

Essays on Game Theory

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Junghyun Oh

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Aldo Rustichini, Advisor

August, 2019

© Junghyun Oh 2019
ALL RIGHTS RESERVED

Acknowledgements

I am deeply indebted to my advisor, Aldo Rustichini, for his continuous guidance, encouragement, support and patience. I am very grateful to David Rahman and Jan Werner for their insightful comments and discussions which have led to the improvement of this thesis.

Dedication

To my parents, Junghee Lim and Moonsik Oh.

Abstract

This dissertation studies mixed strategy equilibria with uncertainty aversion in strategic games. Experimental evidence has found a substantial discrepancy between the prediction of mixed-strategy Nash equilibrium and observed choice frequency. The quantal response equilibrium model of McKelvey and Palfrey (1995) has been successful to explain observed aggregate choice frequencies at least qualitatively. However, its fundamental aspect, inherited from the purification theorem of Harsanyi (1973a), that the mixing is not made intentionally is not supported by experimental studies which confirmed the wide use of the mixing device when it is given to subjects. This paper answers this question by proposing a model which explains players' intentional mixing in strategic games. In the first chapter, we illustrate the main idea of this paper with examples. Then, we compare and contrast the main conclusion of our study with the relevant literature.

In the second chapter, we develop the game theoretic model in which strategic uncertainty is introduced and studies the effect of it to equilibria, especially mixed strategy equilibria, when players are uncertainty averse. The players' perception of uncertainty in opponent players' strategies is modeled with random perturbation. Uncertainty averse players may strictly prefer the properly mixed strategy in order to hedge against this uncertainty. We define an equilibrium under perturbation as a profile of best response strategies to perturbed belief. We show that for any regular mixed-strategy Nash equilibrium in finite games and under the mild assumptions on perturbation structure, there exists a sequence of strict equilibria under perturbations which converges to the given mixed-strategy Nash equilibrium as uncertainty in perturbation vanishes. This result has two implications. First, in almost all finite games, we can find the corresponding equilibrium under perturbation of any given mixed-strategy Nash equilibrium. A discrepancy between them could be substantial if the degree of uncertainty in perturbation is significant. Second, the theory predicts that in equilibrium under perturbation, players intentionally mix their pure strategies due to hedging motive. This implication is clearly distinguished from Harsanyi's purification theorem (and therefore the quantal response equilibrium model's interpretation for mixed equilibrium strategy), and consistent with experimental findings.

In the third chapter, we test our model's predictive power by re-examining the experimental studies; Ochs (1995) and Selten and Chmura (2008). We then compare the performance of our model with the QRE model and show that our model outperforms the QRE model in both re-examinations. In the re-examination of Ochs (1995), our model's predictive power is significantly better than the QRE model in the reasonable range of risk aversion which might not be perfectly controlled in the original experiment of Ochs (1995). Re-examining Selten and Chmura (2008) also shows that the change in equilibrium prediction caused by the payoff transformation is actually observed in the data, which only our model can explain.

Contents

Acknowledgements	i
Dedication	ii
Abstract	iii
List of Figures	vii
1 Introduction	1
1.1 Main Idea of the Dissertation	3
1.1.1 Preference for Randomization	3
1.1.2 Mixed Strategy Equilibria with Uncertainty Aversion in Strategic Games	5
1.2 Related Literature	8
1.2.1 Harsanyi's Purification Theorem and the Quantal Response Equilibrium Model	8
1.2.2 Literature on Uncertainty Averse Preferences	10
1.2.3 Literature on Game Theory with Uncertainty Aversion	11
2 Mixed Strategy Equilibria with Uncertainty Aversion in Strategic Games	14
2.1 Preliminaries	14
2.1.1 Setup	14
2.1.2 Quasi-strictness, Strictness and Regularity of Nash Equilibria	15
2.2 Equilibria under Perturbation	17

2.2.1	Random Perturbation	17
2.2.2	Uncertainty Averse Preferences	19
2.2.3	Equilibria under Perturbation	20
2.2.4	Example Revisited	21
2.3	Nash Equilibria as Limits of Equilibria under Perturbation	24
2.3.1	Parameterized Perturbation	25
2.3.2	Existence and Convergence of Equilibria under Perturbation	27
2.4	Strictness of Equilibria under Perturbation	29
2.4.1	Conditions on Perturbation for Strictness	29
2.4.2	Proof of Strictness	32
3	Testing Mixed-strategy Equilibria with Uncertainty Aversion: a Re- Examination of Experimental Studies	44
3.1	Re-examination of Ochs (1995)	45
3.1.1	Summary of Results in Ochs (1995) and Over-prediction Problem of the QRE model	45
3.1.2	Estimation Results	46
3.2	Re-examination of Selten and Chmura (2008)	51
3.2.1	Experiment Design and Result	51
3.2.2	Payoff Transformation and Equilibrium Change in the Model of Equi- librium with Uncertainty Aversion	52
3.2.3	Estimation Results	55
	References	56
	Appendix A. Appendix to Chapter 2	60

List of Figures

1.1	Generalized matching pennies games in Ochs (1995)	2
1.2	Betting example with unknown urn and two-person symmetric matching pennies game	4
1.3	Best response functions and equilibrium with uncertainty aversion : KMM/MMR case	7
1.4	Game 4	12
1.5	Equilibrium with uncertainty aversion in Game 4 : KMM/MMR case	13
2.1	Best response functions and equilibrium with uncertainty aversion : GS case	24
2.2	Game 5	27
2.3	Construction of $\bar{\epsilon}$	28
2.4	Game 6	35
2.5	Game 7	39
3.1	Nash equilibria and observed frequencies in Ochs (1995)	46
3.2	Choice frequencies in Ochs (1995)	47
3.3	Normalized payoff for estimation in Ochs (1995)	47
3.4	Over-prediction of the QRE model in Ochs (1995)	48
3.5	The best fitting of the QRE model in Ochs (1995)	49
3.6	The best fitting of the equilibrium with uncertainty aversion in Ochs (1995)	50
3.7	The prediction of the QRE model when $r = 0.22$ in Ochs (1995)	51
3.8	Games in Selten and Chmura (2008)	53
3.9	Observations and estimation results in Selten and Chmura (2008)	54

Chapter 1

Introduction

Mixed-strategy Nash equilibrium does not have good predictive power for actual choice behavior in many empirical studies. For instance, in games with unique mixed-strategy Nash equilibrium played recurrently between players (i.e., randomly matched from populations in which one population is for each player role), the aggregate choice frequency of each population converges to particular stationary state. However in many such games, we observed the discrepancy between mixed Nash equilibrium strategy and actual stationary choice frequency.

Mckelvey and Palfrey (1995) proposed the quantal response equilibrium model (henceforth QRE model), based on the seminal work of Harsanyi (1973a), the purification theorem, to explain the observed discrepancy. Harsanyi (1973a) introduces private payoff disturbances in games, which was unmodelled in the original complete information games. Each player's random payoff shock is assumed to be realized and observed only by that player before the player makes her strategy choice. Thus, the realized payoff shock makes the players almost always choose the pure strategies as a strict best response. Stochasticity of choice behavior comes from the randomness of payoff shock, and not the result of deliberate randomization of players. On the other hand, the players only know the probability distribution of the other players' random payoff shocks. As a result, the players only can expect the pure-strategy choice frequency of the other players. They then make their strategy choices based on this belief with their private payoff disturbances. Harsanyi defines the equilibrium in disturbed games as a fixed point of this process. Consequently, the equilibrium is the profile

Figure 1.1: Generalized matching pennies games in Ochs (1995)

		Game 1		Game 2		Game 3	
		<i>L</i>	<i>R</i>	<i>L</i>	<i>R</i>	<i>L</i>	<i>R</i>
<i>U</i>		1	0	4	0	9	0
		0	1	0	1	0	1
<i>D</i>		0	1	0	1	0	1
		1	0	1	0	1	0

of expected frequencies in the minds of other players. This result suggests that the mixed equilibrium strategy does not mean the explicit use of mixed strategy, but the expected frequency of pure-strategy choices in the other players' belief.

The QRE model introduced several acceptable assumptions on payoff disturbance to make the equilibrium model practical. For instance, the logit equilibrium model is the most frequently used specification. With a particular specification, the QRE model suggests that aggregate choice frequency is predicted by the expected frequency in the model. The model seems to be successful because its predictive power of aggregate choice frequency is significantly better than Nash equilibrium. The three games in Figure 1.1 are generalized matching pennies games used in the experimental study of Ochs (1995).¹ The only difference between these three games is player 1's payoff in the upper-left corner. Nash equilibrium predicts that player 1's choice would be the same in all three games since her own payoff change does not affect her equilibrium strategies. But the actual aggregate frequency of player 1 exhibited the "own-payoff effect," that is, it increased significantly with the increase in her own payoff. At the same time, player 2's actual choice frequency deviated from Nash equilibrium prediction. The QRE model predicts this choice pattern well at least qualitatively.

However, it has been found that the QRE model's fundamental aspect which the model inherits from Harsanyi (1973a) is not consistent with the observations made in several

¹Players play games recurrently for around 60 rounds. In the experiment, an explicit mixing device was given to subjects so that they could use the device in each round-game.

experimental studies. As the purification theorem states, the QRE model predicts that when aggregate choice behavior is in a stationary state, players' pure-strategy choices are strict best responses in each repeated single game. Thus even when the explicit mixing device is given to subjects in experiments, they are not supposed to use it to mix their choices. However, for example in Ochs (1995), the extensive use of the mixing device is observed. Among 48 subjects, only 11 subjects were Strict Best Responders, so more than 3/4 of subjects intentionally mix their choices in each round-game.

As a whole, the QRE model proposes a way to resolve the puzzling phenomena that mixed-strategy Nash equilibrium fails to explain. However, its core aspect to deal with stochasticity of choice - choices are almost always (i.e., essentially) pure-strategies in every single game, but in aggregate perspective it is stochastic - is not supported by experimental studies which provide richer observation by giving subjects the explicit mixing device. In those experiments, people intentionally randomize their choices. The purpose of this paper is to provide a new model, which describes such intentional randomization in strategic games. The rest of this chapter is organized as follows. Section 1 illustrates the key idea of this paper with examples. Then, we compare and contrast the main conclusion of our study with the relevant literature in Section 2.

1.1 Main Idea of the Dissertation

1.1.1 Preference for Randomization

Many models of uncertainty aversion adopt a preference for randomization as a central axiom since it was introduced by Schmeidler (1989).² For example, consider the betting example in Figure 1.2. The decision maker faces the urn having two red or black balls with unknown proportions. She bets on the color of ball which will be drawn from the urn. Two bets are denoted by f_R and f_B , respectively. She can also randomize her bets by flipping a fair coin: to choose f_R if it is head, or to choose f_B otherwise, denoted by $\frac{1}{2}f_R + \frac{1}{2}f_B$.³

²We refer to Cerreia-Vioglio et al. (2011) and Battigalli et al. (2017) for the unified understanding of this issue.

³There is an issue for possibly different attitudes towards ex-ante and ex-post randomization. But as Battigalli et al. (2017) argues, commitment makes the distinction between ex-ante and ex-post randomization immaterial. In this paper, we consider the decision and game situations in which committing to

Figure 1.2: Betting example with unknown urn and two-person symmetric matching pennies game

	s_R	s_B
f_R	1	0
f_B	0	1

	a_2^1	a_2^2
a_1^1	1	0
	0	1
a_1^2	0	1
	1	0

In this decision problem, as Cerreia-Vioglio et al. (2011) interprets, the color of drawn ball is a state $s \in S = \{s_R, s_B\}$ and the proportion of colors is a probabilistic *model* $p \in \Sigma \subseteq \Delta(S)$ that govern states' realizations. In this problem, there are three possible models, $\Sigma = \{(1, 0), (\frac{1}{2}, \frac{1}{2}), (0, 1)\}$. Since the proportion is unknown, the decision maker perceives *model* uncertainty. This uncertainty can be expressed by a prior μ over such models. Here it is natural to consider a uniform μ on Σ . If she has neutral attitude toward model uncertainty, she will be indifferent between three possible choices. This is because she considers this unknown urn identical to the urn having one red ball and one black ball. Thus, there is no effect of model uncertainty for uncertainty neutral decision maker. On the other hand, if she is averse to model uncertainty, she will strictly prefer $\frac{1}{2}f_R + \frac{1}{2}f_B$ to f_R and f_B . For instance, suppose that she is extremely averse to uncertainty, so only consider the worst model for each choices. Then, the randomized choice guarantees an expected value of 1/2 regardless of which model is, but both of pure choices only can guarantee 0. Thus the preference for randomization emerges under uncertainty from uncertainty aversion as a hedging motive.

This paper applies this idea to game theoretic setup in order to introduce strategic uncertainty into games. Consider player i in games. In her perspective, the profile of pure strategies of the other players corresponds to the state; the profile of mixed strategies of

randomization (or to a mixed strategy in games) is available. In simultaneous-moves games, the limited time frame to decide the final pure-strategy choice can be thought as an implicit commitment device. We refer to Dominiak and Schnedler (2011), Eichberger et al. (2016), Saito (2015) and Seo (2009) for more detailed studies on this issue.

the other players corresponds to model, which govern realizations of pure strategies. Now, suppose that the player i perceives the model uncertainty, i.e., the uncertainty over the set of the profile of mixed strategies of the other players. The player i thus forms a set of models (i.e. a set of the other players' mixed-strategy profiles) with a prior. The player i 's pure strategy corresponds to pure act, and her mixed strategy corresponds to randomization on pure acts. To fix ideas, consider a two-person game in Figure 1.2. and the unique mixed-strategy Nash equilibrium, in which both players choose the equiprobable mixed-strategy, $(\frac{1}{2}, \frac{1}{2})$. Suppose that the player 1 is concerned that the player 2 may make a slight mistake or *perturbation* from the equilibrium strategy in two ways equally; the player 2 may play $(\frac{1}{2} + \epsilon, \frac{1}{2} - \epsilon)$ or $(\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon)$ with a small ϵ , equally. Thus the player 1 forms the set of two possible models with uniform prior. If ϵ vanishes, then there is no model uncertainty. Thus ϵ captures the degree of uncertainty. Now suppose that the player 1 is averse to uncertainty. Then the equiprobable randomization on two pure strategies (i.e. the mixed strategy $(\frac{1}{2}, \frac{1}{2})$) is strictly preferred to any other randomizations (i.e. mixed-strategies). This is because the equiprobable randomization perfectly hedge against the uncertainty they perceives in the opponent's strategy choice, which is exactly the same reason in the previous betting example.

1.1.2 Mixed Strategy Equilibria with Uncertainty Aversion in Strategic Games

This paper develops the game theoretic model in which strategic uncertainty is introduced and studies the effect of it to equilibria, especially mixed strategy equilibria, when players are uncertainty averse. The player's perception of uncertainty in opponent players' strategies is modeled with *random perturbation* as we have seen in the example of the previous subsection. If she is uncertainty averse, she will be willing to hedge against the uncertainty, and therefore may strictly prefer particular mixed-strategies.

To sketch the main results of this paper, we consider Game 2 in Figure 1.1. In this game, there is a unique mixed-strategy Nash equilibrium; $s^* = ((s_U^*, s_D^*), (s_L^*, s_R^*)) = ((\frac{1}{2}, \frac{1}{2}), (\frac{1}{5}, \frac{4}{5}))$. Suppose player 1 assumes that player 2 would choose the Nash equilibrium strategy, but at the same time, she has some slight doubt if player 2's strategy choice would be perturbed randomly. If she has a neutral attitude toward this randomness, such random

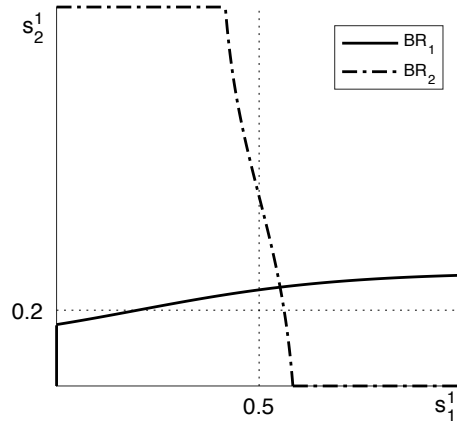
perturbation would not affect her choice, and she would be indifferent to any mixed strategy. But if she is averse to uncertainty, her best response will be $(\frac{1}{5}, \frac{4}{5})$, which is her maximin strategy. She chooses this because the maximin strategy guarantees the expected payoff of $\frac{4}{5}$ regardless of which perturbation is realized, and thus she can perfectly hedge against the uncertainty she perceives in player 2's strategy choice. The presence of uncertainty in belief and the aversion to it forces her to choose the secure strategy.

If player 1's best guess (s_L, s_R) for opponent choice deviates from the Nash equilibrium strategy, her best response reacts to this change. For instance, if $s_L > \frac{1}{5}$, then player 1 is willing to choose a higher value of s_U than $\frac{1}{5}$. Two forces are combined to determine the best response: hedging against the uncertainty versus reaction to the change in s_L . The best response function of player 1 is formed in this way and it is depicted in Figure 1.3. For player 2, the maximin strategy is $(\frac{1}{2}, \frac{1}{2})$ and she is willing to increase s_L when s_U decreases. Thus her best response function is depicted in Figure 1.3.

Now, look at the intersection of the two best response functions. At this intersection, the strategy profile \hat{s} satisfies the consistency condition. This consistency is a relaxed version of the Nash equilibrium consistency condition because each player's choice is the best response to the belief which is formed around the opponent choice with random perturbation, rather than to the singleton of the opponent choice. Thus we call this strategy profile \hat{s} the equilibrium under perturbation.

In \hat{s} , each player's equilibrium mixed strategy is the strict best response. Each player intentionally randomizes between their pure strategies according to equilibrium mixed strategy in order to hedge against the uncertainty she perceives in the other player's strategy choice. Thus the equilibrium is the profile of such intentional randomization. As a result, our model suggests that each player would choose the equilibrium mixed-strategy in each repeated single game (e.g., each round-game in Ochs (1995)) when choice behavior is in a stationary state. In addition, the discrepancy between the mixed Nash equilibrium strategy and the observed aggregate choice frequency can be explained as well. In Figure 1.3, \hat{s} deviates from s^* . Due to hedging motive, equilibrium under perturbation would deviate from the Nash equilibrium. This deviation shrinks as the degree of uncertainty decreases. This is because as the strategic uncertainty decreases the hedging motive reduces, so each player's best response approaches to the case without uncertainty. As a result, \hat{s} converges

Figure 1.3: Best response functions and equilibrium with uncertainty aversion : KMM/MMR case



to s^* as the uncertainty vanishes.

In this paper, we apply this idea to any finite games. We show that for any regular⁴ mixed-strategy Nash equilibrium in finite games and under the mild assumptions on perturbation structure, there exists a sequence of strict equilibria under perturbations which converges to the given mixed-strategy Nash equilibrium as uncertainty in perturbation vanishes. This result has two implications. First, in almost all games, we can find the corresponding equilibrium under perturbation of any given mixed-strategy Nash equilibrium. The discrepancy between them could be substantial if the degree of uncertainty in perturbation is significant. The theory provides the prediction of such discrepancy by formula. Second, the theory predicts that in equilibrium under perturbation, players intentionally randomize between their pure strategies according to the equilibrium strategies. This is the point that our model is clearly distinguished from Harsanyi's purification theorem, and therefore the QRE model's prediction for individual choice behavior.

⁴We discuss the regularity of Nash equilibria in Chapter 2.1.2 following Van Damme (1991).

1.2 Related Literature

1.2.1 Harsanyi's Purification Theorem and the Quantal Response Equilibrium Model

As Morris (2006) states, “Harsanyi’s purification theorem provides the leading interpretation of mixed strategy equilibria among game theorists today.”⁵ Indeed, Harsanyi’s work (1973a) has been motivated to resolve the instability problem of mixed-strategy Nash equilibria. Harsanyi (1973a) claims:

“Equilibrium points in mixed strategies are unstable because any player can deviate without penalty from his equilibrium strategy even if all other players stick to theirs. ... This instability seems to pose a serious problem because many games have only mixed-strategy equilibrium points.” [p.1]

Any properly-mixed Nash equilibrium strategies are weakly preferred. This is the essence of the instability problem that Harsanyi pointed out. To resolve this criticism, he introduces the random payoff disturbances, and assumes that the realized payoff is private information and the other players only know the distribution of it. Thus, each player will choose pure strategies with probability 1 according to their realized payoff, but the players only can know expected frequencies about how the other players choose their own pure strategies. To see this idea more clearly, consider any finite games. Let $U_i(a_i^k, \bar{s}_i)$ denote the expected utility of player i of choosing the k -th pure strategy a_i^k where \bar{s}_i is the profile of the other players’ mixed strategy choices. Suppose now that the random payoff disturbance ϵ_i^k associated with the k -th pure strategy a_i^k is introduced in a following way,

$$\tilde{U}_i(a_i^k, \bar{s}_i) = U_i(a_i^k, \bar{s}_i) + \epsilon_i^k$$

where $\tilde{U}_i(a_i^k, \bar{s}_i)$ is a disturbed expected utility of choosing a_i^k . Player i ’s payoff disturbance vector $\epsilon_i = (\epsilon_i^1, \dots, \epsilon_i^k, \dots, \epsilon_i^{K_i})$ is assumed to be continuously distributed according to a joint probability density function $f_i(\epsilon_i)$. Given \bar{s}_i , the pure strategy a_i^k will be a strict best response if $\tilde{U}_i(a_i^k, \bar{s}_i) > \tilde{U}_i(a_i^{k'}, \bar{s}_i)$ for all $k \neq k'$. Then, due to the assumption that ϵ_i is continuously distributed, the player i will choose pure strategies as a strict best response

⁵We refer to Morris (2008) for more discussion of the purification theorem.

with probability 1. On the other hand, since the probability distribution of ϵ_i is known to the other players, they know the probability π_i^k that the player i chooses a_i^k ,

$$\pi_i^k = \int_{\{\epsilon_i: \tilde{U}_i(a_i^k, \bar{s}_i) > \tilde{U}_i(a_i^{k'}, \bar{s}_i) \text{ for all } k \neq k'\}} f_i(\epsilon_i) d\epsilon_i$$

where π_i^k is the function of \bar{s}_i .

The equilibrium of disturbed game is defined as any strategy profile $s = (s_1, \dots, s_N)$ such that for all player $i = 1, \dots, N$ and $k = 1, \dots, K_i$, $s_i^k = \pi_i^k(\bar{s}_i)$. The purification theorem shows that any regular mixed-strategy Nash equilibria are the limit of equilibria of disturbed games. In disturbed games, the mixed equilibrium strategy is the probability distribution of pure-strategy choices in the other players' belief. It does not mean the explicit use of mixed strategy. At mixed-strategy equilibria in disturbed games, players almost always choose the pure strategies as a strict best response. Therefore, mixed-strategy equilibria in disturbed games are stable.

Based on Harsanyi's work, the QRE model has been proposed to explain the actual choice behavior in many experimental studies which the mixed-strategy Nash equilibria fail to predict. Several assumptions on payoff disturbance ϵ_i are introduced to make the model practical. The most popular logit equilibrium model adopts the assumptions that ϵ_i^k and $\epsilon_i^{k'}$ are i.i.d. for all i and all k with cumulative distribution function $F_i(\epsilon_i^k) = \exp(-\exp(-\lambda\epsilon_i^k - \gamma))$ with $\lambda \geq 0$. Under this specification, the following quantal response function is obtained.

$$\pi_i^k(\bar{s}_i) = \frac{\exp(\lambda U_i(a_i^k, \bar{s}_i))}{\sum_{k=1}^{K_i} \exp(\lambda U_i(a_i^k, \bar{s}_i))}$$

where λ^{-1} is the size of payoff disturbance.

