Essays in Economic Theory

Submitted by Hyun Chang Yi to the University of Exeter
as a dissertation for the degree of
Doctor of Philosophy in Economics
in March 2014

I certify that all material in this dissertation which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: ____

Abstract

This dissertation consists of three research papers on cheap talk game and satisficing behaviour. The first chapter examines the potential for communication via cheap talk between an expert and a decision maker whose type (preferences) is uncertain. The expert privately observes states for each type of the decision maker and wants to persuade the decision maker to choose an action in his favour by informing her of the states. The decision maker privately observes her type and chooses an action. An optimal action for the decision maker depends upon both her type and type-specific states. In equilibrium the expert can *always* inform the decision maker in the form of comparative statements and the decision maker also can partially reveal her type to the expert or public. The second and third chapters build a dynamic model of satisficing behaviour in which an agent's "expected" payoff is explicitly introduced, where this expectation is adaptively formed. If the agent receives a payoff above her satisficing level she continues with the current action, updating her valuation of the action. If she receives a payoff below her satisficing level and her valuation falls below her satisficing level she updates both her action and satisficing level. In the second chapter, we find that in the long run, all players satisfice. In individual decision problems, satisficing behaviour results in cautious, maximin choice and in normal form games like the Prisoner's Dilemma and Stag Hunt, they in the long run play either cooperative or defective outcomes conditional on past plays. In coordination games like the Battle of the Sexes, Choosing Sides and Common Interest, they in the long run coordinate on Pareto optimal outcomes. In the third chapter, we find that satisficing players in the long run play *subgame dominant* paths, which is a refinement of subgame perfection, and identify conditions with which they 'always cooperate' or 'fairly coordinate' in repeated Prisoner's Dilemma and Battle of the Sexes games, respectively, and truthfully communicate in sender-receiver games. Proofs and simulations are provided in appendices.

Acknowledgements

I would like to gratefully and sincerely thank Prof. Rajiv Sarin for his guidance, understanding and patience during my graduate studies at Texas A&M University, University of Birmingham and University of Exeter. He has provided environments within which I could focus on finding and pursuing my research agenda without other worries. In particular, he has helped me develop the satisficing behaviour model, which is joint work of our's, from the very early stage and understand what I am really doing. I would also like to thank Brit Grosskopf, Dieter Balkenborg, Indrajit Ray and Jaideep Roy for their encouragement and comments on my research.

And most importantly I would like to thank my wife EunHye Cha. Her support, encouragement, quiet patience and unwavering love were undeniably the bedrock upon which I could keep working on my research and complete this dissertation. I also thank my parents for their faith in me and encouraging me to be as ambitious as I am.

Table of Contents

List of Tables

List of Figures

Preface

Economic agents are often assumed to have sufficient knowledge about decision problems they face: They know how each action is associated with deterministic/stochastic outcomes and which outcomes are most preferred by them. They are also assumed to have common knowledge of rationality so that they can form precise beliefs about how their opponents would behave. With the help of these knowledges, agents can find optimal or equilibrium actions. This dissertation tries to answer how the agents would behave when those knowledges are not given.

The first chapter analyses cheap talk games in which information transmission between players is required to achieve efficient outcomes. Specifically, it tries to answer how critical the common knowledge of players' preferences is to informative communication. In literature, since the seminal work by Crawford and Sobel (1982), subsequent models have been built on that both the expert and decision maker's types (or preferences) are common knowledge, in particular that the decision maker's preferences are clearly understood by the expert. Given this condition, although the information and the decision making belong to two separate players, the common knowledge that their preferences are not far apart from each other makes it possible for the expert to tell the decision maker to choose an action favourable to the expert as well as the decision maker. Here we try to identify what essentially makes possible for them to communicate in equilibrium by examining whether and how the expert and decision maker communicate each other when the decision maker's preference is not known to the expert.

This equilibrium analysis heavily depends on the "common" knowledge of the environment and rationality. Rational players can calculate which action would be chosen by rational counterparts given the belief that both have the common knowledge. However, the information and cognitive capacity for processing it are limited, and insuring the common knowledge held by all agents is normally too costly to achieve in the real world outside researchers' minds and laboratories.

The second research agenda is how people would behave when they have minimum knowledge of their decision environments: They know only what actions are available and acquire information only through experiences of repeated decision making and realised payoffs. They do not know how actions result in outcomes and even whether they have opponents in case of games. Furthermore, they do not know the nature of the decision problem, whether they are stationary or not. Given these conditions, it could be optimal to take a practical strategy with which agents do not try to find a best action, but guarantee a "good enough" payoff. Satisficing, which was first introduced in economics by Herbert A. Simon, refers to the decision making strategy that attempts to meet some acceptability threshold.

In this regard, the second chapter proposes a dynamic model of satisficing in which the state of an agent, in any period, is given by the action she chooses, her valuation of that action and satisficing level. It is the valuation of the action that depends on the past payoffs it has received. The satisficing level, is thought of as the payoff the agent finds satisfactory. What is satisfactory, in turn, depends on what the agent thinks is the payoff she might get from best outside option. This is adjusted whenever the agent receives any information on this. If the agent's valuation of the current action is above her satisficing level she chooses it again. If it falls below the satisficing level she moves away from it, where her probability of shifting depends upon the amount by which her valuation falls below the satisficing level.

In the second and third chapters, we apply the satisficing behaviour into various decision environments and find that satisficing players in the long run choose maximin options in individual decision problems, play either cooperate only or defect only profiles conditional on past plays in the Prisoner's Dilemma and Stag Hunt games, coordinate on Pareto optimal profiles in the Battle of the Sexes and Common Interest games, and develop languages through which they communicate each other in signalling games.

CHAPTER 1

Persuading Someone You Do Not Know

A lobbyist tries to persuade a government officer to approve a government spending. If their interests coincide, the lobbyist will tell how much is required and the officer will approve as told without doubt. Unfortunately, however, it is usually not the case: The lobbyist prefers more spending to less whereas optimal spending size for the officer differs by her political position as she faces a budget constraint. In addition, the lobbyist and officer both have their own private information. The lobbyist is *expert* on the spending: He knows optimal amounts of spending for all possible positions of the officer, but does not know the officer's position. On the other hand, the officer is *decision maker*: She makes a decision on the spending, but does not know the optimal amount for her.

Ironically, information about the officer's position (or type) does not help the lobbyist. If the officer's position becomes common knowledge,[1] it becomes clear to both that their preferences are totally misaligned. This sort of dilemma between information and communication is not peculiar to the lobbyist and officer. Often, sellers advertise their product without knowing which feature is most valued by potential buyers. Once a buyer tells her most favorate feature, she must not expect any honest explanation about that.

This chapter examines possibilities for informative communication through cheap talk in these circumstances: We analyse two-sided incomplete information cheap talk games between the expert and decision maker in one-way and two-way protocols. When the decision maker remains silent about her private information, the game is one-way and proceeds as follows. First, the expert privately observes states for each type of the decision maker while the decision maker privately observes her own type.[2] Second, the expert sends a costless, non-verifiable message about the states

[1] We assume that there is no way for the lobbyist to identify the officer's preferences privately.

[2] The states for different types of the decision maker is different from the standard term *state of the world* in literature. The state of the world resolves all uncertainty, but, in the current setup, the type specific states and the decision maker's type together constitute the state of the world.

to the decision maker. Third, the decision maker updates her belief about the state for her type, or her state, and chooses an optimal action. The action determines the expert's payoff which does not depend on the realised states and decision maker's type.[3] A two-way game is a one-way game augmented with the decision maker's pre-communication: After players observe their private information, the decision maker talks before the expert does.

Since the seminal work by Crawford and Sobel (1982), most subsequent models in literature have been built upon the assumption that both the expert and decision maker's preferences are common knowledge, in particular that the decision maker's preferences are clearly understood by the expert. Although the information and decision making belong to two separate players, the common knowledge that their preferences are not far apart from each other makes it possible for the expert to convince the decision maker to choose an action favourable to the expert as well as the decision maker.

An exception is the work of Seidmann (1990). He finds an example in which informative communication takes place when the expert's payoff depends only upon the decision maker's action and the expert does not know how the decision maker's payoff is determined. The present chapter generalises this finding. Furthermore, we analyse pre-play communication from the decision maker to the expert. Other related models are as follows: Two-sided incomplete information in Watson (1996) and Chen (2009), Lai (2010), Moreno de Barreda (2010), Ishida and Shimizu (2011); uncertain experts' preferences in Sobel (1985), Morris (2001), Morgan and Stocken (2003), Dimitrakas and Sarafidis (2005), Ottaviani and Sørensen (2006b), Li and Madarasz (2008); and multi-dimensional state space in Battaglini (2002), Chakraborty and Harbaugh (2007), Chakraborty and Harbaugh (2010).

Watson (1996), Chen (2009), Lai (2010), Moreno de Barreda (2010), and Ishida and Shimizu (2011) assume that decision makers also have partial information about the state of the world while their types are common knowledge.[4] Chen (2009) examines two-way communication where the decision maker talks before the expert does, and finds that the decision maker cannot credibly reveal her information.

[3]The expert is said to have *state-independent* preferences.

[4]In their studies, since the expert does not know what the decision maker observes, partial information functions as private information.

Watson (1996) examines a two-sided incomplete information game where the expert has state-independent preferences and the decision maker has partial information of the state. He shows that partial information facilitates informative communication.[5]

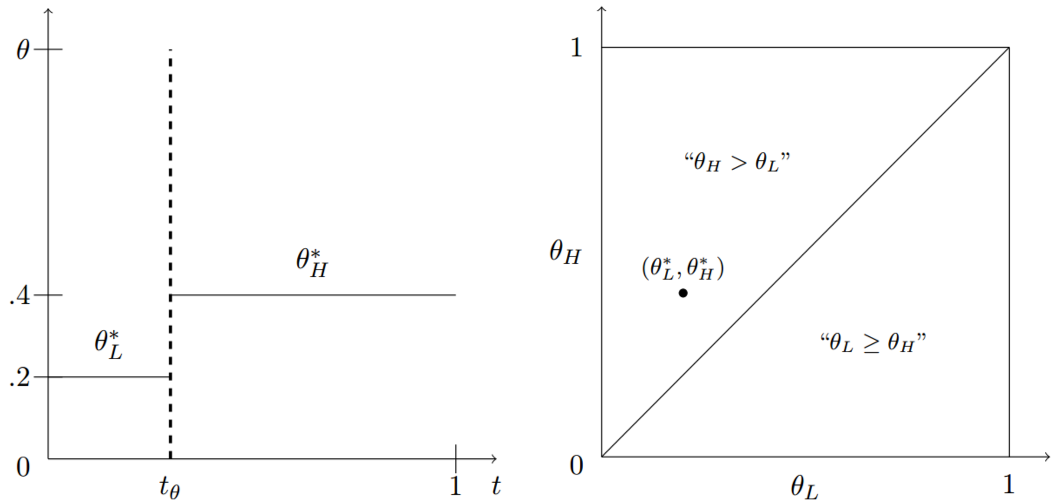
Morgan and Stocken (2003), Dimitrakas and Sarafidis (2005) and Li and Madarasz (2008) show that uncertainty over experts' preferences facilitates informative communication. Morgan and Stocken (2003) find that uncertainty can improve communication compared to cases where the expert has a known, intermediate bias. Dimitrakas and Sarafidis (2005) and Li and Madarasz (2008) show that revelation of the expert's bias can harm communication when the size of the possible bias is uncertain. These results are based on that the expert's preferences are uncertain, but have a similar implication that incomplete information helps players communicate.

In Chakraborty and Harbaugh (2007) and Chakraborty and Harbaugh (2010), an expert observes multiple issues or dimensions of an issue and reports complete or partial rankings of them to a decision maker. Chakraborty and Harbaugh (2010) find that combination of multi-dimensional information and state-independent preferences is sufficient for informative communication through comparative messages, upon which the equilibrium strategy in the one-way game of this chapter is partially based. Because of the type-specific property of his private information, the lobbyist's messages consist of multi-dimensional states for types *and* how states and types are associated. Chakraborty and Harbaugh (2010) address the first part of this.

We find, in both one-way and two-way games, players can communicate. When only the expert talks, he can *always* send an informative message and both the expert and decision maker benefit from the communication for most types of the decision maker. And, when the decision maker talks first, she can *partially* reveal her type without information loss. Furthermore, how the decision maker reveals her type depends on how the expert informs the decision maker.

This chapter is organised as follows. Section 1 explains how communication takes place in the preliminary lobbyist-officer example. Section 2 formally presents a one-way cheap talk game. Section 3 characterises how the expert informs the decision maker in one-way game. Section 4 presents two-way game and constructs

[5]The same feature arises in Ottaviani and Sørensen (2006b), where the *ex post* realised state which is publicly observed plays the role of this partial information obtained by the decision maker. Ottaviani and Sørensen (2006a) briefly discuss this result.



(a) Realised Optimal Spending Profile $(\theta_L^*, \theta_H^*, t_\theta)$ (b) Space of Optimal Spendings

Figure 1.1. Representation of State and Message

equilibrium type-revealing strategies of the decision maker. Section 5 discusses some future works that this study suggests. The Appendix A.1 and B contain proofs and examples of equilibrium strategies.

1.1. The Preliminary Example

In this section, we explain how communication takes place in a simple environment of government spending. The officer's position is represented in the unit interval $[0, 1]$, and optimal amounts of spending for each position are summarised by two numbers θ_L and θ_H and a boundary position t_θ: If the officer's position, say t, is lower (resp. higher) than t_θ, her optimal amount is θ_L (resp. θ_H). The optimal spending profile is represented by a step function

$$\theta(t) = \begin{cases} \theta_H & \text{if } t \geq t_\theta \\ \theta_L & \text{if } t < t_\theta, \end{cases}$$

which is depicted in Figure 1.1. (a). The officer and lobbyist privately observe t and $(\theta_L, \theta_H, t_\theta)$, respectively. All those variables are independently drawn from the common uniform distribution on $[0, 1]$. The lobbyist's payoff is given as the square of the spending approved by the officer.

The lobbyist's strategy consists of messages on two optimal amounts (θ_H, θ_L) and how the two different amounts apply to the officer's types (t_θ). For simplicity, for the moment we assume the expert talks nothing on t_θ, or his message on it is given as $t_\theta \in [0, 1]$.

Consider some messages that the lobbyist could send. First, if the lobbyist sends meaningless messages to the officer, his expected payoff is 1/4 because the unconditional expectation of optimal spending for all types are the same as 1/2. Second, if the lobbyist sends a message that maps each position to a specific value like "the optimal amounts are $\widehat{\theta}_H$ for the right half of positions and $\widehat{\theta}_L$ for the left half." This sort of messages cannot be credible because, if they are, the lobbyist always would say "optimal amounts for all positions are 1" regardless of realised profiles, and in turn this incentive to maximise the lobbyist's payoff makes the officer doubt any message from the lobbyist.

Now, consider two comparative messages: "optimal amounts for *right* political positions are higher than for *left* positions" and "optimal amounts for *right* political positions are lower than for *left* positions" without specifying how political positions are classified into *right* and *left* positions. These messages can be visualised as the upper and lower triangles in the state space of optimal amounts in Figure 1.1. (b). Given $\mathbb{E}[(\theta_L, \theta_H)|\theta_L \geq \theta_H] = (2/3, 1/3)$ and t_θ drawn from $[0, 1]$, the officer's estimate of her optimal amount given the first message is summarised as a function of her type, $\mathbb{E}[\theta(t)|\theta_L \geq \theta_H] = 2(1-t)/3 + t/3$, and the lobbyist's expected payoff is $\mathbb{E}[\mathbb{E}[\theta(t)|\theta_L \geq \theta_H]^2] = \int_0^1 (2(1-t)/3 + t/3)^2 dt$, where the outer expectation is taken over the officer's type t on $[0, 1]$ and the value of the integral is 7/27. Similarly, when the second message is sent, the estimate is $(1-t)/3 + 2t/3$ and the lobbyist's payoff is the same as with the first message. The lobbyist is indifferent between these two messages as long as the officer believes both to be true.

Therefore, it is incentive-compatible for the lobbyist to send a *true* message between the two messages "$\theta_L \geq \theta_H$" and "$\theta_L < \theta_H$," and the officer would believe any message sent by this strategy.[6] Furthermore, this communication makes the lobbyist's expected payoff strictly increase. And, if the officer's type is not 1/2, her estimate is different from the unconditional estimate, that is, she is informed by the lobbyist's message. This result is formulated in Sections 2 and 3.

[6]The term "true" refers to that the sent message is consistent with the realised state, not that the message is equivalent to the state.

1.2. The Model

This section generalises the lobbyist-officer example. We formally describe sources of uncertainty and how the lobbyist sends messages to the officer and how the officer responds to messages.

We have two players, an expert and a decision maker, and two sources of uncertainty, the decision maker's type and type-specific states, or simply state profile. The decision maker's type is represented as $t \in T \equiv [0,1]$ and the state profile as a step function $\theta(t) = \sum_{i=1}^{N} \theta_i \mathbf{1}_{T_i}(t)$ on T with $N \geq 2$, where $T_1 = [0, t_1)$, $T_2 = [t_1, t_2), \ldots, T_N = [t_{N-1}, 1]$ and $\mathbf{1}_{T_i}(t)$ is the indicator function defined on T_i for $i = 1, \ldots, N$.[7] Nature chooses t, $\theta = (\theta_1, \ldots, \theta_N)$ and $\bar{t} = (t_1, \ldots, t_{N-1})$ according to differentiable distributions G, F and H, respectively. F has full support on $\Theta \equiv [-1,1]^N$. $\bar{t} \in \overline{T} \subset [0,1]^{N-1}$ is an $(N-1)$-tuple of ordered statistics obtained from a common uni-variable distribution H that has full support on T and t_i is the ith largest of $N-1$ random variables that are identically and independently drawn from H. Let H_i denote the marginal distribution of t_i. The decision maker privately observes t while the expert privately observes θ and $\bar{t}$. Both players' common prior belief is summarised by G, H and F.

After observation, the expert informs the decision maker about a realised state profile $(\theta, \bar{t}) \in \Theta \times \overline{T}$ by sending a costless, non-verifiable message m, which is chosen by his informing strategy $M : \Theta \times \overline{T} \longrightarrow \Theta \times \overline{T}$. Let Φ denote the state space $\Theta \times \overline{T}$. Suppose the expert strategy is represented by $\{\Phi^j\}_j$, a partition of Φ. If the partition consists of one element, the state space itself, the strategy is babbling. Given such a strategy, if the expert observes a state profile $\phi \in \Phi^j$ for some $j \geq 1$, he randomly draws a state $\phi' \in \Phi^j$ and sends it to the decision maker. With a little abuse of notation, let M denote a message set $\{m^j\}_j$ that has ono-to-one correspondence to the partition. Then, this strategy is equivalent to an abstract strategy that the expert sends a message m^j from the message set M given that m^j is well understood

[7] Many decisions made by economic agents result in discrete outcomes across individuals' types. A few examples include second-degree price discrimination, tier rankings by credit rating agencies and discriminatory subsidies based on firm size or household income. And, if the effects are inherently continuous, they might be approximated by a step function.

to imply that the realised profile belongs to Φ^j.[8] Then, a function $m : \Phi \longrightarrow M$ such that $m^{-1}(m^j) = \Phi^j$ for any j fully describes the strategy generated by $\{\Phi^j\}_j$.

The expert has direct preferences over the decision maker's estimate of her state.[9] His *ex post* payoff u is a convex (or concave) and monotone function of the decision maker's estimate. Since the expert does not know the decision maker's type, he chooses m from M to maximise his expected payoff, $\int_T u(\mathbb{E}[\theta(t)|m])dG(t)$, where $\mathbb{E}[\theta(t)|m]$ is the decision maker's updated estimate given the expert's strategy M and sent message m while $\mathbb{E}[\theta(t)]$ denotes the decision maker's unconditional estimate of the state. The payoff function is common knowledge.

Until we discuss two-way games in which the decision maker also talks, we postpone specifying the decision maker's payoff because it is the expert's strategy that characterises equilibrium outcomes in one-way games. Instead, we define *informative* strategy based on if it makes the decision maker update her *prior* belief in (perfect Bayesian) equilibrium.

Definition 1. *A strategy M is* informative *to type t if* $\mathbb{E}[\theta(t)] \neq \mathbb{E}[\theta(t)|m]$ *for some* $m \in M$.

Accordingly, messages sent by informative strategies are informative messages and, if an equilibrium strategy M is informative to a non-negligible set of types with respect to G, the strategy is said to constitute an informative equilibrium. A babbling equilibrium always exists in which the expert sends arbitrary messages and the decision maker ignores them.

1.3. One-way Communication

In one-way games, only the expert talks to the decision maker about what he observes, how the decision maker's types are classified into distinctive groups and

[8]An expert in Crawford and Sobel (1982) uses a mixed strategy described here. If a realised state belongs to an element of an equilibrium partition of state space, the expert draws and sends an arbitrary state from the element according to a predetermined probability distribution. In perfect Bayesian equilibrium, since the decision maker understands the expert's whole strategy except for the realised state, the sent message or state implies that a realised state belongs to the same element of the equilibrium partition as the sent state. Thus, the equilibrium partition fully describes the equilibrium strategy.

[9]In cheap talk literature, decision makers are assumed to take estimates of the true states as optimal actions. Thus, it is usual that experts have direct preferences over estimates. For discussion of cases where decision makers' estimates and actions are different, see Chen and Olszewski (2011).

states for each group. We find the expert cannot at the same time truthfully announce how types are grouped *and* send informative messages. And, if the expert is silent about types, he can always send informative messages.

A strategy that is truthful on types must be one-to-one on $\overline{T}$. For simplicity, we assume if it is truthful on types, the strategy is identity function on $\overline{T}$.

Definition 2. *A strategy M is truthful on types if $M(\theta, \bar{t}) = (M_\theta(\theta), \bar{t})$ with some function $M_\theta(\theta) : \Theta \longrightarrow \Theta$ for any $(\theta, \bar{t}) \in \Theta \times \overline{T}$.*

A truthful strategy tells which states apply to each type of the decision maker. Thus, as long as the decision maker is receiving the same message on states, she can make a more precise estimate about her state from a truthful strategy. However, the first result of this section says that an informative strategy cannot be truthful on types.

Proposition 1. *If the expert's payoff is strictly convex (or concave), his strategy is either truthful on types or informative.*

If the expert is truthful on types, his expected payoff is bound to be the same as the payoff when there is no communication: If most types happen to belong to a single group,[10] the decision maker's type is practically identified by the expert, thus the expert cannot send informative messages without increasing or decreasing his expected payoff.

If the expert is babbling on Θ, he can be truthful on types without any conflict of interests, but this communication does not benefit either of players. And, if the expert is truthful on types, his strategy cannot be informative. However, once the expert becomes less informative on types, he can find a way to send informative messages to the decision maker.

Proposition 2. *An informative equilibrium always exists. In any informative equilibrium, (1) if the expert's preferences are convex (concave), he is* ex ante *better (worse) off than without communication, and (2) the communication is informative to almost all types of the decision maker.*

[10]For example, in the preliminary example, it is when t_θ is very close to 0 or 1.

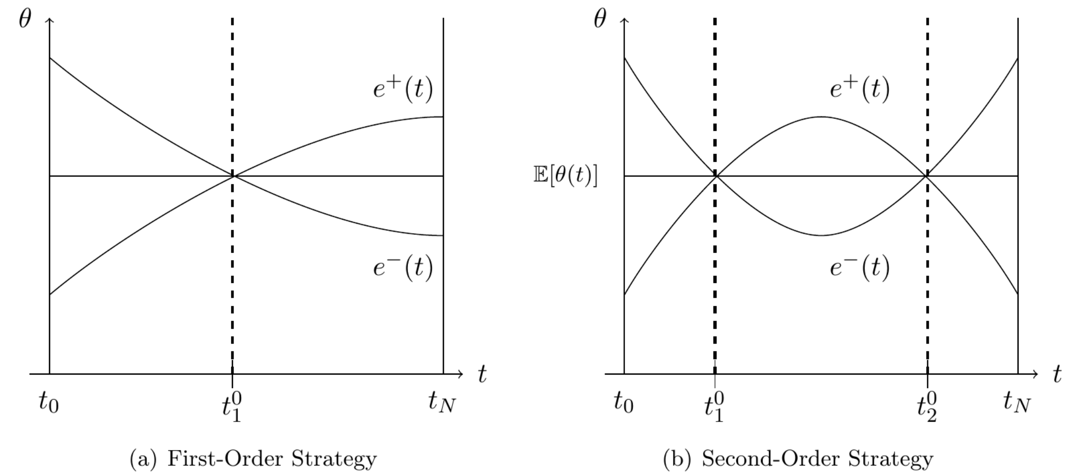


Figure 1.2. Decision Maker's Estimates in Equilibrium under Different Strategies

Figure 1.2 illustrates two equilibria in terms of the decision maker's estimates $e^+(t) = \mathbb{E}[\theta(t)|m^+]$ and $e^-(t) = \mathbb{E}[\theta(t)|m^-]$ induced by a message set $M = \{m^+, m^-\}$ when $\mathbb{E}[\theta(t)]$ is constant for all types. As long as the expert's expected payoffs from the two estimates are the same, it is incentive compatible for the expert to truthfully choose a message from M, thus the decision maker needs not doubt the message. In both equilibria, two messages m^+ and m^- are constructed by partitioning a space of states Θ into two convex subsets. The proof shows that for any priors we can find a pair of messages such that the expert is indifferent between them.

Though the equilibrium strategies are babbling on $\overline{T}$, it is not always necessary. In the preliminary example with the uniform priors, the expert can send informative messages to the decision maker notifying whether the boundary type t_θ is lower than 1/2 or not. For both cases of $t_\theta < 1/2$ and $t_\theta \geq 1/2$, the expert can construct equilibrium strategies for each such that his expected payoffs from both cases are the same.

Note that the decision maker's types are divided into two groups according to whether its informed estimate $\mathbb{E}[\theta(t)|m]$ is higher or not than its unconditional estimate $\mathbb{E}[\theta(t)]$. Accordingly, we define two sets of types $T_m^+ \equiv \{t \in T|\mathbb{E}[\theta(t)|m] > \mathbb{E}[\theta(t)]\}$ and $T_m^- \equiv \{t \in T|\mathbb{E}[\theta(t)|m] < \mathbb{E}[\theta(t)]\}$ for a message m from M. If the decision maker whose type is in T_m^+ receives message m and u is increasing, her estimate is updated more favourably to the expert while for types in T_m^-, her estimate is updated less favourably.

As the proof of Proposition 2 shows, there are only finite number of types for which the decision maker is not informed by the expert's message.[11] Actually, for an equilibrium strategy M, there exist at least one type through which $\mathbb{E}[\theta(t)|m]$ crosses over the uninformed estimate $\mathbb{E}[\theta(t)]$. Given M, the set of those types is denoted by $T_M \equiv \mathrm{cl}(T_m^+) \cap \mathrm{cl}(T_m^-)$, where $\mathrm{cl}(T_m^+)$ is the smallest closed set containing T_m^+, for some $m \in M$. If a type belongs to this set, the strategy is not informative to this type. This set, in particular the size of the set, characterises the expert's strategy.

Definition 3. *A strategy M is* first-order *if T_M has only one element, and* second-order *if T_M has exactly two elements.*

A first-order strategy divides the decision maker's types into two groups as $\{t|t < t_1^0\}$ and $\{t|t > t_1^0\}$ in Figure 1.2. (a). Given a message from this strategy, the direction to which the decision maker's estimate is updated depends on whether her type is to the left of the uninformed type or to the right. And, if a message is sent from a second-order strategy, how the estimate is updated depends on whether her type is in the centre, $\{t|t_1^0 < t < t_2^0\}$ in Figure 1.2. (b).

If state profiles are determined in a symmetric way with respect to types, either of first- and second-order strategies can be employed as an equilibrium strategy.

Proposition 3. *If F is invariant under any permutation of $\theta_1, \ldots, \theta_N$ and $N \geq 3$, the expert can employ either of first-order or second-order strategies.*

In the lobbyist-officer example, this result implies that, no matter how complicated the optimal spending profile is, the lobbyist can persuade the officer by truthfully stating either "the optimal amounts of spending for right positions are higher than for left positions" or "the optimal amounts for centre positions are higher than for others."

Lastly, it is noteworthy that equilibrium strategies developed in Proposition 2 convey very coarse information about state profiles: In equilibrium, sent messages merely tell which element of two subsets in a partition contains a realised state. This observation raises the question whether we can find a finer partition of the

[11] For example, in Figure 1.2. (b), type t_1^0 and t_2^0 decision makers cannot make informed estimates given the expert's strategy and in Figure 1.2. (a), type t_1^0 decision maker cannot make informed estimate.

state space. If possible, the partition could constitute an equilibrium strategy. The following result shows that this is the case.

Proposition 4. *If A 2^k-message informative cheap talk equilibrium always exists for every $k \geq 1$. If the expert's preferences are convex (concave), his* ex ante *payoff increases (decreases) with k.*[12]

Proposition 4 implies that, from any 2-message equilibrium strategy, we can find an infinite sequence of equilibrium strategies with increasing number of messages, and the expert's expected payoff increases (decreases) with the number of messages if u is convex (concave).

1.4. Two-way Communication

Now, we examine whether and how the decision maker reveals her type to the expert or the public in two-way games. For this purpose, first we specify the decision maker's payoff given a strategy of the expert.

The decision maker is assumed to be rewarded for her estimate's accuracy, $-(\theta(t) - e(t))^2$, and maximise her expected payoff. The farther her realised state is from her estimate, the lower her payoff is. Note that the payoff depends only upon the expert's strategy because the decision maker mechanically updates her belief by Bayes' Rule. Since revelation of private information could change the expert's strategy, we need to define the decision maker's payoff as a function of the expert's strategy. The decision maker's interim expected payoff is $\mathbb{E}[-(\theta(t) - e(t))^2|m] = \mathbb{E}[-(\theta(t) - E[\theta(t)|m])^2|m] = -\text{Var}[\theta(t)|m]$ and expected payoff is $\mathbb{E}[-\text{Var}[\theta(t)|m]]$.

We assume that the expert uses 2-message strategies constructed in Proposition 2. Let $M = \{m^+, m^-\}$ be an equilibrium strategy in a one-way game and $\sigma^2(t)$ denote the unconditional variance of $\theta(t)$ and $\mu(\Theta^i) \equiv \Pr(\theta \in \Theta^i)$ given M for $i = +, -$. Without loss of generality, let $\mathbb{E}[\theta_i] = 0$ for all $i = 1, \ldots, N$. Then, the

[12]This result is comparable to Theorem 4 in Chakraborty and Harbaugh (2010). We can construct a 2^2-message equilibrium strategy from a 2-message strategy as follows. Given a 2-message strategy that is induced by a 2-element partition $\{\Phi^1, \Phi^2\}$, we can get four estimates, $e^{1+}(t)$ and $e^{1-}(t)$ from Φ^1 and $e^{2+}(t)$ and $e^{2-}(t)$ from Φ^2 such that $\mathbb{E}[u(e^{1+})] = \mathbb{E}[u(e^{1-})]$ and $\mathbb{E}[u(e^{2+})] = \mathbb{E}[u(e^{2-})]$. If the expert is indifferent to all four estimates, we have found a 2^2-message equilibrium strategy. If not, we can make the messages that are more favourable to the expert a bit noisier so that the resulting estimates become less favourable and the expert becomes indifferent to all four estimates. Repeating this, we can construct an equilibrium strategy with an arbitrary number of messages.

decision maker's expected payoff conditional on the expert's strategy M is increasing in the square of her estimate $e(t)$.

Lemma 1. *$\mathbb{E}[-Var[\theta(t)|m]] = \mu(\Theta^i)e^i(t)^2/\mu(\Theta^j) - \sigma^2(t)$ for $i \neq j$.*

Suppose that the decision maker can reveal her type in a particular way or keep it private. Since the revelation could change the expert's prior belief about the decision maker's type, the expert's equilibrium strategy also might change according to his modified belief and send different messages. If the revelation changes the expert's equilibrium strategy such that the decision maker's estimate moves closer to the uninformed for some type, the decision maker of this type will choose not to reveal it. On the other hand, as long as her estimate remains the same for all types, she might be willing to reveal her type.

