Essays on Games of Strategic Substitutes with Incomplete Information

By

Eric Hoffmann

Submitted to the Department of Economics and the
Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Professor Tarun Sabarwal, Chairperson

Distinguished Professor Bernard Cornet

Committee members

Professor Joseph Sicilian

Professor Jianbo Zhang

Professor David Nualart

Date defended: _____April 21, 2015_____

The Dissertation Committee for Eric Hoffmann certifies
that this is the approved version of the following dissertation :

Essays on Games of Strategic Substitutes with Incomplete Information

_____

Professor Tarun Sabarwal, Chairperson

Date approved: _____ April 21, 2015 _____

Abstract

This dissertation consists of three individual chapters. The first chapter applies lattice theoretic techniques in order to establish fundamental properties of Bayesian games of strategic substitutes (GSS) when the underlying type space is ordered either in increasing or decreasing first-order stochastic dominance. Existence and uniqueness of equilibria is considered, as well as the question of when such equilibria can be guaranteed to be monotone in type, a property which is used to guarantee monotone comparative statics. The second chapter uses the techniques of the first and combines them with the existing results for strategic complements (GSC) in order to extend the literature on global games under both GSC and GSS. In particular, the model of Carlsson and Van Damme (1993) is extended from 2×2 games to GSS or GSC involving a finite amount of players, each having a finite action space. Furthermore, the possibility that groups of players receive the same signal is allowed for, a condition which is new to the literature. It is shown that under this condition, the power of the model to resolve the issue of multiplicity is unambiguously increased. The third chapter considers stability of mixed strategy Nash equilibria in GSS.

Chapter 1 analyzes Bayesian games of strategic substitutes under general conditions. In particular, when beliefs are order either increasingly or decreasingly by first order stochastic dominance, the existence and uniqueness, monotonicity, and comparative statics in this broad class of games are addressed. Unlike their supermodular counterpart, where the effect of an increase in type augments the strategic effect between own strategy and opponent's strategy, submodularity produces competing effects when considering optimal responses. Using adaptive dynamics, conditions are given under

which such games can be guaranteed to exhibit Bayesian Nash equilibria, and it is shown that in many applications these equilibria will be a profile of monotone strategies. Comparative statics of parametrized games is also analyzed using results from submodular games which are extended to incorporate incomplete information. Several examples are provided.

The framework of Chapter 1 is applied to global games in **Chapter 2**. Global games methods are aimed at resolving issues of multiplicity of equilibria and coordination failure that arise in game theoretic models by relaxing common knowledge assumptions about an underlying parameter. These methods have recently received a lot of attention when the underlying complete information game is a GSC. Little has been done in this direction concerning GSS, however. This chapter complements the existing literature in both cases by extending the global games method developed by Carlsson and Van Damme (1993) to multiple player, multiple action GSS and GSC, using a p-dominance condition as the selection criterion. This approach helps circumvent recent criticisms to global games by relaxing some possibly unnatural assumptions on payoffs and parameters necessary to conduct analysis under current methods. The second part of this chapter generalizes the model by allowing groups of players to receive homogenous signals, which, under certain conditions, strengthens the model's power of predictability.

Chapter 3 analyzes the learning and stability of mixed strategy Nash equilibria in GSS, complementing recent work done in the case of GSC. Mixed strategies in GSS are of particular interest because it is well known that such games need not exhibit pure strategy Nash equilibria. First, a bound on the strategy space which indicate where randomizing behavior may occur in equilibrium is established. Second, it is shows that mixed strategy Nash equilibria are generally unstable under a wide variety of learning rules.

Acknowledgements

I would first like to thank those who made this work possible. First and foremost, my adviser and committee chairperson, Professor Tarun Sabarwal, who has taught me what it means to be a professional. Also, Professors Bernard Cornet, Joe Sicilian, Jianbo Zhang, and David Nualart for agreeing to be a part of my committee as well as providing helpful comments and advice during this long process. Lastly, I would like to thank the University of Kansas and the Department of Economics for giving me the opportunity to pursue my goals.

I would like to thank my family, to whom I dedicate Chapter 1: To my parents John and Bonnie Hoffmann, who, by example, have taught me the value of hard work. To my brother, Mark Hoffmann, who has always been there to support me, even while overcoming more obstacles than I ever will in my lifetime. And to Matthew Romans, my climbing partner.

I would like to thank my daughter, Melissa Hoffmann, to whom I dedicate Chapter 2: You have taught me more about life than I will ever be able to teach you. I would not be who I am today without you.

I would like to thank Anne Barthel, to whom I dedicate Chapter 3: Thank you for putting up with all of my craziness, including all of the "bouncy things", "erfs", and "k perp perps". You are an amazing person.

Contents

Chapter 1

Bayesian Games of Strategic Substitutes

1.1 Introduction

In a game of strategic substitutes (GSS), a higher action by one player induces the other players to best respond by taking a lower action. Examples include the simple Cournot oligopoly, common-pool resource games, provisions of public goods, and games of tournaments. With incomplete information, Van Zandt and Vives (2007) show that in games of strategic complements (GSC), where a higher action by one player induces the other players to best respond by taking a higher action, when beliefs are ordered increasingly by first order stochastic dominance (FOSD), extremal monotone Bayesian Nash equilibria (BNE) can be guaranteed to exist, as well as a monotone increase in the equilibria given an upward shock in beliefs in FOSD for the beliefs of each player-type. This chapter studies the properties of GSS under incomplete information, where the beliefs of each type for each player over the types of the others is ordered either increasingly or decreasingly in FOSD. Using generalized adaptive dynamics (GAD), necessary and sufficient conditions for the existence of equilibria can be derived in both cases, and in many cases their monotonicity properties as well. GAD also provides a notion of the stability of such equilibria. The first section of the chapter states the assumptions, and provides motivating examples. The second part addresses the solution concepts offered by GAD, and studies separately the two different assumptions on beliefs.

In the last section, comparative statics and the stability of equilibria is addressed.

1.2   Model and Assumptions

This chapter uses the standard lattice concepts. A game of strategic substitutes (GSS) is given by $\Gamma = (I,(A_i)_{i \in I}, (u_i)_{i \in I})$, whose elements are describe by:

- N players, $I = \{1, 2, ...N\}$

- Each A_i is player i's action set, which we assume is a compact metric lattice (a compact metric space which is a complete lattice, and the lattice operations inf and sup are continuous with respect to underlying metric), implying that all order increasing and decreasing sequences converge in metric to their respective sup and inf.[1] We will sometimes make the following assumption about the order $\geq_{A_i}$ on A_i, which is satisfied, for example, when A_i is a cube in $\mathbb{R}^n$:

  A1. The partial order $\geq_{A_i}$ is closed. That is, $\{(a, b) \in A_i^2 | b \geq_{A_i} a\}$ is closed in the product topology.

- Each player $i \in I$ has utility function $u_i : A \times T_i \to \mathbb{R}$. We assume that u_i is supermodular in a_i, has decreasing differences in (a_i, a_{-i}), increasing differences in (a_i, t_i), is continuous in a for all t_i, and is bounded.

The beliefs of the players and the measurability assumptions are described as follows:

- Each player $i \in I$ has a type space T_i and an associated sigma algebra $\mathscr{F}_i$, and we denote the product measure space by $(T, \mathscr{F})$.

---

[1] See Reny (2011)

- Each player $i \in I$ has a system of beliefs $P_i = (p_i(\cdot|t_i))_{t_i \in T_i}$, where each $p_i(\cdot|t_i) : \mathscr{F}_{-i} \to [0, 1]$ is a probability measure over the types of the other players. We consider two different orderings on P_i, summarized by the two following assumptions:

  A2. P_i is ordered increasingly by the first order stochastic dominance order $\geq_{FOSD}$. That is,
  $$t_i' \geq t_i \Rightarrow p_i(\cdot|t_i') \geq_{FOSD} p_i(\cdot|t_i)$$

  A3. P_i is ordered decreasingly by the first order stochastic dominance order $\geq_{FOSD}$. That is,
  $$t_i' \geq t_i \Rightarrow p_i(\cdot|t_i) \geq_{FOSD} p_i(\cdot|t_i')$$

- $\forall i \in I$, $\forall F_{-i} \in \mathscr{F}_{-i}$, $\varphi_{F_{-i}} : T_i \to [0, 1]$, $t_i \to p_i(F_{-i}|t_i)$, is measurable (intuitively, given an event $F_{-i} \in \mathscr{F}_{-i}$, player i can determine which of his types t_i assign a specific probability $a \in [0, 1]$)

- $\forall i \in I$, $u_i : A \times T_i \to \mathbb{R}$ is measurable in T_i, for all $a \in A$.

With these primitives, a Bayesian game of strategic substitutes (BGSS) is a tuple $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$, which has the following elements:

- For each player $i \in I$, S_i is the space of all measurable strategy functions $s_i : T_i \to A_i$.

- For each player $i \in I$, and profile of strategy functions s_{-i}, interim expected utility is given by $V_i(a_i, s_{-i}, t_i) = \int_{T_{-i}} u_i(a_i, s_{-i}(t_{-i}), t_i) dp_i(t_{-i}|t_i)$

- For each player $i \in I$, and strategy profile s_{-i}, we denote

$$BR_{t_i}(s_{-i}) = \underset{a_i \in A_i}{argmax}(V_i(\cdot, s_{-i}, t_i))$$

as type t_i's best response correspondence. Player $i's$ best response correspondence is

$$BR_i(s_{-i}) = \{s_i : T_i \to A_i \,|\, \forall t_i \in T_i, s_i(t_i) \in BR_{t_i}(s_{-i})\}$$

where each $s_i : T_i \to A_i$ is taken to be a measurable function.

Definition 1. Let $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ be a BGSS. A profile $\widehat{s} \in \mathscr{S}$ is a Bayesian Nash Equilibrium (BNE) if $\forall i \in I$, $\forall t_i \in T_i$,

$$\hat{s}_i(t_i) \in BR_{t_i}(\hat{s}_{-i})$$

Equivalently, $\widehat{s} \in \mathscr{S}$ is a BNE if $\forall i \in I$,

$$\hat{s}_i \in BR_i(\hat{s}_{-i})$$

Lemma 1. *(Monotonicity Theorem). Let X be a lattice, T a partially ordered set, and $f : X \times T \to R$ be such that f satisfies supermodularity in x. If f satisfies decreasing differences in (x, t), then $\varphi(t) = \underset{x \in X}{argmax}(f(x;t))$ is non-increasing in t. In particular, $\vee \varphi(t)$ and $\wedge \varphi(t)$ are non-increasing functions of t.*[2]

Proof. The proof can be found in Roy and Sabarwal(2010). □

The defining feature of an interim Bayesian Game, like the one we are considering, as opposed to an ex-ante Bayesian Game, is that the types of each player are counted among the players in the game, instead of just the *I* original players. This implies that if $\mathscr{S}_i$ is the space of strategy functions for player *i*, then the "everywhere order" ($s_i' \geq_e s_i$ iff $\forall t_i \in T_i$, $s_i'(t_i) \geq_{A_i} s_i(t_i)$) instead of the "almost everywhere order" ($s_i' \geq_{a.e} s_i$ iff $\forall t_i \in T_i$, $s_i'(t_i) \geq_{A_i} s_i(t_i)$ almost everywhere) is the correct partial order to consider on $\mathscr{S}_i$, as it takes the decisions of each player-type into account. However, it can be shown that switching from the $\geq_{a.e.}$ order to the $\geq_e$ order makes S_i no longer a complete lattice. This means that given arbitrary $C \subseteq S_i$, although $\vee C$ and $\wedge C$ are certainly functions, it cannot be guaranteed that they are measurable functions. Despite this, under our measurability assumptions, Van Zandt (2010) shows that, for any given s_{-i}, we can guarantee that $\vee BR_i(s_{-i})$ and $\wedge BR_i(s_{-i})$ are indeed measurable functions. From now on, without mention we will use $\geq$ to mean $\geq_e$. Also, when vectors of strategy functions are considered, we will also use $\geq$ to denote the product order.

---

[2]We say that $\varphi(t')$ is higher than $\varphi(t)$, denoted $\varphi(t') \geq_S \varphi(t)$, if and only if $\forall x \in \varphi(t'), \forall y \in \varphi(t)$, $x \vee y \in \varphi(t')$ and $x \wedge y \in \varphi(t)$

A game of strategic substitutes is defined as a game in which each player's best response is non-increasing in the strategies of the opponents. The following Proposition establishes that under our assumptions, the corresponding Bayesian game will in fact be a game of strategic substitutes.

Proposition 1. *Suppose $\forall i \in I$, u_i is supermodular in a_i and has decreasing differences in (a_i, a_{-i}).*

Then

1. $\forall t_i$, $BR_{t_i}(s_{-i})$ is non-increasing in s_{-i}, in the sense of Lemma 1.

2. BR_i is non-increasing in s_{-i} in the product order.[3] In particular,

3. $\vee BR_i$ and $\wedge BR_i$ are non-increasing functions of s_{-i}.

Proof. By Lemma 1, with $X = A_i$ and $T = \mathscr{S}_{-i}$, and considering

$$BR_{t_i}(s_{-i}) = \underset{a_i \in A_i}{argmax}(\int_{T_{-i}} u_i(a_i, s_{-i}(t_{-i}), t_i) dp_i(t_{-i}|t_i))$$

then since supermodularity and decreasing differences carry through integration, $s'_{-i} \geq s_{-i}$ implies that $BR_{t_i}(s_{-i}) \geq_S BR_{t_i}(s'_{-i})$, giving the first claim. The second claim follows immediately from the definition of product order. To establish the last claim, note that part 1 implies that $\vee BR_{t_i}$ and $\wedge BR_{t_i}$ are non-increasing functions of s_{-i}. Thus the result follows by noticing that

$$\vee BR_i(s_{-i}) = \{s_i : T_i \to A_i | \forall t_i \in T_i, s_i(t_i) \in \vee BR_{t_i}(s_{-i})\}$$

and likewise for $\wedge BR_i(s_{-i})$. □

---

[3]For $s'_{-i} \geq s_{-i}$, then $BR_i(s'_{-i})$ is higher than $BR_i(s_{-i})$ if $\forall t_i$, $BR_i(s'_{-i}) \geq_S BR_{t_i}(s_{-i})$.

1.2.1 Examples

Example 1. 3 Player Cournot. Let $I = \{1, 2, 3\}$, $T_i = \{t_L^i, t_M^i, t_H^i\}$, and let the profit function for player i be given by

$$\pi_i(a, t^i) = (a - b(q_1 + q_2 + q_3))q_i - (c - t^i)q^i$$

It is easily seen that the payoffs satisfy our submodularity assumptions.[4]Suppose that all players have the same beliefs about the others, where player $i's$ beliefs about the types of any other player j are given in the table below. It is easily seen that the first order stochastic dominance ordering on each player $i's$ beliefs is satisfied.

		Player j, $j \neq i$		
		$t_L^j = 1$	$t_M^j = 2$	$t_H^j = 3$
	$t_L^i = 1$	$p_i(t_L^j|t_L^i) = 3/4$	$p_i(t_L^j|t_M^i) = 0$	$p_i(t_L^j|t_H^i) = 1/4$
Player i	$t_M^i = 2$	$p_i(t_M^j|t_L^i) = 1/6$	$p_i(t_M^j|t_M^i) = 1/3$	$p_i(t_M^j|t_H^i) = 1/2$
	$t_H^i = 3$	$p_i(t_H^j|t_L^i) = 1/8$	$p_i(t_H^j|t_M^i) = 1/4$	$p_i(t_H^j|t_H^i) = 5/8$

The best responses of each player is given by

$$BR_{t_i}(s_{-i}) = \left(\frac{a - c + t_i}{2b}\right) - \frac{1}{2}\sum_{j \neq i} E(s_j|t_i)$$

If $a = 20$, $b = 3$, $c = 10$, we have a unique symmetric equilibrium, given by

$$\bar{s}_i(t_i) = \begin{cases} .794, & t^i = t_L^i \\ .986, & t^i = t_M^i \\ 1.2 & t^i = t_H^i \end{cases}, \quad i = 1, 2, 3$$

---

[4]For $f : X \times Y \to \mathbb{R}$ differentiable, f satisfies decreasing differences in (x, y) iff $\frac{\partial f}{\partial x \partial y} \leq 0$, and because $\mathbb{R}$ is totally ordered, supermodularity is automatically satisfied.

It is important to note that in two player BGSS, the game can be analyzed as a Bayesian game of strategic complements simply by reversing the order on one of the player's action sets. In this setting, extremal equilibria in monotone strategies are guaranteed to exist.[5] It is therefore interesting to note that with these specifications, we also have an equilibrium in monotone non-decreasing strategies.

Example 2. Game of Tournaments. Let $I = \{1, 2, 3\}, T_i = \{t_L^i, t_M^i, t_H^i\}$, as above. Suppose that the players are competing for a reward, with market value $r > 0$, where a single winner wins the reward with probability 1, two winners will get the reward with probability $\frac{1}{2}$, and three winners will get the reward with probability $\frac{1}{3}$. Each player $i = 1, 2, 3$ has a private valuation of the reward, t_i, which is a percentage of the market value r. Each player chooses a level of effort $x_i \in [0, 1]$, which corresponds to the probability of success of being a winner, and incurs a cost $\frac{cx_i^2}{2}$ for the level of effort. The expected reward per unit for player i is given by

$$\pi_i(x_i, x_j, x_k) = x_i(1 - x_j)(1 - x_k) + \frac{1}{2}x_ix_j(1 - x_k) + \frac{1}{2}x_ix_k(1 - x_j) + \frac{1}{3}x_ix_jx_k$$

and the payoff to player i is given by $u_i(x_i, x_j, x_k, t_i) = t_i\pi_i(x_i, x_j, x_k) - \frac{cx_i^2}{2}$. We see that $\frac{\partial u_i}{\partial x_i \partial x_j} = t_i(\frac{1}{3}x_k - \frac{1}{2}) \leq 0$, and $\frac{\partial u_i}{\partial x_i \partial t_i} = (1 - x_j)(1 - x_k) + \frac{1}{2}x_j(1 - x_k) + \frac{1}{2}x_k(1 - x_j) + \frac{1}{3}x_jx_k \geq 0$, so the game satisfies the GSS conditions. Assume that each player is certain that all other player's share her same relative valuation of the reward. This belief system, which satisfies **A2**, is written below:

<div align="center">

Player j, $j \neq i$

</div>

		t_j^L	t_j^M	t_j^H
	t_i^L	$p_i(t_j^L\|t_i^L) = 1$	$p_i(t_j^M\|t_i^L) = 0$	$p_i(t_j^H\|t_i^L) = 0$
Player i	t_i^M	$p_i(t_j^L\|t_i^M) = 0$	$p_i(t_j^M\|t_i^M) = 1$	$p_i(t_j^H\|t_i^M) = 0$
	t_i^H	$p_i(t_j^L\|t_i^H) = 0$	$p_i(t_j^M\|t_i^H) = 0$	$p_i(t_j^H\|t_i^H) = 1$

---

[5]See Van Zandt and Vives (2007).