Harsanyi's model and the QRE model are closely related to the seminal works of Luce (1959) and McFadden (1974)'s random utility model. The stochastic nature of choice comes from randomness of payoff disturbance. Thus, the stochastic choice occurs involuntarily, and is not the result of intentional randomization. In contrast, in our model the stochasticity of choice is the result of intentional randomization to hedge the strategic uncertainty as we have seen in Section 1.1.

1.2.2 Literature on Uncertainty Averse Preferences

In this paper, we employ two representations of uncertainty averse preferences.⁶ The first one is the Smooth ambiguity preference of Klibanoff, Marinacci, and Mukerji (2005) (henceforth, KMM). The value of f by KMM is written as

$$V^{KMM}(f) = \int_{\Delta} \phi \left(\int_S u(f(s)) dp(s) \right) d\mu(p).$$

where ϕ is continuous, strictly increasing and strictly concave. The second one is (the second-order) Divergence preference, which is a special case of Variational preference studied by Maccheroni, Marinacci, and Rustichini (2006) (henceforth, MMR). The value of f by MMR is represented by⁷

$$V^{MMR}(f) = \min_{\nu \in \Delta^\sigma(\mathcal{B}(\Delta), \mu)} \left\{ \int_{\Delta} \left(\int_S u(f(s)) dq(s) \right) d\nu(q) + c(\nu) \right\}$$

where $c(\nu) = \theta D_\psi^\omega(\nu \parallel \mu)$ with $\theta > 0$.⁸

Both representations have three common characteristics that are critical in our study. First, they are able to represent aversion to model uncertainty. In our game theoretic framework, the set of pure strategies of the other players corresponds to the set of states; the profile of mixed strategies of the other players corresponds to the probabilistic *model*, which govern states' realizations. Each player forms a set of models with a prior μ , which reflects her perception about the possibility of multiple perturbations in other players' strategy choices. Then, the player's pure strategy corresponds to a pure act, and her mixed strategy corresponds to randomization on pure acts. Second, they are in general smooth. The multiple priors preferences of Gilboa and Schmeidler (1989) (henceforth, GS) is the limit case of both models when the decision maker is extremely averse to model uncertainty.⁹ The non-smoothness of GS model leads the different implication on the

⁶Battigalli et al. (2015) also use these two models in the analysis of recurrent games.

⁷ $\Delta^\sigma(\mathcal{B}(\Delta), \mu)$ is the set of probability measures ν on Δ that are absolutely continuous with respect to μ

⁸ $D_\psi^\omega(\nu \parallel \mu)$ is a ω -weighted ψ divergence of ν with respect μ , which is well-explained in Maccheroni et al. (2006).

⁹For the proof, see Proposition 3 in Klibanoff et al. (2005), and Proposition 22 in Maccheroni et al. (2006).

effect of uncertainty aversion to mixed-strategy equilibria. We will explain this point in the next subsection and in Subsection 2.2.4 as well. Third, both representations adopt a preference for randomization as a central axiom. Thus the players who are represented by these models may choose the randomizations as strict best responses.

1.2.3 Literature on Game Theory with Uncertainty Aversion

This paper is also closely related to a literature on games with uncertainty aversion.¹⁰ The most closely related paper is Klibanoff (1996). His model proposes the following concept of equilibrium with uncertainty aversion. Fix a finite normal form game with a finite set of players $\{1, 2, \dots, N\}$. Let $a_i \in A_i$ denote the pure strategy of player i and $s_i \in S_i$ denote the mixed strategy of player i .

Definition : An *equilibrium with uncertainty aversion* is a $2*N$ -vector $(s_1, \dots, s_N, B_1, \dots, B_N)$ where the B_i are closed, convex subsets of P_{-i} (the set of probability distributions over $\prod_{j \neq i} A_j$) such that, for all i ,

- (1) s_i satisfies $\min_{p \in B_i} \sum_a U_i(a_i, a_{-i}) s_i(a_i) p(a_{-i}) \geq \min_{p \in B_i} \sum_a U_i(a_i, a_{-i}) s'_i(a_i) p(a_{-i})$ for all $s'_i \in S_i$ and
- (2) $\prod_{j \neq i} s_j(a_j) \in B_i$

The condition (2) is the consistency condition of equilibrium saying that the players' beliefs need to contain the truth. The condition (1) means that each player's uncertainty averse preference is represented by the GS model.

Indeed, Klibanoff (1996) claims that "equilibrium with uncertainty aversion can justify mixing as a (strict best) response to strategic uncertainty", which is the direct result of the use of GS model that adopts the axiom of preference for randomization. However, due to the non-smoothness of GS model, Klibanoff (1996) and our model provide the different implication on how the strict mixed strategy equilibria are determined. We will show this point with the following two example games. We first consider Game 2 in Figure 1.1. Under the equilibrium concept in Klibanoff (1996), $(s_U, s_L) = (\frac{1}{5}, \frac{1}{2})$ is the only strict mixed equilibrium strategy profile, as long as the player 1 has some belief

¹⁰See Dow and Werlang (1994), Eichberger and Kelsey (2000, 2002, 2014), Klibanoff (1996), Lo (1996, 2009), and Marinacci (1996).

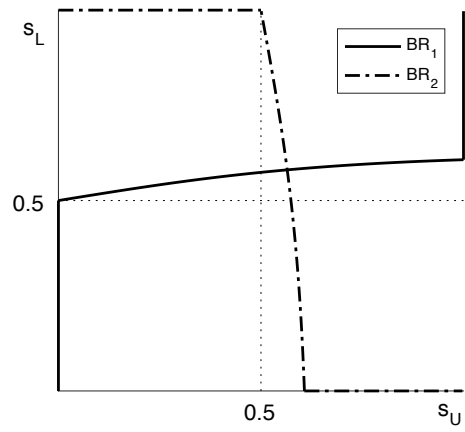
Figure 1.4: Game 4

	L	R
U	2 1	0 2
D	1 1	1 0

assigning $Prob(L) < \frac{1}{5}$ and some belief assigning $Prob(L) > \frac{1}{5}$, and at the same time, the player 2 has some belief assigning $Prob(U) < \frac{1}{2}$ and some belief assigning $Prob(U) > \frac{1}{2}$, respectively. Compared to the analysis with our model in Subsection 1.1.2, there are mainly two prominent differences. First, the predictions of equilibrium are significantly different: our model predicts that $s_U > \frac{1}{2}$ and $s_L \in (\frac{1}{5}, \frac{1}{2})$, whereas Klibanoff (1996)'s prediction is $(s_U, s_L) = (\frac{1}{5}, \frac{1}{2})$. Second, Klibanoff (1996)'s prediction of equilibrium strategy profile is constant no matter how the beliefs are formed as long as the stated conditions on each players' beliefs are satisfied. Thus, there is no such result stating the convergence of equilibrium with uncertainty aversion to Nash equilibrium as uncertainty vanishes, which is the natural conjecture that the equilibrium with uncertainty aversion seems to have to satisfy.

Next, we consider Game 4 in Figure 1.4, in which the unique Nash equilibrium is $(s_U, s_L) = (\frac{1}{2}, \frac{1}{2})$. Our model predicts that there is a strict mixed strategy equilibrium, which is shown in Figure 1.5. As we have seen in the analysis in Subsection 1.1.2, the equilibrium converges to the mixed-strategy Nash equilibrium as the degree of strategic uncertainty vanishes. In contrast, there is no strict mixed equilibrium strategy profile under the equilibrium concept in Klibanoff (1996). Instead, there are two pure strategy equilibrium: The first one is (D, L) with the condition that the player 1 has some belief assigning $Prob(L) < \frac{1}{2}$. The second one is (U, L) with the condition that the player 2 has some belief assigning $Prob(U) < \frac{1}{2}$. This example highlights the difference in equilibrium prediction of two models.

Figure 1.5: Equilibrium with uncertainty aversion in Game 4 : KMM/MMR case



Chapter 2

Mixed Strategy Equilibria with Uncertainty Aversion in Strategic Games

2.1 Preliminaries

2.1.1 Setup

Let Γ be a finite N -person normal form game,

$$\Gamma = (A_1, \dots, A_N; U_1, \dots, U_N)$$

where the nonempty finite set A_i with K_i elements is the set of pure-strategies of player i and the real-valued function U_i defined on $A = \prod_{i=1}^N A_i$ is the utility function of player i . We denote by a_i^k the k th pure strategy of player i . A mixed-strategy s_i of player i is a vector of probabilities assigned to the pure strategies,

$$s_i = (s_i^1, \dots, s_i^{K_i})$$

where $s_i \in \Delta^{K_i-1}$.¹ The set of all mixed-strategies s_i of player i will be denoted by S_i . The set of all possible mixed-strategy profiles $s = (s_1, \dots, s_N)$ will be denoted by S .

¹We write $\Delta^n = \{(t_0, \dots, t_n) \in \mathbb{R}^{n+1} \mid \sum_{k=0}^n t_k = 1 \text{ and } t_k \geq 0 \text{ for all } k\}$.

$\bar{s}_i = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_N)$ is the mixed-strategy profile which represents the strategies of all $(N - 1)$ players other than player i . We will write $s = (s_i, \bar{s}_i)$. The functions U_i are extended to S by

$$U_i(s_1, \dots, s_N) = \sum_{(a_1, \dots, a_N) \in A} U_i(a) \left(\prod_{i=1}^N s_i(a_i) \right)$$

where $s_i(a_i) = s_i^k$ for $a_i = a_i^k$.

We denote by $C(s_i)$ the carrier of s_i

$$C(s_i) := \left\{ a_i^k \in A_i : s_i^k > 0 \right\}.$$

s_i will be called properly mixed if $|C(s_i)| > 1$. The carrier of s is defined by

$$C(s) := \prod_{i=1}^N C(s_i).$$

We define the extended carrier of s_i by $C^e(s_i) := \Delta(C(s_i))$. Thus, $C^e(s_i)$ is the set of all mixed strategies that distribute all probability over the pure strategies in $C(s_i)$. The extended carrier of s is defined by $C^e(s) := \prod_{i=1}^N C^e(s_i)$.

The set of player i 's best responses to \bar{s}_i is denoted by $B_i(\bar{s}_i)$,

$$B_i(\bar{s}_i) := \operatorname{argmax}_{t_i \in S_i} U_i(t_i, \bar{s}_i).$$

We say that $t \in S$ is a best response to s if t_i is a best response to \bar{s}_i for all i . We denote the set of best responses to s by $B(s)$, hence $B(s) := \prod_{i=1}^N B_i(\bar{s}_i)$. The following is the definition of Nash equilibrium s^* ²:

Definition 1: A strategy profile $s^* = (s_1^*, \dots, s_N^*)$ is a Nash equilibrium of Γ if $s^* \in B(s^*)$.

2.1.2 Quasi-strictness, Strictness and Regularity of Nash Equilibria

A Nash equilibrium s^* is called *quasi-strict*³ if $B(s^*) \subseteq C^e(s^*)$ and called *strict* if $B(s^*) = \{s^*\}$. Hence, a strict Nash equilibrium is a quasi-strict Nash equilibrium in pure strategies.

²In the rest of the paper, we denote Nash equilibria by s^* .

³We follow the terminology of Van Damme (1991). See Chapter 2.1. of Van Damme.

Let s^* be a quasi-strict Nash equilibrium of Γ . Suppose that each carrier $C(s_i^*)$ contains γ_i pure strategies, respectively. For notational convenience, we use the following notational convention without loss of generality: We number the pure strategies a_i^k of player i in such a way that the first γ_i pure strategies $a_i^1, \dots, a_i^{\gamma_i}$ are contained in $C(s_i^*)$. Thus we will write

$$C(s_i^*) = \{a_i^1, \dots, a_i^{\gamma_i}\} \text{ for each } i.$$

We can divide the players according to whether s_i^* is properly mixed or not. Let I denote the set of players the strategy of whom is properly mixed, i.e.

$$I = \{i \in \{1, \dots, N\} : \gamma_i > 1\}$$

Consider s with $C(s) = C(s^*)$. For $i \in I$, we denote by $\sigma_i \in \Delta_c^{\gamma_i-1}$ the vector⁴

$$\sigma_i = \left(s_i^1, \dots, s_i^{\gamma_i-1} \right)$$

and let σ denote the vector

$$\sigma = (\sigma_1; \dots; \sigma_i; \dots; \sigma_N), i \in I.$$

Clearly, $\sigma \in \prod_{i \in I} \Delta_c^{\gamma_i-1}$ fully characterizes s with $C(s) = C(s^*)$. We denote σ characterizing the Nash equilibrium s^* by σ^* . We define $\gamma := \sum_{i \in I} (\gamma_i - 1)$ for future notational convenience.

We define $\varphi_i^k : \prod_{j \neq i} S_j \rightarrow \mathbb{R}$ by

$$\varphi_i^k(\bar{s}_i) = U_i(a_i^k, \bar{s}_i) - U_i(a_i^{\gamma_i}, \bar{s}_i) \text{ for } i \in I; k = 1, \dots, \gamma_i - 1.$$

Let φ denote the vector

$$\varphi = \left(\varphi_1^1, \dots, \varphi_1^{\gamma_1-1}; \dots; \varphi_i^1, \dots, \varphi_i^{\gamma_i-1}; \dots; \varphi_N^1, \dots, \varphi_N^{\gamma_N-1} \right), i \in I.$$

φ is a function from S to \mathbb{R}^γ . When the argument of φ is s with $C(s) = C(s^*)$, we will write $\varphi(\sigma)$ with a slight abuse of notation to make the argument of φ more explicit. Let $J_\varphi(\sigma)$ be the Jacobian

⁴We write $\Delta_c^n = \{(t_1, \dots, t_n) \in \mathbb{R}^n \mid \sum_{k=1}^n t_k \leq 1 \text{ and } t_k \geq 0 \text{ for all } k\}$.

$$J_\varphi(\sigma) = \frac{\partial \varphi}{\partial \sigma} = \frac{\partial \left(\varphi_1^1, \dots, \varphi_1^{\gamma_1-1}; \dots; \varphi_i^1, \dots, \varphi_i^{\gamma_i-1}; \dots; \varphi_N^1, \dots, \varphi_N^{\gamma_N-1} \right)}{\partial \left(s_1^1, \dots, s_1^{\gamma_1-1}; \dots; s_i^1, \dots, s_i^{\gamma_i-1}; \dots; s_N^1, \dots, s_N^{\gamma_N-1} \right)}.$$

We denote by J^* the Jacobian $J_\varphi(\sigma)$ evaluated at σ^* . Therefore, we can classify quasi-strict Nash equilibria which are not strict according to whether they have non-singular Jacobian J^* or not. A Nash equilibrium is called *regular* if it is strict or quasi-strict with non-singular Jacobian J^* . Harsanyi (1973b) showed the following theorem.

Theorem (Harsanyi, 1973b): In almost all finite games, all Nash equilibria are regular.

2.2 Equilibria under Perturbation

2.2.1 Random Perturbation

Let $s = (s_1, \dots, s_N)$ be a profile of mixed-strategies. We introduce a perturbation $v_i = (v_i^1, \dots, v_i^{K_i})$ in player i 's strategy s_i . Let p_i denote the perturbed strategy of s_i by v_i ,

$$p_i = s_i + v_i.$$

Of course, v_i should satisfy the condition that $p_i \in \Delta^{K_i-1}$. We interpret p_i in a following way: When the player i is expected to choose s_i , the other players are concerned that the player i makes a small mistake with the perturbation v_i and therefore plays p_i .

Next we introduce randomness in perturbation. We model the situation where the other players are concerned about player i 's perturbation that the multiple cases are possible with probability weights w_i . Random perturbation of player i 's strategy can be written as⁵

$$\Upsilon_i = (v_i[1], \dots, v_i[M_i]; w_i[1], \dots, w_i[M_i])$$

where $M_i \geq 2$, $v_i[m_i] \neq v_i[m'_i]$ for all $m_i \neq m'_i$, and each $v_i[m_i]$ has probability weight $w_i[m_i]$ with $w_i[m_i] > 0$ for $m_i = 1, \dots, M_i$ and $\sum_{m_i=1}^{M_i} w_i[m_i] = 1$. Since we mainly consider perturbations with randomness, when we speak of a perturbation without further

⁵To avoid confusion, we will use $[\cdot]$ instead of (\cdot) for the numbering by m_i .

specification we will always mean a random perturbation. The perturbed strategy of s_i by Υ_i will be written as

$$P_i(s_i, \Upsilon_i) = (p_i[1], \dots, p_i[M_i]; w_i[1], \dots, w_i[M_i])$$

where $p_i[m_i] = s_i + v_i[m_i]$. We denote by Υ the vector of all players' perturbations,

$$\Upsilon = (\Upsilon_1, \dots, \Upsilon_N).$$

$\overline{\Upsilon}_i = (\Upsilon_1, \dots, \Upsilon_{i-1}, \Upsilon_{i+1}, \dots, \Upsilon_N)$ represents the perturbations of all $(N - 1)$ players other than player i .

We will assume that perturbations associated with different players are statistically independent. Due to statistical independence, player i will face $R_i = \prod_{j \neq i} M_j$ cases with probability weight $\prod_{j \neq i} w_j[m_j]$ under Υ . For notational convenience, those cases are numbered with $r_i = 1, \dots, R_i$ with probability weight $W_i[r_i] > 0$. By an abuse of notation, we will write

$$v_j[r_i] = v_j[m_j] \text{ and } p_j[r_i] = p_j[m_j]$$

if $v_j[m_j]$ is involved in the r_i th case. Player i 's belief about the other player's perturbed strategies is represented by

$$\overline{P_i(s_i, \Upsilon_i)} = (P_1(s_1, \Upsilon_1), \dots, P_{i-1}(s_{i-1}, \Upsilon_{i-1}), P_{i+1}(s_{i+1}, \Upsilon_{i+1}), \dots, P_N(s_N, \Upsilon_N)).$$

$\overline{P_i(s_i, \Upsilon_i)}$ will be reformulated in player i 's mind in the form

$$(\overline{p}_i[1], \dots, \overline{p}_i[R_i]; W_i[1], \dots, W_i[R_i])$$

where

$$\overline{p}_i[r_i] = (p_1[r_i], \dots, p_{i-1}[r_i], p_{i+1}[r_i], \dots, p_N[r_i])$$

is the profile of the other players' perturbed strategies at the r_i th case.

2.2.2 Uncertainty Averse Preferences

We introduce uncertainty aversion in players' preference. Two models of uncertainty aversion will be used: KMM and MRR. Firstly by KMM, player i 's value function of choosing $t_i \in S_i$ under $\overline{P_i(s_i, \Upsilon_i)}$ will be written as

$$\Phi_i^{KMM}(t_i, \overline{P_i(s_i, \Upsilon_i)}) = \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \phi_i(U_i(t_i, \overline{p_i}[r_i]))$$

where each ϕ_i is a strictly increasing and strictly concave C^2 function.⁶ Next, the representation by MRR is introduced. For notational simplicity, we denote by W_i the vector of probability weights $(W_i[1], \dots, W_i[R_i]) \in \Delta^{R_i-1}$, and by Q_i any arbitrary vector of probability weights $(Q_i[1], \dots, Q_i[R_i]) \in \Delta^{R_i-1}$. Then, player i 's value function will be written as

$$\Phi_i^{MMR}(t_i, \overline{P_i(s_i, \Upsilon_i)}) = \min_{Q_i \in \Delta^{R_i-1}} \left(\sum_{r_i=1}^{R_i} Q_i[r_i] \cdot U_i(t_i, \overline{p_i}[r_i]) + c_i(Q_i) \right)$$

where

$$\begin{aligned} c_i(Q_i) &= \theta_i D_{\mu_i}^{\omega_i}(Q_i \parallel W_i) \\ &= \theta_i \sum_{r_i=1}^{R_i} \omega_i[r_i] \mu_i \left(\frac{Q_i[r_i]}{W_i[r_i]} \right) W_i[r_i] \end{aligned}$$

with a given ω_i such that $\omega_i[r_i] > 0$ for all r_i and $\sum_{r_i} \omega_i[r_i] W_i[r_i] = 1$ and a strictly convex C^2 function $\mu_i : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ such that $\mu_i(1) = 0$ and $\lim_{t \rightarrow \infty} \mu_i(t)/t = \infty$.⁷ Thus player i 's uncertainty aversion attitude is characterized by ϕ_i in KMM and c_i in MMR. In the rest of paper, when we write the value function Φ_i without superscript, we mean both Φ_i^{KMM} and Φ_i^{MMR} . The following lemma summarizes the characteristic of value functions under uncertainty aversion.

Lemma 1: For a given $\overline{P_i(s_i, \Upsilon_i)}$, the mapping $s_i \rightarrow \Phi_i(s_i, \overline{P_i(s_i, \Upsilon_i)})$ is concave.

⁶We add the technical assumption that ϕ_i is a C^2 function.

⁷ $D_{\psi}^{\omega}(\nu \parallel \mu)$ is a ω -weighted ψ divergence of ν with respect μ . In this paper, we add the technical assumption that μ_i is a C^2 function. We refer to Cerreia-Vioglio et al. (2013) for this specific class of variational preferences.

The next lemmas will be useful in the rest of the paper. Lemma 2 shows that if a pure strategy $a_i^{k'}$ is dominated by the other pure strategy a_i^k at all cases r_i , then the value of choosing a mixed strategy s_i will be increased by moving a probability weight of $a_i^{k'}$ to a_i^k .

Lemma 2: Given $\overline{P_i(s_i, \Upsilon_i)}$, suppose that for $a_i^k, a_i^{k'} \in A_i$,

$$U_i(a_i^k, \overline{p_i[r_i]}) > U_i(a_i^{k'}, \overline{p_i[r_i]}) \text{ for all } r_i.$$

Then for s_i and τ_i such that $\tau_i^l = s_i^l$ for all $l \neq k, k'$ and $\tau_i^k > s_i^k$,

$$\Phi_i(\tau_i, \overline{P_i(s_i, \Upsilon_i)}) > \Phi_i(s_i, \overline{P_i(s_i, \Upsilon_i)}).$$

The following lemma is an immediate result of Lemma 2.

Lemma 3: Suppose that $\overline{P_i(s_i, \Upsilon_i)}$ is given.

(i) If for $a_i^k, a_i^{k'} \in A_i$,

$$U_i(a_i^k, \overline{p_i[r_i]}) > U_i(a_i^{k'}, \overline{p_i[r_i]}) \text{ for all } r_i, \tag{2.1}$$

then $s_i \in \operatorname{argmax}_{t_i \in S_i} \Phi_i(t_i, \overline{P_i(s_i, \Upsilon_i)})$ only if $a_i^{k'} \notin C(s_i)$.

(ii) If for $a_i^k \in A_i$, (2.1) holds for all $a_i^{k'} \neq a_i^k$, then $\{a_i^k\} = \operatorname{argmax}_{t_i \in S_i} \Phi_i(t_i, \overline{P_i(s_i)})$.

2.2.3 Equilibria under Perturbation

In the rest of the paper, we assume that all players' uncertainty aversion attitudes are common knowledge. We define the game Γ^{UA} with uncertainty aversion as $\Gamma^{UA} = (A_1, \dots, A_N; U_1, \dots, U_N; \phi_1, \dots, \phi_N)$ (KMM) and as $\Gamma^{UA} = (A_1, \dots, A_N; U_1, \dots, U_N; c_1, \dots, c_N)$ (MMR). We define the set of player i 's best responses to $\overline{s_i}$ under $\overline{\Upsilon_i}$ by

$$B_i^{UA}(\overline{s_i}, \overline{\Upsilon_i}) = \operatorname{argmax}_{t_i \in S_i} \Phi_i(t_i, \overline{P_i(s_i, \Upsilon_i)}).$$

The set of best responses to s under Υ is denoted by $B^{UA}(s, \Upsilon)$,

$$B^{UA}(s, \Upsilon) := \prod_{i=1}^N B_i^{UA}(\overline{s_i}, \overline{\Upsilon_i}).$$

We define an equilibrium under perturbation Υ as follows.

