

DRAFT

Report #05-033

**Robust Stochastic Games and
Applications to Counter-Terrorism
Strategies**

Kardes, E.

CREATE REPORT
Under **FEMA Grant** N00014-05-0630
December 16, 2005



**Center for Risk and Economic Analysis of Terrorism Events
University of Southern California
Los Angeles, California**

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 N00014-05-0630. 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.

Robust Stochastic Games and Applications to Counter-Terrorism Strategies

CREATE Report

Erim Kardeş

December 16, 2005

Epstein Department of Industrial and Systems Engineering
University of Southern California
Los Angeles, CA, 90089-0193

ABSTRACT

This report presents a new methodology for strategic decision making under uncertainty and presence of adversaries. This investigation is motivated by the need to determine optimal strategies under uncertainty against an adversarial and adaptive opponent. Such problems arise in the context of terrorism threats. To model investment decisions that pertain to homeland security, one should account for both uncertainty and the antagonistic character inherent in the problem. To this end, we propose a novel approach, robust stochastic games. We focus on incomplete information stochastic games and adopt a robust approach to account for uncertainty present in our problem in two dimensions. First, we consider that the adaptive nature of the adversary is uncertain. In other words, we propose a new approach that accounts for the uncertainty in the conversion from one threat category to the other that is based on the alternatives of the adversaries. Second, we consider that payoffs to the opponents are uncertain. We present an interesting new result, existence of equilibrium points in robust stochastic games. A new formulation that uses robust optimization techniques is proposed to solve robust stochastic games. Preliminary results are presented on a simple example with partial unknown data. First, uncertain transition probabilities that belong to convex hull uncertainty sets are considered in this example with exact immediate costs. Next, uncertainty is considered in both transition probabilities and immediate costs. Performance of the nominal solution when parameters attain their worst-case values is compared with the performance of the robust solution when data is certain. It is observed in this small example that the percentage savings resulting from using robust strategies versus the nominal strategies when the parameters attain their worst-case values are higher than the losses caused by using robust strategies when parameters attain their nominal values. It is also observed that compared to the uncertainty in transition probabilities, uncertainty in immediate costs has a greater effect on the robust value of the game. The next phase in this research includes the development of the model for the MANPADS case study, quantification of the model via expert elicitation, and computation of robust optimal strategies.

ACKNOWLEDGMENT

This research was supported by the United States Department of Homeland Security through the Center for Risk and Economic Analysis of Terrorism Events (CREATE), grant number EMW-2004-GR-0112. However, any opinions, findings, and conclusions or recommendations in this document are those of the author's and do not necessarily reflect views of the U.S. Department of Homeland Security.

Contents

1	Introduction	5
1.1	Research Problem and Objectives	7
2	Literature Survey	11
2.1	Probabilistic Risk Analysis (PRA)	11
2.2	Game Theory	12
2.3	Game Theory and Influence Diagrams	15
2.4	Game Theory and Reliability	17
2.5	Game Theory and Robust Optimization	18
2.6	Stochastic Games	19
2.7	Other Approaches	21
2.8	Contributions	21
3	Methodology	23
3.1	Stochastic Games	24
3.2	Robust Optimization	27
3.3	Formulation of Robust Stochastic Games	28
4	Existence of Equilibrium Points in Robust Stochastic Games	31
5	Calculation of an Equilibrium Point	44
5.1	Preliminary Results	49
5.1.1	A Small Numerical Example	49
5.1.2	More Analysis in Example 1	56

1 Introduction

The threat of catastrophic terrorism has motivated multi-billion dollar investments in the United States and elsewhere, with the goal of improving safety and security. Investments of this magnitude demand careful consideration of the costs of implementation, operation and maintenance, as well as the benefits derived from a reduction in exposure to future losses.

Unlike naturally occurring or accidental events – such as floods, earthquakes or system failures – terrorism is fundamentally adversarial and adaptive. Thus investments designed to protect against one type of terrorism threat (e.g., against blasts, biological agents, or radiological devices) have potential to elevate the risk of other types of terrorism. Furthermore, interventions to protect against a certain type of terrorism threat category may influence the terrorists’ selection of alternative methods within the same threat category. On the other hand, investments targeted at reducing the general effectiveness of terrorist organizations, or targeted at the willingness of individuals to engage in terrorism, may protect against multiple types of terrorism.

Other distinctions among accidental events and terrorism are the different requirements they impose on cost/benefit analysis. In the study of accidental events and related safety modeling techniques, it is a common practice to use data, expert judgments, or to use a blend of hard data and expert opinions in suitably constructed models. For instance, in the aviation safety domain in the US, it is fairly easy to access large historical accident / incident data. However, it is extremely difficult in the terrorism case to obtain exact data. This could be attributed to the short history of terrorism in the US. Furthermore, another factor that contributes to the uncertainty in the terrorism domain is the fact that any data to be used in related models are subject to change in the future.

Besides the fact that terrorism is adversarial, the intensity of adversarial intentions is volatile. Although this intensity level could be moderate in some cases, it is a fact that terrorism is fully antagonistic.

Terrorism is also fundamentally different from the risk of warfare among states, especially when it is not state sponsored. Conventional warfare is a less random occurrence than terrorism because state adversaries are more likely to announce their specific intentions and because their actions are more easily monitored through surveillance. States are also more easily deterred through the threat of a specific military response that inflicts losses

on the adversary. In the case of terrorism, the opponent attacks with greater frequency, with greater randomness, and often without the opportunity of deterrence through a direct military response. Thus both the methods of protecting against warfare, and the methods for analyzing the threat, are fundamentally different than they are for terrorism. Last, terrorism also differs from other intentional criminal acts. As witnessed on 9/11, terrorism has potential to produce more catastrophic events than simple criminal acts, as well as to introduce radically new tactics in its effort to produce fear. Tools used to model exposure to criminal losses, which are common in the insurance industry, are not easily modified to evaluate adaptable antagonistic terrorist adversaries.

One goal of the Center for Risk and Economic Analysis of Terrorism Events (CREATE) is to develop tools to guide investments in counter terrorism, accounting for economic costs and benefits, and accounting for non-state-based terrorism risk. Common to such tools are the underlying models used and corresponding input data to these models. In terrorism related security modeling, it is a clear fact that the data necessary to feed corresponding models could not be measured exactly and / or are unknown. Furthermore, even if we assume that certain parts of the input data to security models could be measured, it is another problem that the data is subject to change in the future. Hence, using some estimate data in a certain time period could be useless in future periods. This could result in investments that are no longer applicable to their intended targets and consequently in higher costs. Given that investments could be at a level of multi-billion dollars, the issue of robustness of homeland security investments needs significant attention.

The purpose of this report is to present a new methodology that addresses robustness in homeland security investment decisions with respect to uncertainties present in the problem. The adversarial aspect of this problem is captured by an existing game theoretical approach, namely stochastic games. Stochastic games are briefly introduced in section 3.1. In our new approach, we focus on incomplete information stochastic games and adopt a robust approach to account for uncertainty present in our problem in two dimensions. First, we consider that the adaptive nature of the adversary is uncertain. In other words, we propose a new approach that accounts for the uncertainty in the conversion from one threat category to the other that is based on the alternatives of the adversaries. Second, we consider that resultant payoffs to the opponents based on their choices are uncertain.

1.1 Research Problem and Objectives

Determining investment strategies to improve security against terror differs radically from conventional problems of decision making under uncertainty. This is due to the adaptive adversarial behavior of attackers and also to the lack of exact data needed to model such problems. A way to deal with limited data is to use expert knowledge in models. Of course, this requires expert elicitation, a laborious process even for interviews with a small group of experts. Even if expert knowledge is used in models related to investment decisions, elicitation of experts typically leads to bands of uncertainty that cannot be ignored. If such uncertainty values are used in a numerical evaluation, resulting values may have very large uncertainty bands. On the other hand, continuing to work with point estimates of risk may give a wrong interpretation of any resulting risk value.

Our research objective in this study is to model and solve investment decision problems using a suitable methodology that accounts for the uncertainty inherent in the problem as well as for the uncertain adversarial behavior of opponents. To this end, we follow a game theoretical approach that has a long and illustrious history, namely, stochastic games.

Stochastic games are first introduced to the game theory literature by Lloyd S. Shapley in 1953. In this (non cooperative) two-person zero-sum game, the play proceeds in stages, from one state to the other according to transition probabilities controlled jointly by two opponents. It consists of states and actions associated with each player. Once in a state, each player chooses their respective actions. The play then moves into another state with some probability that is determined by the actions chosen and by the state in which they are chosen. Given that opponents make their respective decisions in a given stage, a cost is incurred to each player. An opponent discounts his projected cost by a factor β , $0 \leq \beta < 1$. The usual interpretation of this factor is that decision makers (or opponents) consider that costs incurred in future stages have less value in the present stage. Another interpretation of this factor in homeland security applications is the interest rate interpretation that determines the return on investment that could have been earned if the decision maker had not invested the funds in security investments.

The first paper on stochastic games in 1953 considers two-person zero-sum stochastic games. Two person indicates that there are two players in the model. Zero sum denotes that a player's (usually player 1) gain is the cost to the other player (usually player 2). Hence, there is a complete utility

transfer from one player to the other and payoffs to the players sum up to zero. The two-person zero-sum scheme is later on extended to the general-sum n -person stochastic games by Fink (1964). In the general-sum n -person stochastic games, one cannot capture the fully antagonistic intentions of the players. Nevertheless, completely antagonistic goals are captured in the two-person zero sum case, a special class of general sum n person stochastic games. This is due to the property that a player's gain is the loss to the other.

A crucial property of stochastic games as also pointed out in [16] is that since costs and transition probabilities depend on the decisions of both decision makers, as well as on the current state, the *fate* of the decision makers is coupled in the process, even though their choices are made independently and secretly of one another. Note that this property is intimately related to the characterization of security investment decisions in the presence of terrorism threats.

In this report, we extend the ideas in [54] by introducing a new approach, *robust stochastic games*. First and foremost, we propose a robust approach to homeland security related decisions, since very small changes in parameters of this problem domain (e.g., any parameters for cost / benefit analysis in homeland security applications) could result in significant changes in ensuing decisions that the decision makers follow as outputs of proper models. The decision environment we deal with in this research has the following aspects to it.

- Clearly, the decisions to be implemented in homeland security applications are regarding very sensitive issues such as, first, human lives, and second, very significant amount of funds to be allocated to security investments.
- The data related to homeland security applications are subject to uncertainty at the time of the decisions to be made. Furthermore, the data are subject to changes in the future.
- Even if we suppose that data could be extracted in some manner from a given source, they cannot be measured, estimated, or computed exactly.

These properties of our decision problem are precisely the character of a decision environment where a robust approach is appealing. Robust models are used to represent uncertainty via uncertainty sets. These uncertainty sets

are not equipped with any probability distributions. In the robust approach, uncertain parameters are not known but it is known that they belong to a set. Determining probability distributions on data is a very difficult and laborious task, especially in the case of homeland security. Hence, robustness, as a distribution-free approach suggests significant promise in terms of reducing modeling costs and of characterizing the uncertainty in our problem via bounded uncertainty sets, rather than probability distributions. In this sense, robust optimization suggests to optimize an objective function with respect to the worst-case values that the uncertain parameters could attain, *with respect to the values the decision variables attain*. In this approach, neither the values of the decision variables, nor those of the uncertain parameters, are known *a priori*.

The worst-case approach may first seem conservative and perhaps intimidating in the case of homeland security decisions. One may claim that considering only the worst-case scenarios is a poor approach since the worst-case may seem to have no limit to it. This concern fails for the following reasons. First, we do not adopt a worst-case approach to determine the worst threat category that could happen. Our robust approach pertains to parameters in each threat category. In this study, we aim to break down the adversarial decision processes into smaller components. For example, if aviation security is a threat category, we could consider the transportation of weapons in this category as an individual step. The step of transportation of weapons encompasses alternatives to the adversaries, corresponding data such as payoffs, and likelihood of conversions to other steps based on the alternatives chosen. We consider uncertainty in the data of such steps that belong to a set of values that are bounded and adopt a worst case approach against possible realizations of the uncertain data. Hence, our endeavor is to determine best strategies against possible realizations of parameters in each threat category.

Second, the worst-case is with respect to the values that the decision variables attain as an outcome of robust models. Hence, if we suppose that a conversion probability to a worse threat category is between 0.2 and 1, this does not mean that our approach chooses the worst-case value 1. As we already pointed out, the uncertain data attain their worst-case values with respect to the attained values of our decision variables. Third, although we present here that the worst-case uncertain data values are attained with respect to the values of the decision variables, this still may be considered as a conservative approach. However, there are ways to deal with the level

of conservatism in the robust optimization literature and these techniques could be applied to our models.

The robustness of the approach proposed in this report not only concerns the future changes in uncertain parameters but also accounts for possible future threats. This is so because the model we propose not only accounts for the current stage of threats but also considers possible future payoffs resulting from intentions to convert to different threat categories. Therefore, two aspects of the new model proposed are the robustness that considers future changes in costs and transition probabilities as well as robustness in the decisions that pertain to threat categories possible to arise in future.

Our main objective then is to determine investment decisions that remain optimal with respect to (future) changes in costs incurred as well as to changes in transition probabilities from one stage to the other. Furthermore, we aim to implement decisions that remain optimal in the presence of future threats and conversions from thereon.

This report is organized as follows. Section 2 introduces survey of literature on decision making under uncertainty in the presence of adversaries. In section 3.1 and 3.2, we introduce the foundations of the new methodology proposed. Specifically, section 3.1 presents stochastic games and 3.2 briefly outlines robust optimization. In section 3.3, we formulate our novel approach, *robust stochastic games*. Section 4 presents interesting new results that enables us to formulate our problem as an optimization problem. We aim to prove in this section that robust stochastic games have equilibrium points in the sense of Nash's pioneering work in [41]. In other words, we seek whether decisions that prescribe us the best response to the other decision makers' best strategies could be found under uncertainty. Second, we intend to solve the natural problem that follows, i.e., to find such optimal robust decisions. To this end, we make use of an emerging body of literature in convex optimization, namely, robust optimization. We concretize the ideas presented in earlier sections in section 5, via a small example. Finally, section 6 concludes the report.

2 Literature Survey

2.1 Probabilistic Risk Analysis (PRA)

The PRA method was initially developed for the purpose of assessing the safety of nuclear reactors. Expert elicitation is typically used as an input to risk models along with historical failure data. Bayesian methods are also frequently used in PRA, which is affected by the assumed prior probability distributions [6]. In this section we focus on the application of PRA to adversarial risks.

A tool is introduced in [29] that is built to assist anti terrorism planners at military installations to draw inferences about the risks of a terrorist attack. This tool allows anti terrorism planners to analyze and manage a large portfolio of risks simultaneously by encoding the knowledge about assets and risks into Bayesian network fragments that could be dynamically combined at run-time into a Bayesian network for assessing risks specific to a given installation and situation. The data sources for this hierarchical network include the planners own subjective assessments, historical database information, analytical model results and simulation results integrated into various nodes on the Bayesian network. The network is dynamically constructed by the tool and is solved and presented to the user for each combination of asset and threat that the user describes.

Laskey and Levitt (2002) provide a practical, computational methodology to encode a distributed library of patterns for automated reasoning about aspects of homeland defense against terrorism. Multi-Entity Bayesian networks (MEBNs) provide a means of encoding repeated patterns and relationships in the form of network fragments. These fragments are combined to form situation specific Bayesian networks. Authors propose the use of MEBNs as the inferential cornerstone of a cumulative national, distributed knowledge base for homeland defense. This paper illustrates the use of MEBNs with an example concerning a multi-city coordinated bio-warfare attack. Authors attempt to show how current trends in the use of on-line reporting by health care and related facilities have the potential to enable opportunistic detection of and response to such an attack.

Weaver et al. (2001) describe a research effort to develop models of terrorist organizations that will permit to stimulate and predict what types of decisions these organizations and their agents might be likely to make. Authors contend that terrorist organizations and individual decision makers

can be described via Markov Decision Processes and repeated Bayesian networks. Another task of this research is to gather literature sources and to assemble a database that contains profiles of a reasonable sample of terrorist organizations and to use this information in conjunction with the models developed.