The best response for player-type t_m^i is then

$$BR_{t_m}(s_j, s_k) = \frac{t_m}{c}(1 - \frac{1}{2}E(s_j|t_m) - \frac{1}{2}E(s_k|t_m)) = \frac{t_m^i}{c}(1 - \frac{1}{2}x_m^j - \frac{1}{2}x_m^k)$$

We then have a symmetric equilibrium $\hat{s}$, which is given by

$$\hat{s}_i(t_m^i) = \frac{t_i^m}{t_i^m + c}, \forall i = 1, 2, 3, \forall m = L, M, H.$$

Example 3. Common Pool Resource Game. Suppose $I = \{1,2,3\}$, and $T_i = \{w_i^L, w_i^H\}$. Each player can either invest in a common resource, or an outside option, both which have diminishing marginal return. Each player invests $x_i \leq b_i$ into the common resource, where $w_i > 0$ is player $i's$ endowment. Player i receives a proportional share of the total return on investment

$$\frac{x_i}{x_i + x_j + x_k}(a(x_i + x_j + x_k) - b(x_i + x_j + x_k)^2)$$

and receives $r(w_i - x_i) - s(w_i - x_i)^2$ from the outside investment. Therefore, the utility to player i is

$$u_i(x_i, x_j, x_k) = r(w_i - x_i) - s(w_i - x_i)^2 + \frac{x_i}{x_i + x_j + x_k}(a(x_i + x_j + x_k) - b(x_i + x_j + x_k)^2)$$

if $x_i + x_j + x_k > 0$, and $rb_i - sb_i^2$ otherwise. Assume that each player is uncertain about whether her opponents have a low endowment, w^L, or a high endowment, w^H, and that beliefs are given by

Player j, $j \neq i$

		w_j^L	w_j^H
	w_i^L	$p_i(w_j^L|w_i^L) = 0$	$p_i(w_j^H|w_i^L) = 1$
Player i	w_i^H	$p_i(w_j^L|w_i^H) = 1$	$p_i(w_j^H|w_i^H) = 0$

8

It is easily seen that the game satisfies the conditions of BGSS, where beliefs are consistent with assumption **A3**, and that each player-type's best response is given by

$$BR_{w_i}(s_j, s_k) = \frac{a - r + 2sb_i}{2b + 2s} - \frac{b}{2b + 2s}(E(s_j|w_i) + E(s_k|w_i))$$

Then the symmetric equilibrium is given by

$$\bar{s}_i(w_i) = \begin{cases} \frac{(2b+2s)(a-r+2sw_i^L)-2b(a-r+2sw_i^H)}{2b^2+8bs+4s^2} & , w_i = w_i^L \\ \frac{(2b+2s)(a-r+2sw_i^H)-2b(a-r+2sw_i^L)}{2b^2+8bs+4s^2} & , w_i = w_i^H \end{cases}, i = 1, 2, 3$$

which are monotone non-decreasing.

1.3 Characterizing Solution Concepts

We now define what it means for a sequence of strategy functions $(s_n)_{n=0}^{\infty} \subseteq 2^{\mathscr{S}}$ to satisfy adaptive dynamics. Unlike Bayesian games with strategic complements, where extremal equilibria are guaranteed to exist in monotone strategies, this is not the case in BGSS. Adaptive dynamics provides us with a way of exploiting the complementarity properties of the game in order to characterize various solution concepts. Intuitively, adaptive dynamics dictates that each player-type eventually behaves in such a way that her chosen actions fall above the lowest best response to the lowest strategy played, and below the highest best response to the highest strategy played. Recall that, under our measurability conditions, the sup and inf of the best response set does in fact lie in the strategy set under this order, a fact which is needed in defining adaptive dynamics. Since a BGSS can be analyzed as a GSS, in some of the proofs of this section, the reader is referred to Roy and Sabarwal (2012), with the exception of cases when new measure theoretic arguments are needed for this setting.

Definition 2. Let $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ be a BGSS. $\forall i \in I, \forall t_i \in T_i, \forall S_{-i} \subseteq \mathscr{S}_{-i}$,

- the set of undominated responses for player-type t_i is given by

$$UR_{t_i}(S_{-i}) = \{a_i \in A_i | \forall a'_i \in A_i, \exists s_{-i} \in \mathscr{S}_{-i}, V_i(a_i, s_{-i}, t_i) \geq V_i(a'_i, s_{-i}, t_i)\}$$

- the set of undominated responses for player i is given by

$$UR_i(S_{-i}) = \{s_i \in \mathscr{S}_i | \forall t_i \in T_i, s_i(t_i) \in UR_{t_i}(S)\}$$

We define the set of undominated responses to $S \subseteq \mathscr{S}$ as $UR(S) = \prod_{i \in I} UR_i(S_{-i})$.

If $\underline{s}, \bar{s} \in \mathscr{S}$ are such that $\bar{s} \geq \underline{s}$, define $\overline{UR}([\underline{s}, \bar{s}]) = [\wedge UR([\underline{s}, \bar{s}]), \vee UR([\underline{s}, \bar{s}])]$. We have the following lemma:

Lemma 2. *Let* $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ *be a BGSS. Then* $\forall s, s \in \mathscr{S}$,

$$\overline{UR}([\underline{s}, \bar{s}]) = [\wedge BR(\bar{s}), \vee BR(\underline{s})]$$

Proof. Let $s' \geq s$. $\forall i \in I, \forall t_i \in T_i, \vee BR_i(\underline{s}_{-i})(t_i) \in BR_{t_i}(\underline{s}_{-i})^6$. If $a_i \in BR_{t_i}(\underline{s}_{-i})$, then $\forall a'_i \in A_i$, $V_i(a_i, \underline{s}_{-i}, t_i) \geq V_i(a'_i, \underline{s}_{-i}, t_i)$. Thus, $\forall a'_i \in A_i, \exists s_{-i} \in [\underline{s}_{-i}, \bar{s}_{-i}]$ (namely $\underline{s}_{-i}$) such that $V_i(a_i, s_{-i}, t_i) \geq V_i(a'_i, s_{-i}, t_i)$. Hence $a_i \in UR_{t_i}([\underline{s}_{-i}, \bar{s}_{-i}])$. Therefore, $\vee BR_i(\underline{s}_{-i})(t_i) \in UR_{t_i}([\underline{s}_{-i}, \bar{s}_{-i}])$, hence $\vee BR_i(\underline{s}_{-i}) \in UR_{t_i}([\underline{s}_{-i}, \bar{s}_{-i}])$, giving $\vee BR_i(\underline{s}_{-i}) \preceq \vee UR_i([\underline{s}_{-i}, \bar{s}_{-i}])$, or $\vee BR(\underline{s}) \preceq \vee UR([\underline{s}, \bar{s}])$. Likewise, $\wedge UR([\underline{s}, \bar{s}]) \preceq \wedge BR(\bar{s})$, giving $[\wedge BR(\bar{s}), \vee BR(\underline{s})] \subseteq \overline{UR}([\underline{s}, \bar{s}])$.

---

[6]Our measurability and continuity assumptions ensure that $BR_{t_i}(\underline{s}_{-i})$ is a complete lattice, see Milgrom and Shannon (1994) Theorem A4.

Alternatively, let $s \notin [\wedge BR(\bar{s}), \vee BR(\underline{s})]$ be such that $\vee BR(\underline{s}) \not\leq s$. Then for some $i \in I$, and $t_i \in T_i$, $\vee BR_i(\underline{s})(t_i) \not\leq s_i(t_i)$. We see that $\hat{a}_i \equiv (\vee BR_i(\underline{s})(t_i)) \wedge (s_i(t_i))$ strictly dominates $s_i(t_i)$. Let $s_{-i} \in [\underline{s}_{-i}, \bar{s}_{-i}]$. Then

$$V_i(s_i(t_i), s_{-i}, t_i) - V_i(\hat{a}_i, s_{-i}, t_i) \leq V_i(s_i(t_i), \underline{s}_{-i}, t_i) - V_i(\hat{a}_i, \underline{s}_{-i}, t_i) \leq$$

$$V_i((s_i(t_i)) \vee (\vee BR_i(\underline{s})(t_i)), \underline{s}_{-i}, t_i) - V_i(\vee BR_i(\underline{s})(t_i), s_{-i}, t_i) < 0$$

where the first inequality follows from decreasing differences of V_i in (a_i, s_{-i}), the second from supermodularity of V_i in a_i, and the third from the definition of $\vee BR_i(\underline{s}_{-i})$. Hence $[\wedge BR(\bar{s}), \vee BR(\underline{s})] \supseteq \overline{UR}([\underline{s}, \bar{s}])$, giving set equality. $\qquad \square$

Definition 3. Let $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ be a BGSS, and $(s_n)_{n=0}^{\infty} \subseteq 2^{\mathscr{S}}$ a sequence of strategy functions. Then $\forall i \in I$, and each n, m such that $\forall n > m$, let $P(m, n) = \{s_j | m \leq j \leq n - 1\}$. Then we define

$$\vee P_i(m, n) = \max_{m \leq j \leq n-1} (s_{i_j})$$

and

$$\vee P(m, n) = \prod_{i \in I} \vee P_i(m, n)$$

Similar definitions are made for $\wedge P(m, n)$.

Definition 4. Let $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ be a BGSS. Then $(s^k)_{k=0}^{\infty} \subseteq 2^{\mathscr{S}}$ is an adaptive dynamic if

$$\forall i \in I, \forall K' > 0, \exists K_i > 0, \forall t_i \in T_i, k \geq K_i \Rightarrow \overline{UR}_{t_i}([\wedge P(K', k), \vee P(K', k)])$$

Thus, adaptive dynamics is satisfied if eventually players choose strategies that fall within these upper and lower bounds of previous play, regardless of the length of history they take into consideration. Note the uniform constant K_i that each player has on each player-type. This reflects the idea that although the player-types t_i represent the beliefs of player i in different environments, it is the player herself who eventually learns to play the game adaptively, and hence in each possible

11

environment. Also note that for a given $K' > 0$, by setting $K = \underset{i}{max}(K_i)$, the above definition of an adaptive dynamic coincides with the one given in RS

The simplest case of an adaptive dynamic is the best response dynamic, which will shown to be a bound on all other adaptive dynamics. This process is now defined.

Definition 5. The best response dynamic starting from $\wedge A$ and $\vee A$ are the sequences of functions $(y^k)_{k=0}^{\infty}$ and $(z^k)_{k=0}^{\infty}$ defined as[7]

- $y^0 = \wedge A$ and $z^0 = \vee A$

- $y^k = \wedge BR(y^{k-1})$ if k is even, $y^k = \vee BR(y^{k-1})$ if k is odd, and
  $z^k = \vee BR(z^{k-1})$ if k is even, and $z^k = \wedge BR(z^{k-1})$ if k is odd

The lower mixture and upper mixtures of $(y^k)_{k=0}^{\infty}$ and $(z^k)_{k=0}^{\infty}$ are the sequences $(\underline{x}^k)_{k=0}^{\infty}$ and $(\bar{x}^k)_{k=0}^{\infty}$ defined as

- $\underline{x}_k = y^k$ if k is even, and $\underline{x}^k = z^k$ if k is odd

- $\bar{x}_k = z^k$ if k is even, and $\bar{x}^k = y^k$ if k is odd.

Lemma 3. *Let $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ be a BGSS, and $(\bar{x}^k)_{k=0}^{\infty}$, $(\underline{x}^k)_{k=0}^{\infty}$ the upper and lower mixtures of the best response dynamics, respectively. Then*

1. The sequence $(\underline{x}^k)_{k=0}^{\infty}$ is non-decreasing, and there exists a strategy function profile $\underline{s} \in \mathscr{S}$ such that $\underline{s}$ is the pointwise limit of $(\underline{x}^k)_{k=0}^{\infty}$.

2. The sequence $(\bar{x}^k)_{k=0}^{\infty}$ is non-increasing, and there exists a strategy function profile $\bar{s} \in \mathscr{S}$ such that $\bar{s}$ is the pointwise limit of $(\bar{x}^k)_{k=0}^{\infty}$.

3. For every k, $\bar{x}^k \succeq \underline{x}^k$.

Proof. Follows the same argument as in RS. Notice that because the pointwise limit of a sequence of measurable functions from a measurable space into a metric space is measurable, we have that $\underline{s}$ and $\bar{s}$ are profiles of measurable functions. $\qquad\square$

---

[7]Noting that$\vee A$ can be viewed as the profile of constant functions $\vee A_i : T_i \to A_i$, $t_i \mapsto \vee A_i$, and likewise for $\wedge A$

Proposition 2. *Let $(\bar{x}^k)_{k=0}^\infty$ and $(\underline{x}^k)_{k=0}^\infty$ be the upper and lower mixtures of the best response dynamics, and $\bar{s}$, $\underline{s}$ their respective limits, and $(s^k)_{k=0}^\infty$ another adaptive dynamic. Then,*

1. $\forall N \geq 0, \exists K_N \geq 0$ such that $k \geq K_N \Rightarrow s^k \in [\underline{x}^N, \bar{x}^N]$.

2. $\underline{s} \preceq liminf(s^k) \preceq limsup(s^k) \preceq \bar{s}$.

3. $\bar{s}$ and $\underline{s}$ are the highest and lowest serially undominated strategies, respectively.

Proof. Follows directly from RS, using only lattice and order properties, and requires no special treatment under our setting. □

We are now ready to investigate the different solution concepts in a BGSS. We build up to one of the main theorems, which states that the converge of best response dynamic $(y^k)_{k=0}^\infty$ or $(z^k)_{k=0}^\infty$ is equivalent to dominance solvability, and the concept of global stability. Recall that although the upper and lower mixtures of the best response dynamics are monotonic sequences, and thus guaranteed to converge, this need not be the case for the best response dynamics themselves. We therefore also provide a necessary and sufficient condition for the convergence of best response dynamics, which relates to the fixed point of the double best response correspondence $BR \circ BR$.

Theorem 1. *Let $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ be a BGSS. Then the following are equivalent:*

1. Best response dynamics starting at $\wedge A$ or $\vee A$ converge.

2. Every adaptive dynamic converges to the same Bayesian Nash equilibrium.

Proof. We see that (1) implies (2). By our previous lemma, we know that $(\underline{x}^k)_{k=0}^\infty$ and $(\bar{x}^k)_{k=0}^\infty$ converge to $\underline{s}$ and $\bar{s}$, respectively. Suppose best response dynamic starting at $\wedge A$, $(y^k)_{k=0}^\infty$, converges, the case when $(z^k)_{k=0}^\infty$ being similar. Since $(y^{2k})_{k=0}^\infty$ and $(y^{2k+1})_{k=0}^\infty$ are subsequences of the convergent sequences $(\underline{x}^k)_{k=0}^\infty$ and $(\bar{x}^k)_{k=0}^\infty$, respectively, we have that $y^{2k} \to \underline{s}$, and $y^{2k+1} \to \bar{s}$. Since $(y^k)_{k=0}^\infty$ converges, we must have $\underline{s} = \bar{s}$.

13

Suppose $\underline{s} \notin BR(\bar{s})$. Then $\exists i \in I$, $\exists t_i \in T_i$, $\exists a_i \in A_i$ such that

$$V_i(a_i, \bar{x}_{-i}, t_i) = \int_{T_{-i}} u_i(a_i, \bar{x}_{-i}(t_{-i}), t_i) dp_i(t_{-i}|t_i) >$$

$$\int_{T_{-i}} u_i(\underline{x}_i(t_i), \bar{x}_{-i}(t_{-i}), t_i) dp_i(t_{-i}|t_i) = V_i(\underline{x}_i(t_i), \bar{x}_{-i}, t_i)$$

By boundedness of u_i, we have that the family $(u_i^k(a_i, t_i))_{k=0}^{\infty}$ defined as $u_i^k(a_i, t_i) = u_i(a_i, \bar{x}_{-i}(t_i), t_i)$, is uniformly bounded. By the continuity of u_i in a_{-i}, and the pointwise convergence of $\bar{x}^k$ to $\bar{x}$, then for all $a_i \in A_i$, $V_i(a_i, \bar{x}_{-i}^k, t_i) \to V_i(a_i, \bar{x}_{-i}, t_i)$. Hence $\exists K$, $\forall k \geq K$, $V_i(a_i, \bar{x}_{-i}^k, t_i) > V_i(\underline{x}_i(t_i), \bar{x}_{-i}^k, t_i)$. By the continuity, and hence upper semi-continuity, of V_i in a_i, we have

$$V_i(\underline{x}_i(t_i), \bar{x}_{-i}^k, t_i) \geq \limsup_{n \to \infty}(V_i(\underline{x}_i^n(t_i), \bar{x}_{-i}^k, t_i)) \geq \lim_{n \to \infty}(V_i(\underline{x}_i^n(t_i), \bar{x}_{-i}^k, t_i))$$

hence for n large, we have $V_i(a_i, \bar{x}_{-i}^k, t_i) > V(\underline{x}_i^n(t_i), \bar{x}_{-i}^k, t_i), \forall k \geq K$. This implies that $V_i(a_i, \bar{x}_{-i}^k, t_i) > V(\underline{x}_i^{k+1}(t_i), \bar{x}_{-i}^k, t_i)$, contradiction the optimality of $\underline{x}_i^{k+1}(t_i)$. Thus $\bar{s} = \underline{s}$ is a Bayesian Nash equilibrium, and by our last proposition, we have that all adaptive dynamics converge to this equilibrium, giving (2).

The fact that (2) implies (1) follows simply from the fact that the best response dynamics are themselves adaptive dynamics. □

Definition 6. A BGSS is globally stable if and only if there exists a Bayesian Nash equilibrium $\tilde{s} \in \mathscr{S}$ such that every adaptive dynamic converges to $\tilde{s}$.

Theorem 2. Let $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ be a BGSS. Then the following are equivalent:

1. Best response dynamic $(y^k)_{k=0}^{\infty}$ or $(z^k)_{k=0}^{\infty}$ converge

2. $BR \circ BR$ has a unique fixed point

3. $\hat{\Gamma}$ is globally stable

4. $\hat{\Gamma}$ is dominance solvable

Proof. Having established Proposition 2 and Theorem 1, the proof follows exactly as that in RS.

$\square$

We now consider what these solution concepts offer us under the different assumptions on beliefs, **A2** and **A3**. Even with the single-crossing like conditions imposed on the utility functions, the problem remains largely intractable without imposing any structure on beliefs. Intuitively, **A2** stipulates that observing a higher type leads a player to put more weight on the possibility of the other players receiving a higher type, and **A3** stipulates that observing a higher type leads a player to put more weight on the possibility of the other players receiving a lower type. In many situations these are rather mild assumptions. Assumption **A3** is studied first.

The following result is variation of the result found in Van Zandt and Vives (2007), but in the case for BGSC, where t_i is suppressed in the utility function. We will assume for the following that for a player i of type t_i, BR_{t_i} is singleton valued, although the same analysis can be done for both $\vee BR_{t_i}$ and $\wedge BR_{t_i}$.