Definition 2: A strategy profile $s = (s_1, \dots, s_N)$ is an equilibrium under perturbation Υ if $s \in B^{UA}(s, \Upsilon)$.

We can naturally generalize the concepts of strictness and quasi-strictness to our equilibrium concept. A equilibrium s under perturbation Υ is called *quasi-strict* if $B^{UA}(s, \Upsilon) \subseteq C^e(s)$ and called *strict* if $B^{UA}(s, \Upsilon) = \{s\}$ or equivalently, for all $i = 1, \dots, N$,

$$\{s_i\} = \operatorname{argmax}_{t_i \in S_i} \Phi_i(t_i, \overline{P_i(s_i, \Upsilon_i)}). \quad (2.2)$$

If perturbation Υ vanishes, that is,

$$v_i[m_i] = 0 \text{ for all } i; \text{ all } m_i \quad (2.3)$$

then there is no effect of uncertainty aversion. The reason is that the best response is identical: If (2.3) holds, then $p_i[m_i] = s_i$ for all i ; all m_i . In KMM, since each ϕ_i is strictly increasing,

$$\operatorname{argmax}_{t_i} \Phi_i^{KMM}(t_i, \overline{P_i(s_i, \Upsilon_i)}) = \operatorname{argmax}_{t_i} \phi_i(U_i(t_i, \overline{s_i})) = \operatorname{argmax}_{t_i} U_i(t_i, \overline{s_i}).$$

In MMR,

$$\begin{aligned} \operatorname{argmax}_{t_i} \Phi_i^{MMR}(t_i, \overline{P_i(s_i, \Upsilon_i)}) &= \operatorname{argmax}_{t_i} \left[\min_{Q_i \in \Delta^{R_i-1}} (U_i(t_i, \overline{s_i}) + c_i(Q_i)) \right] \\ &= \operatorname{argmax}_{t_i} U_i(t_i, \overline{s_i}) + c_i(W_i) = \operatorname{argmax}_{t_i} U_i(t_i, \overline{s_i}) \end{aligned}$$

Thus any Nash equilibrium s^* of Γ is the equilibrium under vanished perturbation of Γ^{UA} .

2.2.4 Example Revisited

In this subsection, we revisit the example of Game 2 in Chapter 1. We will show how the strategic uncertainty is modelled with random perturbation and how the equilibrium is determined under perturbation with uncertainty aversion. We also provide with this example the idea of parameterized perturbation, which will be used in the rest of the paper. Consider Game 2 in Figure 1.1. As we have seen, there is a unique mixed-strategy Nash equilibrium in this game; $s^* = (s_1^*, s_2^*) = ((\frac{1}{2}, \frac{1}{2}), (\frac{1}{5}, \frac{4}{5}))$. Suppose that for $j \in \{1, 2\}$, when the player j is expected to choose a strategy s_j , the opponent player i is concerned that

s_j is perturbed by perturbations $v_j [1] = (\epsilon, -\epsilon)$ or $v_j [2] = (-\epsilon, \epsilon)$ with equal possibility. ϵ is given as a small positive value. To summarize, the random perturbation of player j 's strategy s_j is as follows.

$$\Upsilon_j = \left((\epsilon, -\epsilon), (-\epsilon, \epsilon); \frac{1}{2}, \frac{1}{2} \right)$$

Then, the perturbed strategy of s_j by Υ_j is written as

$$P_j(s_j, \Upsilon_j) = \left(s_j + (\epsilon, -\epsilon), s_j + (-\epsilon, \epsilon); \frac{1}{2}, \frac{1}{2} \right).$$

For the sake of convenience, we will use in this example one particular characterization of uncertainty averse preferences which is the overlap between KMM and MMR. The value of the strategy t_i under $P_j(s_j, \Upsilon_j)$ with uncertainty aversion is represented as follows:

$$\Phi_i(t_i, P_j(s_j, \Upsilon_j)) = \sum_{n=1}^2 \frac{1}{2} \cdot \phi_i(U_i(t_i, s_j + v_j[n]))$$

where $\phi_i(t) = -e^{-\rho_i t}$. Since $s_i^1 + s_i^2 = 1$, it is enough to characterize the best response function $B_i^{UA}(s_j, \Upsilon_j)$ with the first element of it, which will be written as $BR_i(s_j^1, \epsilon)$. The strict concavity of ϕ_i (i.e. strictly positive ρ_i) guarantees that the first order condition is the necessary and sufficient condition for the maximizer of value function. With a simple calculation, we can obtain the best response functions,

$$BR_1(s_2^1, \epsilon) = \frac{1}{5} + \frac{1}{10\epsilon\rho_1} \ln \left(\frac{\epsilon + (s_2^1 - \frac{1}{5})}{\epsilon - (s_2^1 - \frac{1}{5})} \right), \text{ and}$$

$$BR_2(s_1^1, \epsilon) = \frac{1}{2} - \frac{1}{4\epsilon\rho_2} \ln \left(\frac{\epsilon + (s_1^1 - \frac{1}{2})}{\epsilon - (s_1^1 - \frac{1}{2})} \right).$$

Figure 2.1 illustrates the best response functions of both players. Consider player 1. Her best response to s_2^* is $(\frac{1}{5}, \frac{4}{5})$, which is her maximin strategy. The reason is as follows: Suppose $s_2 = s_2^*$. Without perturbation, $(U_1(a_1^1, s_2^*), U_1(a_1^2, s_2^*)) = (\frac{4}{5}, \frac{4}{5})$. If the perturbation is introduced, then $(U_1(a_1^1, s_2^* + v_2[n]), U_1(a_1^2, s_2^* + v_2[n])) = (\frac{4}{5} + 4v_2^1[n], \frac{4}{5} - v_2^1[n])$. By choosing $s_1 = (\frac{1}{5}, \frac{4}{5})$, $U_1(s_1, s_2^* + v_2[n]) = \frac{4}{5}$ for any $n \in \{1, 2\}$. Thus the maximin strategy

guarantees the expected payoff of $\frac{4}{5}$ regardless of which perturbation is, and she can perfectly hedge against the uncertainty. The presense of uncertainty in belief and the aversion to it forces the player to choose the secure strategy.

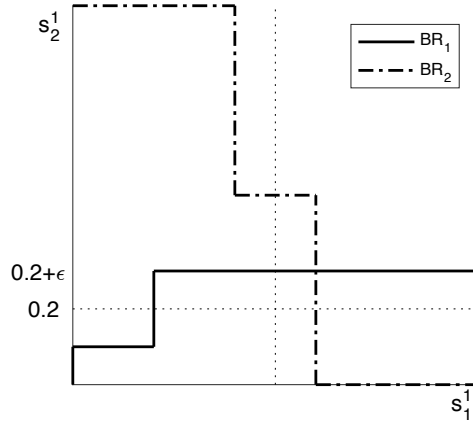
If s_2 changes from s_2^* , the best response reacts to this change. This is because now $(U_1(a_1^1, s_2^* + v_2[n]), U_1(a_1^2, s_2^* + v_2[n])) = (\frac{4}{5} + 4(s_2^1 - \frac{1}{5}) + 4v_2^1[n], \frac{4}{5} - (s_2^1 - \frac{1}{5}) - v_2^1[n])$. If $s_2^1 > \frac{1}{5}$, the player 1 is willing to choose higher value of s_1^1 than $\frac{1}{5}$. Two forces are combined to determine the best response: hedging against the uncertainty versus reaction to the change in s_2 . The best response function of player 1 is formed in this way. For player 2, the maximin strategy is $(\frac{1}{2}, \frac{1}{2})$ and she is willing to increase s_2^1 when s_1^1 decreases. Thus her best response function is as depicted in Figure 2.1.

Now, look at the intersection $(\hat{s}_1^1, \hat{s}_2^1)$ of two best response functions. At this intersection, the following consistency condition is satisfied: $\hat{s}_1^1 = BR_1(\hat{s}_2^1, \epsilon)$ and $\hat{s}_2^1 = BR_2(\hat{s}_1^1, \epsilon)$. As we explained in Section 1, this consistency is a relaxed version of the Nash equilibrium consistency condition. Thus the strategy profile $\hat{s} = (\hat{s}_1, \hat{s}_2)$ is the equilibrium under perturbation $\Upsilon = (\Upsilon_1, \Upsilon_2)$. Here, we emphasize two important aspects of \hat{s} , which are closely related to the main results of this paper. First, \hat{s} deviates from s^* . This discrepancy could be substantial if the degree of uncertainty (i.e. size of ϵ) is significant. The discrepancy shrinks as the degree of uncertainty decreases. This is because as ϵ decreases the hedging motive reduces, so each player's best response approaches to the case without perturbation. As a result, \hat{s} converges to s^* as ϵ vanishes. Second, the equilibrium \hat{s} is strict in the sense that each player's strategy choice is strictly preferred. Thus, in Harsanyi (1973)'s perspective, it is stable. The smoothness property of KMM and MRR is critical for the strictness (or stability) of equilibrium. Suppose that each player's preference is represented by GS model:

$$\Phi_i(s_i, C_j) = \min_{p_j^1 \in C_j} U_i(s_i, p_j)$$

where $C_j = [s_j^1 - \epsilon, s_j^1 + \epsilon] \in \mathbb{R}$, which is the convex hull of the set $\{s_j^1 + v_j^1[n] : n = 1, 2\}$. Figure 2.2 clearly shows that the best response of GS model is non-smooth. Player i sticks to her maximin strategy if $s_j^1 \in (s_j^{*,1} - \epsilon, s_j^{*,1} + \epsilon)$ and jumps to the pure strategies otherwise. This property is due to its extreme aversion to uncertainty. We can find the intersection

Figure 2.1: Best response functions and equilibrium with uncertainty aversion : GS case



of two best response functions. However, in this intersection each player's strategy is not a strict best response. Thus, this strategy profile still has the instability problem for the same reason of instability of mixed-strategy Nash equilibria, which Harsanyi (1973) pointed out.

2.3 Nash Equilibria as Limits of Equilibria under Perturbation

In this section we will show that any regular (i.e. strict or quasi-strict with non-singular Jacobian J^*) Nash equilibrium s^* has its corresponding equilibrium s under perturbation Υ such that s is close enough to s^* if Υ is close enough to zero. The next lemma shows that for any quasi-strict Nash equilibrium s^* of Γ , if an equilibrium s under perturbation Υ is close enough to s^* and Υ is close enough to zero, then $C(s)$ should be the same as $C(s^*)$. To state the result, we need a notion of convergence of perturbation. Consider a sequence $\{\Upsilon(n)\}$ of perturbation. We denote by $M_i(n)$ the total number of cases in perturbation $\Upsilon_i(n)$. Of course, $M_i(n)$ may have different values for different n . We now define $\lim_{n \rightarrow \infty} \Upsilon(n) = 0$ if for any given $\delta > 0$, there exists $n^0 \in \mathbb{N}$ such that for all $n \geq n^0$,

$$\left| v_i^k(n)[m_i] \right| < \delta \text{ for all } i; \text{ all } m_i = 1, \dots, M_i(n); \text{ all } k.$$

Lemma 4: Let s^* be a quasi-strict Nash equilibrium of Γ . Suppose that there exists a sequence $\{s(n)\}$ of equilibria under perturbation $\Upsilon(n)$ such that $\lim_{n \rightarrow \infty} s(n) = s^*$ and $\lim_{n \rightarrow \infty} \Upsilon(n) = 0$. Then, for all n large enough $C(s(n)) = C(s^*)$.

By Lemma 4, we can restrict our attention to strategy profiles s with $C(s) = C(s^*)$. Now, we will consider parameterized perturbation for s with $C(s) = C(s^*)$.

2.3.1 Parameterized Perturbation

We define a perturbation structure $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$ for s with $C(s) = C(s^*)$ as follows: Let $\Upsilon(\epsilon) = (\Upsilon_1(\epsilon), \dots, \Upsilon_N(\epsilon))$ where $\Upsilon_i(\epsilon) = (v_i[1](\epsilon), \dots, v_i[M_i](\epsilon); w_i[1], \dots, w_i[M_i])$.⁸ For all i and all m_i , each $v_i^k[m_i](\epsilon)$ is a continuously differentiable function of ϵ on $[0, \infty)$ such that

- (i) $v_i^k[m_i](0) = 0$ for all k ,
- (ii) $v_i^k[m_i] : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ for $i \in I; k = 1, \dots, \gamma_i$,
- (iii) $v_i^k[m_i] : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^-$ for $i \notin I; k = 1, \dots$,
- (iv) $v_i^k[m_i] : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ for all $i; k = \gamma_i + 1, \dots, K_i$, and
- (v) $\sum_{k=1}^{K_i} v_i^k[m_i](\epsilon) = 0$ for any $\epsilon \in [0, \infty)$.

Note that not all perturbations $\Upsilon(\epsilon)$ in a given perturbation structure $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$ are perturbations, because a perturbation should satisfy that $p_i[m_i] \in \Delta^{K_i-1}$ for all i ; all m_i . But since each $v_i^k[m_i](\epsilon)$ is continuous for ϵ and $v_i^k[m_i](0) = 0$, $v_i^k[m_i](\epsilon)$ is close enough to zero for a small enough ϵ . Consequently, for any perturbation structure $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$, we can always find $\bar{\epsilon} > 0$ such that for all $\epsilon \in [0, \bar{\epsilon})$, $\Upsilon(\epsilon)$ is a perturbation of s that is near s^* . To be precise, suppose s^* is strict and consider s with $C(s) = C(s^*)$. Then $s = s^*$. We can find $\bar{\epsilon} > 0$ such that for all $\epsilon \in [0, \bar{\epsilon})$ and for all i and all m_i ,

$$1 + v_i^1[m_i](\epsilon) \geq 0, \text{ and therefore } p_i[m_i](s_i^*, \epsilon) \in \Delta^{K_i-1}.$$

where $p_i[m_i](s_i^*, \epsilon) = s_i^* + v_i[m_i](\epsilon)$.

Next, suppose s^* is quasi-strict but not strict, and consider s with $C(s) = C(s^*)$. Then we consider the following construction: For each $i \in I$, there exists $\bar{\epsilon}_i > 0$ satisfying

⁸So, it is assumed that for all i, M_i and $(w_i[1], \dots, w_i[M_i])$ are the constants in any $\Upsilon(\epsilon)$.

the following: There exists a neighborhood $N_{\sigma_i^*} \subset \Delta_c^{\gamma_i-1}$ of σ_i^* such that for all $(\sigma_i, \epsilon) \in N_{\sigma_i^*} \times [0, \bar{\epsilon}_i)$ and all m_i ,

$$s_i^k + v_i^k[m_i](\epsilon) > 0 \text{ for } k = 1, \dots, \gamma_i, \text{ and}$$

$$\left| \sum_{k=1}^{\gamma_i} v_i^k[m_i](\epsilon) \right| < 1 \text{ and therefore } p_i[m_i](s_i, \epsilon) \in \Delta^{K_i-1}$$

where $p_i[m_i](s_i, \epsilon) = s_i + v_i[m_i](\epsilon)$. The idea of the above construction of $\bar{\epsilon}_i$ and $N_{\sigma_i^*}$ is the following: Since $\sigma_i^* \in \text{int}(\Delta_c^{\gamma_i-1})$,

$$s_i^{*,k} > 0 \text{ for } k = 1, \dots, \gamma_i.$$

Thus we can find $\bar{\epsilon}_i > 0$ such that for all $\epsilon \in [0, \bar{\epsilon}_i)$ and all m_i ,

$$s_i^{*,k} + v_i^k[m_i](\epsilon) > 0 \text{ for } k = 1, \dots, \gamma_i, \text{ and } \left| \sum_{k=1}^{\gamma_i} v_i^k[m_i](\epsilon) \right| < 1.$$

Then we can find a neighborhood $N_{\sigma_i^*}$ of σ_i^* such that for all $(\sigma_i, \epsilon) \in N_{\sigma_i^*} \times [0, \bar{\epsilon}_i)$ and all m_i ,

$$s_i^k + v_i^k[m_i](\epsilon) > 0 \text{ for } k = 1, \dots, \gamma_i.$$

Also there exists $\bar{\epsilon}_0 > 0$ such that for all $\epsilon \in [0, \bar{\epsilon}_0)$ and for all $i \notin I$ and all m_i , $p_i[m_i](\epsilon) \in \Delta^{K_i-1}$. We define $\bar{\epsilon} = \min\{\bar{\epsilon}_0, \bar{\epsilon}_i : i \in I\}$. Then, for $(\sigma, \epsilon) \in \prod_{i \in I} N_{\sigma_i^*} \times [0, \bar{\epsilon})$, $\Upsilon(\epsilon)$ is a perturbation of s . Thus, in the proof of Theorems 1 and 2, we only consider $\{\Upsilon(\epsilon)\}_{\epsilon \in [0, \bar{\epsilon})}$. For example, consider Game 5 in Figure 2.2. $s^* = ((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}))$ is a unique quasi-strict Nash equilibrium. Suppose that the perturbation structure $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$ is given as follows: For $i \in \{1, 2\}$,

$$v_i[1](\epsilon) = (2\epsilon, -\epsilon, -\epsilon), w_i[1] = \frac{1}{3}$$

$$v_i[2](\epsilon) = (-\epsilon, 2\epsilon, -\epsilon), w_i[2] = \frac{1}{3}$$

$$v_i[3](\epsilon) = (-\epsilon, -\epsilon, 2\epsilon), w_i[3] = \frac{1}{3}$$

Figure 2.2: Game 5

	a_2^1	a_2^2	a_2^3
a_1^1	0 0	1 -1	-1 1
a_1^2	-1 1	0 0	1 -1
a_1^3	1 -1	-1 1	0 0

Figure 2.3 is the mixed-strategy space of player i , Δ^{γ_i-1} . The end point of each blue line from s_i^* represents $p_i[m_i](s_i^*, \bar{\epsilon})$, $m_i \in \{1, 2, 3\}$ for some $\bar{\epsilon} > 0$. For any s_i in the interior of the small grey triangle and any $\epsilon \in [0, \bar{\epsilon})$, $p_i[m_i](s_i, \epsilon) \in \Delta^{\gamma_i-1}$, $m_i \in \{1, 2, 3\}$.

In the rest of the paper, we use the following notation: For a given $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$,

(i) $P_i(s_i, \epsilon) \doteq P_i(s_i, \Upsilon_i(\epsilon))$.

(ii) We write $p_i[m_i](s_i, \epsilon)$ to express that $p_i[m_i]$ is a function of (s_i, ϵ) : $p_i[m_i](s_i, \epsilon) = s_i + v_i[m_i](\epsilon)$.

(iii) Similar to (ii), $\bar{p}_i[r_i](s, \epsilon) = (p_1[r_i](s_1, \epsilon), \dots, p_{i-1}[r_i](s_{i-1}, \epsilon), p_{i+1}[r_i](s_{i+1}, \epsilon), \dots, p_N[r_i](s_N, \epsilon))$.

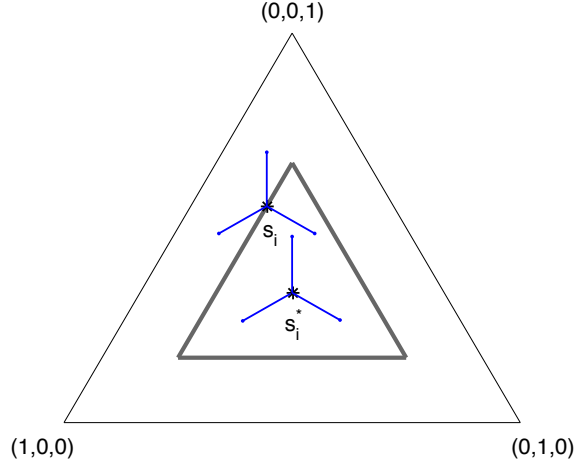
2.3.2 Existence and Convergence of Equilibria under Perturbation

Our question is if there exists an equilibrium $s(\epsilon)$ under perturbation $\Upsilon(\epsilon)$ which is close enough to s^* for small enough ϵ . The following two theorems will show that this is indeed the case for any regular Nash equilibrium s^* and any perturbation structure $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$, and that the sequence $\{s(\epsilon)\}_\epsilon$ of such corresponding equilibria is unique.

Theorem 1: For any strict Nash equilibrium s^* and any perturbation structure $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$, there exists $\epsilon^S > 0$ such that s^* is a strict equilibrium under perturbation $\Upsilon(\epsilon)$ for any $\epsilon \in [0, \epsilon^S)$.

Thus, for strict Nash equilibria, the convergence is trivial because the corresponding equilibria $s(\epsilon) = s^*$.

Figure 2.3: Construction of $\bar{\epsilon}$



Now, we turn to a quasi-strict Nash equilibrium s^* with non-singular Jacobian J^* . We will prove the following theorem.

Theorem 2: For any quasi-strict Nash equilibrium s^* with non-singular Jacobian J^* and any perturbation structure $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$, there exists $\epsilon^{QS} > 0$ and a unique sequence $\{s(\epsilon)\}_{\epsilon \in [0, \epsilon^{QS}]}$ of quasi-strict equilibria under perturbations $\Upsilon(\epsilon)$ such that $\lim_{\epsilon \rightarrow 0} s(\epsilon) = s^*$.

We first show the following lemma.

Lemma 5: For any quasi-strict Nash equilibrium s^* with non-singular Jacobian J^* and any perturbation structure $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$, there exists $\epsilon^{L5} > 0$ and a unique sequence $\{s(\epsilon)\}_{\epsilon \in [0, \epsilon^{L5}]}$ of strategy profiles with $C(s(\epsilon)) = C(s^*)$ such that

(i) $\sigma(\epsilon)$ characterizing $s(\epsilon)$ is a solution of the following system of equations:

$$\frac{\partial \Phi_i}{\partial \sigma_i} = 0 \text{ for } i \in I. \quad (2.4)$$

(ii) $\lim_{\epsilon \rightarrow 0} s(\epsilon) = s^*$.

Based on Lemma 5, we need to show the following to prove the theorem: For small enough ϵ , $s(\epsilon)$ obtained in Lemma 5 is an equilibrium under perturbation $\Upsilon(\epsilon)$. That is, for all i

$$s_i(\epsilon) \in \operatorname{argmax}_{t_i} \Phi_i(t_i, \overline{P_i(s_i(\epsilon), \epsilon)}). \quad (2.5)$$

The complete proof is in the appendix.

2.4 Strictness of Equilibria under Perturbation

In Theorem 1, we showed that any strict Nash equilibrium has its strict corresponding equilibrium under perturbation for small enough ϵ , which is itself. However, for a quasi-strict Nash equilibrium s^* with non-singular J^* , the corresponding equilibrium $s(\epsilon)$ of $\{s(\epsilon)\}_{[0, \epsilon^{QS}]}$ may not be strict. It is notable that the uniqueness of $s(\epsilon)$ for a given $\epsilon \in [0, \epsilon^{QS}]$ does not imply strictness of $s(\epsilon)$. The reason is that such uniqueness may result from the fact that $s(\epsilon)$ should satisfy a *system* of equations, that is,

$$\frac{\partial \Phi_i}{\partial \sigma_i} = 0 \text{ for } i \in I.$$

where $|I| > 1$. But the strictness of $s(\epsilon)$ requires that each $s_i(\epsilon)$ should be a strict best response in $s(\epsilon)$, or equivalently, the unique maximizer of Φ_i . Indeed, this is the essence of the instability problem of mixed-strategy Nash equilibria, which Harsanyi (1973a) argued: A quasi-strict Nash equilibrium s^* with non-singular J^* satisfies the following system of equations,

$$\frac{\partial U_i}{\partial \sigma_i} = 0 \text{ for } i \in I.$$

where $|I| > 1$, and therefore it is locally unique (i.e. isolated). However, for each $i \in I$, s_i^* is not a strict best response in s^* .