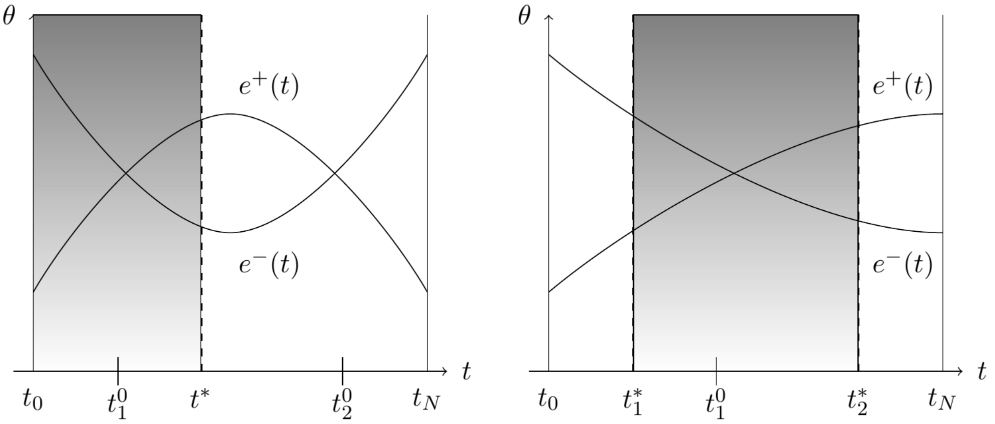
We formulate the decision maker's strategy as for the expert. The decision maker reveals her type through a partition of T, say, $\{T^k\}_k$. If her realised type t is in T^k, she sends a message n^k from a message set $N = \{n^k\}_k$, where n^k is understood to imply that a realised type belongs to T^k. Then, a function $n : T \longrightarrow N$ such that $n^{-1}(n^k) = T^k$ for any k fully describes a strategy of the decision maker. For simplicity, we denote the strategy and a message by N and n^k (or n), respectively.

We also characterise the decision maker's strategy in a similar way to the expert's strategy. Suppose that the decision maker reveals her type through a partition of size two, $N = \{T^1, T^2\}$.

Definition 4. *A strategy N of the decision maker is* first-order *if T^1 and T^2 are two intervals, and* second-order *if only T^1 is an interval.*

Given N and M, a two-way game proceeds as follows. First, the decision maker and expert privately observe her type and a state profile, respectively. Second, the decision maker sends a message about her type according to N. Third, after the revelation, the one-way game follows: The expert informs the decision maker according to M, the decision maker estimates her state and the estimate finalises both player's payoffs.

In a two-way game, if a decision maker completely reveals her type, no communication can occur in a subsequent one-way game. However, if only partial information is revealed so that the decision maker's type is still uncertain, there exists room for



(a) First-Order Revelation Strategy

(b) Second-Order Revelation Strategy

Figure 1.3. Revelation Strategies

the expert to give informative messages to the decision maker. We define a type-revealing equilibrium and show that such equilibrium always exists. Suppose that a two-way game is played by strategies N and M of the decision maker and expert.

Definition 5. *An equilibrium of a two-way game is type-revealing if N contains more than one non-trivial, proper subsets of T.*

If first- or second-order strategies of the decision maker constitute an equilibrium, such equilibrium is type-revealing.

Proposition 5. *For any informative equilibrium in one-way games, there exists a type-revealing equilibrium in the two-way game in which the decision maker partially reveals her type.*

Since there always exists an informative equilibrium by Proposition 2 for any priors, a type-revealing equilibrium also always exists.

Figure 1.3. (a) shows an equilibrium strategy in a two-way game. The decision maker's estimates $e^+(t)$ and $e^-(t)$ are derived by an equilibrium strategy M in a one-way game. Given M, we can find $t^* \in T$ such that $\mathbb{E}[u(e^+(t))|t \in [t_0, t^*)] = \mathbb{E}[u(e^-(t))|t \in [t_0, t^*)]$ and $\mathbb{E}[u(e^+(t))|t \in [t^*, t_N]] = \mathbb{E}[u(e^-(t))|t \in [t^*, t_N]]$. Let $T^1 = [t_0, t^*), T^2 = [t^*, t_N]$ and $N = \{n^1, n^2\}$ be a corresponding message set. If the decision maker's type is in T^1 (or T^2), she sends n^1 (or n^2). Then, message sets N and M constitute a perfect Bayesian equilibrium and the decision maker's *ex ante* payoff is the same as in the one-way game.

In the proof of Proposition 5, we find that the size of T_M plays a critical role in forming a type-revealing strategy. The following result is immediate from the observation.

Proposition 6. *Suppose players are bound to use first- and second-order strategies. If the expert's strategy is first-order (resp. second-order), the decision maker's strategy must be second-order (resp. first-order).*

The right panel of Figure 1.3 shows a type-revealing equilibrium strategy which is second-order. As in the first-order strategy, the decision maker's estimates in the one-way game are $e^+(t)$ and $e^-(t)$, and $T_M = \{t_1^0\}$. We can find $t_1^*, t_2^* \in T$ such that

$$\mathbb{E}[u(e^+(t))|t \in [t_0, t_1^*) \cup [t_2^*, t_N]] = \mathbb{E}[u(e^-(t))|t \in [t_0, t_1^*) \cup [t_2^*, t_N]]$$

and

$$\mathbb{E}[u(e^+(t))|t \in [t_1^*, t_2^*)] = \mathbb{E}[u(e^-(t))|t \in [t_1^*, t_2^*)].$$

Let $T_1 = [t_0, t_1^*) \cup [t_2^*, t_N]$ and $T_2 = [t_1^*, t_2^*)$. Then, strategy $N = \{T_1, T_2\}$ and strategy M in the one-way game constitute a perfect Bayesian equilibrium.

The previous section shows that it is because the decision maker's type is unknown that informative communication can take place and benefit both the expert and decision maker. This section shows, on the other hand, that decision makers may partially reveal their types so that they form groups according to their revealed types and cooperate for common interests without information loss.

1.5. Conclusion

We have examined how the expert and the decision maker communicate with each other when the decision maker's type is uncertain and the expert's payoff depends only upon the decision maker's action, regardless of the realised state and the decision maker's type. There are two main results: (1) the expert can always inform the decision maker in the form of comparative statements; and (2) the decision maker may reveal her private type without loss of informative communication. However, we have not fully exploited the potential for communication in this setup. This study could be further developed in two directions.

The first is for refining equilibria. Studies in cheap talk games have paid much attention to identifying equilibria which are optimal to the expert or the decision maker and to predicting a specific equilibrium in the real world among (infinitely) many equilibria.[13] This chapter has focused mainly on whether and how an expert and a decision maker could talk to each other, but not on what messages would be sent in an equilibrium in a certain environment. That is, there is no answer as to which equilibrium strategies are optimal to players and more likely to be played. Nevertheless, two results of this chapter are worth noticing for the prediction of equilibrium outcomes.

One is that there exist infinitely many equilibria. In particular, there are infinitely many 2-message equilibria that can be constructed from an arbitrary interior point in the state space and infinite sequences of equilibria starting from each 2-message equilibrium. The latter infinity is more problematic than the former in predicting an equilibrium outcome if the expert's preferences are strictly convex. Among 2-message equilibria, we can expect to find an equilibrium that is most favourable to the expert under certain conditions, whereas the expert cannot find an optimal equilibrium strategy from the sequences of equilibria because the expert's *ex ante* payoff strictly increases with the number of messages if the preferences are strictly convex.

The other is that the expert's payoff is state-independent. This allows him to commit to an equilibrium strategy. Suppose that the expert can declare which strategy he will choose among infinitely many equilibrium strategies before a one-way cheap talk game starts. Then, if he finds an optimal strategy which maximises his *ex ante* payoff, declaring that the optimal strategy will be used and sending a message according to the strategy constitute an equilibrium. Thus, once the expert finds his optimal strategy in a certain environment,[14] we could predict that the chosen strategy would be played in equilibrium.

The second direction requires examining the relation between priors and communication. This chapter does not fully discover the relation between players' prior beliefs about their environment and equilibrium outcomes. As we have briefly shown in the government spending example, the distribution function of the decision maker's

[13]In an effort to refine cheap talk equilibria, Chen et al. (2008) identify a condition on equilibrium payoffs, called NITS (no incentive to separate), which selects an equilibrium strategy induced by the biggest partition among CS equilibria.

[14]For example, concave preferences, deliberation costs or conventional communication protocols could restrict the number of messages available in communication.

type determines how well she is informed in equilibrium; if the population is concentrated around her type, she is less informed than otherwise.

This implies that the decision maker has preferences over the distribution of types. That is, the decision maker prefers any distribution where the population is more concentrated on other types than hers to the actual distribution. For example, as shown in Section 4, if the decision maker is located in the left of the type space, she prefers a distribution that has population concentrated on the right. Thus, she has an incentive to disguise not her type but the distribution of types.[15]

[15] In this case, common knowledge among players should be with respect to the probabilities assigned to several distributions of types not to each type itself.

CHAPTER 2

A Model of Satisficing Behaviour

Satisficing, a word introduced into the lexicon of economists' by Simon (1956),[1] refers to a decision making strategy in which the agent stops searching for a better alternative if she is satisfied with the current action and continues to explore if not. Such an agent is, probably, the canonical boundedly rational agent. One who does not optimize, or carefully forms beliefs about her environment, before making a choice. An agent who (implicitly) recognizes that decision making, in single or multi-agent environments, may be complex and difficult and adjusts her behaviour suitably. In this chapter we seek to model and analyse the behaviour of such an agent in decision problems and in games.

Satisficing does not preclude maximizing. It is mainly in complex environments, in which there is considerable lack of information and uncertainty, when optimization given beliefs appears to be implausible, that agents are thought to satisfice. That is, an agent may choose to satisfice in some environments while maximizing in other less complex ones. In simpler, better understood environments, maximizing is probably a reasonable hypothesis. Whereas economists have a good understanding of agents who optimize, there is still little understanding of how agents behave when optimization is not a reasonable hypothesis. In this chapter, we study a particularly suited behavioural hypothesis, when optimization does not seem plausible.

Two immediate challenges confront the modelling of satisficing agents. What is the satisficing level of such an agent and how is it updated? Clearly, the behaviour of any such agent will depend intricately upon her satisficing level. Where does this satisficing level come from and how/when is it updated? We feel, an individual's initial satisficing level probably depends closely on her upbringing and the environment in which she was brought up in: her parents, their peer group or their aspirational social group. However, these same influences are less likely to play a role in how and when it is updated. The experiences of the agent and what she considers to be

[1] The idea was first suggested by Simon (1947).

her "best" outside options at any time probably play a greater role in the updating of what the agent finds to be satisfactory. This clearly depends upon other factors, as well as the endogeneity of this process on that which leads to her choices.

The extant literature on satisficing has largely taken the satisficing level to be exogenously given and fixed (e.g., Posch, 1999; Posch et al., 1999). The literature, then, has studied the consequences of assuming various different satisficing levels. The limited literature which has modelled the satisficing level to be endogenous has taken its level to be adjusted according to the payoff experiences of the agent (e.g., Karandikar et al., 1998; Cho and Matsui, 2005; Chasparis et al., 2013). In particular, the satisficing level is taken to be some average of the historical payoffs achieved by the agent. That is, the satisficing level is often treated as the agent's aspiration level which also adjusts according to past payoff experiences.[2]

In this chapter, we develop a model of satisficing in which the state of an agent, in any period, is given by the action she chooses, her valuation of that action and her satisficing level. It is the valuation of the action that depends on the past payoffs it has received. The satisficing level is thought of as the payoff the agent finds satisfactory. What is satisfactory, in turn, depends on what the agent thinks is the payoff she might get from her best outside option. This is adjusted whenever the agent receives any information on this. If the agent's valuation of the current action is above her satisficing level she chooses it again. If it falls below the satisficing level she moves away from it, where her probability of shifting depends upon the amount by which her valuation falls below the satisficing level. From time to time, the agent experiences shocks or trembles on the action she chooses and on her satisficing level. We study the long run choices of the agent in stochastic decision problems and normal form games.

We show that, in decision problems or normal form games, if the agent never experiences shocks, then she eventually converges to being satisficed. While it is nice to know that players in the long run are satisficed, this still leaves open a great many possible asymptotic outcomes. The addition of noise, which dies out in the long run, helps select among the outcomes predicted for satisficing agents. Additionally, we assume that the weight given to the current action's valuation in

[2]In both cases, the satisficing level has been allowed to experience shocks.

updating the satisficing level declines over time. In decision problems this leads the agent to choose in a very risk averse manner. Specifically, the agent ends up choosing only those actions for which minimum payoff is highest. That is, the agent converges to *maximin actions.*

Taking the same limit in games, in which the agents experience non-stationary distributions over payoffs, leads to some surprising results in which non-maximin profiles are often obtained in the long run. For instance, in the Prisoners' dilemma both the cooperate only and the defect only profiles have positive probability in the long run. Players coordinate in coordination games and converge to Pareto optimal equilibria in pure coordination games whenever they exist. However, in the Stag-Hunt game both cooperate only and defect only emerge in the long run. And, we generalise these results to broader classes of two-player games and to random matching games in which a finite population of players are randomly matched each period.

Lastly, we consider the long run outcomes which arise when we only send the shocks to zero while not decreasing the amount of persistence in the satisficing level. In this case, we find for example, only the defect only outcomes emerges in the Prisoners Dilemma.

Section 1 introduces satisficing behaviour and Section 2 examines its long run property. Section 3 and 4 analyse long run implications of satisficing behaviour in individual decision problems and normal form games. Section 5 summarises the results. All proofs are in Appendix A.2.

2.1. The Model

A finite game $\Gamma = \{\mathcal{I}, A, \pi\}$, where $\mathcal{I} = \{1, \ldots, I\}$ is the set of players,[3] $A = \Pi_{i \in \mathcal{I}} A_i$ with finite A_i for each i is the set of possible action profiles with typical member $a \in A$ and $\pi : A \to \mathbb{R}^I$ is the payoff function with player i's payoff π_i, is played in each period $n \geq 1$. If Nature plays a role, she chooses her action w from a finite set W with fixed probabilities each period, and π becomes a function of w as well as a. With a little abuse of notation, we denote the whole set of payoffs that decision maker i could receive by $\pi_i = \{\pi_i(a, w) \in \mathbb{R} | a \in A, w \in W\}$, and the set of

[3]Player, decision maker and agent are used interchangeably according to context.

payoffs from decision maker i's action $a_i \in A_i$ by $\pi_i(a_i) = \{\pi_i(a_i, a_{-i}, w) \in \mathbb{R} | a_{-i} \in A_{-i}, w \in W\}$.

2.1.1. Satisficing without Trembling

In period n, I decision makers' states are given by their current action profile $a_n \in A$, the corresponding valuations $v_n \in \mathbb{R}^I$ and satisficing levels $s_n \in \mathbb{R}^I$. Let ϕ_n denote the state (a_n, v_n, s_n) with decision maker i's state $\phi_{n,i} = (a_{n,i}, v_{n,i}, s_{n,i})$ for each $i \in \mathcal{I}$. ϕ_0 is assumed given.

At the start of any period $n \in \mathbb{N}$, each decision maker i judges her current action $a_{n,i}$ by comparing its valuation $v_{n,i}$ with her ongoing satisficing level $s_{n,i}$. If valuation $v_{n,i}$ is greater than $s_{n,i}$, she satisfices and continues with the current action and satisficing level. If the valuation falls short of the satisficing level, she switches to an alternative and updates her satisficing level towards the valuation. Specifically, if player i satisfices with $a_{n,i}$, or $v_{n,i} \geq s_{n,i}$,

$$a_{n+1,i} = a_{n,i} \quad \text{and} \quad s_{n+1,i} = s_{n,i},$$

and if $v_{n,i} < s_{n,i}$,

$$a_{n+1,i} = \alpha_{n,i} \quad \text{and} \quad s_{n+1,i} = (1 - \lambda_{n,i}) s_{n,i} + \lambda_{n,i} v_{n,i},$$

where $\alpha_{n,i} \in A_i$ and $\lambda_{n,i} \in [0, \overline{\lambda}]$ for some $\overline{\lambda} \in (0,1)$ are random variables. Each individual i, then, chooses the action so determined and receives a payoff $\pi_i(a_{n+1})$. Before the end of period n, the decision makers revise their valuations by taking weighted averages of their payoffs and valuations as follows. If $a_{n+1,i} = a_{n,i}$,

$$v_{n+1,i} = (1 - \rho_{n,i}) v_{n,i} + \rho_{n,i} \pi_i(a_{n+1}),$$

and, if $a_{n+1,i} \neq a_{n,i}$,

$$v_{n+1,i} = \pi_i(a_{n+1}),$$

where $\rho_{n,i} \in [0,1]$ is a random variable. In the above $\alpha_{n,i}, \rho_{n,i}$ and $\lambda_{n,i}$ have probability measures $\mu_{\alpha,i}, \mu_{\rho,i}$ and $\mu_{\lambda,i}$ with full supports in A_i, $[0,1]$ and $[0, \overline{\lambda}]$, respectively, for all i.[4] We shall refer to $\mu_{\alpha,i}$ as *choice rule*, which is continuous in v and s so that

[4]This assumption implies that $\mu_{\alpha,i}(\{a_i\}) > 0$ for all $a_i \in A_i$ and any open set contained in $[0,1]$ and $[0, \overline{\lambda}]$ have positive probabilities with respect to $\mu_{\rho,i}$ and $\mu_{\lambda,i}$.

$\mu_{\alpha,i}(a) \downarrow 0$ as $v_{n,i} \uparrow s_{n,i}$ for all $a \in A_i \setminus \{a_{n,i}\}$ and $\mu_{\alpha,i}(a_{n,i})$ is bounded away from zero for any $v_{n,i}$ below $s_{n,i}$. $\mu_{\rho,i}$ and $\mu_{\lambda,i}$ are assumed to be absolutely continuous with respect to the Lebesgue measure on the intervals and continuous functions of the current state ϕ_n.[5]

Some remarks are in order. First, notice that the satisficing level is updated only when the individual does not satisfice with the current action. If the agent satisfices she sees no reason to change. The urge to explore alternate actions only arises when dissatisfaction with the current action occurs. In this case, the agent adapts to her environment by choosing an alternate action. She also adapts by lowering her satisficing level. This assumption is motivated by our belief that the satisficing level can be regarded as the payoff a decision maker "expects" from choosing the best alternative action (different from the current).[6] Furthermore, the extent to which the satisficing level adjusts is bounded by the *persistence* parameter $\overline{\lambda} \in (0, 1)$.[7] Since $\overline{\lambda}$ is the same for all decision makers and states, it characterises how fast the decision makers adjust their satisficing level towards their current actions' valuations. We think our interpretation of satisficing level as the expected value of alternatives makes a reasonable economic sense: Satisficing agent satisfices because she does not expect better alternatives.

Second, the weighting parameters λ and ρ are random variables. We assume $\text{supp}(\mu_\rho) = [0, 1]$ so that decision makers could base their valuations of actions on some averages of its past payoffs. In literature, fixed ρ or declining ρ is often used and it is common that decision makers are assumed to respond only to immediate payoffs, or $\rho = 1$.[8] Our choice of random weights is to allow for the bounded rationality our satisficing decision makers exhibit. The actual weighting per period possibly varies according to the attention or subjective importance the decision makers choose to give the current payoff.

---

[5]These two conditions together make the satisficing behaviour continuous so that the state transition over time constitutes a *weak-Feller* Markov chain.

[6]This contrasts with the literature on the subject in which the common interpretation of satisficing level (or aspiration value) is the (best) payoff that a decision maker expects from a decision problem itself (e.g. see Gilboa and Schmeidler (1996), Karandikar et al. and Cho and Matsui (2005)).

[7]Karandikar et al. (1998) use this term. Chasparis et al. (2010) have a parameter that plays a similar role in their model, which they refer to as *step size*.

[8]See Karandikar et al. (1998), Cho and Matsui (2005) and Chasparis et al. (2010). Posch (1999) considers a behaviour rule in which a decision maker switches to another action only if a weighted sum of past payoffs is below an aspiration level.

Lastly, decision makers have inertia to choose their current actions again with some positive probability even when the strategy is not satisfactory. The choice rule makes decision makers repeat actions with positive probabilities that depend on how the valuations fall short of satisficing levels, and with probabilities close to 1 when the valuations are slightly lower than satisficing levels. It makes the transition of the decision maker's states over time continuous and the analysis of the behaviour more tractable.

Let V_i and S_i both denote the convex hull of π_i for all $i \in \mathcal{I}$, and V and S the product spaces of V_i and S_i across all players, respectively.[9] Then, the decision makers' states governed by the above updating rule over time constitute a Markov process $\{\phi_n\}$ on the compact space $\Phi \equiv A \times V \times S$. To distinguish this process from the behaviour with trembling described in the next subsection, we call this *unperturbed* Markov process with transition probability function (hereafter t.p.f.) P.[10] We will refer to P as the Markov process as well as the t.p.f.

We close this subsection with a long run property of satisficing behaviour without trembling: all players eventually settle on an action profile with which they satisfice.

Proposition 7. *All satisficing decision makers without trembles eventually satisfice and choose one action profile forever.*

Given the property of satisficing level, as time goes to infinity, it gets lower and lower until it becomes to support at least one action profile. However, such selected actions are quite arbitrary because any action profile can be supported if initial satisficing levels are sufficiently low. And, even when the initial levels are not so low, most action profiles can be selected with positive probabilities unless they are associated with minimum payoffs.

2.1.2. Trembling Behaviour

In case they tremble, satisficing decision makers experience two types of shocks. The first is trembling hands. When this shock occurs, a decision maker chooses an arbitrary action among all actions the next period, regardless of whether she satisfices or not with the current action. The second shock directly affects satisficing levels

[9]The convex hull of a set is defined as the smallest convex set that contains the set.
[10]Transition probability function is defined in section 2.2.

so that the levels are adjusted towards current actions' valuations. In each period, players independently experience the two shocks, each with positive probability ϵ, which is referred to as *trembling probability* and the same for all satisficing players.[11]

Given $\phi_{n,i}$, if decision maker i experiences only the first shock, which occurs with probability $\epsilon(1-\epsilon)$, her next period action is chosen at random, with

$$a_{n+1,i} = \alpha^{\epsilon}_{n,i},$$

while her satisficing level and valuation of the action are updated according to the unperturbed process described previously. $\alpha^{\epsilon}_{n,i}$ is a random variable that has full support in A_i. On the other hand, if the agent experiences only the second shock, which occurs with probability $\epsilon(1-\epsilon)$, her next period satisficing level is updated as

$$s_{n+1,i} = (1-\lambda^{\epsilon}_{n,i})s_{n,i} + \lambda^{\epsilon}_{n,i}v_{n,i},$$

where $\lambda^{\epsilon}_{n,i}$ is a random variable with probability measure $\mu^{\epsilon}_{\lambda,i}$ with full support in $[0,1]$. Her action and its valuation are updated as in the unperturbed process. Lastly, if she experiences both shocks in the same period, which occurs with probability ϵ^2, her state is updated as if only the first shock occurs, and then just before the next period starts, her satisficing level trembles to be adjusted upwards current valuation.

Let Q_ϵ denote the t.p.f. when at least one decision maker experiences one or both shocks. Then, the unperturbed process P and Q_ϵ together constitute a *perturbed* process P_ϵ, which is the Markov process for satisficing behaviour that we examine to characterise satisficing decision makers' long run behaviour.

2.1.3. Satisficing with Additional Memory

Satisficing decision makers judge actions, choose actions and update satisficing levels only based on latest action profiles and its valuations. This low information, or limited memory, assumption naturally leads to a simple choice rule in which a decision maker assigns equal probabilities to all alternatives when she does not satisfice.

[11] The trembling probabilities may differ across players and by whether the shocks occur on choice of action or satisficing level. In particular, we denote the latter two probabilities by ϵ^a and ϵ^s, respectively if required. And, for any sequence of the probabilities $\{(\epsilon^a_n, \epsilon^s_n)\}_{n\geq 1}$, we assume $\epsilon^a_n/\epsilon^s_n \in (1/K, K)$ for some $K \in \mathbb{N}$ for all $n \geq 1$ and $\epsilon_n \to 0$ implies $\max\{\epsilon^a_n, \epsilon^s_n\} \to 0$. Together with the persistence parameter $\overline{\lambda}$, the probability plays a critical role in the present model.

One may suspect that this assumption is too restrictive in some environments. Here, we model satisficing behaviour in which decision makers might keep track of past actions and valuations so that their choice rules utilise this additional information.

Let a compact metric space Υ represent an auxiliary information, and the Markov process ϕ_n of satisficing behaviour be augmented with a state $\eta_n \in \Upsilon$ so that (ϕ_n, η_n) is a Markov process defined on $\Phi \times \Upsilon$. We assume that (1) η_{n+1} is a continuous function of (ϕ_n, η_n), and (2) $\eta_{n+1} = \eta_n$ if $a_{n+1} = a_n$, i.e., the auxiliary state is updated only when at least one decision maker changes her action.[12] After decision makers' next period state ϕ_{n+1} is determined by the perturbed process P_ϵ with (ϕ_n, η_n) in the place of ϕ_n, the auxiliary state η_n is updated to the next period η_{n+1} by ϕ_{n+1}. To see how the current setup may change behaviour, we introduce an example.

Suppose that each decision maker remembers her all actions' latest valuations. Then, in period n with state (ϕ_n, η_n), decision maker i can base her choice of next period action on this additional information by assigning higher probabilities to actions that are associated with higher valuations. Let $\eta_{n,i} \equiv (v^1_{n,i}, \ldots, v^{|A_i|}_{n,i})$ denote decision maker i's valuations for each action, where $|A_i|$ is the number of actions available to decision maker i. The auxiliary state is updated only when decision makers change their actions. If $a_{n+1,i} \neq a_{n,i}$,

$$v^j_{n+1,i} = \begin{cases} v_{n,i}, & \text{if } a_{n,i} = a^j \\ v^j_{n,i}, & \text{otherwise.} \end{cases}$$

And, if $a_{n+1,i} = a_{n,i}, \eta_{n+1,i} = \eta_{n,i}$.

Accordingly, we can construct a choice rule μ_α that has the inertia and continuity properties.[13] In each period, satisficing decision makers are assumed to experience positive, transitory shocks on her all valuations including the current action's. Let z^j denote the shock on action j, then refer to $v^j + z^j$ as *perceived* valuation of action j, where z^j is an action-specific valuation shock and its probability measure is absolutely continuous w.r.t. the Lebesgue measure and has full support in $\mathbb{R}^+$. A

[12] The first condition is required for the process with the auxiliary state to be weak-Fellar and the second condition is for Lemma 6.

[13] Recall that we require the choice rule, or the probability of choosing each action, is continuous in the valuation of the current action and satisficing level and assigns positive probability to the current action for any valuation and satisficing level.

decision maker judges the current action by comparing its perceived valuation to her satisficing level as modelled previously. The update rules for action, satisficing level and valuation are exactly the same as before except for the perceived valuation. When the decision maker does not satisfice with her current action, she chooses an action with highest perceived valuation as modelled in Sarin and Vahid (1999). Then, the choice rule induced by this procedure satisfies all required properties.[14]

Then, the auxiliary state η_n is a continuous function of ϕ_{n+1} and η_n, and $\eta_{n+1} = \eta_n$ if no decision maker changes her action. With the auxiliary state, we can make choice rule μ_α (and other distributions of z, ρ and λ) depend on η_n so that the actions with higher valuations are chosen more likely by players.[15]

In the next section and corresponding proofs, we show that the introduction of the additional state does not affect asymptotic properties of the original Markov process. Thus, for the expositional simplicity, we refer to the extensive Markov processes also as P and P_ϵ and the auxiliary state η_n is considered implicitly from now on as if P and P_ϵ are t.p.f.s for only ϕ_n without η_n.

2.2. Asymptotic Behaviour: Preliminaries

This section provides two asymptotic results of satisficing behaviour as the number of repetition n approaches infinity and the trembling probability ϵ vanishes while the persistence parameter $\overline{\lambda}$ is fixed.

Some primitive notations and definitions are first introduced. Let Φ be a topological space, and $\mathcal{B}(\Phi)$ denote the Borel σ-field. Then, $P(\phi, B)$ is a t.p.f. if (i) for each $B \in \mathcal{B}(\Phi)$, $P(\cdot, B)$ is a non-negative measurable function on Φ and (ii) for each $\phi \in \Phi$, $P(\phi, \cdot)$ is a probability measure on $\mathcal{B}(\Phi)$. A probability measure (hereafter p.m.) μ is *invariant* with respect to P if $\mu P(B) = \mu(B)$ for all $B \in \mathcal{B}(\Phi)$.[16] Thus, if

[14] If the shocks on actions are independent of each other, when a_n is not satisficing, the probability of the decision maker choosing an alternative a is $P_n(a) = \Pr\{z_n^a - z_n^{a'} < u_n^a - u_n^{a'} \text{ for all } a \neq a', a_n\}$.

[15] If a decision maker has too many available actions to keep track of all actions' valuations, it might be more reasonable to assume that decision maker i remembers valuations of recently chosen actions with the number limited by $K \leq |A_i|$. Let the auxiliary state η_n have the form of $\{(a_{k,i}, v_{k,i})\}_{k=1,\dots,K}$, i.e., the decision maker remembers actions she has chosen recently and corresponding valuations. The state is updated as follows. If $a_{n+1,i} \neq a_{n,i}$,

$$(a_{k,i}, v_{k,i}) = \begin{cases} (a_{k-1,i}, v_{k-1,i}), & \text{if } k \in \{2, \dots, K\} \\ (a_{n,i}, v_{n,i}), & \text{if } k = 1 \end{cases}$$

And, if $a_{n+1,i} = a_{n,i}, \eta_{n+1,i} = \eta_{n,i}$. This also satisfies the required properties.

[16] $\mu P(B) \equiv \int_\Phi \mu(d\phi) P(\phi, B)$.

a Markov process has an invariant p.m. and once reaches it, thereafter the process' state is fully predicted by the invariant p.m. Let P^n denote the n-step t.p.f. of P defined as $P^n = P^{n-1}P$, where multiplication of two t.p.f. P and Q is defined as

$$PQ(\phi, B) \equiv \int_{\Phi} P(\phi, d\phi')Q(\phi', B)$$

for any $\phi \in \Phi, B \in \mathcal{B}(\Phi)$. P^∞ denotes the limit of P^n as n approaches infinity. Lastly, let R_ϵ be the *resolvent* of P, defined as $R_\epsilon \equiv \epsilon \sum_{n=0}^{\infty}(1-\epsilon)^n P^n$ for some $\epsilon \in (0, 1)$.

The first asymptotic result is that the Markov process P_ϵ generated by the satisficing behaviour has a unique invariant p.m. and converges to it regardless of initial choices and satisficing levels.

Proposition 8. *P_ϵ has a unique invariant p.m. μ_ϵ, and for any ϕ, $P_\epsilon^n(\phi, \cdot)$ strongly converges to μ_ϵ as n grows.*

The second result shows how the invariant p.m. changes as the trembling probability ϵ vanishes. Let Q denote the t.p.f. conditional on that only one decision maker experiences only one shock, either the first or the second.[17] Then, QP^∞ is also a t.p.f. and has the same unique invariant p.m. with P_ϵ as ϵ approaches 0.

Proposition 9. *Any weak accumulation point of $\{\mu_\epsilon\}_{\epsilon \downarrow 0}$ is an invariant p.m. of QP^∞.*[18]

This result implies that, as the trembling probability decreases over time, satisficing decision makers' long run behaviour can be approximated by some invariant p.m. of QP^∞. In particular, if QP^∞ has a unique invariant p.m., μ_ϵ weakly converges to it. Therefore, as $P_\epsilon^n(\phi, \cdot)$ converges to μ_ϵ over time, the satisficing decision makers in the long run choose the action profiles that are associated with the invariant sets most of the time.[19] In the following sections, we analyse satisficing behaviour in particular contexts of single person decision problems and normal form games by identifying the invariant sets.

[17]The precise definition of Q is provided in Appendix A.2.

[18]These are derived by Chasparis et al. (2013) and KMRV differently. However, as Chasparis et al. (2013) noted, KMRV wrongly assume that Q is strong-Feller. Here, we follow the approach taken by Chasparis et al. (2013).

[19]In Appendix C, satisficing behaviour is simulated with $\epsilon, \overline{\lambda} = 0.05$ or 0.01 for various decision environments.

2.3. Individual Decision Problems

In this and next sections we analyse satisficing decision makers' long run behaviour when they repeatedly face a decision problem or game as the trembling behaviour dies out and the satisficing level gets more and more persistent. In individual decision problems, in which each action returns random payoffs according to a fixed distribution across periods, we find decision makers end up choosing maximin actions.

First, we examine the long run implication of satisficing behaviour without trembling. A decision maker has J actions, $a^1, \ldots, a^J$, which are indexed by their minimum payoffs so that $\min \pi(a^i) \geq \min \pi(a^j)$ for any $a^i, a^j \in A$ if $i \leq j$. All maximin actions guarantee the highest minimum payoff for any realised state.

Proposition 10. *If the initial satisficing level is sufficiently high, the satisficing decision maker without trembling eventually chooses the maximin actions with probability 1 as $\overline{\lambda}$ tends to* 0.

This result and proof are comparable to Sarin and Vahid (1999). If the satisficing level is initially sufficiently high and very persistent, the decision maker has sufficiently many chances to choose the maximin actions and satisfice before the satisficing level gets too low so that she becomes to satisfice with non-maximin actions.