Proposition 3. *If $u_i : A \to \mathbb{R}$ satisfies supermodularity in a_i, continuity in a_i, and decreasing differences in (a_i, a_{-i}), then $\forall t_i \in T_i$, $V_i(\cdot, s_{-i}, t_i) = \int_{T_{-i}} u_i(\cdot, s_{-i}(t_{-i})) dp_i(t_{-i}|t_i)$ is supermodular and continuous in a_i, and satisfies decreasing differences in (a_i, s_{-i}). Furthermore, under assumption A3, if $s_{-i} \in \mathscr{S}_{-i}$ is a profile of monotone non-decreasing strategies, then $p_i(t_{-i}|t_i') \geq_{FOSD} p_i(t_{-i}|t_i)$, $BR_{t_i'}(s_{-i}) \geq BR_{t_i}(s_{-i}))$.*

Proof. Since supermodularity, continuity, and decreasing differences carry through integration, it follows that $V_i(\cdot, s_{-i}, t_i)$ satisfies supermodularity in a_i, and decreasing differences in (a_i, s_{-i}). Now, suppose u_i satisfies decreasing differences in (a_i, a_{-i}), and $s_{-i} \in \mathscr{S}_{-i}$ a profile of monotone non-decreasing strategies. The decreasing differences of u_i in (a_i, a_{-i}) implies that $u_i(a_i, s_{-i}(t_{-i}))$ satisfies decreasing differences in (a_i, t_{-i}). Let $a_i^H \succeq a_i^L$ and $p_i(t_{-i}|t_i) \geq_{FOSD} p_i(t_{-i}|t_i')$. Define $h(t_{-i}) = u_i(a_i^H, s_{-i}(t_{-i})) - u_i(a_i^L, s_i(t_{-i}))$, which is non-increasing in t_{-i}. By FOSD,

$$V(a_i^H, s_{-i}, t') - V(a_i^L, s_{-i}, t') = \int_{T_{-i}} h(t_{-i}) dp_i(t_{-i}|t_i') \geq$$

15

$$\int\limits_{T_{-i}} h(t_{-i}) dp_i(t_{-i}|t_i) = V(a_i^H, s_{-i}, t_i) - V(a_i^L, s_{-i}, t_i)$$

giving increasing differences in (a_i, t_i). By the monotonicity theorem, $BR_{t_i'}(s_{-i}) \geq_{A_i} BR_{t_i}(s_{-i})$, giving $BR_i(s_{-i}; P_i') \geq BR_i(s_{-i}; P_i)$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

It follows immediately from the above proposition that since best response dynamics begin at monotone non-decreasing strategies $\wedge A$ and $\vee A$, assumption **A3** implies that the set of Nash equilibria is bounded above and below by serially undominated strategies which are monotone non-decreasing. Furthermore, if best response dynamics converge, there is a unique equilibrium which is monotone non-decreasing. It is natural to ask about what can be said if best response dynamics cannot be guaranteed to converge. Under additional assumptions, a monotone non-decreasing Bayesian Nash equilibrium can still be guaranteed to exist, which is now shown. We will say that $(T_i, \mathscr{F}_i, \mu_i, \geq_{T_i})$ is an admissible probability space if the following conditions hold:

1. (T_i, d_i) is a separable metric space, $\mathscr{F}_i$ the corresponding Borel sigma algebra, and μ_i is any probability measure on T_i.

2. $\geq_{T_i}$ is a partial order on T_i such that the sets $\{(a, b) \in T_i^2 \,|\, b \geq_{T_i} a\}$ are measurable in the product sigma algebra.

3. Every atomless set $A \in \mathscr{F}_i$ such that $\mu_i(A) > 0$ contains two strictly ordered points.[8]

When $\mathbb{R}^n$ is ordered with the vector ordering, μ_i is the associated Lebesque measure, and T_i is any cube in $\mathbb{R}^n$ with positive finite measure, then these conditions are automatically satisfied,[9] as well as when T_i is any interval on the real line. Define $M_i = \{f : T_i \to A_i \,|\, f \text{ non-decreasing}\}$. It can be shown[10] that if $(T_i, \mathscr{F}_i, \mu_i, \geq_{T_i})$ is an admissible probability space, and A_i satisfies **A1**, each function $f \in M_i$ is equal $\mu_i - a.e.$ to a measurable, non-decreasing function. Furthermore, under a

---

[8]We say that two points x and x' are strictly ordered iff there exist disjoint open sets $U \ni x$ and $U' \ni x'$ such that for all $(y, z) \in U \times U'$, $z \geq y$.

[9]See the Steinhaus Theorem.

[10]see Reny (2011)

metric $\delta_i : M_i \times M_i \to \mathbb{R}_+$[11], (M_i, δ_i) is a compact metric space such that $\delta_i(f_n, f) \to 0$ if and only if $d_{A_i}(f_n(t_i), f(t_i)) \to 0$ for $\mu_i - a.e.$ We then have the following theorem:

Theorem 3. Let $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ be a BGSS. If $\forall i \in I$, $(T_i, \mathscr{F}_i, \mu_i, \geq_{T_i})$ is an admissible probability space, A_i is convex, satisfies **A1**, P_i satisfies **A3**, and each $p_i(\cdot|t_i) \in P_i$ is absolutely continuous with respect to the product measure $\mu_{-i} : \bigotimes_{j \neq i} \mathscr{F}_j \to [0,1]$. Then if each $BR_{t_i} : \mathscr{S}_{-i} \to A_i$ is singleton-valued and pointwise continuous,[12] there exists a pure strategy Bayesian Nash equilibrium which is a profile of monotone non-decreasing strategies, and the set of Bayesian Nash equilibria is bounded above and below by serially undominated strategies which are also profiles of monotone non-decreasing strategies.

Proof. As above, **A3** automatically implies that the extremal serially undominated strategies are monotone, and since $\forall i \in I$, $(T_i, \mathscr{F}_i, \mu_i, \geq_{T_i})$ is an admissible probability space, **A1** and convexity of A_i implies that every monotone function f_i in the compact, convex space (M_i, δ_i) is equal μ_i-a.e. to a measurable monotone function $\tilde{f}_i$. Let $f_{-i} \in M_{-i}$, and $\tilde{f}_{-i}$ be the measurable representative that f_{-i} is μ_{-i}-a.e. equal to. Since $\forall t_i \in T_i$, $p_i(\cdot|t_i)$ is absolutely continuous with respect to μ_{-i}, f_{-i} is $p_i(\cdot|t_i)$-a.e. equal to $\tilde{f}_{-i}$, and thus by Proposition 1.17, $BR_i : M_{-i} \to M_i$ is well defined. Now let $(f^n_{-i})_{n=1}^\infty \subseteq M_{-i}$ be such that $f^n_{-i} \to f_{-i}$ under the product metric δ_{-i}. Then by above, $f^n_{-i} \to f_{-i}$ μ_{-i}-a.e, and hence $p_i(\cdot|t_i)$-a.e., $\forall t_i$. By continuity of BR_{t_i}, we have $\lim_{n \to \infty}(BR_i(f^n_{-i})(t_i)) \equiv \lim_{n \to \infty}(BR_i(\tilde{f}^n_{-i})(t_i)) = \lim_{n \to \infty}(BR_{t_i}(\tilde{f}^n_{-i})) = BR_{t_i}(\tilde{f}_{-i}) \equiv BR_i(\tilde{f}_{-i})(t_i))$, giving continuity of BR_i. Hence $BR : M \to M$ is a singleton-valued continuous mapping from the compact, convex, metric space (M, δ) into itself, and therefore by Brouwer's fixed point theorem,[13] $Fix(BR) \cap M$ is non-empty and compact, giving the result. $\square$

We now consider a wide class of games when beliefs satisfy the assumption **A3**. In this case, there are two effects that determine the monotonicity of a best response for player i to a strategy s_{-i}.

[11]For $f, f' \in M$, define $\delta_i(f_n, f) = \int_{T_i} d_{A_i}(f_n(t_i), f(t_i)) d\mu_i$.

[12]$s^n_{-i} \to s_{-i}$ pointwise $p_i(\cdot|t_i)$ a.e. $\Rightarrow BR_{t_i}(s^n_{-i}) \to BR_{t_i}(s_{-i})$.

[13]If K is a non-empty, compact, convex subset of a locally convex Hausdorff space, and $f : K \to K$ is continuous, then $Fix(f)$ is non-empty and compact.

Suppose that $s_{-i} : T_{-i} \to A_{-i}$ is a profile of monotone non-decreasing strategies. The competing effects are the following:

- By assumption, u_i has increasing differences between (a_i, t_i). Hence when considering $V_i(a_i, s_{-i}, t_i) = \int u_i(a_i, s_{-i}(t_{-i}), t_i) dp_i(t_{-i}|t_i)$, as t_i increases, the direct effect of this increase through the utility function u_i induces player i to take a higher strategy.

- By assumption, u_i has decreasing differences in (a_i, a_{-i}), or, in consideration of the interim expected utility above, because s_{-i} is monotone increasing, u_i has decreasing differences in (a_i, t_{-i}). Therefore, as t_i rises to t'_i, through the first order stochastic dominance order, t_{-i} has a higher expected value as $p_i(\cdot|t_i)$ increases to $p_i(\cdot|t'_i)$, therefore the decreasing differences in (a_i, t_{-i}) induces player i to take a lower strategy.

In many applications, the direct and indirect effect appear either additively or multiplicatively in the best response. For example, suppose that player $i's$ best response function can be written as $BR_{t_i}(s_{-i}) = g(t_i) - \alpha(\sum_{j \neq i} E(s_{-i}|t_i))$, where $\alpha \geq 0$ and $g(\cdot)$ is an increasing function of t_i, as in our Cournot game. Then, as illustrated in the graphic below, in a two player game, if s_2 is a non-decreasing strategy, it is seen how $g(\cdot)$ captures the direct effect from the increasing differences in (a_1, t_1), and $\alpha(E(s_2|t_1))$ the indirect effect from the first order stochastic dominance assumption on beliefs through the higher expected value of s_2.
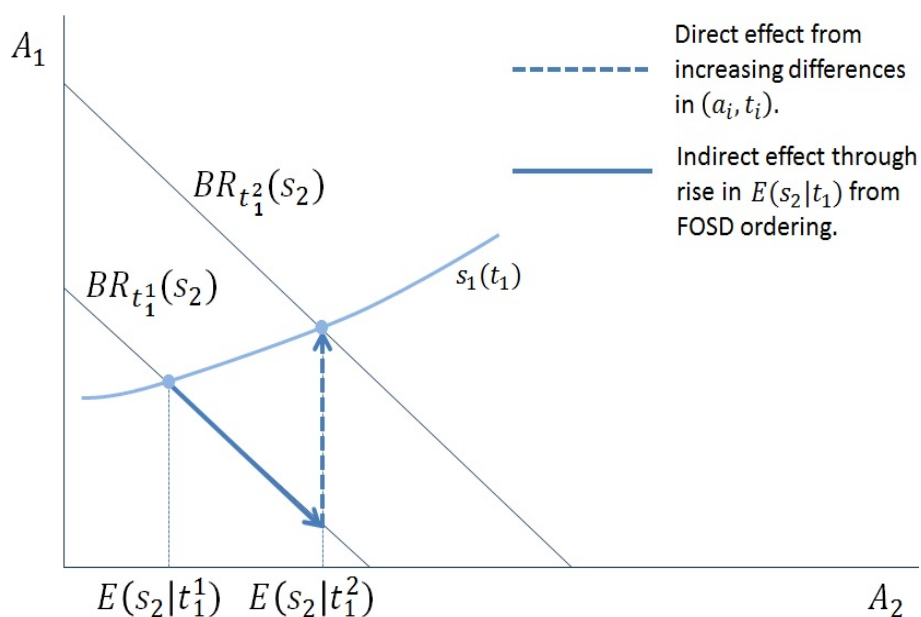
The next theorem gives us sufficient conditions to guarantee the monotonicity of extremal equilibria in many applications, where T_i and A_i are subsets of the real line, by capturing the direct and indirect effects of the best response dynamics. Suppose for the following Theorem that each $A_i \subseteq \mathbb{R}_+$.

Theorem 4. Let $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ be a BGSS. Suppose that *for all $i \in I$, and $t_i \in T_i$ best responses can be written as*

$$BR_{t_i}(s_{-i}) = g_i(t_i) - v_i(t_i)(\sum_{j \neq i} E(s_j | t_i))$$

where v_i is a non-negative function of t_i.[14] *Define $z_i : T_i \to A_i$, $t_i \mapsto BR_{t_i}(g_{-i}(\cdot))$. Then, if $\forall i \in I$,*

1. v_i and z_i are non-decreasing functions of t_i, then best response dynamics starting from a profile of non-decreasing strategies is a sequence whose every even term is a profile of monotone non-decreasing strategies. Likewise, if v_i and z_i are non-increasing functions of t_i, then best response dynamics starting from a profile of non-increasing strategies is a sequence whose every even term is a profile of monotone non-increasing strategies.

2. In addition, suppose $v(t_i) \equiv \alpha \geq 0$. If z_i is monotone non-decreasing, best response dynamics starting from a profile of non-decreasing strategies is a sequence whose every odd term is a profile of monotone non-decreasing strategies. Likewise, best response dynamics starting from a profile of non-increasing strategies is a sequence whose every odd term is a profile of monotone non-increasing strategies.

3. If (1) or (2) are true, then if best response dynamics starting from a profile of monotone non-decreasing (increasing) strategies converges, they converge to a profile of monotone non-decreasing (increasing) strategies.

Proof. The proof is done in the case of starting from a profile of monotone non-decreasing strategies, the other case being similar. Suppose the condition in (1) holds, and consider best response

[14]This ensures that best responses are non-increasing in opponents' strategies, the defining property of a BGSS.

19

dynamics starting from $s \equiv BR^0(s)$, a profile of monotone non-decreasing strategies. Suppose for $k \geq 2$ even, $BR^{k-2}(s)$ is a profile of monotone non-decreasing strategies. Then

$$BR_{t_i}(BR^{k-1}_{-i}(s)) = g_i(t_i) - v_i(t_i)((\sum_{j \neq i} E(BR^{k-1}_j(s))|t_i)) =$$

$$g_i(t_i) - v_i(t_i)(\sum_{j \neq i} E((g_j(t_j) - v_j(t_j)(\sum_{m \neq j} E(BR^{k-2}_m|t_j)|t_i)) =$$

$$g(t_i) - v(t_i)(\sum_{j \neq i} E(g_j(t_j)|t_i)) + v_i(t_i)(\sum_{j \neq i}\sum_{m \neq j} E(v_j(t_j)E(BR^{k-2}_m|t_j)|t_i)$$

Therefore,

$$BR_{t'_i}(BR^{k-1}_{-i}(s)) - BR_{t_i}(BR^{k-1}_{-i}(s)) =$$

$$(g_i(t'_i) - g_i(t_i)) - (v_i(t'_i)\sum_{j \neq i} E(g_j(t_j)|t'_i) - v_i(t_i)\sum_{j \neq i} E(g_j(t_j)|t_i))$$

$$+v_i(t'_i)(\sum_{j \neq i}\sum_{m \neq j} E(v_j(t_j)E(BR^{k-2}_m|t_j)|t'_i)) - v_i(t_i)(\sum_{j \neq i}\sum_{m \neq j} E(v_j(t_j)E(BR^{k-2}_m|t_j)|t_i))$$

Since each BR^{k-2}_m is non-decreasing, by FOSD we have that $E(BR^{k-2}_m|t_j)$ is a non-decreasing function of t_j. Since $v(\cdot)$ is non-decreasing and non-negative, then $v_j(t_j)E(BR^{k-2}_m|t_j)$ is a non-decreasing function of t_j. Again, by FOSD, $E(v(t_j)E(BR^{k-2}_m|t_j)|t_i)$ is a non-decreasing function of t_i, giving the non-negativity of the right-hand side. The left hand side is non-negative by hypothesis, hence $\forall k \geq 0$ even, $BR^k_i(s)$ is non-decreasing, or $BR^k(s)$ is a profile of non-decreasing strategies.

Now suppose that the condition in (2) holds, and consider best response dynamics starting from $s \equiv BR^0(s)$, a monotone non-increasing strategy. Then $\forall i \in I, \forall t_i \in T_i, BR_{t_i}(s_{-i}) = g(t_i) - \alpha \sum_{j \neq i} E(s_j|t_i)$, thus $BR_{t'_i}(s_{-i}) - BR_{t_i}(s_{-i}) = (g(t'_i) - g(t_i)) - \alpha(\sum_{j \neq i} E(s_j|t'_i) - \sum_{j \neq i} E(s_j|t_i))$. By FOSD, we have that since s is a profile of non-increasing strategies, the right hand side is non-negative, and clearly the left hand side is. Hence $BR^1(s)$ is a profile on monotone non-decreasing strategies. Suppose that for $k \geq 2$ odd, $BR^{k-1}(s)$ is a profile of monotone non-decreasing strategies. Then

$\forall i \in I, \forall t_i \in T_i,$

$$BR_{t_i}(BR^k_{-i}) = g(t_i) - \alpha \sum_{j \neq i} E(BR^k_j | t_i) =$$

$$g(t_i) - \alpha \sum_{j \neq i} E(g(t) - \alpha \sum_{m \neq j} E(BR^{k-1}_m | t_j) | t_i) =$$

$$g(t_i) - \alpha \sum_{j \neq i} E(g(t) | t_i) + \alpha^2 (\sum_{j \neq i} \sum_{m \neq j} E(E(BR^{k-2}_m | t_j) | t_i))$$

Therefore, for $t'_i > t_i$,

$$BR_{t'_i}(BR^k_{-i}) - BR_{t_i}(BR^k_{-i}) =$$

$$(g(t'_i) - g(t_i)) - \alpha (\sum_{j \neq i} E(g(t) | t'_i) - \sum_{j \neq i} E(g(t) | t_i)) =$$

$$+ \alpha^2 (\sum_{j \neq i} \sum_{m \neq j} E(E(BR^{k-2}_m | t_j) | t'_i) - \sum_{j \neq i} \sum_{m \neq j} E(E(BR^{k-2}_m | t_j) | t_i))$$

By a similar argument used in (1), from the fact that BR^{k-2}_m is monotone non-decreasing and by FOSD, the right hand side is non-negative, and the left hand side is by assumption. Therefore, $\forall k \geq 0$ odd, $BR^k(s)$ is a profile of monotone non-decreasing strategies.

