG-CPLEX 12.4 branch and select another node (if any) to explore. All generated constraints are added globally; in our experiments there was no significant difference compared to adding them locally.

We restrict our attention to consensus stopping game instances in which each player $i \in \mathscr{N}$ has an individual state $s_i \in S_i$ representing his competitive advantage, where S_i denotes his state space. Also, each player $i \in \mathscr{N}$ has an individual Markovian transition probability matrix P_i, where $P_i(s'_i|s_i)$ shows the probability that player i will be in state $s'_i \in S_i$ at the next period given that he is now in state $s_i \in S_i$. As a result, the game state $s \in \mathscr{S}$ is $(s_1, s_2, \ldots, s_N)$, and the game state space $\mathscr{S}$ is the Cartesian product of $S_1, S_2, \ldots, S_N$ so that $|\mathscr{S}| = \prod_{i\in\mathscr{N}} |S_i|$. Moreover, the game transition probability matrix $\mathscr{P}$ is the Kronecker product of transition probability matrices of all players, i.e, $\mathscr{P}(s'_1,\ldots,s'_N|s_1,\ldots,s_N) = \prod_{i\in\mathscr{N}} P_i(s'_i|s_i)$. Clearly, the equilibrium selection MILP grows rapidly as either the number of players or the size of S_i increases. For each instance, there is an initial state, denoted by $\hat{s}$, and the objective function (3.10a) is set to $\sum_{i\in\mathscr{N}} w_i(\hat{s})$. This is a reasonable objective since it is the sum of all players' expected reward-to-go from the initial state. Moreover, the discount factors are equal for all players, and hence we use both valid inequalities (3.18) and (3.20) in the branch-and-cut algorithm. In computation of the valid inequality (3.18), we let $\alpha \in \mathbb{R}^{|\mathscr{N}|}$ be a vector whose components are all equal to 1.

In order to enhance the performance, we provide all approaches with strategy profiles $\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2$ as a warm start such that for each $s \in \mathscr{S}$,

$$\bar{x}_1(s) = \begin{cases} 1 & \text{if } s = \hat{s}, \\ 0 & \text{otherwise}, \end{cases} \qquad \bar{x}_2(s) = \begin{cases} 1 & \text{if } V_i(s) = u_i(s,1)\, \forall i \in \mathscr{N}, \\ 0 & \text{otherwise}. \end{cases}$$

Proposition 11. *Strategy profile $\bar{\mathbf{x}}_2$ is an equilibrium.*

Proof. If $s \in \mathscr{S}^0(\bar{\mathbf{x}}_2)$, the Bellman-Shapley equation (3.4) is obviously satisfied for all $i \in \mathscr{N}$. Conversely, if $s \in \mathscr{S}^1(\bar{\mathbf{x}}_2)$, then

$$w_i^{\bar{\mathbf{x}}_2}(s) = u_i(s,1) = V_i(s) \geq F_i(s, \mathbf{V}_i) \geq F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}_2}) \qquad \forall i \in \mathscr{N},$$

where the first equality follows from Proposition 2 (i), the second equality follows from the definition of $\bar{\mathbf{x}}^2$, and the first and second inequalities follow from the definition of $\mathbf{V}_i$. Hence, the Bellman-Shapley equation (3.4) is satisfied for all $s \in \mathscr{S}^1(\bar{\mathbf{x}}_2), i \in \mathscr{N}$. □

By Proposition 11, strategy profile $\bar{\mathbf{x}}_2$ is an equilibrium. However, $\bar{\mathbf{x}}_1$ is not necessarily an equilibrium, and in such a case it is automatically eliminated from consideration by the MILP solver.

3.8.2 Computational Results 1: Real Clinical Instances

We used the method described in Kurt et al. (2011) to generate three categories of consensus stopping game instances for their kidney exchange problem. In the first (second) category, there are two players such that the size of S_i is equal to 40 (60) for both players. In the third category, there are three players, and $|S_i| = 15$ for all players.

Tables $7 - 9$ include computational results for the first, second, and third category of instances, respectively. In these tables, we report the number of cuts, number of explored nodes, the best solution, optimality gap (%), and running time (in the hour: minute: second time format) for all approaches. Moreover, let Branch-and-Cut refer to the branch-and-cut algorithm described in Section 3.7, CPLEX refer to solving the original MILP model (3.10a) – (3.10b) through ILOG-CPLEX 12.4, and Benders refer to the state-of-the-art Benders' decomposition (Fischetti et al. 2010). We also sorted the instances in the tables with respect to the optimality gap of Branch-and-Cut, CPLEX, and Benders, respectively. ILOG-CPLEX 12.4 does not generate any internal cuts for Branch-and-Cut, and all cuts used to solve the three categories of instances by Branch-and-Cut are the cuts described in Section 3.7.

In Table 7, both Branch-and-Cut and CPLEX can solve instances $a1 - a10$. On the majority of these instances, Branch-and-Cut is several orders of magnitude faster than CPLEX. Branch-and-Cut needs about 15 minutes total to solve all instances of this subset while CPLEX needs almost 4.5 hours total, i.e., Branch-and-Cut is 18 times faster in processing the whole set. Branch-and-Cut can solve each instance of $a11 - a18$ in less than 19 minutes while CPLEX cannot solve any of them. Neither Branch-and-Cut nor CPLEX is able to solve instances $a19 - a30$. However, Branch-and-Cut provides smaller optimality gaps for all instances. Overall, Branch-and-Cut outperforms CPLEX on all instances. It is also worth noting that the optimality gap of Benders is larger than that of Branch-and-Cut and CPLEX on all instances. In some instances, we observe that the number of explored nodes and the optimality gap are 0, meaning that those instances are solved at the root node. In such instances, $\bar{\mathbf{x}}_1$ is an optimal solution which was provided for the three solution approaches through the warm start; however, establishing optimality can be very challenging,

Table 7: Two player with 40 state-per-player instances, sorted with respect to the optimality gaps of Branch-and-Cut, CPLEX, and Benders, respectively.

	Branch-and-Cut					CPLEX					Benders				
Name	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)
$a1$	31	0	11954.3	0.00	0:08:21	0	0	11954.3	0.00	0:27:23	54	3	11954.3	0.00	1:00:25
$a2$	0	0	11684.2	0.00	0:00:01	1380	0	11684.2	0.00	1:36:08	0	0	11684.2	0.00	0:00:01
$a3$	31	0	11472.6	0.00	0:01:45	0	0	11472.6	0.00	0:19:12	17	0	11472.6	0.00	0:14:19
$a4$	0	0	13057.1	0.00	0:00:01	0	0	13057.1	0.00	0:05:47	0	0	13057.1	0.00	0:00:01
$a5$	0	0	11230.3	0.00	0:00:01	1350	0	11230.3	0.00	2:16:39	0	0	11230.3	0.00	0:00:01
$a6$	0	0	11765	0.00	0:00:01	0	0	11765	0.00	0:05:53	0	0	11765	0.00	0:00:01
$a7$	0	0	11840.4	0.00	0:00:01	0	0	11840.4	0.00	0:02:23	0	0	11840.4	0.00	0:00:01
$a8$	25	0	12125.8	0.00	0:01:37	0	0	12125.8	0.00	0:36:51	125	0	12125.8	1.05	4:00:00
$a9$	13	0	13309.1	0.00	0:01:20	0	0	13309.1	0.00	2:28:57	115	0	13309.1	2.96	4:00:00
$a10$	15	0	11648.1	0.00	0:01:47	0	0	11648.1	0.00	0:49:55	141	0	11648.1	4.85	4:00:00
$a11$	19	0	11045.8	0.00	0:01:29	1707	0	11045.8	0.12	4:00:00	149	1	11045.8	2.60	4:00:00
$a12$	21	0	11682.5	0.00	0:01:32	1639	20	11682.5	0.53	4:00:00	150	1	11682.5	2.32	4:00:00
$a13$	10	0	10429.5	0.00	0:04:32	1513	81	10429.5	0.96	4:00:00	169	1	10429.5	7.55	4:00:00
$a14$	592	0	11833.3	0.00	0:18:10	1321	24	11833.3	1.43	4:00:00	197	1	11833.3	3.69	4:00:00
$a15$	14	0	11911.5	0.00	0:01:27	1883	327	11911.5	2.73	4:00:00	154	1	11911.5	3.64	4:00:00
$a16$	139	0	11540.9	0.00	0:05:10	1416	152	11540.9	2.93	4:00:00	260	0	11540.9	4.74	4:00:00
$a17$	25	0	11395.2	0.00	0:01:37	1498	17	11395.2	3.19	4:00:00	942	1	11395.2	4.79	4:00:00
$a18$	12	0	11447	0.00	0:01:20	1849	2	11447	3.71	4:00:00	158	0	11447	6.88	4:00:00
$a19$	1394	277	12096.9	0.36	4:00:00	1834	0	12096.9	5.23	4:00:00	132	0	12096.9	6.51	4:00:00
$a20$	726	163	13190	2.31	4:00:00	0	0	13124.2	6.67	4:00:00	144	0	13124.2	6.80	4:00:00
$a21$	851	161	12286.7	2.54	4:00:00	1509	97	12276.2	3.60	4:00:00	118	0	12276.2	12.33	4:00:00
$a22$	884	162	12297.7	2.58	4:00:00	1520	71	12226	4.22	4:00:00	166	0	12226	10.34	4:00:00
$a23$	938	167	11769.6	2.72	4:00:00	1623	43	11619	4.79	4:00:00	428	0	11619	13.22	4:00:00
$a24$	799	181	11990.3	2.90	4:00:00	1655	400	11901.7	4.23	4:00:00	252	0	11901.7	13.80	4:00:00
$a25$	759	172	10961.2	3.47	4:00:00	1147	128	10940.4	5.22	4:00:00	45186	1	10940.4	13.89	4:00:00
$a26$	836	185	10315.6	3.49	4:00:00	897	278	10264.7	6.84	4:00:00	223	1	10264.7	12.88	4:00:00
$a27$	719	164	11957.8	3.49	4:00:00	1560	70	11937.8	5.97	4:00:00	152	0	11937.8	12.38	4:00:00
$a28$	796	162	11544.9	4.25	4:00:00	1919	38	11521.1	7.71	4:00:00	214	1	11521.1	12.35	4:00:00
$a29$	815	176	9083.38	4.40	4:00:00	840	607	9052.69	4.74	4:00:00	148	1	9046.38	19.99	4:00:00
$a30$	802	194	11599.7	5.73	4:00:00	1301	30	11558.7	8.59	4:00:00	2	0	11558.7	12.36	4:00:00

Table 8: Two player with 60 state-per-player instances, sorted with respect to the optimality gaps of Branch-and-Cut, CPLEX, and Benders, respectively.

