

---

**Published Articles & Papers**

---

2009

# Discounted Robust Stochastic Games and an Application to Queuing Control

Erim Kardes

*University of Southern California, kardes@usc.edu*

Fernando Ordonez

*University of Southern California, fordon@usc.edu*

Randolph Hall

*University of Southern California, rwhall@usc.edu*

Follow this and additional works at: [http://research.create.usc.edu/published\\_papers](http://research.create.usc.edu/published_papers)

---

## Recommended Citation

Kardes, Erim; Ordonez, Fernando; and Hall, Randolph, "Discounted Robust Stochastic Games and an Application to Queuing Control" (2009). *Published Articles & Papers*. Paper 91.

[http://research.create.usc.edu/published\\_papers/91](http://research.create.usc.edu/published_papers/91)

This Article is brought to you for free and open access by CREATE Research Archive. It has been accepted for inclusion in Published Articles & Papers by an authorized administrator of CREATE Research Archive. For more information, please contact [gribben@usc.edu](mailto:gribben@usc.edu).

# Discounted Robust Stochastic Games and an Application to Queueing Control

Erim Kardeş

University of Southern California, kardes@usc.edu,

Fernando Ordóñez

University of Southern California, fordon@usc.edu,

Randolph W. Hall

University of Southern California, rwhall@usc.edu,

This paper presents a robust optimization model for  $n$ -person finite state/action discounted stochastic games with incomplete information. We consider  $n$ -player, non-zero sum discounted stochastic games in which none of the players knows the true data of the game and each player considers a distribution-free incomplete information stochastic game to be played using robust optimization. We call such games “*discounted robust stochastic games*”. Discounted robust stochastic games allow us to use simple uncertainty sets for the unknown data of the game, and eliminate the need to have an a-priori probability distribution over a set of games. We prove the existence of equilibrium points and propose an explicit mathematical programming formulation for an equilibrium calculation. We illustrate the use of discounted robust stochastic games in a single server queueing control problem.

*Subject classifications:* Games: stochastic; dynamic programming/optimal control: Markov: finite state; queues: optimization.

---

## 1. Introduction

There has been an extensive body of research on stochastic games in various fields including operations research, mathematics, and economics since the 1950s. The first paper on finite state/action, discrete time, two person zero-sum stochastic games was introduced by Shapley (1953). Many extensions to this basic model have been proposed after this seminal paper such as games with infinite states and actions,  $n$ -person games, games with incomplete information, continuous time games, and semi-Markov games among others.

In this paper, we consider  $n$ -player discounted stochastic games where no player knows the transition probabilities and/or payoffs of the game exactly. More specifically, we focus on finite state/action stochastic games, where the uncertain parameters belong to an uncertainty set over which we may or may not have probabilistic information and where each player uses robust optimization to cope with this uncertainty.

Stochastic games where the uncertain data (transition probabilities or payoffs) is only known to belong to a given set can arise when the data is obtained through subjective judgments, there is a lack of data, or there are random measurement errors on the data. A concrete example can occur in substitutable product inventory control problems, where different products are sold by different retailers and customers can switch from one product to the other (see Avşar and Gürsoy (2002)). Substitution rates, given as the probability that a customer switches from one product to the other may be unknown for new products, requiring subjective judgments about the ranges that the substitution rates belong to. Another example can occur in a processor-sharing model, where customers are served simultaneously and each arriving customer observes the current load, and then chooses to join the shared system or uses an alternative service option. Each customer makes an individual decision, wishing to minimize his own service time (see Altman and Shimkin (1998)).

Due to service conditions, presence of other customers in the system, and/or server malfunctions, there may be uncertainty in service times with each customer having his own subjective judgment about the uncertainty.

Furthermore, solutions to stochastic games may be sensitive to payoff and transition probability data for which the estimates could be inaccurate. This stems from the fact that at equilibrium in a stochastic game, each player faces a Markov decision process (MDP), and optimal solutions to MDPs can be sensitive with respect to data (Nilim and Ghaoui (2005)).

Since it considers two or more competing decision makers acting as adversaries, a stochastic game could be viewed as a generalization of a Markov Decision Process (MDP) (Filar and Vrieze (1997)). They could also be viewed as a collection of auxiliary one-shot matrix games (Shapley (1953)). Hence, two lines of related research in the literature are MDPs and one-shot games. Some authors have addressed the issue of uncertainty in the transition probabilities of MDPs. A Bayesian approach is presented by Shapiro and Kleywegt (2002), where a prior distribution on the transition matrix should be known. Satia and Lave (1973), White and Eldeib (1994), and Givan et al. (1997) have modeled an MDP, where the transition matrix lies in a given set, which is most typically a polytope. Nilim and Ghaoui (2005) consider robust control in MDPs, where a proof of the robust value iteration is presented. Bagnell et al. (2001) consider a similar problem and present the robust value iteration without proof. Iyengar (2005) considers robust dynamic programming problems and provides an independent proof of the robust value iteration. It is important to note here that these recent efforts consider the robustness in the context of MDPs, where an opponent player is not modeled explicitly.

Using a worst-case approach has been prevalent in game theory since the “max-min” formulation of von Neumann’s and Morgenstern’s. For instance, Gilboa and Schmeidler (1989), Lo (1996), and Marinacci (2000) have presented a max-min approach to cope with uncertainty in normal form games. Although these authors adopt a worst-case approach, their models are fundamentally probabilistic and are based on prior probability distributions. For instance, Lo (1996) models each player as believing that his opponents’ actions are realizations from an unknown probability distribution, which belongs to a set of known multiple priors. A player then wishes to maximize his minimum expected utility, where the minimization is with respect to the set of multiple priors. Gilboa and Schmeidler (1989), and Marinacci (2000) adopt a similar approach where non-additive probability distributions are used instead of sets of multiple priors. Furthermore, these authors address complete information games and adopt a worst-case approach with respect to players’ behaviors towards each other, rather than addressing uncertainty in the data of a game. Harsanyi (1967,1968) modeled incomplete information games by considering that each player could use a prior probability distribution to obtain a conditional distribution on the data of the game unknown to himself. Unlike these approaches, a robust optimization approach to payoff uncertainty in one shot games is considered by Aghassi and Bertsimas (2006). In that work, the authors prove the existence of a robust equilibrium and formulate the robust game by considering that the payoffs belong to a polytope, yielding a method to compute an equilibrium point.

The incomplete information case within the repeated games is first introduced by Aumann and Maschler (1968). Sorin (1984) and Sorin (1985) consider stochastic games with incomplete information on one side that have a single nonabsorbing state. It is proven in that paper that these games have a min-max and a max-min value. However, there is no explicit computational scheme for these values. In a more recent effort, Rosenberg et al. (2004) consider two-player zero-sum stochastic games with incomplete information where the incomplete information is described by a finite collection of stochastic games, and a game is to be played out of the finite set of games over which a probability distribution is specified. That paper focuses on stochastic games in which one player controls the transitions. In other words, they consider that the evolution of the game is independent of one of the opponents’ actions and only depends on one player’s actions. We

note that the approach adopted in Rosenberg et al. (2004) is based on the approach proposed in Harsanyi (1967,1968), and requires a probability distribution over a set of games.

This paper addresses  $n$ -player, non-zero sum discounted stochastic games in which none of the players knows the true transition probabilities and/or payoffs of the game and each player uses a robust optimization approach to address this uncertainty. We call such games “*discounted robust stochastic games*”. We propose a distribution-free model for discounted stochastic games with incomplete information, relaxing the assumptions on former efforts and present an explicit mathematical programming formulation for equilibrium calculation. Specifically, in section 2, we review basic ideas on discounted stochastic games and robust optimization, followed by the formulation of discounted robust stochastic games. In section 3, we prove the existence of a robust Markov perfect equilibrium point. In section 4, we show that when the uncertainty in the transition data comes from a polytope intersected with the probability simplex, the robust equilibrium can be cast as a multilinear system formulation, the solution of which gives the set of equilibrium points of the discounted robust stochastic game. We also show in this section that if there is data uncertainty in the game, players’ approach to this uncertainty may differ. In section 5, we illustrate the use of our approach on an example from a queueing control problem. Finally, section 6 concludes the paper with remarks and future research directions.

## 2. Problem Setup

### 2.1. Stochastic Games

This section reviews the basics of stochastic game theory, as presented in Shapley (1953) and Fink (1964). In stochastic games, the play proceeds from one state to the other according to transition probabilities controlled jointly by two or more players. It consists of states and actions associated with each player. Once the game starts in a state, each player chooses their respective actions. The play then moves into the next state with some probability and continues from thereon. The probability that the game moves into the next state is determined by the current state and the actions chosen in the current state.

Let the set of states  $S = \{1, \dots, M\}$  and the set of players  $I = \{1, \dots, N\}$  be finite. If the game is in state  $s$ , player  $i$  can choose the action  $a^i$  from his finite set  $A_s^i$  of alternatives in state  $s$ . We assume that each player has  $j$  alternatives in every state, i.e.,  $|A_s^i| = j, \forall i \in I, s \in S$ . The extension to the case where players have different number of alternatives in states is straightforward and involves no new insights; only more complex notation.

Suppose that each player makes a choice in state  $s$ , i.e., we have an action tuple  $a = (a^1, \dots, a^i, \dots, a^N) \in A$ , where  $A$  is the set of all possible action tuples in state  $s$ . Then the game moves into state  $k$  with probability  $P_{sak} \geq 0$ ,  $\sum_{k=1}^M P_{sak} = 1$ .

At each stage, players may consider to use mixed strategies. Let  $\mathbf{x}_s^i$  be the probability distribution over the set  $A_s^i$ . In other words, the probability vector for player  $i$  in state  $s$  is  $\mathbf{x}_s^i = (x_{s1}^i, \dots, x_{sj}^i)$ , where  $x_{sk}^i \geq 0, \sum_{k=1}^j x_{sk}^i = 1$ . If we denote the set of mixed strategies of player  $i$  in state  $s$  by  $X_s^i$ , then  $X_s^i$  is the  $j$ -dimensional probability simplex:

$$X_s^i = \{\mathbf{x}_s^i \in \mathbb{R}_+^j \mid \sum_{k=1}^j x_{sk}^i = 1\}.$$

We consider a certain class of strategies as introduced by Shapley (1953), namely, stationary strategies. Stationary strategies prescribe a player the same probabilities for her choices each time the player visits a certain state, no matter what route she follows to reach that state. Such strategies are natural and have been prevalent in the study of stochastic games since the seminal work by Shapley (1953). Let us represent the stationary strategies of a player  $i$  by  $\mathbf{x}^i = (\mathbf{x}_1^i, \dots, \mathbf{x}_M^i)$  and denote the set of mixed strategies of all players in the state space of the game by

$\mathbf{x} = (\mathbf{x}^1, \dots, \mathbf{x}^N)$ . We denote the mixed strategies of all players for all states except for player  $i$  by  $\mathbf{x}^{-i} = (\mathbf{x}^1, \dots, \mathbf{x}^{i-1}, \mathbf{x}^{i+1}, \dots, \mathbf{x}^N)$ . The following notation is used to distinguish a mixed strategy of player  $i$  from those of others:  $(\mathbf{x}^{-i}, \mathbf{u}^i) = (\mathbf{x}^1, \dots, \mathbf{x}^{i-1}, \mathbf{u}^i, \mathbf{x}^{i+1}, \dots, \mathbf{x}^N)$ .

The probability that an action tuple  $a$  is chosen by the players in state  $s$  is denoted by

$$\pi_s^a(\mathbf{x}^{-i}, \mathbf{u}^i) = \prod_{\substack{m=1 \\ m \neq i}}^N x_{sam}^m u_{sai}^i.$$

Finally, we use the following notation.  $X^i = \prod_{s \in S} X_s^i$ ,  $X_s = \prod_{i \in I} X_s^i$ , and  $X = \prod_{i \in I} X^i$ .

Let  $\{C_t^i\}_{t=0}^\infty$  denote the sequence of costs to player  $i$  throughout the process. The expected cost at stage  $t$  to player  $i$  resulting from the strategy  $\mathbf{x} \in X$  and the initial state  $s$  is denoted by  $E_{sx}(C_t^i)$ . The discounted value of a strategy  $\mathbf{x} \in X$  to player  $i$  is defined by

$$v_{s,\beta}^i(\mathbf{x}) := \sum_{t=0}^{\infty} \beta^t E_{sx}(C_t^i),$$

where  $0 \leq \beta < 1$  is the discount factor. To ease the notation, given  $\mathbf{x} \in X$ , we denote the value of the  $\beta$ -discounted stochastic game to player  $i$  starting in state  $s$  by  $v_s^i$ . The value vector for player  $i$  is denoted by  $\mathbf{v}^i$ .