Singh et al. (2004) develop a tool to detect and track terrorist activity. Authors follow two probabilistic approaches: Hidden Markov Models (HMMs) and Bayesian networks (BNs). Authors assert that HMMs, which are used for modeling partially observed stochastic processes, are an ideal way to make inferences about the evolution of terrorist networks. The HMMs detect the monitored terrorist activity and measure threat levels, whereas BNs combine the likelihoods from many different HMMs to evaluate the cumulative probability of terrorist activity. In other words, BNs represent the overarching terrorist plot and the HMMs, which are related to each BN node, represent detailed terrorist subplots. A case study for the 2004 Olympics is presented in this paper as an example.

Haimes (2002) offers a holistic risk assessment and management framework for modeling the risks of terrorism to the homeland. According to this paper, two major interconnected systems are the homeland and the terrorist network systems. The variables pertaining to the two systems are considered and their interactions are presented in a schematic way. Many other articles on PRA related to security modeling, some of which are cited here, are referred to in [7]. Although PRA is helpful to gain insight about security risks, it lacks the adversarial antagonistic aspect that is extensively present in the homeland security risk modeling domain. Hence, the next section is devoted to another approach that could confront this adaptive perspective.

2.2 Game Theory

Social scientists have written many papers on applications of game theory to terrorism, as explained in [49]. The authors contend that game theory captures the strategic interactions among terrorists and targeted governments, that is between players, where actions are interdependent and neither of the sides can be considered passive. Other reasons include the rationality assumption of the players in games and ability of games to represent gains or losses to a player through payoffs. The main purpose of this paper is to review how game theory has been used in the literature and to present new applications that include terrorists choice of targets, governments choice between

preemption and deterrence, and the government concessionary policy when terrorists are of two minds: hard-liners and moderates. For example, for the choice between preemption and deterrence, a three-player game is played in normal form that includes the US, the UK and the terrorist organization. For the government concessionary problem, a model of bargaining between a government and a terrorist group with moderate and hard-line members is considered.

It is important to note that Sandler and Arce (2003) use a simple game theory model to answer high-level, generic questions. The authors note that the model would benefit from a multi-period analysis of terrorist campaigns, where terrorist resource allocation is studied over time. Another area of future work could be differential games to examine how terrorist organizations are influenced by successful and failed operations. The dynamics of strategic choices of both players can be captured with this approach by modeling for the rate of change over time of resources for each player. Finally, the authors note that cooperative game theory has never been applied to the study of terrorism, which would enable analysis of shared intelligence, training facilities, and operatives to strengthen their abilities.

Sandler et al. (1983) present models that depict the negotiation process between terrorists and government policymakers for incidents where hostages are seized and demands are issued. Lapan and Sandler (1988) present a game in extensive form where the government first chooses the level of deterrence that consequently determines the logistical failure or success of terrorists when they engage in a hostage mission. Atkinson et al. (1987) extend Nash's bargaining game, where time is taken into consideration. Sandler and Siqueira (2002) present an application of game theory that involves terrorists choice of targets for a three-player game involving two targeted governments and a common terrorist threat. Lapan and Sandler (1993) analyze a scenario via a two-period game, where the government is incompletely informed about the terrorists capability. The extent of terrorist attacks in this scenario can provide information to the government about the type of the terrorist group.

Faria (2003) makes two contributions to the literature on terrorism: 1) It presents a model that explains the cyclical characteristic of terrorist attacks, and 2) It improves on the existing theoretical cyclical models since it takes into account terrorists motivations and decision-making explicitly. A differential game is used between terrorists and the government in which terrorists maximize the number of attacks subject to a constraint that combines terrorists resources and government anti-terrorist policies. This model

is a standard microeconomic model, where the representative terrorist group solves a maximization problem based on preferences, actions, incentives, and budget restrictions. The government problem concerns the maximization of national security. The solution of the terrorist problem yields a time path for terrorist activities. The government takes the time path of terrorist activities into account when maximizing national security over time. The solution of the government problem yields a limit cycle between enforcement and terrorist activities. The permanent cyclical paths in enforcement and terror cause national security and terrorist stocks to display cyclical trajectories as well.

D'Artigues and Vignolo (2003) study the emergence of the recent form of terrorism using evolutionary game theory. The model in this paper presents terrorism as the result of competition between countries, when the desire to imitate the leading country is frustrated by the impossibility of doing so. Authors define a multi-country setup where interaction takes place in an international trade game, which is a coordination game. In particular, this paper uses the evolutionary game model to describe the long-run behavior of n countries.

Kunreuther and Heal (2003) consider security as a problem among agents and focus on situations where the security levels of members of a group are interdependent. The main idea in this paper is that the dependence of one agents security on the behavior of others may partially or completely negate the payoffs it receives from its own investment in protective measures. These cross-effects are referred to as contagion. Authors illustrate this argument by reference to an airline that attempts to determine whether to install a baggage checking system. In making this decision, the airline needs to balance the cost of installing and operating such a system with the reduction in the risk of an explosion from a piece of luggage not only from the passengers who check in with it, but also from the bags of passengers who check in on other airlines and then transfer to it. In this example, the incentive to invest in security decreases if other airlines fail to adopt protective measures. As the authors indicate, this paper examines the case where all agents are identical. Heal and Kunreuther (2003) consider situations where the agents have different protection costs and risks, and where the actions creating potential losses are impacted by agents protective decisions. Future research directions suggested in the paper include examining how agents behave in multi-period models and determining appropriate behavioral models of choice that could characterize individuals who make imperfectly rational decisions.

Major (2002) presents another application of game theory, which includes

a simplified model of terrorism risk to develop a probability distribution of losses. However, this effort captures only the severity component of risk that is of potential interest to the insurance professionals. The losses that could occur with certain probabilities are revealed given that an attack is attempted.

Game theory is a suitable way to model adversarial decision-making processes. However, this approach still has limitations and simplifications. Game theory applications could be supported with additional modeling methodologies as described below.

2.3 Game Theory and Influence Diagrams

Pate-Cornell and Guikema (2002) present a generic influence diagram model for setting priorities among threats and among countermeasures. The random variables used in the authors first model is fairly generic and account for types of terrorist groups, their access to materials, cash, types of weapons, and etc. For instance, only one decision variable is used to represent U.S. countermeasures. The authors next model elaborates on the previous one by considering two influence diagrams: one for the terrorist behavior and the other for the U.S. Results pertaining to the influence diagram for terrorist behavior are then used as inputs to the influence diagram for US. Hence, this model is called two-sided. The authors then consider using the two-sided diagram in a dynamic fashion via discrete time steps. At each step, each side updates its beliefs, objectives, and decisions based on the previous step. It is also denoted that each side is uncertain about the other's actions and state of knowledge. According to the authors, another change that needs to be included in the model is the evolution of the organizations involved, the emergence of new groups, or a new structure of existing groups and networks. Although these ideas are put forward, no implementations or quantitative illustrations exist with regards to the dynamic approach or evolutions of organizations.

According to Koller and Milch (2001), the traditional representations of games using the extensive form or the normal form obscure much of the structure that is present in real-world games. Hence, authors propose a new representation language, named multi-agent influence diagrams (MAID), for general multi-player games. This approach extends influence diagrams to a context where more than one decision maker is involved, an idea first examined by Shachter (1986). MAIDs allow the dependencies among variables to

be represented explicitly, whereas both the normal and the extensive form obscure certain important relationships among variables. MAIDs representation extends the Bayesian network formalism [46] and influence diagrams [28] to account for the decision problems involving multiple decision makers. They have defined semantics as non-cooperative games. Just as Bayesian networks make explicit the dependencies among random variables, MAIDs make explicit the dependencies among decision variables. They are also related to the formalism presented by La Mura (2000), where network representation for games is developed. Solutions to MAIDs consider the strategic independence structure on the diagram. Extensions to this research could be establishing the relations among competitive Markov decision processes, stochastic games, and MAIDs. Another extension could be exploring ways to integrate the issue of evolution over time into the MAIDs framework.

Brynielsson and Arnborg (2004) review some military applications of gaming and introduce a game component into an influence diagram example. Authors illustrate the use of Bayesian game-theoretic reasoning for operations planning by transforming a decision problem into a Bayesian game.

Virtanen et al. (2004) describe a multistage influence diagram game for modeling the maneuvering decisions of pilots in one-on-one air combat. Virtanen et al. (2004b) describe an extension of the influence diagram approach into a dynamic multistage setting without any game aspect. Authors contend that this paper is the first elaboration where ideas regarding multi-agent multi-period influence diagrams are combined and implemented. Dynamic programming is considered for the solution of the model in this paper. To cope with the combinatorial explosion, authors trade the solution of the complete game with the computing time and apply a moving horizon control approach, where the horizon of the original influence diagram is truncated and a dynamic game with a shorter planning horizon is solved at each decision instant. Instead of the whole duration of the game, this approach allows the players to update their information about the state of the system at any moment over the limited planning horizon. Virtanen et al.s solution approach is inspired by Cruz et al. (2002), who contend that dynamic game theory is a suitable formulation for problems that involve adversaries interacting with each other over a time period. Cruz et al. (2002) denote that traditional solutions from dynamic game theory that involve optimizing objective functions over the entire time horizon of the system are extremely difficult but not impossible to derive. Hence, the authors discuss a solution approach, where at each step the players limit the computation of their actions to a shorter

time horizon that may involve only the next few time steps. This moving horizon Nash equilibrium solution proves to be useful in near term decisions of the adversaries. An important extension to this research effort could be accounting for the uncertainty in payoffs by combining robust optimization techniques with game theory.

2.4 Game Theory and Reliability

Many of the applications of reliability to security consider the threats against critical infrastructure, such as water supply systems (Haines, 2002). However, many applications do not consider an adaptive adversary. Therefore, incorporating game theory and risk and reliability analysis could be a fruitful approach [7]. Hausken (2002) attempts to combine probabilistic risk analysis (PRA) and game theory by associating each unit in a reliability system with a player. By doing so, a behavioral dimension is introduced into PRA framework. The article demonstrates the different conflicts that arise among players in series, parallel, and summation systems over which players incur costs.

Bier et al. (2004) apply game theory and reliability analysis to identify optimal defenses against intentional threats to system reliability. Various scenarios are considered in this paper such as perfect attacker knowledge of defenses and single attack with constrained defender budget or no attacker knowledge and single attack with unconstrained defender budget. Results of this paper emphasize the value of redundancy as a defensive strategy. According to the authors, future research could include extending this work to combinations of parallel and series systems rather than focusing only on pure parallel or series systems. Finding optimal strategies for arbitrary systems is difficult. Hence, near-optimal heuristic attack and defense strategies could be developed. Another promising area of future research is to extend the models to include time, rather than the current static or snapshot view of system security. This could allow the modeler to consider imperfect attacker information as well as multiple attacks over time. Another interesting future research topic could be the relation of stochastic games and reliability analysis.

In a more recent effort, Azaiez and Bier (2004) extends results for defense of simple systems to combined series/parallel systems of more realistic complexity. This effort sometimes yields counterintuitive results, such as the observation that defending the stronger components in a parallel subsystem

can actually impose greater burdens on prospective attackers than hardening the weaker components. The authors indicate that the approach is limited to cases where the cost of attacks increases linearly with regards to the defensive investments. However, this may hold only for a limited range of defensive investments. Second, an extension could be to relax the budget constraint, and permit the total investment to be optimized based on the value of the system being protected. Third, repeated attacks evolving in time could be investigated.

2.5 Game Theory and Robust Optimization

Combination of game theory and robust optimization techniques is a very new research area. Aghassi and Bertsimas (2004) consider robustness in one-shot general-sum n -person games in [1]. In this interesting paper, authors assume that payoffs to players belong to bounded uncertainty sets and adopt a robust optimization approach. Methods in this paper are the closest to the ones presented in this report, although the two approaches differ significantly.

Hayashi et al. (2004) consider a bimatrix game in which the players can neither evaluate their cost functions exactly nor estimate their opponents' strategies accurately. Note that this is the case in many applications in homeland security research. To formulate such a game, authors introduce the concept of robust Nash equilibrium and prove its existence under some mild conditions. Moreover, authors show that a robust Nash equilibrium in the bimatrix game can be characterized as a solution of a second-order cone problem (SOCP). Some numerical results are presented to illustrate the behavior of robust Nash equilibria. Although Hayashi et al. (2004) considered robustness in a bimatrix game, combining robust optimization techniques with game theory is open to many future research areas. First of all, differential or dynamic games with uncertainty could be worthwhile to study through robust optimization techniques. Furthermore, as the authors indicate, the concept of robust Nash equilibrium could be extended to the general N -person game. For the 2-person bimatrix game studied in this paper, it is sufficient to consider the uncertainty in the cost matrices and the opponent's strategy.

To discuss general N -person games, a more complicated structure should be dealt with. Another issue is to find other sufficient conditions for the existence of robust Nash equilibria. Also, theoretical study on the relation between Nash equilibrium and the robust Nash equilibrium is worthwhile.

For example, it is not known whether the uniqueness of Nash equilibrium is inherited to robust Nash equilibrium. In this paper, authors have formulated several robust Nash equilibrium problems as SOCPs. However, they have only considered the cases where either the cost matrices or the opponent's strategy is uncertain for each player. According to the authors, it seems interesting to study the case where both of them are uncertain, or the uncertainty set is more complicated. In numerical experiments, authors employed an existing algorithm for solving SOCPs. But, there is room for improvement of solution methods. It may be useful to develop a specialized method for solving robust Nash equilibrium problems.

2.6 Stochastic Games

As mentioned in the introduction, unlike naturally occurring or accidental events, terrorism is essentially adversarial. Therefore, investments designed to protect against one type of terrorism (e.g., against blasts, biological agents, or radiological devices) have the potential to elevate the risk of other types of terrorism over a given time period. An approach that accounts for such an evolution over time could be adopted by using stochastic games. There is an extensive amount of research in stochastic games in various fields such as economics, mathematics and operations research since the 1950s. The basic two person zero sum (discrete) stochastic game is played as follows. There are states, and strategy sets for each player and for all states. The system evolves in stages represented by discrete time points. At each state, the system is in one of its states and players 1 and 2 choose their respective actions. There is an immediate payoff as a consequence of the choices of the players. Then, the system moves into another state with some probability determined by the previous state, and by the choices of the players in the previous state. The fundamental question is then to find the optimal strategies that could be adopted by the players that optimize their own (noncooperative) objectives. Shapley (1953) first introduced stochastic games and proved that the value and optimal strategies of the game exist. 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 countless others.

Since publications on stochastic games are usually in the form of research papers and monographs, Filar and Vrieze (1997) devote a single textbook to the topic. The authors study discrete time finite state finite action stochas-

tic games with complete information from the Markov decision processes and mathematical programming points of views, where there are more than one decision makers with conflicting objectives, and use the name Competitive Markov Decision Processes. The authors treat discounted stochastic games, their relation with linear programming and nonlinear programming formulations, and the existence of stationary strategies and equilibria in depth. An important result is that the class of nonstationary strategies cannot achieve a better equilibrium value than the class of stationary strategies. Another observation is that, unlike Markov decision processes problems, general two person zero sum stochastic games cannot be solved by linear programming (LP). However, certain restrictions could be imposed on them in order to convert the problem into a suitable linear programming problem. Two of these restrictions are as follows. First, single controller discounted games lend themselves to LPs [43]. In this model, the system makes a transition into the next state with some probability according to the previous state and the action taken by one of the decision makers in the previous state. Hence, the action of the other player is irrelevant in determining the next state. Second, separable-reward-state-independent-transition discounted stochastic games could be converted to an LP. In this model, the payoff function can be expressed by two components, where one component is dependent only on the current state, and the other component is on the pair of choices made by the decision makers. Also, the transition to the next state is determined only by the pair of actions taken by the opponents and does not depend on the current state of the system [44]. Advances in stochastic games throughout the years could be viewed from two coupled perspectives: game theoretical perspective, and the stochastic processes perspective.