In this section we will show that the corresponding $s(\epsilon)$ to a quasi-strict Nash equilibrium s^* with non-singular J^* is strict under some additional conditions for $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$.

2.4.1 Conditions on Perturbation for Strictness

We introduce the following assumptions on $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$.

Assumption 1: For $i \in I$, all m_i , and all k , $v_i^k[m_i]$ is an analytic function of ϵ on $[0, \bar{\epsilon}]$.

Due to this assumption, each $v_i^k[m_i](\epsilon)$ can be written in a form

$$v_i^k[m_i](\epsilon) = \eta_i^k[m_i] \cdot \epsilon + \nu_i^k[m_i](\epsilon) \text{ for } \epsilon \in [0, \epsilon^A)$$

for some $\epsilon^A \in (0, \epsilon^{QS})$, where $\eta_i^k[m_i]$ is a constant of first-order term and $\nu_i^k[m_i](\epsilon)$ is the remaining higher-order terms. For small enough ϵ , the first order term is dominant, so we will impose the following conditions on $\eta_i^k[m_i]$.

Assumption 2 (Order condition) : For $i \in I$, all m_i , and $k = \gamma_i + 1, \dots, K_i$, $\eta_i^k[m_i] = 0$.

This condition reflects the idea of properness concept of Myerson (1978). As we have seen in the proof of Theorem 2, for any small enough ϵ and any s close enough to s^* ,

$$U_i(a_i^{\gamma_i}, \bar{p}_i[r_i](s, \epsilon)) > U_i(a_i^k, \bar{p}_i[r_i](s, \epsilon)) \text{ for all } i; \text{ all } r_i; k = \gamma_i + 1, \dots, K_i.$$

Thus all pure strategies outside the carrier are strictly dominated in all cases she faces, i.e., $r_i = 1, \dots, R_i$. With the idea of properness we can assume that the other players will make their belief in a following way: The probability for player i to mix pure strategies outside the carrier by mistake is much smaller than the probability she mixes pure strategies in the carrier a little more or less by mistake. Thus, in any possible case of mistake, $m_i = 1, \dots, M_i$, the order of perturbation for all pure strategies outside the carrier should be higher than 1.

We introduce the following vectors and matrixes for efficient expressions. We define

$$\eta_i[m_i] = \left(\eta_i^1[m_i], \dots, \eta_i^{\gamma_i-1}[m_i] \right)^T.$$

By Assumption 2 and the fact that $\eta_i^{\gamma_i}[m_i] = -\sum_{k=1}^{\gamma_i} \eta_i^k[m_i]$, $\eta_i[m_i]$ fully characterizes the first order part of $v_i[m_i]$. The weighted average of $\eta_i[m_i]$ will be denoted by $\bar{\eta}_i$,

$$\bar{\eta}_i = \sum_{m_i=1}^{M_i} w_i[m_i] \eta_i[m_i].$$

We define the normalized vectors $\hat{\eta}_i[m_i]$ by

$$\hat{\eta}_i[m_i] = \eta_i[m_i] - \bar{\eta}_i.$$

Finally, we denote by $\hat{\eta}_i$ the aggregation matrix of $\hat{\eta}_i[m_i]$, $m_i = 1, \dots, M_i - 1$,

$$\hat{\eta}_i = \begin{bmatrix} \hat{\eta}_i[1] & \cdots & \hat{\eta}_i[M_i - 1] \end{bmatrix}.$$

For small enough ϵ , $\hat{\eta}_i$ captures the structure of perturbation in player i 's strategy, whereas ϵ captures the degree of uncertainty. The following assumptions on $\hat{\eta}_i$ requires that perturbation structure should be fully diversified.

Assumption 3 (Richness condition): $\hat{\eta}_i$ has full row rank, $\gamma_i - 1$.

In the rest of this subsection, we introduce several objects which will be helpful to the examples and proofs in the next subsection. First, we define the following vectors

$$\hat{\eta}_i[m_i] = (\eta_i^1[m_i], \dots, \eta_i^{\gamma_i}[m_i])^T \quad \text{and} \quad \hat{\eta}_i[m_i] = \eta_i[m_i] - \bar{\eta}_i$$

where $\bar{\eta}_i = \sum_{m_i=1}^{M_i} w_i[m_i] \hat{\eta}_i[m_i]$, and the aggregation matrices

$$\hat{\eta}_i = \begin{bmatrix} \hat{\eta}_i[1] & \cdots & \hat{\eta}_i[M_i] \end{bmatrix} \quad \text{and} \quad \hat{\hat{\eta}}_i = \begin{bmatrix} \hat{\hat{\eta}}_i[1] & \cdots & \hat{\hat{\eta}}_i[M_i] \end{bmatrix}.$$

We introduce $\hat{\eta}_i$ and $\hat{\hat{\eta}}_i$ because the following might be the most natural steps to think of first order perturbations: We first consider the whole structure of first order perturbation with $\hat{\eta}_i$. Then we consider its normalized one $\hat{\hat{\eta}}_i$. And then we obtain $\hat{\eta}_i$ by eliminating the last column and the last row of $\hat{\hat{\eta}}_i$, since they are redundant for the rank condition (that is, $\text{rank}(\hat{\hat{\eta}}_i) = \text{rank}(\hat{\eta}_i)$). Thus, the full rank condition of $\hat{\eta}_i$ captures the requirement for full diversification of the first order perturbation.

Next, the following objects are introduced for the concise expression. We redefine $v_i[m_i]$ as follows,

$$v_i[m_i] = \left(v_i^1[m_i], \dots, v_i^{\gamma_i-1}[m_i], v_i^{\gamma_i+1}[m_i], \dots, v_i^{K_i}[m_i] \right)^T$$

because it fully characterizes the m_i th perturbation of player i due to the fact that $v_i^{\gamma_i}[m_i] = -\sum_{k \neq \gamma_i} v_i^k[m_i]$. Then we denote by \bar{v}_i the weighted average of $v_i[m_i]$, and by $\hat{v}_i[m_i]$ the normalized $v_i[m_i]$:

$$\bar{v}_i = \sum_{m_i=1}^{M_i} w_i[m_i] v_i[m_i] \quad \text{and} \quad \hat{v}_i[m_i] = v_i[m_i] - \bar{v}_i.$$

Similarly, the following vectors are defined for the higher order part of perturbation: We write

$$\nu_i[m_i] = \left(\nu_i^1[m_i], \dots, \nu_i^{\gamma_i-1}[m_i], \nu_i^{\gamma_i+1}[m_i], \dots, \nu_i^{K_i}[m_i] \right)^T$$

and denote by $\bar{\nu}_i$ the weighted average of $\nu_i[m_i]$, and by $\hat{\nu}_i[m_i]$ the normalized one,

$$\bar{\nu}_i = \sum_{m_i=1}^{M_i} w_i[m_i] \nu_i[m_i] \quad \text{and} \quad \hat{\nu}_i[m_i] = \nu_i[m_i] - \bar{\nu}_i.$$

Finally, we denote by $\hat{\nu}_i$ the aggregation matrix of $\hat{\nu}_i[m_i]$, $m_i = 1, \dots, M_i - 1$,

$$\hat{\nu}_i = \begin{bmatrix} \hat{\nu}_i[1] & \cdots & \hat{\nu}_i[M_i - 1] \end{bmatrix}.$$

2.4.2 Proof of Strictness

In this subsection we prove the following theorem.

Theorem 3: For any quasi-strict Nash equilibrium s^* with non-singular Jacobian J^* and any perturbation structure $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$ satisfying Assumption 1-3, there exists $\epsilon^{QS^*} \in (0, \epsilon^A)$ and a unique sequence $\{s(\epsilon)\}_{\epsilon \in (0, \epsilon^{QS^*})}$ of *strict* equilibria under perturbations $\Upsilon(\epsilon)$ such that $\lim_{\epsilon \rightarrow 0} s(\epsilon) = s^*$.

First, it is an immediate result of Theorem 2 that for any $\epsilon \in (0, \epsilon^{QS})$, (2.2) holds for player $i \notin I$. Next, consider player $i \in I$. In Theorem 2, we have already proven that $\sigma_i(\epsilon)$ is a maximizer of $\Phi_i(\sigma_i, \overline{P_i(s_i(\epsilon), \epsilon)})$ for any $\epsilon \in (0, \epsilon^{QS})$. This is due to concavity of Φ_i and the fact that

$$D_h^1 \left(\Phi_i(\sigma_i(\epsilon), \overline{P_i(s_i(\epsilon), \epsilon)}) \right) = 0 \text{ for any nonzero } h \in \mathbb{R}^{\gamma_i-1}.$$

⁹ (i.e. condition (i) in Lemma 5). But to show that $\sigma_i(\epsilon)$ is a unique maximizer, we need to prove that

$$D_h^2 \left(\Phi_i(\sigma_i(\epsilon), \overline{P_i(s_i(\epsilon), \epsilon)}) \right) < 0 \text{ for any nonzero } h \in \mathbb{R}^{\gamma_i-1}. \quad (2.6)$$

⁹We denote by $D_h^n \left(\Phi_i(\sigma_i, \overline{P_i(s_i)}) \right)$ the n th order directional derivative of $\Phi_i(\sigma_i, \overline{P_i(s_i)})$ for a given $\overline{P_i(s_i)}$ at σ_i in the direction of h .

In the following lemma, we will introduce the equivalent condition for (2.6). To get an intuition, recall the value function of KMM.

$$\Phi_i^{KMM}(\sigma_i, \overline{P_i(s_i(\epsilon), \epsilon)}) = \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \phi_i \left(\sum_{k=1}^{\gamma_i-1} s_i^k \cdot \varphi_i^k(\overline{p_i}[r_i](s(\epsilon), \epsilon)) + U_i(a_i^{\gamma_i}, \overline{p_i}[r_i](s(\epsilon), \epsilon)) \right)$$

For notational simplicity, we denote $\varphi_i^k(\overline{p_i}[r_i](s(\epsilon), \epsilon))$ by $\alpha_i^k[r_i]$. Suppose that there exists some nonzero $h = (h^1, \dots, h^{\gamma_i-1}) \in \mathbb{R}^{\gamma_i-1}$ such that

$$\sum_{k=1}^{\gamma_i-1} h^k \cdot \alpha_i^k[r_i] = 0 \text{ for all } r_i = 1, \dots, R_i$$

Then, $\sigma_i(\epsilon)$ is not a unique maximizer, since

$$\Phi_i^{KMM}(\sigma_i(\epsilon)) = \Phi_i^{KMM}(\sigma_i(\epsilon) + th) \text{ for small enough } t > 0.$$

Thus, the condition (2.6) is equivalent to the condition that there is no nonzero $h \in \mathbb{R}^{\gamma_i-1}$ such that

$$\alpha_i[r_i]^T \cdot h = 0 \text{ for all } r_i. \quad (2.7)$$

where $\alpha_i[r_i] = (\alpha_i^1[r_i], \dots, \alpha_i^{\gamma_i-1}[r_i])^T$. Let α_i denote the aggregation matrix of $\alpha_i[r_i]$,

$$\alpha_i = \begin{bmatrix} \alpha_i[1] & \cdots & \alpha_i[R_i] \end{bmatrix}.$$

Then, the condition (2.7) can be written as follows, $\alpha_i^T \cdot h = 0$. Finally, we obtain the next lemma. The detailed proof is in the appendix.

Lemma 6: For any given $\epsilon \in (0, \epsilon^{QS})$, $D_h^2 \left(\Phi_i(\sigma_i(\epsilon), \overline{P_i(s_i(\epsilon), \epsilon)}) \right) < 0$ for any nonzero $h \in \mathbb{R}^{\gamma_i-1}$ if and only if α_i has full row rank, $\gamma_i - 1$.

We now introduce the sufficient condition for the full rank condition of α_i , which will be directly used in the proof of Theorem 3. To state the lemma, we need some preparations. Given player $i \in I$, without loss of generality we can assume that her R_i th case is a combination of the last case of each player $j \neq i$, that is,

$$\overline{p_i}[R_i] = (p_1[M_1], \dots, p_{i-1}[M_{i-1}], p_{i+1}[M_{i+1}], \dots, p_n[M_n]).$$

We denote by $[R_i \setminus m_j]$ the case which results from the R_i th case by replacing the player j 's last case by j 's m_j th case:

$$\bar{p}_i[R_i \setminus m_j] = (p_1[M_1], \dots, p_j[m_j], \dots, p_{i-1}[M_{i-1}], p_{i+1}[M_{i+1}], \dots, p_n[M_n]).$$

With $\bar{p}_i[R_i \setminus m_j]$, we can express player i 's assessment in which only player j 's perturbation changes from $m_j = 1$ to M_j while all other players' perturbations are fixed at their last cases. We denote by $\bar{\alpha}_{i,j}$ the weighted average of $\alpha_i[R_i \setminus m_j]$ and by $\hat{\alpha}_i[R_i \setminus m_j]$ the normalized one with it:

$$\bar{\alpha}_{i,j} = \sum_{m_j=1}^{M_j} w_j[m_j] \alpha_i[R_i \setminus m_j] \text{ and } \hat{\alpha}_i[R_i \setminus m_j] = \alpha_i[R_i \setminus m_j] - \bar{\alpha}_{i,j}.$$

We denote by $\hat{\alpha}_{i,j}$ the aggregation matrix of $\hat{\alpha}_i[R_i \setminus m_j]$,

$$\hat{\alpha}_{i,j} = \begin{bmatrix} \hat{\alpha}_i[R_i \setminus 1] & \cdots & \hat{\alpha}_i[R_i \setminus M_j - 1] \end{bmatrix}.$$

Finally, we define the following aggregation matrix.

$$\hat{\alpha}_i = \begin{bmatrix} \hat{\alpha}_{i,1} & \cdots & \hat{\alpha}_{i,j} & \cdots & \hat{\alpha}_{i,n} \end{bmatrix}, j \neq i, j \in I.$$

Lemma 7: α_i has full row rank if $\hat{\alpha}_i$ has full row rank.

Guiding Examples for Proof

The following two examples will help us understand how Assumptions 1-3 and the non-singularity of J^* provide strictness to $s(\epsilon)$. The complete proof of Theorem 3 is in the appendix. We first introduce the following lemma, which will be used in the next examples and the proof of Theorem 3.

Lemma 8: Suppose that all the elements of matrix $P(\epsilon)$ is parameterized by ϵ and that $\lim_{\epsilon \rightarrow 0} P(\epsilon) = P^0$ in the sense that all elements in $P(\epsilon)$ converge to the corresponding elements of P^0 as ϵ tends to zero. Then there exists ϵ^0 such that for all $\epsilon \in [0, \epsilon^0)$, $\text{rank}(P(\epsilon)) \geq \text{rank}(P^0)$.

Example 2 : Consider the following 4×4 Game 6.

Figure 2.4: Game 6

	a_2^1	a_2^2	a_2^3	a_2^4
a_1^1	0 0	1 -1	-1 1	1 $-\frac{1}{2}$
a_1^2	-1 1	0 0	1 -1	2 $-\frac{1}{2}$
a_1^3	1 -1	-1 1	0 0	3 $-\frac{1}{2}$
a_1^4	$-\frac{1}{2}$ 1	$-\frac{1}{2}$ 2	$-\frac{1}{2}$ 3	$-\frac{1}{2}$ $-\frac{1}{2}$

$s^* = ((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0), (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0))$ is a quasi-strict equilibrium, since for $i \in \{1, 2\}$ and $j \neq i$,

$$U_i(a_i^1, s_j^*) = U_i(a_i^2, s_j^*) = U_i(a_i^3, s_j^*) = 0 > -\frac{1}{2} = U_i(a_i^4, s_j^*).$$

We write

$$J_{\varphi_i}(\bar{\sigma}_i) = J_{\varphi_i}(\sigma_1; \dots; \sigma_j; \dots; \sigma_n) = \frac{\partial(\varphi_i^1, \dots, \varphi_i^{\gamma_i-1})}{\partial(s_1^1, \dots, s_1^{\gamma_1-1}; \dots; s_j^1, \dots, s_j^{\gamma_j-1}; \dots; s_n^1, \dots, s_n^{\gamma_n-1})}$$

$$= \begin{bmatrix} J_{\varphi_i}(\sigma_1) & \cdots & J_{\varphi_i}(\sigma_j) & \cdots & J_{\varphi_i}(\sigma_n) \end{bmatrix}$$

where $j \neq i$ and $j \in I$. Let J_i^* denote the matrix $J_{\varphi_i}(\bar{\sigma}_i^*)$. Specifically, in any 2-person games, $J_{\varphi_i}(\bar{\sigma}_i)$ is a constant matrix. Also, in 2-person games, we can write J^* in the form

$$J^* = \begin{bmatrix} O & J_1^* \\ J_2^* & O \end{bmatrix}$$

In Game 6, each J_i^* has full row rank,

$$J_1^* = J_2^* = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}.$$

and thus J^* is non-singular.

We now turn to perturbation structure. We assume that the first order perturbation and the weights are given in the following way: For $i \in \{1, 2\}$,

$$\hat{\eta}_i = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix},$$

$$\eta_i^4[m_i] = 0 \text{ for } m_i = 1, 2, 3, \text{ and}$$

$$w_i[m_i] = \frac{1}{3} \text{ for } m_i = 1, 2, 3.$$

Assumption 2 is trivially satisfied. Also, since $\hat{\eta}_i = \hat{\eta}_i$ and $\text{rank}(\hat{\eta}_i) = \text{rank}(\hat{\eta}_i) = 2$, Assumption 3 is satisfied.

We will show that $s_1(\epsilon)$ is a strict best response in $s(\epsilon)$. By Lemma 7, this holds if $\hat{\alpha}_1$ has full row rank, $\gamma_1 - 1$. Since this example is a 2-person game, player 1 has only one opponent player. Thus, this condition can be simply written as

$$\text{rank}(\hat{\alpha}_{1,2}) = \text{rank}\left(\begin{bmatrix} \hat{\alpha}_1[1] & \hat{\alpha}_1[2] \end{bmatrix}\right) = \gamma_1 - 1 = 2.$$

The function $\varphi_i^k(s_j)$ has $K_j - 1$ independent variables, $(s_j^1, \dots, s_j^{\gamma_j-1}, s_j^{\gamma_j+1}, \dots, s_j^{K_j})$, since $s_j^{\gamma_j} = 1 - \sum_{l \neq \gamma_j} s_j^l$. Also, $\varphi_i^k(s_j)$ is linear in each s_j^l , $l = 1, \dots, \gamma_j - 1, \gamma_j + 1, \dots, K_j$. Thus any $\frac{\partial \varphi_i^k}{\partial s_j^l}$ is constant. Consequently, we have the following:

$$\begin{aligned} \hat{\alpha}_i^k[m_j] &= \varphi_i^k(p_j[m_j](s_j(\epsilon), \epsilon)) - \sum_{m'_j=1}^{M_j} w_j[m'_j] \varphi_i^k(p_j[m'_j](s_j(\epsilon), \epsilon)) \\ &= \sum_{l \neq \gamma_i} \left(\frac{\partial \varphi_i^k}{\partial s_j^l} \right) \cdot \left(p_j^l[m_j](s_j(\epsilon), \epsilon) - \sum_{m'_j=1}^{M_j} w_j[m'_j] p_j^l[m'_j](s_j(\epsilon), \epsilon) \right) \\ &= \left[\frac{\partial \varphi_i^k}{\partial s_j^1}, \dots, \frac{\partial \varphi_i^k}{\partial s_j^{\gamma_j-1}}, \frac{\partial \varphi_i^k}{\partial s_j^{\gamma_j+1}}, \dots, \frac{\partial \varphi_i^k}{\partial s_j^{K_j}} \right] \cdot \left(\hat{v}_j^1[m_j](\epsilon), \dots, \hat{v}_j^{\gamma_j-1}[m_j](\epsilon), \hat{v}_j^{\gamma_j+1}[m_j](\epsilon), \dots, \hat{v}_j^{K_j}[m_j](\epsilon) \right)^T \\ &= J_{\varphi_i^k}(s_j) \cdot \hat{v}_j[m_j](\epsilon) \\ &= J_{\varphi_i^k}(s_j) \cdot \hat{\eta}_j[m_j] \cdot \epsilon + J_{\varphi_i^k}(s_j) \cdot \hat{v}_j[m_j](\epsilon) \end{aligned}$$

The third equality holds because

$$\begin{aligned} p_j^l[m_j](s_j(\epsilon), \epsilon) - p_j^l[m'_j](s_j(\epsilon), \epsilon) &= \left(s_j^l(\epsilon) + v_j^l[m_j](\epsilon) \right) - \left(s_j^l(\epsilon) + v_j^l[m'_j](\epsilon) \right) \\ &= v_j^l[m_j](\epsilon) - v_j^l[m'_j](\epsilon) \end{aligned}$$

and therefore,

$$p_j^l[m_j](s_j(\epsilon), \epsilon) - \sum_{m'_j=1}^{M_j} w_j[m'_j] p_j^l[m'_j](s_j(\epsilon), \epsilon) = v_j^l[m_j](\epsilon) - \sum_{m'_j=1}^{M_j} w_j[m'_j] v_j^l[m'_j](\epsilon) = \hat{v}_j^l[m_j](\epsilon).$$

Note that $J_{\varphi_i^k}(s_j)$ is different from $J_{\varphi_i^k}(\sigma_j)$, that is,

$$\begin{aligned} J_{\varphi_i^k}(s_j) &= \left[\frac{\partial \varphi_i^k}{\partial s_j^1}, \dots, \frac{\partial \varphi_i^k}{\partial s_j^{\gamma_j-1}}, \frac{\partial \varphi_i^k}{\partial s_j^{\gamma_j+1}}, \dots, \frac{\partial \varphi_i^k}{\partial s_j^{K_i}} \right], \text{ whereas} \\ J_{\varphi_i^k}(\sigma_j) &= \left[\frac{\partial \varphi_i^k}{\partial s_j^1}, \dots, \frac{\partial \varphi_i^k}{\partial s_j^{\gamma_j-1}} \right]. \end{aligned}$$

Then, we have the following formula:

$$\hat{\alpha}_i[m_j] = \begin{bmatrix} \hat{\alpha}_i^1[m_j] \\ \vdots \\ \hat{\alpha}_i^{\gamma_i-1}[m_j] \end{bmatrix} = J_{\varphi_i}(\sigma_j) \cdot \hat{\eta}_j[m_j] \cdot \epsilon + J_{\varphi_i}(s_j) \cdot \hat{v}_j[m_j](\epsilon).$$

Finally, we have the following.

$$\begin{aligned} \hat{\alpha}_{1,2} &= \begin{bmatrix} \hat{\alpha}_1[1] & \hat{\alpha}_1[2] \end{bmatrix} \\ &= \epsilon \cdot J_1^* \cdot \begin{bmatrix} \hat{\eta}_2[1] & \hat{\eta}_2[2] \end{bmatrix} + J_{\varphi_1}(s_j) \cdot \begin{bmatrix} \hat{v}_2[1](\epsilon) & \hat{v}_2[2](\epsilon) \end{bmatrix} \\ &= \epsilon \cdot J_1^* \cdot \hat{\eta}_2 + J_{\varphi_1}(s_2) \cdot \hat{v}_2(\epsilon). \end{aligned}$$

We can divide the above equation by any non-zero ϵ , and then we have

$$\frac{1}{\epsilon}\hat{\alpha}_{1,2} = J_1^* \cdot \hat{\eta}_2 + R_1(\epsilon).$$

where $R_1(\epsilon) = J_{\varphi_1}(s_2) \cdot \frac{1}{\epsilon}\hat{\nu}_2(\epsilon)$. By Assumption 2, the lowest order of ϵ in any elements of the matrix $\frac{1}{\epsilon}\hat{\nu}_2(\epsilon)$ is equal to 1. Also, $J_{\varphi_1}(s_2)$ is a constant matrix. Thus, $\lim_{\epsilon \rightarrow 0} R_1(\epsilon) = 0$ and we have the following equation.

$$\lim_{\epsilon \rightarrow 0} \left(\frac{1}{\epsilon}\hat{\alpha}_1 \right) = J_1^* \cdot \hat{\eta}_2.$$

By Lemma 8, there exists $\epsilon^{(1)}$ such that for any $\epsilon \in (0, \epsilon^{(1)})$

$$\text{rank} \left(\frac{1}{\epsilon}\hat{\alpha}_1 \right) \geq \text{rank} (J_1^* \cdot \hat{\eta}_2).$$

By Assumption 3, $\hat{\eta}_2$ has full row rank. We use the property that if a matrix Q has full row rank, then $\text{rank}(PQ) = \text{rank}(P)$. By this property, $\text{rank}(J_1^* \cdot \hat{\eta}_2) = \text{rank}(J_1^*) = \gamma_1 - 1$. Consequently, for any $\epsilon \in (0, \epsilon^{(1)})$

$$\text{rank}(\hat{\alpha}_1) = \text{rank} \left(\frac{1}{\epsilon}\hat{\alpha}_1 \right) \geq \gamma_1 - 1,$$

which implies that $\text{rank}(\hat{\alpha}_1)$ is $\gamma_1 - 1$ and thus $s_1(\epsilon)$ is strictly preferred. \square

Example 2 showed how the order condition and richness condition play their roles for $s_i(\epsilon)$ to be a strict best response in $s(\epsilon)$. It also showed that the full row-rank condition of the matrix $J_{\varphi_i}(\bar{\sigma}_i)$ is critical. This condition comes from the fact that

- (i) $J_{\varphi_i}(\bar{\sigma}_i)$ is constant as J_i^* and
- (ii) J_i^* has full row rank.