Let  $C_{sa}^i$  be the immediate cost to player  $i$  induced by the action tuple  $a$  chosen by the players in state  $s$ . Suppose that players minimize their expected overall costs throughout the process and that they choose their actions independently of each other at a given state. It is shown in Fink (1964) that  $\forall i \in I, s \in S$ , given  $\mathbf{x}^{-i}$ , there exists a unique value  $v_s^i$  to player  $i$  in state  $s$  that satisfies

$$v_s^i = \min_{\mathbf{u}_s^i \in X_s^i} \sum_{a \in A} \pi_s^a(\mathbf{x}^{-i}, \mathbf{u}^i) \{C_{sa}^i + \beta_i \sum_{k=1}^M P_{sak} v_k^i\}. \quad (1)$$

Equation (1) is a Bellman-type equation, and it is interpreted as follows: Given all other players' strategies fixed, if player  $i$  knew how to play optimally from the next stage on, then, at the current stage, he would select the strategy that minimizes the expected immediate cost at the current stage plus the total expected future costs. Hence, player  $i$  is not only concerned with the immediate outcome of his actions but also with the future consequences of his strategies in the current stage.

We now define equilibrium points in this setting, which are known as Markov perfect equilibria. We note that these type of equilibria form a small subset of the set of subgame perfect equilibria.

**DEFINITION 1.** A point  $\mathbf{x} \in X$  is a Markov Perfect Equilibrium point in a stochastic game if and only if,  $\forall i \in I, \forall s \in S$ , there exists a value  $v_s^i$  that satisfies Equation 1, such that,

$$\mathbf{x}_s^i \in \operatorname{argmin}_{\mathbf{u}_s^i \in X_s^i} \left( \sum_{a \in A} \pi_s^a(\mathbf{x}^{-i}, \mathbf{u}^i) \{C_{sa}^i + \beta_i \sum_{k=1}^M P_{sak} v_k^i\} \right). \quad (2)$$

This definition states that  $\mathbf{x}_s^i$  is an optimal (stationary) strategy for player  $i$  in state  $s$  if, when Equation (1) is satisfied, the corresponding minimizer of the objective function of player  $i$  is the strategy that other players expect player  $i$  to use. If this statement holds for all players and all states, then no player would wish to deviate from their strategies, resulting in an equilibrium.

## 2.2. Robust Optimization

This section briefly reviews the basics of robust optimization, as introduced in Ben-Tal and Nemirovski (1998). Consider the following optimization problem  $P_\gamma : \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}, \gamma)$  s.t.  $F(\mathbf{x}, \gamma) \in K \subset \mathbb{R}^m$ , where  $\gamma \in \mathbb{R}^M$  is the data vector,  $\mathbf{x} \in \mathbb{R}^n$  is the decision vector, and  $K$  is a convex cone. Suppose that the data of  $P_\gamma$  is uncertain and all that is known about the data is that it belongs to an uncertainty set  $U \in \mathbb{R}^M$ . Now, consider the problem  $P = \{P_\gamma\}_{\gamma \in U}$ , where the constraints  $F(\mathbf{x}, \gamma) \in K$  must be satisfied no matter what the actual realization of  $\gamma \in U$  is. An optimal solution to the uncertain problem  $P$  is defined as a solution that must give the best possible guaranteed value under all possible realizations of constraints. Formally, it should be an optimal solution of the following program:  $P_R : \min_{\mathbf{x} \in \mathbb{R}^n} \{\sup_{\gamma \in U} f(\mathbf{x}, \gamma) \text{ s.t. } F(\mathbf{x}, \gamma) \in K, \forall \gamma \in U\}$ . Problem  $P_R$  is called the robust counterpart of  $P$ , and its feasible and optimal solutions are called robust feasible and robust optimal solutions, respectively (Ben-Tal and Nemirovski (1998)). Prior work (Bertsimas and Sim (2004), Ben-Tal and Nemirovski (1998)) has also shown that for many function types and uncertainty sets, the robust counterpart problem  $P_R$  can be solved as a single optimization problem of size comparable to a deterministic version of the problem.

## 2.3. Formulation of Robust Stochastic Games

In this section, we formalize our robust model for incomplete information stochastic games by considering that both payoffs and transition probabilities of the game belong to respective uncertainty sets. In discounted robust stochastic games, it is assumed that the players commonly know the uncertainty set of payoffs  $C_s$  at each state and the set of transition probabilities  $P_s$  out of each state. Unlike the approach in Rosenberg et al. (2004), players need not have distributional information on the uncertainty sets with respect to which they adopt a worst-case approach. Now, in light of the results summarized in the previous section, we notice the following: If a player knew how to play in the robust stochastic game optimally from the next stage on, then, at the current stage, she would play with such strategies so that she *minimizes the maximum expected immediate cost* at the current stage and also *minimizes the maximum expected costs possibly incurred in future stages*. Hence, if optimal robust values for player  $i$  exist, given  $\mathbf{x}^{-i}$ , they must satisfy the following Bellman-type equation,

$$\omega_s^i = \min_{\mathbf{u}_s^i \in X_s^i} \max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \sum_{a \in A} \pi_s^a(\mathbf{x}^{-i}, \mathbf{u}^i) \{ \tilde{C}_{sa} + \beta_i \sum_{k=1}^M \tilde{P}_{sak} \omega_k^i \}, \quad (3)$$

where the inner maximization problem is with respect to the uncertain transition probabilities and immediate costs. Here, the vectors  $\tilde{C}_s$  and  $\tilde{P}_s$  are elements of  $C_s$  and  $P_s$ , with components  $\tilde{C}_{sa}$  and  $\tilde{P}_{sak}$ ,  $a \in A$ ,  $s, k \in S$ , respectively. We provide the conditions under which the robust values exist (and Equation 3 holds) in the next section when we study the existence of equilibrium (Theorem 2). In our approach, nature is implicitly modeled as the  $(n+1)^{\text{th}}$  player, who makes decisions last and tries to worsen the payoff of every player by selecting the worst-case data at every state, with respect to the players' mixed strategies.

To ease the notation, let us define

$$\psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \omega^i) = \sum_{a \in A} \pi_s^a(\mathbf{x}^{-i}, \mathbf{u}^i) \{ \tilde{C}_{sa} + \beta_i \sum_{k=1}^M \tilde{P}_{sak} \omega_k^i \}.$$

We are now ready to state our definition of robust Markov perfect equilibrium in discounted robust stochastic games.

DEFINITION 2. A point  $\mathbf{x}$  is a *Robust Markov Perfect Equilibrium* point in a discounted robust stochastic game if and only if,  $\exists \boldsymbol{\omega} = (\boldsymbol{\omega}^1, \dots, \boldsymbol{\omega}^N)$  satisfying Equation (3), such that,  $\forall i \in I, \forall s \in S$ ,

$$\mathbf{x}_s^i \in \operatorname{argmin}_{\mathbf{u}_s^i \in X_s^i} \max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i). \quad (4)$$

### 3. Existence of Equilibrium

Our proof of existence of equilibrium points in discounted robust stochastic games parallels Fink's (1964). However, a different point-to-set mapping (correspondence) is defined that takes the robustness into account. This mapping uses a maximum expected total cost function with respect to mixed strategies. The proof is separated into two parts: First part shows that for any strategy  $\mathbf{x}$  of the players, there exists a unique robust value vector for a player. The second part uses Kakutani's fixed point theorem (Theorem 3) to show that the correspondence we consider has a fixed point that coincides with an equilibrium point.

Before we begin, we introduce some additional notation and present Kakutani's fixed point theorem (Kakutani (1941)). In the following proofs,  $C_s$  is a closed and bounded set, and  $P_s$  is a closed set.

Let  $W^i \equiv \{\boldsymbol{\omega}_s^i \in \Re\}_{s \in S}$ ,  $W \equiv \{W^i\}_{i \in I}$ . The infinity norm on  $W$  is defined as follows:  $\|\boldsymbol{\omega} - \boldsymbol{\theta}\|_\infty = \max_{i \in I, s \in S} |\boldsymbol{\omega}_s^i - \boldsymbol{\theta}_s^i|$ . Let  $f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$  be the worst-case expected reward for player  $i$  in state  $s$ , i.e.:

$$f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) = \max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i).$$

Given the strategies of all other players,  $\mathbf{x}_s^{-i}$ , we next define a best response function that takes any robust value vector for player  $i$ ,  $\boldsymbol{\omega}^i$ , and minimizes the maximum expected total cost with respect to the mixed strategies for player  $i$ . Let  $\gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) : W^i \rightarrow \Re$  be defined by

$$\gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) = \min_{\mathbf{u}_s^i \in X_s^i} f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i).$$

*Part I:*

The next theorem shows that the best response function for a player is a contraction mapping and Theorem 2 below shows that a unique robust value vector exists for any given strategy of the players.

THEOREM 1. For  $\mathbf{x} \in X$ , define  $\gamma_{\mathbf{x}}(\boldsymbol{\omega}) : W \rightarrow W$  by  $(\gamma_{\mathbf{x}}(\boldsymbol{\omega}))_{is} = \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)$ . The function  $\gamma_{\mathbf{x}}(\boldsymbol{\omega})$  is a contraction mapping.

When all the other players' strategies are fixed, player  $i$  faces a robust MDP. Hence, proof of Theorem 1 follows directly from Theorem 5 in Iyengar (2005). For completeness, we also give an alternative proof for Theorem 1 in the appendix.

THEOREM 2. *Application of Banach's Theorem.* For any  $\mathbf{x} \in X$ , and  $\forall i \in I, s \in S$ , there exists a unique  $\boldsymbol{\omega}_s^i$  such that  $\boldsymbol{\omega}_s^i = \min_{\mathbf{u}_s^i \in X_s^i} \max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) = \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)$ .

**Proof.**

Note that  $(W, \|\cdot\|_\infty)$  is a complete metric space and by Theorem 1,  $\gamma_{\mathbf{x}} : W \rightarrow W$  is a contraction mapping. Therefore, by Banach's Theorem,  $\gamma_{\mathbf{x}}(\boldsymbol{\omega})$  has a unique fixed point,  $\boldsymbol{\omega}$ . That is, there exists a unique vector,  $\boldsymbol{\omega}$ , such that  $\gamma_{\mathbf{x}}(\boldsymbol{\omega}) = \boldsymbol{\omega}$ , which means

$$\boldsymbol{\omega}_s^i = \min_{\mathbf{u}_s^i \in X_s^i} \max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) = \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i), \quad \forall i \in I, s \in S. \quad (5)$$

□

We define the unique robust values for player  $i$  by

$$\tau^i(\mathbf{x}^{-i}) = \{\omega^i = (\omega_1^i, \dots, \omega_M^i) : \omega_s^i = \min_{\mathbf{u}_s^i \in X_s^i} \max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \omega^i), s \in S\},$$

and denote the  $s^{th}$  element of  $\tau^i(\mathbf{x}^{-i})$  by  $\tau_s^i(\mathbf{x}^{-i}) = \gamma_s^i(\mathbf{x}_s^{-i}, \omega^i)$ .

*Part II:*

We now show the existence of an equilibrium point that satisfies conditions (4) by using Kakutani's fixed point theorem, which we present below.

**DEFINITION 3.** A correspondence  $\phi : S \rightarrow 2^S$  is upper semi-continuous if  $\mathbf{y}^n \in \phi(\mathbf{x}^n)$ ,  $\lim_{n \rightarrow \infty} \mathbf{x}^n = \mathbf{x}$ ,  $\lim_{n \rightarrow \infty} \mathbf{y}^n = \mathbf{y}$  imply that  $\mathbf{y} \in \phi(\mathbf{x})$ .

**THEOREM 3.** (*Kakutani's Fixed Point Theorem*). If  $S$  is a closed, bounded, and convex set in a Euclidean space, and  $\phi$  is an upper semi-continuous correspondence mapping  $S$  into the family of closed, convex subsets of  $S$ , then  $\exists \mathbf{x} \in S$ , s.t.  $\mathbf{x} \in \phi(\mathbf{x})$ .

We will show that the fixed point of a suitably constructed correspondence is an equilibrium point. To this end, let

$$\begin{aligned} \phi(\mathbf{x}) &= \{\mathbf{y} \in X \mid \mathbf{y}_s^i \in \operatorname{argmin}_{\mathbf{u}_s^i \in X_s^i} f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \omega^i), \\ \omega_s^i &= \min_{\mathbf{u}_s^i \in X_s^i} f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \omega^i), \forall s \in S, i \in I\}. \end{aligned}$$

To show that this correspondence satisfies the assumptions of Kakutani's theorem, we first need several technical results. Let us define the metrics:

$$d_{X_s}(\mathbf{x}_s, \mathbf{y}_s) = \max_{i \in I} \|\mathbf{x}_s^i - \mathbf{y}_s^i\|_\infty, \quad d_{W^i}(\omega^i, \theta^i) = \max_{s \in S} \|w_s^i - \theta_s^i\|_\infty.$$

For the strategy tuples  $\mathbf{x}_s$  and  $\mathbf{u}_s$  in state  $\mathbf{s}$ , and for the value vectors  $\omega^i$  and  $\theta^i$ , let  $\mathbf{p} = (\mathbf{x}_s, \mathbf{w}^i)$ ,  $\mathbf{q} = (\mathbf{y}_s, \theta^i)$ .  $d_1(\mathbf{p}, \mathbf{q}) = d_{X_s}(\mathbf{x}_s, \mathbf{y}_s) + d_{W^i}(\mathbf{w}^i, \theta^i)$ . We need the following lemma to show that  $f_s^i$  satisfies the properties needed to use Kakutani's theorem.