An extension to stochastic game models mentioned above is the one with incomplete information. The incomplete information case within the repeated games is first introduced by Aumann and Maschler (1968). Several authors have adopted the approach by Aumann and Maschler (1968) to stochastic games. In a recent paper by Rosenberg, et al. (2004), authors consider stochastic games with incomplete information for one of the players. However, the restriction in this model is that the transitions to the next state are controlled by a single player. Another extension by same authors concerns incomplete information on both sides. A two-player zero-sum stochastic game with incomplete information is described by a finite collection of stochastic games. It is assumed that the games differ only through their payoffs but they all have the same sets of states and actions, and the same

transition matrix. The game is played in stages. A stochastic game is to be played out of the finite set of games over which a probability distribution is specified. Player 1 is informed of the specific game to be played, while player 2 is not. All that the second player knows is that a game is to be chosen randomly from the finite set of games and to be played thereafter. At every stage, the two players choose their actions simultaneously and the system moves into the next state. Both players are informed of their actions and the current state of the system. Note that the actual payoff is not told to player 2 but is known by player 1.

It is important to note that the approach adopted by Rosenberg et al. (2004) is based, in some sense, on the approach proposed by Harsanyi (1967, 1968). In his study that brings him the Nobel Prize in 1994, Harsanyi proved that an incomplete two person zero sum normal form game (I-game) could be converted into a set of complete information games (C-game) that is equivalent to the original I game.

Types of extensions to stochastic games from the stochastic processes perspective include considering nonhomogeneous games [19], continuous time games [38], semi Markov games [30] among numerous others.

2.7 Other Approaches

Faria (2004) contends that terrorist innovations result from the innovation effect that is triggered by counter-terrorist policies. To model this phenomenon, Faria creates a dynamic model of terrorist attacks and innovations. The model consists of a set of differential equations, and is used to compare the effectiveness of three different anti-terrorist policies: deterrence, preemption and intelligence. Hazen (2002) introduces stochastic trees, where chance nodes in a decision tree can be stochastic nodes. This paper also uses stochastic nodes in influence diagrams. By doing so, variables that change state over time are captured in the influence diagram methodology. The authors apply this new methodology to model medical decisions, and specifically, arthritic joint replacement decisions. A possible extension to this methodology could be considering the use of stochastic nodes in games in extensive forms.

2.8 Contributions

Before presenting the contributions of this report, we note some gaps present in the literature.

From an application perspective, multi-stage game theoretical approaches are mentioned by numerous authors as future research directions in the application of game theory to adversarial decision making. Uncertainty inherent in the homeland security applications of game theory and related topics has received very little attention, let alone methods to cope with uncertainty. Robustness of the existing models received very little attention and has been mentioned in several occasions as future research directions. Finally, robustness in a multi-stage stochastic game theoretical setting does not exist.

From a methodological point of view, unlike one-shot games, incomplete information in stochastic games seems to be a fairly new research area in operations research. Very little exists in the literature regarding incomplete information stochastic games. Single controller stochastic games with incomplete information is presented in [48]. Here, authors interpret the incomplete information as partial information on the payoff matrix for one player. Hence, the other player knows the exact payoff matrix. In the future directions section of this paper, authors consider the case where each of the players has partial information on the payoff matrix. The incomplete information scheme in this paper extends the ideas in [3] to stochastic games. Hence, there is some probability distribution associated with the unknown payoff matrix to a player. Stochastic games with incomplete information on one side that have a single non-absorbing state have been studied in [55] and [56].

The key contributions of this report, both from an application and theoretical perspective, are as follows.

1. We present a new approach that accounts for robustness in homeland security decisions with emphasis on uncertainties inherent in the problem domain. Uncertain antagonistic nature of the problem is also addressed in our approach. We extend stochastic games into an incomplete information setting, where we interpret incomplete information as the unknown data of the game for each player that belong to a given uncertainty set.
2. We propose to cope with uncertainty in two ways: First, we consider that the adaptive behavior of the opponent is uncertain at a given stage. That is, our new approach takes into account the uncertainty present in possible conversions from one threat category to the other that are due to alternative selections of the adversaries in a given stage. Second, we treat payoffs to the adversaries as unknown parameters that belong to a bounded set.

3. The robust approach we adopt in this report not only focuses on future perturbations in problem parameters in a given time period, i.e., in payoffs and transition probabilities at a given current time period, but also on perturbations in data associated with threats likely to arise in future time periods.
4. Under suitable interpretations and assumptions, our new methodology readily lends itself to calculation of risk of a given threat category, even though there are uncertainties associated with payoffs to each player and with transition probabilities.
5. We propose a distribution free model for stochastic games with incomplete information on both sides, and on both payoffs and transition probabilities. This model presents an alternative approach to [48] and extends the ideas in [1] to stochastic games. In particular, in section 4, we prove the existence of equilibrium in general-sum n -person robust stochastic games that lends itself to a formulation of an equilibrium point via robust optimization. In section 5, we extend the ideas presented in [16] to their robust counterparts, inspired by the existing literature in robust optimization.
6. Our approach extends the ideas in certain parts of [13]. Specifically, we extend robust Markov decision processes with uncertain transition probabilities to the competitive case where there are more than one players. When there is only one player in the set of players, techniques of this report result in robust Markov decision processes with uncertain parameters that belong to bounded, closed, and convex sets. Furthermore, as mentioned earlier, we consider uncertainty in both transition probabilities and payoffs.

3 Methodology

As presented in section 1, the two important aspects of our problem are uncertainty and adversarial intentions. Moreover, uncertainty is not only associated with the cost / benefit parameters but also with the adversarial character of the opponents. That is, we also face the problem of uncertain adversarial behaviors of the opponent. In this sense, our problem intuitively is an optimization problem, where we wish to determine best strategies to adopt

against an opponent. Besides being adversarial, we face antagonistic intentions that cannot be captured using methods such as influence diagrams or decision trees. Moreover, the question these methods address are fundamentally related to determining optimal decisions against nature, which clearly is not antagonistic. In addition, such methods are suitable to determine pure actions. However, it is very likely for a decision maker to consider not only pure but also mixed strategies that are presented in the sequel.

As also presented in section 1, stochastic games readily lend themselves to modeling adversarial decision processes. They not only could be used to model adversarial behavior, but also to capture antagonistic intentions. As will be discussed in subsequent sections, risks of different threat categories could be obtained using this methodology. Nevertheless, stochastic games by themselves are still not sufficient for our purposes due to the crucial uncertainty aspect of our problem. For the reasons outlined in section 1, we clearly need a more realistic approach to cope with the requirements of our problem.

In this section, we introduce a novel approach, robust stochastic games. To this end, stochastic games with finite state and action sets are briefly introduced.

3.1 Stochastic Games

This section reviews basics of stochastic game theory, as presented in [54] and [17]. 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 play is in state s , player i can choose the action $a_s^i \in A_s^i$, where A_s^i is the set of actions of player i in state s . Suppose that each player makes a choice in state s , i.e., we have $a_s = (a_s^1, \dots, a_s^i, \dots, a_s^N)$. Then the game moves into state k with probability $P_{sa_s k} \geq 0$, $\sum_{k=1}^M P_{sa_s k} = 1$.

In the most general sense, stochastic games could be seen as a sequence of one-shot non-zero sum n person games. Values of the one-shot games to players are accumulated in the process. *Value of a stochastic game* for

player i starting the game in state s is defined as the total value to player i accumulated throughout the process if player i starts the game in state s . In discounted stochastic games player i discounts the values of the one-shot games to be played in the future by a factor β_i , $0 \leq \beta_i < 1$.

At each stage, players may consider using mixed strategies. Let x_s^i be the probability distribution over the set A_s^i with cardinality m_s^i . In other words, the probability vector for player i in state s is $x_s^i = (x_{s,1}^i, \dots, x_{s,m_s^i}^i)$, where $x_{s,k}^i \geq 0$, $\sum_{k=1}^{m_s^i} x_{s,k}^i = 1$. If we denote the set of mixed strategies of player i in state s by X_s^i , then X_s^i is a polytope given by

$$X_s^i = \{x_s^i \in \mathfrak{R}_+^{m_s^i} \mid \sum_{k=1}^{m_s^i} x_{s,k}^i = 1\}.$$

In this proposal, we consider a certain class of strategies as introduced by Shapley (1953), namely, stationary strategies. Stationary strategies prescribe a player the same probabilities for his choices each time the player visits a certain state, no matter what route she follows to reach that state. Let us represent the stationary strategies of a player i by $x^i = (x_1^i, \dots, x_M^i)$ and denote the set of mixed strategies of all players in the state space of the game by $x = (x^1, \dots, x^N)$. We denote mixed strategies of all players for all states except for player i by

$$x^{-i} = (x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^N).$$

The following notation is used to distinguish a mixed strategy of player i from those of others, for all states, and for a specific state s , respectively, as follows.

$$(x^{-i}, u^i) = (x^1, \dots, x^{i-1}, u^i, x^{i+1}, \dots, x^N).$$

Finally, we use the following notation.

$$X^i = \prod_{s \in S} X_s^i, \quad X_s = \prod_{i \in I} X_s^i, \quad \text{and} \quad X = \prod_{i \in I} X^i.$$

Suppose that players play with mixed strategies. Then a probability is associated with each realization of $a_s \in A_s$, where $A_s = \prod_{i=1}^N A_s^i$. Suppose that players choose their actions secretly (independently) at a given state. Then the probability associated with a_s is

$$\prod_{\substack{m=1 \\ m \neq i}}^N x_{s,a_s^m}^m u_{s,a_s^i}^i.$$

Then, expected cost to player i starting in state s is given, $\forall s \in S, i \in I$, by

$$g_s^i(x_s^{-i}, u_s^i; v^i) = \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m u_{s, a_s^i}^i \{C_{sa_s}^i + \beta_i \sum_{k=1}^M P_{sa_s k} v_k^i\}, \quad (1)$$

where $C_{sa_s}^i$ is the immediate cost to player i induced by a_s in state s and v_k^i is the value to player i if the next state is k . We interpret v_k^i as a cost incurred to player i in state k . As it is seen in the above equation, expected cost to player i is composed of his immediate expected cost in state s and expected total values of the games to be played in future stages. In this model, given the strategies of all other players in state s , i.e. x_s^{-i} , player i wishes to minimize his expected cost in s . This minimization, in turn, yields his value of the stochastic game starting in s . Hence, we obtain the following well known condition that the value vector for player i , i.e. $v^i = (v_1^i, \dots, v_M^i)$, must satisfy, if it exists.

$$v_s^i = \min_{u_s^i \in X_s^i} \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m u_{s, a_s^i}^i \{C_{sa_s}^i + \beta_i \sum_{k=1}^M P_{sa_s k} v_k^i\}, \quad \forall s \in S, i \in I. \quad (2)$$

It is in fact another result that, for any $x = (x^1, \dots, x^N) \in X$, there exists a unique corresponding value $v_s^i, \forall i \in I, \forall s \in S$. We are now ready for the following definition.

Definition. A tuple of strategies $x = (x^1, \dots, x^N) \in X$ is a *Nash equilibrium point* in a stochastic game if and only if, $\forall i \in I$ and $\forall s \in S$,

$$v_s^i(x^1, \dots, x^N) \leq v_s^i(x^{-i}, u^i), \forall u^i \in X^i. \quad (3)$$

When the above conditions hold, the value v_s^i is called the *optimal value* of the game to player i starting in state s and x^i is called the *optimal stationary strategies* for i . When (3) holds, we see that player i 's strategy x^i is a best answer against all other players' strategies x^{-i} , for all $i \in I$. Hence neither of the players has an incentive for a deviation from their respective strategies. In other words, once the equilibrium is reached, neither of the players individually wants to deviate from it.

It is now a very well known result that optimal values in stochastic games exist. This result was first found by Shapley (1953) for two-person zero-sum

stochastic games and later on was extended to the general-sum n player stochastic games by Fink (1964).

Equation (2) is a fundamental condition. It states that if a player knew how to play optimally from the next stage on, then, at the current stage, he would play with such strategies so that he minimizes the expected immediate cost at the current stage and also minimizes the expected costs possibly incurred in future stages. 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 next state an equivalent equilibrium definition for the purposes of developments in the sequel.

Definition. A point $x \in X$ is a Nash equilibrium in a stochastic game if and only if, $\forall i \in I$ and $\forall s \in S$, $\exists(x^{-i}, u^i)$ and $v = (v^1, \dots, v^N)$, such that,

$$v_s^i = \min_{u_s^i \in X_s^i} g_s^i(x_s^{-i}, u_s^i; v^i), \quad (4)$$

and

$$x_s^i \in \operatorname{argmin}_{u_s^i \in X_s^i} g_s^i(x_s^{-i}, u_s^i; v^i). \quad (5)$$

This definition states that x_s^{-i} is an optimal (stationary) strategy for player i in state s if, when Equation (4) is satisfied, the corresponding minimizer of the objective function of player i is a strategy that he always wishes to use against all other players' strategies when in state s . If this statement holds for all players and all states, then no player would wish to deviate from their strategies, resulting in an equilibrium. Player i starting in state s can use an arbitrary strategy and obtain Equation (4) and a corresponding value but, in return, other players may change their strategies that forces player i to establish equation (4) again. We look for such strategies, the use of which always makes all players reluctant to deviate from those strategies.

3.2 Robust Optimization

This section briefly reviews the basics of robust optimization, as introduced in [5]. Consider the following optimization problem, P_γ [5].

$$P_\gamma : \quad \min_{x \in \mathfrak{R}^m} f(x, \gamma)$$

$$s.t. \quad F(x, \gamma) \in K \subset \mathfrak{R}^m,$$

where $\gamma \in \mathfrak{R}^M$ is the data vector, $x \in \mathfrak{R}^n$ is the decision vector, and K is a convex cone.

Suppose that

- the data of P_γ is uncertain and all that is known about the data is that it belongs to an uncertainty set $U \in \mathfrak{R}^M$
- the constraints $F(x, \gamma) \in K$ must be satisfied no matter what the actual realization of $\gamma \in U$ is.

Now, consider the problem $P = \{P_\gamma\}_{\gamma \in U}$. 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:

$$P_R : \min_{x \in \mathfrak{R}^n} \{ \sup_{\gamma \in U} f(x, \gamma) \text{ s.t. } F(x, \gamma) \in K, \forall \gamma \in U \}.$$

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 [5].

Optimization of a linear program with column-wise uncertainty in the constraint matrix is first studied in [57]. Soyster's model is equivalent to an LP where all uncertain parameters are fixed to their corresponding worst case values, resulting in an over conservative approach. Ben-Tal and Nemirovski (1998) examine ellipsoidal uncertainty sets that relax the over conservative approach of Soyster's and they show that the robust counterpart of an LP with an ellipsoidal uncertainty set is a second order conic program. Nemirovski explains in [42] that robust counterpart of an optimization problem is not restricted to LPs. Furthermore, Ghaoui et al. (1998) consider semi-definite programs (SDPs) whose data belong to some uncertain set.

3.3 Formulation of Robust Stochastic Games

Starting in this section, we introduce our new approach. The rest of this proposal presents a new methodology and does not exist in the literature.

We have assumed so far in this proposal that costs to players and the transition probabilities of the game are known with certainty. However, this is usually not the case in practical applications. In fact, it is quite reasonable to consider that neither of the players knows exact costs incurred and/or

the transition probabilities to other states. Hence, it is natural to consider uncertainty sets associated with both the cost terms and the transition probabilities. Let us denote the uncertain cost coefficients and the uncertain transition probabilities for player i in state s by $\tilde{C}_{sa_s}^i$ and $\tilde{P}_{sa_s k}$, respectively. We assume that the uncertainty set for immediate payoffs is bounded. That is,

$$\tilde{C}_{sa_s}^i \in C_{sa_s}^i, \forall i \in I \text{ and } s \in S,$$

where $C_{sa_s}^i$ is bounded for every player and state. Furthermore,

$$\tilde{P}_{sa_s k} \in P_{sa_s k}, \forall i \in I \text{ and } s \in S,$$

where $\tilde{P}_{sa_s k}$ is a closed interval in $[0, 1]$, and $\sum_{k \in S} \tilde{P}_{sa_s k} = 1$. We seek whether an equilibrium point and optimal values exist in the case when probability transitions and the immediate costs at any state of the game are not known but it is known that they belong to their respective uncertainty sets. In other words, we are interested in whether a robust equilibrium point and corresponding robust optimal values exist. By a robust equilibrium point, we mean that optimal stationary strategies remain optimal as transition probabilities and immediate costs vary in their respective uncertainty sets. The proofs in this section parallel the ones in [17] to some extent, and establish the existence of an equilibrium point of a robust stochastic game where uncertainty sets are as given above. Note that results of the following proofs hold when the uncertainty set for probabilities belong to a set that is intersected by the probability simplex.