	Branch-and-Cut					CPLEX					Benders				
Name	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)
$b1$	0	0	11736.6	0.00	0:00:01	0	0	11736.6	0.00	1:02:43	0	0	11736.6	0.00	0:00:01
$b2$	0	0	11821	0.00	0:00:01	0	0	11821	0.00	0:54:41	0	0	11821	0.00	0:00:01
$b3$	0	0	11699.1	0.00	0:00:01	0	0	11699.1	0.00	1:00:08	0	0	11699.1	0.00	0:00:01
$b4$	0	0	11189.1	0.00	0:00:01	0	0	11189.1	0.00	0:36:11	0	0	11189.1	0.00	0:00:01
$b5$	0	0	13358.9	0.00	0:00:01	0	0	13358.9	0.00	0:31:29	0	0	13358.9	0.00	0:00:01
$b6$	0	0	11540.9	0.00	0:00:01	0	0	11540.9	0.00	0:35:40	0	0	11540.9	0.00	0:00:01
$b7$	0	0	11699.1	0.00	0:00:01	0	0	11699.1	0.00	1:05:26	0	0	11699.1	0.00	0:00:01
$b8$	18	0	10823	0.00	0:20:22	0	0	10823	0.00	3:01:31	11	0	10823	0.00	4:00:00
$b9$	26	0	11163.8	0.00	0:29:02	81	0	11163.8	0.00	3:02:01	203	0	11163.8	1.20	4:00:00
$b10$	32	0	10593	0.00	0:34:33	0	0	10593	$+\infty$	4:00:00	46	0	10593	0.31	4:00:00
$b11$	4	0	12049.7	0.00	0:02:03	0	0	12049.7	$+\infty$	4:00:00	44	0	12049.7	0.55	4:00:00
$b12$	14	0	11627.7	0.00	0:49:29	0	0	11627.7	$+\infty$	4:00:00	48	0	11627.7	0.55	4:00:00
$b13$	104	0	12071.5	0.00	1:10:28	0	0	12071.5	$+\infty$	4:00:00	61	0	12071.5	1.77	4:00:00
$b14$	123	0	11304.7	0.00	0:58:54	0	0	11304.7	$+\infty$	4:00:00	92	1	11304.7	2.62	4:00:00
$b15$	15	0	11117.6	0.00	0:40:21	0	0	11117.6	$+\infty$	4:00:00	74	1	11117.6	3.75	4:00:00
$b16$	19	0	11964.8	0.00	0:38:07	0	0	11964.8	$+\infty$	4:00:00	65	0	11964.8	4.36	4:00:00
$b17$	172	0	11466.9	0.00	1:19:53	0	0	11466.9	$+\infty$	4:00:00	84	0	11466.9	5.64	4:00:00
$b18$	25	0	11014.6	0.00	0:21:51	0	0	11014.6	$+\infty$	4:00:00	86	0	11014.6	6.14	4:00:00
$b19$	35	0	10352.7	0.00	0:46:17	0	0	10352.7	$+\infty$	4:00:00	96	1	10352.7	7.01	4:00:00
$b20$	295	10	12287.9	0.16	4:00:00	0	0	12287.9	$+\infty$	4:00:00	88	1	12287.9	2.72	4:00:00
$b21$	60	10	11831.1	1.06	4:00:00	0	0	11831.1	$+\infty$	4:00:00	126	1	11831.1	7.31	4:00:00
$b22$	63	10	11623.6	2.33	4:00:00	0	0	11623.6	$+\infty$	4:00:00	120	0	11623.6	9.94	4:00:00
$b23$	83	10	11736.2	2.92	4:00:00	0	0	11598.9	$+\infty$	4:00:00	122	1	11598.9	14.29	4:00:00
$b24$	61	10	10959.8	3.48	4:00:00	0	0	10959.8	$+\infty$	4:00:00	122	1	10959.8	11.28	4:00:00
$b25$	81	10	9621.96	3.56	4:00:00	0	0	9606.76	$+\infty$	4:00:00	97	0	9606.76	30.19	4:00:00
$b26$	92	10	9712.26	3.57	4:00:00	0	0	9709.84	4.65	4:00:00	108	0	9709.84	31.09	4:00:00
$b27$	69	10	9996.52	4.40	4:00:00	3894	0	9991.23	5.79	4:00:00	115	0	9991.23	22.37	4:00:00
$b28$	72	10	10149.9	6.02	4:00:00	0	0	10089.9	$+\infty$	4:00:00	79	1	10089.9	22.94	4:00:00
$b29$	70	10	7982.41	6.09	4:00:00	0	0	7961.32	$+\infty$	4:00:00	31	0	7961.32	29.99	4:00:00
$b30$	204	10	9873.91	10.07	4:00:00	0	0	9852.2	14.08	4:00:00	126	1	9852.2	19.76	4:00:00

Table 9: Three player with 15 state-per-player instances, sorted with respect to the optimality gaps of Branch-and-Cut, CPLEX, and Benders, respectively.

	Branch-and-Cut					CPLEX					Benders				
Name	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)
$c1$	0	0	16647	0.00	0:00:00	0	0	16647	0.00	0:04:50	0	0	16647	0.00	0:00:01
$c2$	0	0	18230.9	0.00	0:00:01	0	0	18230.9	0.00	0:04:19	0	0	18230.9	0.00	0:00:01
$c3$	219	0	17088.9	0.00	0:50:08	0	0	17088.9	0.00	1:00:09	558	0	17088.9	1.78	4:00:00
$c4$	30	0	17862.1	0.00	0:09:42	0	0	17862.1	0.00	0:33:57	830	0	17862.1	3.85	4:00:00
$c5$	26	0	15437	0.00	0:09:32	0	0	15437	0.00	0:35:43	969	0	15437	4.92	4:00:00
$c6$	26	0	16345.6	0.00	0:09:35	0	0	16345.6	0.00	1:46:15	435	0	16345.6	6.42	4:00:00
$c7$	317	0	17582	0.00	1:11:43	0	0	17582	0.70	4:00:00	92	0	17582	0.98	4:00:00
$c8$	194	0	16176	0.00	0:44:55	0	0	16176	1.34	4:00:00	195	0	16176	2.28	4:00:00
$c9$	110	0	16919.5	0.00	0:27:48	0	0	16919.5	$+\infty$	4:00:00	267	0	16919.5	1.11	4:00:00
$c10$	152	0	17886.8	0.00	0:36:09	0	0	17886.8	$+\infty$	4:00:00	196	0	17886.8	1.24	4:00:00
$c11$	80	0	17638.8	0.00	0:21:09	0	0	17638.8	$+\infty$	4:00:00	121	0	17638.8	1.60	4:00:00
$c12$	89	0	17802.4	0.00	0:22:40	0	0	17802.4	$+\infty$	4:00:00	267	0	17802.4	1.73	4:00:00
$c13$	213	0	16755.1	0.00	1:08:11	0	0	16755.1	$+\infty$	4:00:00	405	0	16755.1	1.99	4:00:00
$c14$	275	0	17251.9	0.00	1:02:46	0	0	17251.9	$+\infty$	4:00:00	154	0	17251.9	2.34	4:00:00
$c15$	209	0	17528.5	0.00	0:47:55	0	0	17528.5	$+\infty$	4:00:00	285	0	17528.5	3.25	4:00:00
$c16$	695	40	17817.3	0.51	4:00:00	0	0	17817.3	$+\infty$	4:00:00	141	0	17817.3	6.42	4:00:00
$c17$	288	40	18735.5	0.59	4:00:00	1646	0	18700.1	0.70	4:00:00	624	0	18692.3	10.39	4:00:00
$c18$	983	0	17362.6	0.69	4:00:00	0	0	17362.6	$+\infty$	4:00:00	152	0	17362.6	1.54	4:00:00
$c19$	486	40	15703.3	0.95	4:00:00	0	0	15681.5	$+\infty$	4:00:00	651	0	15677	18.04	4:00:00
$c20$	397	29	16926	1.35	4:00:00	2375	0	16909.6	1.35	4:00:00	564	0	16900.7	12.48	4:00:00
$c21$	290	40	11975.8	2.19	4:00:00	254	17	11514.2	5.16	4:00:00	621	0	11508.1	43.85	4:00:00
$c22$	253	40	14279.6	2.22	4:00:00	2814	0	14109.9	3.18	4:00:00	420	0	14094.7	33.47	4:00:00
$c23$	309	40	15946.9	3.21	4:00:00	1270	0	15908.4	3.55	4:00:00	465	0	15871.2	14.90	4:00:00
$c24$	886	23	13783.7	3.30	4:00:00	2542	0	13783.7	3.00	4:00:00	315	0	13776.8	33.55	4:00:00
$c25$	293	40	12439.3	3.93	4:00:00	243	100	12628	1.29	4:00:00	564	0	12295.4	43.47	4:00:00
$c26$	522	40	15158.6	5.27	4:00:00	2576	0	15141.9	5.60	4:00:00	240	0	15100.6	21.07	4:00:00
$c27$	507	34	15227.8	5.42	4:00:00	2015	0	14842.6	7.77	4:00:00	378	0	14793.1	29.00	4:00:00
$c28$	681	31	17215.1	6.66	4:00:00	3163	0	17215.1	7.86	4:00:00	204	0	17174	14.60	4:00:00
$c29$	349	40	12859.5	11.63	4:00:00	852	0	12457.7	13.48	4:00:00	771	0	12416.1	49.68	4:00:00
$c30$	303	40	12828.7	16.51	4:00:00	1622	0	12828.7	17.44	4:00:00	228	0	12740.1	39.02	4:00:00

e.g., instances $a11-a18$. Moreover, we observe that Branch-and-Cut and Benders solve some instances without adding any cut at the root node, meaning that the inequalities (3.11b) are enough to close the optimality gap. CPLEX also solves some instances without adding any cut at the root node by preprocessing and probing techniques. For instances $a20-a30$, no warm-start solution is optimal. For these instances, finding a better solution than the warm-start solution $\bar{\mathbf{x}}_2$ is quite challenging, and CPLEX is unable to find a better solution while Branch-and-Cut is able to find better solutions and provides smaller optimality gaps. Moreover, we let CPLEX explore these instances in a 24-hour time limit, and observed that CPLEX was only able to find a better solution for half of them, and its performance in terms of optimality gap was dominated by that of Branch-and-Cut for a 4-hour time limit.

Table 8 shows a similar pattern. Both Branch-and-Cut and CPLEX can solve instances $b1-b9$, and Branch-and-Cut is several orders of magnitude faster than CPLEX. Branch-and-Cut can solve instances $b10-b19$ while CPLEX cannot solve any of them. The situation for CPLEX is even worse since it is unable to solve even the LP relaxation and find a bound for these instances. We attempted to tune CPLEX by adjusting the pricing strategy of the LP solver, the algorithm used for the LP solver, and the primal heuristic, but this had little effect. Instances $b20-b30$ cannot be solved by Branch-and-Cut nor by CPLEX. However, Branch-and-Cut provides us with better solutions and smaller optimality gaps. For most of instances $b20-b30$, CPLEX is unable to solve even the LP relaxation, and tuning of CPLEX parameters had little effect just as that for instances $b10-b19$. For instances of $b1-b19$, $\bar{\mathbf{x}}_1$ is optimal, and establishing optimality is the primary challenge in which Branch-and-Cut does remarkably well. In fact, for instances $b15-b19$, CPLEX cannot close the optimality gap even in a week. For instances $b20-b30$, both finding a better solution than $\bar{\mathbf{x}}_2$ and establishing optimality are extremely challenging. For (almost) all of these instances, CPLEX is unable to find a better solution than $\bar{\mathbf{x}}_2$ even within (one week) one day, and its optimality gap is still larger than that of Branch-and-Cut for a 4-hour time limit. Generally speaking, the advantage of Branch-and-Cut over CPLEX is even more apparent with the larger instances. Moreover, Benders outperforms CPLEX in the majority of the instances of Table 8 within the time limit. However, it too is dominated by Branch-and-Cut on all instances. Table 9 shows that Branch-and-Cut outperforms the other solution approaches.