Now suppose the decision maker is subject to both shocks and the trembling probability ϵ and persistence parameter $\overline{\lambda}$ go to 0.[20]

Proposition 11. *The satisficing decision maker in the long run chooses the maximin actions.*

A satisficing agent's choices are results of her initial choice (and satisficing level) and payoff experiences from the decision problem she faces. However, as experiences accumulate, the effect of her initial state in the long run fades away and invariant p.m. μ_ϵ becomes to dominate her behaviour (Proposition 8).

[20] As discussed in the previous section, we analyse the satisficing agent's long run behaviour by characterising the limit of μ_ϵ as ϵ goes to 0. Thus, if both ϵ and $\overline{\lambda}$ go to 0, the behaviour is approximated by the limit invariant distribution of P_ϵ when first ϵ goes to 0, then $\overline{\lambda}$ goes to 0. This is the same for all subsequent analyses when both parameters go to 0.

The proof consists of two arguments. First, by Proposition 10 once the decision maker happens to choose one of the maximin actions and her satisficing level gets close to $\min \pi(a^1)$, after any single tremble either on action or satisficing level, she eventually returns to the maximin actions. In this regard, we say the maximin actions are robust to any single tremble. Second, regardless of initial action and satisficing level, trembling behaviour on action and satisficing level makes her experience the payoffs from the maximin actions and raise her satisficing level up to $\min \pi(a^1)$ through finite trembles.

This result contrasts the current satisficing behaviour model with previous ones.[21] Since models with aspiration level require aspiration level to be updated toward received payoffs every period, it works well in deterministic environments but not in stochastic environments that we consider here. For example, suppose a decision problem with two actions, in which one returns payoffs 3 and 2 with equal probabilities and the other returns payoffs 1 and 0 with equal probabilities. Then, as long as satisficing/aspiration levels are updated towards payoffs every period, agents choose the second action as frequently as the first. In our model, they converge to choose the first action with probability 1.

We look further how the agent of KMRV behave in this simple decision problem.[22] The behaviour is modelled differently from the current in several aspects. First, immediate payoffs are directly compared with satisficing levels, i.e., ρ is fixed at 1. Second, satisficing levels are updated toward immediate payoffs every period with constant weighting parameter λ, which directly plays the role of persistence parameter. Third, agents experience shocks only to satisficing levels, which are perturbed within a range around the current value when shocks occur. In the current stochastic decision problems, an unperturbed process induced by the satisficing behaviour in KMRV does not converge to any pure strategy state for any $\lambda > 0$, rather

[21]Gilboa and Schmeidler (2001) state, "Realism means that the aspiration level is set closer to the best average performance so far experienced." Most satisficing behaviour examined in single-person decision problems are assumed to take an average of past (best) payoffs as aspiration level. See Gilboa and Schmeidler (1996), Pazgal (1997), Kim (1999), Karandikar et al. (1998), Cho and Matsui (2005), Posch (1999), Posch et al. (1999), Napel (2003), Bendor et al. (2009). For survey for different aspiration updating rules, see Bendor et al. (2001).

[22]The satisficing behaviour in Cho and Matsui (2005) is similar to KMRV except that agents in CM experience shocks only to choice of actions instead of satisficing levels (aspiration level in CM). However, the methodology used to model satisficing behaviour in CM cannot be applied to stochastic decision problems. We review the satisficing behaviour in CM in the analysis of 2×2 games.

it alternates between two actions forever because the agent never satisfices with the first action as her satisficing level is raised above 2 infinitely often.

We close this section with short notes on probability matching behaviour and dual risk attitudes induced by satisficing behaviour.

Proposition 11 predicts that, if there are multiple maximin actions, the decision maker will choose any maximin action with positive probabilities. However, this does not tell us the frequencies with which each action is chosen. Consider a decision problem in which two actions a and b have a common minimum payoff, say 0, with probabilities p and $1-p$, respectively, every period. And, as a limit of satisficing behaviour, assume that the decision maker never trembles, the satisficing level is fixed slightly above 0, $\rho = 1$ and does not show inertia behaviour. Then, the frequency of choosing each action is given as the invariant distribution of a simple 2×2 transition probability matrix P in which $P(i, j)$ is the probability of switching from action i to action j, and the long run probability of choosing action a (resp. b) becomes $1-p$ (resp. p).[23]

March (1996) suggests that dual risk attitudes for gains and losses, risk averse when outcomes are positive with respect to a fixed aspiration level and risk seeking when outcomes are negative, could be better explained by learning rather than human traits or utility functions. The same argument can apply to the current satisficing behaviour when initial satisficing level is 0 and $\overline{\lambda}$ is sufficiently small. In Appendix C.1, we simulate satisficing behaviour in binary choice problems between safe and risky options.

2.4. Normal Form Games

In this section, satisficing decision makers repeatedly play a normal form game against other satisficing decision makers.

2.4.1. Two-Player Games

First, we analyse satisficing behaviour in two-player games with *unilaterally competitive* action profiles. A class of (weakly) unilaterally competitive games were first introduced and analysed by Kats and Thisse (1992) as a generalisation of strictly

[23] Börgers and Sarin (2000) shows reinforcement learning with endogenous aspiration level leads to probability matching behaviour.

competitive n-player games.[24] For our analysis, we define the unilateral competitiveness in terms of action profiles rather than games.

Definition 6. *An action profile $a \in A$ of a game Γ is* unilaterally competitive *if for any i and $a_i' \in A_i$,*

$$\begin{aligned} (i) \quad \pi_i(a_i, a_{-i}) &> \pi_i(a_i', a_{-i}) \Longrightarrow \pi_{-i}(a_i, a_{-i}) \leq \pi_{-i}(a_i', a_{-i}) \text{ and} \\ (ii) \quad \pi_i(a_i, a_{-i}) &= \pi_i(a_i', a_{-i}) \Longrightarrow \pi_{-i}(a_i, a_{-i}) = \pi_{-i}(a_i', a_{-i}). \end{aligned}$$

And, the game is unilaterally competitive *if all action profiles are unilaterally competitive.*[25]

In unilaterally competitive games, if a single player deviates to increase her payoff from any action profile, all other players' payoffs weakly decrease. And, if the deviation does not change her payoff, all other players' payoffs also do not change.

An action profile $a \in A$ is a Nash equilibrium if $\pi_i(a_i, a_{-i}) \geq \pi_i(a_i', a_{-i})$ for all $a_i' \in A_i$ and $i \in \mathcal{I}$, and a strict Nash equilibrium if all inequalities are strict. If a Nash equilibrium is unilaterally competitive, it is called a *unilaterally competitive Nash equilibrium.* If it is a strict Nash equilibrium, such equilibrium is unique. And, if a game has more than one unilaterally competitive Nash equilibria, each player's payoffs from the equilibria are the same.[26]

For two action profiles $a, a' \in A$, a is *preferred to a'* if $\pi(a) \geq \pi(a')$ and $\pi_i(a) > \pi_i(a')$ for some $i \in \mathcal{I}$, and a is *strictly preferred to a'* if $\pi(a) \gg \pi(a')$. Then, *Payoff dominant* profiles are defined as profiles that are strictly preferred to unilaterally competitive Nash profiles.[27]

Definition 7. *When a two-player game admits unilaterally competitive Nash equilibria, a set of action profiles are* payoff dominant *if*
(i) any two profiles within the set are not preferred to each other, and
(ii) all those profiles are strictly preferred to its unilaterally competitive Nash profile.

---

[24]In two-player games, any strictly competitive game is also unilaterally competitive.

[25]In Kats and Thisse (1992), a game is unilaterally competitive if

$$\pi_i(a_i, a_{-i}) \geq \pi_i(a_i', a_{-i}) \Longleftrightarrow \pi_{-i}(a_i, a_{-i}) \leq \pi_{-i}(a_i', a_{-i})$$

for all $a \in A$, and the unilaterally competitive game defined here is *weakly unilaterally competitive.*

[26]This property of unilaterally competitive Nash equilibria is generalised for unilaterally competitive games in Lemma 2.

[27]We do not require payoff dominant profiles to be Nash.

	C	D
C	σ, σ	$0, \theta$
D	$\theta, 0$	δ, δ

Table 2.1. Payoffs of 2 $\times$ 2 Game

Consider a 2×2 game with its payoff matrix in Table 2.1. If $0 < \delta < \sigma < \theta$, the game becomes Prisoner's Dilemma game, which is unilaterally competitive. If $0 < \delta < \theta < \sigma$, it is Stag Hunt game and not unilaterally competitive. In terms of action profiles, (D, D) is unilaterally competitive Nash and (C, C) is payoff dominant in both games. And, if $\delta = \sigma < 0 < \theta$ or $0 = \theta < \delta < \sigma$, it is Battle of the Sexes or Common Interest game, respectively, and both games are not unilaterally competitive. The last two games do not have a unilaterally competitive Nash equilibrium, but (C, D) and (D, C) in the Battle of the Sexes and (C, C) in the Common Interest game are defined as Pareto optimal profiles in the next subsection.

In the limit as trembling probability ϵ and persistence parameter $\overline{\lambda}$ tend to 0, not all outcomes are supported by satisficing agents.

Proposition 12. *For two-player games with unilaterally competitive Nash equilibria, satisficing players in the long run play either unilaterally competitive Nash or payoff dominant profiles.*

This result is given as the sum of probabilities assigned to states in which unilaterally competitive Nash and payoff dominant profiles are played converges to 1 as the parameter values go to 0 in the order specified above.

As in the individual decision problems, the proof consists of two arguments: First, the unilaterally competitive Nash and payoff dominant action profiles are robust to single trembles in the sense that once players satisfice with one of the profiles and their satisficing levels are sufficiently close to the corresponding payoffs, any single tremble on either choice of action or satisficing level cannot make players stay away from the initial profile permanently. Second, any other action profile is absorbed into the robust profiles. That is, starting from any initial state in which players satisfice with an action profile that is not robust, players can become to choose and satisfice with one of the robust profiles with positive probability through a finite sequence of single trembles and subsequent infinite plays without trembles.

Since both KMRV and CM analyse Prisoner's Dilemma game, we can directly compare the current satisficing behaviour with theirs in this simple but famous game. In Karandikar et al. (1998), regardless of relative sizes of θ and σ players learn to play the cooperate only profile most of the time whereas in Cho and Matsui (2005), only when the gain from deviation is moderate, i.e., $\theta - \sigma$ is small, players learn to play the cooperate only profile, and otherwise players alternate among three profiles, cooperate only, cooperate-defect (or defect-cooperate) and defect only. Our result lies between them: Regardless of the relative sizes players learn to play both cooperate only and defect only profiles most of the time. As Cho and Matsui explain, the difference between their results is caused by their different assumptions on trembling behaviour. As long as aspiration level is not greater than payoffs from cooperation, KMRV's satisficing players satisfice in cooperation with probability 1 as trembling probability approaches 1. CM players, however, are assumed to experience shocks to choice of actions, even when aspiration levels exceeds payoffs from cooperation, thus players are not kept in cooperate only profile, rather pursue a higher average payoff. Unlike these two, since satisficing levels are not symmetrically updated when players are satisficing, defect only profile becomes robust to trembles but the asymmetric profiles, cooperate-defect and defect-cooperate, do not.

Proposition 12 also predicts that if $\sigma < \delta$ in Table 2.1, i.e., payoff dominant profiles do not exist, players learn to play only the Nash equilibrium most of the time. If we apply the satisficing behaviour models with aspiration levels, players in both become to play the equilibrium profile. In the former, once a player chooses the maximin action D, the other player's aspiration level eventually falls to δ regardless of initial level. Therefore, both players become to satisfice with the Nash equilibrium. In the latter, once both players' aspiration levels are below δ, their action profiles converges to (D, D). If a player's aspiration level is greater than δ and the other's is less than δ, the aspiration level above δ falls until both levels are below δ, and then both players' actions converge to (D, D).

If more than two satisficing players interact, the proposition partially holds. The first argument of the proof still holds but the second does not. That is, profiles other than unilaterally competitive Nash and payoff dominant action profiles are not guaranteed to be absorbed into those robust profiles. In Appendix C.4, we describe public good provision games of Isaac et al. (1984), in which 4 or 10 satisficing players simultaneously decide how much to contribute, and simulate satisficing behaviour in the games. The satisficing players do not show clear tendency of convergence to either full contribution or no contribution, which correspond to payoff dominant and unilaterally competitive Nash profiles, respectively.

Next we consider strategic interaction within finite populations in which randomly matched players play two-player games like the Prisoners' Dilemma and Stage Hunt.[28] In each period, only matched players play the game according to their states, receive payoffs, and update their states while other non-matched players' states stay the same to the next period. The matching is stationary in the sense that each pair is matched with equal probabilities. The result of Proposition 12, which can be interpreted as satisficing behaviour in fixed matching setup, is extended into this random matching environment.

Proposition 13. *For two-player games with unilaterally competitive Nash equilibria, if finite satisficing players are randomly matched to play the games, the whole population in the long run play either unilaterally competitive Nash or payoff dominant profiles.*

In evolutionary contexts, Nowak and Sigmund (1993), Nowak et al. (1995) and Imhof et al. (2007) find that 'Win-Stay, Lose-Shift,' or shortly WSLS, with aspiration level fixed between the payoffs from mutual cooperation and defection outperforms other strategies such as 'Always Cooperate', 'Always Defect' and 'Tit-for-Tat' in Prisoner's Dilemma games. The dominance of the WSLS is analogous to the convergence of the satisficing levels to the cooperative payoff.[29]

Battalio et al. (2001) provide experiment results in Stag Hunt games with various treatments with a single-population (or cohort) random matching. For each treatment, finite subjects in each population are randomly matched to play a game for a number of periods. They find that all subjects in a population converge to either payoff dominant or unilaterally competitive Nash outcomes both with positive probabilities. In particular, subjects choose the payoff dominant action more frequently

[28]For equilibrium analysis of this setup, see Ellison (1994). He shows that a cooperation equilibrium is supported by "contagious" punishments in the repeated Prisoners' Dilemma game in a random matching setup, which makes it difficult for players to observe opponents past behaviour, but they are aware of themselves playing the game against anonymous opponents.

[29]For other WSLS strategies with fixed aspiration levels and the interpretations, see Posch (1999).

when the payoff difference between payoff dominant and unilaterally competitive Nash profiles is larger, which fits well the simulation result in Appendix C.2.[30]

2.4.2. Pareto Optimal Games

This subsection examines a class of coordination games. We define Pareto optimal profiles as a refinement of Nash equilibria, and provide conditions under which only Pareto optimal profiles are selected by satisficing agents.

Definition 8. *A set of action profiles, $A_P \in A$, are* Pareto optimal *if*
(i) any two profiles within the set are not preferred to each other, and
(ii) all those profiles are preferred to any profile that does not belong to the set.

As its name suggests, Pareto optimal action profiles return good enough payoffs to all players compared to non-Pareto profiles, but it is not guaranteed for satisficing agents to choose only Pareto optimal profiles without further conditions. The sufficient condition is given with the help of the following definitions. First, for any $a, a' \in A$ with $s \in S$, we say (a, s) *reaches* a' if $P^n((a, \pi(a), s), \Phi(a')) > \delta$ for some $n < \infty$ and $\delta > 0$, where $\Phi(a')$ is a set of states in which all action profiles are the same as a'. Second, for any $a, a' \in A$, a' is a *unilateral deviation* of a if there exists a player, say i^*, such that $a'_i \neq a_i$ if and only if $i = i^*$ and $\pi(a') \geq \pi(a)$.

Definition 9. *A game* Γ *is* Pareto optimal *if*
(i) for any $a \in A$ and $a^* \in A_P$, $(a, \pi(a^*))$ reaches the Pareto optimal profile a^* and
(ii) for any $a \notin A_P$, there exists a sequence of unilateral deviations that starts with a and ends with $a^* \in A_P$ or there exists i and $a^*_i \in A_i$ such that $\pi(a^*_i, a_{-i}) \ll \pi(a)$.

The first condition makes all Pareto optimal profiles robust to single trembles. Once players' satisficing levels are set to payoffs of a Pareto optimal profile, they visit the profile with certainty during infinite repetitions without trembles. And, the second condition guarantees that all non-Pareto optimal action profiles become absorbed into one of Pareto optimal profiles through finite trembles. Note that the Prisoner's Dilemma game does not have Pareto optimal profiles, and though the

[30] In the appendix, we also simulate satisficing behaviour in various 2×2 games under the fixed and random matching setups. The results suggest that random matching helps players learn faster and converge to a pure strategy profile.

	C	D
C	σ, σ	$0, 0$
D	$0, 0$	δ, δ

Table 2.2. Payoffs of Coordination Game

Stag Hunt game has one Pareto optimal profile (C, C), it does not satisfy the second condition.

The Battle of the Sexes game in Table 2.1 is Pareto optimal with Pareto optimal profiles, (C, D) and (D, C). Choosing Sides ($\sigma = \delta = 1$) and Common Interest games ($\sigma = 2, \delta = 1$) in Table 2.2 are also Pareto optimal. In the Choosing sides game, both diagonal profiles are Pareto optimal whereas in the Common Interest game only (C, C) is Pareto optimal.[31]

Proposition 14. *In Pareto optimal games, satisficing players in the long run coordinate on Pareto optimal profiles.*

This result can be interpreted as how convention, or equilibrium selection, arises when there does not exist a mediator and players do not have proper knowledge about the environment. Though satisficing players do not explicitly form a belief about how the game is played, in the long run they find acceptable actions and values of the game.

This interpretation makes better sense when we consider the interaction within a large population. Suppose a I-player Pareto optimal game is played by players who are each period randomly drawn from I separate, finite populations. Each population corresponds to each role in the game and consists of finite satisficing players, and players are selected to play the game with fixed probabilities. As in the two-player games, only matched players receive payoffs from their action profile and update their states while the others' states stay the same to the next period. Proposition 14 is generalised into this random matching environment.

[31]Chasparis et al. (2013) define a class of coordination games, which includes Network Formation Games and Common-Pool Games, with a bit different conditions. It can be easily verified that the coordination games defined in Chasparis et al. (2013) are also games of Pareto optimality in the current definition. For details about Network Formation Games, see Bala and Goyal (2000).

Proposition 15. *If finite satisficing players are randomly matched to play Pareto optimal games, the whole population in the long run coordinate on Pareto optimal profiles.*

A similar result has been shown by Young (1993), who also considers environments in which players are randomly matched from a large finite population. He finds that *adaptive play*, in which decision makers choose best actions given history of plays like fictitious play but based on finite samples of past plays rather than entire history, in coordination and common interest games selects *stochastically stable equilibria*, which are equivalent to *risk dominant equilibria* in the terminology of Harsanyi and Selten (1988) in 2×2 games.

2.4.3. Unilaterally Competitive Games

Lastly, we analyse unilaterally competitive games when $\overline{\lambda}$ is given constant, i.e., satisficing levels are moderately persistent. Since these games include Prisoner's Dilemma as a special case, this analysis helps understand the role of $\overline{\lambda}$ in the satisficing behaviour.

A useful property of unilaterally competitive games is provided in Lemma 2 without proof, which is given in De Wolf (1999). In short, if a satisficing decision maker chooses one of equilibrium actions, her equilibrium payoff is guaranteed regardless of opponents' choices.

Lemma 2. *In a unilaterally competitive game* Γ*, any two Nash profiles* $a', a'' \in A$ *satisfy* $\pi_i(a') = \min_{a_{-i}} \pi_i(a'_i, a_{-i}) = \pi_i(a'')$ *for all* i*.*

We consider two more properties that unilaterally competitive games might satisfy. Let $a^* \in A$ be an arbitrary Nash profile. (A.1) Any unilateral deviation from a Nash equilibrium returns the same payoff to the deviator, i.e., for any i, $\pi_i(a_i, a^*_{-i}) = \pi_i(a'_i, a^*_{-i}) < \pi_i(a^*)$ if $a_i, a'_i \neq a^*_i$. (A.2) A single tremble of one player's (and, possibly, finite subsequent repetitions of the game without trembles) can make any player receive a lower-than-Nash payoff with positive probability, i.e., for any ϕ_0 and i, $QP^m(\phi_0, B) > 0$ for some $m \in \mathbb{N}$ and $B \subset \Phi$ with $\pi_i(a') < \pi_i(a^*)$ for any $\phi' \in B$.

These somewhat strong conditions are sufficient for more than two players to converge to Nash equilibria in unilaterally competitive games when their satisficing levels are moderately persistent.

Proposition 16. *If a unilaterally competitive game satisfies (A.1) and (A.2), satisficing players in the long run play Nash equilibria for any* $\overline{\lambda} > 0$.

This result is obtained by Proposition 12 and the observation that, if the persistence parameter $\overline{\lambda}$ is fixed constant, payoff dominant profiles in two-player games are not robust to trembles any more: Regardless of how high agents' satisficing levels are, if agents repeatedly experience trembles in bad actions, their satisficing levels get sufficiently low to support a Nash profile, and this transition is not reversible as Nash profiles are robust to trembles.

De Wolf (1999) provides examples of unilaterally competitive games. As example, we provide two games that satisfy the hypothesis of Proposition 16.

The first is a public good provision game in which public good is produced by means of private contribution by 3 agents $i = 1, 2, 3$. Each agent decides whether to contribute a fixed amount of effort, say $a_i = 1$, to the public good or not, $a_i = 0$. The total amount of the public good provided is $a_1 + a_2 + a_3$ and agents' payoffs are given as $\pi_i(a_1 + a_2 + a_3) = \alpha(a_1 + a_2 + a_3) - \beta a_i$ for all i with $0 < \alpha < \beta$. Then, whenever an agent increases her effort from 0 to 1, her payoff decreases while all others' payoffs increase. And, whenever an agent decreases her effort from 1 to 0, her payoff increases while all others' decrease. Thus, this game is unilaterally competitive and has a unique Nash equilibrium, in which $a_i = 0$ for all i.[32]

The second is a price competition model in which a continuum of customers, the measure is normalised at 1, purchase a unit (per consumer) of a homogeneous good. The good is sold by three producers. Each producer has to select a price at which she wants to sell a unit of good. Only three different prices are available, 89, 95 and 99. The consumers buy the good from the producers who offer the minimum price. If more than one producer offer the same minimum price, equal number of consumers purchase from the producers. The producers' payoff matrix is given in Table 2.3. This game's unique Nash equilibrium is $(89, 89, 89)$.

[32]Simulation of public good provision game in a richer setup with bigger action sets and more players is given in Appendix C.4.

	89	95	99	
89	29.7, 29.7, 29.7	44.5, 0, 44.5	44.5, 0, 44.5	
95	0, 44.5, 44.5	0, 0, 89	0, 0, 89	89
99	0, 44.5, 44.5	0, 0, 89	0, 0, 89	

	89	95	99	
89	44.5, 44.5, 44.5	89, 0, 0	89, 0, 0	
95	0, 89, 0	31.7, 31.7, 31.7	47.5, 0, 47.5	95
99	0, 89, 0	0, 47.5, 47.5	0, 0, 95	

	89	95	99	
89	44.5, 44.5, 0	89, 0, 0	89, 0, 0	
95	0, 89, 0	47.5, 47.5, 0	95, 0, 0	99
99	0, 89, 0	0, 95, 0	33, 33, 33	

Table 2.3. Payoffs of Price Competition Model

In two-player games, the hypothesis of Proposition 16 can be loosened while the result still holds. When only two players interact, the competitiveness is required not for the game itself but for only Nash profiles. In the following two corollaries, we consider two-player games with unilaterally competitive Nash equilibria.

Corollary 1. *(i) If a two-player game satisfies (A.1), satisficing players in the long run play unilaterally competitive Nash equilibria.*
(ii) In any 2×2 *game, satisficing players in the long run play unilaterally competitive Nash equilibria.*

The convergence to only unilaterally competitive Nash equilibria in 2×2 games is in contrast with Proposition 12, in which satisficing players converge to either payoff dominant or unilaterally competitive Nash profiles as $\overline{\lambda}$ approaches 0.[33]

This contrast is better illustrated by a simple Markov process that consists of two state, C and D, and a transition probability P: $P(C|C) = 1 - p$, $P(D|C) = p$, $P(C|D) = 0$ and $P(D|D) = 1$ for some $p \in [0, 1]$, where $P(D|C)$ denotes the probability with which the state changes from C to D. This process has a unique invariant p.m. $\mu = (0, 1)$ for some positive p, but as $p \to 0$, its invariant p.m. of the limit process is given as $\mu = (1 - q, q)$ for any $q \in [0, 1]$. The process with fixed p is analogous to the satisficing behaviour when $\overline{\lambda}$ is fixed positive so that players who are satisficing with a payoff dominant profile can switch and settle with the

[33]Beggs (2005) finds that reinforcement learning of Erev and Roth (1998) converges to expected payoff maximising action in individual decision problems and Nash outcome in the Prisoner's Dilemma game.

Nash profile with positive probability.[34] In this process, only the Nash profile is played. On the other hand, the limit process as $p \to 0$ is analogous to the satisficing behaviour in which payoff dominant profiles are robust to trembles as $\overline{\lambda} \to 0$. In this process, both profiles are played.

The simulation results in appendices and the dominance of WSLS strategy with fixed aspiration levels are consistent with these results.

2.5. Conclusion

We separated payoff valuation (expected payoff of action) and satisficing level (expected payoff of the best outside option) in satisficing behaviour. This modification makes it possible for us to analyse satisficing behaviour in individual decision problems and normal form games in a unified way. We also allow agents to tremble in two ways, choice of actions and satisficing levels. Then, we analyse its long run behaviour as trembling probability and persistence parameter decline over time.

We find that in individual decision problems with stationary but stochastic payoffs, satisficing behaviour results in cautious, maximin choice. And, somewhat surprisingly, the same behaviour results in mutual cooperation as well as defection in the Prisoner's Dilemma and Stag Hunt games. Though cooperation is not the maximin action for a single player, satisficing agents cooperate as their satisficing levels are co-evolving. This result applies to a broad class of two-player games that have unilaterally competitive Nash equilibria. And, in Pareto optimal games, which include Battle of the Sexes, Common Interest and Choosing Sides games, satisficing players coordinate on Pareto optimal profiles. On the other hand, when the persistence parameter is fixed, the players play only Nash equilibrium in the Prisoner's Dilemma game.

[34] The fixed $\overline{\lambda} > 0$ makes the first argument in the proof of Proposition 12 invalid.

CHAPTER 3

Satisficing Behaviour in Extensive Form Games

This chapter models satisficing behaviour, proposed in the previous chapter, in extensive form games. In satisficing behaviour, decisions between continuing with the current actions and switching to alternatives are made by comparing valuations to satisficing levels. We apply this behaviour principle to environments in which a single player might have more than one chances, or decision nodes, to choose actions within a single decision problem. At all decision nodes, players evaluate their choices by received payoffs and judge them with respect to satisficing levels.

The satisficing behaviour in extensive form games is the same as in normal form games except that we additionally need to describe off-the-path behaviour in the former case. In normal form games, decision makers have only one decision node so that every period their choices, valuations and satisficing levels are updated by the realised outcomes. In extensive form games, however, decision makers' states at decision nodes off the path are not subject to period by period update.[1]

Compared to individual decision problems and normal form games, extensive form games have hardly been analysed with adaptive learning models. To our knowledge, in particular, satisficing behaviour in extensive form games is first formally modelled and analysed here.[2] A few works of adaptive learning in extensive form games are as follows. Hendon et al. (1996) and Groes et al. (1999) find that fictitious play converges to subgame perfect or sequential equilibrium paths, and Blume and

[1]However, we assume satisficing players experience shocks on choice of action and satisficing level at all decision nodes off the path as well as on the path. This assumption turns out to exclude unreasonable behaviours off the path. As Fudenberg and Levine (1993) noted, in extensive form games, which outcome is played depends not only on the actions on the path but also the planned actions off the path. Since Nash equilibrium does not provide precise predictions in extensive form games, several equilibrium concepts like subgame perfection and sequential equilibrium were introduced with purpose of imposing reasonable restrictions on players' strategies and beliefs off the equilibrium path.

[2]Kim (1995) models "satisficing behaviour" proposed by Gilboa and Schmeidler (1995) and Gilboa and Schmeidler (1996) in extensive form games. Though it is also modelled with endogenously evolving aspiration level, players in the model choose actions that are associated with highest values compared to aspiration level rather than satisfice with "satisfactory" actions. The second chapter of this dissertation discusses the difference between aspiration and satisficing levels.

Arnold (2004) find that fictitious play, defined as backward looking learning rule, with long memory facilitates truthful communication in a sender-receiver game with same interests. On the other hand, Jehiel and Samet (2005), Laslier and Walliser (2005) and Jehiel and Samet (2007) analyse reinforcement learning in extensive form games and find that agents converge to subgame perfect equilibrium paths in perfect information games. Roth and Erev (1995) numerically analyse a cumulative reinforcement learning model for normal form representations of Best Shot and Ultimatum games.

We examine extensive form games such as perfect information and a special class of imperfect information games, repeated games with observed actions and signalling games.

In perfect information games, we introduce a refinement of subgame perfection, *subgame dominance*. At each decision node along a subgame dominant path, every player chooses a best action that returns the highest payoff among all outcomes that follow the decision node. Every subgame dominant path is supported by a subgame perfect equilibrium, but the opposite does not generally hold true. Satisficing players play subgame dominant paths most of the time. And, in 2×2 games with outside options, satisficing behaviour is consistent with what Forward Induction predicts.

And, we analyse finitely repeated 2×2 games.[3] In finitely repeated Prisoner's Dilemma game, 'always cooperate' becomes the only behaviour taken by satisficing players if the payoff from mutual defection is worse enough compared to the payoff from mutual cooperation. This result suggests how a game is perceived by players plays a critical role in predicting outcomes. In finitely repeated Battle of the Sexes game, not only coordination is achieved but also the benefit from coordination is equally shared between players if payoffs from two coordination outcomes are balanced.

Lastly, we analyse two-player signalling games in which both players' payoffs are perfectly aligned: Regardless of the realised type of the first mover, both players' payoffs are the same and the common payoff is determined by the second mover's

[3]Players repeatedly play the finitely repeated games for infinite periods. For example, if satisficing players face a twice repeated Prisoner's Dilemma game, they make choices as if they recognise whether they are at the first decision node, whether they have played mutual cooperation in the first stage, and so on within each period, but do not know which outcome was played in the previous periods.

action. One interpretation of the game is whether communication via cheap talk could arise spontaneously between players who benefit from common understanding of the realised state of the world. We find that satisficing players in the long run develop languages through which they fully communicate.

Section 1 introduces notations in extensive form games and formulate satisficing behaviour. Section 2 defines subgame dominance and characterises satisficing behaviour in perfect information. Section 3 analyses 2×2 games with outside options. Section 4 and 5 analyse repeated games and signalling games, respectively. Section 6 summarises. All proofs are collected in Appendix A.3.

3.1. The Model

Players repeatedly face an extensive form game $\mathcal{G}(\overline{\mathcal{I}}, \mathcal{N}, a, h, \iota, \pi)$ over periods. $\overline{\mathcal{I}} = \{0, 1, \ldots, I\}$ is the set of indices of players that consist of Nature, indexed by 0, and non-Nature players, $1, \ldots, I$. In particular $\mathcal{I}$ denotes the set of non-Nature players. $\mathcal{N}$ is the collection of non-terminal and terminal nodes, $\mathcal{N}^n$ and $\mathcal{N}^t$, respectively. The nodes are partially ordered by an immediate predecessor function $p : \mathcal{N} \to \mathcal{N} \cup \emptyset$. There exists one node $n^0 \in \mathcal{N}$ such that $p(n^0) = \emptyset$, called the root node. We denote a set of immediate successors of n by $s(n) = p^{-1}(n)$ and $n' \in s(n)$ implies that n' follows n, that is, $n = p(n')$. Each node except the root is named by an action function $a : \mathcal{N} \backslash \{n^0\} \to A$, which is one to one on $s(n)$ for each non-terminal node n. $c(n) \equiv \{a \in A | a = a(n') \text{ for some } n' \in s(n)\}$ is the choice set at n, and $h : \mathcal{N}^n \to \mathcal{H}$ is an information function such that for all $n, n' \in \mathcal{N}^n$, $h(n) = h(n')$ implies $c(n) = c(n')$, where $\mathcal{H}$ is a collection of information sets that partition $\mathcal{N}^n$. We denote $c(n)$ by $c(h)$ if $h = h(n)$. A function $\iota : \mathcal{H} \to \overline{\mathcal{I}}$ indicates the player who moves at h. $\mathcal{H}_i \equiv \{h \in \mathcal{H} | \iota(h) = i\}$ is the set of information sets at which player i moves. For each non-terminal node n, $\iota(n)$ and $\iota(h(n))$ are used interchangeably. $\mathcal{G}$ has *perfect recall.*[4] $\pi \equiv \{\pi_i\}_{i \in \mathcal{I}}$ is a payoff function that maps $\mathcal{N}^t$ to $\mathbb{R}^I$ for non-Nature players, $1, \ldots, I$. The set of payoffs that player i would receive in the game is denoted by $\pi_i \equiv \{\pi_i(n) | n \in \mathcal{N}^t\}$. Sets of terminal nodes that follow n and h are denoted by $\mathcal{N}^t(n)$ and $\mathcal{N}^t(h) = \cup_{n \in h} \mathcal{N}^t(n)$, respectively.