Now, in light of the results summarized in the previous section, we should notice the following observation. In our new robust stochastic game model, if robust values for player i exist, given x^{-i} , at optimality, they must satisfy the following, where the inner maximization problem is with respect to the uncertain transition probabilities and uncertain immediate costs.

$$\omega_s^i = \min_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m u_{s, a_s^i}^i \{ \tilde{C}_{sa_s}^i + \beta_i \sum_{k=1}^M \tilde{P}_{sa_s k} \omega_k^i \}. \quad (6)$$

The above equation states that if a player knew how to play in the robust stochastic game optimally from the next stage on, then, at the current stage, he would play with such strategies so that he *minimizes the maximum expected immediate cost* at the current stage and also *minimizes the*

maximum expected costs possibly incurred in future stages. 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. Note that we could have modeled each player as wishing to minimize his expected maximum total cost. However, this approach would be much more conservative than our model since in this case, each uncertain parameter would attain worst-case values in their respective uncertainty sets regardless of the mixed strategies chosen by players. In our robust stochastic game model, uncertain parameters attain their respective worst-case values with respect to the mixed strategies.

To ease the notation, let us define

$$\begin{aligned} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i) = \\ \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m u_{s, a_s^i}^i \{ \tilde{C}_{sa_s}^i + \beta_i \sum_{k=1}^M \tilde{P}_{sa_s k} \omega_k^i \}. \end{aligned}$$

Equation (6) now reads as follows.

$$\omega_s^i = \min_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i). \quad (7)$$

We will in fact show that such robust values exist.

Now, we are ready to state our definition of equilibrium in robust stochastic games.

Definition.

A point $x \in X$ is a *robust Nash equilibrium* point in a robust stochastic game if and only if, $\forall i \in I$ and $s \in S$, $\exists(x^{-i}, u^i)$ and $\omega = (\omega^1, \dots, \omega^N)$, such that,

$$\omega_s^i = \min_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i)$$

and

$$x_s^i \in \operatorname{argmin}_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i).$$

Equivalent to the above definition, a tuple of strategies $x = (x^1, \dots, x^N) \in X$ is a robust Nash equilibrium point in a robust stochastic game if and only if, $\forall i \in I$ and $\forall s \in S$,

$$w_s^i(x^1, \dots, x^N) \leq w_s^i(x^{-i}, u^i), \forall u^i \in X^i. \quad (8)$$

The following is obvious.

Proposition 1

In robust stochastic games, the following must hold, where $v_s^i(x^1, \dots, x^N)$ is the value of the nominal stochastic game to player i in state s .

$$w_s^i(x^1, \dots, x^N) \geq v_s^i(x^1, \dots, x^N), \quad \forall i \in I, s \in S \quad (9)$$

Proposition 1 states that the robust optimal value to player i in a robust stochastic game is always greater than or equal to the optimal value to player i in the respective nominal stochastic game, where data defining the game is known exactly. Inequality (9) must hold since in the robust approach, player i optimizes his objective function against the worst possible data scenario with respect to his mixed strategies. Conversely, in the nominal game with exact data, player i optimizes his objective function with respect to the nominal data values.

4 Existence of Equilibrium Points in Robust Stochastic Games

Our proof of existence of equilibrium points in a robust stochastic game parallels Fink's (1964). However, a different correspondence is defined that takes into account the robustness. Our correspondence uses a maximum expected total cost function with respect to mixed strategies. We show that the fixed point of a suitably constructed correspondence is an equilibrium point. Before this final step, an existing result and a definition, Kakutani's fixed point theorem and the definition of upper-semicontinuity for correspondences are stated. Then, a function pertaining to player i that returns the minimum cost to player i given the strategies of all other players is defined. It is shown that the function we use for robustness is a contracting function and therefore has a unique fixed point, which is the robust value (cost) of the game to player i . Next, it is shown that the robust value to a player

is bounded. Finally, it is shown that a suitably constructed correspondence satisfies the hypotheses of Kakutani's fixed point theorem. The following proofs require basic background on point to set mappings (correspondences) and contraction mappings.

Definition.

A correspondence ϕ that maps a closed, bounded, and convex set S into the family of closed, convex subsets of S is upper semi-continuous if

$$y^n \in \phi(x^n), \quad n = 1, 2, 3, \dots$$

$$\lim_{n \rightarrow \infty} x^n = x$$

$$\lim_{n \rightarrow \infty} y^n = y$$

imply that $y \in \phi(x)$.

Theorem 1 (*Kakutani's Fixed Point Theorem*). *If S is a closed, bounded, and convex set in a Euclidean space, and ϕ is an upper semi-continuous correspondence mapping S into the family of closed, convex subsets of S , then $\exists x \in S$, s.t. $x \in \phi(x)$.*

Let $W^i \equiv \{\omega_s^i \in \mathfrak{R}\}_{s \in S}$, $W \equiv \{W^i\}_{i \in I}$. Note that W is complete. Define the metric on W , $\rho(\omega, \theta) = \max_{i \in I, s \in S} |\omega_s^i - \theta_s^i|$.

Next, a transformation is defined. Given the strategies of all other players and an arbitrary robust value vector for player i , this transformation minimizes the maximum expected total cost with respect to the mixed strategies for player i . As a result of Theorem 2 below, it is justified that such a robust value vector exists for any given x^{-i} .

Theorem 2 *Let $\gamma_{s, x_s^{-i}}^i : W^i \rightarrow \mathfrak{R}$ be defined by*

$$\gamma_{s, x_s^{-i}}^i(\omega^i) = \min_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i).$$

For $x \in X$, define $\gamma_x(\omega) : W \rightarrow W$ by $(\gamma_x(\omega))_{is} = \gamma_{s, x_s^{-i}}^i(\omega^i)$. $\gamma_x(\omega)$ is a contraction mapping.

Proof. Let $\omega, \theta \in W$, $\beta = \max_{i \in I} \beta_i$. For x_s^{-i} fixed, $\forall i \in I, s \in S$,

$$\begin{aligned} \gamma_{s, x_s^{-i}}^i(\omega^i) &= \min_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i) \\ &= \psi_s^i(C_{sa_s}^{*i}(x_s^{-i}, u_s^{*i}), P_{sa_s k}^{*i}(x_s^{-i}, u_s^{*i}); x_s^{-i}, u_s^{*i}; \omega^i), \end{aligned}$$

where u_s^{*i} is the minimizer, and

$$C_{sa_s}^{*i}(x_s^{-i}, u_s^{*i}) \in C_{sa_s}^i \text{ and } P_{sa_s k}^{*i}(x_s^{-i}, u_s^{*i}) \in P_{sa_s k}$$

are the optimizers that now depend on (x_s^{-i}, u_s^{*i}) . Similarly, with z_s^{*i} being the minimizer, and for some

$$C'_{sa_s}{}^i(x_s^{-i}, z_s^{*i}) \in C_{sa_s}^i, P'_{sa_s k}{}^i(x_s^{-i}, z_s^{*i}) \in P_{sa_s k},$$

we have

$$\begin{aligned} \gamma_{s, x_s^{-i}}^i(\theta^i) &= \min_{z_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, z_s^i; \theta^i) \\ &= \psi_s^i(C'_{sa_s}{}^i(x_s^{-i}, z_s^{*i}), P'_{sa_s k}{}^i(x_s^{-i}, z_s^{*i}); x_s^{-i}, z_s^{*i}; \theta^i). \end{aligned}$$

Note that

$$\gamma_{s, x_s^{-i}}^i(\omega_s^i) \leq \psi_s^i(C'_{sa_s}{}^i(x_s^{-i}, z_s^{*i}), P'_{sa_s k}{}^i(x_s^{-i}, z_s^{*i}); x_s^{-i}, z_s^{*i}; \omega^i)$$

and

$$\gamma_{s, x_s^{-i}}^i(\theta_s^i) \leq \psi_s^i(C_{sa_s}^{*i}(x_s^{-i}, u_s^{*i}), P_{sa_s k}^{*i}(x_s^{-i}, u_s^{*i}); x_s^{-i}, u_s^{*i}; \theta^i).$$

Therefore,

$$\begin{aligned} &\gamma_{s, x_s^{-i}}^i(\omega_s^i) - \gamma_{s, x_s^{-i}}^i(\theta_s^i) \\ &= \psi_s^i(C_{sa_s}^{*i}(x_s^{-i}, u_s^{*i}), P_{sa_s k}^{*i}(x_s^{-i}, u_s^{*i}); x_s^{-i}, u_s^{*i}; \omega^i) \\ &\quad - \psi_s^i(C'_{sa_s}{}^i(x_s^{-i}, z_s^{*i}), P'_{sa_s k}{}^i(x_s^{-i}, z_s^{*i}); x_s^{-i}, z_s^{*i}; \theta^i) \\ &\leq \psi_s^i(C'_{sa_s}{}^i(x_s^{-i}, z_s^{*i}), P'_{sa_s k}{}^i(x_s^{-i}, z_s^{*i}); x_s^{-i}, z_s^{*i}; \omega^i) \\ &\quad - \psi_s^i(C'_{sa_s}{}^i(x_s^{-i}, z_s^{*i}), P'_{sa_s k}{}^i(x_s^{-i}, z_s^{*i}); x_s^{-i}, z_s^{*i}; \theta^i) \\ &= \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m z_s^{*i} \{ C'_{sa_s}{}^i(x_s^{-i}, z_s^{*i}) + \beta_i \sum_{k=1}^M P'_{sa_s k}{}^i(x_s^{-i}, z_s^{*i}) \omega_k^i \} \end{aligned}$$

$$\begin{aligned}
& - \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,a_s}^m z_s^{*i} \{C_{sa_s}^{\prime i}(x_s^{-i}, z_s^{*i}) + \beta_i \sum_{k=1}^M P_{sa_s k}^{\prime i}(x_s^{-i}, z_s^{*i}) \theta_k^i\} \\
& = \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,a_s}^m z_s^{*i} \beta_i \left\{ \sum_{k=1}^M P_{sa_s k}^{\prime i}(x_s^{-i}, z_s^{*i}) (\omega_k^i - \theta_k^i) \right\} \\
& \leq \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,a_s}^m z_s^{*i} \beta \left(\sum_{k=1}^M P_{sa_s k}^{\prime i}(x_s^{-i}, z_s^{*i}) \right) \rho(\omega, \theta) = \beta \rho(\omega, \theta).
\end{aligned}$$

Similar to the above arguments, for x_s^{-i} fixed, $\forall i \in I, s \in S$,

$$\begin{aligned}
& \gamma_{s, x_s^{-i}}^i(\theta_s^i) - \gamma_{s, x_s^{-i}}^i(\omega_s^i) \\
& = \psi_s^i(C_{sa_s}^{\prime i}(x_s^{-i}, z_s^{*i}), P_{sa_s k}^{\prime i}(x_s^{-i}, z_s^{*i}); x_s^{-i}, z_s^{*i}; \theta^i) \\
& \quad - \psi_s^i(C_{sa_s}^{*i}(x_s^{-i}, u_s^{*i}), P_{sa_s k}^{*i}(x_s^{-i}, u_s^{*i}); x_s^{-i}, u_s^{*i}; \omega^i) \\
& \leq \psi_s^i(C_{sa_s}^{*i}(x_s^{-i}, u_s^{*i}), P_{sa_s k}^{*i}(x_s^{-i}, u_s^{*i}); x_s^{-i}, u_s^{*i}; \theta^i) \\
& \quad - \psi_s^i(C_{sa_s}^{*i}(x_s^{-i}, u_s^{*i}), P_{sa_s k}^{*i}(x_s^{-i}, u_s^{*i}); x_s^{-i}, u_s^{*i}; \omega^i) \\
& \leq \beta \rho(\omega, \theta).
\end{aligned}$$

Thus,

$$\rho(\gamma_x(\omega), \gamma_x(\theta)) \leq \beta \rho(\omega, \theta).$$

□

The following theorem is an application of a well known result in fixed point theory to robust stochastic games (See [32]).

Theorem 3 *Application of Banach's Contraction Mapping Principle.* *Let $w(x)$ be the robust value vector corresponding to a stationary strategy $x \in X$. Then, for any $x \in X$,*

(a)

$$\omega_s^i(x) = \min_{x_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s; \omega^i(x)).$$

(b) $\omega(x)$ is bounded.

Proof. By Banach's contraction mapping principle $\gamma_x(\omega)$ has a unique fixed point, ω_x . That is, $\exists \omega_x$ such that $\gamma_x(\omega_x) = \omega_x$, which means

$$\omega_s^i(x) = \min_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i).$$

Hence, the first part of the Theorem follows immediately.

For the second part, let the n-fold $\gamma_x(\dots\gamma_x(\gamma_x(\omega))\dots)$ be denoted by $(\gamma_x)^n(\omega)$ and ω_0 be an arbitrary initial value vector of the robust stochastic game. Define the sequence $\omega_{n+1} = \gamma_x(\omega_n)$. Again by Banach's principle,

$$\lim_{n \rightarrow \infty} (\gamma_x)^n(\omega_0) = \lim_{n \rightarrow \infty} \omega_n = \omega_x. \quad (10)$$

Since γ_x is a contraction mapping, we have by Theorem 2 that

$$\rho(w_m, w_{m-1}) \leq \beta \rho(w_{m-1}, w_{m-2}) \leq \dots \leq \beta^{m-1} \rho(w_1, w_0).$$

Hence, using the triangle inequality gives

$$\begin{aligned} \rho(w_m, w_0) &\leq \rho(w_m, w_{m-1}) + \rho(w_{m-1}, w_{m-2}) + \dots + \rho(w_1, w_0) \\ &\leq \beta^{m-1} \rho(w_1, w_0) + \beta^{m-2} \rho(w_1, w_0) + \dots + \rho(w_1, w_0) \\ &= (\beta^{m-1} + \beta^{m-2} + \dots + 1) \rho(w_1, w_0) \leq \frac{1}{1 - \beta} \rho(w_1, w_0), \end{aligned}$$

therefore,

$$\lim_{m \rightarrow \infty} \rho(w_m, w_0) = \rho(w_x, w_0) \leq \frac{1}{1 - \beta} \rho(w_1, w_0).$$

Take $w_0 = 0$. Then,

$$\begin{aligned} \rho(w_x, w_0) &\leq \frac{1}{1 - \beta} \max_{i \in I, s \in S} |\gamma_{s, x_s^{-i}}^i(0)| \\ &= \max_{i \in I, s \in S} \left| \min_{x_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m u_{s, a_s^i}^i \tilde{C}_{sa_s}^i \right| \\ &= \max_{i \in I, s \in S} \left| \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m x_{s, a_s^i}^{*i} C_{sa_s}^{*i}(x^{-i}, x^{*i}) \right|, \end{aligned}$$

where x^{*i} is the optimum, and

$$\begin{aligned}
&\leq \max_{i \in I, s \in S} \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m} x_{s, a_s^i}^{*i} \{ \max_{a_s \in A_s} \{ \sup_{\tilde{C}_{sa_s}^i \in \tilde{C}_{sa_s}^i} |\tilde{C}_{sa_s}^i| \} \} \\
&= \max_{i \in I, s \in S} 1 \{ \max_{a_s \in A_s} \{ \sup_{\tilde{C}_{sa_s}^i \in C_{sa_s}^i} |\tilde{C}_{sa_s}^i| \} \} \\
&= \max_{\substack{i \in I, s \in S \\ a_s \in A_s}} \sup_{\tilde{C}_{sa_s}^i \in C_{sa_s}^i} |\tilde{C}_{sa_s}^i|,
\end{aligned}$$

which is bounded since $C_{sa_s}^i$ is bounded. Thus, $\omega(x)$ is bounded for any $x \in X$. □

The above theorem is a very strong result. It first states that there exists a unique robust value ω_s^i such that when it is given to the transformation $\gamma_{s, x_s^{-i}}^i(\cdot)$ as an argument for any fixed x_s^{-i} , the output of the transformation coincides with the robust value ω_s^i given as an argument to the transformation. Formally, $\forall x^{-i} \in X^{-i} = \prod_{k=1, k \neq i}^N X^k$ and given any arbitrary robust value vector w , minimizing player i 's maximum expected total cost for each state yields back the same value vector w . Second, it states that if, for any fixed x_s^{-i} , we apply the above transformation starting with an arbitrary robust value $\omega_s^i(0)$ over and over again, outputs of these transformations converge to the unique fixed point of the transformation.