(ii) holds because J^* is non-singular. But (i) held in the previous example because Jacobian is constant in 2-person games, to be precise, in games where $|I| = 2$. The next example will show that condition (i) may not hold for games where $|I| > 2$. But $s_i(\epsilon)$ will still be strictly preferred for any small enough ϵ due to the fact that

$$\lim_{\epsilon \rightarrow 0} J_{\varphi_i}(\bar{\sigma}_i) \Big|_{\bar{p}_i[r_i](s(\epsilon), \epsilon)} = J_{\varphi_i}(\bar{\sigma}_i) \Big|_{s_i^*} = J_i^*.$$

Example 3 (3-person game): Consider the following $4 \times 3 \times 2$ Game 7 in Figure 2.5.

Figure 2.5: Game 7

	a_2^1	a_2^2	a_2^3
a_1^1	-3 0 9	-2 -1 9	-4 1 9
a_1^2	-4 1 9	-3 0 9	-2 -1 9
a_1^4	-2 -1 9	-4 1 9	-3 0 9
a_1^3	3 0 -9	3 0 -9	3 0 -9
	a_3^1		

	a_2^1	a_2^2	a_2^3
a_1^1	3 0 -9	4 -1 -9	2 1 -9
a_1^2	2 1 -9	3 0 -9	4 -1 -9
a_1^4	4 -1 -9	2 1 -9	3 0 -9
a_1^3	-3 0 9	-3 0 9	-3 0 9
	a_3^2		

In Game 7, player 1 has an interaction with player 2 and player 3, but player 2 and player 3 are interacted with only player 1, respectively. Player 1's actions can be divided into two groups; playing the constant sum games with player 2, $\{a_1^1, a_1^2, a_1^4\}$, or not playing the game with player 2, $\{a_1^3\}$ ¹⁰. And each group of actions has a conflict with player 3's action choice.

In Game 7, $s^* = ((\frac{1}{6}, \frac{1}{6}, \frac{1}{2}, \frac{1}{6}), (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), (\frac{1}{2}, \frac{1}{2}))$ is a quasi-strict equilibrium. Jacobian $J_\varphi(\sigma)$ is as follows.

$$J_\varphi(\sigma) = J_\varphi(s_1^1, s_1^2, s_1^3; s_2^1, s_2^2; s_3^1) = \begin{bmatrix} & & & 0 & 3 & 0 \\ & O_{3 \times 3} & & -3 & 0 & 0 \\ & & & -2s_3^1 + 1 & 2s_3^1 - 1 & 2(6 - s_2^1 + s_2^2) \\ 0 & -3 & -1 & & & 0 \\ 3 & 0 & 1 & O_{2 \times 2} & & 0 \\ 0 & 0 & -36 & 0 & 0 & O_{1 \times 1} \end{bmatrix}$$

s^* has non-singular Jacobian J^* ,

$$J^* = J_\varphi(\sigma)|_{s^*} = \begin{bmatrix} & & & 0 & 3 & 0 \\ & O_{3 \times 3} & & -3 & 0 & 0 \\ & & & 0 & 0 & 12 \\ 0 & -3 & -1 & & & 0 \\ 3 & 0 & 1 & O_{2 \times 2} & & 0 \\ 0 & 0 & -36 & 0 & 0 & O_{1 \times 1} \end{bmatrix}.$$

We assume that the first order perturbation and the weights for each player are given in the following way under Assumptions 1-3:

¹⁰We intentionally label the player 1's actions in the above way.

$$\hat{\eta}_1 = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}, \hat{\eta}_2 = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}, \hat{\eta}_3 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \text{ and}$$

$$w_i[m_i] = \frac{1}{M_i} \text{ for all } i; \text{ all } m_i.$$

where $M_1 = 4$, $M_2 = 3$ and $M_3 = 2$.

We will show that $\hat{\alpha}_1 = \begin{bmatrix} \hat{\alpha}_{1,2} & \hat{\alpha}_{1,3} \end{bmatrix}$ has full row rank, $\gamma_1 - 1$, which implies that $s_1(\epsilon)$ is strictly preferred. Consider

$$\hat{\alpha}_{1,j} = \begin{bmatrix} \hat{\alpha}_1[R_1 \setminus 1] & \cdots & \hat{\alpha}_1[R_1 \setminus M_j - 1] \end{bmatrix}.$$

To reformulate the m_j th column vector $\hat{\alpha}_1[R_1 \setminus m_j]$, we first consider the k th element of it,

$$\hat{\alpha}_1^k[R_1 \setminus m_j] = \alpha_1^k[R_1 \setminus m_j] - \bar{\alpha}_{1,j}^k.$$

Since $\alpha_1^k[R_1 \setminus m_j]$ and $\alpha_1^k[R_1 \setminus m'_j]$ only differ in player j 's perturbation,

$$\begin{aligned} \hat{\alpha}_1^k[R_1 \setminus m_j] &= \left[\frac{\partial \varphi_1^k}{\partial s_j^1}, \dots, \frac{\partial \varphi_1^k}{\partial s_j^{\gamma_j - 1}} \right]_{\bar{p}_1[R_1](s(\epsilon), \epsilon)} \cdot \left(\hat{v}_j^1[m_j](\epsilon), \dots, \hat{v}_j^{\gamma_j - 1}[m_j](\epsilon) \right)^T \\ &= J_{\varphi_1^k}(\sigma_j) \Big|_{\bar{p}_1[R_1](s(\epsilon), \epsilon)} \cdot \hat{v}_j[m_j](\epsilon) \end{aligned}$$

Note that $J_{\varphi_i^k}(\sigma_j)$ is not a function of σ_j , since it is the first-order partial derivative of φ_i^k which is linear in each s_j^k . Thus we can simply express that $J_{\varphi_1^k}(\sigma_j)$ is evaluated at $\bar{p}_1[R_1](s(\epsilon), \epsilon)$ as above. For instance, consider that $j = 2$. In the expression of $J_{\varphi}(\sigma)$ above, we can find that

$$J_{\varphi_3^k}(\sigma_2) = \begin{bmatrix} -2s_3^1 + 1 & 2s_3^1 - 1 \end{bmatrix}$$

which depends only on s_3^1 and does not involve any term s_2^k . Thus,

$$J_{\varphi_1^3}(\sigma_2) \Big|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} = J_{\varphi_1^3}(\sigma_2) \Big|_{p_3[M_3](s(\epsilon), \epsilon)} = \begin{bmatrix} -2p_3^1[M_3](s(\epsilon), \epsilon) + 1 & 2p_3^1[M_3](s(\epsilon), \epsilon) - 1 \end{bmatrix}.$$

Then, we have that

$$\hat{\alpha}_1[R_1 \setminus m_j] = \begin{bmatrix} \hat{\alpha}_1^1[R_1 \setminus m_j] \\ \vdots \\ \hat{\alpha}_1^{\gamma_1-1}[R_1 \setminus m_j] \end{bmatrix} = J_{\varphi_1}(\sigma_j) \Big|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} \cdot \hat{v}_j[m_j](\epsilon)$$

and that

$$\begin{aligned} \hat{\alpha}_{1,j} &= \begin{bmatrix} \hat{\alpha}_1[R_1 \setminus 1] & \cdots & \hat{\alpha}_1[R_1 \setminus M_j - 1] \end{bmatrix} \\ &= J_{\varphi_1}(\sigma_j) \Big|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} \cdot \begin{bmatrix} \hat{v}_j[1](\epsilon) & \cdots & \hat{v}_j[M_j - 1](\epsilon) \end{bmatrix} \\ &= J_{\varphi_1}(\sigma_j) \Big|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} \cdot \hat{\eta}_j \cdot \epsilon + J_{\varphi_1}(\sigma_j) \Big|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} \cdot \hat{v}_j(\epsilon) \end{aligned}$$

Finally, for any nonzero ϵ , we can merge $\frac{1}{\epsilon}\hat{\alpha}_{1,2}$ and $\frac{1}{\epsilon}\hat{\alpha}_{1,3}$, and obtain $\frac{1}{\epsilon}\hat{\alpha}_1$.

$$\begin{aligned} \frac{1}{\epsilon}\hat{\alpha}_1 &= \begin{bmatrix} \frac{1}{\epsilon}\hat{\alpha}_{1,2} & \frac{1}{\epsilon}\hat{\alpha}_{1,3} \end{bmatrix} \\ &= \begin{bmatrix} J_{\varphi_1}(\sigma_2) & J_{\varphi_1}(\sigma_3) \end{bmatrix} \Big|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} \cdot \begin{bmatrix} \hat{\eta}_2 + \frac{1}{\epsilon}\hat{v}_2(\epsilon) & O \\ O & \hat{\eta}_3 + \frac{1}{\epsilon}\hat{v}_3(\epsilon) \end{bmatrix} \\ &= J_{\varphi_1}(\overline{\sigma_1}) \Big|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} \cdot \Lambda_1 + \begin{bmatrix} R_{1,2}(\epsilon) & R_{1,3}(\epsilon) \end{bmatrix} \end{aligned}$$

where

$$\begin{aligned} \Lambda_1 &= \begin{bmatrix} \hat{\eta}_2 & O \\ O & \hat{\eta}_3 \end{bmatrix}, \text{ and} \\ R_{1,j}(\epsilon) &= J_{\varphi_1}(\sigma_j) \Big|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} \cdot \frac{1}{\epsilon}\hat{v}_j(\epsilon). \end{aligned}$$

Then, we have that $\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon}\hat{\alpha}_1 = J_1^* \cdot \Lambda_1$, because

$$\lim_{\epsilon \rightarrow 0} J_{\varphi_1}(\overline{\sigma_1})|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} = J_{\varphi_1}(\overline{\sigma_1})|_{\overline{\sigma_1^*}} = J_i^*,$$

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \hat{\nu}_j(\epsilon) = 0, \quad \text{and thus} \quad \lim_{\epsilon \rightarrow 0} R_{1,j}(\epsilon) = 0.$$

Therefore, by Lemma 8, there exists $\epsilon^{(1)} > 0$ such for any $\epsilon \in (0, \epsilon^{(1)})$,

$$\text{rank} \left(\frac{1}{\epsilon} \hat{\alpha}_1 \right) \geq \text{rank}(J_1^* \cdot \Lambda_1).$$

Λ_1 has full row rank since $\hat{\eta}_2$ and $\hat{\eta}_3$ have full row rank. Also, J_1^* has full row rank, $\gamma_1 - 1$, since J^* is non-singular. Consequently, for any $\epsilon \in (0, \epsilon^{(1)})$, $\text{rank}(\hat{\alpha}_1) = \gamma_1 - 1$.

Similarly, there exists $\epsilon^{(i)}$ for $i = 2, 3$ such that for any $\epsilon \in (0, \epsilon^{(i)})$, $\text{rank}(\hat{\alpha}_i) = \gamma_i - 1$. Finally, we set $\epsilon^{QS^*} = \min_{i \in I} \{\epsilon^{(i)}\}$. Then for any $\epsilon \in (0, \epsilon^{QS^*})$, each $s_i(\epsilon)$ is strictly preferred and $s(\epsilon)$ is strict.

Chapter 3

Testing Mixed-strategy Equilibria with Uncertainty Aversion: a Re-Examination of Experimental Studies

In this chapter, we test our model's predictive power by re-examining the experimental studies; Ochs (1995) and Selten and Chmura (2008). We then compare the performance of our model with the QRE model and show that our model outperforms the QRE model in both re-examinations. In the re-examination of Ochs (1995), our model's predictive power is significantly better than the QRE model in the reasonable range of risk aversion which might not be perfectly controlled in the original experiment of Ochs (1995). Re-examining Selten and Chmura (2008) also shows that the change in equilibrium prediction caused by the payoff transformation is actually observed in the data, which only our model can explain.

3.1 Re-examination of Ochs (1995)

In this section, we will first report the results of an experimental study by Ochs (1995), and the over-prediction problem of QRE model. Goeree et al. (2003) argues that the over-prediction of the QRE model is because of the failure of risk control in Ochs' experiment, and re-estimates the QRE model with consideration of the risk aversion of the subjects. They found that under the constant relative risk aversion specification, the estimated value of the risk aversion parameter is 0.42. However, this value is much higher than the estimated risk aversion parameter from the lottery-choice experiment in their study, which is 0.31. We show that the equilibrium under perturbation model can explain the observation with much lower risk aversion, 0.22, and that with this risk aversion value, there is a still significant over-prediction in the QRE model.

3.1.1 Summary of Results in Ochs (1995) and Over-prediction Problem of the QRE model

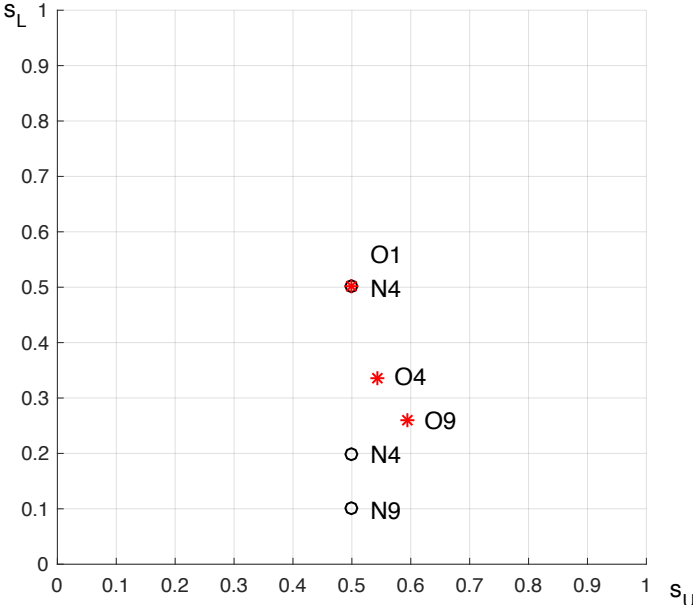
Three games in Figure 1.1 are used in Ochs (1995). Player 1's payoff in the upper-left corner is only different in these three games. We will denote by $A \in \{1, 4, 9\}$ the upper-left corner payoff of player 1. The experiment design is as follows: For each game, 16 subjects participated with fixed roles (player 1 or player 2) for 56-64 rounds. For each round, the subjects were randomly matched. Risk attitude was controlled by the binary lottery procedure. Outcome feedback is given to players after each round's play.

The Nash equilibrium predictions and the observed choice frequencies are depicted in Figure 3.1. For instance, $N9$ and $O9$ are the Nash equilibrium and the observed frequencies when $A = 9$, respectively. As depicted, there is a substantial discrepancy between the Nash equilibrium and actual choice when $A = 4$ or 9 , that is, when the game is asymmetric. As we can see in Figure 3.2, the cumulative frequency converges to the steady state. In Figure 3.2, the black dot is the choice frequency in each round and the blue line is the cumulative frequency. All the frequencies are aggregated for both groups of players.

Next, we show the best-fitting QRE model for asymmetric cases.¹ Before estimation, we first need to convert payoffs in each game into the following form in Figure 3.3. In Ochs

¹We re-estimated the QRE model based on the MATLAB code in Goeree et al. (2016).

Figure 3.1: Nash equilibria and observed frequencies in Ochs (1995)



experiments, the subjects were paid using a binary lottery procedure, and the probability of getting a reward in the lottery was determined by the total percentage of the maximum possible points (payoffs) accumulated over the whole experiment. Thus, we need to convert the original payoffs into the form of Figure 3.3 so that the exchange rate of accumulated payoffs to the expected reward in the lottery is the same for each player. This conversion is necessary for both models (i.e., the QRE model and our model) because the best response structures of both models are affected by a positive scalar multiple of a player’s payoffs.² Figure 3.4 shows the best fitting QRE model. We observe that the QRE model over-predicts player 1’s choice frequency s_U in both game 2 and game 3.

3.1.2 Estimation Results

For estimation of our model, we use the specification of perturbation structure and preferences as in Subsection 2.2.4. Our model has two parameters in the best response, ρ and ϵ .

²See McKelvey and Palfrey (1995) for more detailed explanation.

Figure 3.2: Choice frequencies in Ochs (1995)

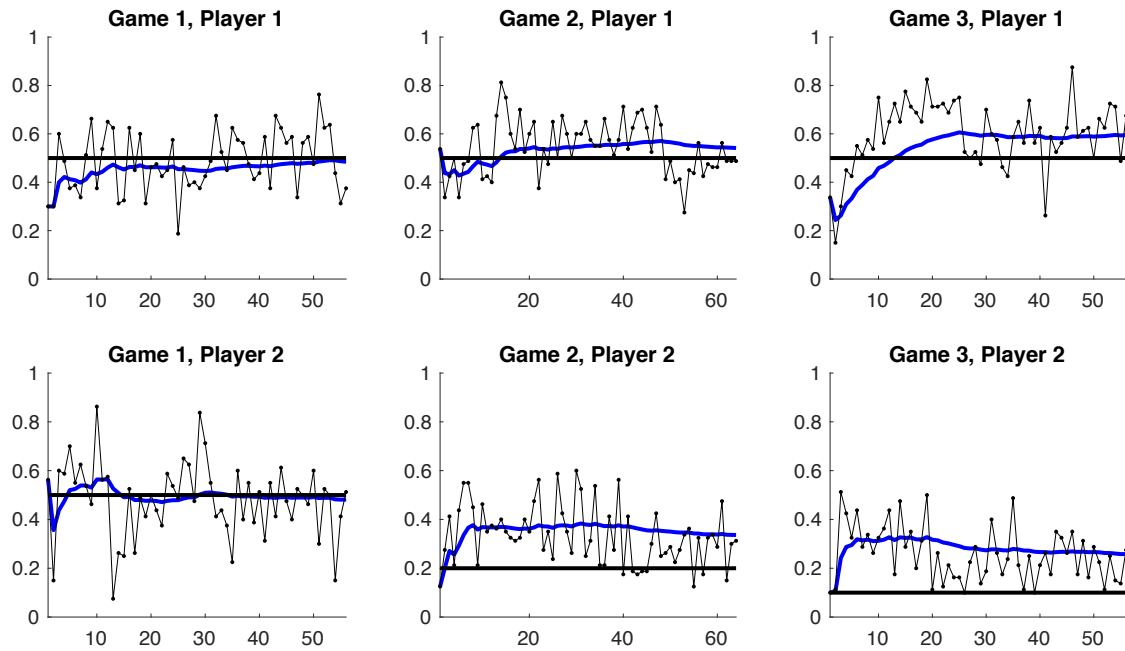
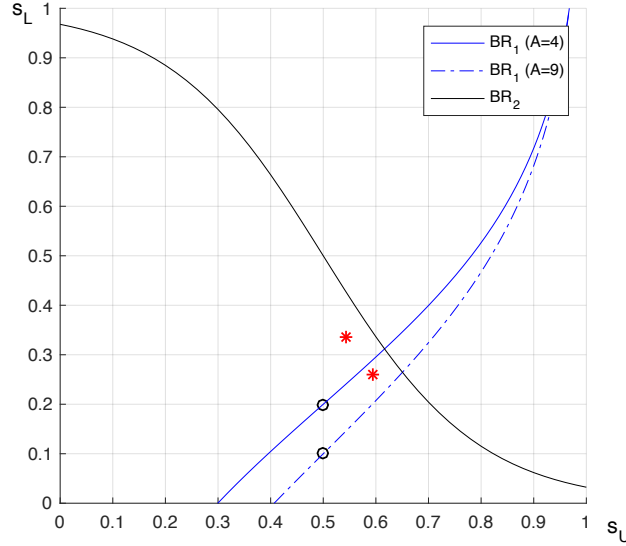


Figure 3.3: Normalized payoff for estimation in Ochs (1995)

		Game 2		Game 3	
		<i>L</i>	<i>R</i>	<i>L</i>	<i>R</i>
<i>U</i>		1	0	1	0
		0	1	0	1
<i>D</i>		0	0.250	0	0.111
		1	0	1	0

Figure 3.4: Over-prediction of the QRE model in Ochs (1995)



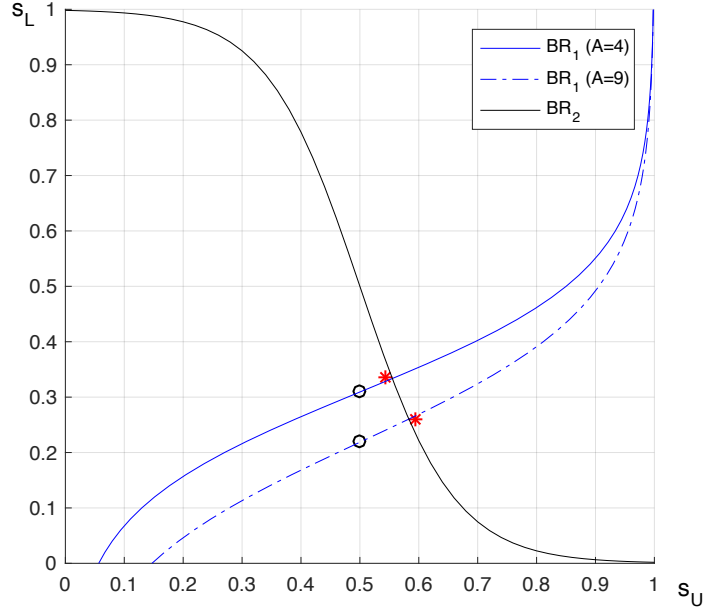
To make the number of parameters to be estimated the same in both models, we assume that $\rho_1 = \rho_2 = 5$, which is the moderate value of uncertainty aversion in the sense that the non-linearity of the best response function is not strong.³ Thus, we estimate ϵ which captures the degree of strategic uncertainty. In the QRE model, the object estimated is the precision parameter.