For (3), assume that (1) holds, and that best response dynamics $(BR^k(s))^{\infty}_{k=0}$ starting from $s \in \mathscr{S}$ monotone non-decreasing converges to a function $\hat{s} \in \mathscr{S}$. Then $\forall i \in I, \forall t_i \in T_i, BR^k_i(t_i) \rightarrow \hat{s}_i(t_i)$. Because $(BR^{2k}_i(t_i))^{\infty}_{k=0}$ is a subsequence of the convergent sequence $(BR^k_i(t_i))^{\infty}_{k=0}$, we have that $\forall t_i \in T_i, \hat{s}_i(t_i) = \lim_{k \to \infty} (BR^{2k}_i(t_i))$. Since $\forall k \geq 0, t'_i > t_i \Rightarrow BR^{2k}_i(t'_i) \geq BR^{2k}_i(t_i)$, it follows that $\hat{s}_i(t'_i) \geq \hat{s}_i(t_i)$ and hence $\hat{s}$ is a profile of monotone non-decreasing strategies. A similar argument holds if (2) holds. $\qquad\square$

It is easily seen that under (2), z_i being monotone non-decreasing is also a necessary condition if the types for player i have homogenous beliefs. The condition can be interpreted as ensuring that even when a player best responds only to the direct effect of a strategy profile, the result will be a monotone non-decreasing strategy, implying that when the indirect effect is taken into account, the result will still be a monotone non-decreasing strategy. The next Corollary connects this result with the solution concepts offered by GAD.

Corollary 1. *If* (1) *or* (2) *hold, then the upper and lower best response mixtures starting from* $\vee A$ *and* $\wedge A$, $(\bar{x}^k)_{k=0}^{\infty}$ *and* $(\underline{x}^k)_{k=0}^{\infty}$, *are profiles whose every even/odd term are monotone non-decreasing strategies, respectively. If best response dynamics converge, the upper and lower mixtures converge to a Bayesian Nash equilibrium profile of monotone non-decreasing strategies.*

Proof. Follows immediately from the last theorem, the fact that $\vee A$ and $\wedge A$ are profiles of constant, and hence monotone non-decreasing/increasing strategies, and the definition of the upper and lower best response dynamic mixtures. $\qquad\square$

Corollary 2. *If* (2) *holds, then there exists a unique Bayesian Nash equilibrium which is a profile of monotone non-decreasing strategies if and only if* $\alpha \in [0, 1)$.

Proof. For each $t_i \in T_i$, define $E^0(g_{-i}|t_i) = 0$, $E^1(g_{-i}|t_i) = \sum_{j \neq i} E(g_j|t_i)$, $E^2(g_{-i}|t_i) = \sum_{j \neq i}\sum_{m \neq j} E(E(g_m|t^j)|t_i)$, and in general, $E^k(g_{-i}|t_i) = \sum_{j \neq i} E(E^{j-1}(g_{-j}|t^j)|t_i)$, where t_j is a specific realization of the variable t^j. Then

$$BR^1_{t_i}(\vee A) = g_i(t_i) + E^0(g_{-i}|t_i) - \alpha(N-1)\vee A$$

Suppose for $k > 0$,

$$BR^k_{t_i}(\vee A) = g_i(t_i) + \sum_{n=1}^{k-1}(-1)^n\alpha^n E^n(g_{-i}|t_i) + (-1)^k\alpha^k(N-1)\vee A$$

Then

$$BR^{k+1}_{t_i}(\vee A) = g_i(t_i) - \alpha\sum_{j \neq i} E(g_j(\cdot) + \sum_{n=1}^{k-1}(-1)^n\alpha^n E^n(g_{-j}|t^j) + (-1)^k\alpha^k(N-1)\vee A)|t_i) =$$

$$g_i(t_i) + \sum_{n=1}^{k}(-1)^n\alpha^n E^n(g_{-i}|t_i) + (-1)^{k+1}\alpha^{k+1}(N-1)\vee A$$

Likewise, for any $k \geq 0$,

$$BR^k_{t_i}(\wedge A) = g_i(t_i) + \sum_{n=1}^{k-1}(-1)^n\alpha^n E^n(g_{-i}|t_i) + (-1)^k\alpha^k(N-1)\wedge A$$

Thus, for any $k \geq 0$, $||BR_{t_i}^k(\vee A) - BR_{t_i}^k(\wedge A)|| = \alpha^k (N-1)(\vee A - \wedge A)$, and thus we see that as $k \to \infty$, best response dynamics starting from $\vee A$ and $\wedge A$ converge, giving the result. $\qquad \square$

Example 4. n-Player Cournot, continuum of types. Consider the n-player extension of the original Cournot game that was presented, but this time with a continuum of types given by $T_i = [0, c)$. Suppose that beliefs are given by the exponential distribution, with cdf $F(x, t_i) = 1 - \frac{1}{t_i} e^{\frac{-x}{t_i}}$, where $x > 0$. It is straightforward to check that $t_i' > t_i \Rightarrow F(\cdot, t_i') \geq_{FOSD} F(\cdot, t_i)$. The best response function for player-type t_i is given by

$$BR_{t_i}(s_{-i}) = \left(\frac{a - c + t_i}{2b}\right) - \frac{1}{2} \sum_{j \neq i} E(s_j | t_i)$$

where $g(t_i) = \left(\frac{a-c+t_i}{2b}\right)$. Since the mean of the exponential distribution is t_i, we have that $\forall i \in I$ and $t_i' > t_i$,

$$BR_{t_i'}(g_{-i}) - BR_{t_i}(g_{-i}) = \left(\frac{n+1}{4b}\right)(t_i' - t_i) > 0$$

and thus a unique monotone non-decreasing BNE can be guaranteed.

1.4 Monotone Comparative Statics

We now consider what happens to the set of Bayesian Nash equilibria if the beliefs of each player-type have an upward shift in first order stochastic dominance. Letting $P = ((p(t_{-i}|t_i))_{i \in I, t_i \in T_i}$, we say $P' \geq_{FOSD} P$ if $\forall i \in I$, $\forall t_i \in T_i$, $p_i'(t_{-i}|t_i) \geq_{FOSD} p_i(t_{-i}|t_i)$. Just as in the case of establishing the monotonicity of a best response, in BGSS there are two competing effects as to whether or not such an upward shift will produce a higher or lower equilibrium. Again, monotonicity of the initial profile plays a crucial role. The competing effects in this case can be summarized as follows:

Let $s \in \mathscr{S}$ be an initial profile of strategies in a BGSS. Then,

1. If s is monotone non-increasing, as $P' \geq_{FOSD} P$, by Proposition 1, each player will want to choose a higher response. Again from Proposition 1, because of the decreasing differences in a (a_i, s_{-i}) each player will subsequently want to take a lower response, producing competing

effects. This is in contrast to BGSC, where the increasing differences in (a_i, s_{-i}) would augment the initial higher response with another higher response (albeit in the case where s is non-decreasing).

2. If s is monotone non-decreasing, as $P' \geq_{FOSD} P$, each player will want to choose a lower response. The deceasing differences in (a_i, s_{-i}) subsequently leads players to choose a higher response, producing competing effects, and in contrast to BGSC where we would again get augmenting effects, but again in the case when s is monotone non-decreasing.

To determine which of the two effects dominates, we use Schauder's fixed point theorem. As is well known, when trying to apply a fixed point theorem in a function space (such as with $\mathscr{S}$), there is a natural tension between finding a topology small enough to support compact order intervals, and large enough to retain the continuity of the best response function, and hence any such argument must be taken on a case by case basis. Despite this, in the next theorem we assume both the compactness of order intervals $[\underline{s}, \overline{s}]$, and the continuity of the best response, and then discuss some simple situations where these are automatically satisfied.

Theorem 5. *Let* $\hat{\Gamma} = (I, (S_i)_{i \in I} (V_i)_{\in I})$ *be a BGSS, and* $BR : \mathscr{S} \to \mathscr{S}$ *be continuous. Let* $P' \geq P$, *and* $s \in NE(P)$. *Then,*

1. *If* s *is monotone non-increasing, and* $BR(BR(s, P')) \geq s$, *there exists* $\hat{s} \in NE(P')$ *such that* $\hat{s} \geq s$.

2. *If* s *is monotone non-decreasing and* $s \geq BR(BR(s, P'))$, *there exists* $\hat{s} \in NE(P')$ *such that* $s \geq \hat{s}$.

Proof. We prove 1, the proof for 2 being nearly identical. Let $s \in NE(P)$ be monotone non-increasing. By Proposition 3, each player-type t_i will choose a higher action given a higher belief, hence BR is non-decreasing in P, giving $BR(s, P') \geq BR(s, P) = s$. Now let $s' \in [s, BR(s, P')]$. Since $BR(s, P') \geq s', BR(s, P') \geq BR(BR(s, P')) \geq s$, where the first inequality follows from non-increasingness in s, and the second from the condition in the theorem. Also, since $s' \geq s$, $BR(s, P') \geq$

$BR(s', P')$. Hence $BR([s, BR(s, P')], P') \subseteq [s, BR(s, P')]$. By continuity and the fact that $[s, BR(s, P')]$ is compact, Schauder's fixed point theorem guarantees some $\hat{s} \in NE(P') \cap [s, BR(s, P')]$, giving the result. $\qquad \square$

Consider the conditions of Theorem 3. Recalling that (M, δ) was guaranteed to be compact, we see that if we replace $[s, BR(s, P')]$ with $[s, BR(s, P')] \cap M$ in the above proof, then $[s, BR(s, P')] \cap M$ is compact. Therefore, under these conditions, we can always do monotone comparative statics. Order intervals cannot be guaranteed to be compact, in general. However, if each T_i is countable, then by considering the topology of pointwise convergence τ_{p_i}, [15] if A_i is metrizable, it follows that τ_{p_i} is, [16] and is thus Hausdorff and also first countable, allowing us to characterize continuity through sequential convergence. $\mathscr{S}_i$ is clearly convex, and order intervals $[\underline{s}, \overline{s}]$ are compact.[17] Hence, if best responses are pointwise continuous, such as the form $BR_{t_i}(s_{-i}) = g(t_i) - v(t_i)(\sum_{j \neq i} E(s_{-i}|t_i))$ given in Theorem 4, the conditions of Theorem 5 can be applied.

To see why the monotonicity of the initial equilibrium is important, consider the following graphical representation of a symmetric game where the initial beliefs for t_i^1 and t_i^2 and the initial Bayesian Nash equilibrium are represented by solid lines, and the new beliefs and equilibrium are represented by dotted lines:

---

[15] Recall that each s_i can be viewed as an element of $\mathscr{A}_i \equiv \prod_{t_i \in T_i} A_i$, and τ_{p_i} is the restriction to $\mathscr{S}_i$ of the corresponding product topology $\tau_{\mathscr{A}_i}$ restricted to $\mathscr{S}_i \subseteq A_i^{T_i}$

[16] $\forall f, g \in \mathscr{S}_i, d_i(f, g) = \sum_{j=1}^{\infty} 2^{-j} \frac{d_j(f(t_j), g(t_j))}{1 + d_j(f(t_j), g(t_j))}$

[17] Because each $[\underline{s}(t_i), \overline{s}(t_i)]$ is compact, by the Tychonoff product theorem, $[\underline{s}, \overline{s}] = \prod_{t_i \in T_i} [\underline{s}(t_i), \overline{s}(t_i)]$ is also.
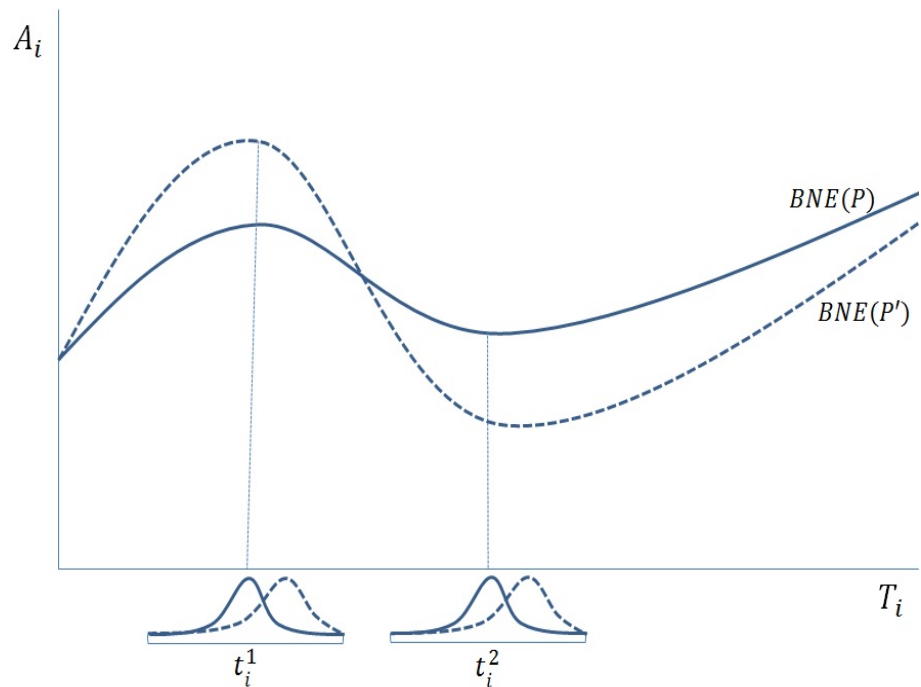
As the belief structure P changes to P', both types' beliefs increase in FOSD. Note, however, that when viewing the old equilibrium profile, type t_i^1 sees unambiguously lower strategies being played by her opponents, and therefore best responds by playing higher by GSS. Type t_i^2, on the other hand, sees unambiguously higher actions being played by her opponents, and best responds by playing lower by GSS. Therefore, the new equilibrium is neither higher nor lower than the old equilibrium, and monotone comparative statics cannot be guaranteed.

Chapter 2

Global Games Selection in Games with Strategic Substitutes or Complements

2.1 Introduction

The global games method serves as an equilibrium selection device for complete information games by embedding them into a class of Bayesian games that exhibit unique equilibrium predictions. This method was pioneered by Carlsson and Van Damme (CvD) (1993) for the case of 2-player, 2-action coordination games. In that paper, a complete information game with multiple equilibria is considered, and instead of players observing a specific parameter in the model directly, they observe noisy signals about the parameter, transforming it into a Bayesian game. As the signals become more precise, a unique serially undominated Bayesian prediction emerges, resolving the original issue of multiplicity by delivering a unique prediction in a slightly "noisy" version of the original complete information game. This method has since been extended by Frankel, Morris, and Pauzner (FMP) (2003) to multiple-player, multiple-action games of strategic complements (GSC), where a higher action from opponents induces a player to also best respond with a higher action. Although this framework encompasses many useful applications, it deviates from that of CvD by requiring the underlying parameter space to produce two "dominance regions" instead of

27

possibly just one. That is, if high parameter values correspond to the highest action being strictly dominant for all players and vice versa for low parameter values, then as long as some extra assumptions on preferences hold, a unique global games prediction emerges as signals become less and less noisy.

Little work in this area has been done in games of strategic substitutes (GSS), where a higher action from opponents induces a player to take a lower action. Morris and Shin (2009) show that this case can be much more complex by giving an example of a global game in the GSS setting which fails to produce a unique outcome from the process of iterated elimination of strictly dominated strategies (IESDS). Harrison (2003) studies a model where this difficulty can be overcome by considering 2-action aggregative GSS with sufficiently heterogeneous players and two overlapping dominance regions. Still, the global games solution can only be guaranteed to be a unique Bayesian Nash equilibrium, not necessarily the dominance solvable solution.

This chapter develops a global games framework which not only allows for the underlying game to be either a GSS or a GSC, but is also much less demanding than both FMP and Harrison in terms of the restrictions that preferences and the underlying parameter space must satisfy. For example, we require the existence of only one dominance region, which need not correspond to the highest or the lowest strategies in the action space. This is opposed to the requirement of two dominance regions, which, as the motivating example in Section 2 illustrates, is often not met. We also dispense of a state monotonicity assumption on preferences which is present in both FMP and Harrison, so that observing a higher parameter need not induce a player to take a higher strategy. Our approach is to directly extend the original 2×2 framework of CvD to multiple-player, multiple-action games by drawing on a common order property present in both GSS and GSC, while also overcoming the computational difficulties present in Harrison by using IESDS as our solution concept. In their original work, CvD use a "risk dominance" criteria to determine which equilibrium will be selected as the global games prediction. We are able to generalize this condition to a $p-$dominance condition which includes theirs as a special case. As in FMP, we also give conditions for when a global games prediction is "noise independent", so that it is robust to the

specification of the prior distribution on the parameter space. In fact, our $p-$dominance condition for selection is similar to the one given in FMP as a sufficient condition for noise independence, and equivalent in games with symmetric payoffs.

Lastly, our $p-$dominance condition for selection becomes more restrictive as the number of players grows larger. This can be motivated by the idea in the presence of more opponents and subsequently more actions to consider, a player must be more sure that a specific equilibrium is being played in order for her to reciprocate uniquely by playing her part in the equilibrium. In order to overcome this difficulty, we allow for the possibility that instead of a player believing that her opponents' information is independent conditional on her own information, she may instead perceive some correlation in the information of "groups" of others. This effectively reduces the amount of uncertainty present in the game, and in some situations may restore the full power of the $2-$player $p-$dominance condition even in the presence of an arbitrarily large amount of players.

The chapter is organized as follows: Section 2 presents a motivating example which highlights the contributions of this chapter in relation to the current literature. Section 3 lays out the basic model and assumptions, and presents the first of three main results. Section 4 introduces the grouping method which allows us to relax the assumption of conditionally independent beliefs, and generalizes the results in Section 3. Section 5 presents the results for noise independence and common valuations.

2.2   Motivating Example

Because our model is one of the first to address global games analysis in finite GSS, we use the following motivating example to highlight the contributions to the well-established literature in the GSC setting. Consider a slightly modified version of the technology adoption model considered in Keser, Suleymanova, and Wey (2012). Three agents must mutually decide on whether to adopt an inferior technology *A*, or a superior technology *B*. The benefit to each player *i* of adopting a

29

specific technology $t = A, B$ is given by

$$U_t(N_t) = v_t + \gamma_t(N_t - 1)$$

where N_t is the total number of players using technology t, v_t is the stand-alone benefit from using technology t, and γ_t is the benefit derived from the network effect of adopting the technology of others. It is assumed that $v_B > v_A$ in order to distinguish B as the superior technology. Assume for simplicity that $v_A = \gamma_A = 1$, and $\gamma_B = 3$. Letting $v_B = x$, we have the following payoff matrix:

		A	B
	A	3, 3, 3	2, x, 2
P1	B	x, 2, 2	x+3, x+3, 1

P3 = A (P2 header: A, B)

		A	B
	A	2, 2, x	1, x+3, x+3
P1	B	x+3, 1, x+3	x+6, x+6, x+6

P3 = B (P2 header: A, B)

For $x \in [1, 3]$, both (A, A, A) and (B, B, B) are strict Nash equilibria, and for $x > 3$ we have that (B, B, B) is the strictly dominant action profile, giving an "upper dominance region". Suppose that a modeler wishes to use the global games approach to resolving the issue of multiplicity on $[1, 3]$: By allowing each player i to observe only a noisy signal $x_i = x + v\varepsilon_i$ of the true parameter x, we obtain a Bayesian game with noise parameter v. FMP show that under certain assumptions, GSC like the game above eventually exhibit an "essentially unique" Bayesian prediction $s : \mathbb{R} \to A$ as the noise parameter $v \to 0$. That is, when noise is arbitrarily small about an observation x, the corresponding complete information game is approximated and we have a unique "global games prediction" given by $s(x)$.