Table 10: Performance of Branch-and-Cut when each type of valid inequality is deactivated.

		Original Branch-and-Cut	Inequality (3.7)	Inequality (3.8)	Inequality (3.12)	Inequality (3.18)	Inequality (3.20)
Instances of Table	Number of solved problems	18	18	18	18	8	18
	Average of gap (%)	1.27	1.30	1.46	1.37	2.57	1.28
Instances of Table	Number of solved problems	19	19	19	19	10	17
	Average of gap (%)	1.45	1.45	1.55	1.53	2.65	1.70
Instances of Table	Number of solved problems	15	15	16	16	3	15
	Average of gap (%)	2.14	2.19	2.67	2.57	3.05	2.21

In order to evaluate efficiency of the five families of valid inequalities applied in Branch-and-Cut, we deactivated each type of valid inequality in Branch-and-Cut one at a time and collected the numerical results. In Table 10, we report the number of solved problems and average of the optimality gaps in rows $1-2$, $3-4$, and $5-6$ for instances of Table $7-9$, respectively. In column 1, we report the statistics for Branch-and-Cut. In columns $2-6$, we report the statistics for Branch-and-Cut after deactivation of the associated valid inequalities. This table shows that the valid inequalities (3.18) and (3.20) have a significant effect in closing the optimality gap.

3.8.3 Computational Results 2: More General Instances

All instances studied in the preceding subsection had non-negative rewards. In this subsection, we investigate a set of random instances including both negative and non-negative rewards to test the computational performance of our approach under this setting. In our synthetic test bed, each instance consists of four components: Players' individual transition matrices, rewards, discount factors for each player, and the initial state of the game. All these components, but discount factors, are randomly generated for two categories of instances as follows.

We generated two categories of two-player consensus stopping game instances. For the first (second) category, the size of S_i is equal to 40 (60) for both players. Recall from Subsection 3.8.1 that (1) the game state is $s = (s_1, s_2)$, i.e., $\mathscr{S} = S_1 \times S_2$, and (2) the game transition probability matrix $\mathscr{P}$ is the Kronecker product of the individual transition probability matrices, i.e., $\mathscr{P}(s_1', s_2'|s_1, s_2) = P_i(s_1'|s_1)P_i(s_2'|s_2)$. In many practical applications of consensus stopping games such as war termination and organ exchange, which were discussed in Section 3.1, transition in state of the game is slow, i.e., the game most likely remains in the same state at the next period as that of the current period. For this reason, we randomly generated a set of individual transition probability matrices that are highly diagonal, i.e., the diagonal entries are close to 1. Specifically, in generating our transition matrices we used the notion of increasing failure rate (IFR) property. The IFR property has its origins in maintenance optimization and reliability literature (Barlow and Proschan 1965), but it has been recently shown that data in varying real contexts, primarily in healthcare and service operations, empirically exhibit IFR property (Alagoz 2004). The transition matrices we generated are designed to be moderately sparse but do not have any diagonal entry that is less than 0.99. Such transition matrices can be encountered in real-life dynamic settings where decision epochs are spaced very close to each other so that leaving the state of the system in one period is not very likely. It is also common to see sparse transition matrices with large diagonal entries when solutions of a large-scale dynamic decision-making problem are approximated through state aggregation. To ensure our transition matrices have the IFR property and the specified threshold probabilities in their diagonals, we simulated their entries iteratively starting from the top row. In each particular row, we simulated the entries from left to right in column order after randomly fixing the diagonal entry in the specified range. All entries of the same row are generated from a uniform distribution whose boundaries are imposed by a corresponding partial row sum from the previous row due to the IFR property. While such order restrictions can disallow some entries to be positive, we also allowed each nondiagonal entry to be 0 with probability 0.01. Across all transition matrices we generated, on average, the transition matrices for 40-state-per-player instances were 53% sparse whereas the transition matrices for 60-state-per-player instances were 63% sparse.

For each $i \in \{1,2\}$ and $s \in \mathscr{S}$, the rewards $u_i(s,0)$ and $u_i(s,1)$ only depend on s_i. For each $i \in \{1,2\}$ and $s \in \{(s_1,s_2) \in \mathscr{S} : s_i \neq |S_i|\}$, $u_i(s,0)$ is generated according to a uniform distribution on the interval $[-150,50]$, and $u_i(s,1)$ is equal to $\min_{s\in\mathscr{S}} d_i(s) + rand(s_i)\frac{|S_i|-s_i}{|S_i|}\max_{s\in\mathscr{S}} d_i(s)$, where $rand(s_i)$ is a random number from a uniform distribution on $[0,1]$. In addition, the last state (i.e., $s_i = |S_i|$) is absorbing, and its continuation and stopping rewards are 0 and $-\infty$, respectively. Finally, the initial state of the game $\hat{s}$ is generated by a discrete uniform distribution between $\frac{|S_i|}{4}$ and $\frac{|S_i|}{2}$.

We report our computational results for the first and second categories of instances in Tables 11 – 12, respectively. These tables illustrate that Branch-and-Cut greatly outperforms CPLEX and Benders. Note that CPLEX incorrectly finds many instances infeasible or unbounded. It also removes the optimal solution for instance $d2$ in Table 11. We attempted to circumvent these numerical failures by setting the feasibility tolerance parameter (CPX_PARAM_EPRHS) to its highest allowed value (0.1), but this had little effect. There are a couple of reasons behind these failures: (1) Large variability in the coefficients of the formulation Δ due to the existence of the transition probability matrix as well as the big-M type coefficients $\langle V_i(s)\rangle_{s\in\mathscr{S},i\in\mathscr{N}}$ and $\langle d_i(s)\rangle_{s\in\mathscr{S},i\in\mathscr{N}}$ raises the possibility of numerical instability. (2) For each $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$, if $x(s) = 0$ ($x(s) = 1$), then inequalities (3.9a) – (3.9b) ((3.9c) – (3.9d)) must hold as equalities, which are more numerically unstable.

Table 11: Negative reward instances with two players and 40 state-per-player.

Name	Branch-and-Cut					CPLEX					Benders				
	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)
*d*1	31	0	-43791.2	0.00	0:01:24	0	0	-43791.2	0.00	0:00:07	2523	0	-43791.2	0.14	4:00:00
*d*2	6593	173	-42298.1	0.19	4:00:00	287	0	-42315.4	0.00	0:14:16	1530	2	-42315.4	20.69	4:00:00
*d*3	225	0	-50115	0.00	0:06:32	0	0	INF	N/A	0:00:05	1660	2	-50115	4.54	4:00:00
*d*4	2866	181	-43871.9	0.27	4:00:00	0	0	-44002.2	UNB	0:00:03	14372	0	-44002.2	37.37	4:00:00
*d*5	3149	251	-61003.2	0.01	4:00:00	0	0	INF	N/A	0:00:04	2861	0	-61003.2	8.08	4:00:00
*d*6	31	0	-41204.1	0.00	0:00:43	0	0	INF	N/A	0:00:03	1969	0	-41204.1	1.33	4:00:00
*d*7	4531	178	-40924.1	0.16	4:00:00	0	0	INF	N/A	0:00:02	1380	1	-41415.6	16.59	4:00:00
*d*8	7929	318	-38364.6	0.67	4:00:00	945	0	-38414.4	0.71	4:00:00	1924	1	-38414.4	15.68	4:00:00
*d*9	121	0	-46703.7	0.00	0:04:20	0	0	-46703.7	UNB	0:00:02	4330	0	-46703.7	1.00	4:00:00
*d*10	131	0	-47275.9	0.01	0:01:43	0	0	-47277.6	0.00	0:01:26	1876	1	-47277.6	24.57	4:00:00
*d*11	481	17	-69717	0.01	0:20:24	0	0	-69850.1	UNB	0:00:05	2510	0	-69850.1	9.66	4:00:00
*d*12	113	0	-36144.2	0.00	0:05:45	0	0	INF	N/A	0:00:03	5552	0	-36144.2	1.60	4:00:00
*d*13	1411	44	-56391.4	0.01	0:55:14	0	0	-56431.1	UNB	0:00:05	1828	0	-56431.1	10.92	4:00:00
*d*14	3819	242	-38651.9	0.06	4:00:00	0	0	-38672.8	UNB	0:00:05	1931	0	-38672.8	18.96	4:00:00
*d*15	5424	197	-55730.8	0.05	4:00:00	0	0	INF	N/A	0:00:04	3191	1	-55919.2	11.34	4:00:00
*d*16	7099	315	-33109.3	0.62	4:00:00	366	0	-33621.2	2.02	4:00:00	2006	2	-33621.2	45.65	4:00:00
*d*17	6513	209	-60506.6	0.02	4:00:00	0	0	-62539.4	UNB	0:00:04	2446	1	-62539.4	13.17	4:00:00
*d*18	47	0	-55398.6	0.00	0:01:14	0	0	-55398.6	0.00	0:00:10	2682	0	-55398.6	7.29	4:00:00
*d*19	439	28	-49544.6	0.00	0:20:57	0	0	INF	N/A	0:00:05	3142	1	-50244.9	25.84	4:00:00
*d*20	2104	146	-50500	0.01	1:27:08	0	0	-50662.4	UNB	0:00:03	2869	0	-50662.4	10.87	4:00:00
*d*21	147	0	-54442.8	0.00	0:06:21	0	0	-54442.8	0.00	0:00:07	2179	2	-54442.8	7.10	4:00:00
*d*22	43	0	-26155	0.00	0:03:32	0	0	INF	N/A	0:00:03	355	0	-26155	4.40	4:00:00
*d*23	119	0	-42945.8	0.00	0:03:38	0	0	-43032.7	UNB	0:00:02	5848	1	-43032.7	20.81	4:00:00
*d*24	99	0	-52390.7	0.00	0:04:30	0	0	INF	N/A	0:00:05	4969	0	-52390.7	5.52	4:00:00
*d*25	149	0	-40318	0.00	0:05:09	0	0	-40327.2	UNB	0:00:04	892	1	-40327.2	7.25	4:00:00
*d*26	5001	154	-74479.8	0.52	4:00:00	206	75	-74506.1	0.25	4:00:00	2138	2	-74506.1	10.14	4:00:00
*d*27	168	0	-50124.8	0.00	0:07:36	0	0	-54372.4	UNB	0:00:03	6163	1	-50124.8	9.00	4:00:00
*d*28	0	0	-47153	0.00	0:00:00	0	0	INF	N/A	0:00:03	0	0	-47153	0.00	0:00:00
*d*29	4005	154	-71754.9	0.02	4:00:00	0	0	INF	N/A	0:00:06	5838	0	-71972	35.42	4:00:00
*d*30	104	0	-32799.1	0.00	0:05:00	0	0	INF	N/A	0:00:05	2030	0	-32799.1	6.26	4:00:00

• INF: Infeasible • UNB: Unbounded • N/A: Not applicable

Table 12: Negative reward instances with two players and 60 state-per-player.