**LEMMA 1.** Given  $\epsilon > 0$ ,  $\exists \delta(\epsilon) > 0$  such that if for any  $\mathbf{p}, \mathbf{q} \in X_s \times W^i$ ,  $d_1(\mathbf{p}, \mathbf{q}) < \delta(\epsilon)$ , then,  $\forall \tilde{C}_s \in C_s$ ,  $\forall \tilde{P}_s \in P_s$ ,  $|\psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s, \omega^i) - \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{y}_s, \theta^i)| < \epsilon$ .

**Proof.**

Since,  $\tilde{C}_s \in C_s$  and  $C_s$  is bounded  $\forall s \in S$ , we have  $|\tilde{C}_{sa}| \leq K$ , where  $K < \infty$ . It is clear that robust values are bounded. Hence, we have  $\forall i \in I, s \in S$ , that  $|\omega_s^i| \leq W$ , where  $W < \infty$ . Note that

$$\begin{aligned} & \left| \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s, \omega^i) - \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{y}_s, \theta^i) \right| = \\ & \left| \sum_{a \in A} \prod_{m=1}^N x_{sa^m}^m \tilde{C}_{sa} + \beta_i \sum_{a \in A} \left( \prod_{m=1}^N x_{sa^m}^m \right) \left( \sum_{k=1}^M \tilde{P}_{sak} \omega_k^i \right) \right. \\ & \left. - \sum_{a \in A} \prod_{m=1}^N y_{sa^m}^m \tilde{C}_{sa} - \beta_i \sum_{a \in A} \left( \prod_{m=1}^N y_{sa^m}^m \right) \left( \sum_{k=1}^M \tilde{P}_{sak} \theta_k^i \right) \right| \end{aligned}$$



$$\begin{aligned}
&= \left| \sum_{a \in A} \tilde{C}_{sa} \left( \prod_{m=1}^N x_{sa^m}^m - \prod_{m=1}^N y_{sa^m}^m \right) + \beta_i \sum_{a \in A} \sum_{k=1}^M \tilde{P}_{sak} \left( \prod_{m=1}^N x_{sa^m}^m \omega_k^i - \prod_{m=1}^N y_{sa^m}^m \theta_k^i \right) \right| \\
&\leq \left| \sum_{a \in A} \tilde{C}_{sa} \left( \prod_{m=1}^N x_{sa^m}^m - \prod_{m=1}^N y_{sa^m}^m \right) \right| + \left| \beta_i \sum_{a \in A} \sum_{k=1}^M \tilde{P}_{sak} \left( \prod_{m=1}^N x_{sa^m}^m \omega_k^i - \prod_{m=1}^N y_{sa^m}^m \theta_k^i \right) \right| \\
&\leq K \sum_{a \in A} \left| \prod_{m=1}^N x_{sa^m}^m - \prod_{m=1}^N y_{sa^m}^m \right| + \beta_i \sum_{a \in A} \sum_{k=1}^M \left| \prod_{m=1}^N x_{sa^m}^m \omega_k^i - \prod_{m=1}^N y_{sa^m}^m \theta_k^i \right|.
\end{aligned}$$

The second to last inequality above follows from the triangle inequality. The last inequality follows since we have  $|\tilde{C}_{sa}| \leq K$  and  $\tilde{P}_{sak} \leq 1, \forall s \in S, a \in A, k \in S$ .

Let

$$\delta_1(\epsilon) = \frac{\min\{\epsilon, 1\}}{3K(2^N - 1)j^N}, \delta_2(\epsilon) = \frac{\min\{\epsilon, 1\}}{3M\beta_i j^N}, \delta_3(\epsilon) = \frac{\min\{\epsilon, 1\}}{3WM\beta_i(2^N - 1)j^N},$$

and let  $\delta(\epsilon) = \min\{\delta_1(\epsilon), \delta_2(\epsilon), \delta_3(\epsilon)\}$ . Now,  $d_1(\mathbf{p}, \mathbf{q}) < \delta(\epsilon)$  implies that,  $\forall i \in I, s \in S$ , and for all actions  $a^i$ ,  $x_{sa^m}^m = y_{sa^m}^m + \alpha_{sa^m}^m$  and  $\omega_s^i = \theta_s^i + \gamma_s^i$ , where  $|\alpha_{sa^m}^m| < \delta(\epsilon)$ , and  $|\gamma_s^i| < \delta(\epsilon)$ . We will make use of the following algebraic identity.

$$\left| \prod_{m=1}^N (y_{sa^m}^m + \alpha_{sa^m}^m) - \prod_{m=1}^N y_{sa^m}^m \right| = \left| \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left( \prod_{m \in I} \alpha_{sa^m}^m \right) \left( \prod_{m \in I^C} y_{sa^m}^m \right) \right|,$$

where  $I^C = \{1, \dots, N\} \setminus I$ . Note that  $\prod_{m \in I} |\alpha_{sa^m}^m| < (\delta_1(\epsilon))^{|I|} \leq \delta_1(\epsilon)$ , and that

$$\left| \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left( \prod_{m \in I} \alpha_{sa^m}^m \right) \left( \prod_{m \in I^C} y_{sa^m}^m \right) \right| \leq \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{sa^m}^m \right| \left| \prod_{m \in I^C} y_{sa^m}^m \right|.$$

Hence, we have

$$\begin{aligned}
K \sum_{a \in A} \left| \prod_{m=1}^N (y_{sa^m}^m + \alpha_{sa^m}^m) - \prod_{m=1}^N y_{sa^m}^m \right| &\leq K \sum_{a \in A} \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{sa^m}^m \right| \left| \prod_{m \in I^C} y_{sa^m}^m \right| \\
&\leq K \sum_{a \in A} \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{sa^m}^m \right| < K \sum_{a \in A} \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \gamma_1(\epsilon) = \frac{\epsilon}{3}.
\end{aligned}$$

We also have

$$\begin{aligned}
\beta_i \sum_{a \in A} \sum_{k=1}^M \left| \prod_{m=1}^N x_{sa^m}^m \omega_k^i - \prod_{m=1}^N y_{sa^m}^m \theta_k^i \right| &= \beta_i \sum_{a \in A} \sum_{k=1}^M \left| \prod_{m=1}^N (y_{sa^m}^m + \alpha_{sa^m}^m) \omega_k^i - \prod_{m=1}^N y_{sa^m}^m \theta_k^i \right| \\
&= \beta_i \sum_{a \in A} \sum_{k=1}^M \left| \prod_{m=1}^N y_{sa^m}^m (\omega_k^i - \theta_k^i) + \omega_k^i \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \prod_{m \in I} \alpha_{sa^m}^m \prod_{m \in I^C} y_{sa^m}^m \right| \\
&\leq \beta_i \sum_{a \in A} \sum_{k=1}^M \left| \prod_{m=1}^N y_{sa^m}^m \right| |\omega_k^i - \theta_k^i| + \beta_i W \sum_{a \in A} \sum_{k=1}^M \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{sa^m}^m \right| \left| \prod_{m \in I^C} y_{sa^m}^m \right| \tag{6}
\end{aligned}$$

$$\leq \beta_i \sum_{a \in A} \sum_{k=1}^M |\gamma_s^i| + \beta_i W \sum_{a \in A} \sum_{k=1}^M \sum_{\substack{I \subseteq 1, \dots, N \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{sa^m}^m \right| \quad (7)$$

$$\begin{aligned} &< \beta_i \sum_{a \in A} \sum_{k=1}^M \delta_2(\epsilon) + \beta_i W \sum_{a \in A} \sum_{k=1}^M \sum_{\substack{I \subseteq 1, \dots, N \\ |I| \geq 1}} \delta_3(\epsilon) \\ &= \frac{\epsilon}{3} + \frac{\epsilon}{3} = 2\frac{\epsilon}{3}. \end{aligned} \quad (8)$$

The inequality ( $\leq$ ) in (6) above follows from triangle inequality, and from the fact that the robust values are bounded and from the facts

$$\left| \prod_{m=1}^N y_{sa^m}^m (\omega_k^i - \theta_k^i) \right| \leq \left| \prod_{m=1}^N y_{sa^m}^m \right| |(\omega_k^i - \theta_k^i)|$$

and

$$\left| \sum_{\substack{I \subseteq 1, \dots, N \\ |I| \geq 1}} \prod_{m \in I} \alpha_{sa^m}^m \prod_{m \in I^C} y_{sa^m}^m \right| \leq \sum_{\substack{I \subseteq 1, \dots, N \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{sa^m}^m \right| \left| \prod_{m \in I^C} y_{sa^m}^m \right|.$$

The inequality ( $\leq$ ) in (7) above follows since  $|\prod_{m \in I^C} y_{sa^m}^m| \leq 1$ , and  $\omega_s^i - \theta_s^i = \gamma_s^i$ . The inequality ( $<$ ) in (8) follows since

$$|\gamma_s^i| < \delta(\epsilon) \leq \delta_2(\epsilon), \text{ and } \left| \prod_{m \in I} \alpha_{sa^m}^m \right| \leq |\alpha_{sa^m}^m| < \delta(\epsilon) \leq \delta_3(\epsilon).$$

Thus,

$$\begin{aligned} &\left| \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s, \boldsymbol{\omega}^i) - \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{y}_s, \boldsymbol{\theta}^i) \right| \leq \\ &K \sum_{a \in A} \left| \prod_{m=1}^N x_{sa^m}^m - \prod_{m=1}^N y_{sa^m}^m \right| + \beta_i \sum_{a \in A} \sum_{k=1}^M \left| \prod_{m=1}^N x_{sa^m}^m \omega_k^i - \prod_{m=1}^N y_{sa^m}^m \theta_k^i \right| < \frac{\epsilon}{3} + 2\frac{\epsilon}{3} = \epsilon. \end{aligned}$$

□

Lemma 1 proves the equicontinuity of the set of functions  $\{\psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s, \boldsymbol{\omega}^i), \tilde{C}_s \in C_s, \tilde{P}_s \in P_s\}$ . This is a key result in our existence proof. It is first used to show the continuity of the function  $f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$  in Lemma 2. The continuity of the function  $f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$  is then used in the main existence theorem below (Theorem 4). Lemma 1 is also needed to establish the upper semi-continuity result used in Theorem 4.

Lemma 2 below follows from the basic real analysis result that states that the pointwise maximum of a family of equicontinuous functions is continuous. We also provide a detailed proof for our specific functions  $f_s^i$  and  $\psi_s^i$  in the appendix for completeness. Lemma 3 below follows from the definition of  $f_s^i$ .

LEMMA 2. *The function  $f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$  is continuous in all of its variables  $\forall i \in I$ , and  $s \in S$ .*

LEMMA 3.  *$f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$  is convex in  $\mathbf{u}_s^i$  for fixed  $\mathbf{x}_s^{-i}$  and  $\boldsymbol{\omega}^i$ .*

The following two technical results are the final ingredient needed to show the upper semi-continuity of the correspondence  $\phi(\mathbf{x})$ . Proof of Lemma 4 below follows directly from Fink (1964) and Lemma 1 above. Proof of Lemma 5 follows directly from Lemma 4 as shown in Fink (1964). These proofs are presented using our notation in the Appendix. Lemma 4 is used to prove Lemma 5, and Lemma 5 is used to show the upper semi-continuity result required by Kakutani's fixed point theorem.

LEMMA 4.  $\gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)$  is continuous on  $X_s^{-i}$ . Furthermore, the set  $\{\gamma_s^i(\cdot, \boldsymbol{\omega}^i) | \boldsymbol{\omega}^i \text{ is bounded}\}$  is equicontinuous.

LEMMA 5. If  $\mathbf{x}^{-i,n} \rightarrow \mathbf{x}^{-i}$  and  $\tau_s^i(\mathbf{x}^{-i,n}) \rightarrow \omega_s^i$  as  $n \rightarrow \infty$ , then  $\tau_s^i(\mathbf{x}^{-i}) = \omega_s^i$ .

We are now ready to prove the main result of this section.

THEOREM 4. (Existence of Equilibrium in Robust Stochastic Games)

Suppose that uncertain transition probabilities and payoffs in a discounted robust stochastic game belong to compact sets and that the set of actions and players, who use stationary strategies, are finite. Then, an equilibrium point of this game exists.

**Proof.**

By Lemma 2,  $f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$  is continuous in its variables. Since the minimum of this continuous function on a compact set  $X_s^i$  exists,  $\arg\min_{\mathbf{u}_s^i \in X_s^i} f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) \neq \emptyset$ .