[4]Players remember any information they once knew, and actions they have chosen previously. That is, (i) if $h(n) = h(n')$, then $n \neq p(n')$ and $n' \neq p(n)$ and (ii) if $h(n') = h(n'')$, $n = p(n')$ and $\iota(n) = \iota(n')$, then there exists a node n^* such that $h(n) = h(n^*)$, $n^* = p(n'')$ and $a(n') = a(n'')$.

A behaviour of player i in the game is represented by a function $b_i : \mathcal{H}_i \to A$ such that $b_i(h) \in c(h)$ for all $h \in \mathcal{H}_i$ and $i \in \overline{\mathcal{I}}$. Nature has a probability distribution $\nu : \mathcal{H}_0 \times A \to [0, 1]$ such that $\nu(h, a) = 0$ if $a \notin c(h)$ and $\sum_{a \in c(h)} \nu(h, a) = 1$ for all $h \in \mathcal{H}_0$. Nature's behaviour b_0 is chosen at random from the stationary distribution at each repetition of the game. B_i denotes the set of all possible behaviours of player $i \in \overline{\mathcal{I}}$ with $\overline{B} \equiv \Pi_{i \in \overline{\mathcal{I}}} B_i$ for all players and $B \equiv \Pi_{i \in \mathcal{I}} B_i$ for non-Nature players. A set of nodes is called a *path* if it can be ordered, say $n_0, n_1, \ldots, n_J$, so that n_0 is the root, n_J is a terminal node and n_j follows n_{j-1} for all $j = 1, \ldots, J$. A behaviour profile $b = (b_0, b_1, \ldots, b_I) \in \overline{B}$ specifies a path $\mathcal{N}(b) \subset \mathcal{N}$ with a terminal node $\mathcal{N}^t(b)$, a set of non-Nature players who have turns to move along the path $\mathcal{I}(b) \equiv \{i \in \mathcal{I} | i = \iota(n)$ for some $n \in \mathcal{N}(b)\}$ and a collection of information sets through which the path passes $\mathcal{H}(b) \equiv \{h \in \mathcal{H} | h = h(n)$ for some $n \in \mathcal{N}(b)\}$ with $\mathcal{H}_i(b) = \mathcal{H}(b) \cap \mathcal{H}_i$. A set of payoffs from b is denoted by $\pi(b) \equiv \pi(\mathcal{N}^t(b))$. With explicit notation of a path $p = \mathcal{N}(b)$, we denote these sets by $\mathcal{N}^t(p), \mathcal{I}(p), \mathcal{H}(p), \mathcal{H}_i(p)$ and $\pi(p)$.

3.1.1. Satisficing without Trembling

Non-Nature players satisfice rather than optimise in the extensive form games. In period n, player i's state is described by her current behaviour $b_{n,i}$, valuations $v_{n,i} \in \mathbb{R}^{|\mathcal{H}_i|}$ and satisficing levels $s_{n,i} \in \mathbb{R}^{|\mathcal{H}_i|}$, where $v_{n,i}$ and $s_{n,i}$ map $\mathcal{H}_i$ to $\mathbb{R}^{|\mathcal{H}_i|}$. Shortly, $\phi_{n,i} = (b_{n,i}, v_{n,i}, s_{n,i})$ is used for each $i \in \mathcal{I}$ and the I players' states are denoted by $\phi_n = \{\phi_{n,i}\}_{i \in \mathcal{I}}$ with $b_n = \{b_{n,i}\}_{i \in \mathcal{I}}$. ϕ_0 is assumed given.

At the start of any period $n \in \mathbb{N}$, each player i judges her current actions $b_{n,i}(h)$ by comparing its valuation $v_{n,i}(h)$ with her ongoing satisficing level $s_{n,i}(h)$ at each information set h. If valuation $v_{n,i}(h)$ is greater than satisficing level $s_{n,i}(h)$, she satisfices and continues with the current action and satisficing level at h. If the valuation falls short of the satisficing level, she switches to an alternative and updates her satisficing level towards the lower valuation. Specifically, if $i \notin \mathcal{I}(b_n)$, $b_{n+1,i} = b_{n,i}$ and $s_{n+1,i} = s_{n,i}$. If $i \in \mathcal{I}(b_n)$, for $h \in \mathcal{H}_i(b_n)$ such that $v_{n,i}(h) < s_{n,i}(h)$,

$$b_{n+1,i}(h) = \alpha_{n,i}(h) \quad \text{and} \quad s_{n+1,i}(h) = (1 - \lambda_{n,i}(h))s_{n,i}(h) + \lambda_{n,i}(h)v_{n,i}(h),$$

where $\alpha_{n,i}(h)$ is chosen at random from $c(h)$ with positive probabilities for all alternatives in the choice set, and $\lambda_{n,i}(h)$ is from $[0, \overline{\lambda}]$. For all other $h \in \mathcal{H}_i$ for $i \in \mathcal{I}(b_n)$,

$b_{n+1,i}(h) = b_{n,i}(h)$ and $s_{n+1,i}(h) = s_{n,i}(h)$. Each individual, then, chooses the behaviour $b_{n,i}$ so determined and receives a payoff $\pi_i(b_{n+1})$. Before the end of period n, the players revise their valuations by taking weighted averages of their payoffs and previous valuations as follows. If $i \notin \mathcal{I}(b_{n+1})$, $v_{n+1,i} = v_{n,i}$. For $h \in \mathcal{H}_i(b_{n+1})$ with $i \in \mathcal{I}(b_{n+1})$, if $b_{n+1,i}(h) = b_{n,i}(h)$,

$$u_{n+1,i}(h) = (1 - \rho_{n,i}(h))v_{n,i}(h) + \rho_{n,i}(h)\pi_i(b_{n+1}),$$

where $\rho_{n,i}(h) \in [0, 1]$ and otherwise $v_{n+1,i}(h) = \pi_i(b_{n+1})$. For $h \notin \mathcal{H}_i(b_{n+1})$ with $i \in \mathcal{I}(b_{n+1})$, $v_{n+1,i}(h) = v_{n,i}(h)$.

In the above $\alpha_{n,i}, \rho_{n,i}$ and $\lambda_{n,i}$ at each information set are random variables with corresponding measures $\mu_{\alpha,i}, \mu_{\rho,i}$ and $\mu_{\lambda,i}$ with full supports in $c(h)$,$[0, 1]$ and $[0, \overline{\lambda}]$ for some $\overline{\lambda} \in (0, 1)$, respectively, for all i. In particular, the distribution of the next period action μ_α is referred to as *choice rule* and assumed to be continuous functions of $v_{n,i}$ and $s_{n,i}$ so that at all h, $\mu_{\alpha,i}(a) \downarrow 0$ as $v_{n,i} \uparrow s_{n,i}$ for all $a \in c(h)\backslash\{b_{n,i}(h)\}$ and $\mu_{\alpha,i}(a_{n,i})$ is bounded away from zero for any $v_{n,i}$ below $s_{n,i}$. The other two probability distributions of ρ and λ reflect how players weigh recent information, on which we do not put any specific condition. And, the upper bound on the weighting coefficient on satisficing level $\overline{\lambda}$ is referred to as *persistence parameter* that determines how persistent players' satisficing levels are. As $\overline{\lambda}$ gets lower, satisficing levels are adjusted more slowly. Unless otherwise stated, the measures are assumed to be absolutely continuous with respect to the Lebesgue measure on the intervals and continuous in the current state ϕ_n.[5]

3.1.2. Trembling Behaviour

In case they tremble, satisficing players experience two types of shocks as in the second chapter with *trembling probability* ϵ. Given $\phi_{n,i}$, if player i experiences only the first shock, which occurs with probability $\epsilon(1 - \epsilon)$ at each information set in each period, her next period action at the information set is determined as follows. For all $i \in \mathcal{I}$ and $h \in \mathcal{H}_i$,

$$b_{n+1,i}(h) = \alpha^{\epsilon}_{n,i}(h),$$

[5]For more discussion about the roles of the random variables and assumptions on these, refer to the second chapter.

while her satisficing level and valuation at h are updated as before. $\alpha^{\epsilon}_{n,i}(h)$ is similarly defined as $\alpha_{n,i}(h)$. On the other hand, if the player experiences only the second shock at an information set, which occurs with probability $\epsilon(1-\epsilon)$, her next period satisficing level at the information set is adjusted towards the valuation. For all $i \in \mathcal{I}$ and $h \in \mathcal{H}_i$,[6]

$$s_{n+1,i}(h) = (1 - \lambda^{\epsilon}_{n,i}(h))s_{n,i}(h) + \lambda^{\epsilon}_{n,i}(h)v_{n,i}(h)$$

where $\lambda^{\epsilon}_{n,i}(h)$ is a random variable with probability measure $\mu^{\epsilon}_{\lambda,i}(h)$ with full support in $[0,1]$. Her action and its valuation are revised by the satisficing behaviour without trembling. Lastly, if she experiences both shocks at an information set h in one period, which occurs with probability ϵ^2, her state is updated as if only the first shock occurs, and then just before the next period starts, her satisficing level trembles to be adjusted towards valuations.

Let $V_i(h)$ and $S_i(h)$ both denote the convex hull of $\{\pi_i(n)|n \in \mathcal{N}^t(h)\}$ for all $i \in \mathcal{I}$ and $h \in \mathcal{H}_i$, and V and S the product spaces of $\Pi_{h\in\mathcal{H}_i}V_i$ and $\Pi_{h\in\mathcal{H}_i}S_i$ across all satisficing players, respectively. Then, the players' states governed by the satisficing behaviour with trembling constitutes a Markov process $\{\phi_n\}$ on the compact space $\Phi \equiv A \times V \times S$. Asymptotic properties of the Markov process are the same as in the second chapter. And, we also can augment the satisficing behaviour in extensive form games with additional memory by allowing player to keep track of valuations of all actions at each information set without changing the asymptotic behaviour.

3.2. Perfect Information

Here, we analyse satisficing behaviour in perfect information games without Nature playing a role, i.e., $\mathcal{G}(\mathcal{I}, \mathcal{N}, a, h, \iota, \pi)$ with $h\,(n) \neq h\,(n')$ for any $n \neq n'$. We first introduce a refinement of subgame perfection in terms of outcomes and prove the refined outcomes are played most of the time by satisficing players. A path p is said to be subgame perfect if there exists a subgame perfect equilibrium b of $\mathcal{G}$ such that $p = \mathcal{N}(b)$.

[6]Unlike behaviour without trembling, we assume that all agents experience shocks at all information sets even when they did not move in the previous round.

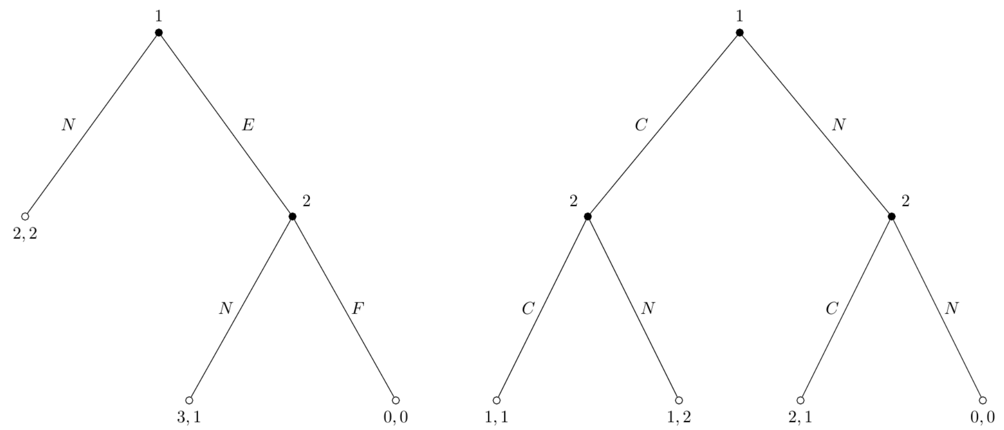


Figure 3.1. Subgame Perfect and Dominant Paths

Definition 10. *A path p is* weakly dominant *if for each $n \in p \cap \mathcal{N}^n$ with $\iota(n) \in \mathcal{I}$, $\pi_{\iota(n)}(p) \geq \max\{\pi_{\iota(n)}(n') | n' \in \mathcal{N}^t(n) \backslash \mathcal{N}^t(p)\}$. It is* subgame dominant *if all inequalities are strict.*

Along a subgame dominant path, no player can increase her payoff by deviation at any information set. If a game admits a subgame dominant path, it is unique but it might not exist in some games. Note that a subgame dominant path is also subgame perfect, but a subgame perfect path might not be subgame dominant.

Consider two extensive form games in Figure 3.1. The left game is a stage game of the Chainstore paradox. The first mover is the entrant and the second is the incumbent. There are three paths in the game: the entrant does not enter the market, the entrant enters and the incumbent fights, and the entrant enters and the incumbent does not fight, $(N, \cdot), (E, F)$ and (E, N), respectively. The game has two Nash paths (N, F) and (E, N) while there exists only one subgame perfect path (E, N). Among the two Nash equilibria, only the subgame perfect (E, N) is subgame dominant.

The other game is a simplified Best Shot game. The first mover determines whether to contribute to produce a public good, then observing the move by the first, the second chooses whether to contribute. The public good is produced when at least one player contributes. Contribution is equally costly for both players. Two paths (C, N) and (N, C) are supported by Nash equilibria while only (N, C) is subgame perfect and subgame dominant.

Next, consider extensive form games that have a subgame perfect but not dominant paths. In Figure 3.2, the left game has two Nash paths (C, N) and (N, C) but only (N, C) is subgame perfect. Unlike the games in Figure 3.1, the subgame perfect path (N, C) is not subgame dominant because the first mover can receive the higher payoff, 3, from a non-equilibrium path (C, C). The Centipede game in the right panel has a unique subgame perfect equilibrium in which the first mover takes all stake and ends the game at the first decision node, but the equilibrium path is not subgame dominant.

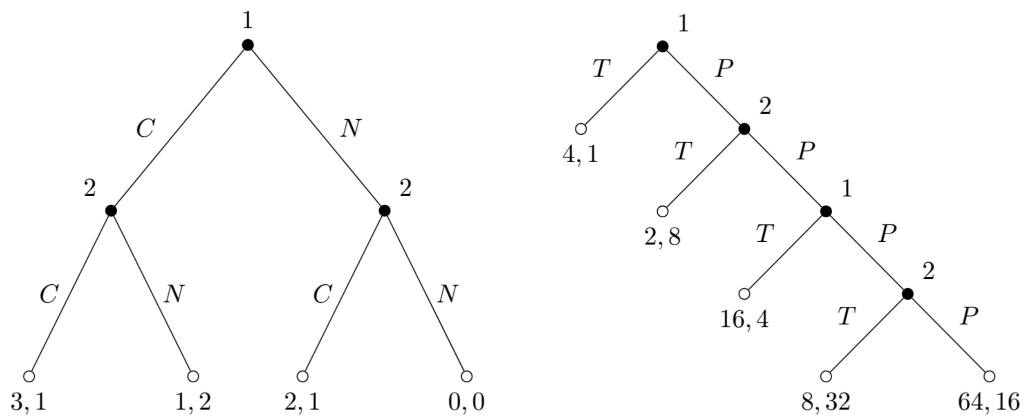


Figure 3.2. Subgame Perfect but not Dominant Paths

The first result of this chapter is that satisficing players will eventually play subgame perfect outcomes only when it is also subgame dominant. This predicts that subgame perfect but not dominant outcomes will be less likely observed than subgame dominant ones if players satisfice rather than optimise.

Proposition 17. *If a perfect information game has a subgame dominant path, satisficing players in the long run play the path.*

As in the second chapter, satisficing agents' long run behaviours derived in the current chapter are given in the limit as trembling probability ϵ and persistence parameter $\overline{\lambda}$ tend to 0 in the order specified previously.[7]

The proof can be sketched by simple backward induction like solving subgame perfect equilibrium. Movers at the last non-terminal nodes of any path would choose actions which return highest payoffs and then raise their satisficing levels up to the payoffs. Then movers at the second last non-terminal nodes would choose actions

[7]For more related discussion, refer to the note after Proposition 11.

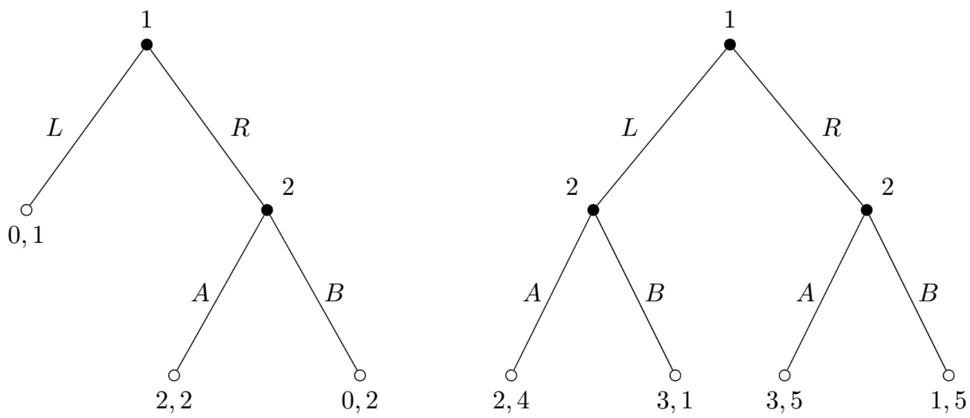


Figure 3.3. Weakly Dominant Paths

which return highest payoffs among the payoffs determined by the previous movers and raise their satisficing levels up to the payoffs. If this process is rolled up to the first non-Nature player, all players conform to a subgame dominant path. And, by the property of the path, it is robust to single trembles. If only weakly dominant paths exist, satisficing players play those paths with positive probabilities.

In the stage game of Chainstore paradox in Figure 3.1 if the entrant and incumbent are satisficing players, the entrant will enter and the incumbent will accommodate it. And, in the Best Shot game, the first mover does not contribute while the second contributes. These two predictions are consistent with experimental results of Schotter et al. (1994) and Harrison and Hirshleifer (1989).

It is noteworthy why satisficing players need not converge to subgame perfect paths. In the Centipede game in Figure 3.2, suppose players have settled in with the equilibrium path with satisficing levels (4,1). While the first mover satisfices with 'Take' at her first turn, the second (resp. the first) mover could tremble to "plan" to choose 'Pass' (resp. 'Take') at his first (resp. her second) turn. Then, once the first mover trembles to 'Pass', outcome (16,4) is reached and both players satisfice.[8]

If a subgame dominant path does not exist, both dominant and non-dominant paths could be played by satisficing players. Consider the game in Figure 3.3, which

[8] In the long run, the second player's satisficing level at the information set is strictly higher than 4 with probability 1 by the trembling behaviour. However, when $\overline{\lambda}$ is fixed above 0, the players could become to satisfice with the path as we assume the probability of choosing the current action converges to 1 as the valuation approaches satisficing level. For more details, see Karandikar et al. (1998). Since we assume the weighting coefficient are random variable here, we require additional assumption on the infinite sum of moments of λ. For experimental results, see McKelvey and Palfrey (1992) and Palacios-Huerta and Volij (2009).

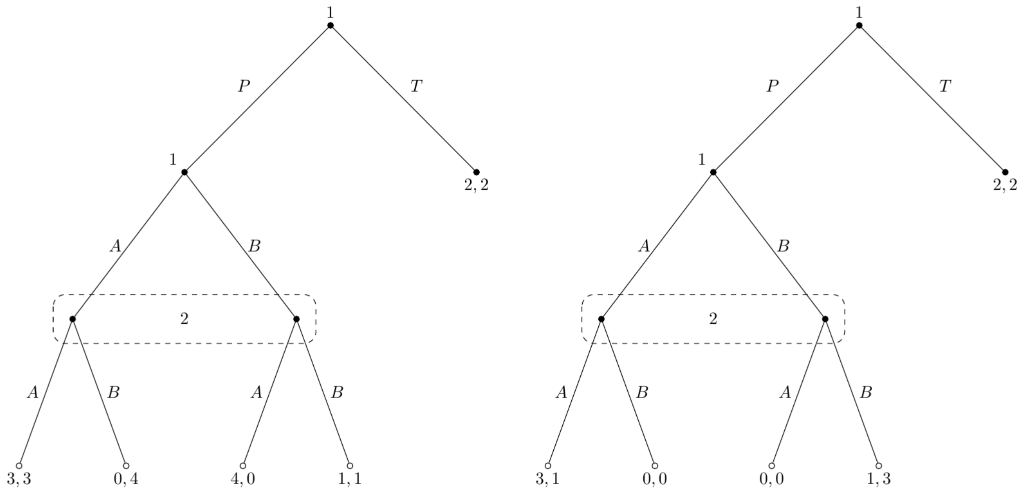


Figure 3.4. Dominant Paths of Imperfect Information Game

has one weakly dominant path (R, A). Satisficing players will play not only the dominant path but also (L, A) with positive probabilities in the long run: satisficing with (R, A) with satisficing level $(3, 5)$, once player 2 trembles to choose B, player 1 keeps switching between L and R until he satisfices with a non-dominant path (L, A). Appendix C.5 provides simulation results of Best Shot and Ultimatum games.

3.3. Games with Outside Option

In Figure 3.4, player 1 is given an *outside option* which guarantees a fixed payoff, and if he chooses the option, the game ends, and otherwise a 2×2 game follows. The Subgame Perfection predicts that both $((P, A), A)$ and $((T, B), B)$ (and an additional equilibrium with a mixed equilibrium in the subgame) would be played whereas Forward Induction predicts that since the choice of P by player 1 signals that he expects payoff more than 2, both players in the subgame would play only (A, A).[9]

The following result shows how satisficing players behave in such games. Without loss of generality, we assume outside options return intermediate payoffs to both players: e.g., between the payoffs from two coordination outcomes in the Battle of the Sexes or from the mutual cooperation and defection in the Prisoner's Dilemma. If the guaranteed payoffs are above the ranges, satisficing players will always play the outside options. And, if below the ranges, the options do not affect their choices.

[9]The right-hand game is used by Kohlberg and Mertens (1986) to introduce the idea of forward induction. Brandts and Holt (1994) provide experiment results in the Battle of the Sexes with an outside option.

Proposition 18. *In the Battle of the Sexes game with an outside option, satisficing players in the long run coordinate so that player 1 receives the higher payoff than the option.*

Proposition 19. *In the Prisoner's Dilemma with an outside option, satisficing players in the long run play the cooperate only profile.*

Without the outside option, players would play both mutual cooperation and defection in the Prisoner's Dilemma. However, once player 1 is given the option of securing a payoff higher than mutual defection, the defection outcome becomes less likely to be played because otherwise player 1 takes the option. And, once the option is chosen by player 1, nothing binds players to the mutual defection outcome, thus both players can agree on the mutual cooperation with finite trembles off the path.

3.4. Finitely Repeated Games

Next, we consider repeated games with observed actions. Payoffs are simply summed up over stages without discount. The payoff matrix of a Prisoner's Dilemma game is given in Table 3.1. For simplicity, the payoffs are given by two parameters b and c with $b > c > 0$, for mutual cooperation and defection, respectively. A single defector receives $b + c$ while leaves the other player 0. If players 'always cooperate' over two stages, both receive $2b$. If players alternate between mutual cooperation and defection, like (C, C) in the first stage and (D, D) in the second or (D, D) in the first and (C, C) in the second, both receive $b + c$. If player 1 always cooperates and player 2 always defects, player 1 receives 0 and player 2 receives $2b + 2c$, which is the highest payoff that a player can achieve in the twice repeated game.

	C	D
C	b, b	$0, b + c$
D	$b + c, 0$	c, c

Table 3.1. Prisoner's Dilemma

When satisficing players repeat the stage Prisoner's Dilemma game, the second chapter shows that both mutual cooperation and defection are played. However, if they repeat a multi-stage game in which an earlier action affects players' perception in later stages, cooperation over all stages could arise as a unique outcome. The following result shows that cooperation can take place if the benefit from mutual cooperation is significantly higher than mutual defection.

Proposition 20. *In the n-times repeated Prisoner's Dilemma game, satisficing players in the long run 'always cooperate' if* $nc \leq b$.

The condition $nc \leq b$, which is derived from $(n-1)b + nc \leq nb$, implies that once both players satisfice with the 'always cooperate' path and their satisficing levels are pushed up to the corresponding payoffs (nb, nb), none of them satisfices with any path that starts with mutual defection in the first stage.[10] In particular, if this condition holds in the twice repeated game, 'always cooperation' returns higher payoffs to both players than a single defection followed by mutual defection.

The other repeated game is the Battle of the Sexes. The payoff matrix is given in Table 3.2. The payoffs are also determined by two parameters b and c for coordinated outcomes with $b > c > 0$ and 0 for coordination failure. If players alternate between two coordinated outcomes, like (F, F) in the first stage and (O, O) in the second or (O, O) in the first and (F, F) in the second, both receive $b + c$. If players repeat one coordinated outcome twice, one player receives the highest payoff $2b$ and the other receives $2c$, which is still better than that from coordination failures.

	F	O
F	b, c	$0, 0$
O	$0, 0$	c, b

Table 3.2. Battle of the Sexes

In the stage game, satisficing players converge to one of two coordination outcomes, which favours one player more than the other. That is, the benefit from coordination is not equally shared between two players. However, if the game is repeated over several stages, another behaviour may arise in which players share the benefit in a fair way like alternating between two coordination outcomes (F, F) and (O, O) as if a mediator advises so.

[10]If players keep choosing (C, D) in every subgame after (D, D) in the first stage and choose (C, C) (resp. $(C, D), (D, C)$ and (D, D)), they receive $(b + c, (n-1)b + (n-1)c)$ (resp. $(c, (n-1)b + nc)$, $(b + 2c, (n-2)b + (n-1)c)$ and $(2c, (n-2)b + (n-1)c)$). In this last subgame, player 1's satisficing level is greater than c. Thus, if the hypothesis of the proposition holds, whenever players reach this subgame, both do not satisfice at the same time.

Proposition 21. *In the $2n$-times repeated Battle of the Sexes, satisficing players in the long run equally play both coordination outcomes if $\frac{n+1}{n}c \leq b < \frac{n}{n-1}2c$.*

This result says that satisficing players fairly coordinate in the finitely repeated Battle of the Sexes game without any device or mediator if the payoffs from coordination outcomes are balanced. On the other hand, it is implied that, if the "balanced payoff" condition does not hold, one player could get the most of the benefit from coordination as they do when the stage game is infinitely repeated.

3.5. Signalling Games

A special class of imperfect information games, which we analyse here, are signalling games in which Nature chooses an action (or type) with fixed probabilities every period, player 1 (Sender) observes the type and chooses his action (or message), then player 2 (Receiver) observes the message taken by Sender but not the type chosen by Nature and chooses her action. Both Sender and Receiver's payoffs are the same at every terminal node and determined by whether the type chosen by Nature and the action by Receiver match each other or not.[11] If the type and action are matched like rain and umbrella (probably via Sender's message "it will rain"), both players receive higher payoffs than otherwise. We assume that Sender's (resp. Receiver's) choice sets are the same for each type (resp. message) and the numbers of elements of both players' choice sets are bigger than Nature's.

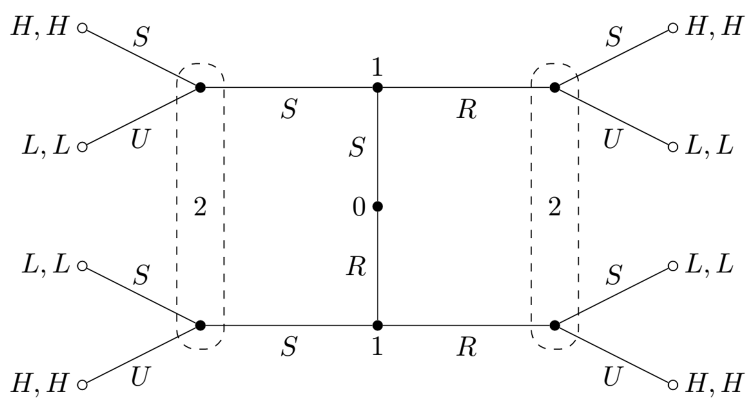


Figure 3.5. A Signalling Game with Two States

[11]When players' interests are not aligned, equilibrium strategies are normally given as mixed strategies. Appendix C.6 provides simulation results of signaling games between players whose interests misaligned.

Consider a simple signalling game in which each choice set consists of two elements depicted in Figure 3.5 and its normal form representation in Table 3.3. Nature chooses tomorrow's weather between *Rainy* and *Sunny*. Sender, the row player in the Table, chooses a message between R and S for each type, and Receiver, the column player, chooses an action between *Umbrella* and *Sunglasses* for each message. The left table is the payoff matrix when the type is *Rainy* and the other is when the type is *Sunny*. Both players can receive a higher payoff H regardless of the realised type if both coordinate on either $\{(R, S), (U, S)\}$ or $\{(S, R), (S, U)\}$, which are defined as *complete* profiles below. We show that satisficing players eventually develop complete communication strategies through which they achieve the highest payoffs for any realised type.

	$t = R$					$t = S$			
	U, U	U, S	S, U	S, S		U, U	U, S	S, U	S, S
R, R	H, H	H, H	L, L	L, L	R, R	L, L	L, L	H, H	H, H
R, S	H, H	H, H	L, L	L, L	R, S	L, L	H, H	L, L	H, H
S, R	H, H	L, L	H, H	L, L	S, R	L, L	L, L	H, H	H, H
S, S	H, H	L, L	H, H	L, L	S, S	L, L	H, H	L, L	H, H

Table 3.3. A Signalling Game with Two States

We formally describe the above signalling game $\mathcal{G}_S$. Let T, M and A denote the choice sets of Nature, Sender and Receiver, respectively. The cardinalities of M and A are not smaller than that of T. At the root node, Nature chooses a type $t \in T$ with positive probabilities for all types. Then, Sender observes the type and chooses a message $m \in M$. Lastly, Receiver observes only the message and chooses an action $a \in A$. There exists an injective function $f : T \to \mathcal{N}^t$ such that for any $t \in T$, $\pi_1(p) = \pi_2(p) = H$ if $t \in p$ implies $f(t) \in p$ and $\pi_1(p) = \pi_2(p) = L$ otherwise, where $H > L$. In words, for any type t, if Receiver chooses a unique corresponding action $a(f(t)) \in A$, both players receive H. Otherwise, both receive L. With a little abuse of notation. let Nature and Sender' actions represent the corresponding information set, then $t \in T$ represents an information set $h_t \equiv \{n \in \mathcal{N} | a(n) = t\}$ and $m \in M$ represents $h_m \equiv \{n \in \mathcal{N} | a(n) = m\}$. The complete behaviour profiles are defined as follows.

Definition 11. *A behaviour profile b of $\mathcal{G}_S$ is* complete *if $b_1 : T \to M$ and $b_2 : M \to A$ are injective functions such that $b_2(b_1(t)) = f(t)$ for any $t \in T$.*

The above signalling game with two types has two complete behaviour profiles $\{(R,S),(U,S)\}$ or $\{(S,R),(S,U)\}$. In general, if Nature chooses a state from N different types and Sender chooses a message from M different messages, there are $M!/(M-N)!$ complete behaviour profiles. However, under all complete behaviour profiles, players' payoffs are the same.

Proposition 22. *In the signalling games, satisficing players in the long run play complete profiles most of the time.*

As the normal form representation suggests, the way players develop a complete behaviour profile is similar to how players coordinate on Pareto optimal outcomes. For any realised type, if Sender employs a new message and Receiver happens to interpret the message correctly once, both satisfice and continue with the behaviour at the type and message. And, as the proof in Appendix implies, in the long run satisficing players are not stuck in a specific complete profile, rather they easily switch from one to other between complete profiles.

3.6. Conclusion

This chapter models the satisficing behaviour, proposed in the second chapter, in extensive form games. A player at each decision node is modelled as an independent agent: A player satisfices with actions with respect to satisficing levels at each decision node and experiences shocks independently across all decision nodes on and off the path.

In perfect information games, we introduce a refinement of subgame perfection, *subgame dominance*. Satisficing players play subgame dominant paths most of the time in perfect information games. And, we identify conditions under which satisficing players 'always cooperate' in repeated Prisoner's Dilemma games and 'fairly coordinate' in repeated Battle of the Sexes games. Lastly, we find that sender and receiver with same interests develop complete communication strategies in signalling games.