Name	Branch-and-Cut					CPLEX					Benders				
	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)	# of cuts	# of nodes	Best solution	Gap (%)	Time (h:m:s)
*e*1	0	0	-9859.9	0.00	0:00:01	0	0	-9859.9	UNB	0:00:24	0	0	-9859.9	0.00	0:00:01
*e*2	62	0	-40272.1	0.00	0:57:54	0	0	INF	N/A	0:00:01	200	1	-40272.1	37.51	4:00:00
*e*3	575	29	-16238.9	6.78	4:00:00	0	0	INF	N/A	0:00:09	200	2	-16529.6	47.49	4:00:00
*e*4	0	0	5491.23	0.00	0:00:01	0	0	5491.23	0.00	0:00:22	0	0	5491.23	0.00	0:00:00
*e*5	1259	65	-14832.3	3.19	4:00:00	0	0	INF	N/A	0:00:02	306	1	-17333.6	33.34	4:00:00
*e*6	380	11	-36677.5	7.94	4:00:00	0	0	INF	N/A	4:00:00	176	1	-36887	29.06	4:00:00
*e*7	0	0	-16584.5	0.00	0:00:01	3846	0	INF	N/A	4:00:00	0	0	-16584.5	0.00	0:00:01
*e*8	423	5	-21916.7	1.68	4:00:00	0	0	INF	N/A	0:00:02	198	2	-22595.7	33.55	4:00:00
*e*9	383	5	-23243.7	0.20	4:00:00	0	0	INF	N/A	0:00:23	218	1	-24911.8	57.43	4:00:00
*e*10	91	0	-16321.9	0.00	1:56:24	0	0	INF	N/A	0:00:10	172	0	-16321.9	1.11	4:00:00
*e*11	416	20	-28559.8	4.38	4:00:00	0	0	INF	N/A	0:00:10	138	0	-29452	18.11	4:00:00
*e*12	67	0	-8071.9	0.00	1:00:08	3864	0	INF	N/A	4:00:00	193	2	-8071.9	10.69	4:00:00
*e*13	913	36	-23466.3	2.69	4:00:00	0	0	INF	N/A	0:00:01	362	1	-23545.8	74.65	4:00:00
*e*14	34	0	-8939.3	0.00	0:34:48	0	0	INF	N/A	0:00:02	171	0	-8939.3	33.58	4:00:00
*e*15	392	24	-9643.13	19.25	4:00:00	0	0	INF	N/A	4:00:00	160	2	-10294.1	61.04	4:00:00
*e*16	220	0	-18861.3	0.00	2:19:10	0	0	-27315.1	$+\infty$	4:00:00	199	1	-18861.3	11.96	4:00:00
*e*17	0	0	-911.841	0.00	0:00:01	0	0	INF	N/A	0:00:09	0	0	-911.841	0.00	0:00:00
*e*18	0	0	5865.58	0.00	0:00:01	3980	0	INF	N/A	4:00:00	0	0	5865.58	0.00	0:00:01
*e*19	0	0	-4283.5	0.00	0:00:01	0	0	-4283.5	0.00	0:00:16	0	0	-4283.5	0.00	0:00:00
*e*20	431	13	-25861.9	5.56	4:00:00	3495	0	-27120.3	11.70	4:00:00	192	1	-27120.3	23.69	4:00:00
*e*21	0	0	-10039.5	0.00	0:00:01	0	0	INF	N/A	0:00:01	0	0	-10039.5	0.00	0:00:01
*e*22	121	0	1717.15	0.00	2:04:38	0	0	INF	N/A	4:00:00	187	1	1717.15	135.90	4:00:00
*e*23	0	0	208.367	0.00	0:00:01	0	0	INF	N/A	0:00:09	0	0	208.367	0.00	0:00:01
*e*24	199	0	-9698.95	0.00	3:18:37	0	0	-17128.8	$+\infty$	4:00:00	222	0	-9698.95	59.03	4:00:00
*e*25	600	14	-15213.2	11.09	4:00:00	0	0	-19753.8	$+\infty$	4:00:00	190	1	-19753.8	54.39	4:00:00
*e*26	0	0	6784.35	0.00	0:00:01	0	0	INF	N/A	0:00:02	0	0	6784.35	0.00	0:00:00
*e*27	0	0	8851.23	0.00	0:00:01	16	0	INF	N/A	4:00:00	0	0	8851.23	0.00	0:00:01
*e*28	373	12	-14376.5	17.56	4:00:00	3509	0	-14376.5	28.64	4:00:00	204	1	-14376.5	35.64	4:00:00
*e*29	351	7	-34220.7	0.67	4:00:00	0	0	INF	N/A	0:00:09	158	0	-34240.8	37.86	4:00:00
*e*30	365	20	-39354.9	7.64	4:00:00	0	0	INF	N/A	1:45:26	180	1	-40643.6	47.62	4:00:00

- INF: Infeasible
- UNB: Unbounded
- N/A: Not applicable

4.0 CHARACTERIZING ENTRY AND EXIT FOR STATIONARY EQUILIBRIA OF A DYNAMIC OLIGOPOLY MODEL

4.1 INTRODUCTION

Industrial organization is a field of economics that studies structures of firms and markets. The computational branch of industrial organization which is connected with analyze stationary equilibria of stochastic games arising in industries, has received a considerable attention since the seminal work of Ericson and Pakes (1995). However, technical difficulties related to computability of stationary equilibria have restricted its applicability. Doraszelski and Satterthwaite (2010) provided a more computationally tractable model of an oligopolistic industry with investment, entry, and exit by imposing symmetry restrictions on firms' behaviors. To enhance the computational tractability, we characterize entry and exit behaviors of firms under stationary equilibria for the model of Doraszelski and Satterthwaite (2010) while relaxing the symmetry restrictions.

4.2 MODEL REVIEW

This section introduces the oligopoly model of Doraszelski and Satterthwaite (2010) that investigates evolution of an industry with heterogeneous firms under a competitive dynamic discrete-time infinite horizon setting. There are two groups of firms: *incumbent* firms and *potential entrants*. In each period, every incumbent firm has to decide whether to remain in or leave the industry. If it chooses to remain, it must decide how much to invest in this period. A potential entrant has to decide whether to enter the industry or not. If it chooses to enter the industry, it must decide how much to invest. When these decisions are made, product market competition occurs and each firm accrues an immediate profit in this period. Next, the industry moves into a new state according to a Markovian transition, and so on. We formalize the evolution of the industry in the following.

States and firms. Let $\mathscr{N}$ and $\mathscr{N}^e$ denote the sets of incumbent and entrant firms, respectively. Let N denote the total number of firms, which is fixed. Firm n is described by its state $\omega_n \in \{0, 1, \ldots, M\}$. States $1, \ldots, M$ describe the product quality of a firm that is active in the product market, i.e., an incumbent firm, while state 0 identifies a firm as being inactive, i.e., a potential entrant. When incumbent (potential entrant) firm n decides to exit (enter), its state in the next period will be $\omega_n = 0$ ($\omega_n \neq 0$). The vector of firms' states is $\omega = (\omega_1, \ldots, \omega_N)$, which characterizes the industry at any point. Ω is the set of all possible states.

Let N^* be the number of incumbent firms (i.e., active firms), so that there are $N - N^*$ potential entrants (i.e., inactive firms). In other words, the number of incumbent firms and the number of potential entrants may vary from period to period, but the total number of incumbent and potential entrants is fixed. Thus, once an incumbent firm exits the industry, a potential entrant automatically takes its slot in the competition and has to decide whether to enter the industry. Potential entrants are drawn from a large pool. They are short-lived and base their entry decisions on the net present value of entering today; potential entrants do not have the option of delaying entry, that is if a potential entrant does not enter in this period, it perishes. In contrast, incumbent firms are long-lived and solve an infinite-horizon maximization problems to reach their exit decisions. They discount future payoffs by a factor of β.

Incumbent firms. Consider incumbent firm n, so that $\omega_n \neq 0$. We assume that at the beginning of each period each incumbent firm draws a random salvage value ϕ_n from a distribution $F(\cdot)$ with $\mathbb{E}(\phi_n) = \phi$. Salvage values are independently and identically distributed across firms and periods. Incumbent firm n learns its salvage value ϕ_n prior to deciding about its exit and investment, but the salvage values of its rivals remain unknown to it. Let $\chi_n(\omega, \phi_n) = 1$ indicate that the decision of incumbent firm n, who has drawn salvage value ϕ_n, is to remain in the industry in state ω and let $\chi_n(\omega, \phi_n) = 0$ indicate that its decision is to exit the industry, collect the salvage value ϕ_n, and perish. Since this decision is conditioned on its private ϕ_n, it is a random variable from the perspective of other firms. We use $\xi_n(\omega) = \int \chi_n(\omega, \phi_n) dF(\phi_n)$ to denote the probability that incumbent firm n remains in the industry in state ω.

If an incumbent firm n remains in the industry at state ω, it competes in the product market, and accrues a current profit of $\pi_n(\omega)$. In addition to receiving the current profit, the incumbent incurs the investment $x_n(\omega) \in [0, \overline{x}]$ that it decided on at the beginning of the period and moves from state ω_n to state $\omega'_n \neq 0$ in accordance with the transition probabilities specified subsequently.

Potential entrants. Suppose that $\omega_n = 0$ and consider potential entrant n. We assume that at the beginning of each period each potential entrant draws a random setup cost ϕ^e_n from a distribution $F^e(\cdot)$ with $\mathbb{E}(\phi^e_n) = \phi^e$. Like salvage values, setup costs are independently and identically distributed across firms and periods, and setup cost of a firm is private information of the firm. If potential entrant n enters the industry, it incurs the setup cost ϕ^e_n. If it stays out, it receives nothing and perishes. We use $\chi^e_n(\omega, \phi^e_n) = 1$ to indicate that the decision of potential entrant n, who has drawn setup cost ϕ^e_n, is to enter the industry in state ω and $\chi^e_n(\omega, \phi^e_n) = 0$ to indicate that its decision is to stay out. From the perspective of other firms $\xi_n(\omega) = \int \chi^e_n(\omega, \phi^e_n) dF^e(\phi_n)$ denotes the probability that potential entrant n enters the industry in state ω.

Unlike an incumbent, the entrant does not compete in the product market. Instead it undergoes a setup period upon committing to entry. The entrant incurs its previously chosen investment $x^e_n \in [0, \overline{x}^e]$ and moves to state $w'_n \neq 0$. Hence, at the end of the setup period, the entrant becomes an incumbent.

Transition probabilities. The probability that the industry transitions from today's state ω to tomorrow's state ω' is determined jointly by the investment decisions of the incumbent firms that remain in the industry and the potential entrants that enter the industry. Thus, $P(\omega' \mid \omega, \chi(\omega, \phi), \chi^e(\omega, \phi^e), x(\omega), x^e(\omega))$ is the probability that the industry moves from state ω to state ω' given that the incumbent firms' exit decision are $\chi(\omega, \phi)$, their investment decisions are $x(\omega)$, etc.

We denote a strategy profile by (x, ξ), that specifies decisions of all players for all states. Following the notation convention, (x_{-n}, ξ_{-n}) denotes strategies of all players except for n.