Let us now recall the vector function $\psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s, \omega^i)$ defined previously as

$$\psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s, \omega^i) = \sum_{a_s \in A_s} \prod_{m=1}^N x_{s, a_s^m} \{ \tilde{C}_{sa_s}^i + \beta_i \sum_{k=1}^M \tilde{P}_{sa_s k} \omega_k^i \}.$$

We need the following lemma to show that the maximum expected cost functions we use satisfy the properties needed to use Kakutani's theorem. A portion of the proof of the following lemma makes use of an algebraic identity used in [1].

Lemma 1 *Let $p = (x_s, w^i)$, $q = (u_s, \theta^i)$. Define the metrics*

$$d_{X_s}(x_s, u_s) = \max_{i \in I} |x_s^i - u_s^i|, \quad d_{W^i}(w^i, \theta^i) = \max_{s \in S} |w_s^i - \theta_s^i|,$$

and

$$d_1(p, q) = d_{X_s}(x_s, u_s) + d_{W^i}(w^i, \theta^i).$$

Given $\epsilon > 0$, $\exists \delta(\epsilon) > 0$ such that if for any $p, q \in X_s \times \mathfrak{R}^M$,

$$d_1(p, q) < \delta(\epsilon),$$

then, $\forall \tilde{C}_{sa_s}^i \in C_{sa_s}^i$, $\forall \tilde{P}_{sa_s k} \in P_{sa_s k}$,

$$\left| \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s, \omega^i) - \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; u_s, \theta^i) \right| < \epsilon.$$

Proof. Since, $\tilde{C}_{sa_s}^i \in C_{sa_s}^i$ and $C_{sa_s}^i$ is bounded $\forall i \in I, s \in S$, we have

$$\left| \tilde{C}_{sa_s}^i \right| \leq K,$$

where $1 < K < \infty$.

Note that by Theorem 3, robust values are bounded. Hence, we have $\forall i \in I, s \in S$, that

$$\left| \omega_s^i \right| \leq W,$$

where $1 < W < \infty$. Note that

$$\begin{aligned} & \left| \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s, \omega^i) - \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; u_s, \theta^i) \right| \\ &= \left| \sum_{a_s \in A_s} \prod_{m=1}^N x_{s, a_s^m}^m \tilde{C}_{sa_s}^i + \beta_i \sum_{a_s \in A_s} \left(\prod_{m=1}^N x_{s, a_s^m}^m \right) \left(\sum_{k=1}^M \tilde{P}_{sa_s k} \omega_k^i \right) \right. \\ & \quad \left. - \sum_{a_s \in A_s} \prod_{m=1}^N u_{s, a_s^m}^m \tilde{C}_{sa_s}^i - \beta_i \sum_{a_s \in A_s} \left(\prod_{m=1}^N u_{s, a_s^m}^m \right) \left(\sum_{k=1}^M \tilde{P}_{sa_s k} \theta_k^i \right) \right| \\ &= \left| \sum_{a_s \in A_s} \tilde{C}_{sa_s}^i \left(\prod_{m=1}^N x_{s, a_s^m}^m - \prod_{m=1}^N u_{s, a_s^m}^m \right) + \beta_i \sum_{a_s \in A_s} \sum_{k=1}^M \tilde{P}_{sa_s k} \left(\prod_{m=1}^N x_{s, a_s^m}^m \omega_k^i - \prod_{m=1}^N u_{s, a_s^m}^m \theta_k^i \right) \right| \\ &\leq \left| \sum_{a_s \in A_s} \tilde{C}_{sa_s}^i \left(\prod_{m=1}^N x_{s, a_s^m}^m - \prod_{m=1}^N u_{s, a_s^m}^m \right) \right| + \left| \beta_i \sum_{a_s \in A_s} \sum_{k=1}^M \tilde{P}_{sa_s k} \left(\prod_{m=1}^N x_{s, a_s^m}^m \omega_k^i - \prod_{m=1}^N u_{s, a_s^m}^m \theta_k^i \right) \right| \\ &\leq K \sum_{a_s \in A_s} \left| \prod_{m=1}^N x_{s, a_s^m}^m - \prod_{m=1}^N u_{s, a_s^m}^m \right| + \beta_i \sum_{a_s \in A_s} \sum_{k=1}^M \left| \prod_{m=1}^N x_{s, a_s^m}^m \omega_k^i - \prod_{m=1}^N u_{s, a_s^m}^m \theta_k^i \right|. \end{aligned} \tag{11}$$

Let

$$\delta_1(\epsilon) = \frac{\min\{\epsilon, 1\}}{3K(2^N - 1) \prod_{i=1}^N m_s^i}, \delta_2(\epsilon) = \frac{\min\{\epsilon, 1\}}{3M\beta_i \prod_{i=1}^N m_s^i},$$

$$\delta_3(\epsilon) = \frac{\min\{\epsilon, 1\}}{3WM\beta_i(2^N - 1) \prod_{i=1}^N m_s^i},$$

and let

$$\delta(\epsilon) = \min\{\delta_1(\epsilon), \delta_2(\epsilon), \delta_3(\epsilon)\}.$$

Now,

$$d_1(p, q) < \delta(\epsilon)$$

implies that, $\forall i \in I, s \in S$, and $\forall a_s^i \in A_s^i$,

$$x_{s,a_s^m}^m = u_{s,a_s^m}^m + \alpha_{s,a_s^m}^m \quad \text{and} \quad \omega_s^i = \theta_s^i + \gamma_s^i,$$

where $|\alpha_{s,a_s^m}^m| < \delta(\epsilon)$, and $|\gamma_s^i| < \delta(\epsilon)$. Furthermore, as Aghassi and Bertsimas (2004) show, we have the following algebraic identity.

$$\left| \prod_{m=1}^N (u_{s,a_s^m}^m + \alpha_{s,a_s^m}^m) - \prod_{m=1}^N u_{s,a_s^m}^m \right| = \left| \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left(\prod_{m \in I} \alpha_{s,a_s^m}^m \right) \left(\prod_{m \in I'} u_{s,a_s^m}^m \right) \right|,$$

where $I' = \{1, \dots, N\} \setminus I$. Note that

$$\prod_{m \in I} |\alpha_{s,a_s^m}^m| < (\gamma_1(\epsilon))^{|I|} \leq \gamma_1(\epsilon),$$

and that

$$\left| \prod_{m=1}^N (u_{s,a_s^m}^m + \alpha_{s,a_s^m}^m) - \prod_{m=1}^N u_{s,a_s^m}^m \right| \leq \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{s,a_s^m}^m \right| \left| \prod_{m \in I'} u_{s,a_s^m}^m \right|.$$

Hence, for the first term in (12), we have

$$K \sum_{a_s \in A_s} \left| \prod_{m=1}^N (u_{s,a_s^m}^m + \alpha_{s,a_s^m}^m) - \prod_{m=1}^N u_{s,a_s^m}^m \right|$$

$$\begin{aligned}
&\leq K \sum_{a_s \in A_s} \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{s, a_s^m}^m \right| \left| \prod_{m \in I'} u_{s, a_s^m}^m \right| \\
&\leq K \sum_{a_s \in A_s} \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{s, a_s^m}^m \right| < K \sum_{a_s \in A_s} \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \gamma_1(\epsilon) = \frac{\epsilon}{3}.
\end{aligned}$$

Now consider the second term in (12). We have

$$\begin{aligned}
&\beta_i \sum_{a_s \in A_s} \sum_{k=1}^M \left| \prod_{m=1}^N x_{s, a_s^m}^m \omega_k^i - \prod_{m=1}^N u_{s, a_s^m}^m \theta_k^i \right| \\
&= \beta_i \sum_{a_s \in A_s} \sum_{k=1}^M \left| \prod_{m=1}^N (u_{s, a_s^m}^m + \alpha_{s, a_s^m}^m) \omega_k^i - \prod_{m=1}^N u_{s, a_s^m}^m \theta_k^i \right| \\
&= \beta_i \sum_{a_s \in A_s} \sum_{k=1}^M \left| \prod_{m=1}^N u_{s, a_s^m}^m (\omega_k^i - \theta_k^i) + \omega_k^i \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \prod_{m \in I} \alpha_{s, a_s^m}^m \prod_{m \in I'} u_{s, a_s^m}^m \right| \\
&\leq \beta_i \sum_{a_s \in A_s} \sum_{k=1}^M \left| \prod_{m=1}^N u_{s, a_s^m}^m \right| |\omega_k^i - \theta_k^i| \\
&\quad + \beta_i W \sum_{a_s \in A_s} \sum_{k=1}^M \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{s, a_s^m}^m \right| \left| \prod_{m \in I'} u_{s, a_s^m}^m \right| \\
&\leq \beta_i \sum_{a_s \in A_s} \sum_{k=1}^M |\gamma_s^i| + \beta_i W \sum_{a_s \in A_s} \sum_{k=1}^M \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \left| \prod_{m \in I} \alpha_{s, a_s^m}^m \right| \\
&< \beta_i \sum_{a_s \in A_s} \sum_{k=1}^M \delta_2(\epsilon) + \beta_i W \sum_{a_s \in A_s} \sum_{k=1}^M \sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \delta_3(\epsilon) \\
&= \frac{\epsilon}{3} + \frac{\epsilon}{3} = 2\frac{\epsilon}{3}.
\end{aligned}$$

Thus,

$$\begin{aligned}
& \left| \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s, \omega^i) - \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; u_s, \theta^i) \right| \\
& \leq K \sum_{a_s \in A_s} \left| \prod_{m=1}^N x_{s, a_s^m}^m - \prod_{m=1}^N u_{s, a_s^m}^m \right| + \beta_i \sum_{a_s \in A_s} \sum_{k=1}^M \left| \prod_{m=1}^N x_{s, a_s^m}^m \omega_k^i - \prod_{m=1}^N u_{s, a_s^m}^m \theta_k^i \right| \\
& < \frac{\epsilon}{3} + 2\frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

□

Let the vector function $f(x_s^{-i}, u_s^i; \omega^i)$ be defined as follows.

$$f(x_s^{-i}, u_s^i; \omega^i) = \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i).$$

Lemma 2 *By Lemma 1, it immediately follows that $f(x_s^{-i}, u_s^i; \omega^i)$ is continuous $\forall i \in I$, and $s \in S$.*

Lemma 3 *$f(x_s^{-i}, u_s^i; \omega^i)$ is convex in u_s^i for fixed x_s^{-i} and ω^i .*

Proof. Suppose that $y_s^i, z_s^i \in X_s^i$. Note that, $\forall \lambda \in [0, 1]$,

$$\begin{aligned}
& f(x_s^{-i}, (\lambda y_s^i + (1-\lambda)z_s^i); \omega^i) \\
& = \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m (\lambda y_s^i + (1-\lambda)z_s^i) \\
& \quad \times \{C_{sa_s}^{*i}(x_s^{-i}, \lambda y_s^i + (1-\lambda)z_s^i) + \beta \sum_{k=1}^M P_{sa_s k}^*(x_s^{-i}, \lambda y_s^i + (1-\lambda)z_s^i) \omega^i\},
\end{aligned}$$

where optimizers $C_{sa_s}^{*i}$ and $P_{sa_s k}^*$ now depend on $(x_s^{-i}, \lambda y_s^i + (1-\lambda)z_s^i)$. Hence,

$$\begin{aligned}
& f(x_s^{-i}, (\lambda y_s^i + (1-\lambda)z_s^i); \omega^i) \\
& = \lambda \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m y_{s, a_s^i}^i \{C_{sa_s}^{*i}(x_s^{-i}, \lambda y_s^i + (1-\lambda)z_s^i) + \beta \sum_{k=1}^M P_{sa_s k}^*(x_s^{-i}, \lambda y_s^i + (1-\lambda)z_s^i) \omega^i\} \\
& \quad + (1-\lambda) \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s, a_s^m}^m z_{s, a_s^i}^i \{C_{sa_s}^{*i}(x_s^{-i}, \lambda y_s^i + (1-\lambda)z_s^i) + \beta \sum_{k=1}^M P_{sa_s k}^*(x_s^{-i}, \lambda y_s^i + (1-\lambda)z_s^i) \omega^i\}
\end{aligned}$$

$$\begin{aligned}
&\leq \lambda \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,a_s^m}^m y_{s,a_s^i}^i \{C_{sa_s}^{*i}(x_s^{-i}, y_s^i) + \beta \sum_{k=1}^M P_{sa_s k}^*(x_s^{-i}, y_s^i) w^i\} \\
&+(1-\lambda) \sum_{a_s \in A_s} \prod_{\substack{m=1 \\ m \neq i}}^N x_{s,a_s^m}^m z_{s,a_s^i}^i \{C_{sa_s}^{*i}(x_s^m, z_s^i) + \beta \sum_{k=1}^M P_{sa_s k}^*(x_s^m, z_s^i) w^i\} \\
&= \lambda f(x_s^{-i}, y_s^i; \omega^i) + (1-\lambda) f(x_s^{-i}, z_s^i; \omega^i).
\end{aligned}$$

□

Next we define a function for player i that returns the fixed point of $\gamma_{s, x_s^{-i}}^i$ given the strategies of all other players except i .

Define

$$\begin{aligned}
\beta^i(x^{-i}) &= \{\omega^i = (\omega_1^i, \dots, \omega_M^i)\} \\
\omega_s^i &= \min_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i), \quad s = 1, \dots, M,
\end{aligned}$$

and denote the s^{th} element of $\beta^i(x^{-i})$ by $\beta_s^i(x^{-i})$. By earlier results stating that the value of a robust stochastic game is bounded, we obtain by definition that $\beta_s^i(x^{-i})$ is bounded $\forall i \in I, \forall s \in S, \forall x^{-i} \in X^{-i}$.

Lemma 4 *If $x_s^{i, n_k} \rightarrow x_s^i$ and $\beta_s^i(x^{-i, n_k}) \rightarrow \omega_s^i$, then the robust value ω_s^i is a fixed point of $\gamma_{s, x_s^{-i}}^i$, i.e., $\beta_s^i(x^{-i}) = \omega_s^i$.*

Theorem 4 (Existence of Equilibrium in Robust Stochastic Games)

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

Proof. We now apply Kakutani's fixed point theorem. To this end, let

$$Y^i = \{y^i = (y_1^i, \dots, y_M^i) \in \prod_{s=1}^M X_s^i\}$$

such that

$$y_s^i \in \operatorname{argmin}_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i),$$

and

$$\omega_s^i = \min_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i), s = 1, \dots, M\}.$$

Define the correspondence $\phi(x^1, \dots, x^N)$

$$= \{(y^1, \dots, y^N) \in \prod_{i=1}^N X^i \mid y^i \in Y^i, i = 1, \dots, N\}.$$

Now, we show that $\phi(x^1, \dots, x^N) \neq \emptyset$. Note that, by Theorem 3, $\forall i \in I$, $\forall s \in S$, and for any fixed x_s^{-i} , there exist unique robust values ω_s^i that satisfies

$$\omega_s^i = \min_{u_s^i \in X_s^i} \max_{\substack{\tilde{C}_{sa_s}^i \in C_{sa_s}^i \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^i(\tilde{C}_{sa_s}^i, \tilde{P}_{sa_s k}; x_s^{-i}, u_s^i; \omega^i).$$

Note furthermore that, by lemmas 2 and 3, $f(x_s^{-i}, u_s^i; \omega^i)$ is convex in u_s^i and continuous. By Bolzano-Weierstrass theorem that states that a continuous function on a non-empty compact set such as $X_s^i \subset \mathfrak{R}_+^{m_s^i}$ achieves a minimum in X_s^i , we obtain that $Y^i \neq \emptyset$. Therefore, the result that $\phi(x^1, \dots, x^N) \neq \emptyset$ follows.

Note that by definition, $\phi(x) \subseteq X$, $\forall x \in X$.