However, notice that because we have the parameter restriction $x = v_B > v_A = 1$, no lower dominance region can be established,[1] and therefore the framework of FMP cannot be applied. One possible resolution of this problem has been given in Basteck, Daniëls, and Heinemann (2013), which shows that any GSC can be re-parametrized so that two dominance regions are established

---

[1]Likewise, if common knowledge about v_A is relaxed, no upper dominance region would be established.

and the conditions of FMP are met, producing a subsequent global games prediction. But this procedure can also be problematic, due to a recent observation by Weinstein and Yildiz (WY) (2007). In games of incomplete information, rationality arguments rely on analyzing a player's hierarchy of beliefs, that is, their belief about the parameter space, what they believe their opponents believe about the parameter space, and so on. In general situations, this information can be identified as a player's type, or beliefs over the parameter space and the types of others. WY shows that if the parameter space is "rich" enough, so that any given rationalizable strategy a_i^* for player i is strictly dominant at *some* parameter in the model, then the (degenerate) complete information beliefs can be slightly perturbed in such a way so as to make a_i^* the *unique* rationalizable action for player i. This poses a serious criticism to global games analysis: If players' beliefs can be slightly perturbed in a specific way so that any rationalizable strategy can be justified as the unique rationalizable strategy, how does a modeler know if the global games method is the "right way"? Notice that a re-parametrization of the above model à la Basteck, Daniëls, and Heinemann which produces upper and lower dominance regions automatically satisfies the richness condition of WY, making the global games selection ad hoc in this case.

By relaxing the need for two dominance regions, as well as requiring no state monotonicity assumptions, our framework allows us to expand global games analysis to cases like the one above by allowing for a more "natural" way to add uncertainty to the model. That is, by avoiding an arbitrary re-parametrization, we can use those parameters which are motivated in the description of the model itself, in this case v_B or v_A.

2.3 Model and Assumptions

The chapter will be stated in the case of GSS. When it is needed, the adjustments that are necessary for the results to hold for GSC will be pointed out.

Definition 7. A game $G = (\mathscr{I}, (A_i)_{i \in \mathscr{I}}, (u_i)_{i \in \mathscr{I}})$ of strategic substitutes has the following elements:

- The number of players is finite and given by the set $\mathscr{I} = \{1, 2, ..., N\}$.

- Each player i's action set is denoted by A_i and is finite and linearly ordered. Let $\bar{a}_i$ and $\underline{a}_i$ denote the largest and the smallest elements in A_i, respectively. Also, for a specific $\tilde{a}_i \in A_i$, denote $\tilde{a}_i^+ = \{a_i \in A_i \,|\, a_i > \tilde{a}_i\}$ and $\tilde{a}_i^- = \{a_i \in A_i \,|\, \tilde{a}_i > a_i\}$.

- Each player's utility function is given by $u_i : A \to \mathbb{R}$, and is continuous in all arguments.

- (Strategic Substitutes) For each player i, if $a_i' \geq a_i$ and $a_{-i}' \geq a_{-i}$, then

$$u_i(a_i, a_{-i}') - u_i(a_i', a_{-i}') \geq u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})$$

We will restrict our attention to games that exhibit multiple equilibria, and will require that the specific equilibrium under consideration is "strict" in the following sense:

Definition 8. Let $\tilde{a}$ be a Nash equilibrium. Then $\tilde{a}$ is a strict Nash equilibrium if for all i, and for all a_i,

$$u_i(\tilde{a}_i, \tilde{a}_{-i}) > u_i(a_i, \tilde{a}_{-i})$$

Simply, a Nash equilibrium is strict if each player is best responding uniquely to the other players when they play their part of the equilibrium. In order to resolve the issue of multiple equilibria in games where a specific $\tilde{a} \in A$ is a strict Nash equilibrium, we will need to "embed" such a game into a family of games such that $\tilde{a}$ is a strict Nash equilibrium on some neighboring range of parameters.

Definition 9. An $\tilde{a}-$based parametrized game of strategic substitutes $G_X = (\mathscr{I}, X, (A_i)_{i \in I}, (u_i)_{i \in I})$ has the following elements:

- $\forall x \in \mathbb{R}$, we denote $G_X(x)$ to be the unparametrized game when x is realized. We assume that for all x, $G_X(x)$ is a game of strategic substitutes. Also, denote by $NE_X(x)$ the set of strict Nash equilibria in $G_X(x)$.

- $X = [\underline{X}, \overline{X}]$ is a closed interval of $\mathbb{R}$ such that $\forall x \in X$, $\tilde{a} \in NE_X(x)$. We also make the following convention that $\forall x \geq \overline{X}$, $u_i(a, x) = u_i(a, \overline{X})$, and likewise $\forall x \leq \underline{X}$, $u_i(a, x) = u_i(a, \underline{X})$.[2]

- $\forall i \in \mathscr{I}, \forall a \in A$, $u_i(a, \cdot) : \mathbb{R} \to \mathbb{R}$ is a continuous function of x.

With the addition of a noise structure, a parametrized game of strategic substitutes becomes a Bayesian game. We will call a Bayesian game a global game if it has the specific noise structure defined below, and has the payoff properties as defined in Definition 9.

---

[2]Because our analysis will be focused on the interior of X, this is only for simplicity.

Definition 10. A global game $G(v) = (G_X, f, (\varphi_i)_{i \in \mathscr{I}})$ is a Bayesian game with the following elements:

- G_X is an $\tilde{a}$-based parametrized game of strategic substitutes for some $\tilde{a} \in A$.

- $f : \mathbb{R} \to [0, 1]$ is any continuous pdf with connected support.

- $\{\varphi_i\}_{i \in I}$ are uniform densities with support $[-1, 1]$, where $(f, \{\varphi_i\}_{i \in I})$ are mutually independent.

- Each player i receives a signal $x_i = x + v\varepsilon_i$, where x is distributed according to f and each ε_i is distributed according to φ_i. After Bayesian updating, players form a pdf over payoff parameters and the signals received by opponents, which is given by $f_i(\cdot|x_i, v) : \mathbb{R} \times \mathbb{R}^{N-1} \to [0, 1]$, with support[3]

$$supp(f_i(\cdot|x_i, v)) \subseteq [x_i - v, x_i + v] \times [x_i - 2v, x_i + 2v]^{N-1}$$

We let $F_i(\cdot|x_i, v)$ denote the corresponding cdf, and μ_{F_i} denote the Lebesgue–Stieltjes measure induced by $F_i(\cdot|x_i, v)$.[4]

Note that each global game $G(v)$ is characterized by the noise level v of the signal the players receive. In order to resolve multiplicity at any $x \in X$ (at which $\tilde{a}$ is a strict Nash equilibrium), the following process can then be followed: At each noise level v, the upper and lower serially undominated strategies $\bar{s}^v$ and $\underline{s}^v$ can be calculated.[5] Suppose that as $v \to 0$, the upper and lower serially undominated strategies agree, so that for each x, $s(x) \equiv \bar{s}^v(x) = \underline{s}^v(x)$ provides a unique

---

[3] Since $x_i = x + v\varepsilon_i$ and $x_j = x + v\varepsilon_j$, $x_j = x_i - v\varepsilon_i + v\varepsilon_j$.
[4] That is, the unique measure on $\mathbb{R} \times \mathbb{R}^{N-1}$ such that $\forall \overrightarrow{x} = (x, x_{-i}) \in \mathbb{R} \times \mathbb{R}^{N-1}$, $F_i(\overrightarrow{x}|x_i, v) = \mu_{F_i}((-\infty, \overrightarrow{x}]|x_i, v)$.
[5] In Hoffmann (2014) it is established that any Bayesian game of strategic substitutes has a smallest and a largest strategy profile surviving iterated deletion of dominated strategies.

prediction in a slightly noisy version of the complete information game $G_X(x)$. Because this noise can be made arbitrarily small, we will be justified in choosing this equilibrium in the complete information game.

Once the signal is received, player i chooses a strategy, hence forming a strategy function $s_i : \mathbb{R} \to A_i$. We denote all of player $i's$ strategy functions by the set S_i. Player i's expected utility from playing strategy a_i against the strategy function s_{-i} after receiving x_i is given by

$$\pi_{x,x_{-i}}(a_i, s_{-i}, x_i) = \int_{\mathbb{R}^{N-1}} \int_{\mathbb{R}} u_i(a_i, s_{-i}(x_{-i}), x) d\mu_{F_i}(x_{-i}, x | x_i, v)$$

In order to analyze $G(v)$, it will be useful to analyze a simplified version $G^*(v)$ which is described below.

Definition 11. For any global game $G(v)$, the simplified global game $G^*(v)$ is defined as in Definition 10 with the following alterations:

- $f^* : \mathbb{R} \to [0, 1]$ is distributed uniformly on some interval containing $[\underline{X} - 2v, \bar{X} + 2v]$.

- $\forall i \in \mathscr{I}, \forall a \in A, u_i(a, \cdot)$ depends only on the signal x_i received, and not the underlying state x.

Therefore, if $F_i^*(x, x_{-i} | x_i, v)$ is the corresponding cdf over (x, x_{-i}) after receiving x_i and $F_i^*(x_{-i} | x_i, v)$ is the corresponding marginal cdf, then by denoting $\mu_{F_i}^*$ as the measure induced by $F_i^*(x_{-i} | x_i, v)$, player i's expected utility from playing strategy a_i against the strategy function s_{-i} after receiving x_i in $G^*(v)$ is given by

$$\pi_{x_{-i}}(a_i, s_{-i}, x_i) = \int_{\mathbb{R}^{N-1}} u_i(a_i, s_{-i}(x_{-i}), x_i) d\mu_{F_i}^*(x_{-i} | x_i, v)$$

We will proceed by establishing all results for a simplified global game $G^*(v)$, and then show in Section 5 that these results can be extended to the underlying $G(v)$. To simplify notation, we let $\triangle u_i(a'_i, a_i, a_{-i}, x) = u_i(a'_i, a_{-i}, x) - u_i(a_i, a_{-i}, x)$ be player $i's$ advantage of playing a'_i over a_i when facing a_{-i} at a given x . Similarly, we write $\triangle \pi_{x_{-i}}(a'_i, a_i, s_{-i}, x)$ for player i's expected advantage from playing a'_i against s_{-i} after receiving signal x.

Much of our analysis will involve characterizing the set of serially undominated strategies in a global game.

Definition 12. Let $G^*(v)$ be a simplified global game. For each player $i \in \mathscr{I}$, and each $a_i \in A_i$, define the following:

- $\mathscr{P}_{i,a_i}^{v,0} = \emptyset$, $\mathscr{S}_i^{v,0} = S_i$

- $\forall n > 0$,
$$\mathscr{P}_{i,a_i}^{v,n} = \{x \in X \,|\, \forall a'_i \in A_i, \forall s_{-i} \in \mathscr{S}_{-i}^{v,n-1}, \triangle \pi_{x_{-i}}(a_i, a'_i, s_{-i}, x) > 0\},$$
$$\mathscr{S}_i^{v,n} = \{s_i \in \mathscr{S}_i^{v,n-1} \,|\, \forall a_i, s_i \,|_{\mathscr{P}_{i,a_i}^{v,n}} = a_i\}$$

- $\mathscr{P}_{i,a_i}^v = \underset{n \geq 0}{\cup} \mathscr{P}_{i,a_i}^{v,n}$, $\mathscr{S}_i^v = \underset{n \geq 0}{\cap} \mathscr{S}_i^{v,n}$

It is an easy fact to check that for each a_i and each n, $\mathscr{P}_{i,a_i}^{v,n} \subseteq \mathscr{P}_{i,a_i}^{v,n+1}$, $\mathscr{S}_i^{v,n+1} \subseteq \mathscr{S}_i^{v,n}$, and that the set of serially undominated strategies for player i in a global game is a subset of the set $\mathscr{S}_i^v$.

P-dominance

In their 2×2 formulation, CvD show that the "risk dominant" equilibrium will be the one that is selected in the global games procedure. In multiple-player, multiple-action games this notion can be extended to $p-$dominance, which we now define.[6]

Definition 13. Let $\widetilde{a}$ be a Nash equilibrium at some $x \in \mathbb{R}$. Then $\widetilde{a}$ is $p(x)-$dominant, where $p(x) = (p_1(x), p_2(x), ..., p_N(x))$, if for each player i, $p_i(x)$ is the smallest value satisfying

$$l_i(a_i, \lambda_i, x) \equiv \sum_{a_{-i}} \triangle u_i(\widetilde{a}_i, a_i, a_{-i}, x) \lambda_i(a_{-i}) \geq 0.$$

for each $a_i \in A_i$ and $\lambda_i \in \triangle(A_{-i})$ such that $\lambda_i(\tilde{a}_{-i}) \geq p_i(x)$.

Taking the smallest of all such values is natural because any value larger than $p_i(x)$ will also satisfy the definition. It will also be useful to recast Definition 13 involving a fixed $a_i \in A_i$ for player i, which we denote by $p_i(a_i, x)$: The lowest value that player i must see $\tilde{a}_{-i}$ being played at x so that $\tilde{a}_i$ does at least as good as a_i. Note that any strictly dominant strategy $\tilde{a}_i$ is $0-$dominant, and any part of a Nash equilibrium is $1-$dominant. Therefore, the lower the $p_i(x)$, the more dominant a strategy $\widetilde{a}_i$ is for player i at x. Also, its easily seen that each l_i is a continuous function of both λ_i and x.

The global games method relies heavily on the presence of "dominance regions". These are defined as subsets of the parameter space on which it is strictly dominant for some or all players to choose strategies consistent with some profile $\tilde{a}$. Unlike FMP (2003), Harrison (2003), and Morris and Shin (2000, 2009), we require the presense of only one dominance region associated with only one action profile $\tilde{a}$. Furthemore, we directly generalize CvD by requiring that only $N - 1$ of the players have a dominant action in the dominance region. The concepts are defined below:

---

[6]Risk-dominance is simply $p-$dominance in the case of a 2×2 game where each $p_i = \frac{1}{2}$.

Let G_X be a parameterized GSS, and $\tilde{a} \in A$ be an action profile. For each player i and strategy $a_i \neq \tilde{a}_i$, define

$$D^{\tilde{a}_i, a_i} = \{x \in \mathbb{R} \,|\, \forall a_{-i}, \triangle u_i(\widetilde{a}_i, a_i, a_{-i}, x) > 0\}$$

Then player i's dominance region is given by $D^{\tilde{a}_i} = \underset{a_i}{\cap} D^{\tilde{a}_i, a_i}$. Also, let us define $D^{\tilde{a}}$ by a set of x's in a parametrized game at which $\tilde{a}$ is a strict Nash equilibrium and strictly dominant for $N-1$ of the players. That is,

$$D^{\tilde{a}} = \{x \in X \,|\, \tilde{a} \in NE_X(x) \cap D^{\tilde{a}_i} \; for \; N-1 \; of \; the \; i \in \mathscr{I}\}$$

We will assume from here on that $D^{\tilde{a}}$ is an open interval, which, by our continuity assumptions, is WLOG.

Finally, we will require that a certain $p-$dominance condition holds in order to make our equilibrium selection. We distinguish between the $2-$player case and the multiple-player case by defining $P(N)$ as follows:

- For $N = 2$,

$$P(N) = \left\{ x \in X \,|\, \forall i, j \in I, \, p_i(x) + p_j(x) < 1 \right\}$$

- For $N > 2$,

$$P(N) = \left\{ x \in X \,|\, \forall i \in I, \, p_i(x) < \frac{1}{N} \right\}$$

We now state the first of two main theorems:

Theorem 6. *Let $G^*(v) = (G_X, f^*, \{\varphi_i\}_{i \in I})$ be a simple global game, where G_X is an $\tilde{a}$-based parameterized game of strategic substitutes for some $\tilde{a} \in A$. If*

1. *$\tilde{a}$ is $p-$dominant on $P(N)$*

2. *$I = (a, b) \subseteq P(N)$ is an open interval such that*

    (a) *$I \cap D^{\widetilde{a}} \neq \emptyset$.*

    (b) *$\exists \alpha_I > 0$ such that $[a - \alpha_I, b + \alpha_I] \subseteq int(P(N))$*

3. *$P(N) \subseteq X$*

Then for each $x \in I$, there exists a $\tilde{v} > 0$ such that for all $v \in (0, \tilde{v}]$, $\underline{s}^v(x) = \bar{s}^v(x) = \tilde{a}$.

That is, for v small, because action spaces are linearly ordered, any serially undominated strategy in $G^*(v)$ selects $\tilde{a}$ at any x satisfying the conditions of Theorem 6. Note that the requirements of $P(N)$ become more and more demanding as the number of players gets larger. The next section of this chapter considers a method for resolving this issue.

Below we establish some useful facts which will allow for a sketch of the proof of Theorem 6 before proving it in more generality (Theorem 7). The following Lemma highlights the role of strategic substitutes in the model. In particular, they allow us to characterize the iterated deletion of strictly dominated strategies.

Lemma 4. *For each player $i \in \mathscr{I}$, define*

$$\bar{s}_i^{v,n} = \begin{cases} a_i & , if \, x \in \mathscr{P}_{i,a_i}^{v,n} \\ \bar{a}_i & , otherwise \end{cases} \quad and \quad \underline{s}_i^{v,n} = \begin{cases} a_i & , if \, x \in \mathscr{P}_{i,a_i}^{v,n} \\ \underline{a}_i & , otherwise \end{cases}$$

and similarly $\bar{s}_i^v$ and $\underline{s}_i^v$ by replacing the $\mathscr{P}_{i,a_i}^{v,n}$ with $\mathscr{P}_{i,a_i}^v$. Then,

1. *$\forall n, \; \underline{s}_i^{v,n} \leq \underline{s}_i^v \leq \bar{s}_i^v \leq \bar{s}_i^{v,n}$.*

2. *$\bar{s}_i^{v,n} \to \bar{s}_i^v$ and $\underline{s}_i^{v,n} \to \underline{s}_i^v$ pointwise as $n \to \infty$.*

3. *For a given $\tilde{a}_i \in A_i$, then $x \in \mathscr{P}_{i,\tilde{a}_i}^v$ if and only if* [7]

   (a) *$\forall a_i \in \tilde{a}_i^+, \; \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}_{-i}^v, x) > 0$ and*

   (b) *$\forall a_i \in \tilde{a}_i^-, \; \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \bar{s}_{-i}^v, x) > 0$*

4. *Moreover, if $\underline{x} \in D^{\tilde{a}}$ and $\bar{v} > 0$ are such that $\forall x \in B(\underline{x}, 2\bar{v})$, $B(x, 2\bar{v}) \subseteq D^{\tilde{a}}$. Then $i \in \mathscr{I}$, $\forall v \in (0, \bar{v}]$,*

$$B(\underline{x}, 2v) \subseteq \mathscr{P}_{i,\tilde{a}_i}^v$$

Proof. For the first claim, suppose that for some a_i, $x \in \mathscr{P}_{i,a_i}^{v,n}$. Then $x \in \bigcup_{n \geq 0} \mathscr{P}_{i,a_i}^{v,n} = \mathscr{P}_{i,a_i}^v$, so that $\underline{s}_i^{v,n}(x) = \underline{s}_i^v(x)$. Therefore if $\underline{s}_i^{v,n}(x) > \underline{s}_i^v(x)$, we must have that $x \in (\bigcup_{a_i} \mathscr{P}_{i,a_i}^{v,n})^C$. But then $\underline{s}_i^{v,n}(x) = \underline{a}_i$, a contradiction. The same argument applies to show $\bar{s}_i^v \leq \bar{s}_i^{v,n}$, and obviously $\underline{s}_i^v \leq \bar{s}_i^v$.