Proposition 12. *Given a strategy profile (x, ξ) with associated payoff profile $V^{x,\xi}$, the following hold:*

$$V_n^{x,\xi}(\omega,\phi_n) = \pi_n(\omega) + \chi_n(\omega,\phi_n)\big(- x_n(\omega) + \beta \mathbb{E}\left\{ V_n^{x,\xi}(\omega')|\omega, \xi_{-n}(\omega), x(\omega)\right\} \big) + \big(1-\chi_n(\omega,\phi_n)\big)\phi_n \quad \forall \omega \in \Omega, n \in \mathscr{N}, \tag{4.1a}$$

$$V_n^{x,\xi}(\omega) = \mathbb{E}\{V_n^{x,\xi}(\omega,\phi_n)\} \quad \forall \omega \in \Omega, n \in \mathscr{N}, \tag{4.1b}$$

$$V_n^{x,\xi}(\omega,\phi_n^e) = \chi_n^e(\omega,\phi_n^e)\big(- \phi_n^e - x_n(\omega) + \beta \mathbb{E}\left\{ V_n^{x,\xi}(\omega')|\omega, \xi_{-n}(\omega), x(\omega)\right\} \big) \quad \forall \omega \in \Omega, n \in \mathscr{N}^e, \tag{4.1c}$$

$$V_n^{x,\xi}(\omega) = \mathbb{E}\{V_n^{x,\xi}(\omega,\phi_n^e)\} \quad \forall \omega \in \Omega, n \in \mathscr{N}^e. \tag{4.1d}$$

Furthermore, $V^{x,\xi}$ is the unique solution of (4.1a) – (4.1d).

Let

$$g_n\big(\omega, \chi_n(\omega,\phi_n), x_n(\omega), V_n\big) := \pi_n(\omega) + \chi_n(\omega,\phi_n)\big(- x_n(\omega) + \beta \mathbb{E}\left\{ V_n(\omega')|\omega, \xi_{-n}(\omega), x(\omega)\right\} \big) + \big(1-\chi_n(\omega,\phi_n)\big)\phi_n \quad \forall \omega \in \Omega, n \in \mathscr{N}, \tag{4.2a}$$

$$g_n\big(\omega, \chi_n(\omega,\phi_n), x_n(\omega), V_n\big) := \chi_n(\omega,\phi_n)\big(- \phi_n - x_n(\omega) + \beta \mathbb{E}\left\{ V_n(\omega')|\omega, \xi_{-n}(\omega), x(\omega)\right\} \big) \quad \forall \omega \in \Omega, n \in \mathscr{N}^e. \tag{4.2b}$$

Proposition 13. (x,ξ) *is an equilibrium if and only if*

$$V_n^{x,\xi}(\omega,\phi_n) = \sup_{\substack{\tilde{\chi}_n(\omega,\phi_n)\in\{0,1\}\\ \tilde{x}_n\in[0,\overline{x}]}} \left\{g_n\big(\omega,\tilde{\chi}_n(\omega,\phi_n),\tilde{x}_n(\omega),V_n^{x,\xi}\big)\right\} \qquad \forall\omega\in\Omega, n\in\mathscr{N}, \tag{4.3a}$$

$$V_n^{x,\xi}(\omega) = \mathbb{E}\{V_n^{x,\xi}(\omega,\phi_n)\} \qquad \forall\omega\in\Omega, n\in\mathscr{N}, \tag{4.3b}$$

$$V_n^{x,\xi}(\omega,\phi_n^e) = \sup_{\substack{\tilde{\chi}_n^e(\omega,\phi_n)\in\{0,1\}\\ \tilde{x}_n^e\in[0,\overline{x}^e]}} \left\{g_n\big(\omega,\tilde{\chi}_n^e(\omega,\phi_n^e),\tilde{x}_n(\omega),V_n^{x,\xi}\big)\right\} \qquad \forall\omega\in\Omega, n\in\mathscr{N}^e, \tag{4.3c}$$

$$V_n^{x,\xi}(\omega) = \mathbb{E}\{V_n^{x,\xi}(\omega,\phi_n^e)\} \qquad \forall\omega\in\Omega, n\in\mathscr{N}^e. \tag{4.3d}$$

Propositions 12 − 13 are standard results in stochastic games. For further discussion on these results, see Doraszelski and Satterthwaite (2010).

4.3 CHARACTERIZATION OF ENTRY AND EXIT IN STATIONARY EQUILIBRIA

In this section, we provide several characterizations of stationary equilibrium behavior of firms for the model of Doraszelski and Satterthwaite (2010). Specifically, our first characterization presents a set of strategy profiles for which the payoff of a specific player dominates, or is dominated by, that of a given strategy profile. The second characterization provides a set of states in which the Bellman-Shapley equation is violated. The third characterization yields a set of non-equilibrium strategy profiles in a "neighborhood" of a given strategy profile. These properties are expected to facilitate the process of computing stationary equilibria. In what follows, unless otherwise stated, we use the terms strategy and equilibrium to refer to stationary strategy and stationary equilibrium, respectively.

For each strategy profile (x,ξ), define

$$C_n(\omega,x,\xi) := -x_n(\omega) + \beta\mathbb{E}\left\{V_n^{x,\xi}(\omega')|\omega,\xi_{-n}(\omega),x(\omega)\right\} \qquad \forall\omega\in\Omega, \text{ if } n\in\mathscr{N}, \tag{4.4a}$$

$$C_n^e(\omega,x,\xi) := -x_n(\omega) + \beta\mathbb{E}\left\{V_n^{x,\xi}(\omega')|\omega,\xi_{-n}(\omega),x(\omega)\right\} \qquad \forall\omega\in\Omega, \text{ if } n\in\mathscr{N}^e. \tag{4.4b}$$

For an incumbent firm n, $C_n(\omega,x,\xi)$ represents the expected discounted profit if it decides to stay in the market in this period while the other firms always follow the strategies induced by (x, ξ), and firm n follows the strategy induced by (x, ξ) from the next period on. For an entrant firm n, $C_n(\omega, x, \xi)$ represents the maximum expected discounted profit if it enters the market in this period without paying a setup cost while the other firms always follow the strategies induced by (x, ξ) and firm n follows the strategy induced by (x, ξ), from the next period on. For $a \in [0,1]$, the generalized inverse distribution functions $F^{-1}(a)$ and $F^{e-1}(a)$ are as follows:

$$F^{-1}(a) := \inf\{y \in \mathbb{R} | F(y) \geq a\},$$
$$F^{e-1}(a) := \inf\{y \in \mathbb{R} | F^e(y) \geq a\}.$$

For each strategy profile (x, ξ), let

$$\overline{\Omega}_n(x,\xi) := \left\{\omega \in \Omega \mid C_n(\omega, x, \xi) > F^{-1}(\xi_n(\omega))\right\} \qquad \text{if } n \in \mathscr{N}, \tag{4.5a}$$
$$\underline{\Omega}_n(x,\xi) := \left\{\omega \in \Omega \mid C_n(\omega, x, \xi) \leq F^{-1}(\xi_n(\omega))\right\} \qquad \text{if } n \in \mathscr{N}, \tag{4.5b}$$
$$\overline{\Omega}^e_n(x,\xi) := \left\{\omega \in \Omega \mid C^e_n(\omega, x, \xi) \geq F^{e-1}(\xi_n(\omega))\right\} \qquad \text{if } n \in \mathscr{N}^e, \tag{4.5c}$$
$$\underline{\Omega}^e_n(x,\xi) := \left\{\omega \in \Omega \mid C^e_n(\omega, x, \xi) < F^{e-1}(\xi_n(\omega))\right\} \qquad \text{if } n \in \mathscr{N}^e. \tag{4.5d}$$

Proposition 14. *Given a strategy profile $(\hat{x}, \hat{\xi})$ with associated payoff profile $V^{\hat{x},\hat{\xi}}$:*

(i) For $n \in \mathscr{N} \cup \mathscr{N}^e$ and a strategy profile (x, ξ) with associated payoff profile $V^{x,\xi}$, suppose that

$$\hat{\xi}_n(\omega) \leq \xi_n(\omega) \leq F(C_n(\omega, \hat{x}, \hat{\xi})) \qquad \textit{if } \omega \in \overline{\Omega}_n(\hat{x}, \hat{\xi}), \tag{4.6a}$$
$$F(C_n(\omega, \hat{x}, \hat{\xi})) \leq \xi_n(\omega) \leq \hat{\xi}_n(\omega) \qquad \textit{if } \omega \in \underline{\Omega}_n(\hat{x}, \hat{\xi}), \tag{4.6b}$$
$$\hat{\xi}_n(\omega) \leq \xi_n(\omega) \leq F^e(C^e_n(\omega, \hat{x}, \hat{\xi})) \qquad \textit{if } \omega \in \overline{\Omega}^e_n(\hat{x}, \hat{\xi}), \tag{4.6c}$$
$$F^e(C^e_n(\omega, \hat{x}, \hat{\xi})) \leq \xi_n(\omega) \leq \hat{\xi}_n(\omega) \qquad \textit{if } \omega \in \underline{\Omega}^e_n(\hat{x}, \hat{\xi}), \tag{4.6d}$$

and

$$\begin{aligned} &- x_n(\omega) + \beta \mathbb{E}\left\{V^{\hat{x},\hat{\xi}}_n(\omega') | \omega, \xi_{-n}(\omega), x(\omega)\right\} \geq \\ &- \hat{x}_n(\omega) + \beta \mathbb{E}\left\{V^{\hat{x},\hat{\xi}}_n(\omega') | \omega, \hat{\xi}_{-n}(\omega), \hat{x}(\omega)\right\} \qquad \forall \omega \in \Omega. \end{aligned} \tag{4.7}$$

Then $V^{x,\xi}_n(\omega) \geq V^{\hat{x},\hat{\xi}}_n(\omega)$ for all $\omega \in \Omega$.

(ii) For $n \in \mathcal{N} \cup \mathcal{N}^e$ and a strategy profile (x, ξ) with associated payoff profile $V^{x,\xi}$, suppose that

$$\xi_n(\omega) \leq \hat{\xi}_n(\omega) \qquad \text{if } \omega \in \overline{\Omega}_n(\hat{x}, \hat{\xi}), \tag{4.8a}$$

$$\hat{\xi}_n(\omega) \leq \xi_n(\omega) \qquad \text{if } \omega \in \underline{\Omega}_n(\hat{x}, \hat{\xi}), \tag{4.8b}$$

$$\xi_n(\omega) \leq \hat{\xi}_n(\omega) \qquad \text{if } \omega \in \overline{\Omega}^e_n(\hat{x}, \hat{\xi}), \tag{4.8c}$$

$$\hat{\xi}_n(\omega) \leq \xi_n(\omega) \qquad \text{if } \omega \in \underline{\Omega}^e_n(\hat{x}, \hat{\xi}), \tag{4.8d}$$

and

$$\begin{aligned} &-x_n(\omega) + \beta \mathbb{E}\left\{V_n^{\hat{x},\hat{\xi}}(\omega') | \omega, \xi_{-n}(\omega), x(\omega)\right\} \leq \\ &-\hat{x}_n(\omega) + \beta \mathbb{E}\left\{V_n^{\hat{x},\hat{\xi}}(\omega') | \omega, \hat{\xi}_{-n}(\omega), \hat{x}(\omega)\right\} \qquad \forall \omega \in \Omega. \end{aligned} \tag{4.9}$$

Then $V_n^{x,\xi}(\omega) \leq V_n^{\hat{x},\hat{\xi}}(\omega)$ for all $\omega \in \Omega$.