For both models, as Goeree et al. (2003), we assume that risk attitude in the experiment is not perfectly controlled by the binary lottery procedure. That is, all payoffs in Figure 2.3 are the utility unit, only when the risk attitudes of the players are well controlled as a risk-neutral. We transform the payoffs in Figure 2.3 into utility units by using the constant relative risk aversion, $U(x) = x^{1-r}$ as in Goeree et al. (2003). To summarize, the risk aversion parameter r is the object that we estimate additionally.

The results obtained are as follows: as Goeree et al. (2003) reported, the best fitting QRE model requires the high value of risk aversion parameter, $r = 0.42$. Figure 2.5 is the result. The reason for the high risk aversion estimate is straightforward. If the risk aversion is high, then the Nash equilibrium prediction of s_L increases, and thus gets closer to the

³We conformed that for $\rho \in [3, 7]$, the result is robust.

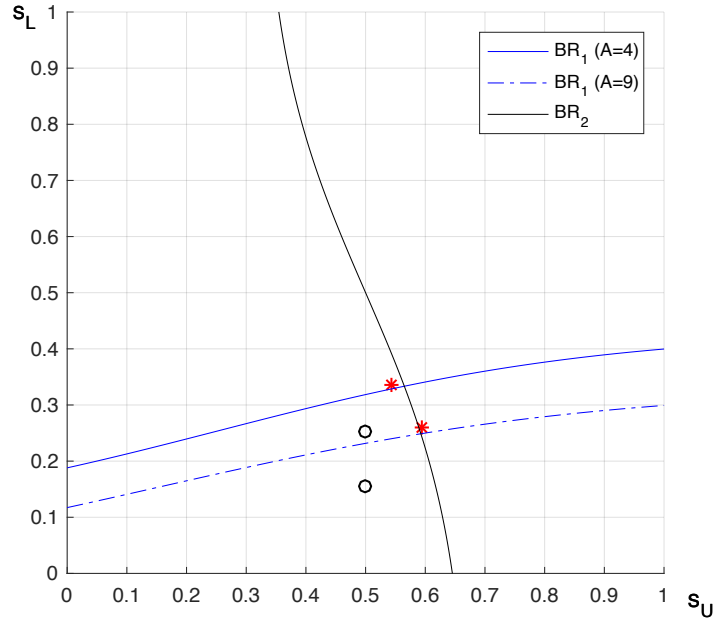
Figure 3.5: The best fitting of the QRE model in Ochs (1995)



observed value of s_L . We can observe this by comparing Figure 2.5 to Figure 2.4. Thus, the over-prediction of s_U by the QRE model can be mitigated by the high risk aversion, since it was the result of the relatively high discrepancy between the observed s_L and the Nash equilibrium prediction of it. But the problem is, again, the estimated risk-aversion is unreasonably high, because Ochs (1995) used the binary lottery procedure to control the risk attitude. Even though the risk attitude control in Ochs' experiment was not perfect, it would be expected that the controlled risk aversion in Ochs (1995) would be lower than the uncontrolled risk aversion in Goeree et al, (2003), $r = 0.31$. Harrison et al. (2013) shows in their experimental study that the binary lottery procedure induces a statistically significant shift toward risk neutrality.

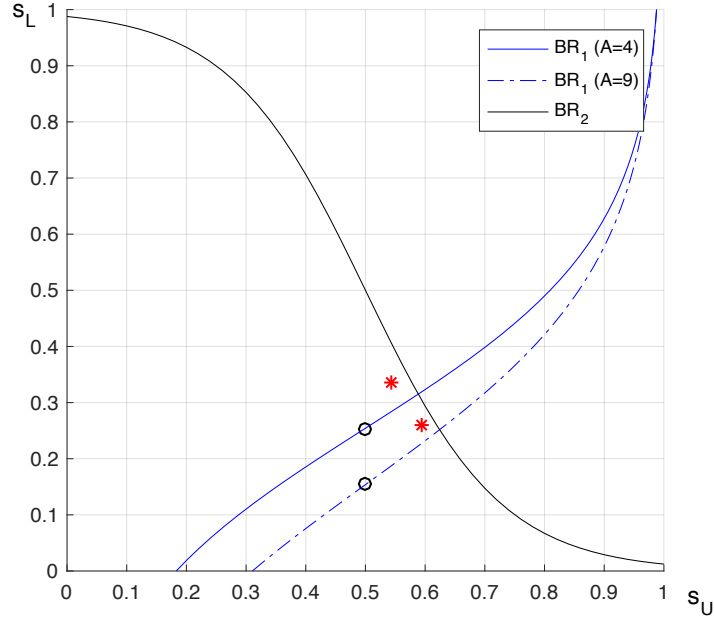
Next, we show that with our model, the estimated risk aversion, 0.22, is much lower than the case of the QRE model. Figure 3.6 is the result. The different feature of our model to the QRE model is that each player chooses her maximin strategy when the opponent player's choice is the Nash equilibrium strategy, rather than the equiprobable strategy as in the QRE model. To be specific, in this experiment design, player 2's maximin strategy

Figure 3.6: The best fitting of the equilibrium with uncertainty aversion in Ochs (1995)



is the equiprobable strategy, and thus player 2's best response functions of two models are very similar. (In linear approximation, they are identical in the sense that they are pivoting from $(s_U, s_L) = (\frac{1}{2}, \frac{1}{2})$). But player 1's maximin strategy is different from the equiprobable strategy, so player 1's best response of our model is almost a parallel shift of the best response function of the QRE model. As a result, this experiment gives us a practical way to test the two models' performance the most clearly. In Game 2 and Game 3, the best response function of our model shifts to the left from the QRE model, because in each game, the maximin strategy s_U is smaller than $1/2$. This aspect makes the quantitative difference between the two models' prediction. Consequently, the QRE model's over-prediction problem can be resolved without the unreasonably high value of risk aversion parameter estimated in the QRE model. Lastly, we show in Figure 3.7 that with $r = 0.22$, the over-prediction problem of the QRE model remains substantial, which shows the significant difference of two models.

Figure 3.7: The prediction of the QRE model when $r = 0.22$ in Ochs (1995)



3.2 Re-examination of Selten and Chmura (2008)

In this section, we first introduce the experiment design of Selten and Chmura (2008). In this study, there are interesting observations that the actual choice frequencies change as the payoffs are linearly transformed, which the mixed strategy Nash equilibria and the QRE model fail to explain. We then explain our model predicts such equilibrium changes according to linear payoff transformations. Lastly, we show the estimation results using our model and the QRE model, and conclude that our model's predictive power for the data in Selten and Chmura (2008) is significantly better than the QRE model's.

3.2.1 Experiment Design and Result

12 games in Figure 3.8 are used in the experiment of Selten and Chmura (2008). They are designed in pairs with linear payoff transformations. For instance, the game $G1'$ is the transformed game of the original game $G1$. Similar to Ochs (1995), each subjects participated with fixed roles for 200 rounds (in the original games) or 100 rounds (in the

transformed games), and they were randomly matched.

Figure 3.9 shows the result of this experiment. The first row shows the results of the pair of $G1$ and $G1'$, the pair of $G2$ and $G2'$, and the pair of $G3$ and $G3'$, respectively. The three sub-figures in the second row show the results of remainder. In each sub-figure, the black circle is the prediction of mixed-strategy Nash equilibrium, and the black triangle is the best-fitted prediction of the QRE model. The red dot and the orange dot are the actual choice frequency in the original game and the transformed game, respectively. Clearly, there are significant changes in choice frequencies between the original game and the transformed game, especially in the first four pairs of games. However, the mixed-strategy Nash equilibria and the QRE model fail to explain these equilibrium changes, since theoretically both model predict that there should not be equilibrium changes when there are such linear payoff transformations in the experiment design of Selten and Chmura (2008).

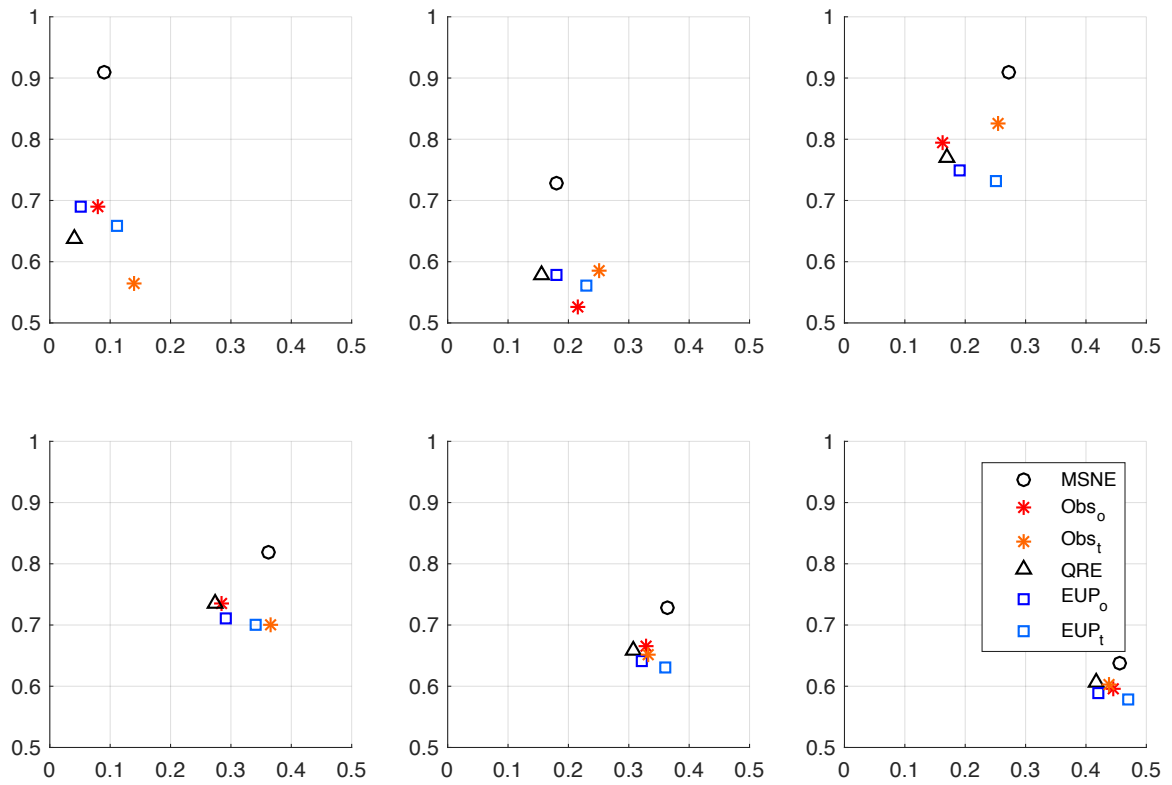
3.2.2 Payoff Transformation and Equilibrium Change in the Model of Equilibrium with Uncertainty Aversion

Different from Mixed-strategy Nash equilibria and the QRE model, our model predicts the equilibrium change when there is a linear payoff transformation. We will explain this point with Game 1 in Figure 1.1 and Game 4 in Figure 1.4. Game 4 is the transformed game of Game 1 in the same way of Selten and Chmura. Both games have $(s_U, s_L) = (\frac{1}{2}, \frac{1}{2})$ as a mixed strategy Nash equilibrium and the quantal response equilibrium under the logit equilibrium model specification. However, as we have seen in Figure 1.5, the prediction of equilibrium in Game 4 with our model differs from $(s_U, s_L) = (\frac{1}{2}, \frac{1}{2})$. The reason is as follows: In Game 4, by the linear payoff transformation, the maximin strategies of each players change from $s_U = \frac{1}{2}$ to $s_U = 0$ for player 1 and from $s_L = \frac{1}{2}$ to $s_L = 1$ for player 2, respectively. These changes make the shifts of each best response functions, which are depicted in Figure 1.5. To summarize, the payoff transformation changes the hedging behavior of players, which shifts the best response functions of players. As a result, the equilibrium, which is the intersection of best response functions, changes according to payoff transformation.

Figure 3.8: Games in Selten and Chmura (2008)

<i>G1 & G1'</i>	<i>G2 & G2'</i>	<i>G3 & G3'</i>																								
<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">10</td><td style="border: 1px solid black; padding: 2px;">0</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">8</td><td style="border: 1px solid black; padding: 2px;">18</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">9</td><td style="border: 1px solid black; padding: 2px;">10</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">9</td><td style="border: 1px solid black; padding: 2px;">8</td></tr> </table>	10	0	8	18	9	10	9	8	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">9</td><td style="border: 1px solid black; padding: 2px;">0</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">4</td><td style="border: 1px solid black; padding: 2px;">13</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">6</td><td style="border: 1px solid black; padding: 2px;">8</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">5</td></tr> </table>	9	0	4	13	6	8	7	5	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">8</td><td style="border: 1px solid black; padding: 2px;">0</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">6</td><td style="border: 1px solid black; padding: 2px;">14</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">10</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">4</td></tr> </table>	8	0	6	14	7	10	7	4
10	0																									
8	18																									
9	10																									
9	8																									
9	0																									
4	13																									
6	8																									
7	5																									
8	0																									
6	14																									
7	10																									
7	4																									
<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">10</td><td style="border: 1px solid black; padding: 2px;">4</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">12</td><td style="border: 1px solid black; padding: 2px;">22</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">9</td><td style="border: 1px solid black; padding: 2px;">14</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">9</td><td style="border: 1px solid black; padding: 2px;">8</td></tr> </table>	10	4	12	22	9	14	9	8	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">9</td><td style="border: 1px solid black; padding: 2px;">3</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">16</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">6</td><td style="border: 1px solid black; padding: 2px;">11</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">5</td></tr> </table>	9	3	7	16	6	11	7	5	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">8</td><td style="border: 1px solid black; padding: 2px;">3</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">9</td><td style="border: 1px solid black; padding: 2px;">17</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">13</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">4</td></tr> </table>	8	3	9	17	7	13	7	4
10	4																									
12	22																									
9	14																									
9	8																									
9	3																									
7	16																									
6	11																									
7	5																									
8	3																									
9	17																									
7	13																									
7	4																									
<i>G4 & G4'</i>	<i>G5 & G5'</i>	<i>G6 & G6'</i>																								
<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">0</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">4</td><td style="border: 1px solid black; padding: 2px;">11</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">5</td><td style="border: 1px solid black; padding: 2px;">9</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">6</td><td style="border: 1px solid black; padding: 2px;">2</td></tr> </table>	7	0	4	11	5	9	6	2	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">0</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">2</td><td style="border: 1px solid black; padding: 2px;">9</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">4</td><td style="border: 1px solid black; padding: 2px;">8</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">5</td><td style="border: 1px solid black; padding: 2px;">1</td></tr> </table>	7	0	2	9	4	8	5	1	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">1</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">1</td><td style="border: 1px solid black; padding: 2px;">7</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">3</td><td style="border: 1px solid black; padding: 2px;">8</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">5</td><td style="border: 1px solid black; padding: 2px;">0</td></tr> </table>	7	1	1	7	3	8	5	0
7	0																									
4	11																									
5	9																									
6	2																									
7	0																									
2	9																									
4	8																									
5	1																									
7	1																									
1	7																									
3	8																									
5	0																									
<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">2</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">6</td><td style="border: 1px solid black; padding: 2px;">13</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">5</td><td style="border: 1px solid black; padding: 2px;">11</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">6</td><td style="border: 1px solid black; padding: 2px;">2</td></tr> </table>	7	2	6	13	5	11	6	2	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">2</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">4</td><td style="border: 1px solid black; padding: 2px;">11</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">4</td><td style="border: 1px solid black; padding: 2px;">10</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">5</td><td style="border: 1px solid black; padding: 2px;">1</td></tr> </table>	7	2	4	11	4	10	5	1	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border: 1px solid black; padding: 2px;">7</td><td style="border: 1px solid black; padding: 2px;">3</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">3</td><td style="border: 1px solid black; padding: 2px;">9</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">3</td><td style="border: 1px solid black; padding: 2px;">10</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">5</td><td style="border: 1px solid black; padding: 2px;">0</td></tr> </table>	7	3	3	9	3	10	5	0
7	2																									
6	13																									
5	11																									
6	2																									
7	2																									
4	11																									
4	10																									
5	1																									
7	3																									
3	9																									
3	10																									
5	0																									

Figure 3.9: Observations and estimation results in Selten and Chmura (2008)



3.2.3 Estimation Results

Figure 3.9 summarized the estimation results. In each sub-figures, the blue dot and the skyblue dot are the best-fitted predictions of our model in the original game and the transformed game, respectively. Since we impose the strong restriction that the parameter representing the degree of uncertainty is identical in all 6 pairs of games, the predictive power of our model seem to be lower than in the estimation of Ochs (1995). However, compared to the QRE model, our model, at least in part, explains the change in choice frequencies; except for the frequency change of player 2 in $G2 - G2'$ pair and $G3 - G3'$ pair, our model explains the choice behavior at least qualitatively. To confirm that our model is statistically significant than the QRE model, we performed the Wilcoxon matched-pairs signed rank test as in Selten and Chmura, and found the significant difference in performances of two model with significant level of 0.05.

References

Battigalli, P., S. Cerreia Vioglio, F. Maccheroni, and M. Marinacci (2015). Self Conforming Equilibrium and Model Uncertainty, *American Economic Review*, 105(2), 646-77.

Battigalli, P., S. Cerreia Vioglio, F. Maccheroni, and M. Marinacci (2017). Mixed Extensions of Decision Problems under Uncertainty, *Economic Theory*, 63, 827-866.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011). Uncertainty Averse Preferences, *Journal of Economic Theory*, 146, 1275–1330.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2013). Ambiguity and Robust Statistics, *Journal of Economic Theory*, 148 (3), 974–1049.

Dominiak, A., and W. Schmedler (2011). Attitudes toward Uncertainty and Randomization: an Experimental Study, *Economic Theory*, 48, 289–312.

Dow, J. and D. Werlang (1994). Nash equilibrium under uncertainty: Breaking Down Backward Induction, *Journal of Economic Theory*, 64, 305–324.

Eichberger, J., S. Grant, and D. Kelsey (2016). Randomization and Dynamic Consistency, *Economic Theory*, 62, 547–566.

Eichberger, J, and D. Kelsey (2000). Non-additive Beliefs and Strategic Equilibria, *Games and Economic Behavior*, 30, 183–215.

Eichberger, J, and D. Kelsey (2002). Strategic Complements, Substitutes and Ambiguity: The Implications for Public Goods, *Journal of Economic Theory*, 106, 436–66.

- Eichberger, J, and D. Kelsey (2014). Optimism and Pessimism in Games, *International Economic Review*, 55, 483-505.
- Gilboa, I., and D. Schmeidler (1989). Maxmin Expected Utility with Non-unique Prior, *Journal of Mathematical Economics*, 18, 141-153.
- Goeree, J., C. Holt, and T. Palfrey (2003). Risk Averse Behavior in Generalized Matching Pennies Games, *Games and Economic Behavior*, 45, 97-113.
- Goeree, J., C. Holt, and T. Palfrey (2016). *Quantal Response Equilibrium*, Princeton University Press.
- Harrison, G., J. Martinez-Correab, and T. Swarthout (2013). Inducing Risk Neutral Preferences with Binary Lotteries: A Reconsideration, *Journal of Economic Behavior & Organization*, 94, 145-159.
- Harsanyi, J. (1973a). Games with Randomly Disturbed Payoffs, *International Journal of Game Theory*, 2, 1-23.
- Harsanyi, J. (1973b). Oddness of the Number of Equilibrium Points: a New Proof, *International Journal of Game Theory*, 2, 235-250.
- Klibanoff, P. (1996). Uncertainty, Decision and Normal Form Games, Technical report, Northwestern University.
- Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A Smooth Model of Decision Making Under Ambiguity, *Econometrica*, 73, 1849-1892.
- Lo, K.C., (1996). Equilibrium in Beliefs Under Uncertainty, *Journal of Economic Theory*, 71, 443-484.
- Lo, K.C., (2009). Correlated Nash equilibrium, *Journal of Economic Theory*, 144, 722-743.
- Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley: New York.
- Maccheroni, F., M. Marinacci, and A. Rustichini (2006). Ambiguity Aversion, Robustness and the Variational Representation of Preferences, *Econometrica*, 74, 1447-1498.

- Marinacci, M. (2000). Ambiguous Games, *Games and Economic Behavior*, 31, 191-219.
- McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior, In P. Zarembka, ed., *Frontiers in Econometrics*, New York: Academic Press.
- McKelvey, R., and T. Palfrey (1995). Quantal Response Equilibria for Normal Form Games, *Games and Economic Behavior*, 10, 6-38.
- Morris, S. (2008): "Purification," In *New Palgrave Dictionary of Economics*, ed. by S. Durlauf, and L. Blume. MacMillan.
- Myerson, R. B. (1978). Refinements of the Nash Equilibrium Concept, *International Journal of Game Theory*, 7 (2), 73-80.
- Ochs, J. (1995). Games with Unique Mixed Strategy Equilibria: An Experimental Study, *Games and Economic Behavior*, 10, 202-217.
- Saito, K. (2015). Preferences for Flexibility and Randomization Under Uncertainty, *American Economic Review*, 105, 1246-71.
- Schmeidler, D (1989). Subjective Probability and Expected Utility without Additivity, *Econometrica*, 57, 571-587.
- Seo, K (2001). Ambiguity and Second-order Belief, *Econometrica*, 77, 1575-1605.
- Sundaram, R. (1996). *A First Course in Optimization Theory*. Cambridge: Cambridge University Press.
- Van Damme, E. (1991). *Stability and the Perfection of Nash Equilibria*. Berlin: Springer-Verlag.

Appendix A

Appendix to Chapter 2

Proof of Lemma 1:

(a) For KMM : Recall that

$$\Phi_i^{KMM}(s_i, \overline{P_i(s_i, \Upsilon_i)}) = \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \phi_i(U_i(s_i, \overline{p_i}[r_i]))$$

Given $\overline{P_i(s_i, \Upsilon_i)}$, each $\phi_i(U_i(s_i, \overline{p_i}[r_i]))$ is concave in s_i , since it is a composite function of the concave function ϕ_i and the linear function U_i . Therefore, Φ_i is concave in s_i since it is a positive linear combination of concave functions.

(b) For MMR: Recall that

$$\Phi_i^{MMR}(s_i, \overline{P_i(s_i, \Upsilon_i)}) = \min_{Q_i \in \Delta^{R_i-1}} \left(\sum_{r_i=1}^{R_i} Q_i[r_i] \cdot U_i(s_i, \overline{p_i}[r_i]) + c_i(Q_i) \right)$$

Let $F : \Delta^{R_i-1} \times \mathbb{R}^{R_i}$ denote the function

$$F(Q_i, u) = \sum_{r_i=1}^{R_i} Q_i[r_i] \cdot u[r_i] + c_i(Q_i)$$

where $u = (u[1], \dots, u[R_i])$. Then, $F^*(u) \doteq \min_{Q_i \in \Delta^{R_i-1}} (F(Q_i, u))$ is continuous and concave in u by the maximum theorem¹, since F is convex on $\Delta^{R_i-1} \times \mathbb{R}^{R_i}$ and Δ^{R_i-1} is a convex set.

¹See Chapter 9 of Sundaram (1996).

Given $\overline{P_i(s_i, \Upsilon_i)}$, $\Phi_i^{MMR}(s_i, \overline{P_i(s_i, \Upsilon_i)})$ is a composite function of a concave function $F^*(u)$ and a linear function $u(s_i) = (u[1](s_i), \dots, u[R_i](s_i))$, where $u[r_i](s_i) = U_i(s_i, \overline{p_i}[r_i])$, thus it is concave in s_i .