Secondly, let x be given. If for some a_i, $x \in \mathscr{P}_{i,a_i}^v$, then since $\mathscr{P}_{i,a_i}^v = \bigcup_{n \geq 0} \mathscr{P}_{i,a_i}^{v,n}$, and the $\mathscr{P}_{i,a_i}^{v,n}$

---

[7] By interchanging $\underline{s}_{-i}^v$ and $\bar{s}_{-i}^v$ in conditions (a) and (b), we get the corresponding result for GSC. This is essentially the only point of difference between the two cases.

are an increasing sequence of sets, $\exists N, \forall n \geq N, x \in \mathscr{P}_{i,a_i}^{v,n}$, so that $\underline{s}_i^{v,n}(x) \to \underline{s}_i^v(x)$. If $x \in (\cup_{a_i} \mathscr{P}_{i,a_i}^v)^C$, then since for all n, $\cup_{a_i} \mathscr{P}_{i,a_i}^{v,n} \subseteq \cup_{a_i} \mathscr{P}_{i,a_i}^v$, we must have that $x \in (\cup_{a_i} \mathscr{P}_{i,a_i}^v)^C$, giving convergence. The same arguments can be made to show that $\bar{s}_i^{v,n} \to \bar{s}_i^v$.

For the third claim, suppose $x \in \mathscr{P}_{i,\tilde{a}_i}^v$. Since $\mathscr{P}_{i,\tilde{a}_i}^v = \cup_{n \geq 0} \mathscr{P}_{i,\tilde{a}_i}^{v,n}$, $\exists N > 0$ such that $x \in \mathscr{P}_{i,\tilde{a}_i}^{v,N}$. We now show that for all n, $\bar{s}_i^v$ and $\underline{s}_i^v$ are in $\mathscr{S}_i^{v,n}$. Suppose this is not the case. Then $\exists n$, $\exists a_i$, $\exists x' \in \mathscr{P}_{i,a_i}^{v,n}$ such that $\bar{s}_i^v(x') \neq a_i$ or $\underline{s}_i^v(x') \neq a_i$. Since $\forall n$, $\mathscr{P}_{i,a_i}^{v,n} \subseteq \mathscr{P}_{i,a_i}^v$, then $x' \in \mathscr{P}_{i,a_i}^v$ but $\bar{s}_i^v(x') \neq a_i$ or $\underline{s}_i^v(x') \neq a_i$, a contradiction. Thus, since this holds for all n, $\bar{s}_i^v$ and $\underline{s}_i^v$ are in $\mathscr{S}_i^{v,N-1}$, and since $x \in \mathscr{P}_{i,\tilde{a}_i}^{v,N}$, $a_i \in \tilde{a}_i^+ \Rightarrow \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}_{-i}^v, x) > 0$ and $\forall a_i \in \tilde{a}_i^- \Rightarrow \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \bar{s}_{-i}^v, x) > 0$.

Conversely, suppose that $\forall a_i \in \tilde{a}_i^+$, $\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}_{-i}^v, x) > 0$ and $\forall a_i \in \tilde{a}_i^-$, $\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \bar{s}_{-i}^v, x) > 0$. Suppose by way of contradiction that $x \notin \mathscr{P}_{i,\tilde{a}_i}^v$. Then $\forall n \geq 1$, there exists a $s_{-i} \in \mathscr{S}_{-i}^{v,n-1}$ such that $\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, s_{-i}, x) \leq 0$ for either some $a_i \in \tilde{a}_i^+$ or $a_i \in \tilde{a}_i^-$. Let $(s_{-i}^n)_{n=1}^\infty$ be any arbitrary collection of such s_{-i} at each n, and notice that for each n, $\bar{s}_{-i}^{v,n} \geq s_{-i}^n \geq \underline{s}_{-i}^{v,n}$. Suppose that $a_i \in \tilde{a}_i^+$. By GSS, we must have that $\forall n$, $\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, s_{-i}^n, x) \geq \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}_{-i}^{v,n}, x)$. By hypothesis, $\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}_{-i}^v, x) > 0$, and hence by continuity and the fact that $\underline{s}_i^{v,n} \to \underline{s}_i^v$, we must have that for n large, $\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}_{-i}^{v,n}, x) > 0$. Therefore, for n large, we must have $\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, s_{-i}^n, x) \geq \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}_{-i}^{v,n}, x) > 0$, or $\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, s_{-i}^n, x) > 0$. Because the same argument can be made for any $a_i \in \tilde{a}_i^-$ using the fact that $\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, s_{-i}^n, x) \geq \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \bar{s}_{-i}^{v,n}, x) > 0$, we contradict the construction of $(s_{-i}^n)_{n=1}^\infty$, and hence $x \in \mathscr{P}_{i,\tilde{a}_i}^v$.

For part 4, let $v \in (0, \bar{v}]$ and i be the player not among the $N-1$ associated with $D^{\tilde{a}}$. Suppose $x \in B(\underline{x}, 2v)$. By hypothesis, we have that $B(x, 2v) \subseteq D^{\tilde{a}}$. Therefore, $\forall j \neq i$, $\forall n \geq 1$, $B(x, 2v) \subseteq D^{\tilde{a}_j} \subseteq \mathscr{P}_{j,\tilde{a}_j}^{v,n}$, so that $B(x, 2v) \subseteq \mathscr{P}_{j,\tilde{a}_j}^v$. For player i, suppose $x \in B(\underline{x}, 2v)$, $n \geq 2$, and let $s_{-i} \in \mathscr{S}_{-i}^{v,n-1}$ be arbitrary. Since for each $\forall j \neq i$ we must have $B(x, 2v) \subseteq \mathscr{P}_{i,\tilde{a}_j}^{v,n-1}$, then each $s_j |_{B(x,2v)} = \tilde{a}_j$. Therefore, since $\tilde{a}$ is a strict Nash equilibrium at each such x, we have that $\forall a_i \in A_i$,

$$\triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, s_{-i}, x) = \triangle u_i(\tilde{a}_i, a_i, \tilde{a}_{-i}, x) > 0$$

41

That is, $B(\underline{x}, 2v) \subseteq \mathscr{P}^{v,n}_{i,\tilde{a}_i} \subseteq \mathscr{P}^{v}_{i,\tilde{a}_i}$.

$\square$

The next result can be helpful even beyond the scope of the results presented here. It not only provides us with a method for calculating the individual p_i for which an equilibrium is p-dominant, but also shows that in our setting, the p_i satisfy a useful continuity property when viewed as a function of x. We first calculate such a value of p_i with a fixed a_i. When a parameter space X is mentioned, assume an arbitrary global games embedding.

Proposition 4. *Let $\tilde{a}$ be a strict Nash equilibrium on X. For each player $i \in I$, and $a_i \in A_i$, we have that*

1. $p_i(a_i, x) = \begin{cases} \displaystyle\max_{\substack{\lambda_i \in \triangle(A_{-i}) \\ l_i(a_i, \lambda_i, x) = 0}} (\lambda_i(\tilde{a}_{-i})), & x \notin D^{\tilde{a}_i, a_i} \\ \\ 0, & x \in D^{\tilde{a}_i, a_i} \end{cases}$

2. $\forall x \in X,\ p_i(x) = \displaystyle\max_{a_i}(p_i(a_i, x))$

3. p_i *is an upper semi-continuous function on X.*

Proof. Appendix.

$\square$

In the $2-$player case, we will use a particular "trick" with players' beliefs that was developed by CvD, which is given in Lemma 5. To this end, we will need to be able to ensure that if both players receive a signal close enough to each other, so that the complete information games that they respectively observe are approximately equal in terms of payoffs, then the $p-$dominance condition continues to hold. The following Corollary tells us that this will be true:

Corollary 3. *Suppose the conditions of Theorem 6 hold in the case of $N = 2$. If $[a, b] \subseteq int\,(P(2))$,*

then $\exists \bar{v} > 0$, $\forall v \in (0, \bar{v}]$, $\forall x, y \in [a, b]$,

$$d(x, y) < v \implies p_i(x) + p_j(y) < 1$$

Proof. Let $v*$ be such that $\forall x \in [a, b]$, $B(x, v*) \subseteq P$. For a contradiction, suppose that $\forall v \in (0, v*]$,

$\exists v' \leq v$, $\exists x_{v'}, y_{v'} \in [a, b]$ such that $d(x_{v'}, y_{v'}) < v'$ and $p_i(x_{v'}) + p_j(y_{v'}) \geq 1$. Collecting all such

$(x_{v'}, y_{v'})_{v>0}$, then since for all v', $x_{v'} \in [a, b]$, by passing to a subsequence if necessary, we may

assume that $x_{v'} \to x* \in [a, b]$. Since $\forall v'$, $d(x_{v'}, y_{v'}) < v'$, then then $y_{v'} \to x*$. By Proposition 1,

since p_i and p_j are upper semi-continuous,

$$p_i(x*) + p_j(x*) \geq \limsup_{v' \to 0} (p_i(x_{v'}) + p_j(y_{v'})) \geq 1$$

contradicting the fact that $x* \in P(2)$. Therefore, there exists a $\bar{v} > 0$ satisfying the hypothesis.

$\square$

Finally, we must establish how conditional beliefs are formed. Again, this is broken up into the

$2-$player and multiple-player cases. For any signal $x_i \in X$, denote by $\vec{x}_{-i}$ the $(N-1) \times 1$ vector of

oppenents signals, each equal to x_i.

Lemma 5. *Suppose the conditions of Theorem 6 hold. Then for each $i \in \mathcal{I}$ and $x_i, x_j \in I$, we have:*

- For $N = 2$,

$$F_i^* (x_j | x_i, v) + F_i^* (x_i | x_j, v) = 1$$

- For $N \geq 2$,

$$F_i^* (\vec{x}_{-i} | x_i, v) \geq \frac{1}{N}$$

Proof. Appendix.

□

Below is a heuristic sketch of the proof of Theorem 6, the full proof in more generality being relegated to the next section.

Proof. (Sketch of Theorem 6) Suppose the conditions of Theorem 6 hold. Let $(-\infty, \underline{x})$ be a region on which $\tilde{a}$ is strictly dominant for each player, and $[\underline{x}, \infty) \subseteq P(N)$. For a contradiction, suppose that there is some $\tilde{x} \in I$ such that for all $v > 0$, there is some serially undominated strategy s such that $s(\tilde{x}) \neq \tilde{a}$, and for each player i consider the points $x_i^v = sup(x \,|\, [\underline{x}, x) \subseteq \mathscr{P}_{i,\tilde{a}_i}^v)$. Recalling that since $\mathscr{P}_{i,\tilde{a}_i}^v$ contains those $x's$ at which *every* serially undominated strategy s_i plays $\tilde{a}_i$ for player i, $s(\tilde{x}) \neq \tilde{a}$ implies that for some i, $x_i^v < \infty$. We now consider both the $2-$player and the multiple-player cases:

- 2 Players: Because $x_i^v < \infty$, it must be the case that at x_i^v, player i does not observe $\tilde{a}_j$ being played with high enough probability ($> p_i(x_i^v)$) to unambiguously best respond with $\tilde{a}_i$. This implies that player j must also discontinue playing $\tilde{a}_j$ somewhere near x_i^v, or $x_j^v \in B(x_i^v, 2v)$, so that player j also does not observe $\tilde{a}_i$ being played with high enough probability ($> p_j(x_j^v)$) to unambiguously best respond with $\tilde{a}_j$. If we suppose that x_i^v is the lowest of the two points, so that $x_j^v = x_i^v + d$, we get the following graphical representation:
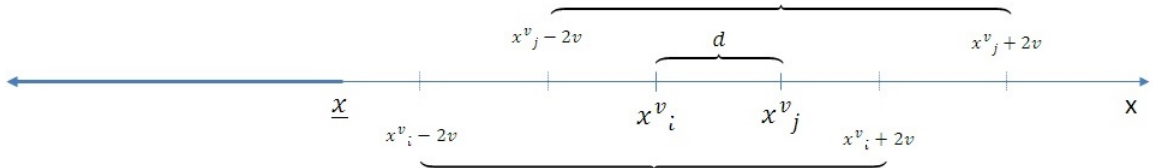
Because v could have been taken to be small enough to satisfy Corollary 3, we have that $p_i(x_i^v) + p_j(x_j^v) < 1$. However, because player 1 sees $\tilde{a}_j$ being played (approximately) with at

least probability $\left(\frac{1}{2} + \frac{d}{2v}\right)$ and player 2 sees $\tilde{a}_j$ being played (approximately) with at least probability $\left(\frac{1}{2} - \frac{d}{2v}\right)$, by the above argument we must have that

$$p_l(x_l^v) + p_j(x_j^v) \geq \left(\frac{1}{2} + \frac{d}{2v}\right) + \left(\frac{1}{2} - \frac{d}{2v}\right) = 1$$

- $N > 2$ Players: Following along the same argument as in the $2-$player case, let x_i^v be the lowest of all such points among all players. Since at x_i^v player i does not observe $\tilde{a}_{-i}$ being played with high enough probability to unambiguously best respond with $\tilde{a}_i$, and since x_i^v being the the lowest of all x_j^v implies that player i observes $\tilde{a}_{-i}$ being played with at least probability $F_i^*\left(\vec{x}_{-i}^v | x_i^v, v\right)$, then we musts have that $p_i(x_i^v) \geq F_i^*\left(\vec{x}_{-i}^v | x_i^v, v\right)$. However, by Lemma 9, $F_i^*\left(\vec{x}_{-i}^v | x_i^v, v\right) \geq \frac{1}{N}$, contradicting the fact that since $x_i^v \in P(N)$, $p_i(x_i^v) < \frac{1}{N}$.

$\square$

We now consider some examples:

Example 5. Consider again the modified version of the technology adoption model considered in Keser, Suleymanova, and Wey (2012) from the introduction. Again allowing $v_B = x$ to represent the parameter of uncertainty, and using the same values for v_A, γ_A, and γ_B as before, the payoff matrix is given by the following:

		A P3				B	
		P2				P2	
		A	B			A	B
P1	A	3, 3, 3	2, x, 2		A	2, 2, x	1, x+3, x+3
P1	B	x, 2, 2	x+3, x+3, 1		B	x+3, 1, x+3	x+6, x+6, x+6

For $x \in [1, 3]$, (A, A, A) and (B, B, B) are strict Nash equilibria, and for $x > 3$, B is the strictly dominant action for each player. Recall that because we have the parameter restriction $x = v_B > v_A = 1$, no lower dominance region can be established as in the FMP framework, and the same is true about the upper dominance region if CK about v_A is relaxed. Hence the FMP framework does not apply to any "natural" parameters present in the model.

Applying Theorem 6, we have that for each $i = 1, 2, 3$,

$$
p_i(x) = \begin{cases} \frac{3-x}{8} & 1 < x < 3 \\ \\ 0 & x \geq 3 \end{cases}
$$

In order to satisfy $p_i(x) < \frac{1}{3}$ for all $i \in I$, we have that (B, B, B) is the global games prediction for any $x > \frac{1}{3}$. Because 1 was the lower bound for x in this model, indeterminacy is completely resolved.

Example 6. Consider a scenario of deterrence between two countries (Player 1 and Player 2), often modeled by a game of Chicken. In this game, both countries must decide between an aggressive strategy (A) or capitulation (C). We follow the formulation of Baliga and Sjostrom (2009) by allowing h_i to be Player i's preference for aggression, c and d the respective costs of being aggressive or capitulating in the face of an aggressive opponent, and normalize the payoffs of mutual capitulation to 0. By assuming that $0 < h_i < c - d$ for each player, we have a game of Chicken (GSS) with two strict Nash equilibrium (A, C) and (C, A), represented below:

$$P2$$

		A	C
	A	$h_1 - c, h_2 - c$	$h_1, -d$
$P1$	C	$-d, h_2$	$0, 0$

Examples of such situations where capitulating to an attacking opponent is preferred to mutual aggression are numerous, including the Cuban missile crisis, the Munich crisis of 1938, and the

Berlin crisis of 1948. In each of these examples, only one of the two equilibria emerges, with one party capitulating to the other. Goldmann (1994) suggests that the prevailing party will be the one that is able to express a stronger preference for aggression. Along these lines, Kilgour and Zagare (1991) formulate a model in which each player is uncertain about the opponent's preferences and conclude that a player will capitulate if they perceive a high enough probability of aggression from the other party, which they deem a "credible threat".

Suppose that we allow Player 1 to "send a signal" of their preference for aggression by allowing $h_1 = x$. Then for $0 < x \le c - d$ we have multiple equilibria as before, but for $x > c - d$ we have that A is strictly preferred for Player 1. Calculating $p_1(x)$ and $p_2(x)$ gives $p_2(x) = \frac{h_2}{c-d}$ and

$$
p_1(x) = \begin{cases} \frac{c-d-x}{c-d} & 0 < x < c-d \\ \\ 0 & x \ge c-d \end{cases}
$$

If we instead define a credible threat to be any value of x that is ex-ante plausible in some situation (within the support of the prior), then as long as Player 1 can make credible threats that are sufficiently high ($x > c - d$), we can apply Theorem 6 to the profile (A, C). We find that the condition

$$
p_1(x) + p_2(x) < 1
$$

is satisfied for all $x > h_2$. The interpretation is in line with previous hypotheses: If Player 1 is able to express a slightly stronger preference for aggression, she will be able to prevail by getting the opponent to capitulate. In fact, this suggests that the capitulating party need not perceive a high probability of aggression from the other opponent as in Kilgour and Zagare (1991), all that is needed is a slight difference in preference for aggression along with a "kernel of doubt" in beliefs.