References

Bala, Venkatesh and Sanjeev Goyal, "A Noncooperative Model of Network Formation," *Econometrica*, 2000, *68* (5), 1181–1229.

Banks, Jeffrey S, Colin F Camerer, and David Porter, "An Experimental Analysis of Nash Refinements in Signaling Games," *Games and Economic Behavior*, January 1994, *6* (1), 1–31.

Battaglini, Marco, "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, July 2002, *70* (4), 1379–1401.

Battalio, Raymond C, Larry Samuelson, and John B Van Huyck, "Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games," *Econometrica*, May 2001, *69* (3), 749–764.

Beggs, Alan W, "On the convergence of reinforcement learning," *Journal of Economic Theory*, May 2005, *122* (1), 1–36.

Bendor, Jonathan, Dilip Mookherjee, and Debraj Ray, "Aspiration-Based Reinforcement Learning in Repeated Interaction Games: An Overview," *International Game Theory Review*, 2001, *3* (2-3), 159–174.

______, Sunil Kumar, and David A. Siegel, "Satisficing: a 'Pretty Good' Heuristic," *B.E. Journal of Theoretical Economics*, 2009, *9* (1), Art. 9–38.

Billingsley, Patrick, *Convergence of Probability Measures*, Wiley, 1999.

Blume, Andreas and Tone Arnold, "Learning to communicate in cheap-talk games," *Games and Economic Behavior*, 2004, *46* (2), 240–259.

Börgers, Tilman and Rajiv Sarin, "Naive reinforcement learning with endogenous aspirations," *International Economic Review*, 2000, *41* (4), 921–950.

Chakraborty, Archishman and Rick Harbaugh, "Comparative Cheap Talk," *Journal of Economic Theory*, 2007, *132* (1), 70–94.

______ and ______, "Persuasion by Cheap Talk," *American Economic Review*, 2010, *100* (5), 2361–2382.

Chasparis, Georgios C., Ari Arapostathis, and Jeff S Shamma, "Aspiration

Learning in Coordination Games," *Siam Journal on Control and Optimization*, January 2013, *51* (1), 465–490.

Chen, Ying, "Communication with Two-Sided Asymmetric Information," *SSRN eLibrary*, February 2009.

______ **and Wojciech Olszewski**, "Effective Persuasion," *SSRN eLibrary*, February 2011.

______**, Navin Kartik, and Joel Sobel**, "Selecting Cheap-Talk Equilibria," *Econometrica*, 2008, *76* (1), 117–136.

Cho, In-Koo and Akihiko Matsui, "Learning Aspiration in Repeated Games," *Journal of Economic Theory*, 2005, *124* (2), 171–201.

Crawford, Vincent and Joel Sobel, "Strategic Information Transmission," *Econometrica*, 1982, *50* (6), 1431–1451.

de Barreda, Inés Moreno, "Cheap Talk with Two-Sided Private Information," *Job Market Paper, London School of Economics and STICERD*, 2010.

De Wolf, O, "Optimal Strategies in n-Person Unilaterally Competitive Games," *Working Paper*, 1999.

Dimitrakas, Vassilios and Yianis Sarafidis, "Advice from an Expert with Unknown Motives," *mimeo, INSEAD*, 2005.

Ellison, Glenn, "Cooperation in the Prisoners-Dilemma with Anonymous Random Matching," *Review of Economic Studies*, June 1994, *61* (3), 567–588.

Erev, Ido and Alvin E Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, September 1998, *88* (4), 848–881.

Fudenberg, Drew and David K Levine, "Self-Confirming Equilibrium," *Econometrica*, May 1993, *61* (3), 523–545.

Gilboa, Itzhak and David Schmeidler, "Case-Based Decision Theory," *Quarterly Journal of Economics*, 1995, *110* (3), 605–639.

______ **and** ______, "Case-Based Optimization," *Games and Economic Behavior*, 1996, *15* (1), 1–26.

______ **and** ______, *A Theory of Case-Based Decisions*, Cambridge University Press, July 2001.

Goeree, Jacob K, Charles A Holt, and Thomas R Palfrey, "Risk averse behavior in generalized matching pennies games," *Games and Economic Behavior*, September 2003, *45* (1), 17–17.

Groes, Ebbe, Hans Jørgen Jacobsen, and Birgitte Sloth, "Adaptive learning in extensive form games and sequential equilibrium," *Economic Theory*, 1999, *13* (1), 125–142.

Harrison, Glenn W and Jack Hirshleifer, "An Experimental Evaluation of Weakest Link/Best Shot Models of Public Goods," *Journal of Political Economy*, February 1989, *97* (1), 201–225.

Harsanyi, John C and Reinhard Selten, *A General Theory of Equilibrium Selection in Games*, MIT Press, 1988.

Hendon, Ebbe, Hans Jørgen Jacobsen, and Birgitte Sloth, "Fictitious play in extensive form games," *Games and Economic Behavior*, 1996, *15* (2), 177–202.

Hernández-Lerma, Onésimo and Jean-Bernard 1953 Lasserre, *Markov Chains and Invariant Probabilities*, Basel ; Boston : Birkhäuser, 2003.

Imhof, Lorens A., Drew Fudenberg, and Martin A Nowak, "Tit-for-tat or win-stay, lose-shift?," *Journal of theoretical biology*, 2007, *247* (3), 574–580.

Isaac, R Mark, James M Walker, and Susan H Thomas, "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations," *Public Choice*, January 1984, *43* (2), 113–149.

Ishida, Junichiro and Takashi Shimizu, "Cheap Talk with an Informed Receiver," *ISER Discussion Paper, Institute of Social and Economic Research, Osaka University*, June 2011.

Jehiel, Philippe and Dov Samet, "Learning to play games in extensive form by valuation," *Journal of Economic Theory*, October 2005, *124* (2), 129–148.

——— **and** ———, "Valuation Equilibria," *Theoretical Economics*, June 2007, *2*, 163–185.

Karandikar, Rajeeva, Dilip Mookherjee, Debraj Ray, and Fernando Vega-Redondo, "Evolving Aspirations and Cooperation," *Journal of Economic Theory*, 1998, *80*, 292–331.

Kats, A and J F Thisse, "Unilaterally Competitive Games," *International Journal of Game Theory*, 1992, *21* (3), 291–299.

Kim, Youngse, "A Satisficing Model of Learning in Extensive Form Games," Queen Mary & Westfield College, University of London 1995.

———, "Satisficing and Optimality in 2 X 2 Common Interest Games," *Economic Theory*, 1999, *13* (2), 365–375.

Lai, Ernest K, "Expert Advice for Amateurs," *mimeo, Lehigh University*, August 2010.

Laslier, Jean-François and Bernard Walliser, "A reinforcement learning process in extensive form games," *International Journal of Game Theory*, 2005, *33* (2), 219–227.

Ledyard, John O, "Public goods: A survey of experimental research," in "Working Paper," EconWPA, 1994.

Li, Ming and Kristof Madarasz, "When Mandatory Disclosure Hurts: Expert Advice and Conflicting Interests," *Journal of Economic Theory*, 2008, *139* (1), 47–74.

March, James G, "Learning to be risk averse.," *Psychological Review*, April 1996, *103* (2), 309–319.

McKelvey, Richard D and Thomas R Palfrey, "An Experimental Study of the Centipede Game," *Econometrica*, July 1992, *60* (4), 803–836.

——— **and** ———, "Quantal Response Equilibria for Extensive Form Games," *Experimental Economics*, June 1998, *1* (1), 9–41.

Meyn, Sean and Richard L. Tweedie, *Markov Chains and Stochastic Stability* Communications and control engineering, second ed., Cambridge: Cambridge University Press, 1993.

Morgan, John and Phillip Stocken, "An Analysis of Stock Recommendations," *Rand Journal of Economics*, 2003, *34* (1), 183–203.

Morris, Stephen, "Political Correctness," *Journal of Political Economy*, April 2001, *109* (2), 231–265.

Napel, Stefan, "Aspiration Adaptation in the Ultimatum Minigame," *Games and Economic Behavior*, 2003, *43* (1), 86–106.

Nowak, Martin A and Karl Sigmund, "A Strategy of Win Stay, Lose Shift That Outperforms Tit-for-Tat in the Prisoners-Dilemma Game," *nature*, 1993, *364* (6432), 56–58.

———, ———, and Esam El-Sedy, "Automata, repeated games and noise," *Journal of Mathematical Biology*, July 1995, *33* (7), 703–722.

Ottaviani, Marco and Peter Sørensen, "Professional advice," *Journal of Economic Theory*, 2006, *126* (1), 120–142.

______ **and** ______, "Reputational Cheap Talk," *Rand Journal of Economics*, 2006, *37* (1), 155–175.

Palacios-Huerta, Ignacio and Oscar Volij, "Field Centipedes," *American Economic Review*, September 2009, *99* (4), 1619–1635.

Pazgal, Amit, "Satisficing Leads to Cooperation in Mutual Interests Games," *International Journal of Game Theory*, 1997, *26* (4), 439–453.

Posch, Martin, "Win-Stay, Lose-Shift Strategies for Repeated Games - Memory Length, Aspiration Levels and Noise," *Journal of theoretical biology*, 1999, *198* (2), 183–195.

______**, Alexander Pichler, and Karl Sigmund**, "The Efficiency of Adapting Aspiration Levels," *Proceedings: Biological Sciences*, July 1999, *266* (1427), 1427–1435.

Roth, Alvin E and Ido Erev, "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 1995, *8* (1), 164–212.

Sarin, Rajiv and Farshid Vahid, "Payoff Assessments without Probabilities: A Simple Dynamic Model of Choice," *Games and Economic Behavior*, 1999, *28* (2), 294–309.

Schotter, Andrew, Keith Weigelt, and Charles Wilson, "A Laboratory Investigation of Multiperson Rationality and Presentation Effects," *Games and Economic Behavior*, May 1994, *6* (3), 445–468.

Seidmann, Daniel, "Effective Cheap Talk with Conflicting Interests," *Journal of Economic Theory*, 1990, *50* (2), 445–458.

Simon, Herbert A, *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*, University of Michigan Press, 1947.

______, "Rational Choice and the Structure of the Environment," *Psychological Review*, 1956, *63* (2), 129–138.

Sobel, Joel, "A Theory of Credibility," *Review of Economic Studies*, October 1985, *52* (4), 557–573.

Watson, Joel, "Information Transmission When the Informed Party Is Confused," *Games and Economic Behavior*, 1996, *12* (1), 143–161.

Young, H. Peyton, "The Evolution of Conventions," *Econometrica*, 1993, *61* (1), 57–84.

APPENDIX A

Proofs

A.1. Persuading Someone You Do Not Know

Proof of Proposition 1 Without loss of generality, suppose the decision maker's types are classified into two groups, i.e., $N = 2$, with $\theta = (\theta_1, \theta_2)$ and $\bar{t} = (t_1)$, and u is strictly convex. If M is an equilibrium strategy, the expert must be indifferent between any two messages sent by M: $\int_0^1 u(\mathbb{E}[\theta(t)|m])dG(t) = \int_0^1 u(\mathbb{E}[\theta(t)|m'])dG(t)$ for any $(\theta, \bar{t})$ and $(\theta', \bar{t}')$, where $m = M(\theta, \bar{t})$ and $m' = M(\theta', \bar{t}')$. And, since M is truthful on $\overline{T}$, the expert's expected payoff is $\int_0^1 u(\mathbb{E}[\theta(t)])dG(t)$. To see this, assume t_1 approaches 1. Then, the decision maker's type becomes identified, thus the expert cannot send credible messages to the decision maker. And, the above equality implies that $\int_0^1 u(\mathbb{E}[\theta(t)|m])dG(t) = \int_0^1 u(\mathbb{E}[\theta(t)])dG(t)$ for any $(\theta, \bar{t})$. Now suppose that for a non-negligible (with respect to G) set of types, M is informative. Let P denote the probability measure of m that is induced by players' priors F and H and strategy M. Then, we have

$$
\begin{aligned}
\int_0^1 u(\mathbb{E}[\theta(t)|m])dG(t) &= \int_0^1 u(\mathbb{E}[\theta(t)|m])dG(t) \cdot \int_\Phi dP(m) \\
&= \int_\Phi \int_0^1 u(\mathbb{E}[\theta(t)|m])dG(t)dP(m) \\
&= \int_0^1 \int_\Phi u(\mathbb{E}[\theta(t)|m])dP(m)dG(t) \\
&> \int_0^1 u\left(\int_\Phi \mathbb{E}[\theta(t)|m]dP(m)\right) dG(t) \\
&= \int_0^1 u\left(\mathbb{E}[\theta(t)]\right) dG(t).
\end{aligned}
$$

This contradicts. Thus, we can conclude that $\mathbb{E}[\theta(t)|m] = \mathbb{E}[\theta(t)]$ for almost all types. ■

Receiving a message from the expert, the decision maker estimates $\theta(t)$ based on her priors F and H, the expert's strategy M, and sent message m according to

Bayes' rule. Let $e^j(t) = E[\theta(t)|m^j]$, $e(t) = E[\theta(t)|m]$, $\theta^j = E[\theta|m^j]$, $\theta_i^j = E[\theta_i|m^j]$, and $H_0(t) = 1$ and $H_N(t) = 0$ for all $t \in T$. Then, the decision maker's estimate given m^j takes the following form.

Lemma 3. $e^j(t) = \sum_{i=1}^N \theta_i^j (H_{i-1}(t) - H_i(t))$.

Proof 3 By the law of iterated expectation, we have

$$\begin{aligned} e^j(t) &= \mathbb{E}[\theta(t)|m^j] = \mathbb{E}[\mathbb{E}[\theta(t)|\theta]|m^j] \\ &= \mathbb{E}[\textstyle\sum_{i=1}^N \theta_i (H_{i-1}(t) - H_i(t))|m^j] \\ &= \textstyle\sum_{i=1}^N \mathbb{E}[\theta_i|m^j](H_{i-1}(t) - H_i(t)) \\ &= \textstyle\sum_{i=1}^N \theta_i^j (H_{i-1}(t) - H_i(t)), \end{aligned}$$

where we want to show that the second equality holds. Since a realised event is either $t < t_{i-1}, t_{i-1} \leq t < t_i$, or $t_i \leq t$, it holds that $\Pr(t_{i-1} \leq t \text{ and } t < t_i) = 1 - \Pr(t_{i-1} > t) - \Pr(t_i \leq t) = H_{i-1}(t) - H_i(t)$. And, by definition, $\Pr(t < t_1) = 1 - H_1(t) = H_0(t) - H_1(t)$ and $\Pr(t \geq t_{N-1}) = H_{N-1}(t) = H_{N-1}(t) - H_N(t)$. Thus, $\theta(t)$ equals to θ_i^j with probability $H_{i-1}(t) - H_i(t)$ for all $i = 1, \ldots, N$ and the second equality follows. ■

To prove Proposition 2, we first introduce a few definitions and lemmas. Let $\mathbb{S}^{N-1}$ denote the boundary of the N-dimensional unit ball $B^N \subset \mathbb{R}^N$. Then a hyperplane $h(s, c)$ of orientation $s \in \mathbb{S}^{N-1}$ passing through an interior point $c \in \Theta$ partitions Θ into two nonempty sets $\Theta^+(s, c)$ and its complement $\Theta^-(s, c)$,[1] with corresponding estimates $\theta^i(s, c) = \mathbb{E}[\theta|\theta \in \Theta^i(s, c)]$ for $i = +, -$. Hereafter, we omit c for notational simplicity. Let Θ^+ be in the half-space that contains the point $s + c$. As discussed above, a message set induced by a partition of the state space could function as a strategy and each element of the partition is interpreted as a message.

Let π_i be the coordinate map from Θ to Θ_i defined as $\pi_i(\theta) = \theta_i$ for all $i = 1, \ldots, N$, and $\Pi(N)$ be the set of all permutation functions defined on $\{1, \ldots, N\}$. The following lemma shows that there exists a hyperplane such that two corresponding estimates of θ are identical except for one coordinate.

Lemma 4. *There exists* $s \in \mathbb{S}^{N-1}$ *such that for any* $p \in \Pi(N)$, $\pi_{p(i)}(\theta^+(s)) = \pi_{p(i)}(\theta^-(s))$ *for* $i = 1, \ldots, N - 1$.

[1]This construction of a partition of the state space was originally used in Chakraborty and Harbaugh (2010).

Proof First, note that, since F has full support on Θ, $\theta^+(s) \in \text{int}\Theta^+(s)$ and $\theta^-(s) \in \text{int}\Theta^-(s)$ so that $\theta^+(s) \neq \theta^-(s)$. Note next that for any two opposite orientations $-s, s \in \mathbb{S}^{N-1}$, $\Theta^+(s) = \Theta^-(-s)$ and $\Theta^-(s) = \Theta^+(-s)$. It follows that $\theta^+(s) = \theta^-(-s)$, implying that a map $\rho : \mathbb{S}^{N-1} \to \mathbb{R}^{N-1}$ which consists of $\rho_i(s) = \pi_i(\theta^-(s)) - \pi_i(\theta^+(s))$ for $i \in \{p(j)|j = 1, \dots, N-1 \text{ and } p \in \Pi(N)\}$ is a continuous odd function of s for any $p \in \Pi(N)$.[2] By the Borsuk-Ulam theorem, there exists $s^p \in \mathbb{S}^{N-1}$ such that $\rho(s^p) = 0$. ■

From Lemma 4, we can find a hyperplane such that the expert is indifferent between two messages induced by the hyperplane. Let $e^+(t; s)$ and $e^-(t; s)$ denote the estimate of $\theta(t)$ when the messages are $\Theta^+(s)$ and $\Theta^-(s)$, respectively.

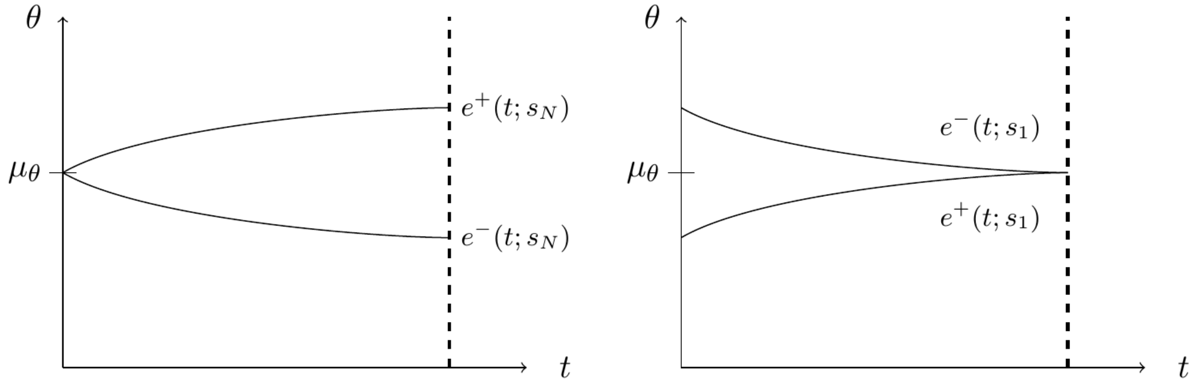
Lemma 5. *There exists $s \in \mathbb{S}^{N-1}$ such that $\mathbb{E}[u(e^+(t; s))] = \mathbb{E}[u(e^-(t; s))]$.* [3]

Proof Consider a function $\Delta(s) \equiv \int_T [U(e^+(t; s)) - U(e^-(t; s))] \, \mathrm{d}G(t)$. We need to find $s^* \in \mathbb{S}^{N-1}$ that satisfies $\Delta(s^*) = 0$. By Lemma 4, there exist $s_N, s_1 \in \mathbb{S}^{N-1}$ such that $\theta_i^+(s_N) = \theta_i^-(s_N)$ for $i \in \{1, \dots, N-1\}$ and $\theta_i^+(s_1) = \theta_i^-(s_1)$ for $i \in \{2, \dots, N\}$. Since $\theta^+(s) \neq \theta^-(s)$ for any $s \in \mathbb{S}^{N-1}$, $\theta_N^+(s_N) \neq \theta_N^-(s_N)$ and $\theta_1^+(s_1) \neq \theta_1^-(s_1)$. Suppose that $\theta_N^+(s_N) > \theta_N^-(s_N)$.[4] Then, since H has full support on the type space, $e^+(t; s_N) > e^-(t; s_N)$ for all $t \in \text{int}T$ and $\Delta(s_N) > 0$ as in the left panel of Figure A.1. Similarly, suppose $\theta_1^+(s_1) < \theta_1^-(s_1)$. Then $e^+(t; s_1) < e^-(t; s_1)$ for all $t \in \text{int}T$ and $\Delta(s_1) < 0$ as in the right panel of Figure A.1. Consider a continuous map $s : [0, 1] \to \mathbb{S}^{N-1}$ such that $s(0) = s_N$ and $s(1) = s_1$. Then $\Delta(s(x))$ is a continuous function such that $\Delta(s(0)) > 0$ and $\Delta(s(1)) < 0$. By continuity, there exists $x^* \in (0, 1)$ such that $\Delta(s(x^*)) = 0$. ■

[2]It can be simply shown that $\theta^+(s)$ and $\theta^-(s)$ are continuous functions of s by the dominated convergence theorem.

[3]Since $\mathbb{E}[u(e^+(t; s))] - \mathbb{E}[u(e^-(t; s))]$ is a continuous odd function of s, this lemma can be proved without Lemma 4 as in the proof of Theorem 1 of Chakraborty and Harbaugh (2010). However, the current approach better illustrates the expert's strategies and the proof of Proposition 3.

[4]If the inequality does not hold for s_N, then the desired inequality holds for $-s_N$.



(a) s_N (b) s_1

Figure A.1. Estimates Induced by Two Orientations s_N and s_1

Last, we can easily check that $\mathbb{E}[\theta(t)]$ is a convex combination of $e^+(t;s)$ and $e^-(t;s)$ with weights $P^+(s)$ and $P^-(s)$ for any t, where $P^i(s) \equiv \Pr(\theta \in \Theta^i(s))$ for $i = +, -$.[5]

Proof of Proposition 2 Choose any $c \in \text{int}\Theta$. Then, by Lemma 5, we can find $s^* \in \mathbb{S}^{N-1}$ such that $\mathbb{E}[u(e^+(t;s^*))] = \mathbb{E}[u(e^-(t;s^*))]$. Consider a simple strategy as follows. If the expert observes $\theta \in \Theta^+(s^*)$ $(\Theta^-(s^*))$, the expert draws a state from $\Theta^+(s^*)$ $(\Theta^-(s^*))$ according to a probability distribution on the subset and sends the decision maker the drawn state. This fully specifies the expert's strategy because $\Theta^+(s^*) \cup \Theta^-(s^*) = \Theta$. Given this strategy, the decision maker estimates the expected $\theta(t)$ based on the sent message. The estimate is $e^+(t;s^*)$ $(e^-(t;s^*))$ if the alleged state is in $\Theta^+(s^*)$ $(\Theta^-(s^*))$. Since $\mathbb{E}[u(e^+(t;s^*))] = \mathbb{E}[u(e^-(t;s^*))]$, the expert's strategy maximises his *ex ante* payoff.

[5]For any $t \in T$,

$$\begin{aligned} P^+(s)e^+(t;s) + P^-(s)e^-(t;s) &= \sum_{i=1}^{N} (P^+(s)\theta_i^+(s) + P^-(s)\theta_i^-(s))(H_{i-1}(t) - H_i(t)) \\ &= \sum_{i=1}^{N} \mathbb{E}[\theta_i](H_{i-1}(t) - H_i(t)) \\ &= \mathbb{E}[\theta(t)]. \end{aligned}$$

Upon the strategy, the expert's expected payoff is

$$\begin{aligned}\mathbb{E}[u(e^+(t;s^*))] &= (\mathrm{P}^+(s^*)+\mathrm{P}^-(s^*))\int_T u(e^+(t;s^*))\,\mathrm{d}G(t)\\ &= \int_T[\mathrm{P}^+(s^*)u(e^+(t;s^*))+\mathrm{P}^-(s^*)u(e^-(t;s^*))]\,\mathrm{d}G(t)\\ &\geq \int_T u(\mathrm{P}^+(s^*)e^+(t;s^*)+\mathrm{P}^-(s^*)e^-(t;s^*))\,\mathrm{d}G(t)\\ &= \mathbb{E}[u(\mathbb{E}[\theta(t)])],\end{aligned}$$

where the inequality holds if u is convex, and the last term is the expert's expected payoff without communication. Thus, the expert is not worse off with the communication for any realised state so long as u is convex. If u is strictly convex, the expert is strictly better off with the equilibrium communication. The opposite holds if u is (strictly) concave.

Now, consider a system of simultaneous equations $e^+(t)=\mathbb{E}[\theta(t)]$ or $\sum_{i=1}^N(\theta_i^+-\mu_i)(H_{i-1}(t)-H_i(t))=0$ for $t\in T_0\subset T$ for some finite set T_0. Since H has full support on T, the problem consists of $|T_0|$ equations. Then, only when $|T_0|\leq N$, the system admits a solution $\{\hat{\theta}_i\}_{i=1}^N$ such that $\sum_{i=1}^N(\hat{\theta}_i-\mu_i)(H_{i-1}(t)-H_i(t))=0$ for $t\in T_0$. Thus, we have $\Pr[t\in T_0]=0$. Since $e^+(t)\neq\mathbb{E}[\theta(t)]$ whenever $t\notin T_0$, the decision maker can make an informed decision with probability 1. This completes the proof. ■

Proof of Proposition 3 An interior point $c\in\mathrm{int}\Theta$ and an orientation $s\in\mathbb{S}^{N-1}$ are represented in coordinates by $c=(c_1,\dots,c_N)$ and $s=(a_1,\dots,a_N)$. Then, (c,s) defines two subsets of the state space, $\Theta^+=\{\theta\in\Theta|\sum_{i=1}^N a_i(\theta_i-c_i)\geq 0\}$ and $\Theta^-=\{\theta\in\Theta|\sum_{i=1}^N a_i(\theta_i-c_i)<0\}$ and corresponding estimates $e^+(t;c,s)$ and $e^-(t;c,s)$. Choose $a>0$. Let $a_i=a$ for $i=1,\dots,n$ for some $1<n<N$ and $a_i=-1$ for $i=n+1,\dots,N$ and $e^+(t;a)$ and $e^-(t;a)$ denote the corresponding estimates. Since F is invariant, we have $\theta_i^+=\theta_a^+>\mathbb{E}[\theta_i]$ for $i=1,\dots,n$ and $\theta_i^+=\theta_{-1}^+<\mathbb{E}[\theta_i]$ for $i=n+1,\dots,N$. Thus, we have $e^+(t;a)=\theta_a^++(\theta_{-1}^+-\theta_a^+)H_n(t)$, which is decreasing in t with $e^+(t_0;a)>0>e^+(t_N;a)$. Similarly, $e^-(t;a)=\theta_a^-+(\theta_{-1}^--\theta_a^-)H_n(t)$, which is increasing in t with $e^-(t_0;a)<0<e^-(t_N;a)$. Suppose $\mathbb{E}[e^+(t;a)]\geq\mathbb{E}[e^-(t;a)]$. If $\mathbb{E}[e^+(t;a)]=\mathbb{E}[e^-(t;a)]$, we have found a strategy which is first-order. If $\mathbb{E}[e^+(t;a)]>\mathbb{E}[e^-(t;a)]$, we can find $a'\in(0,a)$ such that $\mathbb{E}[e^+(t;a')]=\mathbb{E}[e^-(t;a')]$. For a second-order strategy, let $a_i=a$ for $i=n,\dots,n+k$ for some $k\in\mathbb{N}\cup\{0\}$ such that $n+k<N$ and $a_i=-1$ for $i=1,\dots,n-1,n+k+1,\dots,N$.

Then, $e^+(t;a) = \theta^+_{-1} + (\theta^+_a - \theta^+_{-1})(H_{n-1}(t) - H_{n+k}(t))$ is a uni-modal function for $\frac{\mathrm{d}}{\mathrm{d}t}(H_{n-1}(t) - H_{n+k}(t)) = (n+k)H(t)^{n-2}h(t)(\frac{n-1}{n+k} - H(t)^{k+1})$, where $\frac{\mathrm{d}H}{\mathrm{d}t} = h$. Thus, once we find a' such that $\mathbb{E}[e^+(t;a')] = \mathbb{E}[e^-(t;a')]$ in the previous manner, the corresponding strategy is second-order. ■

Proof of Proposition 4 The proposition is directly derived by Proposition 2 and the inductive argument used in the proof of Theorem 4 in Chakraborty and Harbaugh (2010). ■

Proof of Proposition 5 Let M be a 2-message informative equilibrium strategy in a one-way cheap talk game, and messages in M lead to decision maker's estimates $e^+(t)$ and $e^-(t)$. Then, $T_M \cap \mathrm{int}T \neq \emptyset$. Otherwise, $e^+(t) > \mathbb{E}[\theta(t)] > e^-(t)$ or $e^-(t) > \mathbb{E}[\theta(t)] > e^+(t)$ for any $t \in \mathrm{int}T$ so that $\mathbb{E}[u(e^+(t))] \neq \mathbb{E}[u(e^-(t))]$, which contradicts that M constitutes an equilibrium.

Choose consecutive types $p, q, r \in T_M \cup \{t_0, t_N\}$ such that $q \in T_M \setminus \{t_0, t_N\}$ and $p < q < r$, where t_0 and t_N are the end points of T (Figure A.2), and the sign of $e^+(t) - e^-(t)$ reverses only at q in (p, r). Without loss of generality, suppose that $e^+(t) > e^-(t)$ for $t \in (p, q)$ and $e^+(t) < e^-(t)$ for $t \in (q, r)$ and

$$\int_{(p,r)} g(t)u(e^+(t))\,\mathrm{d}t \leq \int_{(p,r)} g(t)u(e^-(t))\,\mathrm{d}t.$$

Then, for any $t^*_1 \in (p, q)$, there exists $t^*_2 \in (q, r)$ such that

$$\int_{(t^*_1,t^*_2)} g(t)u(e^+(t))\,\mathrm{d}t = \int_{(t^*_1,t^*_2)} g(t)u(e^-(t))\,\mathrm{d}t$$

because e^+ and e^- are continuous functions.

Repeating this procedure, we can find a finite set of exclusive intervals $\{T_1, \ldots, T_p\}$ such that $\int_{T_i} g(t)u(e^+(t))\,\mathrm{d}t = \int_{T_i} g(t)u(e^-(t))\,\mathrm{d}t$ for $i = 1, \ldots, p$, where $1 \leq p \leq |T_M|$. Consider a strategy N which consists of a partition $\{T_1, \ldots, T_p, T \setminus \cup^p_{i=1} T_i\}$ in which a decision maker, whose type belongs to one element of the partition, draws an arbitrary type from the subset according to a probability distribution and sends it to the expert. Then, the expert is still indifferent between two messages in the original strategy M. Therefore, two strategies N and M constitute an equilibrium in the two-way game. ■

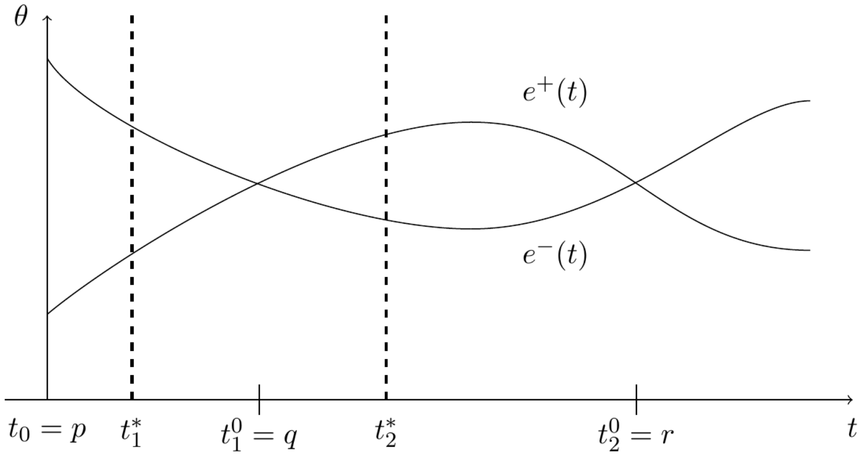


Figure A.2. Construction of Revelation Strategy

Proof of Proposition 6 If $T_M = \{t_1^0\}$ with $t_1^0 \in \text{int}T$, the possible form of strategy is $\{T_1, T \setminus T_1\}$ where $t_1^0 \in T_1 \subset T$ and $t_0, t_N \notin T_1$. Therefore, if $T_1 \neq T$, the strategy is not first-order.