Next, we show that $\phi(x^1, \dots, x^N)$ is a convex set. Suppose that

$$(z^1, \dots, z^N), (v^1, \dots, v^N) \in \phi(x^1, \dots, x^N).$$

Then, $\forall u_s^i$, and $s \in S, i \in I$,

$$\begin{aligned} f(x_s^{-i}, z_s^i; \omega^i) &= f(x_s^{-i}, v_s^i; \omega^i) \\ &\leq f(x_s^{-i}, u_s^i; \omega^i) \end{aligned}$$

Hence, for any $\lambda \in [0, 1]$ and $\forall i \in I, s \in S$,

$$\begin{aligned} \lambda f(x_s^{-i}, z_s^i; \omega^i) + (1 - \lambda) f(x_s^{-i}, v_s^i; \omega^i) \\ \leq f(x_s^{-i}, u_s^i; \omega^i) \end{aligned}$$

By the convexity of $f(x_s^{-i}, u_s^i; \omega^i)$, we obtain

$$\begin{aligned} & f(x_s^{-i}, ((\lambda)z_s^i + (1 - \lambda)v_s^i); \omega^i) \\ & \leq \lambda f(x_s^{-i}, z_s^i; \omega^i) + (1 - \lambda)f(x_s^{-i}, v_s^i; \omega^i) \\ & \leq f(x_s^{-i}, u_s^i; \omega^i), \end{aligned}$$

and hence,

$$(\lambda)(z^1, \dots, z^N) + (1 - \lambda)(v^1, \dots, v^N) \in \phi(x^1, \dots, x^N).$$

Finally, we must show that $\phi(x^1, \dots, x^N)$ is an upper semi-continuous correspondence. (A good treatment of correspondences could be found in [27] and [9], along with Kakutani's fixed point theorem.)

Suppose for $n = 1, 2, \dots$,

$$\begin{aligned} & (x^{1,n}, \dots, x^{N,n}) \in X, \\ & (y^{1,n}, \dots, y^{N,n}) \in \phi(x^{1,n}, \dots, x^{N,n}), \\ & \lim_{n \rightarrow \infty} (x^{1,n}, \dots, x^{N,n}) = (u^1, \dots, u^N) \in X, \\ & \lim_{n \rightarrow \infty} (y^{1,n}, \dots, y^{N,n}) = (q^1, \dots, q^N) \in X. \end{aligned}$$

Now, $\forall r_s^i \in X_s^i, s = 1, \dots, M$,

$$f(x_s^{-i,n}, y_s^{i,n}; \beta(x^{-i,n})) \leq f(x_s^{-i,n}, r_s^i; \beta(x^{-i,n})).$$

It follows by the continuity of $f(x_s^{-i}, u_s^i; \omega^i)$ and by Lemma 4 that,

$$\begin{aligned} & \lim_{n \rightarrow \infty} f(x_s^{-i,n}, y_s^{i,n}; \beta(x^{-i,n})) = f(u_s^{-i}, q_s^i; \omega_s^i) \\ & \leq f(u_s^{-i}, r_s^i; \omega_s^i) = \lim_{n \rightarrow \infty} f(x_s^{-i,n}, r_s^i; \beta(x^{-i,n})), \end{aligned}$$

therefore,

$$(q^1, \dots, q^N) \in \phi(u^1, \dots, u^N),$$

which completes the proof that ϕ is an upper semi-continuous correspondence.

Note that ϕ is an upper semi-continuous correspondence that maps the closed, bounded, and convex set X into the family of closed, convex subsets of X . Therefore, ϕ satisfies the assumptions of Kakutani's fixed point theorem and the proof is complete. \square

5 Calculation of an Equilibrium Point

In general-sum n -person stochastic games, one cannot capture the fully antagonistic intentions of the players. Nevertheless, completely antagonistic goals make sense in the two-person zero-sum case, a special case of general-sum n -person stochastic games. In such games, a player's (usually player 1) gain is the cost to the other player (usually player 2). Hence, there is a complete utility transfer from one player to the other and payoffs to the players sum up to zero.

Let $\omega_{(x^1, x^2)}$ denote the value of the game to player 2, induced by arbitrary stationary strategies (x^1, x^2) . Now that $\omega_{(x^1, x^2)}^1 = -\omega_{(x^1, x^2)}^2$ in the zero-sum case, similar to nominal stochastic games (with known parameters), equilibrium conditions for robust two-person zero-sum stochastic games (condition 8) reduce to

$$\omega_{(x^1, x^{*2})} \leq \omega_{(x^{*1}, x^{*2})} \leq \omega_{(x^{*1}, x^2)}, \quad (12)$$

where $\omega_{(x^{*1}, x^{*2})}$ denote the optimal value of the game, induced by the optimal stationary strategies (x^{*1}, x^{*2}) . In this section, we extend the nonlinear programming formulations for two-person zero-sum stochastic games given by Filar and Vrieze (1997) to the programs for robust two-person zero-sum stochastic games. Now, as we have seen in the previous section, we have the following for robust stochastic games, where ω_s is used to represent robust value of the game to player 2 starting in state s . We should bear in mind that any cost incurred to player 2 is a gain to player 1.

$$\omega_s = \min_{x_s^2 \in S_s^2} \max_{\substack{\tilde{C}_{sa_s}^2 \in C_{sa_s}^2 \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \psi_s^2(\tilde{C}_{sa_s}^2, \tilde{P}_{sa_s k}; x_s^1, x_s^2; \omega)$$

or

$$\begin{aligned} \omega_s &= \min_{x_s^2 \in S_s^2} \max_{\substack{\tilde{C}_{sa_s}^2 \in C_{sa_s}^2 \\ \tilde{P}_{sa_s k} \in P_{sa_s k}}} \sum_{a_s \in A_s} x_{s, a_s}^1 x_{s, a_s}^2 \left\{ \tilde{C}_{sa_s}^2 + \beta \sum_{k=1}^M \tilde{P}_{sa_s k} \omega_k \right\} \\ &= \sum_{a_s \in A_s} x_{s, a_s}^1 x_{s, a_s}^{*2} \left\{ C_{sa_s}^{*2}(x_s^1, x_s^{*2}) + \beta \sum_{k=1}^M P_{sa_s k}^*(x_s^1, x_s^{*2}) \omega_k \right\}, \end{aligned}$$

where, optimizers $C_{sa_s}^{*2}$ and $P_{sa_s k}^*$ now depend on strategies. Consider the

right hand side.

$$\begin{aligned}
& \sum_{a_s \in A_s} x_{s,a_s^1}^1 x_{s,a_s^2}^{*2} \{ C_{sa_s}^{*2}(x_s^1, x_s^{*2}) + \beta \sum_{k=1}^M P_{sa_s k}^*(x_s^1, x_s^{*2}) w_k \} \\
&= \sum_{a_s \in A_s} x_{s,a_s^1}^1 x_{s,a_s^2}^{*2} C_{sa_s}^{*2}(x_s^1, x_s^{*2}) + \\
& \beta \sum_{k=1}^M \left(\sum_{a_s \in A_s} x_{s,a_s^1}^1 x_{s,a_s^2}^{*2} P_{sa_s k}^*(x_s^1, x_s^{*2}) \right) \left\{ \sum_{a_k \in A_k} x_{k,a_k^1}^1 x_{k,a_k^i}^{*2} C_{ka_k}^{*2}(x_k^1, x_k^{*2}) \right. \\
& \quad \left. + \beta \sum_{k'=1}^M \sum_{a_k \in A_k} x_{k,a_k^1}^1 x_{k,a_k^i}^{*2} P_{ka_s k'}^*(x_k^1, x_k^{*2}) w_{k'} \right\} \\
&= \sum_{a_s \in A_s} x_{s,a_s^1}^1 x_{s,a_s^2}^{*2} C_{sa_s}^{*2}(x_s^1, x_s^{*2}) \\
& + \beta \sum_{k=1}^M \left(\sum_{a_s \in A_s} x_{s,a_s^1}^1 x_{s,a_s^2}^{*2} P_{sa_s k}^*(x_s^1, x_s^{*2}) \right) \left(\sum_{a_k \in A_k} x_{k,a_k^1}^1 x_{k,a_k^i}^{*2} C_{ka_k}^{*2}(x_k^1, x_k^{*2}) \right) \\
& + \beta^2 \sum_{k=1}^M \left(\sum_{a_s \in A_s} x_{s,a_s^1}^1 x_{s,a_s^2}^{*2} P_{sa_s k}^*(x_s^1, x_s^{*2}) \right) \left(\sum_{k'=1}^M \sum_{a_k \in A_k} x_{k,a_k^1}^1 x_{k,a_k^2}^{*2} P_{ka_s k'}^*(x_k^1, x_k^{*2}) w_{k'} \right).
\end{aligned}$$

The last term in the last equality could be rewritten as

$$\beta^2 \sum_{k=1}^M \left\{ \left(\sum_{a_s \in A_s} x_{s,a_s^1}^1 x_{s,a_s^2}^{*2} P_{sa_s k}^*(x_s^1, x_s^{*2}) \right) \sum_{k'=1}^M \left(\sum_{a_k \in A_k} x_{k,a_k^1}^1 x_{k,a_k^i}^{*2} P_{ka_s k'}^*(x_k^1, x_k^{*2}) \right) w_{k'} \right\},$$

where there is a single number in each of the two most inner parenthesis. The inner parentheses are in fact one-step transition probabilities in their respective uncertainty sets from one state to the other, induced by the mixed strategies of the players in state s . Let us denote the one step transition probabilities induced by (x_s^1, x_s^{*2}) by

$$\Gamma_{sa_s k}^*(x_s^1, x_s^{*2}) \in P_{sa_s k}, \quad a_s \in A_s.$$

Now, the last term could be rewritten as follows.

$$\begin{aligned} & \beta^2 \sum_{k'=1}^M \sum_{k=1}^M \Gamma_{sa_s k}^*(x_s^1, x_s^{*2}) \Gamma_{ka_k k'}^*(x_k^1, x_k^{*2}) w_{k'} \\ &= \beta^2 \sum_{k'=1}^M \Gamma_{sa_s k'}^{*2}(x^1, x^{*2}) w_{k'}, \end{aligned}$$

where $\Gamma_{sa_s k'}^{*2}(x^1, x^{*2})$ is the 2-step transition probabilities from s to k' . Note that the first term in the RHS is in fact the expected immediate payoff. Let the expected immediate payoff induced by (x^1, x^{*2}) be denoted, in matrix form, by

$$E_s^2[C^*(x^1, x^{*2})] = \left[x_{s,a_s^1}^1 x_{s,a_s^2}^{*2} C_{sa_s}^{*2}(x_s^1, x_s^{*2}) \right]_{a_s^1=1, a_s^2=1}^{m_s^1, m_s^2}.$$

Let

$$P^*(x^1, x^{*2}) = \left[\sum_{a_s \in A_s} x_{s,a_s^1}^1 x_{s,a_s^2}^{*2} P_{sa_s k}^*(x_s^1, x_s^{*2}) \right]_{s=1, k=1}^{M, M}.$$

Define the following $M \times 1$ vector:

$$E^2[C^*(x^1, x^{*2})] = [\mathbf{1}^T E_s^2[C^*(x^1, x^{*2})] \mathbf{1}]_{s=1}^M,$$

where $\mathbf{1}$ is a vector of ones of the appropriate dimension. Using this notation, the RHS reduces to the following in matrix notation.

$$E^2[C^*(x^1, x^{*2})] + \beta P^*(x^1, x^{*2}) E^2[C^*(x^1, x^{*2})] + \beta^2 P^{*2}(x^1, x^{*2}) \omega,$$

If we substituted the above equation into itself n times, we would have obtained

$$\begin{aligned} \omega &= E^2[C^*(x^1, x^{*2})] + \beta P^*(x^1, x^{*2}) E^2[C^*(x^1, x^{*2})] \\ &+ \beta^2 P^{*2}(x^1, x^{*2}) E^2[C^*(x^1, x^{*2})] + \dots + \beta^n P^{*n}(x^1, x^{*2}) \omega. \end{aligned}$$

Let $n \rightarrow \infty$ to obtain

$$\omega_{(x^1, x^{*2})} = [I - \beta P^*(x^1, x^{*2})]^{-1} E^2[C^*(x^1, x^{*2})], \quad (13)$$

where I is the $M \times M$ identity matrix and $\omega_{(x^1, x^{*2})}$ is the robust value of the game to player 2, induced by strategies (x^1, x^{*2}) .

Suppose that players play with (x^1, x^{*2}) . Let $R_{t,(x^1, x^{*2})}$, $t = 0, 1, \dots$ denote sequence of costs at each stage of the game. Then, starting in state s , expected immediate cost at stages $0, 1, 2, \dots, n$ would, in matrix form, be as follows.

$$\begin{aligned} E^2[R_{0,(x^1, x^{*2})}] &= E^2[C^*(x^1, x^{*2})] \\ E^2[R_{1,(x^1, x^{*2})}] &= P^*(x^1, x^{*2})E^2[C^*(x^1, x^{*2})] \\ E^2[R_{2,(x^1, x^{*2})}] &= P^{*2}(x^1, x^{*2})E^2[C^*(x^1, x^{*2})] \\ &\vdots \\ E^2[R_{n,(x^1, x^{*2})}] &= P^{*n}(x^1, x^{*2})E^2[C^*(x^1, x^{*2})] \end{aligned}$$

Now, we can alternatively calculate the robust value of the game induced by (x^{-i}, x^{*2}) to player 2 starting in state s as follows.

$$\begin{aligned} \omega_{(x^1, x^{*2})} &= \sum_{t=0}^{\infty} \beta^t E^2[C^*(x^1, x^{*2})] P^{*t}(x^1, x^{*2}) \\ &= E^2[C^*(x^1, x^{*2})] + \beta P^*(x^1, x^{*2})E^2[C^*(x^1, x^{*2})] \\ &\quad + \beta^2 P^{*2}(x^1, x^{*2})E^2[C^*(x^1, x^{*2})] + \dots \\ &= E^2[C^*(x^1, x^{*2})] \{I + \beta P^*(x^1, x^{*2}) + \beta^2 P^{*2}(x^1, x^{*2}) + \dots\}, \end{aligned}$$

where I is the $M \times M$ identity matrix, and hence,

$$\omega_{(x^1, x^{*2})} = [I - \beta P^*(x^1, x^{*2})]^{-1} E^2[C^*(x^1, x^{*2})],$$

which is same as the solution found in Eq (14). Now, suppose that for the arbitrary robust value vector θ^i and arbitrary x we have

$$E^2[C^*(x)] + \beta P^*(x)\theta \leq \theta.$$

Iterating the left-hand-side of the above equation into itself infinitely many times, we obtain

$$E^2[C^*(x)] + \beta P^*(x)E^2[C^*(x)] + \beta^2 P^{*2}(x)E^2[C^*(x)] + \dots \leq \theta.$$

Now, let $n \rightarrow \infty$ to obtain,

$$\omega_{(x)} \leq \theta,$$

and we see that the arbitrary robust value vector θ is an upper bound on the value vector of the game induced by arbitrary x . This observation may lead

us to an optimization problem, where we would like to minimize an arbitrary (variable) robust value vector subject to some constraints that will prescribe us robust equilibrium conditions (13).

Define

$$E_s^2[\tilde{C}(x^2)] = \left[x_{s,a_s^2}^2 \tilde{C}_{sa_s} \right]_{a_s^1=1, a_s^2=1}^{m_s^1, m_s^2}.$$

Define

$$T(s, \theta) = \left[\sum_{k \in S} \tilde{P}_{sa_s k} \theta_k \right]_{a_s^1=1, a_s^2=1}^{m_s^1, m_s^2}.$$

Claim: Consider the following uncertain nonlinear program (RNL_1) in matrix form, bearing in mind that w and x^2 are the decision variables and the superscript denotes player 2:

$$\min \sum_{k=1}^M \omega_k$$

s.t.

$$\begin{aligned} \max_{\tilde{C}_{sa_s} \in C_{sa_s}} E_s^2[\tilde{C}(x^2)] \mathbf{1} + \beta \max_{\tilde{P}_{sa_s k} \in P_{sa_s k}} T(s, \omega) x_s^2 \leq \omega_s \mathbf{1}, \forall s \in S \quad (14) \\ (x_s^1, x_s^2) \in S_s^1 \times S_s^2, \quad \forall s \in S, \end{aligned}$$

where $\mathbf{1}$ is a vector of ones of appropriate dimension, and ω and x^2 represent an arbitrary (variable) robust value vector (for player 2) and an arbitrary (variable) robust stationary strategies. Let $\omega_{(x^{*1}, x^{*2})}^*$ and x^{*2} be the optimal robust value vector and robust optimal stationary strategy for player 2, respectively. Then, ω^* and x^{*2} form a global minimum of RNL_1 .