2.4 Groups

In this section we relax the classical global games assumption of conditionally independent beliefs. That is, after receiving signal x_i, we allow for the possibility that player i perceives correlation among the signals of her opponents. There are at least two main motivations for doing this. First, in many situations it is more natural to assume that players do in fact receive correlated signals. For example, consider an n-firm Cournot economy, where two groups of firms are separated geographically from one another. If common knowledge of a weather forecast-based parameter is relaxed, it is more plausible to assume that the firms in one region receive the same forecast. Second, there is strong evidence in the social psychology literature which suggests that decision makers often exhibit "stereotyping" behavior, which may lead them to infer more correlation about opponents' types than is actually present.[8] As Macrae and Bodenhausen (2000) write, "Given basic cognitive limitations and a challenging stimulus world, perceivers need some way to simplify and structure the person perception process. This they achieve through the activation and implementation of categorical thinking (Allport 1954, Bodenhausen & Macrae 1998, Brewer 1988, Bruner 1957, Fiske & Neuberg 1990)." They go on to write that, "The principal function of activated categorical representations is to provide the perceiver with expectancies that can guide the processing of subsequently encountered information (Olson et al 1996). As previously noted, there are two primary ways that expectancies can influence subsequent information processing. First, they can serve as frameworks for the assimilation and integration of expectancy-consistent information, leading the perceiver to emphasize stereotype-consistent information to a greater extent than he or she would have in the absence of category activation. (e.g. Fiske 1998, Fyock & Stangor 1994, Macrae et al 1994b,c).". That is, categorical thinking by a player i may lead to a partitioning of the space $\mathscr{I} / \{i\}$, in which each "group" of opponents is ascribed the same informational attribute, or signal in this context.

---

[8]See Healy (2007) for a comprehensive literature review.

The next definition formalizes this thinking:

Definition 14. Let $\widetilde{a} \in A$. Then $\mathscr{G}^{\widetilde{a}} = \{g^1, g^2, ..., g^N\}$ is an $\widetilde{a}-$based partitioning of I if $\forall i \in \mathscr{I}$:

1. $g^i = g^i_i \cup g^i_{-i}$ is a partition of $\mathscr{I}$, where $g^i_i = \{i\}$ and g^i_{-i} partitions $\mathscr{I}/\{i\}$. This can be expressed as $g^i = \left(\{g^i_m\}_{m=1}^{|g^i|} \right)$, where g^i_m is the element containing player m and $|g^i|$ denotes the total number of elements.

2. It is common knowledge that Player i updates her beliefs by assuming that each player in each group g^i_j receives $x_{g^i_j} = x + v\varepsilon_{g^i_j}$, each $\varepsilon_{g^i_j}$ being distributed according to uniform density $\varphi_{g^i_j}$ with support $[-1, 1]$, where $\left(f^*, \left\{ \varphi_{g^i_j} \right\}_{i \in I, g^i_j \in g^i} \right)$ are mutually independent.

One extreme case is when each g^i is the trivial partitioning consisting of singletons, which reduces to the traditional global games assumption and the formulation in Theorem 6. The other extreme case is when a g^i consists of only two elements, placing all other opponents into the same group. This may be a more natural assumption in the presence of a large number of players, or in more mentally demanding situations, as there is evidence that judgment becomes more stereo-typical under cognitive load (Macrae and Bodenhausen (2000)). Absent in the above definition is the notion of "self-stereotyping", that is, players never include themselves in any other opponent's partition element. Allowing for self-stereotyping can be achieved under additional assumptions and a slightly more complicated proof, which can be found in the online Appendix. Finally, for any player i and signal x_i, we will abuse notation by allowing x_{-i} by the signals received by the groups in i's partitioning.[9]

---

[9]For example, if $I = \{1, 2, 3, 4\}$ and $g^1 = \{\{1\}, \{2\}, \{3, 4\}\}$, then $x_{-1} = \left(x_{g^1_2}, x_{g^1_3} \right)$, where $g^1_2 = \{2\}$ and $g^1_3 = \{3, 4\}$. In this way, a function $s_{-1}(x_{-1})$ is evaluated as $\left(s_2(x_{g^1_2}), s_3(x_{g^1_3}), s_4(x_{g^1_3}) \right)$.

By defining $N_g = \max_{i \in I} (|g^i|)$, we again distinguish the $p-$dominance condition that must be satisfied between the $2-$player case and the multiple-player case by defining $P(N_g)$ as follows:

- For $N_g = 2$,

$$P(N_g) = \{x \in X \mid \forall i, j \in I, \, p_i(x) + p_j(x) < 1\}$$

- For $N_g > 2$,

$$P(N_g) = \left\{x \in X \mid \forall i \in I, \, p_i(x) < \frac{1}{N_g}\right\}$$

The second main Theorem is stated below:

Theorem 7. *Let* $G^*(v) = \left(G_X, f^*, \left\{\varphi_{g_j^i}\right\}_{i \in I, g_j^i \in g^i}\right)$ *be a simple global game, where* G_X *is an* $\tilde{a}$*-based parameterized game of strategic substitutes for some* $\tilde{a} \in A$*, and* $\mathscr{G}^{\tilde{a}} = \{g^1, g^2, ..., g^N\}$ *is an* $\tilde{a}-$*based partitioning of* $\mathscr{I}$*. If*

1. $\tilde{a}$ *is* $p-$*dominant on* $P(N_g)$

2. $I = (a, b) \subseteq P(N_g)$ *is an open interval such that*

    (a) $I \cap D^{\tilde{a}} \neq \emptyset$.

    (b) $\exists \alpha_I > 0$ *such that* $[a - \alpha_I, b + \alpha_I] \subseteq int(P(N_g))$

3. $P(N_g) \subseteq X$.

Then for each $x \in I$*, there exists a* $\tilde{v} > 0$ *such that for all* $v \in (0, \tilde{v}]$*,* $\underline{s}^v(x) = \bar{s}^v(x) = \tilde{a}$.

In what follows, we will let $s_{-i}^v(x_i)$ denote the probability with which player i believes that her opponents will play $\tilde{a}_{-i}$ according to s_{-i} if x_i is observed. Specifically,

$$s_{-i}^v(x_i) = \int_{\mathbb{R}^{|g^i|-1}} (1_{\{s_{-i} = \tilde{a}_{-i}\}}) d\mu_{F_i}^*(x_{-i} | x_i, v)$$

50

Proposition 5. *Suppose the conditions of Theorem 7 hold, let $i \in \mathscr{I}$, and $x_i \in P$. Suppose it's the case that $0 \geq \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}^v_{-i}, x_i)$ for some $a_i \in \tilde{a}_i^+$ or $0 \geq \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \bar{s}^v_{-i}, x_i)$ for some $a_i \in \tilde{a}_i^-$.*

Then:

1. $\exists j \neq i$ such that $B(x_i, 2v) \not\subseteq P^v_{j, \tilde{a}_j}$.

2. $p_i(x_i) \geq \underline{s}^v_{-i}(x_i)$ or $p_i(x_i) \geq \bar{s}^v_{-i}(x_i)$, respectively.

Proof. Suppose that $0 \geq \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}^v_{-i}, x_i)$ for some $a_i \in \tilde{a}_i^+$, the proof of the other case being identical. For the first part, suppose that $\forall j \neq i$, $B(x_i, 2v) \subseteq P^v_{j, \tilde{a}_j}$. Then after receiving x_i, player i knows that $\widetilde{a}_{-i}$ is played for sure, and thus we have

$$0 \geq \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}^v_{-i}, x_i) =$$

$$\int_{\mathbb{R}^{|g^i|-1}} \triangle u_i(\tilde{a}_i, a_i, \underline{s}^v_{-i}, x_i) d\mu^*_{F_i}(x_{-i}|x_i, v) = \triangle u_i(\tilde{a}_i, a_i, \widetilde{a}_{-i}, x_i) > 0$$

a contradiction.

For the second part, we have that

$$0 \geq \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}^v_{-i}, x_i) = \int_{\mathbb{R}^{|g^i|-1}} \triangle u_i(\tilde{a}_i, a_i, \underline{s}^v_{-i}, x_i) d\mu^*_{F_i}(x_{-i}|x_i, v)$$

$$= \sum_{a_{-i}} \triangle u_i(\tilde{a}_i, a_i, a_{-i}, x_i) \left(\int_{\mathbb{R}^{|g^i|-1}} (1_{\{\underline{s}^v_{-i}=a_{-i}\}}) d\mu^*_{F_i}(x_{-i}|x_i, v) \right)$$

If we define $\lambda'_i \in \triangle(A_{-i})$ by $\lambda'_i(a_{-i}) = \int_{\mathbb{R}^{|g_i|-1}} (1_{\{\underline{s}^v_{-i}=a_{-i}\}}) d\mu^*_{F_i}(x_{-i}|x_i)$, we have that $0 \geq l_i(a_i, \lambda'_i, x_i)$.

By the second half of Lemma 10 in the Appendix,[10] $0 \geq l_i(a_i, \lambda'_i, x_i) \Rightarrow p_i(a_i, x) \geq \lambda'_i(\tilde{a}_{-i})$. There-

---

[10] $0 \geq \triangle \pi_{x_{-i}}(\tilde{a}_i, a_i, \underline{s}^v_{-i}, x_i)$ implies $x_i \notin D^{\tilde{a}_i, a_i}$ and hence Lemma 10 can be applied.

fore,

$$p_i(x) = \max_{a_i}(p_i(a_i, x_i)) \geq p_i(a_i, x_i) \geq$$

$$\lambda_i'(\tilde{a}_{-i}) = \int_{\mathbb{R}^{|g^i|-1}} (1_{\{\underline{s}^v_{-i} = \tilde{a}_{-i}\}}) d\mu^*_{F_i}(x_{-i}|x_i, v) = \underline{s}^v_{-i}(x_i)$$

completing the proof.

□

Again if we let $\vec{x}_{-i}$ denote the x_{-i} whose elements are given by x_i, then we have the following generalization of Lemma 5:

Lemma 6. *Suppose the conditions of Theorem 7 hold. Then for each $i \in I$ and $x_i, x_j \in I$, we have:*

- For $N_g = 2$,

$$F_i^* \left(x_j | x_i, v \right) + F_i^* \left(x_i | x_j, v \right) = 1$$

- For $N_g \geq 2$,

$$F_i^* \left(\vec{x}_{-i} | x_{g_i^i}, v \right) \geq \frac{1}{|g^i|}$$

Proof. Identical to the proof in Lemma 5, but with the $\varphi^v_{g^i_j}$'s in place of the φ^v_i's.

□

Theorem 7 is now proven:

Proof. (Of Theorem 7) Suppose $\underline{x} \in I \cap D^{\tilde{a}}$. Let $v' < \frac{\alpha_I}{2}$ be such that $B(\underline{x}, 2v') \subseteq I \cap D^{\tilde{a}}$, v'' satisfy the conditions of Corollary 3, and $\bar{v} < min(\frac{v'}{2}, \frac{v''}{2})$. For a contradiction, let $v \in (0, \bar{v}]$ violate the hypothesis of Theorem 7, so that for some serially undominated strategy s and some $\tilde{x} \in I$, $s(\tilde{x}) \neq \tilde{a}$.

52

Since $\tilde{x} \in I/D^{\tilde{a}}$ and $D^{\tilde{a}}$ is an interval, we can assume WLOG that $\tilde{x}$ lies to the right of $D^{\tilde{a}}$.[11] For each player j, define $x_j^v = sup(x \,|\, [\underline{x}, x) \subseteq \mathscr{P}_{j,\tilde{a}_j}^v)$, noting that since $\bar{v} > v$ and $\underline{x}$ satisfy the conditions of Lemma 4 (4), these are well-defined. From now on, for each player i, let $x_{g_m^i}^v = \min_{j \in g_m^i} \left(x_j^v \right)$.

Now, for each finite x_j^v, we must have that $p_j(x_j^v) \geq \bar{s}_{-j}^v(x_j^v)$ or $p_j(x_j^v) \geq \underline{s}_{-j}^v(x_j^v)$. To see this, we show that there is some $a_j \in \tilde{a}_j^+$ such that $0 \geq \triangle\pi_{x_{-j}}(\tilde{a}_j, a_j, \underline{s}_{-j}^v, x_j^v)$ or $a_j \in \tilde{a}_j^-$ such that $0 \geq \triangle\pi_{x_{-j}}(\tilde{a}_j, a_j, \bar{s}_{-j}^v, x_j^v)$. If this is not true, then by continuity in x, if $\triangle\pi_{x_{-j}}(\tilde{a}_j, a_j, \underline{s}_{-j}^v, x_j^v) > 0$ for all $a_j \in \tilde{a}_j^+$ and $\triangle\pi_{x_{-j}}(\tilde{a}_j, a_j, \bar{s}_{-j}^v, x_j^v) > 0$ for all $a_j \in \tilde{a}_j^-$, there exists some $\bar{\varepsilon} > 0$ such that for each $\varepsilon \in (0, \bar{\varepsilon}]$, we again have that $\triangle\pi_{x_{-j}}(\tilde{a}_j, a_j, \underline{s}_{-j}^v, x_j^v + \varepsilon) > 0$ and $\triangle\pi_{x_{-j}}(\tilde{a}_j, a_j, \bar{s}_{-j}^v, x_j^v + \varepsilon) > 0$ for all $a_j \in \tilde{a}_j^+$ and $a_j \in \tilde{a}_j^-$, respectively. By Lemma 4(3), $\mathscr{P}_{j,\tilde{a}_j}^v$ could then be extended to the right of x_j^v, contradicting the definition of x_j^v. Thus, WLOG suppose that there is some $a_j \in \tilde{a}_j^-$ such that $0 \geq \triangle\pi_{x_{-j}}(\tilde{a}_i, a_i, \bar{s}_j^v, x_j^v)$.

By Proposition 5, we have that $p_i(x_j^v) \geq \bar{s}_{-j}^v(x_j^v)$.

Let x_l^v be the smallest of the x_j^v, which must be finite by hypothesis.[12] Since by the above discussion we may assume $0 \geq \triangle\pi_{x_{-l}}(\tilde{a}_l, a_l, \bar{s}_{-l}^v, x_l^v)$ for some $a_l \in \tilde{a}_l^-$, then by Proposition 5(1), the set

$$L_l^v \equiv \left\{ j \neq l \,|\, \exists x \in B(x_l^v, 2v)/\mathscr{P}_{j,\tilde{a}_j}^v \right\}$$

is non-empty. Let $\hat{x} \equiv \inf_{j \in L_l^v} \left(x \in B(x_l^v, 2v)/\mathscr{P}_{j,\tilde{a}_j}^v \right)$, and $j \in L_l^v$ be such that $\hat{x} = \inf \left(x \in B(x_l^v, 2v)/\mathscr{P}_{j,\tilde{a}_j}^v \right)$. Then $\hat{x}$ satisfies the following properties:

1. $\hat{x} = x_j^v$: Suppose $x_j^v > \hat{x}$. By the definition of $\hat{x}$, for every $\varepsilon > 0$ there is some $x \notin \mathscr{P}_{j,\tilde{a}_j}^v$ such that $x < \hat{x} + \varepsilon$. Setting $\varepsilon = x_j^v - \hat{x} > 0$, we have the existence of some $x \notin \mathscr{P}_{j,\tilde{a}_j}^v$ such that $x < \hat{x} + (x_j^v - \hat{x})$, or $x < x_j^v$, contradicting the definition of x_j^v. Thus $\hat{x} \geq x_j^v$. If $\hat{x} > x_j^v$, then

---

[11]Or, $\forall x \in D^{\tilde{a}}, \tilde{x} > x$. In the case of an upper dominance region, simply consider the parameter space $\tilde{X} = -X$ and $\tilde{u}_i = u_i(\cdot, -x)$.

[12]Since the set of serially undominated strategies for player j is contained in $\mathscr{S}_j^v$, by hypothesis there exists some $s_j \in \mathscr{S}_j^v$ such that $s_j(\tilde{x}) \neq \tilde{a}_j$. This implies that $\tilde{x} \notin \mathscr{P}_{j,\tilde{a}_j}^v$, and hence $x_j^v \leq \tilde{x} < \infty$.

since $\hat{x} \in [x_l^v, x_l^v + 2v)$, we have that $x_l^v + 2v > \hat{x} > x_j^v \geq x_l^v$. By definition of x_j^v there must be some $x' \in [x_j^v, \hat{x})$ such that $x' \notin \mathscr{P}_{j,\tilde{a}_j}^v$, contradicting the definition of $\hat{x}$, giving the result.

2. $p_j(\hat{x}) \geq \bar{s}_{-j}^v(\hat{x})$ or $p_j(\hat{x}) \geq \underline{s}_{-j}^v(\hat{x})$: Since $\hat{x} = x_j^v$, this follows from the discussion above.

3. $\forall m \neq l$, $\hat{x} \leq x_m^v$: Suppose for some m we have $\hat{x} > x_m^v$. The contradiction is the same as in the second half of part 1.

Denote $\hat{x} = x_j^v$, and note that since $x_l^v \in [\underline{x}, \tilde{x}]$ and $x_j^v \in B(x_l^v, 2v) \subseteq I \subseteq int(P)$, Corollary 3 may be applied to x_l^v and x_j^v. Since x_j^v is also finite, assume that $p_j(x_j^v) \geq \bar{s}_{-j}^v(x_j^v)$ WLOG for the remainder of the proof. We have that

$$F_l^* \left(\left(x_{g_m^l}^v \right)_{g_m^l \in g_{-l}^l} \middle| x_l^v, v \right) = \int_{\mathbb{R}^{|g^l|-1}} 1_{\left\{ x_{-l} \leq \left(x_{g_m^l}^v \right)_{g_m^l \in g_{-l}^l} \right\}} d\mu_{F_l}^*(x_{-l}|x_l^v, v) \leq$$

$$\int_{\mathbb{R}^{|g^l|-1}} 1_{\{\bar{s}_{-l}^v = \tilde{a}_{-l}\}} d\mu_{F_l}^*(x_{-l}|x_l^v, v) = \bar{s}_{-l}^v(x_l^v)$$

and likewise $F_j^* \left(\left(x_{g_m^j}^v \right)_{g_m^j \in g_{-j}^j} \middle| x_j^v, v \right) \leq \bar{s}_{-j}^v(x_j^v)$. We now consider two cases:

Case 1: $N_g = 2$: By the above, we have that

$$p_l(x_l^v) + p_j(x_j^v) \geq \bar{s}_{-l}^v(x_l^v) + \bar{s}_{-j}^v(x_j^v) \geq F_l^* \left(x_j^v | x_l^v, v \right) + F_j^* \left(x_l^v | x_j^v, v \right) \geq 1$$

contradicting the fact that since v satisfies the conditions of Corollary 3, we must have

$$p_l(x_l^v) + p_j(x_j^v) < 1$$

54

Case 2: $N_g > 2$: Again by the above argument, we have that

$$p_l(x_l^v) \geq \bar{s}_{-l}^v(x_l^v) \geq F_l^* \left(\left(x_{g_m^l}^v \right)_{g_m^l \in g_{-l}^l} |x_l^v, v \right) \overset{(\star)}{\geq} F_l^* \left(\vec{x}_{-l}^v |x_l^v, v \right) \geq \frac{1}{|g^l|}$$

where inequality $(\star)$ follows from the fact that x_l^v is the lowest of all x_j^v. But this contradicts the fact that we must have $p_l(x_l^v) < \frac{1}{N_g} \leq \frac{1}{|g^l|}$, completing the proof.