Proof. Note that payoffs of firm n under a given strategy profile may be calculated by value iteration (Denardo 1967). Specifically, let $[V_n^{x,\xi}(\omega, \phi_n)]^m$ and $[V_n^{x,\xi}(\omega)]^m$ be the m^{th} iteration of the value iteration algorithm under strategy profile (x, ξ) where we initialize our value iteration by payoffs of firm n under strategy profile $(\hat{x}, \hat{\xi})$.

(i) First, suppose that $n \in \mathcal{N}$. From Proposition 12, there are following possible cases:

- If $\omega \in \overline{\Omega}_n(\hat{x}, \hat{\xi})$ and $\phi_n < F^{-1}(\hat{\xi}_n(\omega))$, then

$$\begin{aligned} [V_n^{x,\xi}(\omega, \phi_n)]^1 &= \pi_n(\omega) + \big(-x_n(\omega) + \beta \mathbb{E}\left\{[V_n^{x,\xi}(\omega')]^0 | \omega, \xi_{-n}(\omega), x(\omega)\right\} \big) \\ &= \pi_n(\omega) + \big(-x_n(\omega) + \beta \mathbb{E}\left\{V_n^{\hat{x},\hat{\xi}}(\omega') | \omega, \xi_{-n}(\omega), x(\omega)\right\} \big) \\ &\geq \pi_n(\omega) + \big(-\hat{x}_n(\omega) + \beta \mathbb{E}\left\{V_n^{\hat{x},\hat{\xi}}(\omega') | \omega, \hat{\xi}_{-n}(\omega), \hat{x}(\omega)\right\} \big) \\ &= V_n^{\hat{x},\hat{\xi}}(\omega, \phi_n) \\ &= [V_n^{x,\xi}(\omega, \phi_n)]^0. \end{aligned}$$

- If $\omega \in \overline{\Omega}_n(\hat{x}, \hat{\xi})$ and $F^{-1}(\hat{\xi}_n(\omega)) \leq \phi_n < F^{-1}(\xi_n(\omega))$, then

$$\begin{aligned}
[V_n^{x,\xi}(\omega, \phi_n)]^1 &= \pi_n(\omega) + \big(- x_n(\omega) + \beta \mathbb{E} \left\{ [V_n^{x,\xi}(\omega')]^0 | \omega, \xi_{-n}(\omega), x(\omega) \right\} \big) \\
&= \pi_n(\omega) + \big(- x_n(\omega) + \beta \mathbb{E} \left\{ V_n^{\hat{x},\hat{\xi}}(\omega') | \omega, \xi_{-n}(\omega), x(\omega) \right\} \big) \\
&\geq \pi_n(\omega) + \big(- \hat{x}_n(\omega) + \beta \mathbb{E} \left\{ V_n^{\hat{x},\hat{\xi}}(\omega') | \omega, \hat{\xi}_{-n}(\omega), \hat{x}(\omega) \right\} \big) \\
&= C_n(\omega, \hat{x}, \hat{\xi}) \\
&\geq F^{-1}(\xi_n(\omega)) \\
&\geq \phi_n \\
&= V_n^{\hat{x},\hat{\xi}}(\omega, \phi_n) \\
&= [V_n^{x,\xi}(\omega, \phi_n)]^0.
\end{aligned}$$

- If $\omega \in \overline{\Omega}_n(\hat{x}, \hat{\xi})$ and $F^{-1}(\xi_n(\omega)) \leq \phi_n$, then

$$\begin{aligned}
[V_n^{x,\xi}(\omega, \phi_n)]^1 &= \phi_n \\
&= V_n^{\hat{x},\hat{\xi}}(\omega, \phi_n) \\
&= [V_n^{x,\xi}(\omega, \phi_n)]^0.
\end{aligned}$$

- If $\omega \in \underline{\Omega}_n(\hat{x}, \hat{\xi})$ and $\phi_n < F^{-1}(\xi_n(\omega))$, then

$$\begin{aligned}
[V_n^{x,\xi}(\omega, \phi_n)]^1 &= \pi_n(\omega) + \big(- x_n(\omega) + \beta \mathbb{E} \left\{ [V_n^{x,\xi}(\omega')]^0 | \omega, \xi_{-n}(\omega), x(\omega) \right\} \big) \\
&= \pi_n(\omega) + \big(- x_n(\omega) + \beta \mathbb{E} \left\{ V_n^{\hat{x},\hat{\xi}}(\omega') | \omega, \xi_{-n}(\omega), x(\omega) \right\} \big) \\
&\geq \pi_n(\omega) + \big(- \hat{x}_n(\omega) + \beta \mathbb{E} \left\{ V_n^{\hat{x},\hat{\xi}}(\omega') | \omega, \hat{\xi}_{-n}(\omega), \hat{x}(\omega) \right\} \big) \\
&= V_n^{\hat{x},\hat{\xi}}(\omega, \phi_n) \\
&= [V_n^{x,\xi}(\omega, \phi_n)]^0.
\end{aligned}$$

- If $\omega \in \underline{\Omega}_n(\hat{x}, \hat{\xi})$ and $F^{-1}(\xi_n(\omega)) \leq \phi_n < F^{-1}(\hat{\xi}_n(\omega))$, then

$$\begin{aligned}
[V_n^{x,\xi}(\omega, \phi_n)]^1 &= \phi_n \\
&\geq F^{-1}(\xi_n(\omega)) \\
&\geq C_n(\omega, \hat{x}, \hat{\xi}) \\
&= \pi_n(\omega) + \big(- \hat{x}_n(\omega) + \beta \mathbb{E}\left\{ V_n^{\hat{x},\hat{\xi}}(\omega') | \omega, \hat{\xi}_{-n}(\omega), \hat{x}(\omega) \right\} \big) \\
&= V_n^{\hat{x},\hat{\xi}}(\omega, \phi_n) \\
&= [V_n^{x,\xi}(\omega, \phi_n)]^0.
\end{aligned}$$

- If $\omega \in \underline{\Omega}_n(\hat{x}, \hat{\xi})$ and $F^{-1}(\hat{\xi}_n(\omega)) \leq \phi_n$, then

$$\begin{aligned}
[V_n^{x,\xi}(\omega, \phi_n)]^1 &= \phi_n \\
&= V_n^{\hat{x},\hat{\xi}}(\omega, \phi_n) \\
&= [V_n^{x,\xi}(\omega, \phi_n)]^0.
\end{aligned}$$

Our arguments show that if $n \in \mathscr{N}$, then $[V_n^{x,\xi}(\omega, \phi_n)]^1 \geq [V_n^{x,\xi}(\omega, \phi_n)]^0$ for all $\omega \in \Omega$. Otherwise, if $n \in \mathscr{N}^e$, it can similarly be shown that $[V_n^{x,\xi}(\omega, \phi_n)]^1 \geq [V_n^{x,\xi}(\omega, \phi_n)]^0$ for all $\omega \in \Omega$. Therefore, it follows from Proposition 12 that $[V_n^{x,\xi}(\omega)]^1 \geq [V_n^{x,\xi}(\omega)]^0$ for all $\omega \in \Omega$. By the monotonicity of the dynamic programming operator induced by strategy profile (x, ξ) for firm n (Blackwell 1965), it follows that $[V_n^{x,\xi}(\omega)]^{m+1} \geq [V_n^{x,\xi}(\omega)]^m$ for all $\omega \in \Omega$. Consequently, $V_n^{x,\xi}(\omega) = \lim_{m \to +\infty} [V_n^{x,\xi}(\omega)]^m \geq [V_n^{x,\xi}(\omega)]^0 = V_n^{\hat{x},\hat{\xi}}(\omega)$ for all $\omega \in \Omega$.

(ii) The proof is similar to part (i).

□

Proposition 15. (i) *For a strategy profile (x, ξ), the Bellman-Shapley equation (4.3a) is violated for an incumbent firm $n \in \mathscr{N}$ in state (ω, ϕ_n) if the following conditions hold:*

$$F^{-1}(\xi_n(\omega)) \leq \phi_n < C_n(\omega, x, \xi) \qquad \text{if } \omega \in \overline{\Omega}_n(x, \xi), \tag{4.10a}$$

$$C_n(\omega, x, \xi) < \phi_n \leq F^{-1}(\xi_n(\omega)) \qquad \text{if } \omega \in \underline{\Omega}_n(x, \xi). \tag{4.10b}$$

(ii) For a strategy profile (x, ξ), the Bellman-Shapley equation (4.3c) is violated for an entrant firm $n \in \mathcal{N}^e$ in state (ω, ϕ_n) if the following conditions hold:

$$F^{e-1}(\xi_n(\omega)) \leq \phi_n < C_n^e(\omega, x, \xi) \qquad \text{if } \omega \in \overline{\Omega}_n^e(x, \xi), \tag{4.11a}$$

$$C_n^e(\omega, x, \xi) < \phi_n \leq F^{e-1}(\xi_n(\omega)) \qquad \text{if } \omega \in \underline{\Omega}_n^e(x, \xi). \tag{4.11b}$$

Proof. (i) Note that when (ω, ϕ_n) satisfies the condition (4.10a), firm n decides to stay in the market in state (ω, ϕ_n) under strategy profile (x, ξ) (i.e., $\chi_n(\omega, \phi_n) = 1$). The payoff of this decision is equal to $C_n(\omega, x, \xi)$ which is strictly smaller than that of deciding to leave the market, ϕ_n. Hence, the Bellman-Shapley equation (4.3a) is violated in (ω, ϕ_n).

If (ω, ϕ_n) satisfies the condition (4.10b), then firm n decides to leave the market in state (ω, ϕ_n) under strategy profile (x, ξ) (i.e., $\chi_n(\omega, \phi_n) = 0$). The payoff of such a decision is ϕ_n which is strictly smaller than that of deciding to stay in the market, $C_n(\omega, x, \xi)$. Therefore, the Bellman-Shapley equation (4.3a) is violated in (ω, ϕ_n).

(ii) The proof is similar to part (i).

□

Proposition 16. *Given a strategy profile $(\hat{x}, \hat{\xi})$:*

(i) Suppose that there exist $n \in \mathcal{N}$ and $\breve{\omega} \in \Omega$ for which $F^{-1}(\hat{\xi}_n(\breve{\omega})) < C_n(\breve{\omega}, \hat{x}, \hat{\xi})$. Any strategy profile (x, ξ) satisfying conditions (4.6a) – (4.7) and $F^{-1}(\hat{\xi}_n(\breve{\omega})) \leq F^{-1}(\xi_n(\breve{\omega})) < C_n(\breve{\omega}, \hat{x}, \hat{\xi})$ is not an equilibrium.

(ii) Suppose that there exist $n \in \mathcal{N}$ and $\breve{\omega} \in \Omega$ for which $C_n(\breve{\omega}, \hat{x}, \hat{\xi}) < F^{-1}(\hat{\xi}_n(\breve{\omega}))$. Any strategy profile (x, ξ) satisfying conditions (4.8a) – (4.9) is not an equilibrium.