Proof of Lemma 2:

We first observe that for all r_i ,

$$\begin{aligned}
U_i(\tau_i, \overline{p_i}[r_i]) &= \sum_{l=1}^{K_i} \tau_i^l \cdot U_i(a_i^l, \overline{p_i}[r_i]) \\
&= \left[\sum_{l=1}^{K_i} s_i^l \cdot U_i(a_i^l, \overline{p_i}[r_i]) \right] + (\tau_i^k - s_i^k) \cdot U_i(a_i^k, \overline{p_i}[r_i]) + (\tau_i^{k'} - s_i^{k'}) \cdot U_i(a_i^{k'}, \overline{p_i}[r_i]) \\
&= \left[\sum_{l=1}^{K_i} s_i^l \cdot U_i(a_i^l, \overline{p_i}[r_i]) \right] + (\tau_i^k - s_i^k) \cdot (U_i(a_i^k, \overline{p_i}[r_i]) - U_i(a_i^{k'}, \overline{p_i}[r_i])) \\
&> \sum_{l=1}^{K_i} s_i^l \cdot U_i(a_i^l, \overline{p_i}[r_i]) = U_i(s_i, \overline{p_i}[r_i]).
\end{aligned}$$

Then, for KMM,

$$\begin{aligned}
\Phi_i^{KMM}(\tau_i, \overline{P_i(s_i, \Upsilon_i)}) &= \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \phi_i(U_i(\tau_i, \overline{p_i}[r_i])) \\
&> \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \phi_i(U_i(s_i, \overline{p_i}[r_i])) = \Phi_i^{KMM}(s_i, \overline{P_i(s_i, \Upsilon_i)}).
\end{aligned}$$

For MMR, let \tilde{W}_i denote the distorted probability for τ_i . Then,

$$\begin{aligned}
\Phi_i^{MMR}(\tau_i, \overline{P_i(s_i, \Upsilon_i)}) &= \sum_{r_i=1}^{R_i} \tilde{W}_i[r_i] \cdot U_i(\tau_i, \overline{p_i}[r_i]) + c_i(\tilde{W}_i) \\
&> \sum_{r_i=1}^{R_i} \tilde{W}_i[r_i] \cdot U_i(s_i, \overline{p_i}[r_i]) + c_i(\tilde{W}_i) \geq \Phi_i^{MMR}(s_i, \overline{P_i(s_i, \Upsilon_i)})
\end{aligned}$$

Proof of Lemma 4:

We first clarify the notation: For a given (s, n) , we write

$$\overline{p}_i[r_i](s, n) = (p_j[r_i](s_j, n))_{j \neq i}$$

where each $p_j[r_i](s_j, n) = s_j + v_j[r_i](n)$, and $v_j[r_i](n)$ is an element of $\Upsilon(n)$.

It is clear that there exists \underline{n} such that for all i

$$C(s_i(n)) \supseteq C(s_i^*) = \{a_i^1, \dots, a_i^{\gamma_i}\} \text{ for all } n > \underline{n}. \quad (\text{A.1})$$

Next, since s^* is quasi-strict,

$$U_i(a_i^{\gamma_i}, \overline{s}_i^*) > U_i(a_i^k, \overline{s}_i^*) \text{ for } k = \gamma_i + 1, \dots, K_i. \quad (\text{A.2})$$

Thus, for all n large enough,

$$U_i(a_i^{\gamma_i}, \overline{p}_i[r_i](s(n), n)) > U_i(a_i^k, \overline{p}_i[r_i](s(n), n)) \text{ for all } r_i; k = \gamma_i + 1, \dots, K_i.$$

The reason is as follows: We can write

$$\begin{aligned} & U_i(a_i^{\gamma_i}, \overline{p}_i[r_i](s(n), n)) - U_i(a_i^k, \overline{p}_i[r_i](s(n), n)) \\ &= \left[U_i(a_i^{\gamma_i}, \overline{s}_i^*) - U_i(a_i^k, \overline{s}_i^*) \right] - \left[U_i(a_i^k, \overline{p}_i[r_i](s(n), n)) - U_i(a_i^k, \overline{s}_i^*) \right] \\ &+ \left[U_i(a_i^{\gamma_i}, \overline{p}_i[r_i](s(n), n)) - U_i(a_i^{\gamma_i}, \overline{s}_i^*) \right]. \end{aligned}$$

In the second line, the first term is positive due to (A.2). The second and the third term converge to zero since for all $j \neq i$ and for all r_i

$$\lim_{n \rightarrow \infty} p_j[r_i](s_j(n), n) = \lim_{n \rightarrow \infty} (s_j(n) + v_j[r_i](n)) = s_j^*$$

and $U_i(a_i^k, \overline{p}_i[r_i](s(n), n))$ is continuous in each $p_j[r_i](s(n), n)$. Consequently, by Lemma 3, for all n large enough,

$$a_i^k \notin C(s_i(n)) \text{ for } k = \gamma_i + 1, \dots, K_i. \quad (\text{A.3})$$

Finally, by (A.1) and (A.3), for all n large enough and for all i , $C(s_i(n)) = C(s_i^*)$.

Proof of Theorem 1:

We first clarify the following notation: For a given (s, ϵ) , we write

$$\overline{p}_i[r_i](s, \epsilon) = (p_j[r_i](s_j, \epsilon))_{j \neq i}$$

where each $p_j[r_i](s_j, \epsilon) = s_j(\epsilon) + v_j[r_i](\epsilon)$.

Since $s^* = (a_1^1, \dots, a_n^1)$ is strict, for all i

$$U_i(a_i^1, \overline{s}_i^*) > U_i(a_i^k, \overline{s}_i^*) \text{ for } k = 2, \dots, K_i. \quad (\text{A.4})$$

Then, we have

$$\lim_{\epsilon \rightarrow 0} U_i(a_i^k, \overline{p}_i[r_i](s^*, \epsilon)) = U_i(a_i^k, \overline{s}_i^*) \text{ for all } k; \text{ all } r_i. \quad (\text{A.5})$$

since for all $j \neq i$ and for all r_i

$$\lim_{\epsilon \rightarrow 0} p_j[r_i](s^*, \epsilon) = s_j^* + \lim_{\epsilon \rightarrow 0} v_j[r_i](s^*, \epsilon) = s_j^*.$$

By (A.4) and (A.5), there exists $\epsilon^S \in (0, \bar{\epsilon})$ such that for all $\epsilon \in [0, \epsilon^S)$

$$U_i(a_i^1, \overline{p}_i[r_i](s^*, \epsilon)) > U_i(a_i^k, \overline{p}_i[r_i](s^*, \epsilon)) \text{ for all } i; k = 2, \dots, K_i; \text{ all } r_i.$$

Consequently, by Lemma 3-(ii), for all $\epsilon \in [0, \epsilon^S)$

$$\{s_i^*\} = \{a_i^1\} = \operatorname{argmax}_{t_i \in S_i} \Phi_i(t_i, \overline{P}_i(s_i^*, \epsilon)) \text{ for all } i.$$

Thus for all $\epsilon \in [0, \epsilon^S)$, s^* is an equilibrium under perturbation $\Upsilon(\epsilon)$.

Proof of Lemma 5:

(i) For KMM: For a strategy profiles s with $C(s) = C(s^*)$,

$$\begin{aligned}
\Phi_i^{KMM}(s_i, \overline{P_i(s_i, \Upsilon_i)}) &= \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \phi_i(U_i(s_i, \overline{p_i}[r_i])) \\
&= \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \phi_i\left(\sum_{k=1}^{\gamma_i-1} s_i^k \cdot \varphi_i^k(\overline{p_i}[r_i]) + U_i(a_i^{\gamma_i}, \overline{p_i}[r_i])\right)
\end{aligned}$$

Thus for $k = 1, \dots, \gamma_i - 1$,

$$\frac{\partial \Phi_i^{KMM}}{\partial s_i^k} = \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \varphi_i^k(\overline{p_i}[r_i](\sigma, \epsilon)) \cdot \phi_i'(U_i(s_i, \overline{p_i}[r_i](\sigma, \epsilon))).$$

We define $\psi_i^k : \prod_{i \in I} N_{\sigma_i^*} \times [0, \bar{\epsilon}] \rightarrow \mathbb{R}$ by

$$\begin{aligned}
\psi_i^k(\sigma, \epsilon) &= \frac{1}{\phi_i'(U_i(s^*))} \frac{\partial \Phi_i^{KMM}}{\partial s_i^k} \\
&= \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \varphi_i^k(\overline{p_i}[r_i](\sigma, \epsilon)) \cdot \frac{\phi_i'(U_i(s_i, \overline{p_i}[r_i](\sigma, \epsilon)))}{\phi_i'(U_i(s^*))} \text{ for } i \in I; k = 1, \dots, \gamma_i - 1.
\end{aligned}$$

Since ϕ_i' is strictly positive, ψ_i^k is well-defined. Let ψ denote the vector

$$\psi = \left(\psi_1^1, \dots, \psi_1^{\gamma_1-1}; \dots; \psi_i^1, \dots, \psi_i^{\gamma_i-1}; \dots; \psi_N^1, \dots, \psi_N^{\gamma_N-1} \right), i \in I.$$

$\psi : \prod_{i \in I} N_{\sigma_i^*} \times [0, \bar{\epsilon}] \rightarrow \mathbb{R}^\gamma$ is continuously differentiable, since we assumed that ϕ_i is twice continuously differentiable and that $v_i[m_i](\epsilon)$ is continuously differentiable. The equation (2.4) in Lemma 5 is equivalent to the following equation:

$$\psi(\sigma, \epsilon) = 0. \tag{A.6}$$

We know the trivial solution of equation (A.6), $(\sigma, \epsilon) = (\sigma^*, 0)$. We will show that

$$\frac{\partial \psi}{\partial \sigma}(\sigma^*, 0) = J^*. \tag{A.7}$$

To prove this, observe that

$$\frac{\partial \psi_i^k}{\partial s_i^{k'}}(\sigma, 0) = \frac{\partial}{\partial s_i^{k'}} \left(\varphi_i^k(\bar{s}_i) \cdot \frac{\phi_i'(U_i(s_i, \bar{s}_i))}{\phi_i'(U_i(s^*))} \right)$$

Since $\varphi_i^k(\bar{s}_i^*) = 0$ for $i \in I; k = 1, \dots, \gamma_i - 1$, by evaluating the above equation at $\sigma = \sigma^*$,

$$\frac{\partial \psi_i^k}{\partial s_i^{k'}}(\sigma^*, 0) = \frac{\partial \varphi_i^k}{\partial s_i^{k'}}(\sigma^*) \cdot \left(\frac{\phi_i'(U_i(s^*))}{\phi_i'(U_i(s^*))} \right) + \varphi_i^k(\bar{s}_i^*) \cdot \frac{\partial}{\partial s_i^{k'}} \left(\frac{\phi_i'(U_i(s_i, \bar{s}_i))}{\phi_i'(U_i(s^*))} \right) \Bigg|_{\sigma^*} = \frac{\partial \varphi_i^k}{\partial s_i^{k'}}(\sigma^*).$$

Therefore we have (A.7). J^* is assumed non-singular, thus $J_\psi(\sigma^*; 0)$ is non-singular. Consequently, by the implicit function theorem, there exists c and there exists a unique continuously differentiable function $\sigma : [0, \epsilon^{L5}) \rightarrow \prod_{i \in I} N_{\sigma_i^*}$ such that $\psi(\sigma(\epsilon), \epsilon) = 0$ for all $\epsilon \in [0, \epsilon^{L5})$, and $\sigma(0) = \sigma^*$.

(ii) For MMR: Recall that for $\sigma_i = (s_i^1, \dots, s_i^{\gamma_i-1})$,

$$\Phi_i^{MMR}(\sigma_i, \overline{P_i(s_i, \Upsilon_i)}) = \min_{Q_i} \left(\sum_{r_i=1}^{R_i} Q_i[r_i] \cdot U_i(\sigma_i, \bar{p}_i[r_i]) + c_i(Q_i) \right)$$

Consider $\sigma_i' = \sigma_i + t a_i^k = (s_i^1, \dots, t + s_i^k, \dots, s_i^{\gamma_i-1})$ for sufficiently small $t > 0$. Then,

$$\Phi_i^{MMR}(\sigma_i', \overline{P_i(s_i')}) = \min_{Q_i} \left(\sum_{r_i=1}^{R_i} Q_i[r_i] \cdot \left(U_i(\sigma_i, \bar{p}_i[r_i]) + t \varphi_i^k(\bar{p}_i[r_i]) \right) + c_i(Q_i) \right).$$

Thus,

$$\frac{\partial \Phi_i^{MMR}}{\partial s_i^k} = \lim_{t \rightarrow 0} \frac{\Phi_i^{MMR}(\sigma_i', \overline{P_i(s_i')}) - \Phi_i^{MMR}(\sigma_i, \overline{P_i(s_i, \Upsilon_i)})}{t} = \sum_{r_i=1}^{R_i} \tilde{Q}_i[r_i] \cdot \varphi_i^k(\bar{p}_i[r_i])$$

where \tilde{Q}_i is the distorted probability for σ_i . The above equation holds since μ_i (and therefore c_i) is strictly convex, and thus Φ_i^{MMR} is differentiable.

The remaining part is almost identical to the case of KMM model, but we need to show that for all $i \in I$, \tilde{Q}_i is continuously differentiable on a neighborhood of $(\sigma^*, 0)$ in $\prod_{i \in I} N_{\sigma_i^*} \times$

$[0, \bar{\epsilon})$. With an abuse of notation, we sometimes denote $(Q_i[1], \dots, Q_i[R_i - 1]) \in \Delta_c^{R_i-1}$ by Q_i , since it fully characterizes $(Q_i[1], \dots, Q_i[R_i]) \in \Delta^{R_i-1}$. Then, we can rewrite \tilde{Q}_i as

$$\begin{aligned} \left\{ \tilde{Q}_i \right\} &= \operatorname{argmin}_{Q_i \in \Delta_c^{R_i-1}} \left(\sum_{r_i=1}^{R_i-1} Q_i[r_i] \cdot \xi_i[r_i] + U_i(\sigma_i, \bar{p}_i[R_i]) + c_i(Q_i) \right) \\ &= \operatorname{argmin}_{Q_i \in \Delta_c^{R_i-1}} \left(\sum_{r_i=1}^{R_i-1} Q_i[r_i] \cdot \xi_i[r_i] + c_i(Q_i) \right) \end{aligned}$$

where $\xi_i[r_i] = U_i(\sigma_i, \bar{p}_i[r_i]) - U_i(\sigma_i, \bar{p}_i[R_i])$ for $r_i = 1, \dots, R_i - 1$. By the maximum theorem, \tilde{Q}_i is a continuous function in $\xi_i = (\xi_i[1], \dots, \xi_i[R_i - 1]) \in \mathbb{R}^{R_i-1}$. We now show that for ξ_i close enough to zero, \tilde{Q}_i is in the interior of simplex and moreover continuously differentiable in ξ_i . First, due to continuity and the fact that $\tilde{Q}_i(\xi_i = 0) = W_i \in \operatorname{int}(\Delta_c^{R_i-1})$, there exists an open ball $B(0; \delta) \subseteq \mathbb{R}^{R_i-1}$ such that for any $\xi_i \in B(0; \delta)$, $\tilde{Q}_i(\xi_i) \in \operatorname{int}(\Delta_c^{R_i-1})$. Then, for $\xi_i \in B(0; \delta)$, \tilde{Q}_i is a solution of the following first order condition with respect to Q_i :

$$\varsigma_i[r_i] \doteq \xi_i[r_i] + \theta_i \left(\omega_i[r_i] \mu_i' \left(\frac{Q_i[r_i]}{W_i[r_i]} \right) - \omega_i[R_i] \mu_i' \left(\frac{Q_i[R_i]}{W_i[R_i]} \right) \right) = 0 \text{ for } r_i = 1, \dots, R_i - 1. \quad (\text{A.8})$$

We denote $(\varsigma_i[1], \dots, \varsigma_i[R_i - 1])$ by ς_i . Then, $\varsigma_i : B(0; \delta) \times \Delta_c^{R_i-1} \rightarrow \mathbb{R}^{R_i-1}$ is a continuously differentiable function of (ξ_i, Q_i) , since μ_i is twice continuously differentiable. Observe that $\varsigma_i(0, W_i) = 0$ since $\mu_i'(1) = 0$. We also have that

$$\frac{\partial \varsigma_i}{\partial Q_i} = \begin{bmatrix} \beta_i[1] + \beta_i[R_i] & \beta_i[R_i] & \cdots & \beta_i[R_i] \\ \beta_i[R_i] & \beta_i[2] + \beta_i[R_i] & \cdots & \beta_i[R_i] \\ \vdots & \vdots & & \vdots \\ \beta_i[R_i] & \beta_i[R_i] & \cdots & \beta_i[R_i - 1] + \beta_i[R_i] \end{bmatrix}$$

where $\beta_i[r_i] = \theta_i \omega_i[r_i] \mu_i'' \left(\frac{Q_i[r_i]}{W_i[r_i]} \right) \frac{1}{W_i[r_i]}$ for $r_i = 1, \dots, R_i - 1$, which are strictly positive. Since $\frac{\partial \varsigma_i}{\partial Q_i}$ is positive definite, it is non-singular. Consequently, by the implicit function theorem, there exists an open ball $B(0; \delta') \subseteq B(0; \delta)$ such that for any $\xi_i \in B(0; \delta')$, the solution $\tilde{Q}_i(\xi_i)$ of the equation (A.8) is continuously differentiable in ξ_i .

Then, due to continuity of ξ_i in (σ, ϵ) and the fact that $\xi_i(\sigma^*, 0) = 0$, there exists an open ball $B((\sigma^*, 0); \delta_i) \subseteq \prod_{i \in I} N_{\sigma_i^*} \times [0, \bar{\epsilon})$ such that $(\sigma, \epsilon) \in B((\sigma^*, 0); \delta_i)$ implies $\xi_i \in B(0; \delta')$. Also, since ξ_i is continuously differentiable for (σ, ϵ) in $\prod_{i \in I} N_{\sigma_i^*} \times [0, \bar{\epsilon})$, \tilde{Q}_i is continuously differentiable for (σ, ϵ) in $B((\sigma^*, 0); \delta_i)$. Lastly, we take $\bar{\delta} = \min_{i \in I} \{\delta_i\}$. Then, for all $i \in I$, \tilde{Q}_i is continuously differentiable on an open set $U = B((\sigma^*, 0); \bar{\delta}) \subseteq \prod_{i \in I} N_{\sigma_i^*} \times [0, \bar{\epsilon})$. Next, we define $\psi_i^k : U \rightarrow \mathbb{R}$ by

$$\psi_i^k(\sigma, \epsilon) = \frac{\partial \Phi_i^{MMR}}{\partial s_i^k} = \sum_{r_i=1}^{R_i} \tilde{Q}_i[r_i] \cdot \varphi_i^k(\bar{p}_i[r_i](\epsilon)) \text{ for } i \in I; k = 1, \dots, \gamma_i - 1.$$

Let ψ denote the vector

$$\psi = \left(\psi_1^1, \dots, \psi_1^{\gamma_1-1}; \dots; \psi_i^1, \dots, \psi_i^{\gamma_i-1}; \dots; \psi_N^1, \dots, \psi_N^{\gamma_N-1} \right), i \in I.$$

Then $\psi : U \rightarrow \mathbb{R}^\gamma$ is a continuously differentiable function with $\psi(\sigma^*, 0) = 0$. we also have that

$$\frac{\partial \psi}{\partial \sigma}(\sigma^*, 0) = J^*,$$

since

$$\psi_i^k(\sigma, 0) = \sum_{r_i=1}^{R_i} W_i[r_i] \cdot \varphi_i^k(\bar{s}_i) = \varphi_i^k(\bar{s}_i) \text{ for } i \in I; k = 1, \dots, \gamma_i - 1.$$

Consequently, by the implicit function theorem, there exist $\epsilon^{L5} \in (0, \bar{\epsilon})$ and a unique continuously differentiable function $\sigma : [0, \epsilon^{L5}) \rightarrow \prod_{i \in I} N_{\sigma_i^*}$ such that $\psi(\sigma(\epsilon), \epsilon) = 0$ for all $\epsilon \in [0, \epsilon^{L5})$ and $\sigma(0) = \sigma^*$.

Proof of Theorem 2:

Suppose that for a given quasi-strict equilibrium s^* and given $\{\Upsilon(\epsilon)\}_{\epsilon \geq 0}$, we have found $\epsilon^{L5} > 0$ and a unique family $\{s(\epsilon)\}_{\epsilon \in (0, \epsilon^{L5})}$ of strategy profiles with $C(s(\epsilon)) = C(s^*)$ which satisfy the condition (i) and (ii) in Lemma 5. Then, there exists $\epsilon^{QS} \in (0, \epsilon^{L5})$ such that for any $\epsilon \in [0, \epsilon^{QS})$,

$$U_i(a_i^{\gamma_i}, \overline{p_i[r_i]}(s(\epsilon), \epsilon)) > U_i(a_i^k, \overline{p_i[r_i]}(s(\epsilon), \epsilon)) \text{ for all } i; \text{ all } r_i; k = \gamma_i + 1, \dots, K_i.$$

The reason is that

$$U_i(a_i^{\gamma_i}, \overline{s_i^*}) > U_i(a_i^k, \overline{s_i^*}) \text{ for all } i; k = \gamma_i + 1, \dots, K_i$$

and

$$\lim_{\epsilon \rightarrow 0} U_i(a_i^k, \overline{p_i[r_i]}(s(\epsilon), \epsilon)) = U_i(a_i^k, \overline{s_i^*}) \text{ for all } i; \text{ all } r_i; \text{ all } k.$$

Thus, for player $i \notin I$, $s_i(\epsilon) = a_i^1$ satisfies (2.5) by Lemma 3-(ii). For player $i \in I$, by Lemma 3-(i), the condition that

$$C(s_i(\epsilon)) \subseteq \{a_i^1, \dots, a_i^{\gamma_i}\}$$

is a necessary condition for $s_i(\epsilon)$ to satisfy (2.5). Thus, the condition (i) in Lemma 5 is the first order condition of (2.5). This first order condition is sufficeint due to concavity of Φ_i which is proven in Lemma 1.

Proof of Lemma 6:

(i) For KMM: $D_h^2 \left(\Phi_i(\sigma_i(\epsilon), \overline{P_i(s_i(\epsilon), \epsilon)}) \right) < 0$ for any $h \neq 0$ if and only if

$$h^T \cdot H_{\Phi_i}(s(\epsilon)) \cdot h < 0 \text{ for any } h \neq 0. \tag{A.9}$$

Due to the particular structure of Φ_i , we have

$$H_{\Phi_i}(s(\epsilon)) = \sum_{r_i=1}^{R_i} W_i[r_i] \cdot H_{\phi_i[r_i]}(s(\epsilon))$$

where $\phi_i[r_i] = \phi_i(U_i(\sigma_i, \overline{p_i[r_i]}(s(\epsilon), \epsilon))) = \phi_i\left(\sum_{k=1}^{\gamma_i-1} s_i^k \cdot \alpha_i^k[r_i] + U_i(a_i^{\gamma_i}, \overline{p_i[r_i]}(s(\epsilon), \epsilon))\right)$.