If $T_M = \{t_1^0, t_2^0\}$ with $t_1^0, t_2^0 \in \text{int}T$, there exists $t^* \in (t_1^0, t_2^0)$ such that

$$\int_{(t_0,t^*)} g(t)u(e^+(t))\,\mathrm{d}t = \int_{(t_0,t^*)} g(t)u(e^-(t))\,\mathrm{d}t.$$

Thus, the decision maker's strategy $N = \{[t_0, t^*), [t^*, t_N]\}$ is first-order and constitutes an equilibrium. ■

A.2. A Model of Satisficing Behaviour

Proof of Proposition 7 Given a transition probability function P for (ϕ_n, η_n) on $\Phi \times \Upsilon$, let $(\Omega, \mathcal{F}, \mathbb{P}_{\phi,\eta})$ be the corresponding filtered space, where Ω is the canonical sample space $(\Phi \times \Upsilon)^\infty$. Then, since a sequence $\{s_{n,i}(\omega)\}$ is monotone and bounded for all i and $\omega \in \Omega$, it has a limit $s_{\infty,i} \in S_i$. Suppose the proposition does not hold, or at least one satisficing decision maker infinitely often does not satisfice with her actions. This implies that the sequence of the decision maker's satisficing level does not have a limit, which contradicts. ■

We introduce definitions and results in Markov process literature that are essential for the following proofs. A Markov process ϕ_n (or P) is *weak-Feller* if P maps the space of bounded and continuous functions on Φ into itself. A process

ϕ_n is *φ-irreducible* if there exists a non-trivial measure φ on $\mathcal{B}(\Phi)$ such that, whenever $\varphi(B) > 0$, $U(\phi, B) \equiv \sum_{n=1}^{\infty} P^n(\phi, B) > 0$ for any ϕ. Furthermore, a set $C \in \mathcal{B}(\Phi)$ is called a *small set* if there exists $m \in \mathbb{N}$ and a non-trivial measure ν on $\mathcal{B}(\Phi)$ such that for all $\phi \in C$ and $B \in \mathcal{B}(\Phi)$, $P^m(\phi, B) \geq \nu(B)$. And, if ϕ_n is φ-irreducible, then by Proposition 4.2.2 in Meyn and Tweedie (1993) (henceafter MT), there exists a *maximal irreducibility* probability measure ψ on $\mathcal{B}(\Phi)$ such that (i) ϕ_n is ψ-irreducible, (ii) for any other measure φ', the process ϕ_n is φ'-irreducible if and only if φ' is absolutely continuous with respect to ψ, and (iii) if we let $\mathcal{B}^+(\Phi) \equiv \{B \in \mathcal{B}(\Phi) | \psi(B) > 0\}$ for the sets of positive ψ-measure, $\mathcal{B}^+(\Phi)$ is unique. ϕ is called *recurrent* if it is ψ-irreducible and $U(\Phi, B) = \infty$ for every ϕ and every $B \in \mathcal{B}^+(\Phi)$. Lastly, a set $B \in \mathcal{B}(\Phi)$ is *Harris recurrent* if $P_\phi(\eta_B = \infty) = 1$ for all $\phi \in B$, where $\eta_B \equiv \sum_{n=1}^{\infty} \mathbf{1}\{\phi_n \in B\}$, and a process ϕ_n is Harris recurrent if it is ψ-irreducible and every set $B \in \mathcal{B}^+(\Phi)$ is Harris recurrent. If in addition the process ϕ_n admits an invariant probability measure, then it is positive Harris recurrent.

It is known that if a positive Harris recurrent Markov process is aperiodic and has a unique invariant probability measure, its n-step t.p.f. strongly converges to the measure: for any ϕ, $P_\epsilon^n(\phi, \cdot)$ converges to μ_ϵ in the total variation norm.[6] A process is periodic if there exists a minimum periods of time $d > 1$ such that once the process leaves a subset, it takes multiples of d for the process to return to the set.[7] If $n = 1$, the process is aperiodic. Because of the trembling behaviour, if ϕ_n returns to a subset in n periods, it also could return in $n+1$ periods. Therefore, ϕ_n is aperiodic. The following proof shows that (ϕ_n, η_n) is positive Harris recurrent for any positive ϵ and $\overline{\lambda}$.

Proof of Proposition 8 Since the process (ϕ_n, η_n) is weak-Feller and $\Phi \times \Upsilon$ is a compact metric space, P_ϵ admits an invariant probability measure μ by Theorem 7.2.3. in Hernández-Lerma and Lasserre (2003) (henceafter HL). And, since both shocks occur independently to all decision makers with positive probabilities and the state space is compact, we can construct a non-trivial measure φ such that the

[6]The norm is defined as

$$||\mu|| \equiv \sup_{A \in \mathcal{B}(\Phi)} \mu(A) - \inf_{A \in \mathcal{B}(\Phi)} \mu(A).$$

[7]For the formal definition and related properties, refer to Theorem 5.4.4 in MT.

process (ϕ_n, η_n) is φ-irreducible, and furthermore, the whole state space is a small set with respect to φ. Thus, the process is ψ-irreducible, which implies that μ is the unique invariant probability measure for P_ϵ by Proposition 4.2.2. in HL. And, by Proposition 9.1.7. in MT, (ϕ_n, η_n) is positive Harris recurrent. Thus, by Theorem 4.3.4. in HL, we have the result. ■

The next asymptotic result is when ϵ approaches 0. To characterise the limit of the invariant p.m. μ_ϵ as $\epsilon \downarrow 0$, we decompose the perturbed process P_ϵ into components as

$$P_\epsilon = (1 - \zeta(\epsilon))P + \zeta(\epsilon)Q_\epsilon$$

where $\zeta(\epsilon)$ is the probability that at least one player experiences a shock, and Q_ϵ is the conditional t.p.f. And, Q_ϵ can be further decomposed as

$$Q_\epsilon = (1 - \xi(\epsilon))Q + \xi(\epsilon)\widetilde{Q}$$

where $1 - \xi(\epsilon)$ is the conditional probability of only one player experiencing only one shock, either the first or the second, given that at least one player experiences a shock, and Q is the conditional t.p.f.[8] $\widetilde{Q}$ is the t.p.f. when more than one shock occur.

Before proceeding to characterise the limit of μ_ϵ in the proof of Proposition 9, we analyse the limit of P^n as n grows.

Lemma 6. *There exists a weak-Feller t.p.f. P^∞ such that, for any $f \in C(\Phi \times \Upsilon), |R_\epsilon f - P^\infty f| \to 0$ as $\epsilon \to 0$.*

Proof Consider a sequence of functions $\{P^n f\}_{n \geq 0}$ for $f \in C(\Phi \times \Upsilon)$. Let $A_n \equiv \{\omega | s_n \leq \min \pi(a_n)\}$. $\mathbb{E}_{\phi,\eta}(\cdot)$ is the expected value with respect to $\mathbb{P}_{\phi,\eta}$. For any

[8]With ϵ^a and ϵ^s for trembling probabilities on action and satisficing levels,

$$\begin{aligned} \zeta(\epsilon) &= 1 - (1-\epsilon^a)^I(1-\epsilon^s)^I \\ \xi(\epsilon) &= 1 - \frac{I(\epsilon^a(1-\epsilon^a)^{I-1}(1-\epsilon^s)^I + \epsilon^s(1-\epsilon^a)^I(1-\epsilon^s)^{I-1})}{\zeta(\epsilon)}. \end{aligned}$$

$n, m \geq 0$,

$$\begin{aligned}
& |P^{2n} f(\phi, \eta) - P^{2n+m} f(\phi, \eta)| \\
= \ & |\mathbb{E}_{\phi,\eta}[f(\phi_{2n}, \eta_{2n}) - f(\phi_{2n+m}, \eta_{2n+m})]| \\
\leq \ & |\mathbb{E}_{\phi,\eta}[(f(\phi_{2n}, \eta_{2n}) - f(\phi_{2n+m}, \eta_{2n+m}))\mathbf{1}_{A_n}]| \\
& + |\mathbb{E}_{\phi,\eta}[(f(\phi_{2n}, \eta_{2n}) - f(\phi_{2n+m}, \eta_{2n+m}))\mathbf{1}_{\Omega \setminus A_n}]| \\
\leq \ & \sup_{s \leq u} |\mathbb{E}_{\phi,\eta}[f(\phi_n, \eta_n) - f(\phi_{n+m}, \eta_{n+m})]| \\
& + 2\mathbb{P}_{\phi,\eta}(\Omega \setminus A_n)||f|| \\
= \ & \sup_{s \leq u} |P^n f(\phi, \eta) - P^{n+m} f(\phi, \eta)| + 2\mathbb{P}_{\phi,\eta}(\Omega \setminus A_n)||f||.
\end{aligned}$$

If $s \leq u$, by the property of the satisficing behaviour and the second condition imposed on the auxiliary state, the subsequent process reduces to the dynamics of valuations $\{v_n\}_{n \geq 0}$ of a certain action profile a with a t.p.f. induced by μ_ρ, and, if applied, the distribution of $w \in W$. Since the outcome is stationary conditional on a, the reduced process is positive Harris recurrent and aperiodic. Thus, for any (ϕ, η) with $s \leq u$,

$$\sup_{s \leq u} |P^n f(\phi, \eta) - P^{n+m} f(\phi, \eta)| \leq ||f|| \sup_{s \leq u} |P^n((\phi, \eta), \cdot) - P^{n+m}((\phi, \eta), \cdot)|,$$

which approaches 0 as n approaches infinity. And, as shown in Proposition 7, $\mathbb{P}_{\phi,\eta}(A_n) \to 1$ as $n \to \infty$ for any $(\phi, \eta) \in \Phi \times \Upsilon$. Therefore, we have

$$\sup_{m > 0} ||P^{2n} f(\phi, \eta) - P^{2n+m} f(\phi, \eta)|| \to 0 \text{ as } n \to 0.$$

Since $\{P^n f\}_{n \geq 0}$ is a Cauchy sequence in $C(\Phi \times \Upsilon)$ equipped with the sup norm, and hence converges in $C(\Phi \times \Upsilon)$, by Proposition 3.4 in Chasparis et al. (2013), there exists a weak-Feller t.p.f. P^∞ such that for any $f \in C(\Phi \times \Upsilon)$

$$||P^n f - P^\infty f|| \to 0 \text{ and } ||R_\epsilon f - P^\infty f|| \to 0 \text{ as } n \to \infty,$$

where $R_\epsilon \equiv \zeta(\epsilon) \sum_{n=0}^{\infty} (1 - \zeta(\epsilon))^n P^n$. ■

Proof of Proposition 9 (1) Post-multiplying both sides of $P_\epsilon = (1 - \zeta(\epsilon))P + \zeta(\epsilon) Q_\epsilon$ by R_ϵ, $P_\epsilon R_\epsilon = R_\epsilon - \zeta(\epsilon)\boldsymbol{I} + \zeta(\epsilon) Q_\epsilon R_\epsilon$, where $\boldsymbol{I}$ denotes the identity operator. And, pre-multiplying an invariant p.m. μ_ϵ of P_ϵ, $\mu_\epsilon R_\epsilon = \mu_\epsilon R_\epsilon - \zeta(\epsilon)\mu_\epsilon + \zeta(\epsilon)\mu_\epsilon Q_\epsilon R_\epsilon$, which is simplified to $\mu_\epsilon = \mu_\epsilon Q_\epsilon R_\epsilon$. (2) For any $f \in C(\Phi \times \Upsilon)$,

$$\begin{aligned}\|Q_\epsilon R_\epsilon f - QP^\infty f\| &\leq \|Q_\epsilon(R_\epsilon f - P^\infty f)\| + \|Q_\epsilon P^\infty f - QP^\infty f\| \\ &\leq \|R_\epsilon f - P^\infty f\| + \|Q_\epsilon P^\infty f - QP^\infty f\|,\end{aligned}$$

where the first and second terms on the right hand side tend to 0 as $\epsilon \downarrow 0$ by Lemma 6 and the definition of Q_ϵ. (3) Let μ_0 be an accumulation point of $\{\mu_\epsilon\}_{\epsilon\downarrow 0}$, which exists by Banach-Alaoglu theorem. For any $f \in C(\Phi \times \Upsilon)$,

$$\|\mu_0 f - \mu_0 QP^\infty f\| \leq \|\mu_0 f - \mu_\epsilon f\| + \|\mu_\epsilon(Q_\epsilon R_\epsilon f - QP^\infty f)\| + \|\mu_\epsilon QP^\infty f - \mu_0 QP^\infty f\|,$$

where the first and third terms tend to 0 as $\epsilon \downarrow 0$ along some subsequence by the weak convergence of $\{\mu_\epsilon\}_{\epsilon\downarrow 0}$ to μ_0, while the second term is dominated by $\|Q_\epsilon R_\epsilon f - QP^\infty f\|$ which also tends to 0 as shown above. Thus, we have $\mu_0 = \mu_0 QP^\infty$ by Theorem 1.2 in Billingsley (1999). ■

Proof of Proposition 10 Without loss of generality, we assume that a^1 is the unique maximin action and $s_0 \leq \min \pi(a^1)$, which is guaranteed by Proposition 7. Let $E_\infty(a^j) \equiv \{\omega | a_n = a^j \text{ for all } n \geq m \text{ for some } m \in \mathbb{N}\}$. We show $\mathbb{P}_{\phi_0}(\cup_{j=2}^J E_\infty(a^j)) \to 0$ as $\overline{\lambda} \to 0$ if $s_0 \in (\min \pi(a^2), \min \pi(a^1)]$. If $a_0 = a^1$, the result is trivial. Suppose that $a_0 \neq a^1$. Since μ_α assigns positive probabilities to all alternatives, the maximum probability which the decision maker assigns to non-maximum actions is bounded above by some constant lower than 1 at any time, that is, $\overline{\delta} \equiv \max_\phi\{\sum_{j=2}^J \mu_\alpha(\{a^j\}; \phi)\} < 1$. Let $g(\overline{\lambda}; s_0)$ denote the minimum number of switching actions required for the satisficing level to get as low as $\min \pi(a^2)$ given ϕ_0 and $\overline{\lambda}$. Since the number can be minimised when the decision maker repeatedly receives the worst payoff, $\min \pi$, $g(\overline{\lambda}; s_0) = \lceil (s_0 - \min \pi(a^2))/(\overline{\lambda}(s_0 - \min \pi)) \rceil$, where $\lceil x \rceil$ is the ceiling function returning the smallest integer not less than $x \in \mathbb{R}$. Let $C_n \equiv \cap_{m=0}^n \{\omega | a_m \neq a^1\}$ denote the event in which the maximin action is not chosen for the first $n+1$ times. Let $D \equiv \{\phi | s \leq \min \pi(a^2)\}$ and $\tau_D \equiv \{n \geq 1 | \phi_n \in D\}$.

Then, for any $\overline{\lambda} > 0$,

$$\begin{aligned}
\mathbb{P}_{\phi_0}(\cup_{j=2}^{J} E_\infty(a^j)) &= \mathbb{P}_{\phi_0}(\cup_{m=g(\overline{\lambda};s_0)}^{\infty}\{\omega \in \Omega | \tau_D = m\}) \\
&\leq \sum_{m=g(\overline{\lambda};s_0)}^{\infty} \mathbb{P}_{\phi_0}(\{\omega | \tau_D = m\}) \\
&\leq \mathbb{P}_{\phi_0}(C_{g(\overline{\lambda};s_0)}) \sum_{m=g(\overline{\lambda};s_0)}^{\infty} \mathbb{P}_{\phi_0}(\{\omega | \tau_D = m\} | C_{g(\overline{\lambda};s_0)}) \\
&\leq \overline{\delta}^{g(\overline{\lambda};s_0)},
\end{aligned}$$

which vanishes as $\overline{\lambda} \to 0$ for any $s_0 \in (\min \pi(a^2), \min \pi(a^1)]$. This completes the proof. ■

Proof of Proposition 11 Without loss of generality, we assume that there exists a unique maximin action a^1. First, we show that QP^∞ has a unique invariant p.m. μ_0. Let $C \equiv \{\phi | a = a^1, s \leq \min \pi(a^1)\}$ and $D \equiv \{\phi | a = a^1, s \geq \min \pi(a^2)\}$. Then, $Q(\phi, C \cup D) \geq \delta_1$ for any ϕ for some $\delta_1 > 0$. And, for any $\phi \in D$, as shown in the proof of Proposition 10, $P^\infty(\phi, C) \geq \delta_2$ for some $\delta_2 > 0$. Thus, if we let $\delta \equiv \delta_1 \cdot \delta_2$, $QP^\infty(\phi, C) \geq \delta > 0$ for any ϕ. This implies that QP^∞ is ψ-irreducible. And, since QP^∞ is weak Feller, it admits an invariant probability measure. Then, by Proposition 4.2.2 in HL, an invariant p.m. of QP^∞ is unique. Second, let $C^* \equiv \{\phi | a = a^1 \text{ and } s \in (\min \pi(a^2), \min \pi(a^1)]\}$. Then, $QP^\infty(\phi, C^*)$ converges to 1 as $\overline{\lambda} \to 0$ for any $\phi \in C^*$ as shown in Proposition 10 and $QP^\infty(\phi, C^*) \geq \delta$ for any ϕ for some $\delta > 0$ as shown above. Therefore, $\mu_0(C^*) \to 1$ as $\overline{\lambda} \to 0$. ■

Proof of Proposition 12 Let A_C and A_D denote the sets of unilaterally competitive Nash and payoff dominant profiles, respectively. Without loss generality, we assume the unilaterally competitive Nash profile is strict, i.e., there exists a unique such equilibrium. We show (i) for any $a^* \in A_C \cup A_D$, $\Phi(a^*) \equiv \{\phi | a_i = a_i^*, s_i \in (\widehat{\pi}_i(a^*), \pi_i(a^*)] \text{ for } i = 1, 2\}$ becomes invariant in QP^∞ as $\overline{\lambda} \to 0$, where $\widehat{\pi}_i(a^*)$ is defined as $\max_{a' \neq a^*} \pi_i(a_i', a_{-i}^*)$ if $a^* \in A_C$ and $\max_{a' \in A_C} \pi_i(a')$ if $a^* \in A_D$, and (ii) if a state does not belong to one of the invariant sets, it *enters* $\Phi(a)$ for some $a \in A_C \cup A_D$ with positive probability in finite repetitions of QP^∞, that is, for any

ϕ with $a \notin A_C \cup A_D$ and $s \leq \pi(a)$, $(QP^{\infty})^n(\phi, \Phi(a^*)) \geq \delta$ for some $a^* \in A_C \cup A_D$, $\delta > 0$ and $n < \infty$.

(i) If a^* is unilaterally competitive Nash, it is trivial to show $\Phi(a^*)$ becomes invariant as $\overline{\lambda} \to 0$ because any single tremble by a player cannot change the other's action. Suppose a^* is payoff dominant. It is enough to show that starting from any state ϕ with $a_i \neq a_i^*$ for some i the process enters $\Phi(a^*)$ in finite repetitions of P with positive probability as long as all players' satisficing levels are fixed at $\pi(a^*)$. As long as the current action profile differs from a^*, at least one player, say player 1, does not satisfice by the definition of the payoff dominance. Then, the following period, with positive probability, player 1 could either switch to or continue with an action a_1' such that $(a_1', a_2') \in A_C$ for some a_2'. If player 2's action is not a_2', she does not satisfice thus either switch to or continue with a_2' the following period while player 1 chooses a_1' again with positive probability. Then, both do not satisfice and switch to a^* with positive probability. This sequence of states shows that once a state belongs to $\Phi(a^*)$, following any single tremble, players will return to the initial action profile with probability 1 as $\overline{\lambda} \longrightarrow 0$.

(ii) Now suppose that players satisfice with an action profile $a^0 \in A$, which is neither unilaterally competitive Nash nor payoff dominant.[9] Without loss of generality, we assume their satisficing levels are equal to its payoffs, i.e., $s = \pi(a^0)$. We show that starting from these states the process enters $\Phi(a^*)$ for $a^* \in A_C \cup A_D$ in finite repetitions of QP^{∞} with positive probability. First, suppose that a^0 is strictly preferred to any $a \in A_C$. If one player trembles to choose a unilaterally competitive Nash action. Then, since both players do not satisfice, the process could enter $\Phi(a^*)$ for some payoff dominant profile a^* with positive probability: in the first following period, players sequentially switch to the payoff dominant actions a_1^* and a_2^* and in the subsequent repetitions of QP^{∞}, their satisficing levels are raised up to $\pi(a^*)$ with finite trembles. Second, suppose at least one player, say player 1, will satisfice with a unilaterally competitive Nash profile a^*, that is, $s_1^0 \leq \pi_1(a^*)$. If the player trembles to choose the equilibrium action a_1^*, through the subsequent infinite repetitions of

[9]If players satisfice with $a \in A_C \cup A_D$ but the state ϕ^0 does not belong to $\Phi(a^*)$, as each player trembles to raise their satisficing levels sequentially, the state can enter the invariant set in 2 repetitions of QP^{∞}.

P, if not, player 2's satisficing level will decrease as low as $\pi_2(a^*)$ while player 1 satisfices with a_1^* for any level of $\overline{\lambda}$. ■

Proof of Proposition 13 Suppose players for each role, row and column, are randomly drawn from two separated populations $\mathcal{I} = \{1_i, \ldots, I_i\}$ and $\mathcal{J} =\{1_j, \ldots, I_j\}$, respectively.[10] Without loss of generality, we assume each period only one pair of players are picked, play the two-player game and update their states while the other players' states stay the same. Let $\phi_n \equiv (\phi_{n,1_i}, \ldots, \phi_{n,I_i}; \phi_{n,1_j}, \ldots, \phi_{n,J_j}) \in \Phi$ denote period-n state of all satisficing decision makers. We show (i) for any $a^* \in A_C \cup A_D$, $\Phi(a^*) \equiv \{\phi | a_i = a_1^*, a_j = a_2^*, s_i \in (\widehat{\pi}_1(a^*), \pi_1(a^*)], s_j \in (\widehat{\pi}_2(a^*), \pi_2(a^*)]$ for all $i \in \mathcal{I}$ and $j \in \mathcal{J}\}$ becomes an invariant set of QP^∞ as $\overline{\lambda} \to 0$ and (ii) if a state does not belong to the invariant set, it enters into one of the sets with positive probability in finite repetitions of QP^∞.

(i) If a^* is unilaterally competitive Nash, it is trivial to show $\Phi(a^*)$ becomes invariant as $\overline{\lambda} \to 0$. Suppose that a^* is payoff dominant and all players' satisficing levels are fixed at $\pi_1(a^*)$ or $\pi_2(a^*)$ according to their roles. Then, consider a sequence through which a pair of players whose action profile is not a^* are picked consecutively until they switch to a^*, then another pair of players whose action profile is not a^* are picked until they switch to a^*, and so on. By Proposition 12**,** each pair can switch to a^* within finite repetitions of P with positive probability. Then, since the populations are finite, the whole sequence can be completed with all players playing a^* within finite periods with positive probability.

(ii) The second argument is similar to the second part of the proof of Proposition 12. If all players' satisficing levels are higher than the payoffs from the unilaterally competitive Nash profiles, the same argument in the previous proof applies. Suppose at least one player's satisficing level is lower than or equal to a payoff from a unilaterally competitive Nash profile. If the player, say a row player, trembles to choose a corresponding action of the Nash profile, in the infinite sequence of plays, all column players will be matched with the row player infinitely often until their

[10]The following arguments equally apply to the case in which players are drawn from a single population to play a symmetric game.

satisficing levels get as low as the Nash outcome and satisfice with the Nash action, during which all row players also get their satisficing levels as low as the Nash outcome and satisfice with the Nash action. ■

Since Proposition 14 is a special case of Proposition 15, we just prove the latter.

Proof of Proposition 15 We show that (i) only the sets that are associated with Pareto optimal action profiles are invariant respect to QP^{∞} in the limit as $\overline{\lambda} \to 0$, and (ii) any state in which the action profile is not Pareto optimal and all satisficing decision makers satisfice enters one of the invariant sets with positive probability in finite repetitions of QP^{∞}.

(i) Choose a Pareto optimal action profile a^*. We need to show that starting from any state ϕ in which all players' satisficing levels are very close to $\pi(a^*)$ but their some decision makers' current actions are different from a^*, the process returns to $\Phi(a^*)$ in finite repetitions of P with positive probability. Suppose that for sufficiently long, but finite periods, the same I players are repeatedly selected and play the game. Then, by the first condition of Definition 9**,** the player becomes to choose a^* with positive probability in the repetitions without trembles. Since each population is finite, this procedure applies to all other players so that the whole population can become to choose a^* with positive probability within finite repetitions without trembles.

(ii) Consider a state ϕ in which all players satisfice but their action profile is not Pareto optimal. This implies that there exists a non-Pareto optimal action profile a^1 with which all players satisfice and, furthermore, all players can move to a^1 through finite trembles. Choose a Pareto optimal action profile a^* of the game. By the definition of Pareto optimality, there exists a sequence of unilateral deviations $(a^1, \ldots, a^M = a^*)$ for some finite M. Then the whole population of players can sequentially tremble to coordinate on $a^* = a^M$. ■

Proof of Proposition 16 We show that unilaterally competitive Nash profiles constitute an invariant set of QP^{∞}. First, we show $\Phi(A_C) \equiv \cup_{a \in A_C} \Phi(a)$, where $\Phi(a)$ is defined as in Proposition 12, is invariant. If player i deviates to $a_i \notin \{a_i | a \in A_C\}$, she never satisfices, thus keeps switching actions until she returns to the original action while other players keep choosing their initial action profile. Second, we show that from any state ϕ with $a \notin A_C$, the process enters $\Phi(a^*)$ for some $a^* \in A_C$ with positive probability. Consider an initial state $\phi^1 = (\phi_1^1, \ldots, \phi_I^1) \notin \Phi(a^*)$. By (A.2), without loss of generality we can assume $s_i^1 \leq \pi_i(a^*)$ for all i. If player i trembles to choose a_i^*, since her satisficing level is not higher than her equilibrium payoff, she satisfices with a_i^* regardless of others' choices in the following periods governed by P^∞ by Lemma 2. Other players also can tremble to their equilibrium actions in this way. ∎

A.3. Satisficing Behaviour in Extensive Form Games

Since all satisficing players without trembles eventually satisfice and they experience upward perturbations in satisficing levels, when we characterise invariant sets of QP^∞, we just need to consider states on the paths of which all players satisfice and their satisficing levels are the same as the payoffs from the path if they are following the path every repetition of a game. And, in the proofs, when a player trembles on the path, we implicitly suppose some trembles off the path as required in context have been made before the tremble occurs on the path.

Proof of Proposition 17 We show that if exists, only subgame dominant paths are played by satisficing players, which also implies that weakly dominant paths are played with positive probabilities. Let p^* be a subgame dominant path and define $\Phi^* = \{\phi \in \Phi | p^* = \mathcal{N}(b), v_i(h) = s_i(h) = \pi_i(p^*) \text{ for all } i \in \mathcal{I}(p^*) \text{ and } h \in \mathcal{H}_i(p^*)\}$ and $\overline{\Phi}^* = \{\phi \in \Phi | s_i(h) = \pi_i(p^*) \text{ for all } i \in \mathcal{I}(p^*) \text{ and } h \in \mathcal{H}_i(p^*)\}$. We show that as $\overline{\lambda} \to 0$, $(QP^\infty)(\phi, \Phi^*) \to 1$ for any $\phi \in \overline{\Phi}^*$ and $(QP^\infty)^n(\phi, \Phi^*) > \delta$ for some $n < \infty$ and $\delta > 0$ for all $\phi \in \Phi$.

(i) Suppose the current state is $\phi^* \in \Phi^*$. Without trembles, satisficing players will stay with the path p^* forever. Suppose player $i \in \mathcal{I}(p^*)$ trembles on the path. We show that for any $\phi \in \overline{\Phi}^*$, $P^\infty(\phi, \Phi^*) = 1$ as $\overline{\lambda} \longrightarrow 1$. Pick an arbitrary path p^0. At n^0, $\iota(n^0)$ can choose $a(n^1)$ such that $n^1 \in p^*$ and keep choosing the action for the next finite repetitions of the game with some positive probability δ^0, regardless of whether the player satisfices or not. And, in the following period, $\iota(n^1)$ may choose $a(n^2)$ such that $n^2 \in p^*$ at n^1 and keep choosing the action for the following repetitions of the game with some positive probability δ^1, regardless of whether the

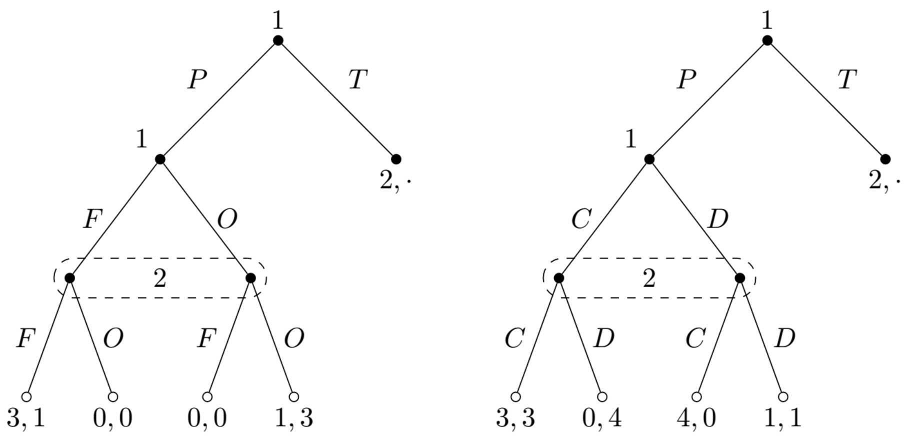


Figure A.3. 2 × 2 Games with Outside Options

player satisfice or not. Since the game is finite, all players who have turns along the path p^* can become to choose actions on the path within finite repetitions of P with some probability at least $\delta^0 \cdots \delta^J > 0$ for some $J < \infty$.

(ii) For the second part, we show that from any state ϕ in which a path is played and satisficing players who move along the path satisfice, players become to play the dominant path p^* in finite repetitions of QP^∞ with positive probability. Suppose that players are satisficing with a path p^0. And, let n' be the first node of p^0 from which p^* differs, i.e., there exists n'' such that $n'' \in p^*$ but $n' \notin p^*$. While players are satisficing with p^0, with a finite sequence of trembles, their behaviours can change so that at all nodes along the dominant path except those of p^0 players choose actions along the dominant path. Then, once player $\iota(n'')$ trembles to choose an action on the dominant path, all subsequent moves will be determined by p^* and all the players who move along the path will satisfice. After then, if those players' satisficing levels are raised up to the payoff $\pi(p^*)$, the transition is completed. For weakly dominant paths, the second argument applies. ■

Proof of Proposition 18 Consider the Battle of the Sexes game with an outside option in the left panel of Figure A.3. We show that only path $((P, F), F)$ is robust to single trembles and attracts players on all other paths. (i) Suppose that both players are satisficing with the path and their satisficing levels are 3 at player 1's two information sets and 1 at player 2's information set. To show that any single tremble cannot make player stays off the path permanently, it is enough to show that from any state, they return to the path as long as their satisficing levels are fixed at 3 and 1 in finite repetitions with positive probability. Suppose that they happen to play $((P, O), O)$. Then, player 1 does not satisfice so that in the following period they may play $((P, F), O)$ and both players do not satisfice, which leads to $((P, F), F)$. At $((T, \cdot), \cdot)$, player 1 does not satisfice and choose P, then they will play $((P, F), F)$ again. (ii) Suppose both players satisfice with $((P, O), O)$ with satisficing levels 1 and 3 at corresponding information sets. If player 1 trembles to choose T at the first information set, player 1 satisfices with the outcome $((T, \cdot), \cdot)$ for the following infinite repetitions of the game without trembles. Then, both players can tremble to choose $((P, F), O)$ or $((P, O), F)$ so that player 2's satisficing level at her information set gets as low as 1 while player 1 keeps choosing T. Then, if both players tremble to choose $((P, F), F)$, both satisfice. ■

Proof of Proposition 19 Consider the Prisoner's Dilemma game with an outside option in the right panel of Figure A.3. We show only one profile $((P, C), C)$ is robust to single trembles and attracts players on all other paths. As in the Battle of the Sexes game, once players satisfice with $((P, C), C)$ with corresponding satisficing levels 3 at all information sets, any single tremble cannot make players stay off the path permanently. Thus, we only need to show that from any path with which players satisfice, they become to play $((P, C), C)$ within finite repetitions of a single tremble and subsequent infinite plays. From $((P, C), D)$ and $((P, D), D)$, players could reach $((P, C), C)$ via $((T, \cdot), \cdot)$ as shown above in the Battle of the Sexes game. Suppose that players satisfice with $((P, D), C)$ with satisficing levels 4 and 0 at corresponding information sets. If player 2 trembles to choose D, since player 2 satisfices with the action at her information set, player 1 keeps switching between $(T, \cdot), (P, C)$ and (P, D) until his satisficing level at his first information set gets as low as 2 and player 1 satisfices with $((T, \cdot), \cdot)$. Then, if both players tremble to choose $((P, C), C)$, both satisfice. ■

Proof of Proposition 20 A two stage Prisoner's Dilemma game with $b = 3$ and $c = 1$ is depicted in Figure A.4. We show that (i) 'always cooperate' path is robust to single trembles and (ii) all other paths are not given the hypothesis of the proposition.