Also, by Theorem 2, the equality in the expression  $\omega_s^i = \min_{\mathbf{u}_s^i \in X_s^i} f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$  can be established. Therefore,  $\phi(\mathbf{x}) \neq \emptyset$ . Note that by definition,  $\phi(\mathbf{x}) \subseteq X$ ,  $\forall \mathbf{x} \in X$ .

Next, we show that  $\phi(\mathbf{x})$  is a convex set. Suppose that  $(\mathbf{z}^1, \dots, \mathbf{z}^N), (\mathbf{v}^1, \dots, \mathbf{v}^N) \in \phi(\mathbf{x}^1, \dots, \mathbf{x}^N)$ . Then,  $\forall \mathbf{u}_s^i$ , and  $s \in S, i \in I$ ,  $\omega_s^i = f_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^i; \boldsymbol{\omega}^i) = f_s^i(\mathbf{x}_s^{-i}, \mathbf{v}_s^i; \boldsymbol{\omega}^i) \leq f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$ . Hence, for any  $\lambda \in [0, 1]$  and  $\forall i \in I, s \in S$ ,

$$\omega_s^i = \lambda f_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^i; \boldsymbol{\omega}^i) + (1 - \lambda) f_s^i(\mathbf{x}_s^{-i}, \mathbf{v}_s^i; \boldsymbol{\omega}^i) \leq f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$$

By the convexity of  $f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$ , we obtain

$$f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) \geq \omega_s^i = \lambda f_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^i; \boldsymbol{\omega}^i) + (1 - \lambda) f_s^i(\mathbf{x}_s^{-i}, \mathbf{v}_s^i; \boldsymbol{\omega}^i) \geq f_s^i(\mathbf{x}_s^{-i}, ((\lambda)\mathbf{z}_s^i + (1 - \lambda)\mathbf{v}_s^i); \boldsymbol{\omega}^i) \geq \omega_s^i,$$

and hence,  $(\lambda)(\mathbf{z}^1, \dots, \mathbf{z}^N) + (1 - \lambda)(\mathbf{v}^1, \dots, \mathbf{v}^N) \in \phi(\mathbf{x}^1, \dots, \mathbf{x}^N)$ .

Finally, we must show that  $\phi(\mathbf{x})$  is an upper semi-continuous correspondence. Suppose  $\mathbf{x}_n \rightarrow \mathbf{x}$ ,  $\mathbf{y}_n \rightarrow \mathbf{y}$ , and  $\mathbf{y}_n \in \phi(\mathbf{x}_n)$ . Taking a subsequence, we can consider  $\tau_s^i(\mathbf{x}^{-i,n}) \rightarrow \omega_s^i$ . Using the triangle inequality, we have  $\forall i \in I, s \in S$  that

$$\begin{aligned} |f_s^i(\mathbf{x}_s^{-i}, \mathbf{y}_s^i; \boldsymbol{\omega}^i) - \omega_s^i| &\leq |f_s^i(\mathbf{x}_s^{-i}, \mathbf{y}_s^i; \boldsymbol{\omega}^i) - f_s^i(\mathbf{x}_s^{-i,n}, \mathbf{y}_s^{i,n}; \boldsymbol{\tau}^i(\mathbf{x}^{-i,n}))| + |f_s^i(\mathbf{x}_s^{-i,n}, \mathbf{y}_s^{i,n}; \boldsymbol{\tau}^i(\mathbf{x}^{-i,n})) - \omega_s^i| \\ &= |f_s^i(\mathbf{x}_s^{-i}, \mathbf{y}_s^i; \boldsymbol{\omega}^i) - f_s^i(\mathbf{x}_s^{-i,n}, \mathbf{y}_s^{i,n}; \boldsymbol{\tau}^i(\mathbf{x}^{-i,n}))| + |\tau_s^i(\mathbf{x}^{-i,n}) - \omega_s^i| \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

Therefore,  $\omega_s^i = f_s^i(\mathbf{x}_s^{-i}, \mathbf{y}_s^i; \boldsymbol{\omega}^i)$ . By Lemma 5, we also have  $\tau_s^i(\mathbf{x}^{-i}) = \omega_s^i$ . Thus, we obtain that

$$\omega_s^i = f_s^i(\mathbf{x}_s^{-i}, \mathbf{y}_s^i; \boldsymbol{\omega}^i) = \tau_s^i(\mathbf{x}^{-i}) = \min_{\mathbf{u}_s^i \in X_s^i} f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i).$$

Therefore,  $\mathbf{y} \in \phi(\mathbf{x})$ , completing the proof that  $\phi$  is an upper semi-continuous correspondence. The fact that  $\phi(\mathbf{x})$  is a closed set for any  $\mathbf{x}$  follows from the fact that it is an upper-semi continuous correspondence. Therefore,  $\phi$  satisfies the assumptions of Kakutani's fixed point theorem.  $\square$

#### 4. A Multilinear System Formulation for Robust Markov Perfect Equilibria

Now that we have proved the existence of an equilibrium point in a discounted robust stochastic game, our next step is to calculate such a point. We will show that when the uncertainty in the probability transition data of the game belongs to a polytope intersected with the probability simplex, the problem of finding an equilibrium point could be cast as a multilinear system formulation. The characterization result we present here generalizes a previous result for normal form games by Aghassi and Bertsimas (2006) to stochastic games. For simplicity, we only consider uncertainty in

the transition data in constructing a formulation that characterizes equilibrium points. An analogous approach can be used to consider uncertainty both in payoffs and the transition data of a game.

Recall the definition of a robust Markov perfect equilibrium and conditions (4). These conditions are equivalent to the requirement that  $\forall i \in I, s \in S, \exists q_s^i \in \mathfrak{R}$  such that  $(\mathbf{x}_s^i, q_s^i)$  is an optimizer of the following robust mathematical program  $P_R$  with the objective value at optimality being equal to  $\omega_s^i$ :

$$P_R := \{\omega_s^i = \min_{u_s^i, q_s^i} q_s^i : q_s^i \geq \max_{\tilde{P}_s \in P_s} \psi_s^i(C_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \omega^i), \mathbf{1} \mathbf{u}_s^i = 1, \mathbf{u}_s^i \geq \mathbf{0}\}.$$

Here,  $(\mathbf{x}^{-i}, \omega^i)$  is treated as data. Define the uncertain probability transition matrix induced by a strategy  $(\mathbf{x}^{-i}, \mathbf{u}^i)$ :

$$\tilde{\mathbf{P}}(\mathbf{x}^{-i}, \mathbf{u}^i) = \left[ \sum_{a \in A} \prod_{\substack{m=1 \\ m \neq i}}^N x_{sa}^m u_{sa}^i \tilde{P}_{sak} \right]_{s=1, k=1}^{M, M}.$$

Denote the  $s^{th}$  row of  $\tilde{\mathbf{P}}(\mathbf{x}^{-i}, \mathbf{u}^i)$  by  $\tilde{\mathbf{p}}_s(\mathbf{x}^{-i}, \mathbf{u}^i)$ . Let  $\tilde{\mathbf{p}}_s$  denote the uncertain transition probability vector associated with the starting state  $s$ , that is,  $\tilde{\mathbf{p}}_s = [\tilde{P}_{sak}]_{a \in A; k \in S}$ . Let  $\mathbf{1}$  be a vector of ones of appropriate dimension. Let  $\mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \in \mathfrak{R}^{j^N}$  denote the matrix a row of which is given by the vector

$$\left[ \prod_{\substack{m=1 \\ m \neq i}}^N x_{sa}^m C_{s, a^{-i}, a}^i \right]_{a^i \in 1, \dots, j}.$$

Note that we have the following requirement in  $P_R$ :

$$q_s^i \geq \max_{\tilde{P}_s \in P_s} \psi_s^i(C_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \omega^i) = \max_{\tilde{\mathbf{p}}_s} \beta \tilde{\mathbf{p}}_s(\mathbf{x}^{-i}, \mathbf{u}^i) \omega^i + \mathbf{1}^T \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{u}_s^i \quad (9)$$

We assume that for any alternative combination of the players, the uncertain transition probabilities belong to a polytope intersected with the probability simplex.

Let  $\mathbf{Q}_s \in \mathfrak{R}^{j^{2NM}}$  be a matrix of 0s and 1s (such that each of its rows that corresponds to a pure strategy tuple  $\mathbf{a} \in A$  satisfies  $\sum_{k \in S} \tilde{P}_{sak} = 1$ ). In other words, we assume that the transition probabilities belong to the following uncertainty set:  $P = \{\tilde{\mathbf{p}}_s, s \in S : \mathbf{A}_s \tilde{\mathbf{p}}_s \geq \mathbf{b}_s, \mathbf{Q}_s \tilde{\mathbf{p}}_s = \mathbf{1}, \tilde{\mathbf{p}}_s \geq \mathbf{0}\}$ , where  $\mathbf{A}_s \in \mathfrak{R}^{Mj^N}$ .

Consider the maximization problem in  $P_R$ , where  $(\mathbf{x}_s^{-i}, \mathbf{u}_s^i, \omega^i)$  is regarded as data. Given that the uncertainty set is as stated, for fixed  $(\mathbf{x}_s^{-i}, \mathbf{u}_s^i, \omega^i)$ , this maximization problem is equivalent to the following LP:

$$\{\max_{\tilde{\mathbf{p}}_s} \beta \tilde{\mathbf{p}}_s(\mathbf{x}^{-i}, \mathbf{u}^i) \omega^i : \mathbf{A}_s \tilde{\mathbf{p}}_s \geq \mathbf{b}_s, \mathbf{Q}_s \tilde{\mathbf{p}}_s = \mathbf{1}, \tilde{\mathbf{p}}_s \geq \mathbf{0}\}. \quad (10)$$

Define the column vector  $\mathbf{z}_s^i = [\prod_{\substack{m=1 \\ m \neq i}}^N x_{sa}^m u_{sa}^i \omega_k^i]_{a^m \in A_s^m; a^i \in A_s^i; k \in S}$  such that the indices of  $\mathbf{z}_s^i$  match the ones of  $\tilde{\mathbf{p}}_s$ . Let  $\mathbf{Y}_s^i(\mathbf{x}_s^{-i}, \omega^i) \in \mathfrak{R}^{j^{N+1}M}$  be the matrix such that  $\mathbf{Y}_s^i(\mathbf{x}_s^{-i}, \omega^i) \mathbf{u}_s^i = \mathbf{z}_s^i$ . Let  $\mathbf{m}_s^i$  and  $\mathbf{n}_s^i$  be the dual variable vectors of problem (10). The dual of problem (10) is

$$\{\min_{\mathbf{m}_s, \mathbf{n}_s} [\mathbf{b}_s]'; [\mathbf{1}]' \left[ \begin{matrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{matrix} \right] : \left[ \begin{matrix} \mathbf{A}_s \\ \mathbf{Q}_s \end{matrix} \right]' \left[ \begin{matrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{matrix} \right] \geq \beta \mathbf{Y}_s^i(\mathbf{x}_s^{-i}, \omega^i) \mathbf{u}_s^i, \mathbf{m}_s^i \leq \mathbf{0}\}. \quad (11)$$

By the definition of our uncertainty set, problem (10) is feasible and it is clear that it is bounded. By strong duality, problem (11) is bounded and feasible and its optimal objective value is equal to

that of problem (10). Therefore, if  $(\mathbf{x}^{-i}, \mathbf{u}^i, \boldsymbol{\omega}^i)$  satisfies condition (9), then (9) is equivalent to the condition that  $\exists \mathbf{m}_s^i \in \mathbb{R}^l$  and  $\exists \mathbf{n}_s^i \in \mathbb{R}^{j^N}$  such that:

$$\begin{aligned} q_s^i - \mathbf{1}^T \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{u}_s^i &\geq [\mathbf{b}_s]'; [\mathbf{1}]' \begin{bmatrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{bmatrix} \\ \begin{bmatrix} \mathbf{A}_s \\ \mathbf{1}_s \end{bmatrix}' \begin{bmatrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{bmatrix} &\geq \beta \mathbf{Y}_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) \mathbf{u}_s^i \\ \mathbf{m}_s^i &\leq 0 \end{aligned} \quad (12)$$

Conversely, if condition (12) is satisfied, then problem (11) is feasible. Then by weak duality, any feasible solution  $[\mathbf{b}_s]'; [\mathbf{1}]' \begin{bmatrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{bmatrix}$  of problem (11) is greater than or equal to any solution  $\beta \tilde{\mathbf{p}}_s(\mathbf{x}^{-i}, \mathbf{u}^i) \boldsymbol{\omega}^i$  of problem (10), so,

$$q_s^i - \mathbf{1}^T \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{u}_s^i \geq [\mathbf{b}_s]'; [\mathbf{1}]' \begin{bmatrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{bmatrix} \geq \max_{\tilde{\mathbf{p}}_s} \beta \tilde{\mathbf{p}}_s(\mathbf{x}^{-i}, \mathbf{u}^i) \boldsymbol{\omega}^i.$$

Therefore conditions (9) and (12) are equivalent. This proves:

LEMMA 6. *Condition (9) is equivalent to condition (12).*

Let

$$\mathbf{T}^i(\mathbf{x}) = \left[ \sum_{a \in A} \prod_{m=1}^N x_{sam}^m t_{sak}^i \right]_{s=1, k=1}^{M, M},$$

and denote the  $s^{th}$  row of  $\mathbf{T}^i(\mathbf{x})$  by  $\mathbf{t}_s^i(\mathbf{x})$ . Let  $\mathbf{t}_s^i$  denote the variables representing the transition probabilities adopted by player  $i$  according to  $i$ 's worst-case perspective, associated with the starting state  $s$ . That is,  $\mathbf{t}_s^i = [t_{sak}^i]_{a \in A; k \in S}$ .