Thus,

$$\begin{aligned}
H_{\phi_i[r_i]}(s(\epsilon)) &= \begin{bmatrix} \frac{\partial^2 \phi_i[r_i]}{\partial (s_i^1)^2} & \cdots & \frac{\partial^2 \phi_i[r_i]}{\partial s_i^1 \partial s_i^{\gamma_i-1}} \\ \vdots & & \vdots \\ \frac{\partial^2 \phi_i[r_i]}{\partial s_i^{\gamma_i-1} \partial s_i^1} & \cdots & \frac{\partial^2 \phi_i[r_i]}{\partial (s_i^{\gamma_i-1})^2} \end{bmatrix}_{s(\epsilon)} \\
&= \begin{bmatrix} (\alpha_i^1[r_i])^2 & \cdots & \alpha_i^1[r_i] \alpha_i^{\gamma_i-1}[r_i] \\ \vdots & & \vdots \\ \alpha_i^{\gamma_i-1}[r_i] \alpha_i^1[r_i] & \cdots & (\alpha_i^{\gamma_i-1}[r_i])^2 \end{bmatrix} \cdot \phi_i''[r_i] \Big|_{s(\epsilon)}
\end{aligned}$$

where $\phi_i''[r_i] = \phi_i''(U_i(\sigma_i, \overline{p_i}[r_i](s(\epsilon), \epsilon)))$. And the condition (A.9) is equivalent to the condition that

$$\sum_{r_i=1}^{R_i} W_i[r_i] \cdot [h^T \cdot H_{\phi_i[r_i]}(s(\epsilon)) \cdot h] < 0 \text{ for any } h \neq 0. \quad (\text{A.10})$$

Let $h = (h^1, \dots, h^{\gamma_i-1})^T \neq 0$. Then,

$$\begin{aligned}
h^T \cdot H_{\phi_i[r_i]}(s(\epsilon)) \cdot h &= \left(\sum_{k=1}^{\gamma_i-1} \alpha_i^k[r_i] \cdot h^k \right)^2 \cdot \phi_i''[r_i] \Big|_{s(\epsilon)} \\
&= (\alpha_i[r_i]^T \cdot h)^2 \cdot \phi_i''[r_i] \Big|_{s(\epsilon)}
\end{aligned}$$

Consequently, we have

$$\sum_{r_i=1}^{R_i} W_i[r_i] \cdot [h^T \cdot H_{\phi_i[r_i]}(s(\epsilon)) \cdot h] = \sum_{r_i=1}^{R_i} W_i[r_i] \cdot (\alpha_i[r_i]^T \cdot h)^2 \cdot \phi_i''[r_i] \Big|_{s(\epsilon)}.$$

Since W_i is strictly positive and ϕ_i'' is strictly negative in any case, the condition (A.10) holds if and only if there is no $h \neq 0$ such that $\alpha_i[r_i]^T \cdot h = 0$ for all r_i , or equivalently, $\alpha_i^T \cdot h = 0$. Again, this condition is equivalent to the following condition: α_i has full row rank, $\gamma_i - 1$.

(ii) For MMR: For notational simplicity, we denote $\Phi_i^{MMR}(\sigma_i, \overline{P_i}(s_i, \Upsilon_i))$ for a given $\overline{P_i}(s_i, \Upsilon_i)$ by $V(\sigma_i)$,

$$V(\sigma_i) = \min_{Q_i} \left(\sum_{r_i=1}^{R_i} Q_i[r_i] \cdot U_i(\sigma_i, \bar{p}_i[r_i]) + c_i(Q_i) \right).$$

In the proof of Lemma 5, we obtained that

$$\frac{\partial \Phi_i^{MMR}}{\partial s_i^k} = \sum_{r_i=1}^{R_i} \tilde{Q}_i[r_i] \cdot \varphi_i^k(\bar{p}_i[r_i]) \quad (\text{A.11})$$

where \tilde{Q}_i is the distorted probability for σ_i . $\frac{\partial \Phi_i^{MMR}}{\partial s_i^k}$ is a directional derivative of V in the direction of e_k .² Thus, we can write (A.11) in the form

$$V'(\sigma_i; e_k) = \sum_{r_i=1}^{R_i} \tilde{Q}_i[r_i] \cdot \varphi_i^k(\bar{p}_i[r_i]).$$

Therefore we have that for any non-zero $h \in \mathbb{R}^{\gamma_i-1}$,

$$V'(\sigma_i; h) = \sum_{r_i=1}^{R_i} \tilde{Q}_i[r_i] \cdot \left(\varphi_i(\bar{p}_i[r_i])^T \cdot h \right)$$

where $\varphi_i(\bar{p}_i[r_i]) = \left(\varphi_i^1(\bar{p}_i[r_i]), \dots, \varphi_i^{\gamma_i-1}(\bar{p}_i[r_i]) \right)^T$. Given that $\overline{P_i(s_i, \Upsilon_i)} = \overline{P_i(s_i(\epsilon), \epsilon)}$,

$$V'(\sigma_i; h) = \sum_{r_i=1}^{R_i} \tilde{Q}_i[r_i] \cdot (\alpha_i[r_i]^T \cdot h). \quad (\text{A.12})$$

We have shown in Lemma 5 that

$$V'(\sigma_i(\epsilon); h) = \sum_{r_i=1}^{R_i} \tilde{W}_i[r_i] \cdot (\alpha_i[r_i]^T \cdot h) = 0 \text{ for any } h \neq 0$$

where \tilde{W}_i is the distorted probability for $\sigma_i(\epsilon)$.

We want to prove that

$$\lim_{t \rightarrow 0} \frac{V'(\sigma_i(\epsilon) + th; h) - V'(\sigma_i(\epsilon); h)}{t} < 0 \text{ for any } h \neq 0. \quad (\text{A.13})$$

Observe that by (A.12),

²We denote by e_k the k th standard unit vector of \mathbb{R}^{γ_i-1} .

$$V'(\sigma_i(\epsilon) + th; h) = \sum_{r_i=1}^{R_i} \tilde{W}'_i[r_i] \cdot (\alpha_i[r_i]^T \cdot h)$$

where \tilde{W}_i is the distorted probability for $\sigma_i + th$. Thus,

$$\frac{V'(\sigma_i(\epsilon) + th; h) - V'(\sigma_i(\epsilon); h)}{t} = \sum_{r_i=1}^{R_i} \left(\frac{\tilde{W}'_i[r_i] - \tilde{W}_i[r_i]}{t} \right) \cdot (\alpha_i[r_i]^T \cdot h).$$

To prove (A.13), we first have to show that the second directional derivative at $\sigma_i(\epsilon)$ exists. As we have shown in the proof of Lemma 5, for any $\epsilon \in (0, \epsilon^{QS})$, \tilde{W}_i is differentiable. Thus $\lim_{t \rightarrow 0} \frac{\tilde{W}'_i - \tilde{W}_i}{t}$ exists and the second directional derivative of V at $\sigma_i(\epsilon)$ exists. To be more specific, recall the equation (A.8) in the proof of Lemma 5. Both \tilde{W}_i and \tilde{W}'_i are solutions of this equation of Q_i where $\xi_i[r_i] = U_i(\sigma_i(\epsilon), \bar{p}_i[r_i]) - U_i(\sigma_i(\epsilon), \bar{p}_i[R_i])$ and $\xi_i[r_i] = U_i(\sigma_i(\epsilon) + th, \bar{p}_i[r_i]) - U_i(\sigma_i(\epsilon) + th, \bar{p}_i[R_i])$, respectively. Since $\frac{\partial \xi_i}{\partial Q_i}$ evaluated at \tilde{W}_i is non-singular, by the implicit function theorem, we have the following result:

$$\lim_{t \rightarrow 0} \frac{\tilde{W}'_i - \tilde{W}_i}{t} = - \left[\left(\frac{\partial \xi_i}{\partial Q_i} \right) \Big|_{\tilde{W}_i} \right]^{-1} \cdot (\alpha_i^T \cdot h) \quad (\text{A.14})$$

Consequently, the condition (A.13) is equivalent to the following condition: For any sufficiently small $t > 0$,

$$V'(\sigma_i(\epsilon) + th; h) - V'(\sigma_i(\epsilon); h) < 0 \text{ for any } h \neq 0$$

or equivalently,

$$\sum_{r_i=1}^{R_i} \left(\tilde{W}'_i[r_i] - \tilde{W}_i[r_i] \right) \cdot (\alpha_i[r_i]^T \cdot h) < 0 \text{ for any } h \neq 0. \quad (\text{A.15})$$

We will show two things:

- (a) $\sum_{r_i=1}^{R_i} \left(\tilde{W}'_i[r_i] - \tilde{W}_i[r_i] \right) \cdot (\alpha_i[r_i]^T \cdot h) \leq 0$ for any h .
- (b)

$$\sum_{r_i=1}^{R_i} \left(\tilde{W}'_i[r_i] - \tilde{W}_i[r_i] \right) \cdot (\alpha_i[r_i]^T \cdot h) = 0 \quad (\text{A.16})$$

if and only if $\alpha_i[r_i]^T \cdot h = 0$ for all r_i , or equivalently, $\alpha_i^T \cdot h = 0$.

To show (a), observe that

$$V(\sigma_i(\epsilon)) = \sum_{r_i=1}^{R_i} \tilde{W}_i[r_i] \cdot U_i(\sigma_i(\epsilon), \bar{p}_i[r_i]) + c_i(\tilde{W}_i) \leq \sum_{r_i=1}^{R_i} \tilde{W}'_i[r_i] \cdot U_i(\sigma_i(\epsilon), \bar{p}_i[r_i]) + c_i(\tilde{W}'_i). \quad (\text{A.17})$$

and that

$$\begin{aligned} V(\sigma_i(\epsilon) + th) &= \sum_{r_i=1}^{R_i} \tilde{W}'_i[r_i] \cdot (U_i(\sigma_i(\epsilon), \bar{p}_i[r_i]) + t\alpha_i[r_i]^T \cdot h) + c_i(\tilde{W}'_i) \\ &\leq \sum_{r_i=1}^{R_i} \tilde{W}_i[r_i] \cdot (U_i(\sigma_i(\epsilon), \bar{p}_i[r_i]) + t\alpha_i[r_i]^T \cdot h) + c_i(\tilde{W}_i). \end{aligned} \quad (\text{A.18})$$

Since $t > 0$, we obtain (a) by combining (A.17) and (A.18). Next, (A.16) holds if and only if both inequalities in (A.17) and (A.18) hold as equality if and only if $\tilde{W}'_i = \tilde{W}_i$ due to strict convexity of c_i . From (A.14), for any sufficiently small $t > 0$,

$$\tilde{W}'_i = \tilde{W}_i - \left[\left(\frac{\partial \varsigma_i}{\partial Q_i} \right) \Big|_{\tilde{W}_i} \right]^{-1} \cdot (\alpha_i^T \cdot h) \cdot t.$$

Thus, $\tilde{W}'_i = \tilde{W}_i$ if and only if $\alpha_i^T \cdot h = 0$. Consequently, the condition (A.15) holds if and only if the solution of the following equation of h , $\alpha_i^T \cdot h = 0$, is only zero.

Proof of Lemma 7:

Let $\alpha_{i,j}$ denote the aggregation matrix of $\alpha_i[R_i \setminus m_j]$, $m_j = 1, \dots, M_j - 1$:

$$\alpha_{i,j} = \left[\alpha_i[R_i \setminus 1] \quad \cdots \quad \alpha_i[R_i \setminus M_j - 1] \right].$$

Let

$$\begin{aligned} P &= \left[\alpha_{i,1} \quad \cdots \quad \alpha_{i,j} \quad \cdots \quad \alpha_{i,n} \quad \alpha_i(R_i) \right], j \neq i, j \in I, \text{ and} \\ Q &= \left[\hat{\alpha}_{i,1} \quad \cdots \quad \hat{\alpha}_{i,j} \quad \cdots \quad \hat{\alpha}_{i,n} \quad \alpha_i(R_i) \right] = \left[\hat{\alpha}_i \quad \alpha_i(R_i) \right], j \neq i, j \in I. \end{aligned}$$

We will show that $\text{rank}(\alpha_i) \geq \text{rank}(P) = \text{rank}(Q) \geq \text{rank}(\hat{\alpha}_i)$. The first inequality holds because we obtain P from α_i by selecting the particular columns of α_i . The last inequality also holds trivially. We prove the second equality. We first show that we can obtain $\begin{bmatrix} \hat{\alpha}_{i,j} & \alpha_i(R_i) \end{bmatrix}$ from $\begin{bmatrix} \alpha_{i,j} & \alpha_i(R_i) \end{bmatrix}$ by elementary column operations:

- (i) By $c_{M_j} \rightarrow \sum_{m_j=1}^{M_j} w_j[m_j]c_{m_j}$, $\begin{bmatrix} \alpha_{i,j} & \alpha_i(R_i) \end{bmatrix} \rightarrow \begin{bmatrix} \alpha_{i,j} & \bar{\alpha}_{i,j} \end{bmatrix}$.
- (ii) By $c_{m_j} \rightarrow c_{m_j} - c_{M_j}$ for $m_j = 1, \dots, M_j - 1$, $\begin{bmatrix} \alpha_{i,j} & \bar{\alpha}_{i,j} \end{bmatrix} \rightarrow \begin{bmatrix} \hat{\alpha}_{i,j} & \bar{\alpha}_{i,j} \end{bmatrix}$.
- (iii) By $c_{M_j} \rightarrow c_{M_j} - \frac{1}{w_j[M_j]} \sum_{m_j=1}^{M_j-1} w_j[m_j]c_{m_j}$, $\begin{bmatrix} \hat{\alpha}_{i,j} & \bar{\alpha}_{i,j} \end{bmatrix} \rightarrow \begin{bmatrix} \hat{\alpha}_{i,j} & \alpha_i(R_i) \end{bmatrix}$.

Therefore we can obtain Q from P by repeating the above elementary column operations.

Proof of Lemma 8:

Suppose that $m \times n$ matrix P^0 has rank $p \leq \min(m, n)$. Then we can make $p \times p$ matrix Q^0 by selecting particular rows and columns from P^0 so that Q^0 is non-singular, i.e. $\det(Q^0) \neq 0$. With the same selected rows and columns, we make $p \times p$ matrix $Q(\epsilon)$ from $P(\epsilon)$. It is obvious that $\lim_{\epsilon \rightarrow 0} Q(\epsilon) = Q^0$. Thus, there exists $\epsilon^0 > 0$ such that for all $\epsilon \in [0, \epsilon^0)$, $\det(Q(\epsilon)) \neq 0$. This is because the determinant of matrix is a continuous function of all elements of the matrix, and all elements of $Q(\epsilon)$ converges to the elements of Q^0 . Thus, for all $\epsilon \in [0, \epsilon^0)$, $\text{rank}(Q(\epsilon)) = p$. Clearly, $\text{rank}(P(\epsilon)) \geq \text{rank}(Q(\epsilon)) = p$.

Proof of Theorem 3:

First, for player $i \notin I$, it is an immediate result of Theorem 2 that (2.2) holds for any $\epsilon \in (0, \epsilon^{QS})$. Next, consider player $i \in I$. By Lemma 7, we aim to show that there exists $\epsilon^{QS*} \in (0, \epsilon^A)$ such that for any $\epsilon \in (0, \epsilon^{QS*})$, $\hat{\alpha}_i$ has full column rank, $\gamma_i - 1$, for all $i \in I$. The sketch of proof already has been shown in Examples 2 and 3. Recall that

$$\hat{\alpha}_i = \begin{bmatrix} \hat{\alpha}_{i,1} & \cdots & \hat{\alpha}_{i,j} & \cdots & \hat{\alpha}_{i,n} \end{bmatrix}, j \neq i, j \in I.$$

We will reformulate $\hat{\alpha}_{i,j}$ for each j . Since $\hat{\alpha}_{i,j} = \begin{bmatrix} \hat{\alpha}_i[R_i \setminus 1] & \cdots & \hat{\alpha}_i[R_i \setminus M_j - 1] \end{bmatrix}$, we first need to reformulate the m_j th column vector $\hat{\alpha}_i[R_i \setminus m_j]$. Consider the k th element of it.

$$\begin{aligned}
& \hat{\alpha}_i^k [R_i \setminus m_j] \\
&= \alpha_i^k [R_i \setminus m_j] - \bar{\alpha}_{i,j}^k \\
&= \left[\frac{\partial \varphi_i^k}{\partial s_j^1}, \dots, \frac{\partial \varphi_i^k}{\partial s_j^{\gamma_j-1}}, \frac{\partial \varphi_i^k}{\partial s_j^{\gamma_j+1}}, \dots, \frac{\partial \varphi_i^k}{\partial s_j^{K_i}} \right]_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \left(\hat{v}_j^1[m_j](\epsilon), \dots, \hat{v}_j^{\gamma_j-1}[m_j](\epsilon), \hat{v}_j^{\gamma_j+1}[m_j](\epsilon), \dots, \hat{v}_j^{K_j}[m_j](\epsilon) \right)^T \\
&= \epsilon \cdot J_{\varphi_i^k}(\sigma_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \left(\hat{\eta}_j^1[m_j], \dots, \hat{\eta}_j^{\gamma_j-1}[m_j] \right)^T \\
&\quad + J_{\varphi_i^k}(s_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \left(\hat{v}_j^1[m_j](\epsilon), \dots, \hat{v}_j^{\gamma_j-1}[m_j](\epsilon), \hat{v}_j^{\gamma_j+1}[m_j](\epsilon), \dots, \hat{v}_j^{K_j}[m_j](\epsilon) \right)^T \\
&= \epsilon \cdot J_{\varphi_i^k}(\sigma_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \hat{\eta}_j[m_j] + J_{\varphi_i^k}(s_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \hat{v}_j[m_j](\epsilon)
\end{aligned}$$

Thus,

$$\begin{aligned}
\hat{\alpha}_i [R_i \setminus m_j] &= \begin{bmatrix} \hat{\alpha}_i^1 [R_i \setminus m_j] \\ \vdots \\ \hat{\alpha}_i^{\gamma_i-1} [R_i \setminus m_j] \end{bmatrix} \\
&= \epsilon \cdot J_{\varphi_i}(\sigma_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \hat{\eta}_j[m_j] + J_{\varphi_i}(s_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \hat{v}_j[m_j](\epsilon).
\end{aligned}$$

Then,

$$\begin{aligned}
\hat{\alpha}_{i,j} &= \begin{bmatrix} \hat{\alpha}_i [R_i \setminus 1] & \cdots & \hat{\alpha}_i [R_i \setminus M_j - 1] \end{bmatrix} \\
&= \epsilon \cdot J_{\varphi_i}(\sigma_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \begin{bmatrix} \hat{\eta}_j[1] & \cdots & \hat{\eta}_j[M_j - 1] \end{bmatrix} + J_{\varphi_i}(s_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \begin{bmatrix} \hat{v}_j[1](\epsilon) & \cdots & \hat{v}_j[M_j - 1](\epsilon) \end{bmatrix} \\
&= \epsilon \cdot J_{\varphi_i}(\sigma_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \hat{\eta}_j + J_{\varphi_i}(s_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \hat{v}_j(\epsilon)
\end{aligned}$$

We can divide the equation by any non-zero ϵ , and we have

$$\frac{1}{\epsilon} \hat{\alpha}_{i,j} = J_{\varphi_i}(\sigma_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \hat{\eta}_j + R_{i,j}(\epsilon)$$

where $R_{i,j}(\epsilon) = J_{\varphi_i}(s_j) \Big|_{\bar{p}_i[R_i](s(\epsilon), \epsilon)} \cdot \frac{1}{\epsilon} \hat{v}_j(\epsilon)$. Consequently, we have that

$$\frac{1}{\epsilon} \hat{\alpha}_i = \begin{bmatrix} \frac{1}{\epsilon} \hat{\alpha}_{i,1} & \cdots & \frac{1}{\epsilon} \hat{\alpha}_{i,j} & \cdots & \frac{1}{\epsilon} \hat{\alpha}_{i,n} \end{bmatrix} = J_{\varphi_i}(\bar{\sigma}_i) \Big|_{\bar{p}_1[R_1](s(\epsilon), \epsilon)} \cdot \Lambda_i + R_i(\epsilon)$$

where

$$J_{\varphi_i}(\overline{\sigma_i}) = \begin{bmatrix} J_{\varphi_i}(\sigma_1) & \cdots & J_{\varphi_i}(\sigma_j) & \cdots & J_{\varphi_i}(\sigma_n) \end{bmatrix},$$

$$\Lambda_i = \begin{bmatrix} \hat{\eta}_1 & \cdots & O & \cdots & O \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ O & \cdots & \hat{\eta}_j & \cdots & O \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ O & \cdots & O & \cdots & \hat{\eta}_m \end{bmatrix}, j \neq i, j \in I, \text{ and}$$

$$R_i(\epsilon) = \begin{bmatrix} R_{i,1}(\epsilon) & \cdots & R_{i,j}(\epsilon) & \cdots & R_{i,n}(\epsilon) \end{bmatrix}, j \neq i, j \in I.$$

Then, we have that $\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \hat{\alpha}_i = J_i^* \cdot \Lambda_i$, because

$$\lim_{\epsilon \rightarrow 0} J_{\varphi_i}(\overline{\sigma_i})|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} = J_{\varphi_i}(\overline{\sigma_i})|_{s_i^*} = J_i^*,$$

$$\lim_{\epsilon \rightarrow 0} J_{\varphi_i}(s_j)|_{\overline{p_1}[R_1](s(\epsilon), \epsilon)} = J_{\varphi_i}(s_j)|_{s_i^*}$$

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \hat{\nu}_j(\epsilon) = 0, \text{ and thus } \lim_{\epsilon \rightarrow 0} R_{i,j}(\epsilon) = 0 \text{ and } \lim_{\epsilon \rightarrow 0} R_i(\epsilon) = 0.$$

Therefore, by Lemma 8, there exists $\epsilon^{(i)} \in (0, \epsilon^A)$ such for any $\epsilon \in (0, \epsilon^{(i)})$

$$\text{rank} \left(\frac{1}{\epsilon} \hat{\alpha}_i \right) \geq \text{rank} (J_i^* \cdot \Lambda_i).$$

Λ_i has full row rank since each $\hat{\eta}_j$ has full row rank. Also J_i^* has full row rank, $\gamma_i - 1$, since J^* is non-singular. To be specific, we can write

$$J_{\varphi}(\sigma) = \begin{bmatrix} J_{\varphi_1}(\sigma) \\ \vdots \\ J_{\varphi_i}(\sigma) \\ \vdots \\ J_{\varphi_n}(\sigma) \end{bmatrix}$$

where $i \in I$. Thus $J_{\varphi}(\sigma)$ is non-singular only if each $J_{\varphi_i}(\sigma)$ has full row rank. Also, $\text{rank}(J_{\varphi_i}(\sigma)) = \text{rank}(J_{\varphi_i}(\overline{\sigma_i}))$, since $J_{\varphi_i}(\sigma_i)$, which is the part of $J_{\varphi_i}(\sigma)$, is trivially a zero matrix. Thus J^* is non-singular only if each J_i^* has full row rank.

Consequently, for any $\epsilon \in (0, \epsilon^{(i)})$

$$\text{rank}(\hat{\alpha}_i) = \text{rank}\left(\frac{1}{\epsilon}\hat{\alpha}_i\right) = \gamma_i - 1. \quad (\text{A.19})$$

Finally, we define ϵ^{QS^*} by $\epsilon^{QS^*} = \min_{i \in I} \{\epsilon^{(i)}\}$. Then for any $\epsilon \in (0, \epsilon^{QS^*})$, (A.19) holds for all $i \in I$. Since $\epsilon^{QS^*} < \epsilon^{QS}$, for any $\epsilon \in (0, \epsilon^{QS^*})$, $s_i(\epsilon)$ is strictly preferred for all $i = 1, \dots, N$. Thus $s(\epsilon)$ is strictly stable. We complete the proof.