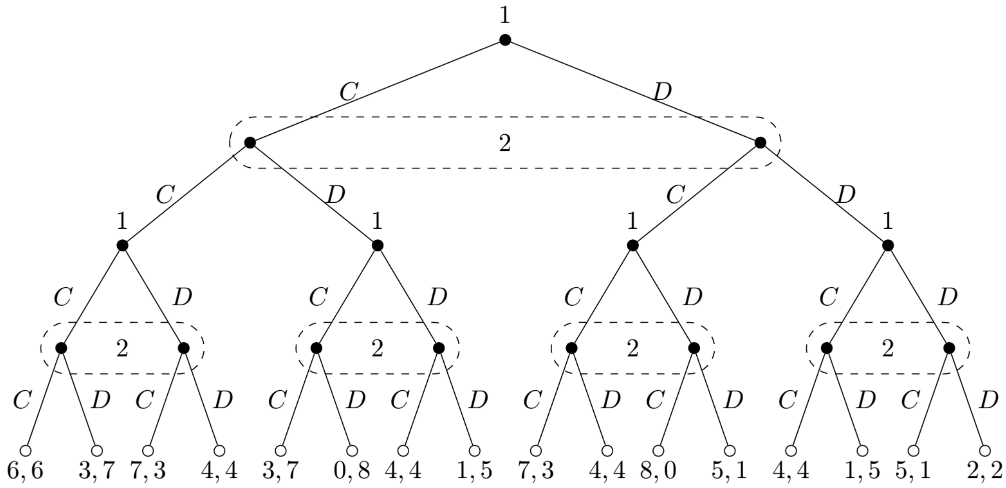


Figure A.4. Repeated Prisoner's Dilemma

(i) It is enough to show that if players' satisficing levels at all information sets on the 'always cooperate' path are fixed at (nb, nb), in any subgame that has its root node on the 'always cooperate' path, players choose mutual cooperation within finite repetitions of the game with positive probability.[11] In those subgames, if players currently choose (C, D) or (D, C) and proceed to terminal nodes, at least one player, player 1 in the case of (C, D) and player 2 in the case of (D, C) by the condition of the proposition, does not satisfice, and the following period players could proceed to (D, D) in the same subgame with positive probability. Then, again by the condition both player do not satisfice and choose (C, C) in the subgame the following period with positive probability.

(ii) Suppose that players are currently satisficing with a path from which both receive payoffs lower than $2b$, the payoffs from the 'always cooperate' path. Among those, the highest payoff for player 2 is $(n-1)b + 2c$ and player 1's corresponding payoff is $(n-2)b + c$, which can be achieved from a path, say p^0, that involves playing (D, D) in the first stage, (C, C) in $n-2$ stages and (C, D) in only one stage. The same payoffs also can be received from a path, say p^1, that starts with (C, D) in the first stage. Then, a single tremble by player 1 can make players switch from p^0 to p^1, where both satisfice, and another tremble by player 2 can make players

[11]For example, in Figure A.4, only the subgame that follows the mutual cooperation in the first stage game is on the 'always cooperate' path but the other proper subgames are not and as long as players' satisficing levels are fixed at 6 at the information sets in the first stage and in the subgame following (C, C) in the first stage, the probability of both players choosing (C, C) is positive in the first stage and the subgame regardless of their current behaviour.

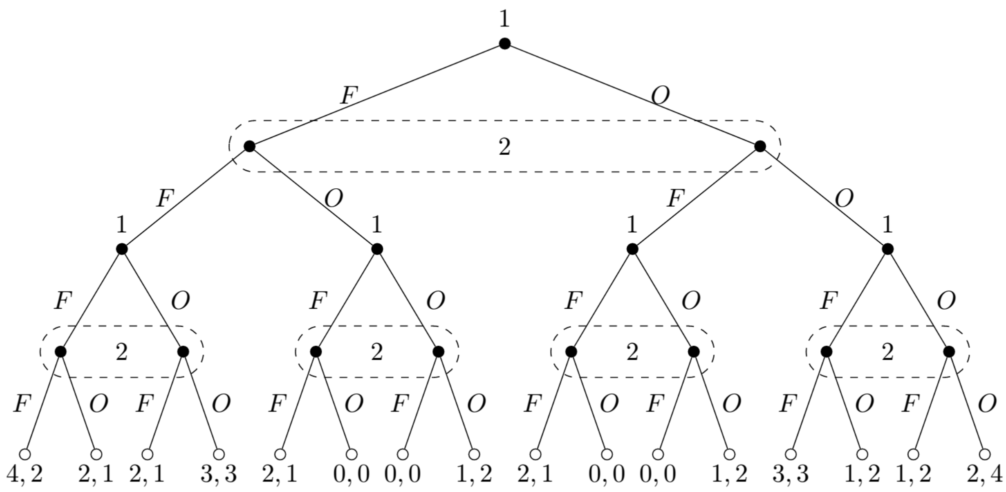


Figure A.5. Repeated Battle of the Sexes

switch from p^1 to the 'always cooperate' path. And, the second highest payoff for player 2 is $(n-1)b+c$, which is also the payoff for player 1 from a path of $n-1$ times of (C, C) and one initial (D, D) in the case players choose (D, D) in the first stage. For these payoffs, we also can find a path that starts with (D, C) in the first stage. Then, the same argument applies. Similarly, we can show that starting from any path that returns payoffs lower than (nb, nb), the process can reach the 'always cooperate' path with finite single trembles.

(iii) Now suppose that players are currently satisficing with a path p^0, from which one player, say player 2, receives a payoff higher than nb. This path must start with (C, C) or (C, D) in the first stage given the hypothesis of the proposition. We consider two paths p^1 and p^2: p^1 (resp. p^2) starts with (D, C) (resp. (D, D)) in the first stage and both return payoffs $(n-1)b+2c$ for player 1 and $(n-2)b+c$ for player 2. Suppose that while players are satisficing with p^0, finite trembles make them choose and satisfice with p^1 or p^2 off the path. Then, a single tremble from p^0 can make player 2 keep switching between C and D in the first stage infinitely until player 2's satisficing level at her information set in the first stage game gets as low as $(n-2)b+c$ while player 1 satisfices with p^1 and p^2. Then, (ii) applies. ■

Proof of Proposition 21 A two stage Battle of the Sexes game with $b=2$ and $c=1$ is depicted in Figure A.5. We show that if the game is repeated $2n$ times, each of (F, F) and (O, O) is selected n times and players' payoffs are $n(b+c)$ most of the times given the hypothesis. The proof is similar to that of Proposition 20. Let P^* denote the set of paths in which players receive the same payoffs $n(b+c)$.

(i) Suppose players' satisficing levels at their information sets along a path $p^* \in P^*$ are the same as $n(b+c)$. Note that if players fail to coordinate in any single stage, the maximum payoff that a player can achieve is $(2n-1)b$. Thus, the condition $b < \frac{n}{n-1}c$ implies that once they fail to coordinate on either (F,F) or (O,O) in any stage, no one can satisfice with respect to the satisficing level $n(b+c)$. Also note that no path exists such that both players receive higher payoffs than $n(b+c)$. Now we construct a sequence through which starting from any path players return to the path $p^* \in P^*$ without trembles within finite repetitions of the game with positive probability as long as their satisficing levels are fixed at $n(b+c)$. Choose an arbitrary path $p^0 \notin P^*$. Along the path, at least one player, say player 1, receives a lower payoff than his satisficing level $n(b+c)$ at the information set in the first stage. The following period in the first stage player 1 continues or switches to the action on p^* in the first stage. If player 2 chooses the other action in the first stage, her payoff will be lower than her satisficing level $n(b+c)$ at the information set, thus the following period she also choose the same action as player 1. This procedure can be repeated downward along the path p^*.

(ii) Now consider the cases in which players satisfice with a path. First, suppose players are satisficing with a path from which both players receive payoffs lower than $n(b+c)$. Then, we can simply find a sequence of paths through which players move to a path in P^* with finite trembles: in every subgame where players fail to coordinate, they could tremble to coordinate so that both players' payoffs are not greater than $n(b+c)$. This works because the Battle of the Sexes games has Pareto optimal outcomes that can be reached with one single tremble from any non-Pareto outcome. Second, suppose that players are satisficing with a path p^0 in which a player, say player 2, receives a payoff higher than $n(b+c)$. Then, player 1's maximum payoff is $(n-1)b+(n+1)c$. Now consider two paths p^1 and p^2 that start with (O,F) (or (F,O)) and (O,O) in the first stage and return payoffs $((2n-1)b, c)$ and $(n(b+c), n(b+c))$, respectively. While players are satisficing with p^0, some trembles can make players choose and satisfice with p^1 and p^2 once they reach the paths. Then, because of the condition $\frac{n+1}{n}c \leq b$, a single tremble from p^0 can make player 2 keep switching between F and O in the first stage infinitely until player 2's satisficing level at her information set in the first stage game gets as low as $n(b+c)$. Then, the above argument applies. ■

Proof of Proposition 22 Let $\mathcal{G}_S$ be a signalling game and B^* be the set of all complete profiles. We show that starting from any profile, players become to choose a complete profile within finite repetitions of QP^∞, and from any complete profile, a single tremble and subsequent infinite repetitions of the game without trembles lead to the initial or another complete profile with probability 1 as $\overline{\lambda}$ goes to 0.

(i) Let $T^H(b^0) \in T$ denote the set of states from which both players achieve the higher payoffs H given a profile b^0, i.e., $b_2^0(b_1^0(t)) = f(t)$ for $t \in T^H(b^0)$. We show that starting from an arbitrary ϕ^0 with b^0 and $T^H(b^0) \neq T$ in which both players satisfice for all $t \in T$, which does not mean their payoffs are H, through finite repetitions of QP^∞, the process reaches a set of states in which a complete profile is played and both players' satisficing levels are higher than the lower payoff L. Choose a Nature's action $t^0 \notin T^H(b^0)$ from which both players receive the lower payoffs according to their profile b^0. Suppose Sender trembles to choose an action $m^0 \notin \{m \in M | b_1^0(t) = m, t \in T^H(b^0)\}$ at the information set t^0. If $b_2^0(m) = f(t^0)$, the modified profile is referred to as b^1 and $t^0 \in T^H(b^1)$. Otherwise, through the infinite repetitions of P, both players satisfice with b^0. Then, if Receiver trembles to choose the action $a = f(t^0)$ at the information set m^0, the modified profile is referred to as b^1, and $t^0 \in T^H(b^1)$. Subsequently, both players' satisficing levels are raised up to H from L by two consecutive repetitions of QP^∞. This process can be repeated whenever $T^H(b') \neq T$ for some incomplete profile b'.

(ii) For the second part, it is enough to show that starting from any state in which both players' satisficing levels at all information sets are fixed and higher than L, the process will reach a state in which their profile is complete in infinite repetitions of the game with probability 1. The first part of proof shows that it takes place with finite repetitions of QP^∞. Similarly, if the satisficing levels are higher then L, the players become to play a complete profile in finite repetitions of the game without trembles with positive probability. Therefore, as long as their satisficing levels are fixed above L, they eventually play a complete profile with probability 1. ■

APPENDIX B

Equilibrium Strategies of Informative Communication

B.1. Equilibrium Outcomes with Different Distributions of Types

We construct equilibrium strategies under different priors to see the relation between equilibrium strategy and distribution of the decision maker's type. Equilibrium strategies for different distributions of types are depicted in Figure B.1. The distribution functions are Beta(α, β) with different sets of parameters, such as $(1, 3), (2, 3)$ and $(2, 2)$.[1] In Beta(1,3), the population is concentrated on left types or the distribution is skewed to the right. In Beta(2,2), the population is symmetrically distributed and concentrated on centre types. All distributions are uni-modal.[2] The other priors F and H are fixed as $\mathrm{U}[0, 1]^2$ and Beta(2,2), respectively.

Each equilibrium has been solved numerically by searching for a slope $\Delta \in [0, 2\pi]$ such that the expert is indifferent between two messages generated by the hyperplane with the slope when the interior point c is fixed at $(1/2, 1/2)$. Thus, given an equilibrium slope Δ^*, two messages m^+, m^- are "$\theta_L - \frac{1}{2} \leq \tan \Delta^*(\theta_H - \frac{1}{2})$" and "$\theta_L - \frac{1}{2} > \tan \Delta^*(\theta_H - \frac{1}{2})$," respectively. The solutions are as follows. First, when $\alpha = 1$ and $\beta = 3$, $\Delta \approx \frac{1}{10}\pi$, $\mathbb{E}[(\theta_L, \theta_H)|m^+] = (0.4402, 0.7393)$ and $\mathbb{E}[(\theta_L, \theta_H)|m^-] = (0.5598, 0.2607)$. Second, when $\alpha = 2$ and $\beta = 3$, $\Delta \approx \frac{1}{5}\pi$, $\mathbb{E}[(\theta_L, \theta_H)|m^+] = (0.3785, 0.7057)$ and $\mathbb{E}[(\theta_L, \theta_H)|m^-] = (0.6215, 0.2943)$. Last, when $\alpha = 2$ and $\beta = 2$, $\Delta = \frac{1}{4}\pi$, $\mathbb{E}[(\theta_L, \theta_H)|m^+] = (0.3333, 0.6667)$ and $\mathbb{E}[(\theta_L, \theta_H)|m^-] = (0.6667, 0.3333)$. The estimates derived from these are shown in the right-hand column of Figure B.1.

Comparing the figures, a relation between the distribution of types and the decision maker's estimate induced by the equilibrium messages is noticeable. As the population is more concentrated on left types, the estimates $e^+(t)$ and $e^-(t)$ by the decision maker of left types move closer to the uninformed. For message m, we can interpret the difference between estimates with and without the message $|\mathbb{E}[\theta(t)|m] - \mathbb{E}[\theta(t)]|$ as the amount of information contained in the message. This

[1] The beta distribution is flexible in representing various distributions on the unit interval such as uni-/bi-modal, skewed to the left/right and symmetric distributions.

[2] See the graphs on the left in the figure.

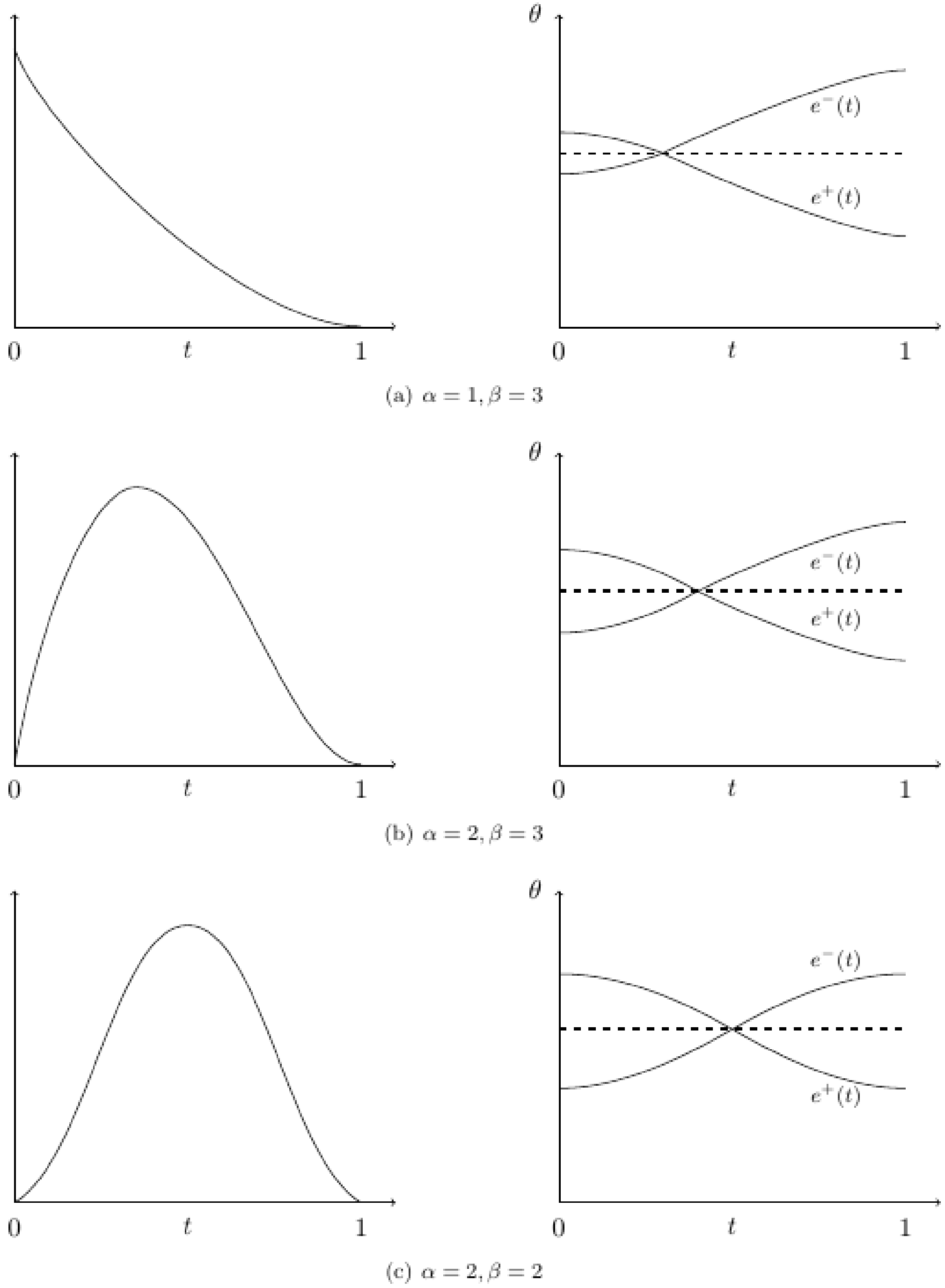


Figure B.1. Distributions of Types and Equilibrium Outcomes

suggests that a message from the expert is more valuable to the decision maker of right types than left types if the population is skewed to the right. The opposite also holds true when the population is skewed to the left.

APPENDIX C

Simulation of Satisficing Behaviour

Here we describe the common simulation setup and provide simulation results for individual decision problems, normal form games and extensive form games. All simulations are calculated with the same distributions and parameter values except where otherwise stated.

The trembling probabilities and persistence parameter are set to 0.05 or 0.01 and the same parameter values are applied to all satisficing players in each simulation. We let the persistence parameter have higher values, around 0.7, in early periods, then decreases over time to 0.05 or 0.01. The distributions of weighting coefficients ρ, λ and λ^{ϵ} are the same at all states as $\rho_{n,i} = p^{\gamma}, \lambda_{n,i} = q^{\gamma}\overline{\lambda}$ and $\lambda^{\epsilon}_{n,i} = r^{\gamma}$ where p, q and r are independently drawn from the Uniform distribution on $[0, 1]$ and γ is set to 1/2. Initial actions are chosen with equal probabilities for all actions, initial satisficing levels are set to 0 or higher than maximum payoffs of the given decision problems[1] and initial valuations are 0. Satisficing players satisfice only when valuations are not lower than satisficing levels.[2]

Each period satisficing players in games are matched with fixed or random opponents. In fixed matching, players play with the same opponents for all periods while in random matching players are uniformly randomly matched within each cohort each period. For example, in 2×2 games under random matching in Appendix C.2, 500 satisficing players are generated and each period 250 pairs are randomly matched and play the games. And, in Best Shot and Ultimatum games in Appendix C.5, 6 players for each role within a cohort, in total 12, are generated and all players are randomly matched according to their roles each period.

[1]Since the persistence parameter values are high in early periods, the high initial satisficing levels generally do not have effects on the long run behaviour of satisficing players.
[2]We assume the choice rule allows agents to choose their current actions with probabilties close to 1 when valuations are slightly lower than satisficing levels. We do not explicitly incorporate the property into the simulations but this does not violate the assumption because the simulations are calculated only with finite precision.

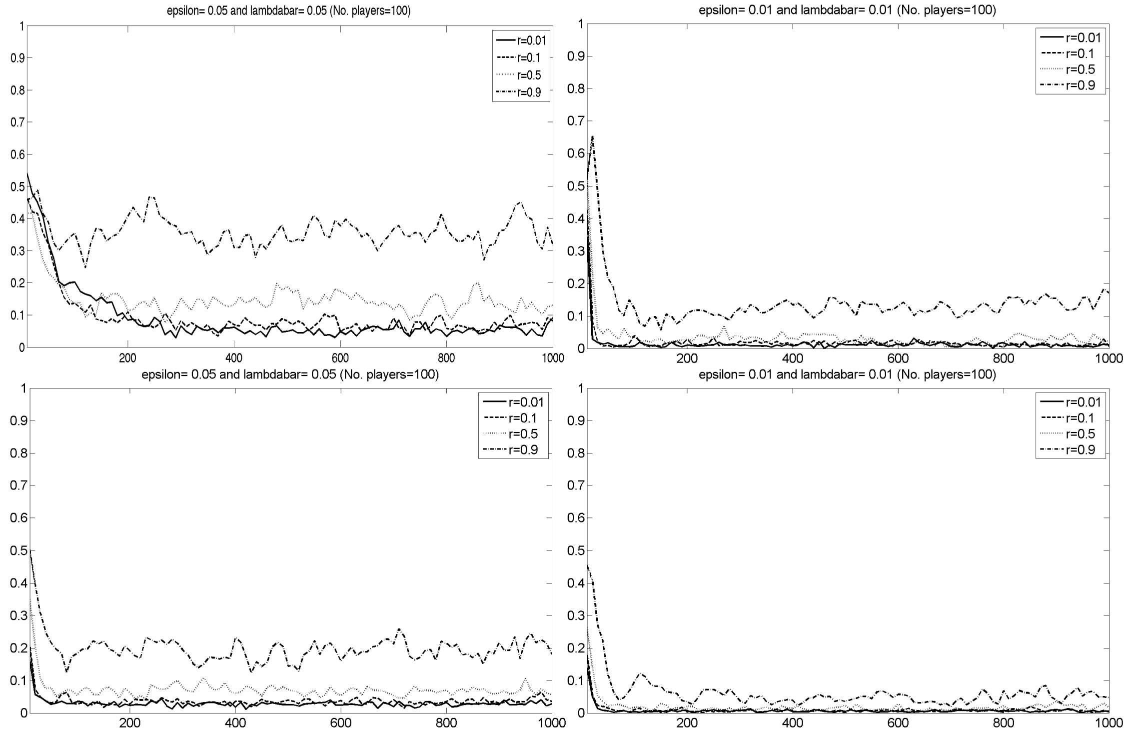


Figure C.1. Proportions of Risky Choice for Gains

Satisficing players may have memory in the form of valuations of all actions as described in the subsection 2.1.3. With the memory, valuations are updated with the same distributions of random coefficients ρ and when players do not satisfice, they choose actions, including the current, that are associated with highest perceived valuations, which are given as $v^j + z^j$, where v^j is valuation of action j and z^j is a shock on the valuation drawn from the normal distribution with mean 0 and standard deviation 1/2. Shocks are independent across actions and players.

C.1. Individual Decision Problem

Following March (1996), we consider binary decision problems between safe and risky options for gains or losses: the safe option returns payoff 1 (or -1 for loss) with certainty and the risky option returns payoff $1/r$ (or $-1/r$ for loss) and 0 with probabilities r and $1-r$, respectively. r is given as one of $0.01, 0.1, 0.5$ and 0.9. Both options have the same expected payoff, 1. The higher r is, the more risky the option is for both gains and losses.

The proportions of risky options being chosen for gains and losses are given in Figure C.1 and C.2. The left charts are with $\epsilon = \overline{\lambda} = 0.05$ and the right charts are with $\epsilon = \overline{\lambda} = 0.01$, and the top charts are without memory and the bottom charts

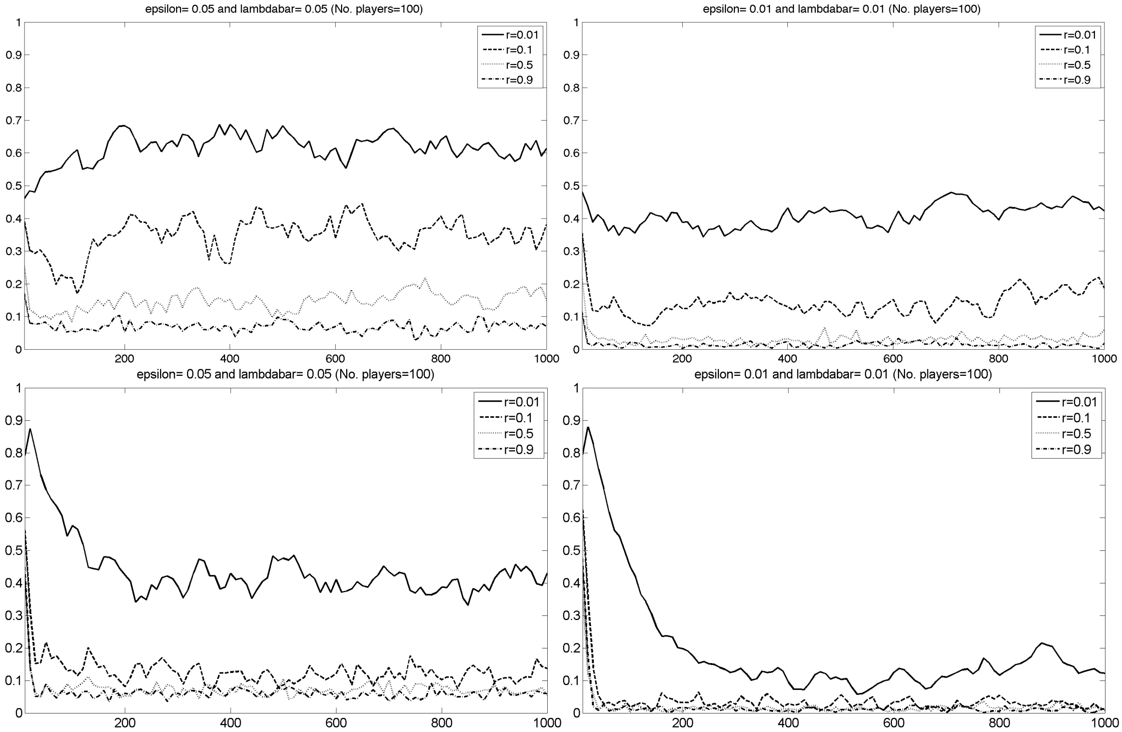


Figure C.2. Proportions of Risky Choice for Losses

are with memory. The initial satisficing levels are set to 10. As noted in March (1996), the simulated satisficing players seem to show dual risk attitudes, i.e., risk aversion for gains and risk seeking for losses. For gains, the safe option is always preferred to risky options and the less risky options are preferred to the more risky options. In contrast, for losses, the more risky options are preferred to the less risky options, but only the most risky option is clearly preferred to the safe option when $\epsilon = \overline{\lambda} = 0.05$. And, as the parameter values lower or memory is augmented, satisficing decision makers more prefer the safe option to the risky options.

C.2. Prisoner's Dilemma, Stag Hunt and Common Interest

For simulation in normal form games, we first consider Prisoner's Dilemma, a modified Prisoner's Dilemma, Stag Hunt and Common Interest games under fixed and random matching setups. The payoff matrices are given in Table C.1. In the modified Prisoner's Dilemma game, a single defector receives higher payoff, 5, than the Prisoner's Dilemma.

Under fixed matching, 500 pairs of satisficing players are simulated and each player plays the games against the same opponent for 8,000 periods. Figure C.3 shows frequencies of mutual cooperation and defection under fixed matching. And,

Figure C.4 shows frequencies of choosing 'cooperate' and 'defect' under random matching setup. In both setups, initial satisficing levels are set to 0. In both figures, the left is calculated with $\epsilon = \overline{\lambda} = 0.05$ and the right is with $\epsilon = \overline{\lambda} = 0.01$.

	C	D
C	$2,2$	$0,3$
D	$3,0$	$1,1$

	C	D
C	$2,2$	$0,5$
D	$5,0$	$1,1$

	C	D
C	$2,2$	$0,2$
D	$2,0$	$1,1$

	C	D
C	$2,2$	$0,0$
D	$0,0$	$1,1$

Table C.1. Payoffs of Prisoner's Dilemma, Stag Hunt and Common Interest Games

Under the fixed matching setup, players in the long run manage to coordinate on mutual cooperation and defection. In particular, when $\epsilon = \overline{\lambda} = 0.01$, players choose mutual cooperation or defection most of the time, above 95%, in all games. The proportions of cooperate and defect differ across games. In the Common Interest game, 'cooperate' is played more often than 'defect,' 70% with $\epsilon = \overline{\lambda} = 0.05$ and 92% with $\epsilon = \overline{\lambda} = 0.01$. In other games, 'defect' is played more often. Among the Prisoner's Dilemma and Stag Hung games, 'cooperate' is played most often in the Prisoner's Dilemma and least often in the Stag Hunt game.[3]

Under the random matching setup, 'cooperate' is most preferred in the Common Interest game, and among others the frequency of choosing 'cooperate' is highest in the modified Prisoner's Dilemma and lowest in the Stag Hunt, that is, as benefit from a single defection from mutual cooperation increases, players more often choose 'cooperate' action rather than 'defect' action.

Next we simulate satisficing plays in three Stag Hunt games of Battalio et al. (2001). The payoff matrices are given in Table C.2 and the frequencies of choosing X and Y in the first and last, 75th, period is in Table C.3. They define *optimization premium* as the difference between the payoffs of two responses given opponent's action, which is highest in $2R$ treatment and lowest in $0.6R$ treatment. And the experiment results conform to their predictions that "behaviour will converge to an equilibrium more quickly the larger is the optimization premium" and "behaviour

[3]Higher frequency of cooperate in the Prisoner's Dilemma than modified Prisoner's Dilemma is consistent with Cho and Matsui (2005). However, under random matching, the order reverses.

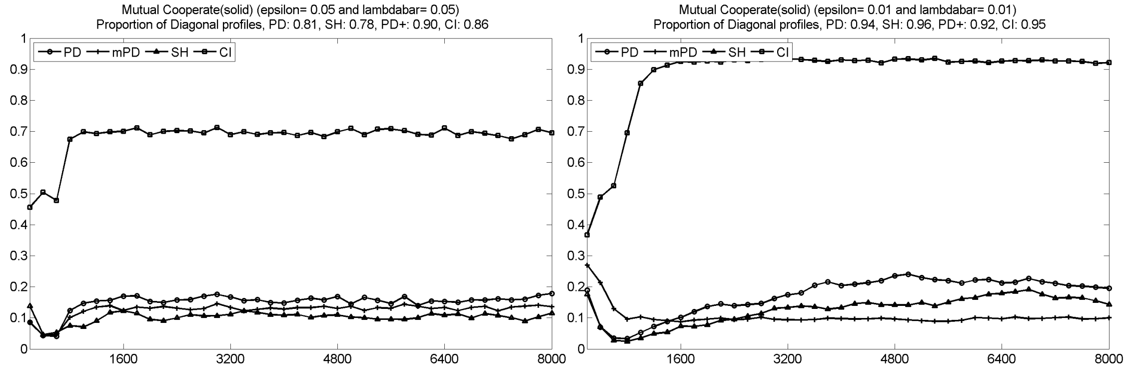


Figure C.3. Proportions of Mutual Cooperate under Fixed Matching

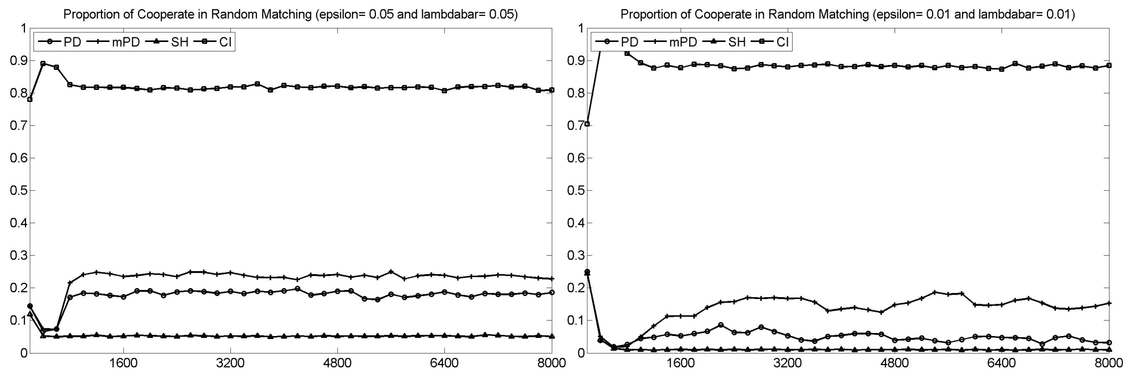


Figure C.4. Proportions of Cooperate under Random Matching

is more likely to converge to the payoff dominant equilibrium the smaller is the optimization premium."