THEOREM 5. *A stationary strategy  $\mathbf{x}$  is a robust Markov perfect equilibrium point with the robust value vector  $\boldsymbol{\omega}^i$ , iff  $\forall i \in I, s \in S, \exists \mathbf{m}_s^i \in \mathbb{R}^l, \mathbf{n}_s^i \in \mathbb{R}^{j^N}, \mathbf{t}_s^i \in \mathbb{R}^{Mj^N}$  such that for  $k = 1, \dots, j$ ,  $(\boldsymbol{\omega}^i, \mathbf{x}_s, \mathbf{m}_s^i, \mathbf{n}_s^i, \mathbf{t}_s^i)$  satisfies*

$$\boldsymbol{\omega}_s^i = \mathbf{1}^T \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{x}_s^i + \beta \mathbf{t}_s^i(\mathbf{x}) \boldsymbol{\omega}^i$$

$$[\mathbf{e}_{sk}^i]' \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{1} + \beta [\mathbf{e}_{sk}^i]' \mathbf{Y}_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) \mathbf{t}_s^i \geq \boldsymbol{\omega}_s^i \geq [\mathbf{b}_s]'; [\mathbf{1}]' \begin{bmatrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{bmatrix} + \mathbf{1}^T \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{x}_s^i$$

$$\begin{bmatrix} \mathbf{A}_s \\ \mathbf{Q}_s \end{bmatrix}' \begin{bmatrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{bmatrix} - \beta \mathbf{Y}_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) \mathbf{x}_s^i \geq 0$$

$$\mathbf{1} \mathbf{x}_s^i = 1, \quad \mathbf{m}_s^i \leq 0, \quad \mathbf{x}_s^i \geq 0, \quad \mathbf{A}_s \mathbf{t}_s^i \geq \mathbf{b}_s, \quad \mathbf{Q}_s \mathbf{t}_s^i = \mathbf{1}.$$

**Proof.**

Recall problem  $P_R$ . By Lemma 6, if  $\mathbf{x}$  is a robust Markov perfect equilibrium point,  $\forall i \in I, s \in S, \exists q_s^i \in \mathbb{R}, \mathbf{m}_s^i \in \mathbb{R}^l, \mathbf{n}_s^i \in \mathbb{R}^{j^N}$  such that  $(\mathbf{x}_s^i, q_s^i, \mathbf{m}_s^i, \mathbf{n}_s^i)$  is an optimizer of

$$\begin{aligned} \boldsymbol{\omega}_s^i &= \min_{\mathbf{u}_s^i, q_s^i, \mathbf{m}_s^i, \mathbf{n}_s^i} q_s^i \\ q_s^i - \mathbf{1}^T \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{u}_s^i &\geq [\mathbf{b}_s]'; [\mathbf{1}]' \begin{bmatrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{bmatrix} \\ \begin{bmatrix} \mathbf{A}_s \\ \mathbf{Q}_s \end{bmatrix}' \begin{bmatrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{bmatrix} &\geq \beta \mathbf{Y}_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) \mathbf{u}_s^i, \quad \mathbf{m}_s^i \leq 0, \quad \mathbf{1} \mathbf{u}_s^i = 1, \quad \mathbf{u}_s^i \geq 0 \end{aligned} \quad (13)$$

Let  $\mathbf{e}_{sk}^i$  be the  $k^{th}$  unit vector. Dual of the above is:

$$\max_{\nu_s^i, \mathbf{t}_s^i} \nu_s^i : \mathbf{A}_s \mathbf{t}_s^i \geq \mathbf{b}_s, \quad \mathbf{Q}_s \mathbf{t}_s^i = \mathbf{1}, \quad \nu_s^i \leq [\mathbf{e}_{sk}^i]' \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{1} + \beta [\mathbf{e}_{sk}^i]' \mathbf{Y}_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) \mathbf{t}_s^i, k = 1, \dots, j. \quad (14)$$

The statement in the theorem follows from strong duality and Theorem 2. For the other direction, now suppose a given  $\mathbf{x}_s, \boldsymbol{\omega}^i$ , and that  $\forall i \in I, s \in S, \exists \mathbf{m}_s^i \in \mathcal{R}^I, \mathbf{n}_s^i \in \mathcal{R}^{j^N}, \mathbf{t}_s^i \in \mathcal{R}^{Mj^N}$ , such that  $(\boldsymbol{\omega}^i, \mathbf{x}_s, \mathbf{m}_s^i, \mathbf{n}_s^i, \mathbf{t}_s^i)$  satisfies the above system. Let

$$\nu_s^i = \min_{k \in 1, \dots, j} [\mathbf{e}_{sk}^i]' \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{1} + \beta [\mathbf{e}_{sk}^i]' \mathbf{Y}_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) \mathbf{t}_s^i, \quad k = 1, \dots, j,$$

$$q_s^i = \left[ [\mathbf{b}_s]'; [\mathbf{1}]' \right] \begin{bmatrix} \mathbf{m}_s^i \\ \mathbf{n}_s^i \end{bmatrix} + \mathbf{1}^T \mathbf{E}_s^i(\mathbf{x}_s^{-i}, C^i) \mathbf{u}_s^i$$

Then, for  $(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)$ ,  $(\mathbf{x}_s^i, q_s^i, \mathbf{m}_s^i, \mathbf{n}_s^i)$  is feasible for problem (13), and  $(\nu_s^i, \mathbf{t}_s^i)$  is feasible for problem (14) with  $\nu_s^i \geq q_s^i$ . By weak duality  $\nu_s^i \leq q_s^i$ , so  $\nu_s^i = q_s^i$ . Hence,  $(\mathbf{x}_s^i, q_s^i, \mathbf{m}_s^i, \mathbf{n}_s^i)$  is optimal for problem (13). Hence,  $(\mathbf{x}_s^i, q_s^i)$  is optimal in  $P_R$ . Therefore,  $\mathbf{x}$  is an equilibrium point of the discounted robust stochastic game.  $\square$

Theorem 5 states that when the uncertainty in the probability transition data of the game belongs to a polytope intersected with the probability simplex, the set of robust Markov perfect equilibria can be characterized as a multilinear system formulation.

Due to the nonlinearity in the Bellman-type optimality equation (1), solving an  $n$ -person nonzero sum complete information discounted game is regarded as a difficult task (Filar and Vrieze (1997)). Our formulation in Theorem 5 is not easier compared to the formulations for nominal stochastic games. In particular, for polytopic uncertainty sets, Theorem 5 states that solving the robust problem does not incur a significant additional computational effort. This is also the case in robust Markov decision processes, when the uncertainty set on the transition data is defined as a polytope (Nilim and Ghaoui (2005)). In the next section we solve Theorem 5 to compute a sample Markov perfect equilibrium point using an existing solver (LOQO) for nonlinear, nonconvex problems (Vanderbei (2006)). Solving the problem given by Theorem 5 requires solving an MPEC, hence, this approach can be used to solve large scale problems to the extent that we can solve large scale MPEC problems. However, MPECs can be notoriously difficult to solve and an active area of research is the search for efficient algorithms. A survey of algorithms for different kinds of complete information stochastic game models is presented in Raghavan and Filar (1991). New methods that can be used to solve MPEC problems are presented in Ferris et al. (2005).

In the next section, we illustrate the use of discounted robust stochastic games in a queueing control problem.

## 5. A Queueing Control Application

Game theoretical analysis has been widely applied to queueing control problems (Altman and Shimkin (1998), Heyman (1968), Sobel (1969), Stidham and Weber (1989), Yechiali (1971), Altman (1994b), Altman (1994a), Altman and Hordijk (1995)). In this section, we present an application of our robust model for incomplete information stochastic games to a single server exponential queueing system.

Consider a queueing control problem encountered in packet switched networks. The most well-known packet switched networks are the Internet and local area networks (Peterson (2007)). For instance, in packet switched networks, packets (blocks of data) are routed among nodes over data links shared with other traffic. Here, a node is a server connected to the network. For example, a server in a network can be a computer, or a personal digital assistant. The service rate of a server is controlled by a service provider (player 1) (for instance, by increasing or decreasing the processing

capacity of a server at a node). Although the service rate can be set to different levels, it may change in time in a way unpredictable to the router. Uncertainty in service times is a common assumption in the flow control literature (see Altman and Shimkin (1998), Altman (1994b), Altman (1994a), Altman and Hordijk (1995), Altman and Shimkin (1993)). The reasons for this can be the imprecision during the implementation of an intended service rate due to operating conditions, service conditions, presence of other packets in the system, and/or server malfunctions. In our model, having uncertainty in the service rates results in uncertainty in the transition probabilities. This affects both players and the players combat with the uncertainty using a robust optimization approach.

In packet switched networks, packets are routed by a programmable physical device, called a router (player 2). A router dynamically controls the flow of arriving packets into a finite buffer at a server. The rates that the service provider and the router choose depend on the number of packets in the system. This allows them to choose rates so that they can address congestion and throughput in the system (Altman (1994b)). In fact, it is to the benefit of a service provider to increase the throughput (amount of packets processed). However, increase in throughput may result in an increase in packets' waiting times in the buffer (called latency) and routers are programmed to operate in a way to minimize packets' waiting times. Due to the the service provider and the router having conflicting objectives, the game theoretical nature of the problem arises. We model this problem as a zero-sum stochastic game, where a player is incurred payoffs that are modeled as being paid to the other player. Next we present the details of our model.

Let  $X_t$  represent the number of packets in the system at time  $t$ ,  $t = 0, 1, 2, \dots$ . Hence, if there are  $x$  packets in the system at time  $t$ , we have  $X_t = x$ . Only one packet can be in service at any time, with the remaining  $x - 1$  packets in the buffer, waiting for service. The state space is denoted by  $\mathbf{X} = \{0, 1, \dots, L\}$ , where  $L < \infty$  is the maximum number of packets that can be present in the system. The router admits one packet into the system at a time. When there are  $x$  packets present, the service provider and the router simultaneously set the service rate and the admission rate to  $\mu_x$  and  $\lambda_x$ , respectively. As mentioned above, we consider that  $\mu_x$  can deviate unpredictably. For instance, if the service provider intends to have a service rate of 1 packet per 20 seconds, in practice he may observe that the actual service rates vary between 1 packet per 19 seconds to 1 packet per 21 seconds. Suppose that there are  $x$  packets in the system and the service rate and the flow rate are set to  $\mu_x$  and  $\lambda_x$ , respectively. Then the router is incurred a holding cost  $h(x)$ , and a cost  $\theta(\mu_x, \lambda_x)$  associated with having packets served at rate  $\mu_x$  when it admits packets at rate  $\lambda_x$ . If there are no packets in the system, this cost represents the set-up cost of the server. These payoffs are modeled as being paid to the service provider, since the players' objectives are in conflict. The service provider, in turn, pays the router  $\rho(\mu_x, \lambda_x)$ , which represents the reward to the router for admitting a packet. It can also be interpreted as the set-up cost of the router.

We assume that the time until the admission of a new packet and the next service completion are both exponentially distributed with means  $1/\lambda_x$  and  $1/\mu_x$ , respectively, when there are  $x$  packets in the system. We can therefore represent the number of packets in the system with a birth and death process, which have the following state transition probabilities:

$$\tilde{p}(y|x, \mu_x, \lambda_x) = \begin{cases} \mu_x/(\lambda_x + \mu_x), & 1 \leq x \leq L, \quad y = x - 1 \\ \lambda_x/(\lambda_x + \mu_x), & 0 \leq x < L - 1, \quad y = x + 1 \\ 1, & x = 0, \quad y = 1 \\ 1, & x = L, \quad y = L - 1 \end{cases} \quad (15)$$

Although there is uncertainty only in the service rates, both players face uncertainty in the transition data among the states. This is so because the transition probabilities depend not only on the admission rates but also on the service rates. For instance, if the players choose  $\mu_x$  and

$\lambda_x$  in state  $x$ , with  $\mu_x$  in the form of an interval, then the transition probability to the next state would be minimum  $\lambda_x/(\lambda_x + \mu_x^{max})$  and maximum  $\lambda_x/(\lambda_x + \mu_x^{min})$ , where  $\mu_x^{max}$  and  $\mu_x^{min}$  represent the end points of the interval for  $\mu_x$ .