Proof. By the existence theorem for robust stochastic games proved in the previous section, ω^* and x^{*2} exist. Multiply the objectives to be maximized in (15) from the left by arbitrary x_s^1 to get

$$\begin{aligned} \max_{\tilde{C}_{sa_s} \in C_{sa_s}} [x_s^1]^T E_s^2[\tilde{C}(x^{*2})] \mathbf{1} + \beta \max_{\tilde{P}_{sa_s k} \in P_{sa_s k}} [x_s^1]^T T(s, \omega^*) x_s^{*2} \\ \leq [x_s^1]^T \omega_s^* \mathbf{1}, \forall s \in S, \end{aligned}$$

which, since \tilde{C} and \tilde{P} belong to bounded sets, could be rewritten as

$$E^2[C^*(x^1, x^{*2})] + \beta P^*(x^1, x^{*2}) \omega^* \leq \omega^*,$$

where $C^*(x^1, x^{*2})$ and $P^*(x^1, x^{*2})$ are optimizers that depend on strategies (x^1, x^{*2}) . Iterate the above equation into itself infinitely many times to get

$$\omega_{(x^1, x^{*2})}^* \leq \omega^*,$$

which holds by condition (13). Therefore, ω^* and x^{*2} form a feasible point.

Now suppose that ω and x^2 form an arbitrary feasible point. If we follow the above procedure by multiplying the constraint by an arbitrary x^1 , we obtain

$$E^2[C^*(x^1, x^2)] + \beta P^*(x^1, x^2)\omega \leq \omega$$

By the same iterative procedure we obtain

$$\omega_{(x^1, x^2)} \leq \omega, \quad \forall x^1 \in S^1,$$

which also holds with x^{*1} since x^1 is arbitrary, that is,

$$\omega_{(x^{*1}, x^2)} \leq \omega.$$

But by (13), we also have

$$\omega_{(x^{*1}, x^{*2})} \leq \omega_{(x^{*1}, x^2)} \leq \omega.$$

The proof is now complete, since ω is arbitrary and equilibrium points exist by the existence theorem for robust stochastic games. \square

5.1 Preliminary Results

5.1.1 A Small Numerical Example

Example 1: Consider the robust stochastic game in Figure 1.

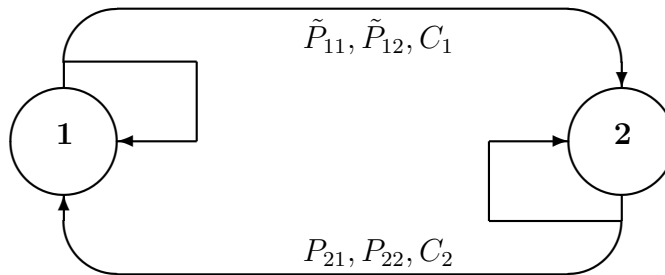


Figure 1. Robust Stochastic Game Example 1

In this example, we interpret states as threat categories and actions in each state as alternatives. In Figure 1, the uncertain probability transition matrices \tilde{P}_{1j} , $j = 1, 2$, and other certain transition matrices are as follows.

$$\tilde{P}_{12} = \begin{pmatrix} (0.3, 0.7) & (0.4, 0.7) \\ 0 & 0 \end{pmatrix}, P_{21} = [0.5], P_{22} = [0.5]$$

$$\tilde{P}_{11} = \begin{pmatrix} (0.4, 0.8) & (0.2, 0.5) \\ 1 & 1 \end{pmatrix}.$$

We label the rows of P_{11} and P_{12} by a and b , respectively, to denote the alternatives of player 1. We label their columns by c and d , to represent alternatives of player 2. Hence, in the second state, each player has one alternative that we label by e and f for player 1 and player 2, respectively. Whenever there is a vector as an entry of the above matrices, we mean that the corresponding transition probability is uncertain but is within the upper and lower bounds defined by the entries of that vector. For example, if players choose alternatives a and d in state 1, then the lower bound on the transition probability to state 1 again is 0.2. The upper bound in this case is 0.5. If players choose a and c in state 1, then the game moves into state 2 with a lower bound of 0.3 and an upper bound of 0.7. If there is a scalar as an entry of any matrix, this means that the corresponding transition probability is known with certainty.

In this example, for simplicity, we define the uncertainty set for transition probabilities simply to be the convex hull of lower and upper bounds defined in the above matrices. This implies that the uncertain transition probabilities could be any convex combinations of lower and upper bounds. We observe that an uncertain probability not necessarily attains its respective upper or lower bound. In other words, the probabilities that maximize the expected cost to be minimized by player 2 are convex combinations of the bounds. Therefore, the set we consider for probabilities are, $\forall s \in S, k \in S, \forall a_s \in A_s$, as follows.

$$P_{sa_s k} = \left\{ \tilde{P} = P_{sa_s k}^0 + \sum_{l=1}^2 \lambda_{sa_s k}^l P_{sa_s k}^l \mid \lambda_{sa_s k}^l \geq 0, \sum_{l=1}^2 \lambda_{sa_s k}^l \leq 1 \right\}.$$

Note that in our model, $P_{sa_s k}^0$ are all equal to zero for simplicity. However, we could have modeled this example by determining two bounds and a nominal

probability P^0 , and then by requiring the uncertain probabilities to belong to the convex hull of these three values.

We will also demonstrate later in this research that these probabilities could be obtained by a suitably constructed linear program, as an outcome of the nonlinear program given below.

For simplicity, we assume that the immediate costs in this example are known with certainty. Let

$$C_1 = \begin{pmatrix} 4 & 6 \\ 7 & 3 \end{pmatrix}, C_2 = [2],$$

where labels on the rows and columns of C_1 and C_2 are the same as any probability transition matrix out of state 1 and 2, respectively. The entries in the matrices denote the immediate costs to player 2 that correspond to the actions taken. For example, if a and d are chosen by the players in state 1, then player 2's cost is 6 units.

The nonlinear program RNL_1 for Example 1, is as follows. Note that the following formulation is small since some transitions are assumed to be certain in Example 1.

$$\min \omega_1 + \omega_2$$

s.t.

$$\sum_{a_s^2 \in A_s^2} C_{s(a_s^1, a_s^2)} x_{s, a_s^2}^2 + \beta \max_{\tilde{P} \in P_{s(a_1, a_2)k}} \sum_{k=1}^M \sum_{a_s^2 \in A_s^2} \tilde{P}_{s(a_1, a_2)k} x_{s, a_s^2}^2 \omega_k \leq \omega_k, \forall s \in S, \forall a_s^1 \in A_s^1$$

$$(x_s^1, x_s^2) \in S_s^1 \times S_s^2, \quad \forall s \in S$$

Next, we rewrite the above by considering the uncertainty set that the probabilities belong to, and by appending the additional constraints that define the uncertainty set.

$$RNL_2 := \min \omega_1 + \omega_2$$

s.t.

$$\begin{aligned} & \max_{\lambda^l, l=1,2} \sum_{a_s^2 \in A_s^2} C_{s(a_s^1, a_s^2)} x_{s, a_s^2}^2 \\ & + \beta \sum_{k=1}^M \sum_{a_s^2 \in A_s^2} \left(\sum_{l=1}^2 \lambda_{sa_s k}^l P_{sa_s k}^l \right) x_{s, a_s^2}^2 \omega_k \leq \omega_s, \quad \forall s \in S, \forall a_s^1 \in A_s^1 \end{aligned}$$

$$\begin{aligned}
& \sum_{l=1}^2 \lambda_{sa_s k}^l \leq 1 \\
& \lambda_{sa_s k}^l \geq 0, \quad l = 1, 2 \\
& (x_s^1, x_s^2) \in S_s^1 \times S_s^2, \quad \forall s \in S \\
& \sum_{k=1}^M \left(\sum_{l=1}^2 \lambda_{sa_s k}^l P_{sa_s k} \right) = 1, \quad \forall s \in S, \forall a_s \in A_s
\end{aligned}$$

Now, it is important to note that uncertain probabilities out of any given state must sum up to 1. Hence, we must add the last constraint above that captures this requirement.

Consider the inner maximization problems:

$$\begin{aligned}
& \max_{\lambda^l, l=1,2} \sum_{a_s^2 \in A_s^2} C_{s(a_s^1, a_s^2)} x_{s, a_s^2}^2 \\
& + \beta \sum_{k=1}^M \sum_{a_s^2 \in A_s^2} \left(\sum_{l=1}^2 \lambda_{sa_s k}^l P_{sa_s k}^l \right) x_{s, a_s^2}^2 \omega_k \leq \omega_s, \quad \forall s \in S, \forall a_s^1 \in A_s^1
\end{aligned}$$

s.t.

$$\sum_{k=1}^M \left(\sum_{l=1}^2 \lambda_{sa_s k}^l P_{sa_s k} \right) = 1, \quad \forall s \in S, \forall a_s \in A_s \quad (15)$$

$$\sum_{l=1}^2 \lambda_{sa_s k}^l \leq 1, \quad \forall s \in S, \forall a_s \in A_s \quad (16)$$

$$\lambda_{sa_s k}^l \geq 0, \quad l = 1, 2, \quad \forall s \in S, \forall a_s \in A_s$$

The number of inner maximization problems is equal to the cardinality $|A_s^1|$ at state s . Note that the constraints are linear in λ . We next associate the dual variables corresponding to (16) and (17).

$$(16) \rightarrow \mu_{s(a_s^1, a_s^2)}, \quad \forall s \in S, \forall a_s \in A_s,$$

$$(17) \rightarrow \mu_{s(a_s^1, a_s^2)k}, \quad \forall s, k \in S, \forall a_s \in A_s,$$

and write the set of duals D_1 as follows.

$$\begin{aligned} & \min \sum_{a_s^2 \in A_s^2} \mu_{s(a_s^1, a_s^2)} + \sum_{k=1}^M \sum_{a_s^2 \in A_s^2} \mu_{s(a_s^1, a_s^2)k} \\ & \leq \omega_s - \sum_{a_s^2 \in A_s^2} C_{s(a_s^1, a_s^2)} x_{s, a_s^2}^2, \quad \forall s \in S, \forall a_s^1 \in A_s^1 \end{aligned}$$

s.t.

$$\begin{aligned} & P_{s(a_s^1, a_s^2)k}^l \mu_{s(a_s^1, a_s^2)} + \mu_{s(a_s^1, a_s^2)k} \geq \\ & \beta P_{s(a_s^1, a_s^2)k}^l x_{s, a_s^2}^2 \omega_s, \quad l = 1, 2, \forall s, k \in S, \forall a_s^1 \in A_s^1, \forall a_s^2 \in A_s^2 \\ & \mu_{s(a_s^1, a_s^2)k} \geq 0, \forall s, k \in S, \forall a_s^1 \in A_s^1, \forall a_s^2 \in A_s^2. \end{aligned}$$

Substituting D_1 into RNL_2 , we obtain the following.

$$RNL_3 := \min \omega_1 + \omega_2$$

s.t.

$$\begin{aligned} & \sum_{a_s^2 \in A_s^2} \mu_{s(a_s^1, a_s^2)} + \sum_{k=1}^M \sum_{a_s^2 \in A_s^2} \mu_{s(a_s^1, a_s^2)k} \leq \omega_s - \sum_{a_s^2 \in A_s^2} C_{s(a_s^1, a_s^2)} x_{s, a_s^2}^2, \quad \forall s \in S, \forall a_s^1 \in A_s^1 \\ & P_{s(a_s^1, a_s^2)k}^l \mu_{s(a_s^1, a_s^2)} + \mu_{s(a_s^1, a_s^2)k} \geq \\ & \beta P_{s(a_s^1, a_s^2)k}^l x_{s, a_s^2}^2 \omega_s, \quad l = 1, 2, \forall s, k \in S, \forall a_s^1 \in A_s^1, \forall a_s^2 \in A_s^2 \\ & \mu_{s(a_s^1, a_s^2)k} \geq 0, \forall s, k \in S, \forall a_s^1 \in A_s^1, \forall a_s^2 \in A_s^2. \end{aligned}$$

Returning to our example, note that it only has uncertain transition probabilities in state 1 and they are only associated with our opponent's alternative a . Therefore, RNL_3 results only in one dual problem with additional original formulation constraints corresponding to alternatives and states with certain data. Using the data as given above, RNL_3 is numerically as follows.

$$\min \omega_1 + \omega_2$$

s.t.

$$\mu_{1ac} + \mu_{1ad} + \mu_{1ac1} + \mu_{1ac2} + \mu_{1ad1} + \mu_{1ad2} \leq w_1 - 4x_{1c}^2 - 6x_{1d}^2$$

$$\begin{aligned}
0.4\mu_{1ac} + \mu_{1ac1} &\geq 0.3x_{1c}^2w_1 \\
0.8\mu_{1ac} + \mu_{1ac1} &\geq 0.6x_{1c}^2w_1 \\
0.3\mu_{1ac} + \mu_{1ac2} &\geq 0.225x_{1c}^2w_2 \\
0.7\mu_{1ac} + \mu_{1ac2} &\geq 0.525x_{1c}^2w_2 \\
0.2\mu_{1ad} + \mu_{1ad1} &\geq 0.15x_{1d}^2w_1 \\
0.5\mu_{1ad} + \mu_{1ad1} &\geq 0.375x_{1d}^2w_1 \\
0.4\mu_{1ad} + \mu_{1ad2} &\geq 0.3x_{1d}^2w_2 \\
0.7\mu_{1ad} + \mu_{1ad2} &\geq 0.525x_{1d}^2w_2 \\
\mu_{1ac1} \geq 0, \mu_{1ac2} \geq 0, \mu_{1ad1} \geq 0, \mu_{1ad2} \geq 0 \\
7x_{1c}^2 + 3x_{1d}^2 + (0.75)1w_1x_{1c}^2 + (0.75)1w_1x_{1d}^2 &\leq w_1 \\
2x_{2f}^2 + (0.375)x_{2f}^2w_2 + (0.375)x_{2f}^2w_1 &\leq w_2 \\
x_{1c}^2 \geq 0, x_{1d}^2 \geq 0, x_{2f}^2 \geq 0 \\
x_{1c}^2 + x_{1d}^2 &= 1 \\
x_{2f}^2 &= 1.
\end{aligned}$$

Example 1 is the perturbed version of the same example with certain (nominal) transition matrices as follows.

$$\begin{aligned}
P_{12} &= \begin{pmatrix} 0.4 & 0.666 \\ 0 & 0 \end{pmatrix}, P_{21} = [0.5], P_{22} = [0.5] \\
P_{11} &= \begin{pmatrix} 0.6 & 0.333 \\ 1 & 1 \end{pmatrix}.
\end{aligned}$$

This nominal problem has the following nonlinear formulation that is presented here using the (nominal) formulations in [16].

$$\begin{aligned}
&\min \omega_1 + \omega_2 \\
&\text{s.t.} \\
&4x_{1c}^2 + 6x_{1d}^2 + 0.75(0.6)\omega_1x_{1c}^2 + 0.75(0.4)\omega_2x_{1c}^2 + 0.75(0.333)\omega_1x_{1d}^2 + 0.75(0.666)\omega_2x_{1d}^2 \leq \omega_1; \\
&7x_{1c}^2 + 3x_{1d}^2 + 0.75(1)\omega_1x_{1c}^2 + 0.75(1)\omega_1x_{1d}^2 \leq \omega_1
\end{aligned}$$

Table 1: Solutions to nominal and robust stochastic games

Decision Variables	Nominal Stochastic Game	Robust Stochastic Game
objective value	28.96241	30.50092
ω_1	16.10151	17.06308
ω_2	12.86090	13.43785
x_{1c}^2	0.2563442	0.3164422
x_{1d}^2	0.7436558	0.6835578
x_{2f}^2	1	1
μ_{1ac}	NA	3.189226
μ_{1ac}	NA	6.889158
μ_{1ac1}	NA	0.6883055
μ_{1ac2}	NA	0
μ_{1ad1}	NA	0.9292704
μ_{1ad2}	NA	0
global optimal	28.96241	30.50092
local optimal	28.96241	30.50092

$$2x_{2f}^2 + 0.375x_{2f}^2\omega_2 + 0.375x_{2f}^2\omega_1 \leq \omega_2$$

$$x_{1c}^2 \geq 0, x_{1d}^2 \geq 0, x_{2f}^2 \geq 0$$

$$x_{1c}^2 + x_{1d}^2 = 1, x_{2f}^2 = 1$$

The above problems are solved using the nonlinear (convex, nonconvex) optimization software LINGO. Table 1 depicts the optimal values associated with the nominal and the robust stochastic game solutions.