	X	Y		X	Y		X	Y
X	$45, 45$	$0, 42$	X	$45, 45$	$0, 40$	X	$45, 45$	$0, 35$
Y	$42, 0$	$12, 12$	Y	$40, 0$	$20, 20$	Y	$35, 0$	$40, 40$

Table C.2. Payoffs of Stag Hunt games of Battalio et al. (2001): $0.6R$, R and $2R$

	X	Y		X	Y
$0.6R$	64%	36%	$0.6R$	44%	56%
R	70%	30%	R	25%	75%
$2R$	53%	47%	$2R$	5%	95%
Total	63%	37%	Total	24%	76%

Table C.3. Contingency Table of Stag Hung games: Period 1 and 75

In simulation, as in the experiments, 4 pairs of satisficing players are randomly matched within each cohort of size 8. All other parameters are the same as before.

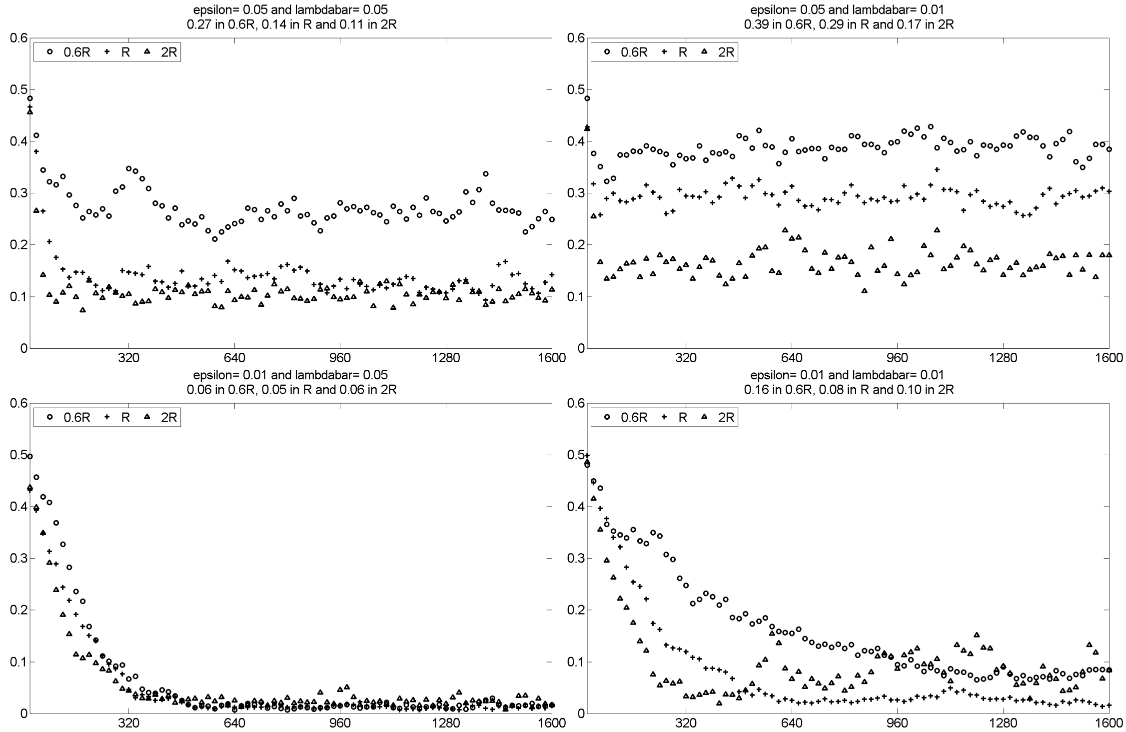


Figure C.5. Proportions of Cooperate in Stag Hung games

The simulation results in Figure C.5 are averages of 30 cohorts. The simulation results are also consistent with the experimental results as well as the predictions of Battalio et al. (2001).

C.3. Matching Pennies

Goeree et al. (2003) consider a 2×2 game in Table C.4, which has a unique mixed Nash equilibrium but its observed choice frequencies systematically differ from the Nash equilibrium. Furthermore, any quantal response equilibrium is far away from the observed data: Quantal response equilibria predict the risky action, R, would be played more than 50% of the time for any parameter value, but participants in the experiment chose the safe action, L, about 65% of the time. Goeree et al. (2003) show that if risk aversion is incorporated in the quantal response equilibrium, it explains the data very well.

	L	R
U	$200, 160$	$160, 10$
D	$370, 200$	$10, 370$

Table C.4. Payoffs of Matching Pennies game of Goeree et al. (2003)

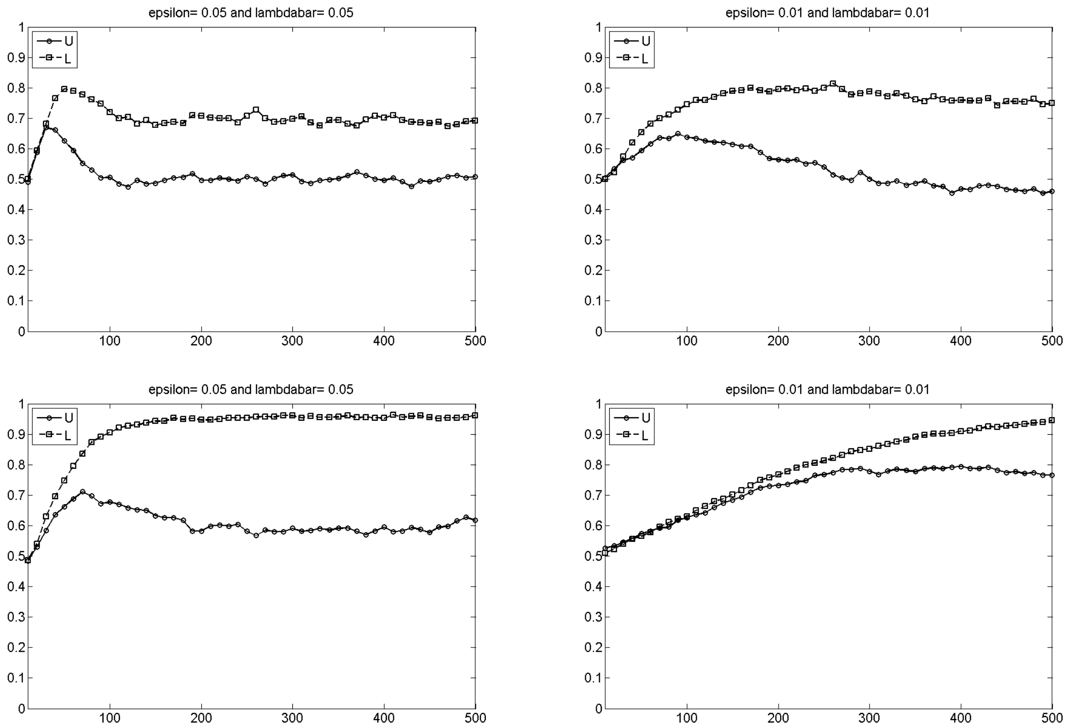


Figure C.6. Simulation of Matching Pennies of Goeree et al. (2003)

The experiment was run with 6 cohorts of 10-12 participants each and they were randomly matched for 10 periods with a fixed role. The observed choice frequencies of the row player choosing U and the column player choosing L are 0.47 and 0.65, respectively. The simulation results with cohort size fixed at 10 under the random matching setup are given in Figure C.6. Initial satisficing levels are set to 10. The top charts are without memory and the bottom charts are with memory. Satisficing behaviour without memory fits the observed data relatively well in its magnitude and direction, which could be explained by that the satisficing behaviour inherently has risk aversion property as shown in the individual decision problems. It seems that memory does not play a role in this 2×2 game.

C.4. Public Good Provision

We simulate satisficing behaviour in public good provision games in the setup of Isaac et al. (1984). They designed experiments in which each participant chooses how much to contribute to public good for 10 periods with payoff function $\pi_i = E - c_i + a(\sum_{i=1}^{N} c_i)/N$, where $E = 62, 25$ or 10 are individual endowments, c_i is the amount of contribution made by player i, $a = 1.2, 3$ or 7.5 are total return from public good and $N = 4$ or 10 are cohort sizes. Return from public good per capita is defined as $M = a/N$. For inexperienced and experienced participants, four experiments with $(N, M, a) = (4, 0.3, 1.2), (4, 0.3, 1.2), (10, 0.3, 3)$ and $(10, 0.75, 7.5)$ were run. The percentages of total contribution over the total endowment $100 \times \sum_{i=1}^{N} c_i/(N \cdot E)$ over all periods are given in Table C.5.

(%)	$M = .3$	$M = .75$
$N = 4$	26.36	65.1
$N = 10$	32.88	65.1

(%)	$M = .3$	$M = .75$
$N = 4$	12.08	49.6
$N = 10$	33.76	53.8

Table C.5. Observed Contributions by Inexperienced and Experienced in Isaac et al. (1984)

In simulation, we allow satisficing players to choose an arbitrary integer between 0 and E as their contribution level each period. Thus, the size of players' action sets are $E + 1$. We simulate satisficing behaviour with four different sets of (N, M, a) as described above without distinguishing inexperienced and experienced players. For details of each experiment setup, refer to Table 1 of Isaac et al. (1984). The initial satisficing levels are set to 0. The simulation results without and with memory are given in the top and bottom of Figure C.7.

Ledyard (1994) summarises the experiment results of Isaac et al. (1984) as (1) increasing M from 0.3 to 0.75 increases the rate of contribution in all cases, (2) inexperienced subjects contribute more and (3) repetition decreases and group size increases contributions for low $M = 0.3$ but neither seem to have any effect if $M = 0.75$. The simulation results seem to partially fit the experiment results: The fitness depends on the set of parameter values, memory and the experimental setups. Note that in the Public Good Provision game, convergence to either full contribution or no contribution outcomes, which is the unique Nash outcome, is not guaranteed by satisficing players unlike its two-player version, Prisoner's Dilemma game.

C.5. Extensive Form Games: Best Shot and Ultimatum

We consider two extensive form games: Best Shot and Ultimatum. Best Shot game is a sequential public good provision game in which player 1 chooses how much to contribute, then after observing player 1's choice, player 2 chooses how

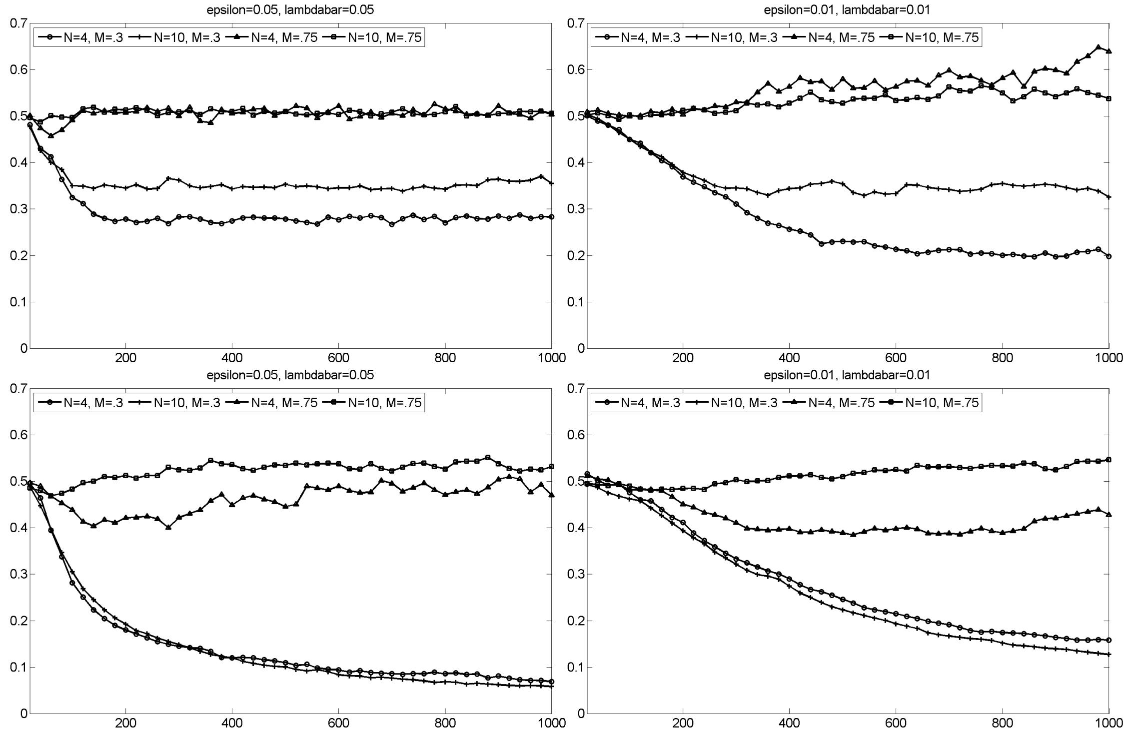


Figure C.7. Simulation of Public Good Provision of Isaac et al. (1984)

much to contribute. Their payoffs are given in Table 1 in Roth and Erev (1995). The subgame perfect outcome is for player 1 to contribute nothing and for player 2 to contribute 4 units and in experiments most participants choose the equilibrium outcome. We model satisficing players to choose $x \in \{0, \ldots, 9\}$. First, at his unique decision node, player 1 chooses his contribution level x_1, then at the decision node following x_1, player 2 chooses x_2. Player 2 has 10 decision nodes. Thus, player 1's behaviour is given as $x_1 \in \{0, \ldots, 9\}$ with one satisficing level while player 2's behaviour is given as a function from $\{0, \ldots, 9\}$ to $\{0, \ldots, 9\}$ with satisficing levels for each decision node.

In Ultimatum game, player 1 offers $x \in \{1, \ldots, 9\}$ to player 2, then player 2 chooses whether to accept or reject the offer. If the offer is accepted, player 1 and 2's payoffs are $10 - x$ and x, respectively, and otherwise both receive 0. Player 1 has one decision node and his behaviour is given as $x \in \{1, \ldots, 9\}$ while player 2's behaviour is given as a function from $\{1, \ldots, 9\}$ to $\{accept, reject\}$, which is different from the cut-off strategy used by Roth and Erev (1995).[4] This assumption

[4]In a cut-off strategy, player 2 sets a threshold value and accepts player 1's offer only when the offer is not smaller than the value. This strategy transforms the Ultimatum game into a normal form game.

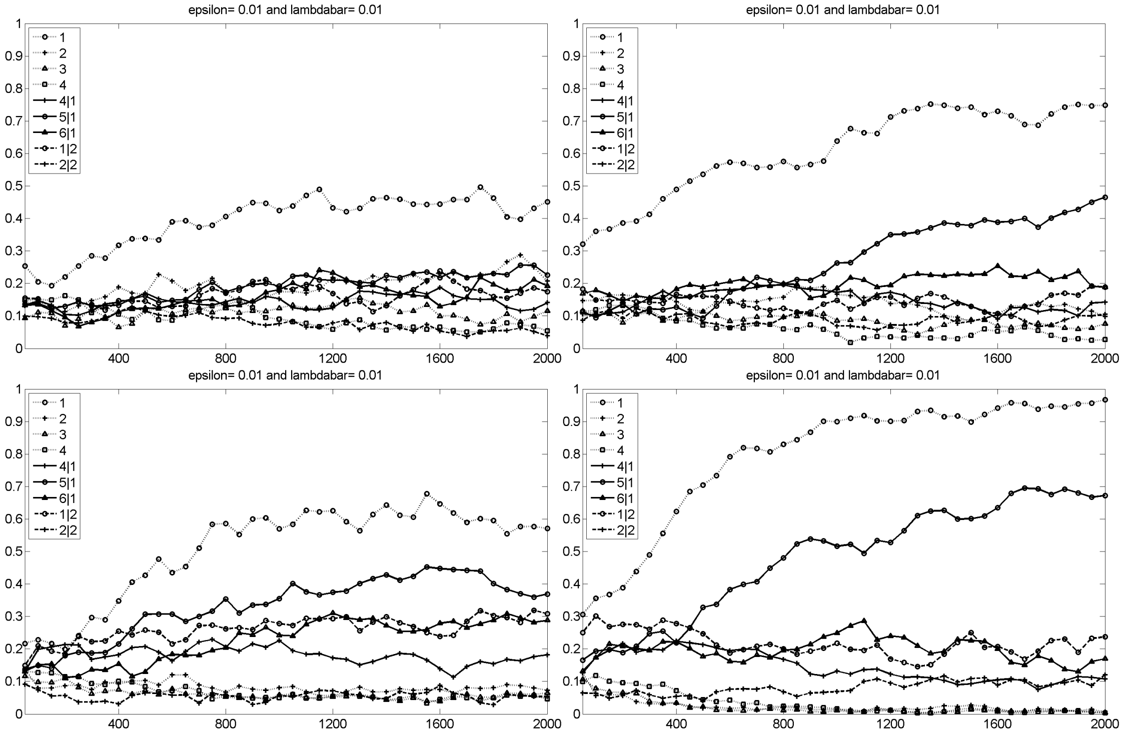


Figure C.8. Simulation of Best Shot

on player 2's behaviours does not impose any restriction on player 2's choices across her decision nodes, thus the same player could accept x but reject y with $x < y$. Initial satisficing levels are set to 5.

The simulation results are given in Figure C.8 and C.9. The top charts are under fixed matching setup and the bottom charts are under random matching, and the left charts are without memory and the right charts are with memory. In the Best Shot game, series $1, 2, 3$ and 4 represent the frequencies of player 1 contributing $0, 1, 2$ and 3 units, respectively, and series 4|1,5|1 and 6|1 represent the frequencies of player 2 contributing $3, 4$ and 5 units after observing player 1 has contributed 0 unit. In the Ultimatum game, series x represents the frequencies of player 1 offering $x - 1$ units. Cohort size is set to 6 and the results are averaged over 30 cohorts.

In both games, simulation results with memory fit well the observed behaviour: subgame perfect outcome in the Best Shot game and non-subgame perfect outcomes, intermediate shares being most frequently offered, in the Ultimatum game.

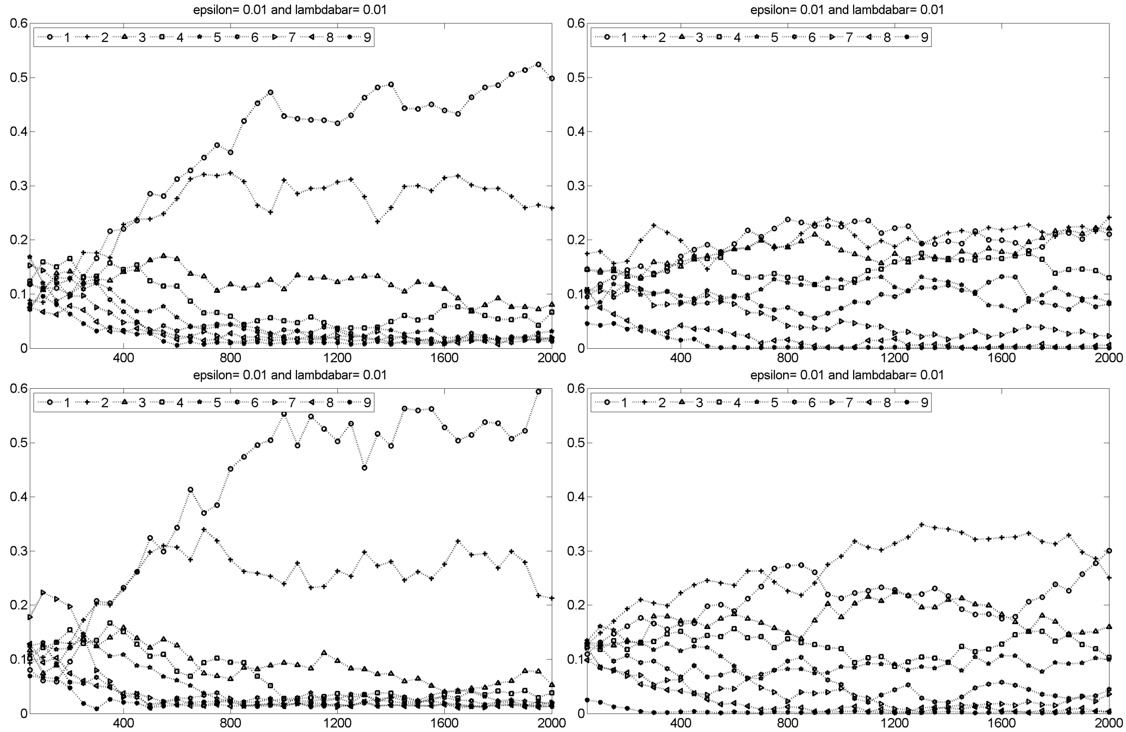


Figure C.9. Simulation of Ultimatum

C.6. Extensive Form Games: Signalling

Following McKelvey and Palfrey (1998), we simulate three signalling games, Game 2, 3 and 4 of Banks et al. (1994), who investigate refinements of Nash equilibrium in two-person signalling games like Bayes-Nash, sequential and intuitive, and find experiment subjects select the more refined equilibria. In all games, Nature chooses a type, Sender privately observes the type and sends a message, and Receiver chooses an action, which determines both Sender and Receiver's payoffs as in Table C.6. The observed frequencies of individual choice at all information sets are summarised by McKelvey and Palfrey (1998) in Table C.7.

Figure C.10, C.11 and C.12 show the proportions of satisficing players' individual choices at all decision node. In figures, '$m|t$' and '$a|m$' denote choice of m at decision node t and choice of a at decision node m, respectively. In each simulation, 6 satisficing players for each role are simulated for each cohort, in total 30 cohorts, and each period 3 pairs of players are randomly matched and play the signalling games for 1,000 periods with memory and $\epsilon = \overline{\lambda} = 0.05$. Generally, the simulation results seem to fit well the observed data.

m_1	a_1	a_2	a_3	m_2	a_1	a_2	a_3	m_3	a_1	a_2	a_3
BCP # 2 (Nash versus sequential)*											
t_1	1, 2	2, 2	0, 3	t_1	1, 2	1, 1	2, 1	t_1	3, 1	0, 0	2, 1
t_2	2, 2	1, 4	3, 2	t_2	2, 2	0, 4	3, 1	t_2	2, 2	0, 0	2, 1
BCP # 3 (sequential versus intuitive)†											
t_1	0, 3	2, 2	2, 1	t_1	1, 2	2, 1	3, 0	t_1	1, 6	4, 1	2, 0
t_2	1, 0	3, 2	2, 1	t_2	0, 1	3, 1	2, 6	t_2	0, 0	4, 1	0, 6
BCP # 4 (intuitive versus divine)‡											
t_1	4, 0	0, 3	0, 4	t_1	2, 0	0, 3	3, 2	t_1	2, 3	1, 0	1, 2
t_2	3, 4	3, 3	1, 0	t_2	0, 3	0, 0	2, 2	t_2	4, 3	0, 4	3, 0

*Nash: (m_1, a_2, a_2, a_2) Seq: (m_3, a_2, a_2, a_1).
†Sequential: (m_2, a_1, a_3, a_1) Intuitive: (m_1, a_2, a_1, a_1).
‡Intuitive: (m_2, a_3, a_3, a_2) Divine: (m_3, a_2, a_2, a_1).

Table C.6. Payoffs of Signalling Games in Banks et. al. (1994), reproduced by McKelvey and Palfrey (1998)

		BCP #2				BCP #3				BCP #4			
		n	f_i	QRE	NNM	n	f_i	QRE	NNM	n	f_i	QRE	NNM
t_1	m_1	7	.184	.103	.092	20	.500	.497	.817	2	.067	.274	.217
	m_2	0	.000	.021	.092	14	.350	.339	.092	9	.300	.221	.217
	m_3	31	.816	.875	.817	6	.150	.165	.092	19	.633	.504	.567
t_2	m_1	11	.324	.176	.092	45	.900	.927	.817	21	.700	.422	.217
	m_2	0	.000	.026	.092	4	.080	.057	.092	0	.000	.032	.217
	m_3	23	.677	.798	.817	1	.020	.016	.092	9	.300	.546	.567
m_1	a_1	0	.000	.049	.092	4	.062	.153	.092	9	.391	.298	.217
	a_2	18	1.000	.837	.817	61	.939	.704	.817	13	.565	.593	.567
	a_3	0	.000	.113	.092	0	.000	.142	.092	1	.044	.109	.217
m_2	a_1	0		.184	.092	14	.778	.686	.817	0	.000	.044	.217
	a_2	0		.797	.817	0	.000	.174	.092	4	.444	.647	.567
	a_3	0		.019	.092	4	.222	.140	.092	5	.556	.308	.217
m_3	a_1	53	.841	.726	.817	7	1.000	.999	.817	23	.821	.704	.567
	a_2	0	.000	.026	.092	0	.000	.001	.092	5	.179	.235	.217
	a_3	10	.159	.248	.092	0	.000	.000	.092	0	.000	.062	.217
	$\lambda \mid \gamma$			2.249	.725			1.598	.725			1.193	.350
	$\lambda_{lo} \mid \gamma_{lo}$			1.901	.627			1.441	.634			.099	.216
	$\lambda_{hi} \mid \gamma_{hi}$			2.715	.809			1.864	.802			1.470	.479
	$-\mathcal{L}^*$			83.390	92.224			101.050	108.628			95.65	118.151

Table C.7. Observed Outcomes in Signalling Games in Banks et. al. (1994), reproduced by McKelvey and Palfrey (1998)

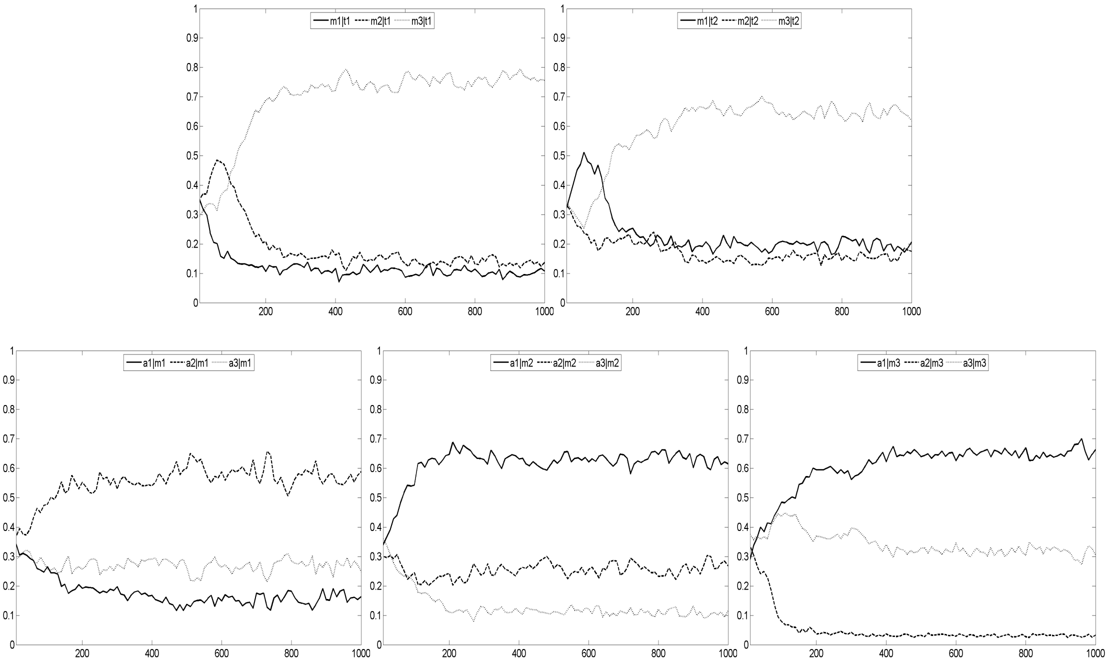


Figure C.10. Simulation of Signalling Game 2 of Banks et. al. (1994)

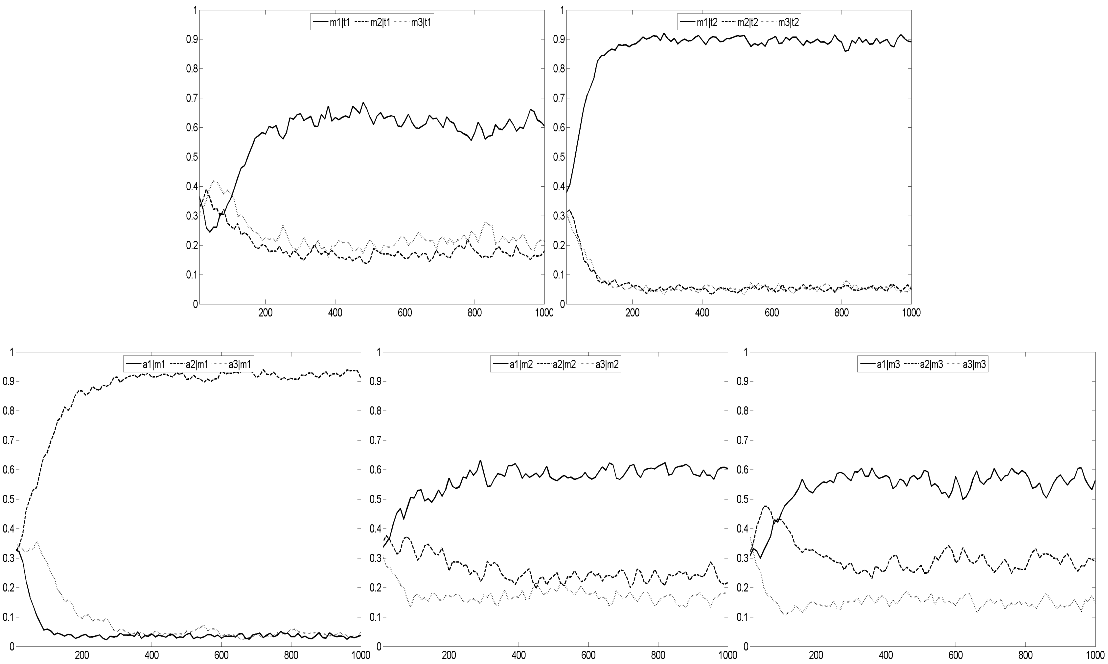


Figure C.11. Simulation of Signalling Game 3 of Banks et. al. (1994)

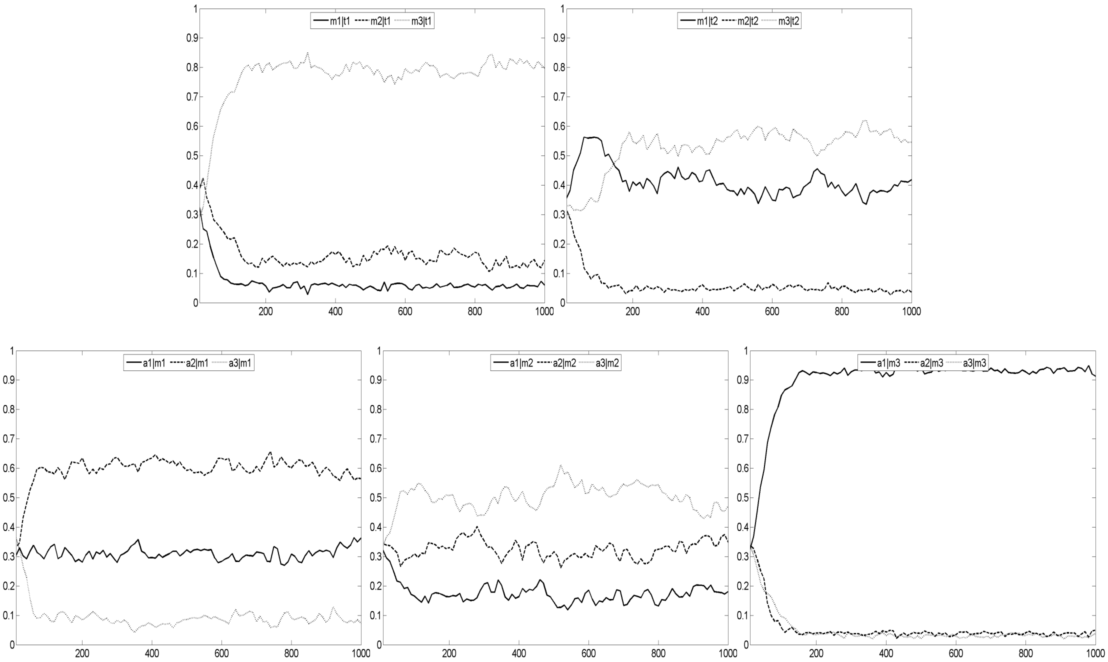


Figure C.12. Simulation of Signalling Game 4 of Banks et. al. (1994)