For every fixed strategy of the players in a state, the uncertainty in service rates results in interval uncertainty on the transition probabilities. Note that since this uncertainty is resolved after the players select their strategies, the entire transition probability interval is valid. Therefore, the uncertainty set on the transition probabilities out of each state is represented by a polyhedron formed by the intersection of these intervals and the probability simplex.

### 5.1. Numerical Example

#### *Problem Description*

We set up an example of the above problem with the router and the service provider having two pure strategies (two alternatives) for each state. Players are not confined only to use pure strategies, and they may choose to use any rate in between the two pure strategies. For simplicity, we keep the two pure strategies for each player the same for every state  $x$ . However, since the players can use mixed strategies, the rates they choose in effect may differ. Router's first alternative, denoted by  $\bar{\lambda}$ , is to admit a packet into the system every 10 seconds. Its second alternative,  $\underline{\lambda}$ , is to admit a packet every 25 seconds. Service provider's first alternative is to serve a packet every 11 seconds. We call this rate the intended service rate. For instance, due to the uncertainty, if the service rate is set to 11 seconds, it may vary between 10 to 12 seconds. We denote his first alternative by  $\bar{\mu}$ ; hence we have  $\bar{\mu} \in [1/12, 1/10]$ . Service provider's second alternative is to serve a packet every 20 seconds, which may vary between 19 to 21 seconds due to uncertainty. Therefore, we have  $\underline{\mu} \in [1/21, 1/19]$ . Hence, the *interval length* within which the service rate varies is 2 seconds per packet for each pure strategy.

In this example, we use an exponential holding cost  $h(x) = ab^{\alpha x}$ ,  $x \geq 1$  with  $a = 1.2$ ,  $b = 1.9$ , and  $\alpha = 0.2$ . The holding cost when there are no packets in the system is 0. For each state, we let  $\theta(\bar{\mu}, \bar{\lambda}) = \theta(\bar{\mu}, \underline{\lambda}) = 110$ ,  $\theta(\underline{\mu}, \bar{\lambda}) = \theta(\underline{\mu}, \underline{\lambda}) = 90$ ,  $\rho(\bar{\mu}, \bar{\lambda}) = 60$ ,  $\rho(\bar{\mu}, \underline{\lambda}) = 30$ ,  $\rho(\underline{\mu}, \bar{\lambda}) = 20$ ,  $\rho(\underline{\mu}, \underline{\lambda}) = 70$ .

This payoff scenario indicates that the router pays the service provider more when the service rate is higher. It also indicates that the reward that the router receives from an admitted packet is higher when both players choose their first alternatives (higher rates), or second alternatives (lower rates). In other words, the router receives less reward when the admission and service rates are inconsistent; that is, when the service rate is relatively higher compared to the admission rate, or vice versa.

We use  $L = 30$ . Alternatives for the router at the boundary state 30 represent the set-up options for the router so that it can control admissions at specified rates when the system is not full. When the system is full, packets are dropped, and  $\rho$  represents the set-up cost of the router. Similarly, the alternatives of the server at state 0 represent the set-up options. If there are no packets in the system,  $\theta$  represents the set-up cost of the server so that, given an admission rate, it can be controlled to operate at a required rate at the next state. When the system is full or empty, players still choose a strategy based on the set-up costs. However, their strategies at the boundary states 0 and 30 does not affect the transition probabilities out of these states, which are 1 according to formula (15). Therefore, there can be at most 30 packets in the system.

We generate four more instances of this problem by increasing the interval length for service rates to 6, 10, 14 and 18 seconds per packet, keeping the intended service rates the same. Intervals of service rates for each instance are depicted in Table 1. For instance, when the interval length is 10 seconds per packet (for the instance number 3), we have  $\bar{\mu} \in [1/16, 1/6]$ , and  $\underline{\mu} \in [1/25, 1/15]$  for states 1 through 30. Therefore, given the players choose an alternative pair in a given state, the difference between the minimum and maximum transition probabilities to another state becomes larger for larger interval lengths. For instance, when the interval length is 10, if the players choose

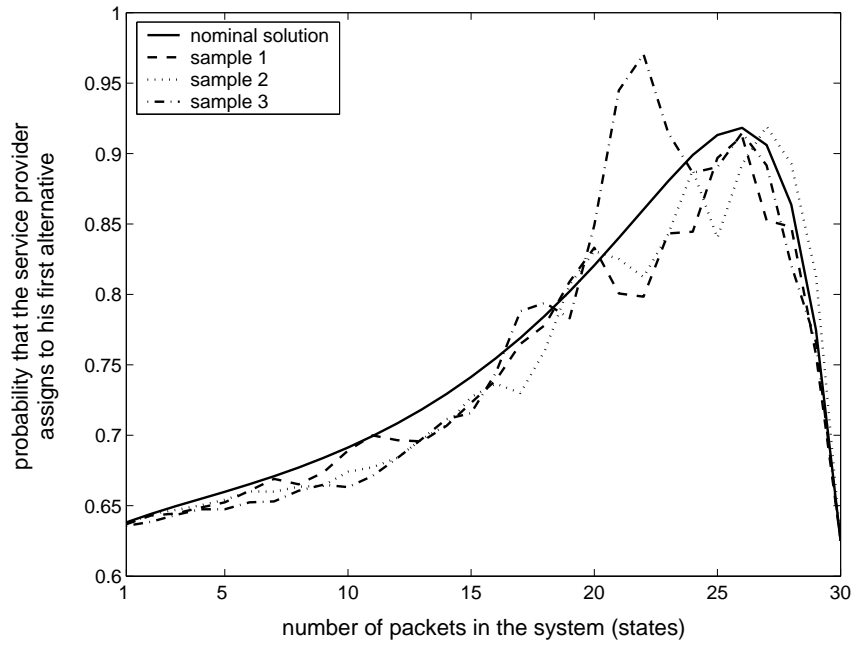


**Table 1** Intervals of rates for different instances

instance number	1	2	3	4	5
$\bar{\mu}$	[1/12,1/10]	[1/14,1/8]	[1/16,1/6]	[1/18,1/4]	[1/20,1/2]
$\underline{\mu}$	[1/21,1/19]	[1/23,1/17]	[1/25,1/15]	[1/27,1/13]	[1/29,1/11]

their second alternatives in a state, the transition probability to the next state would be minimum 0.3753, and maximum 0.5.

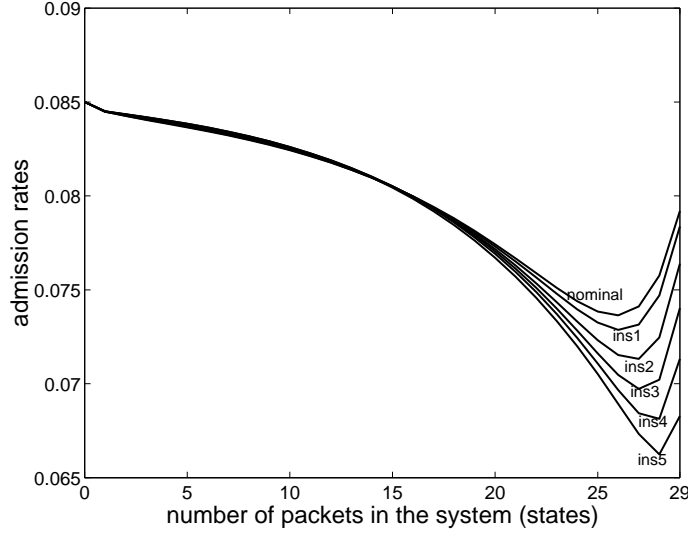
In Figure 1, for the nominal game and for three different sampled rates, we plot the server's equilibrium strategy for his first alternative. In the nominal game, no uncertainty is considered. The rates at every state are sampled from uniform distributions on the intervals in instance 3 ([1/16,1/6] and [1/25,1/15]). We observe that the equilibrium server strategy depends on the different samples significantly.

**Figure 1** Sensitivity of the nominal solution

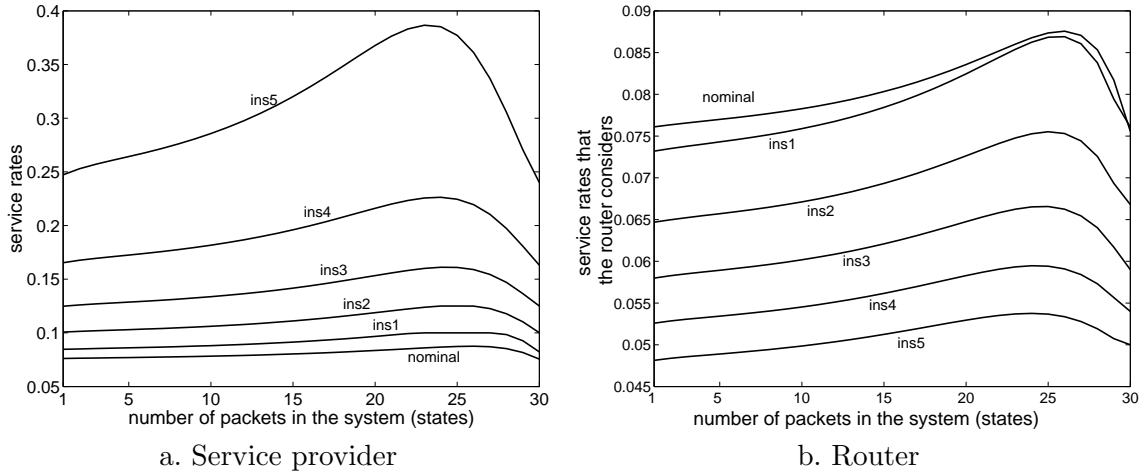
### *Equilibrium Solution*

Now we investigate how the robust and the nominal solutions differ for different instances of our problem. To this end, for each instance, we calculate the intervals that the transition probabilities belong to. We then solve each instance using Theorem 5 with the calculated lower and upper bounds on the transition data, the payoffs given above, and with the discount factor  $\beta = 0.95$ . A feature of Theorem 5 is that it yields the equilibrium strategies  $(x_s^i)$ , and the probabilities  $(t_s^i)$  from the uncertainty set that each player considers when playing the game.

Figure 2 depicts the equilibrium admission rates for the nominal and robust solutions. This admission rate is obtained by combining the two possible admission rates with the optimal router strategy. The service rate cannot be computed directly from the server strategy as it also depends on the uncertainty assumption made by each player, which influences the value of the service rate. Therefore, because of the different uncertainty assumptions of the router and server, each will see a different service rate. We present in Figure 3 the equilibrium service rates seen by the server and



**Figure 2** Admission rates



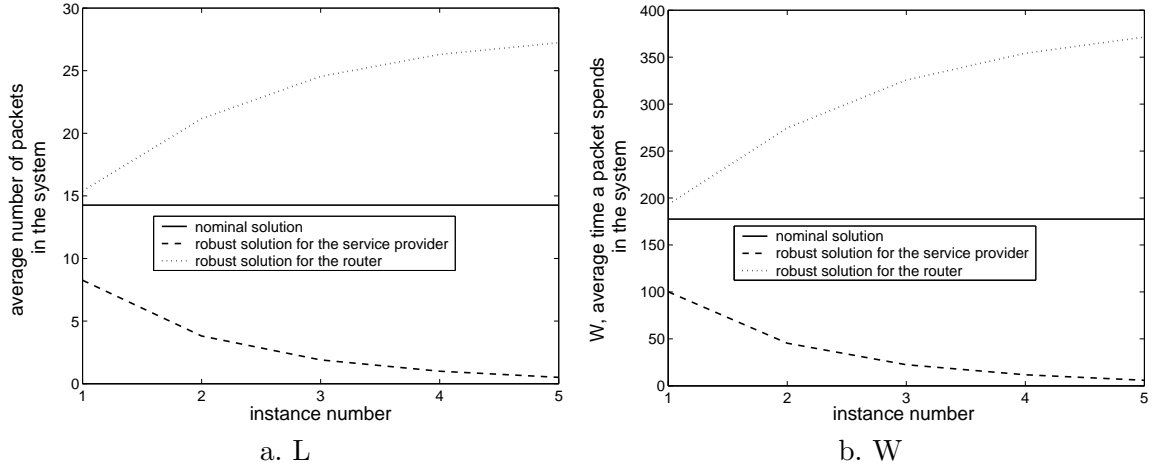
**Figure 3** Service rates seen by the service provider and the router

router, respectively. These rates are computed by using the formula of transition probabilities for a birth death process (15) along with the solution weighted transition probabilities of each player given by  $T^i(x)$  in Lemma 6.