Some observations on Table 1 are in order:

1. First and foremost, it is a result that every local optima in a nominal stochastic game is also a global optima. Hence, global and local optimal values for a nominal stochastic game coincide. We see in the above table that global and local optimal values for our robust example coincide. A natural question that needs attention is whether the same nice property for nominal problems holds for robust stochastic games.
2. We note in the robust stochastic game formulation section that Proposition 1 must hold. It holds for our example, raising the question of

formally proving this property.

3. It is possible to obtain the values for uncertain parameters the robust model uses, using the dual variables in Table 1, and constructing a suitable linear program.
4. The optimal strategies for the two problems differ. Robust solutions suggest player 2 to put slightly more emphasis on his first alternative, and consequently, a little less prominence on his second alternative in state 1.
5. The change in optimal strategies in this example could be deemed to be small. Suppose that our decision maker in this example interprets optimal strategy values as percentages and wishes to allocate her funds to his alternatives based on these measures. The difference between her first alternatives in state 1 is 0.060098. Clearly, given that the investments could be at very significant monetary magnitudes, even very small changes in the output of this small model could result in significant changes in investment decisions.
6. Risk is defined as the multiplication of likelihood of an event and its severity. Using suitable interpretations and assumptions, we could calculate risk readily from this model as follows. One could run the underlying Markov chain in Example 1 starting with the transitions matrix induced by the optimal strategies of the opponents. Limiting probabilities in each state then could be multiplied by the optimal strategy values obtained from our nonlinear robust optimization model for each state to obtain risk values.

5.1.2 More Analysis in Example 1

We next consider uncertain immediate costs in state 1, when player 1 chooses to play with his alternative a . Immediate cost data is now as follows.

$$C_1 = \begin{pmatrix} (2, 7) & (4, 10) \\ 7 & 3 \end{pmatrix}, C_2 = [2].$$

Table 2: Robust vs Nominal Worst Case Solutions

Decision Variables	Robust Game	Nom. W-Case Perf.	Perc Saving
objective value	43.2215	46.08051	6.6 %
ω_1	25.0134	26.8003	7.1 %
ω_2	18.2080	19.2801	5.8 %
x_{1c}^2	0.8133	0.2563	NA
x_{1d}^2	0.1866	0.7436	NA
x_{2f}^2	1	1	NA

Similarly, uncertain immediate costs are assumed to belong to convex combinations of their respective lower and upper bounds. That is, we have

$$C_{sa_s} = \{\tilde{C} = C_{sa_s}^0 + \sum_{l=1}^2 \tau_{sa_s}^l C_{sa_s}^l \mid \tau_{sa_s}^l \geq 0, \sum_{l=1}^2 \tau_{sa_s}^l \leq 1\}.$$

It is assumed that $C_{sa_s}^0 = 0, \forall s \in S, a_s \in A_s$. This version of the problem is solved using duality and similar arguments used in the previous version. The data and the solution for the nominal game in this version is the same as in the previous one. Results are summarized in Tables 2 and 3. Robust Game column in Table 2 refers to the solution of the robust stochastic game with given uncertain immediate costs and transitions probabilities. The next column refers to the performance of the nominal solution when parameters attain their worst-case values with respect to it. The worst-case parameter values with respect to the nominal solution are obtained from a suitably constructed linear program that could be extracted from the program to solve the new version of Example 1. Percentage savings resulting from the use of the robust solution versus the use of nominal solution under worst-case data scenarios are depicted in the last column.

Table 3 depicts the performance of the robust solution when data is certain. If the data were certain and had we adopted the robust solution, resulting percentage loss values due to the use of the robust solution under the nominal data scenario are presented in the last column. We see in this example that the percentage savings are higher than respective percentage losses. It is important also to note that uncertainty is present in a small portion of the data defining this example. It is reasonable to expect that more significant savings could result in a model with fully uncertain data.

Table 3: Performance of Robust Solution under Nominal Data

Decision Variables	Nom Game	Rob Perf w/ Nom Data	Perc Loss
objective value	28.9624	30.4927	5.3 %
ω_1	16.1015	17.0579	5.9 %
ω_2	12.8609	13.4347	4.4 %
x_{1c}^2	0.2563	0.8133	NA
x_{1d}^2	0.7436	0.1866	NA
x_{2f}^2	1	1	NA

6 Conclusions

Challenges in modeling homeland security related decision making problems arise from the antagonistic character inherent in the problem and the uncertainty associated with it. The *uncertain* adaptive behaviour of the attacker in such problems requires a modeling technique that could also capture the uncertainty inherent in the problem. In most cases, point estimate values are not easy to obtain from experts. For instance, it is very natural that an expert would be much more comfortable to give intervals of probabilities and costs, rather than giving single numerical values. The transition probabilities from a given threat to the other based on the decisions made and the costs of these decisions are difficult to predict. The new technique introduced in this report suggests a way to cope with such uncertainties and also accounts for the antagonism in the problem.

This report demonstrates existence of equilibrium points in robust stochastic games and suggests a computation method based on duality. This method is demonstrated on a small numerical example and results are summarized. It is observed in this small example that the percentage savings resulting from using robust strategies versus the nominal strategies when the parameters attain their worst-case values are higher than the losses caused by using robust strategies when parameters attain their nominal values. It is also observed that compared to the uncertainty in transition probabilities, uncertainty in immediate costs has a greater effect on the value of the game.

The future work in this research includes a quantified model of the MANPADS case study. A qualitative model that includes defined state-action pairs for a two player zero-sum robust stochastic game is in development phase. MANPADS model is proposed to be quantified via expert elicitation. The next step after developing the model is the computation of the robust optimal

strategies.

References

- [1] Aghassi, M., Bertsimas, D. (2004). “Robust Game Theory,” To appear, revised November, 2004.
- [2] Atkinson, S. E., Sandler, T., and Tschirhart, J. T. (1987). “Terrorism in a bargaining framework”, *Journal of Law and Economics*, Vol.30, No.1.
- [3] Aumann, R. J., and Maschler, M. B. (1968). “Repeated games of incomplete information: The zerosum extensive case,” in *Report of the U.S. Arms Control and Disarmament Agency ST-143*, Washington, D.C., Chapter III, pp. 37116.
- [4] Azaiez, N., Bier, V. M. (2004). “Optimal Resource Allocation for Security in Reliability Systems,” Working Paper, Industrial Engineering Department, University of Wisconsin Madison.
- [5] Ben-Tal, A., Nemirovski, A. (1998). “Robust Convex Optimization,” *Mathematics of Operations Research*, Vol.23, No.4.
- [6] Bernard, H. (2004). “Mathematical Methods in Combating Terrorism,” *Risk Analysis*, Vol.24, No.4.
- [7] Bier, V. (2004). “Game-Theoretic and Reliability Methods in Counter-Terrorism and Security,” Technical Report, Center for Human Performance and Risk Analysis, University of Wisconsin Madison.
- [8] Bier, V., Nagaraj, A., Abhichandani, V. (2004). “Protection of simple series and parallel systems with components of different values,” *Reliability Engineering and Systems Safety*, Vol.10, No.27.
- [9] Border, Kim C. (1985). *Fixed Point Theorems with Applications to Economics and Game Theory*, Cambridge University Press.
- [10] Brynielsson, J., Arnborg, S. (2004). “Bayesian Games for Threat Prediction and Situation Analysis,” Technical Report, Department of Numerical Analysis and Computer Science, Royal Institute of Technology, Stockholm, Sweden.

- [11] Cruz, Jr., J. B., Simaan M. A., Gacic, A., and Liu, Y. (2002). “Moving horizon Nash strategies for a military air operation,” *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, No.3.
- [12] DArtigues, A., Vignolo, T. (2003). “Why Global Integration May Lead to Terrorism: An Evolutionary Theory of Mimetic Rivalry,” *Economics Bulletin*, Vol.6, No.11.
- [13] L. El Ghaoui, A. Nilim. (2004). “Robust solutions to markov decision problems with uncertain transition matrices,” *Operations Research*, 2004. To appear.
- [14] Faria, J. R. (2003). “Terror Cycles,” *Studies in Nonlinear Dynamics and Econometrics*, Vol.7, No.1.
- [15] Faria, J. R. (2004). “Terrorist Innovations and Anti-Terrorist Policies”, Political Economy Working Paper, School of Social Sciences, University of Texas at Dallas.
- [16] Filar, J., Vrieze, K. (1997). *Competitive Markov Decision Processes*, Springer-Verlag, New York.
- [17] Fink, A.M. (1964). “Equilibrium in a Stochastic N-Person Game,” *Journal of Science in Hiroshima University*, Series A-I, vol.28 pp.89-93.
- [18] Ghaoui, L. E., Oustry, F., and Lebret, H. (1998). “Robust Solutions to Uncertain Semidefinite Programs,” *SIAM J. Optim.*, Vol.9, No.1, pp.33-52.
- [19] Guo, X., Hernandez, O. (2004). “Zero Sum Games for Nonhomogeneous Markov Chains with an Expected Average Payoff Criterion,” *Appl. Comput. Math*, Vol. 3, No.1.
- [20] Haimes, Y. Y. (2002). “Roadmap for Modeling Risks of Terrorism to the Homeland,” *Journal of Infrastructure Systems*, June 2002.
- [21] Haimes, Y. Y. (2002). “Strategic Responses to Risks of Terrorism to Water Resources,” *Journal of Water Resources Planning and Management*, Vol.128.

- [22] Harsanyi, J. C. (1967, 1968). “Games with Incomplete Information Played by Bayesian Players, I-III,” *Management Science*, Vol. 14, pp.159-182, pp.320-334, pp.486-502.
- [23] Hausken, K. (2002). “Probabilistic risk analysis and game theory,” *Risk Analysis*, Vol.22.
- [24] Hayashi, S., Yamashita, N., and Fukushima, M. (2004). “Robust Nash equilibria and second-order cone complementarity problems,” Technical Report 2004-004, Department of Applied Mathematics and Physics, Kyoto University, April, 2004.
- [25] Hazen, G. B. (2002). “Stochastic Trees and the StoTree Modeling Environment: Models and Software for Medical Decision Analysis,” *Journal of Medical Systems*, Vol.26.
- [26] Heal, G., Kunreuther, H. (2003). “You Only Die Once: Managing Discrete Interdependent Risks”, Working Paper Columbia Business School and Wharton Risk Management and Decision Processes Center.
- [27] Hildenbrand W., Kirman, A.P. (1976). *Introduction to Equilibrium Analysis*. North Hollan Publishing Company.
- [28] Howard, R. A., Matheson, J.E. (1981). “Influence Diagrams,” *Principles and Applications of Decision Analysis*, Vol.2.
- [29] Hudson, L. D., Ware, B. S., Laskey, K. B., and Mahoney, S. M. (2002). “An Application of Bayesian Networks to Antiterrorism Risk Management for Military Planners,” Technical Report, Digital Sandbox, Inc.
- [30] Jaskiewicz, A. (2002). “Zero Sum Semi-Markov Games,” *SIAM Journal on Control and Optimization*, Vol. 41, No. 3.
- [31] Kakutani, S. (1941). “A Generalization of Brouwer’s Fixed Point Theorem,” *Duke Mathematical Journal*, Vol.8, pp.457-59.
- [32] Kirk, B., W. A., Sims, B. (2001). *Handbook of Metric Fixed Point Theory*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- [33] Koller, D.,Milch, B. (2001). “Multi-agent influence diagrams for representing and solving games,” in *Proceedings of the 17th International Joint Conference of Artificial Intelligence*.

- [34] Kunreuther, H., Heal, G. (2003). "Interdependent security," *Journal of Risk and Uncertainty*, Vol.26, No.2.
- [35] LaMura, P. (2000). "Game Networks," in *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*.
- [36] Lapan, H. E., Sandler, T. (1988). "To bargain or not to bargain: That is the question," *American Economic Review*, Vol.78, No.2.
- [37] Lapan, H. E., Sandler, T. (1993). "Terrorism and signaling," *European Journal of Political Economy*, Vol.9, No.3.
- [38] Laraki, R., Solan, E. (2002). "Stopping Games in Continuous Time," *SIAM Journal on Control and Optimization*, to appear.
- [39] T. S. Laskey, K. B., Levitt T. S. (2002). "Multisource Fusion for Opportunistic Detection and Probabilistic Assessment of Homeland Terrorist Threats," Technical Report, SEOR Department, George Mason University.
- [40] Major, J. (2002). "Advanced techniques for modeling terrorism risk," *Journal of Risk Finance*, Vol.4.
- [41] Nash, J. F. (1950). "Equilibrium Points in n-Person Games," *Proceedings of the National Academy of Sciences of the United States of America*, Vol.36, No.1, pp.48-49.
- [42] Nemirovski, A. *Lectures on Modern Convex Optimization*. Technion, Israel Institute of Technology.
- [43] Parthasarathy, T., Raghavan, T.E.S. (1981). "An Orderfield Property for Stochastic Games When One Player Controls Transition Probabilities," *Journal of Optimization Theory and Applications*, Vol. 33, 375-392.
- [44] Parthasarathy, T., Tijis, S.H., and Vrieze, O. J. (1984). *Stochastic Games with State Independent Transition and Separable Rewards*, Selected topics in OR and Mathematical Economics, Springer-Verlag, Lecture Note Series 226, New York.

- [45] Pate-Cornell, E., Guikema, S. (2002). "Probabilistic modeling of terrorist threats: A systems analysis approach to setting priorities among countermeasures," *Military Operations Research*, Vol.7, No.4.
- [46] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, San Francisco.
- [47] Robust Optimization Lecture Notes. Epstein Department of Industrial and Systems Eng., Fall 2004.
- [48] Rosenberg, D., Solan, E., and Vieille, N. (2004). "Stochastic Games with a Single Controller and Incomplete Information," *SIAM Journal of Control and Optimization*, Vol. 43, No.1.
- [49] Sandler, T., Arce M., D. G. (2003). "Terrorism and Game Theory," *Simulation and Gaming*, Vol.34, No.3.
- [50] Sandler, T., Tschirhart, T. T., and Cauley, J. (1983). "A theoretical analysis of transnational terrorism," *American Political Science Review*, Vol.77, No.4.
- [51] Sandler, T., Siqueira, K. (2002). "Global terrorism: Deterrence versus preemption," Unpublished manuscript, University of Southern California.
- [52] Shachter, R. D. (1986). "Evaluating influence diagrams," *Operations Research*, Vol.34, No.6.
- [53] Singh, S., Tu, H., Allanach, J., Pattipati, K.R., and Willett, P. (2004). "Stochastic Modeling of a Terrorist Event via the ASAM System", in *Proceedings of the International Conference on Systems, Man and Cybernetics*, The Hague, The Netherlands.
- [54] Shapley, L.S. (1953). "Stochastic Games," *Proceeding of the National Academy of Science*, 39, 1095-1100.
- [55] Sorin, S. (1984). "Big Match with Lack of Information on One Side (Part 1)," *International Journal of Game Theory*, Vol.13, 201-255.
- [56] Sorin, S. (1985). "Big Match with Lack of Information on One Side (Part 2)," *International Journal of Game Theory*, Vol.14, 173-204.

- [57] Soyster, A.L. (1973). “Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming,” *Operations Research*, Vol.21, pp.1154-57.
- [58] Virtanen, K., Karelahti, J., Raivio, T., and Hamalainen, R. P. (2004). “Modeling Air Combat by a Moving Horizon Influence Diagram Game,” Technical Report, Systems Analysis Laboratory, Helsinki University of Technology, Finland.
- [59] Virtanen, K., T. Raivio, R. P. Hmlinen. (2004b). “Modeling pilots sequential maneuvering decisions by a multistage influence diagram,” to appear in *Journal of Guidance and Control*.
- [60] Vrieze, O. J. (1987). “Stochastic Games with Finite State and Action Spaces,” *CWI Tracts 33*, Amsterdam.
- [61] Weaver, R., Silverman, B. G., Shin, H., and Dubois, R. (2001). “Modeling and Simulating Terrorist Decision-Making: A Performance Moderator Function Approach to Generating Virtual Opponents,” *10th Conference On Computer Generated Forces and Behavioral Representation*, SISO, May, 2001.