(iii) Suppose that there exist $n \in \mathcal{N}^e$ and $\breve{\omega} \in \Omega$ for which $F^{e-1}(\hat{\xi}_n(\breve{\omega})) < C_n^e(\breve{\omega}, \hat{x}, \hat{\xi})$. Any strategy profile (x, ξ) satisfying conditions (4.6a)–(4.7) and $F^{e-1}(\hat{\xi}_n(\breve{\omega})) \leq F^{e-1}(\xi_n(\breve{\omega})) < C_n^e(\breve{\omega}, \hat{x}, \hat{\xi})$ is not an equilibrium.

(iv) Suppose that there exist $n \in \mathcal{N}^e$ and $\breve{\omega} \in \Omega$ for which $C_n^e(\breve{\omega}, \hat{x}, \hat{\xi}) < F^{e-1}(\hat{\xi}_n(\breve{\omega}))$. Any strategy profile (x, ξ) satisfying conditions (4.8a) – (4.9) is not an equilibrium.

Proof. (i) Since strategy profile (x, ξ) satisfies conditions (4.6a)–(4.7), it follows from Proposition 14 (i) that $V_n^{x,\xi}(\omega) \geq V_n^{\hat{x},\hat{\xi}}(\omega)$ for all $\omega \in \Omega$. Therefore,

$$\mathbb{E}\left\{V_n^{x,\xi}(\omega')|\omega, \xi_{-n}(\omega), x(\omega)\right\} \geq \mathbb{E}\left\{V_n^{\hat{x},\hat{\xi}}(\omega')|\omega, \xi_{-n}(\omega), x(\omega)\right\} \qquad \forall \omega \in \Omega. \tag{4.12}$$

Inequalities (4.7) and (4.12) imply that:

$$
\begin{aligned}
&-x_n(\omega)+\beta\mathbb{E}\left\{V_n^{x,\xi}(\omega')|\omega,\xi_{-n}(\omega),x(\omega)\right\}\geq\\
&-x_n(\omega)+\beta\mathbb{E}\left\{V_n^{\hat{x},\hat{\xi}}(\omega')|\omega,\xi_{-n}(\omega),x(\omega)\right\}\geq\\
&-\hat{x}_n(\omega)+\beta\mathbb{E}\left\{V_n^{\hat{x},\hat{\xi}}(\omega')|\omega,\hat{\xi}_{-n}(\omega),\hat{x}(\omega)\right\} \qquad \forall\omega\in\Omega.
\end{aligned}
$$

Therefore, $C_n(\omega,x,\xi)\geq C_n(\omega,\hat{x},\hat{\xi})$ for all $\omega\in\Omega$. In particular, $C_n(\breve{\omega},x,\xi)\geq C_n(\breve{\omega},\hat{x},\hat{\xi})$. Thus, it follows that $F^{-1}(\xi_n(\breve{\omega}))<C_n(\breve{\omega},x,\xi)$ by the hypothesis that $F^{-1}(\hat{\xi}_n(\breve{\omega}))\leq F^{-1}(\xi_n(\breve{\omega}))<C_n(\breve{\omega},\hat{x},\hat{\xi})$. By Proposition 15, the Bellman-Shapley equation (4.3a) is violated in state $(\breve{\omega},\phi_n)$ for all $\phi_n\in[F^{-1}(\xi_n(\breve{\omega})),C_n(\breve{\omega},x,\xi))$. This implies that strategy profile (x,ξ) is not an equilibrium.

(ii) Since strategy profile (x,ξ) satisfies conditions (4.8a) − (4.9), it follows from Proposition 14 (ii) that $V_n^{x,\xi}(\omega)\leq V_n^{\hat{x},\hat{\xi}}(\omega)$ for all $\omega\in\Omega$. Hence, similar to the proof of part (i), it can be shown that $C_n(\omega,x,\xi)\leq C_n(\omega,\hat{x},\hat{\xi})$ for all $\omega\in\Omega$. In particular, $C_n(\breve{\omega},x,\xi)\leq C_n(\breve{\omega},\hat{x},\hat{\xi})$. As (x,ξ) satisfies conditions (4.8a) − (4.9) and $\breve{\omega}\in\underline{\Omega}_n(\hat{x},\hat{\xi})$, it follows that $\hat{\xi}_n(\breve{\omega})\leq\xi_n(\breve{\omega})$, which in turn implies that $F^{-1}(\hat{\xi}_n(\breve{\omega}))\leq F^{-1}(\xi_n(\breve{\omega}))$. From the hypothesis that $C_n(\breve{\omega},\hat{x},\hat{\xi})<F^{-1}(\hat{\xi}_n(\breve{\omega}))$, it follows that $C_n(\omega,x,\xi)\leq C_n(\breve{\omega},\hat{x},\hat{\xi})<F^{-1}(\hat{\xi}_n(\breve{\omega}))\leq F^{-1}(\hat{\xi}_n(\breve{\omega}))$. As a result, strategy profile (x,ξ) is not an equilibrium since by Proposition 15, the Bellman-Shapley equation (4.3a) is violated in state $(\breve{\omega},\phi_n)$ for all $\phi_n\in(C_n(\breve{\omega},x,\xi),F^{-1}(\xi_n(\breve{\omega}))]$.

(iii) The proof is similar to part (i).

(iv) The proof is similar to part (ii).

□

5.0 CONCLUSIONS

5.1 AN OPTIMAL INCENTIVE ALIGNMENT FOR A NATIONAL KIDNEY EXCHANGE PROGRAM

PKE is a growing clinical practice to provide transplantable kidneys for ESRD patients with willing incompatible living-donors. Currently, the Netherlands and South Korea are conducting a national PKE program, and the United States is considering creating a national PKE program since such a program will provide more and better transplants. In Chapter 2, we investigated barriers to establish a national PKE program in the United States. Specifically, we addressed how to incentivize transplant centers by a payment strategy to participate in a national PKE program. We developed a principal-agent framework to model how a national PKE program and transplant centers interact through the payment strategy. To find an equilibrium payment strategy, we developed a bilevel program that can be solved through a transformation into an MILP. We calibrated our model and provided several data-driven insights regarding an equilibrium payment strategy and benefits of a national PKE program. In particular, we demonstrated that there exists a "win-win" equilibrium payment strategy under which all participants- consisting of patient-donor pairs, insurance companies, and transplant centers- benefit from creation of a national PKE program.

5.2 OPTIMIZING OVER PURE STATIONARY EQUILIBRIA IN CONSENSUS STOPPING GAMES

Stochastic games are a powerful tool to model competition of several players in a dynamic setting. However, their applications have been limited since a stochastic game, in general, possesses multiple stationary equilibria, and that makes it hard to predict how players behave. To solve this issue, it is inherently interesting to find a socially optimal stationary equilibrium. However, the problem of finding a socially optimal stationary equilibrium is often computationally intractable. We showed that such a problem is amenable to MILP approaches for special cases. In Chapter 3, we considered consensus stopping games, a broad class of stochastic stopping games. We studied the problem of finding a best pure stationary equilibrium for this class of games, which we showed to be NP-hard. We presented an MILP formulation for the problem of finding a best pure stationary equilibrium. We characterized the pure stationary equilibria of the game, and developed several families of valid inequalities. We developed an algorithm to solve the problem and demonstrated its efficiency by our computational experiments. The majority of results in this chapter can be applied to nonlinear objective functions of payoffs. In particular, the valid inequalities (3.7) and (3.8) may be applied to any nondecreasing and general nonlinear objective function of payoffs, respectively. The approach of this chapter might also be amenable to analyze stationary equilibria of other types of stochastic games. We leave this extension as a topic for future research.

5.3 CHARACTERIZING ENTRY AND EXIT FOR STATIONARY EQUILIBRIA OF A DYNAMIC OLIGOPOLY MODEL

Industrial organization is a field of economics to study how firms behave in an industry. Recently, a class of stochastic games is adopted to model behaviors of firms in the literature of industrial organization. Complicated structures of this class of stochastic games make it difficult to compute stationary equilibria. In Chapter 4, we investigated this class of stochastic games. To overcome challenges in computation of stationary equilibria, we developed several characterizations of stationary equilibria. We expect these characterizations to facilitate such computations. In fact, we anticipate that such characterizations may be adopted in a framework similar to Chapter 3 to solve the problem of finding a best stationary equilibrium for this class of stochastic games. We leave this direction for future research.

BIBLIOGRAPHY

V. Aguirregabiria and P. Mira. Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53, 2007.

O. Alagoz. *Optimal policies for the acceptance of living-and cadaveric-donor livers*. PhD thesis, University of Pittsburgh, 2004.

American Hospital Association. Hospital facts to know, Accessed May 11, 2015. URL http://www.aha.org/research/rc/stat-studies/Studies.shtml.

I. Ashlagi and A. E. Roth. New challenges in multihospital kidney exchange. *American Economic Review*, 102(3):354–359, 2012.

I. Ashlagi and A. E. Roth. Free riding and participation in large scale, multi-hospital kidney exchange. *Theoretical Economics*, 9(3):817–863, 2014.

I. Ashlagi, F. Fischer, I. A. Kash, and A. D. Procaccia. Mix and match: A strategyproof mechanism for multi-hospital kidney exchange. *Games and Economic Behavior*, 2013.

P. Awasthi and T. Sandholm. Online stochastic optimization in the large: Application to kidney exchange. In *Proceedings of the International Joint Conference on Artificial Intelligence*, volume 9, pages 405–411, 2009.

E. Balas. Disjunctive programming. *Annals of Discrete Mathematics*, 5:3–51, 1979.

E. Balas. Projection, lifting and extended formulation in integer and combinatorial optimization. *Annals of Operations Research*, 140(1):125–161, 2005.

R. E. Barlow and F. Proschan. *Mathematical Theory of Reliability*. John Wiley & Sons, 1965.

J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4(1):238–252, 1962.

T. S. Bentley. 2014 U.S. organ and tissue transplant cost estimates and discussion. Technical report, Milliman, 12 2014.

D. Blackwell. Discounted dynamic programming. *The Annals of Mathematical Statistics*, 36(1): 226–235, 1965.

R. N. Borkovsky, U. Doraszelski, and Y. Kryukov. A user's guide to solving dynamic stochastic games using the homotopy method. *Operations Research*, 58(4-Part-2):1116–1132, 2010.

Bureau of Labor Statistics, Accessed May 09, 2015. URL http://www.bls.gov/data/inflation_calculator.htm.

V. Conitzer and T. Sandholm. New complexity results about Nash equilibria. *Games and Economic Behavior*, 63(2):621–641, 2008.

D. E. Cunningham. Veto players and civil war duration. *American Journal of Political Science*, 50 (4):875–892, 2006.

A. Dehghanian, M. Kurt, and A. J. Schaefer. Optimizing over pure stationary equilibria in consensus stopping games. *Under review*, 2015.

S. Dempe. *Foundations of Bilevel Programming*. Springer, 2002.

E. V. Denardo. Contraction mappings in the theory underlying dynamic programming. *SIAM Review*, 9(2):165–177, 1967.

S. DeNegre and T. K. Ralphs. A branch-and-cut algorithm for bilevel integer programming. In *Proceedings of the Eleventh INFORMS Computing Society Meeting*, pages 65–78, 2009. URL http://coral.ie.lehigh.edu/~ted/files/papers/BILEVEL08.pdf.