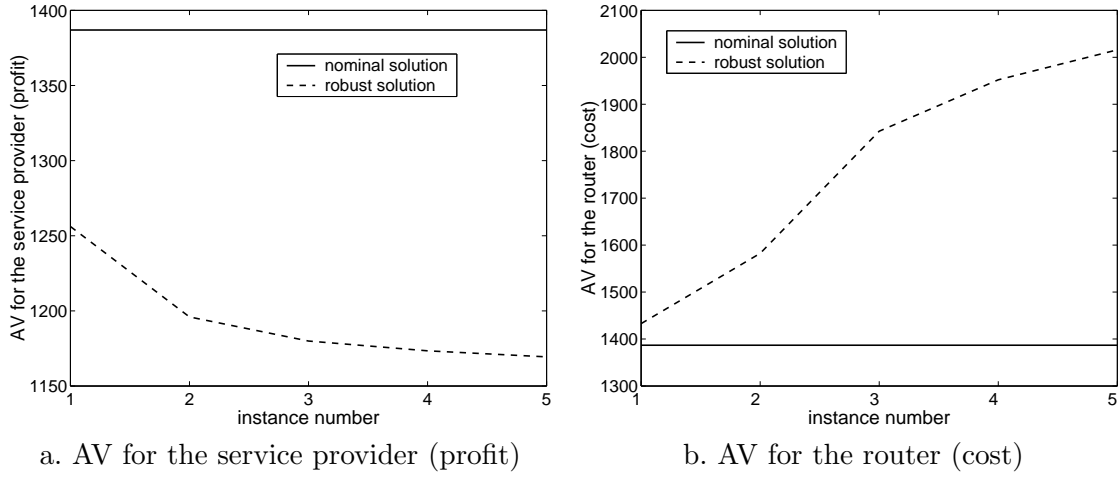
We note from these figures that the router considers that the service provider decreases the service rates as the uncertainty gets larger, and therefore tends to decrease the admission rates to protect himself against congestion. On the other hand, the service rates the service provider considers increase as the uncertainty set becomes larger, which is a pessimistic approach since this may result in a decrease in the queue size, and consequently, in a decrease in his overall profit.

Next, we calculate steady state probabilities, the average number of packets in the system ( $L$ ), the average amount of time a packet spends in the system ( $W$ ), and the average value ( $AV$ ) of the game from the point of view of each player.  $AV$  for a player is calculated by weighting the value of a state to a player by the respective steady state probabilities from that player's perspective, and taking the summation over the states.

The results are depicted in Figures 4 and 5. Figure 4a indicates that as far as the service provider is concerned, the average number of customers in the system is less than that of the nominal



**Figure 4** Average number of packets in the system (L) and average time a packet spends in the system (W)



**Figure 5** Average values for the players

solution when there is uncertainty in the system. From the router's perspective,  $L$  increases as the intervals becomes larger. Note that these are pessimistic points of views for both players since an increase in the number of packets in the system would be an advantage for the service provider, whereas it would be a disadvantage for the router. Figure 4b indicates that the average waiting time for a packet decreases from the service provider's perspective, and increases from the router's perspective. The reason for this is that in equilibrium, the service provider assumes the pessimistic perspective of having less customers in the system, whereas the router assumes the opposite. Accordingly, as the uncertainty sets get larger, AV of the game to the service provider (i.e. the overall profit that he makes) decreases, whereas AV to the router (i.e. the overall cost of the game to the router) increases. Note that, although this stochastic game is zero-sum, the resulting values differ for each player when they play robustly. This example illustrates that although the problem is a zero-sum game, formulations for zero-sum games cannot be used to solve discounted robust stochastic games, despite the fact that one player pays the other player a fixed amount.

### Simulation

For each instance, we fix the robust and nominal solution strategies for both players, simulate the service rates, and calculate the corresponding values by solving equations (16). Service rates are simulated from uniform distributions given by the corresponding interval for each alternative. Our sample size in the simulations is 600 for each instance. This study allows us to compare the performances of the robust and nominal solutions when the service rates are uniformly distributed on their respective intervals.

$$\omega_s^i = \sum_{a \in A} \prod_{m=1}^N x_{sa}^m \{C_{sa} + \beta \sum_{k=1}^M P_{sak} \omega_k^i\}. \quad (16)$$

For each sample, we calculate the average value of the game (AV). Means and variances of the AVs of each instance is depicted in Table 2.

Instance Number	Nominal		Robust		% Variance Reduction
	Mean	Variance	Mean	Variance	
1	1380.74	195.77	1373.81	186.39	4.8
2	1315.39	998.81	1303.74	865.94	13.3
3	1233.60	590.20	1227.66	475.41	19.5
4	1189.43	62.76	1188.36	53.70	14.4
5	1175.48	8.69	1175.29	8.67	0.2

**Table 2** Mean and Variance Values

Table 2 depicts that using robust strategies yields slightly lower means for each instance. Note that lower means less profit for the service provider. This could be interpreted in two ways: making less profit is to the disadvantage of the service provider. On the other hand, this also could be interpreted as an advantage to the service provider because robustness comes at an extra price, and the price he pays in this example is relatively minor compared to using nominal strategies. Similarly, from the router's perspective, using robust strategies yields slightly less overall costs and is to its advantage.

More important than the mean values, and what makes the robust solution's conservative strategies preferable in this example is the variances of the samples. The variances obtained by using robust strategies are lower than the variances obtained by using nominal strategies. This is to the benefit of both players as lower AV variances imply less sensitivity to perturbations in the data, or more stability in the system. We also observe for both players that although the uncertainty becomes larger, AV variances first increase for instances 1 and 2, and then decrease. The reason for this is that after the second instance, the steady state probabilities in the samples become higher for the initial states. We also notice that the variances for the values of states (obtained from Equation (16)) decrease as the number of states decrease. Because the steady state probabilities are higher in initial states, the AV variances are mostly obtained from the variances of the values of the initial states, and therefore they decrease.

## 6. Concluding Remarks and Future Research

In this paper, we consider n-player, non-zero sum discounted stochastic games in which none of the players knows the true transition probabilities and/or payoffs of the game and each player adopts a robust optimization approach to data uncertainty. We offer an alternative equilibrium concept for stochastic games with incomplete information. We propose a distribution-free model that lends itself to computational results via a multilinear system formulation that characterizes equilibrium

points. We finally illustrate the use of discounted robust stochastic games in a queueing control example.

We determined several properties of discounted robust stochastic games in this research. First, an equilibrium exists even if there exists players who do not adopt a robust optimization approach. This stems from the fact that when there are no uncertainty sets for the data of a stochastic game, best response functions are already continuous, as shown in Fink (1964). Hence, we can construct a correspondence that satisfies Kakutani's theorem and that includes players who may play non-robustly. Second, in our existence proof, we have assumed that both players have the common uncertainty set of payoffs  $C_s$  at each state. The results of this existence theorem hold, even if each player has different sets for uncertainty. Third, if there is uncertainty in the discounted stochastic game and the uncertainty is a common set for the players, the zero sum property of the equilibrium values do not necessarily hold. This is because if there is uncertainty in any data of the game, the players' approach to this uncertainty may differ. This is the case even if the game is zero-sum and there is uncertainty only in transition data. Therefore, formulations for zero-sum stochastic games could not be used for analyses and properties that pertain to zero-sum games cannot be expected to hold in the presence of uncertainty.

In the example of Section 5, the robust solution's conservative strategies yield desirable properties compared to those of the nominal solution's. This may not always be the case since a robust solution can cause the players to adopt overly conservative strategies or can have instances that are not sensitive to data uncertainty. When players play overly conservative and the service rates are sampled from large intervals, there may be substantial discrepancies between the sampled rates and the rates according to which the players play. Therefore, high variances can be observed as a result of overly conservative strategies. Future research should investigate in more detail the problem conditions and/or a-priori measures to identify when a robust formulation of a stochastic game is preferable. Another research direction would be taking the players' risk attitudes into account and investigating the player behavior for which a robust approach is most beneficial. The extension of the ideas in this paper to limiting average stochastic games would also be a promising future research direction.

## Appendix

**Theorem** (Banach's Contraction Mapping Theorem).

Let  $(W, \rho)$  be a complete metric space and let  $\gamma: W \rightarrow W$  be a contraction mapping. Then there exists a unique fixed point of the function  $\gamma$ .

### Lemma 2

The function  $f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$  is continuous in all of its variables  $\forall i \in I$ , and  $s \in S$ .

#### Proof.

Recall that  $f_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) = \max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$ .

Let  $\max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) = a$ .

Let  $\mathbf{p} = (\mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i)$ ,  $\mathbf{q} = (\mathbf{y}_s^{-i}, \mathbf{z}_s^i; \boldsymbol{\theta}^i)$ . Lemma 1 states that given  $\epsilon > 0$ ,  $\exists \delta(\epsilon) > 0$  such that for any  $\mathbf{p}, \mathbf{q} \in X_s \times W^i$ , with  $d_1(\mathbf{p}, \mathbf{q}) < \delta(\epsilon)$ , then,  $\forall \tilde{C}_s \in C_s$ ,  $\forall \tilde{P}_s \in P_s$ ,  $|\psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) - \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{y}_s^{-i}, \mathbf{z}_s^i; \boldsymbol{\theta}^i)| < \epsilon$ .

Therefore,  $\forall \tilde{C}_s \in C_s$ ,  $\forall \tilde{P}_s \in P_s$ ,  $\psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{y}_s^{-i}, \mathbf{z}_s^i; \boldsymbol{\theta}^i) \leq \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) + \epsilon \leq a + \epsilon$ .

Hence,  $\max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{y}_s^{-i}, \mathbf{z}_s^i; \boldsymbol{\theta}^i) \leq a + \epsilon$ .

For the other direction, take  $\tilde{C}_s^0 \in C_s$  and  $\tilde{P}_s^0 \in P_s$  such that  $\psi_s^i(\tilde{C}_s^0, \tilde{P}_s^0; \mathbf{x}_s^{-i}, \mathbf{u}_s^i, \boldsymbol{\omega}^i) \geq a - \epsilon/2$ . From Lemma 1, we have that since  $d_1(\mathbf{p}, \mathbf{q}) < \delta(\epsilon)$ ,  $\left| \psi_s^i(\tilde{C}_s^0, \tilde{P}_s^0; \mathbf{y}_s^{-i}, \mathbf{z}_s^i, \boldsymbol{\theta}^i) - \psi_s^i(\tilde{C}_s^0, \tilde{P}_s^0; \mathbf{x}_s^{-i}, \mathbf{u}_s^i, \boldsymbol{\omega}^i) \right| < \epsilon$ . Then,

$$\max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{y}_s^{-i}, \mathbf{z}_s^i, \boldsymbol{\theta}^i) \geq \psi_s^i(\tilde{C}_s^0, \tilde{P}_s^0; \mathbf{y}_s^{-i}, \mathbf{z}_s^i, \boldsymbol{\theta}^i) \geq \psi_s^i(\tilde{C}_s^0, \tilde{P}_s^0; \mathbf{x}_s^{-i}, \mathbf{u}_s^i, \boldsymbol{\omega}^i) - \epsilon \geq a - 3\epsilon/2$$

In conclusion, we have that  $\forall \epsilon > 0, \exists \delta(\epsilon)$  such that if  $d_1(\mathbf{p}, \mathbf{q}) < \delta(\epsilon)$  then

$$a - 2\epsilon \leq a - 3\epsilon/2 \leq \max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{y}_s^{-i}, \mathbf{z}_s^i, \boldsymbol{\theta}^i) \leq a + \epsilon \leq a + 2\epsilon,$$

which completes the proof.  $\square$

**Lemma 4.** (Fink (1964))

$\gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)$  is continuous on  $X_s^{-i}$ . Furthermore, the set  $\{\gamma_s^i(\cdot, \boldsymbol{\omega}^i) | \boldsymbol{\omega}^i \text{ is bounded}\}$  is equicontinuous.

**Proof.**

Let

$$\begin{aligned} \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) &= \psi_s^i(C_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}), P_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}); \mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}; \boldsymbol{\omega}^i), \\ \gamma_s^i(\mathbf{y}_s^{-i}, \boldsymbol{\omega}^i) &= \psi_s^i(C_s^i(\mathbf{y}_s^{-i}, \mathbf{z}_s^{*i}), P_s^i(\mathbf{y}_s^{-i}, \mathbf{z}_s^{*i}); \mathbf{y}_s^{-i}, \mathbf{z}_s^{*i}; \boldsymbol{\omega}^i). \end{aligned}$$

Furthermore,

$$\begin{aligned} \gamma_s^i(\mathbf{y}_s^{-i}, \boldsymbol{\omega}^i) - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) &\leq \psi_s^i(C_s^i(\mathbf{y}_s^{-i}, \mathbf{u}_s^{*i}), P_s^i(\mathbf{y}_s^{-i}, \mathbf{u}_s^{*i}); \mathbf{y}_s^{-i}, \mathbf{u}_s^{*i}; \boldsymbol{\omega}^i) \\ &\quad - \psi_s^i(C_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}), P_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}); \mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}; \boldsymbol{\omega}^i), \\ \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) - \gamma_s^i(\mathbf{y}_s^{-i}, \boldsymbol{\omega}^i) &\leq \psi_s^i(C_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}), P_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}); \mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}; \boldsymbol{\omega}^i) \\ &\quad - \psi_s^i(C_s^i(\mathbf{y}_s^{-i}, \mathbf{z}_s^{*i}), P_s^i(\mathbf{y}_s^{-i}, \mathbf{z}_s^{*i}); \mathbf{y}_s^{-i}, \mathbf{z}_s^{*i}; \boldsymbol{\omega}^i). \end{aligned}$$

If  $\boldsymbol{\omega}^i$  is restrained to be in a bounded region, then the right hand sides can be made uniformly small because of the uniform continuity of  $\psi_s^i$  on compact sets.  $\square$

**Lemma 5.** (Fink(1964)) If  $\mathbf{x}^{-i,n} \rightarrow \mathbf{x}^{-i}$  and  $\tau_s^i(\mathbf{x}^{-i,n}) \rightarrow \omega_s^i$  as  $n \rightarrow \infty$ , then  $\tau_s^i(\mathbf{x}^{-i}) = \omega_s^i$ .