U. Doraszelski and J. F. Escobar. A theory of regular Markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification. *Theoretical Economics*, 5(3):369–402, 2010.

U. Doraszelski and M. Satterthwaite. Computable Markov-perfect industry dynamics. *The RAND Journal of Economics*, 41(2):215–243, 2010.

J. Eliashberg and R. L. Winkler. Risk sharing and group decision making. *Management Science*, 27(11):1221–1235, 1981.

H. Ergin, T. Sonmez, and M. U. Unver. Lung exchange. Technical report, Boston College, Department of Economics, 2015.

R. Ericson and A. Pakes. Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, 62(1):53–82, 1995.

J. A. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer-Verlag New York, Inc., 1997.

J. A. Filar, T. A. Schultz, F. Thuijsman, and O. J. Vrieze. Nonlinear programming and stationary equilibria in stochastic games. *Mathematical Programming*, 50(1-3):227–237, 1991.

D. Filson and S. Werner. A bargaining model of war and peace: Anticipating the onset, duration, and outcome of war. *American Journal of Political Science*, 46(4):819–837, 2002.

A. M. Fink. Equilibrium in a stochastic n-person game. *Journal of Science of the Hiroshima University, Series AI*, 28(1):89–93, 1964.

M. Fischetti, D. Salvagnin, and A. Zanette. A note on the selection of Benders' cuts. *Mathematical Programming*, 124(1-2):175–182, 2010.

L. R. Ford and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, 1962.

J. Fortuny-Amat and B. McCarl. A representation and economic interpretation of a two-level programming problem. *Journal of the Operational Research Society*, 32(9):783–792, 1981.

M. Gairing, T. Lücking, M. Mavronicolas, B. Monien, and M. Rode. Nash equilibria in discrete routing games with convex latency functions. In *Automata, Languages and Programming*, pages 645–657. Springer, 2004.

K. M. Glorie, J. J. van de Klundert, and A.P.M. Wagelmans. Kidney exchange with long chains: An efficient pricing algorithm for clearing barter exchanges with branch-and-price. *Manufacturing and Service Operations Management*, 2014.

F. Glover. Improved linear integer programming formulations of nonlinear integer problems. *Management Science*, 22(4):455–460, 1975.

Y. Heller. Sequential correlated equilibria in stopping games. *Operations Research*, 60(1):209–224, 2012.

P. Herings and R.J.A.P. Peeters. Stationary equilibria in stochastic games: Structure, selection, and computation. *Journal of Economic Theory*, 118(1):32–60, 2004.

J. Hörner, T. Sugaya, S. Takahashi, and N. Vieille. Recursive methods in discounted stochastic games: An algorithm for $\delta \to 1$ and a folk theorem. *Econometrica*, 79(4):1277–1318, 2011.

ILOG-CPLEX 12.4. IBM ILOG CPLEX Optimization Studio 12.4 Information Center, 2012. URL http://pic.dhe.ibm.com/infocenter/cosinfoc/v12r4/index.jsp.

R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller, J. W. Thatcher, and J. D. Bohlinger, editors, *Complexity of Computer Computations*, The IBM Research Symposia Series, pages 85–103. Springer US, 1972.

M. Köppe, C. T. Ryan, and M. Queyranne. Rational generating functions and integer programming games. *Operations Research*, 59(6):1445–1460, 2011.

M. Kurt. *Dynamic decision models for managing the major complications of diabetes*. PhD thesis, University of Pittsburgh, 2012.

M. Kurt, M. S. Roberts, A. J. Schaefer, and M. U. Unver. Valuing prearranged paired kidney exchanges: A stochastic game approach. Technical report, Boston College, Department of Economics, 2011.

A. Laupacis, P. Keown, N. Pus, H. Krueger, B. Ferguson, C. Wong, and N. Muirhead. A study of the quality of life and cost-utility of renal transplantation. *Kidney international*, 50(1): 235–242, 1996.

T. L. Magnanti and R. T. Wong. Accelerating Benders decomposition: Algorithmic enhancement and model selection criteria. *Operations Research*, 29(3):464–484, 1981.

E. Maskin and J. Tirole. Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory*, 100(2):191–219, 2001.

A. J. Matas and M. Schnitzler. Payment for living donor (vendor) kidneys: A cost-effectiveness analysis. *American Journal of Transplantation*, 4(2):216–221, 2003.

M. L. Melcher, C. D. Blosser, L. A. Baxter-Lowe, F. L. Delmonico, S. E. Gentry, R. Leishman, G. A. Knoll, M. S. Leffell, A. B. Leichtman, D. A. Mast, P. W. Nickerson, E. F. Reed, M. A. Rees, J. R. Rodrigue, D. L. Segev, D. Serur, S. G. Tullius, E. Y. Zavala, and S. Feng. Dynamic challenges inhibiting optimal adoption of kidney paired donation: Findings of a consensus conference. *American Journal of Transplantation*, 13(4):851–860, 2013. ISSN 1600-6143.

R. A. Montgomery, A. A. Zachary, L. E. Ratner, D. L. Segev, J. M. Hiller, J. Houp, M. Cooper, L. Kavoussi, T. Jarrett, J. Burdick, W. R. Maley, J. K. Melancon, T. Kozlowski, C. E. Simpkins, M. Phillips, A. Desai, B. Reeb, E. Kraus, H. Rabb, M. S. Leffell, and D. S. Warren. Clinical results from transplanting incompatible live kidney donor/recipient pairs using kidney paired donation. *Journal of the American Medical Association*, 294(13):1655–1663, 2005a.

R. A. Montgomery, A. A. Zachary, L. E. Ratner, D. L. Segev, J. M. Hiller, J. Houp, M. Cooper, L. Kavoussi, T. Jarrett, J. Burdick, W. R. Maley, J. K. Melancon, T. Kozlowski, C. E. Simpkins, M. Phillips, A. Desai, B. Reeb, E. Kraus, H. Rabb, M. S. Leffell, and D. S. Warren. Clinical results from transplanting incompatible live kidney donor/recipient pairs using kidney paired donation. *The Journal of the American Medical Association*, 294(13):1655–1663, 2005b.

G. L. Nemhauser and L. E. Trotter Jr. Properties of vertex packing and independence system polyhedra. *Mathematical Programming*, 6(1):48–61, 1974.

A. S. Nowak and K. Szajowski. *Advances in Dynamic Games*. Springer, 2005.

OPTN Data, Accessed March 15, 2015. URL http://optn.transplant.hrsa.gov/converge/latestData/stateData.asp?type=center.

W. B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, 2007.

F. Qiu, S. Ahmed, S. S. Dey, and L. A. Wolsey. Covering linear programming with violations. *INFORMS Journal on Computing*, 26(3):531–546, 2014.

T.E.S. Raghavan and Z. Syed. A policy-improvement type algorithm for solving zero-sum two-person stochastic games of perfect information. *Mathematical Programming*, 95(3):513–532, 2003.

M. A. Rees, M. A. Schnitzler, E. Zavala, J. A. Cutler, A. E. Roth, F. D. Irwin, S. W. Crawford, and A. B. Leichtman. Call to develop a standard acquisition charge model for kidney paired donation. *American Journal of Transplantation*, 12(6):1392–1397, 2012.

S. M. Robinson. Generalized equations and their solutions, part I: Basic theory. In P. Huard, editor, *Point-to-Set Maps and Mathematical Programming*, volume 10 of *Mathematical Programming Studies*, pages 128–141. Springer Berlin Heidelberg, 1979.

A. E. Roth, T. Sönmez, and M. U. Ünver. Kidney exchange. *Quarterly Journal of Economics*, 119 (2):457–488, 2004.

A. E. Roth, T. Sönmez, and M. U. Ünver. Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences. *American economic review*, 97(3):828–851, 2007.

S. L. Saidman, A. E. Roth, T. Sönmez, M. U. Ünver, and F. L. Delmonico. Increasing the opportunity of live kidney donation by matching for two- and three-way exchanges. *Transplantation*, 81(5):773–782, 2006.

D. L. Segev, S. E. Gentry, D. S. Warren, B. Reeb, and R. A. Montgomery. Kidney paired donation and optimizing the use of live donor organs. *Journal of the American Medical Association*, 293(15):1883–1890, 2005.

L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America*, 39(10):1095, 1953.

E. Shmaya and E. Solan. Two-player nonzero-sum stopping games in discrete time. *The Annals of Probability*, 32(3):2733–2764, 2004.

J. D. Singer. Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985. *International Interactions*, 14(2):115–132, 1988.

J. D. Singer, S. Bremer, and J. Stuckey. Capability distribution, uncertainty, and major power war, 1820-1965. In B. Russett, editor, *Peace, War, and Numbers*, pages 19–48. Sage Beverly Hills, 1972.

M. J. Sobel. Noncooperative stochastic games. *The Annals of Mathematical Statistics*, 42(6): 1930–1935, 1971.

E. Solan. Stochastic games. In R. A. Meyers, editor, *Computational Complexity*, pages 3064–3074. Springer New York, 2012.

E. Solan and N. Vieille. Quitting games. *Mathematics of Operations Research*, 26(2):265–285, 2001.

T. Sonmez and M. U. Ünver. Enhancing the efficiency of and equity in transplant organ allocation via incentivized exchange. *Available at SSRN 2551344*, 2015.

H. Sperber. How to find Nash equilibria with extreme total latency in network congestion games? *Mathematical Methods of Operations Research*, 71(2):245–265, 2010.

SRTR. The 2012 Annual Data Report of the U.S. Organ Procurement and Transplantation Network and the Scientific Registry of Transplant Recipients: Transplant Data 1998-2012. Department of Health and Human Services, Health Resources and Services Administration, Healthcare Systems Bureau, Division of Transplantation, Rockville, MD; United Network for Organ Sharing, Richmond, VA; University Renal Research and Education Association, Ann Arbor, MI., Accessed March 15, 2015. URL http://srtr.transplant.hrsa.gov/annual_reports/2012/Default.aspx.

R. H. Steinberg. In the shadow of law or power? Consensus-based bargaining and outcomes in the GATT/WTO. *International Organization*, 56(2):339–374, 2002.

Y. Tang, J. P. Richard, and J. C. Smith. A class of algorithms for mixed-integer bilevel min-max optimization. *Journal of Global Optimization*, pages 1–38, 2015.

A. Tevar. Personal communication, May 2015.

P. Toulis and D. C. Parkes. Design and analysis of multi-hospital kidney exchange mechanisms using random graphs. *Games and Economic Behavior*, 2015.

G. Tsebelis. *Veto Players: How Political Institutions Work*. Princeton University Press, 2002.

United States Renal Data System. 2014 Annual Data Report: Epidemiology of Kidney Disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2014., Accessed June 11, 2015. URL http://www.usrds.org/2014/view/v2_01.aspx.

M. U. Ünver. Dynamic kidney exchange. *Review of Economic Studies*, 77(1):372–414, 2010.

G. Y. Weintraub, C. L. Benkard, and B. Van Roy. Markov perfect industry dynamics with many firms. *Econometrica*, 76(6):1375–1411, 2008.

G. Y. Weintraub, C. L. Benkard, and B. Van Roy. Computational methods for oblivious equilibrium. *Operations Research*, 58(4-Part-2):1247–1265, 2010.