**Proof.**

We have

$$|\omega_s^i - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)| \leq |\omega_s^i - \tau_s^i(\mathbf{x}^{-i,n})| + |\tau_s^i(\mathbf{x}^{-i,n}) - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\tau}^i(\mathbf{x}^{-i,n}))| + |\gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\tau}^i(\mathbf{x}^{-i,n})) - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)|.$$

Now, by assumption, as  $n \rightarrow \infty$   $|\omega_s^i - \tau_s^i(\mathbf{x}^{-i,n})| \rightarrow 0$  and  $|\gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\tau}^i(\mathbf{x}^{-i,n})) - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)| \rightarrow 0$ . Note that  $|\tau_s^i(\mathbf{x}^{-i,n}) - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\tau}^i(\mathbf{x}^{-i,n}))| = |\gamma_s^i(\mathbf{x}_s^{-i,n}, \boldsymbol{\tau}^i(\mathbf{x}^{-i,n})) - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\tau}^i(\mathbf{x}^{-i,n}))| \rightarrow 0$  as  $n \rightarrow \infty$  by Lemma 4. Hence,  $|\omega_s^i - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)| \rightarrow 0$  as  $n \rightarrow \infty$ .  $\square$

**Alternative proof for Theorem 1:**

**Theorem 1**

For  $\mathbf{x} \in X$ , define  $\gamma_{\mathbf{x}}(\boldsymbol{\omega}) : W \rightarrow W$  by  $(\gamma_{\mathbf{x}}(\boldsymbol{\omega}))_{is} = \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i)$ . The function  $\gamma_{\mathbf{x}}(\boldsymbol{\omega})$  is a contraction mapping.

**Proof.** Let  $\boldsymbol{\omega}, \boldsymbol{\theta} \in W$ . For  $\mathbf{x}_s^{-i}$  fixed,  $\forall i \in I, s \in S$ ,

$$\begin{aligned} \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) &= \min_{\substack{\mathbf{u}_s^i \in X_s^i \\ \tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \max \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{u}_s^i; \boldsymbol{\omega}^i) \\ &= \psi_s^i(C_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}), P_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}); \mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}; \boldsymbol{\omega}^i), \end{aligned}$$

where  $\mathbf{u}_s^{*i}$  is the minimizer, and  $C_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}) \in C_s$  and  $P_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}, \boldsymbol{\omega}^i) \in P_s$  are the optimizers that now depend on  $(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i})$ . Similarly, with  $\mathbf{z}_s^{*i}$  and  $C_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}) \in C_s$ ,  $P_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \boldsymbol{\theta}^i) \in P_s$ , we have

$$\begin{aligned}\gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\theta}^i) &= \min_{\mathbf{z}_s^i \in X_s^i} \max_{\substack{\tilde{C}_s \in C_s \\ \tilde{P}_s \in P_s}} \psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{z}_s^i; \boldsymbol{\theta}^i) \\ &= \psi_s^i(C_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}), P_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \boldsymbol{\theta}^i); \mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}; \boldsymbol{\theta}^i).\end{aligned}$$

Now,

$$\begin{aligned}&\gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\theta}^i) \\ &= \psi_s^i(C_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}), P_s^i(\mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}, \boldsymbol{\omega}^i); \mathbf{x}_s^{-i}, \mathbf{u}_s^{*i}; \boldsymbol{\omega}^i) - \psi_s^i(C_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}), P_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \boldsymbol{\theta}^i); \mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}; \boldsymbol{\theta}^i) \\ &\leq \psi_s^i(C_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}), P_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \boldsymbol{\omega}^i); \mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}; \boldsymbol{\omega}^i) - \psi_s^i(C_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}), P_s^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \boldsymbol{\theta}^i); \mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}; \boldsymbol{\theta}^i) \\ &= \sum_{a \in A} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,am}^m z_{s,a}^{*i} \{C_{sa}^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}) + \beta \sum_{k=1}^M P_{sak}^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \omega_k^i) \omega_k^i\} \\ &\quad - \sum_{a \in A} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,am}^m z_{s,a}^{*i} \{C_{sa}^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}) + \beta \sum_{k=1}^M P_{sak}^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \theta_k^i) \theta_k^i\} \\ &\leq \sum_{a \in A} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,am}^m z_{s,a}^{*i} \beta \left\{ \sum_{k=1}^M P_{sak}^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \omega_k^i) (\omega_k^i - \theta_k^i) \right\} \\ &\leq \sum_{a \in A} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,am}^m z_{s,a}^{*i} \beta \left( \sum_{k=1}^M P_{sak}^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \omega_k^i) \right) \|\boldsymbol{\omega} - \boldsymbol{\theta}\|_\infty = \beta \|\boldsymbol{\omega} - \boldsymbol{\theta}\|_\infty.\end{aligned}$$

The second to the last inequality above follows from the fact that

$$\sum_{a \in A} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,am}^m z_{s,a}^{*i} \sum_{k=1}^M P_{sak}^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \omega_k^i) \theta_k^i \leq \sum_{a \in A} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,am}^m z_{s,a}^{*i} \sum_{k=1}^M P_{sak}^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \theta_k^i) \theta_k^i,$$

because for a given  $(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \theta_k^i)$ ,  $[P_{sak}^i(\mathbf{x}_s^{-i}, \mathbf{z}_s^{*i}, \theta_k^i)]_{k=1, \dots, M}$  is the maximizer of  $\psi_s^i(\tilde{C}_s, \tilde{P}_s; \mathbf{x}_s^{-i}, \mathbf{z}_s^i; \boldsymbol{\theta}^i)$  over  $\tilde{P}_s \in P_s$ .

Similar to the above arguments, we have for  $\mathbf{x}_s^{-i}$  fixed that,  $\forall i \in I, s \in S$ ,

$$\gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\theta}^i) - \gamma_s^i(\mathbf{x}_s^{-i}, \boldsymbol{\omega}^i) \leq \beta \|\boldsymbol{\omega} - \boldsymbol{\theta}\|_\infty.$$

Thus,  $\|\gamma_x(\boldsymbol{\omega}) - \gamma_x(\boldsymbol{\theta})\|_\infty \leq \beta \|\boldsymbol{\omega} - \boldsymbol{\theta}\|_\infty$ . □

## Acknowledgments

This research was supported by the United States Department of Homeland Security through the Center for Risk and Economic Analysis of Terrorism Events (CREATE) under grant number 2007-ST-061-000001. However, any opinions, findings, and conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the United States Department of Homeland Security. I also would like to acknowledge and thank Michelle Aghassi for the insights she provided on the existence proof for robust one-shot  $n$ -player normal form games.

## References

- Aghassi, M., D. Bertsimas. 2006. Robust game theory. *Mathematical Programming (Series B)* **107** 231–273.
- Altman, E. 1994a. Flow control using the theory of zero-sum markov games. *IEEE Trans. Automatic Control* **39**(4) 814–818.
- Altman, E. 1994b. Monotonicity of optimal policies in a zero sum game: A flow control model. *Advances in Dynamic Games and Applications*, vol. 1. Birkhauser.
- Altman, E., A. Hordijk. 1995. Zero-sum markov games and worst-case optimal control of queueing systems. *QUESTA* **21** 415–447.
- Altman, E., N. Shimkin. 1993. Worst-case and nash routing policies in parallel queues with uncertain service allocations. Tech. rep., Institute for Mathematics and Applications, University of Minnesota. IMA Preprint No. 1120, Minneapolis.
- Altman, E., N. Shimkin. 1998. Individually optimal dynamic routing in a processor sharing system: Stochastic game analysis. *Operations Research* **46**(6) 776–784.
- Aumann, R. J., M. B. Maschler. 1968. Repeated games of incomplete information: The zero-sum extensive case. Tech. rep., U.S. Arms Control and Disarmament Agency. ST-143, Washington D.C., Chapter III, 37116.
- Avşar, Z. M., M. B. Gürsoy. 2002. Inventory control under substitutable demand: A stochastic game application. *Naval Research Logistics* **49** 359–375.
- Bagnell, J., A. Ng, J. Schneider. 2001. Solving uncertain markov decision problems. Tech. rep., Robotics Institute, Carnegie Mellon University. Technical Report CMU-RI-TR-01-25, Pittsburgh, PA.
- Ben-Tal, A., A. Nemirovski. 1998. Robust convex optimization. *Mathematics of Operations Research* **23**(4).
- Bertsimas, D., M. Sim. 2004. The price of robustness. *Operations Research* **52** 35–53.
- Ferris, M. C., S. P. Dirkse, A. Meeraus. 2005. Mathematical programs with equilibrium constraints: Automatic reformulation and solution via constrained optimization. *Frontiers in Applied General Equilibrium Modeling*. Cambridge University Press, 67–93.
- Filar, J., K. Vrieze. 1997. *Competitive Markov Decision Processes*. Springer-Verlag.
- Fink, A. M. 1964. Equilibrium in a stochastic n-person game. *Journal of Science in Hiroshima University, Series A-I* **28** 89–93.
- Gilboa, I., D. Schmeidler. 1989. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics* **18** 141–153.
- Givan, R., S. Leach, T. Dean. 1997. Bounded parameter markov decision processes. *Fourth European Conf. Planning*. 234–246.
- Harsanyi, J. C. 1967, 1968. Games with incomplete information played by bayesian players, i-iii. *Management Science* **14**.
- Heyman, D. P. 1968. Optimal operating policies for m/g/1 queueing systems. *Operations Research* **16** 362–282.
- Iyengar, G. 2005. Robust dynamic programming. *Mathematics of Operations Research* **30**(2) 1–21.
- Kakutani, S. 1941. A generalization of brouwer's fixed point theorem. *Duke Mathematical Journal* **8** 457–459.
- Lo, K. C. 1996. Equilibrium in beliefs under uncertainty. *Journal of Economic Theory* **71** 443–484.
- Marinacci, M. 2000. Ambiguous games. *Games and Economic Behavior* **31** 191–219.
- Nilim, A., L. El Ghaoui. 2005. Robust control of markov decision processes with uncertain transition matrices. *Operations Research* **53**(5) 780–798.
- Peterson, L. L. 2007. *Computer Networks: A systems approach*. Morgan Kaufmann.
- Raghavan, T.E.S., J. Filar. 1991. Algorithms for stochastic games - a survey. *Methods and Models of Operations Research* **35** 437–472.
- Rosenberg, D., E. Solan, N. Vieille. 2004. Stochastic games with a single controller and incomplete information. *SIAM Journal of Control and Optimization* **43**(1).



- Satia, J. K., R. L. Lave. 1973. Markov decision processes with uncertain transition probabilities. *Operations Research* **21**(3) 728–740.
- Shapiro, A., A. J. Kleywegt. 2002. Minimax analysis of stochastic problems. *Optim. Methods Software* **17**(1) 523–592.
- Shapley, L. S. 1953. Stochastic games. *Proceedings of the National Academy of Science*, vol. 39. 1095–1100.
- Sobel, M. J. 1969. Optimal average-cost policy for a queue with start-up and shut-down costs. *Operations Research* **17** 145–162.
- Sorin, S. 1984. Big match with lack of information on one side i. *International journal of Game Theory* **13** 201–255.
- Sorin, S. 1985. Big match with lack of information on one side ii. *International journal of Game Theory* **14** 173–204.
- Stidham, S., R. R. Weber. 1989. Monotonic and insensitive optimal policies for control of queues with undiscounted costs. *Operations Research* **87**(4) 611–625.
- Vanderbei, R.J. 2006. Loqo user’s manual - version 4.05. Tech. rep., Operations Research and Financial Engineering, Princeton University.
- White, C. C., H. K. Eldeib. 1994. Markov decision processes with imprecise transition probabilities. *Operations Research* **42**(4) 739–749.
- Yechiali, U. 1971. On optimal balking rules and toll charges in a gi/m/1 queueing process. *Operations Research* **19** 349–370.