$\square$

Note that Theorem 6 follows immediately, which has the same set-up as Theorem 7 but with the "trivial" partitioning consisting of singletons. We now consider an example.

Example 7. Consider the following version of the Brander-Spencer model, where a foreign firm (F_f) decides whether to remain in (R) or leave (L) a market consisting of two domestic firms (F_d), who must decide whether to enter (E) or stay out (S) of the market. The domestic firms receives a government subsidy $s \geq 0$ whereas the foreign firm does not. Suppose we have a simplified payoff matrix given by the following:

		R			F_f		L	
		F_d					F_d	
		E	S				E	S
F_d	E	$-3+s, -3+s, -3$	$2+s, 0, 2$		F_d	E	$2+s, 2+s, 0$	$3+s, 0, 0$
	S	$0, 2+s, 2$	$0, 0, 3$			S	$0, 3+s, 0$	$0, 0, 0$

This is a game of strategic substitutes parametrized by $s \geq 0$. For $s \in [0, 3]$, the Nash equilibria are given by (E, S, R), (S, E, R), and (E, E, L), and for $s > 3$, (E, E, L) is a strict Nash equilibrium, where E is strictly dominant for the two domestic firms. Note that the parameter restriction $s \geq 0$ prevents us from establishing a lower dominance region, thus traditional global games methods cannot be applied. In order to apply Theorem 6, we require that

$$p_i(s) < \frac{1}{3}, \forall i \in I$$

Calculating the p_i functions gives

$$p_1(s) = p_2(s) = \begin{cases} \frac{3-s}{5} & 0 \le s < 3 \\ \\ 0 & s \ge 3 \end{cases}$$

$$p_3(s) = \frac{1}{2}, \ \forall s.$$

Therefore, Theorem 6 is also violated by p_3. However, if each player i has stereotypical beliefs so that each g^i_{-i} is a singleton, then $N_g = 2$. According to Theorem 7, it is easy to check that $p_i(s) + p_j(s) < 1$ holds for all $s > \frac{1}{2}$, and therefore multiplicity can be resolved on $(\frac{1}{2}, 3]$.

2.5   Common Valuations and Arbitrary Prior

In this section our main results are extended beyond the simple global game $G^*(v)$ to the original global game $G(v)$ with common valuations and an arbitrary prior. In order to do so, a useful variable transformation is used to express expected utility over normalized signal differences z_{-i} rather than the signals of the opponents themselves. For any player i and signal x_i, denote by $\overrightarrow{x_i}$ the $(|g^i| - 1) \times 1$ vector whose elements are given by x_i, and define $g_{x_i} : \mathbb{R}^{|g^i|-1} \to \mathbb{R}^{|g^i|-1}$ by

$$g_{x_i}(x_{-i}) = \left(\frac{x_{-i} - \overrightarrow{x_i}}{v} \right)$$

Let $Z_{x_i} = g_{x_i}(\mathbb{R}^{|g^i|-1})$ denote the set of all normalized signal differences about x_i. After receiving signal x_i, player i's beliefs over $Z_{x_i} \times \mathbb{R}$ are given by $\tilde{\mu}_i(z_{-i}, x | x_i, z) = \mu_{F_i}(g_{x_i}^{-1}(z_{-i}), x | x_i, v)$.

Applying a change of variables, player i's expected utility after receiving signal x_i against strategy s_{-i} is given by

$$\pi_{x,x_{-i}}(a_i, s_{-i}, x_i) = \int\limits_{\mathbb{R}^{|g^i|-1}} \int\limits_{\mathbb{R}} u_i(a_i, s_{-i}(x_{-i}), x) d\mu_{F_i}(x_{-i}, x | x_i, v) =$$

$$\int\limits_{Z_{x_i}} \int\limits_{\mathbb{R}} u_i(a_i, s_{-i}(\overrightarrow{x_i} + vz_{-i}), x) d\tilde{\mu}_i(z_{-i}, x | x_i, v)$$

We can make the same transformation in the simple global game $G^*(v)$ by similarly defining $\tilde{\mu}_i^*(z_{-i}|x_i, v) \equiv \mu_{F_i}^*(g^{-1}(z_{-i})|x_i, v)$. The connection between the games $G(v)$ and $G^*(v)$ will be made through an intermediate global game, $G^{**}(v)$, defined below.

Definition 15. Let $G(v)$ be a global game as in Definition 10, but alternatively with beliefs $\tilde{\mu}_i(z_{-i}, x|x_i, v)$ as defined above for each player i and each signal x_i. Define the intermediate game $G^{**}(v)$ as a private valuations global game with payoffs as in $G(v)$ but with beliefs over Z_{x_i} defined as $\forall i \in I$, $\forall x_i$,

$$\tilde{\mu}_i^{**}(z_{-i}|x_i, v) = marg_{Z_{x_i}}(\tilde{\mu}_i(z_{-i}, x|x_i, v))$$

The next Lemma makes a connection between the two games $G^*(v)$ and $G^{**}(v)$. That is, as $v \to 0$, beliefs in $G^{**}(v)$ approximate those in $G^*(v)$.

Lemma 7. *Let $G(v)$ be a global game. Then as $v \to 0$, the beliefs $\tilde{\mu}_i^{**}(\cdot|x_i, v)$ in $G^{**}(v)$ converge to those beliefs $\tilde{\mu}_i^*(\cdot|x_i, v)$ in $G^*(v)$.*[13]

---

[13]Convergence is in the following sense: For all $\varepsilon > 0$, and compact interval $B = [b_0, b_1]$ such that $[b_0 - 2v, b_1 + 2v]$

Proof. See Lemma A.2 in FMP.

$\square$

From here on, we will assume without mention that $\forall i, \forall a \in A$, $u_i(a, \cdot)$ is Lipschitz continuous with respect to x. Taking the maximum of all Lipschitz constants guarantees some $K > 0$ such that $\forall i, \forall a \in A$,

$$d(x, y) < v \Rightarrow |u_i(a, x) - u_i(a, y)| \leq Kv$$

The last aspect in our approximation lies in the following definition.

Definition 16. Let $G(v)$ be a global game, and let $\varepsilon \geq 0$. Define the following:

1. $\forall i \in I, \forall x_i \in \mathbb{R}, \forall M \subseteq S_{-i}$, the set of $\varepsilon-$undominated responses for Player i given x_i is

$$UR_{x_i}(M, \varepsilon) = \{a_i \in A_i | \forall a_i' \in A_i, \exists s_{-i} \in M, \triangle \pi_{x, z_{-i}}(a_i, a_i' s_{-i}, x_i) \geq -\varepsilon\}$$

2. $\forall i \in I, \forall M \subseteq S_{-i}$, the set of $\varepsilon-$undominated responses for Player i is

$$UR_i(M, \varepsilon) = \{s_i \in S_i | \forall x_i \in \mathbb{R}, s_i(x_i) \in UR_{x_i}(M, \varepsilon)\}$$

3. $\forall M \subseteq S$, the set of $\varepsilon-$undominated responses is

$$UR(M, \varepsilon) = \prod_{i \in I} UR_i(M_{-i}, \varepsilon)$$

---

lies in the interior of the support of f, there exists a $\tilde{v} > 0$ such that for all $x_i \in B$, and all $v \in (0, \tilde{v}]$,

$$\tilde{\mu}_i^{**}(\cdot | x_i, v) \in \left\{ \mu \in \triangle(Z_{x_i}) | \sup_{S \subseteq Z_{x_i}} (|\mu(S) - \tilde{\mu}_i^*(S | x_i, v)|) \leq \varepsilon \right\}$$

58

By defining $S_0 = S$, and $\forall n \geq 0$, $S_{n+1}(\varepsilon) = UR(S_n, \varepsilon)$, the set of $\varepsilon-$serially undominated strate-gies is given by $SU(G(v), \varepsilon) = \underset{n \geq 0}{\cap} S_n(\varepsilon)$. It is easy to verify that for all $\varepsilon' \geq \varepsilon$, $SU(G(v), \varepsilon) \subseteq SU(G(v), \varepsilon')$. Thus, since $\varepsilon = 0$ corresponds to the set of serially undominated strategies $SU(G(v), 0)$, which is non-empty, we have that for all $\varepsilon \geq 0$, $SU(G(v), \varepsilon)$ is non-empty as well. We can like-wise recast these definitions in terms of an intermediate global game $G^{**}(v)$ by replacing the $UR_i(M, \varepsilon)$'s with $UR_i^{**}(M, \varepsilon)$, $S_n(\varepsilon)$'s with $S_n^{**}(\varepsilon)$, and $SU(G(v), \varepsilon)$ with $SU(G^{**}(v), \varepsilon)$. We then have the following connection between $G(v)$ and $G^{**}(v)$.

Theorem 8. *Let $G(v)$ be a global game and $G^{**}(v)$ the corresponding intermediate game. Then $\forall \varepsilon > 0$, $\exists v(\varepsilon) > 0$, $\forall v \in (0, v(\varepsilon)]$,*

$$SU(G(v), 0) \subseteq SU(G^{**}(v), \varepsilon)$$

Proof. Appendix.

$\square$

Theorem 8 says that $SU(G(v), 0)$ is approximatelly contained in $SU(G^{**}(v), 0)$, which itself approximates $SU(G^*(v), 0)$ by Lemma 4, as as $v \to 0$. In this sense, for v small, $SU(G(v), 0)$ becomes consistent with the global games solutions offered by Theorem 6 and Theorem 7.

Chapter 3

On the Learning and Stability of Mixed Strategy Nash Equilibria in Games of Strategic Substitutes

3.1 Introduction

Many positive results have been established in the literature on games of strategic substitutes (GSS) in terms of the characterization of solution sets, adaptive learning processes, and comparative statics properties.[1] The analysis of this wide class of games, however, has concentrated mainly on situations where players are assumed to play pure strategies only, and although it is well known that such games need not exhibit pure strategy Nash equilibria (PSNE), the role of mixed strategies has largely been ignored. It is therefore important to ask under what conditions players may find it optimal to randomize over their set of actions, and if mixed strategy Nash equilibria (MSNE) offer good long-run predictions of behavior. By drawing on a connection in GSS between learning in repeated play and rationalizability, the first part of this question is answered by determining a bound for the support of any such mixed behavior. As a consequence, a new characterization

---

[1] See Dubey, Haimanko, and Zapechelnyuk (2006), and Roy and Sabarwal (2008, 2010, 2012), for example.

of global stability and a sufficient condition for the fictitious play property in this class of games is obtained. The second part of this question confirms that MSNE do not generally offer good predictions by showing that they are unstable under a range of widely-used learning procedures in a repeated games framework.

The validity of MSNE as an equilibrium prediction has always been a topic of debate in economics. The classical argument against them is as follows: If opponents are behaving in such a way as to make a player indifferent between a subset of her actions, why would randomizing be preferred to simply choosing a pure strategy best response? One response to this argument has been by way of Harsayni's Purification Theorem, which proves that if players privately observe a sequence of i.i.d. random shocks to their payoffs, then a mixed equilibrium emerges in the resulting game of incomplete information which approximates the original mixed equilibrium.[2] More recent studies have asked whether, if randomizing behavior is to be understood in the framework of players committing to a distribution over their actions when an underlying game is repeated over time, players can eventually learn to play according to an equilibrium distribution. Work along these lines has been conducted in a variety of game-theoretic settings. Crawford (1985) shows that purely mixed strategy Nash equilibria are always unstable under gradient dynamics.[3] Fudenberg and Kreps (1993), Kaniovski and Young (1995), and Benaim and Hirsch (1997) study the convergence to mixed equilibria in 2×2 games and 3×2 games, whereas Ellison and Fudenberg (2000) study the stability of MSNE in 3×3 games. Hofbauer and Hopkins (2005) investigates such stability in 2-player, finite-action games under a smooth fictitious play learning process, and Benaim, Hofbauer, and Hopkins (2009) studies convergence in games whose Nash equilibria are mixed and unstable under fictitious play-like learning.

This chapter is most closely related to Echenique and Edlin (EE)(2004), which considers the stability of purely mixed strategy Nash equilibria in games of strategic complements (GSC) when the set of players is finite and action spaces are a complete lattice. The heart of the analysis lies in exploiting a complementarity between the order structure inherit in GSC and a quite general as-

---

[2]See Govindan, Reny, and Robson (2003) for a shorter and more general proof of this result.

[3]As opposed to the best-response dynamics studied here. See Jordan (1993) for a discussion.

sumption on how players update their beliefs, which includes Cournot and fictitious play learning. Specifically, if a player makes a small mistake in her beliefs about equilibrium behavior by shifting an arbitrarily small amount of probability towards the largest action in the support of opponents' MSNE profile, then this upward shift (in FOSD) of beliefs implies that she will best respond by playing a strategy higher than her equilibrium mixed strategy. A subsequent update in beliefs again results in even higher shift in beliefs, resulting in an even higher response. This pattern continues on indefinitely, so that intended play never returns to the original MSNE. A similar argument can be made when the underlying game is a GSS, as the next example illustrates:

Consider the following slight variation to the 3-player Dove-Hawk-Chicken game presented in Sabarwal and Roy (2010):

$$P3$$

		D			H	
		$P2$			$P2$	
	D	H		D	H
D	$1, 1, 1$	$\varepsilon, 2, 1$	D	$1, \varepsilon, 2$	$1, 1, \varepsilon$
H	$2, 1, \varepsilon$	$1, \varepsilon, 1$	H	$\varepsilon, 1, 1$	$0, 0, 0$

(with $P1$ labeling the rows on both matrices)

where $\varepsilon \in (0, \frac{1}{2})$. This is a GSS which has no PSNE. One would hope, therefore, that a MSNE would provide a good prediction of play. Calculating the best-response functions, we obtain

$$BR_1 = \begin{cases} D & \sigma_3(D) < \left(\frac{1-\varepsilon\sigma_2(D)}{2-\varepsilon}\right) \\ [0,1] & \sigma_3(D) = \left(\frac{1-\varepsilon\sigma_2(D)}{2-\varepsilon}\right) \\ H & \sigma_3(D) > \left(\frac{1-\varepsilon\sigma_2(D)}{2-\varepsilon}\right) \end{cases}, BR_2 = \begin{cases} D & \sigma_1(D) < \left(\frac{1-\varepsilon\sigma_3(D)}{2-\varepsilon}\right) \\ [0,1] & \sigma_1(D) = \left(\frac{1-\varepsilon\sigma_3(D)}{2-\varepsilon}\right) \\ H & \sigma_1(D) > \left(\frac{1-\varepsilon\sigma_3(D)}{2-\varepsilon}\right) \end{cases}, BR_3 = \begin{cases} D & \sigma_2(D) < \left(\frac{1-\varepsilon\sigma_1(D)}{2-\varepsilon}\right) \\ [0,1] & \sigma_2(D) = \left(\frac{1-\varepsilon\sigma_1(D)}{2-\varepsilon}\right) \\ H & \sigma_2(D) > \left(\frac{1-\varepsilon\sigma_1(D)}{2-\varepsilon}\right) \end{cases}$$

We see that when player i believes that opponents are playing $\hat{\sigma}_{-i} = ((\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}))$, then $\hat{\sigma}_i = (\frac{1}{2}, \frac{1}{2})$ is a best response. That is, $\hat{\sigma}$ defined by $\forall i, \hat{\sigma}_i = (\frac{1}{2}, \frac{1}{2})$ is a MSNE. Now suppose that players make a slight error in their judgments about the behavior of others, and that for $\alpha > 0$ small, player i believes that all other players j will play $\sigma_j = (\frac{1}{2} - \alpha, \frac{1}{2} + \alpha)$. Then in the first round, each player i best-responds uniquely by playing D, or $\sigma_i = (1, 0)$. If players are Cournot learners, so that in each successive round they best-respond only to the profile played in the previous round, then

in the second round, each player i best responds uniquely by playing H, or $\sigma_i = (0, 1)$. Continuing in this manner, we see that play therefore enters a cycle: (D, D, D), (H, H, H), (D, D, D), ...etc., and it never again becomes optimal to best respond by mixing evenly among the two actions. One immediate question arises when we consider the possibility that players are not purely myopic in their best responses, and ask what happens if they can anticipate this cyclic behavior to their advantage. Even if we assume that this happens after a number of rounds of play, we see that if play ever approaches the original MSNE, a small misspecification of beliefs will once again disrupt convergence.

This chapter is organized as follows: Section 2 and 3 provide the relevant definitions and the setup of the model, and introduce the assumptions about how players update their beliefs between periods of play. Section 4 establishes an interval containing the limit of all intended (mixed) play, giving new results on global convergence and the fictitious play property. In Section 5 it is shown that under a wide range of learning rules, truly mixed play over this interval can at best result in unstable equilibria.

3.2   Model and Assumptions

The standard lattice concepts are used throughout this chapter.[4] A game of strategic substitutes will be defined as follows:

Definition 17. A (strict) GSS $\Gamma = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ consists of the following elements:

1. I is a finite set of players, $I = \{1, 2, ..., N\}$.

2. Each player i has an action set denoted by A_i. Each A_i is assumed to be a complete lattice with ordering $\succeq_i$. A_i is endowed with the order interval topology, which is assumed to be Hausdorff, and the corresponding Borel sigma-algebra $\mathscr{F}_{A_i}$. Also, $A = \prod_{i \in I} A_i$ and $A_{-i} = \prod_{j \neq i} A_i$ are given the product topologies and corresponding Borel sigma-algebras. We will abuse

---

[4]For an overview, see Topkis (1998).

notation by letting $\succeq$ denote the ordering for any A_i, A_{-i}, and A in the respective context, where the ordering for product sets are understood to be product orders.

3. Each player i has a utility function given by $u_i : A \to \mathbb{R}$. We assume that

   (a) For each $a_{-i} \in A_{-i}$, each u_i is continuous, bounded, and supermodular in a_i

   (b) Each u_i satisfies (strict) decreasing differences in (a_i, a_{-i}).[5]

For each player i, we let $\triangle(A_i)$ denote the set of probability measures (mixed strategies) μ_i over A_i, where $supp(\mu_i)$ denotes the support of any μ_i,[6] and $\triangle(A) = \prod_{i \in I} \triangle(A_i)$ denotes the set of all mixed strategy profiles. Player i's beliefs over the actions of her opponents is given by a probability measure $\mu \in \triangle(A_{-i})$. We will also endow each $\triangle(A_i)$ or $\triangle(A_{-i})$ with the weak* topology for probability measures, or, in the case of finite actions, the topology on the probability simplex as a subset of Euclidean space.

When player i holds belief $\mu \in \triangle(A_{-i})$ over opposing actions, the expected utility from playing action $a_i \in A_i$ is given by

$$U_i(a_i, \mu) = \int_{A_{-i}} u_i(a_i, a_{-i})d\mu$$

Player i's best response correspondence $BR_i : \triangle(A_{-i}) \twoheadrightarrow A_i$ is then given by

$$BR_i(\mu) = \underset{a_i \in A_i}{argmax}(U_i(a_i, \mu))$$

Using this notation, we have the following definition.

Definition 18. Let $\Gamma = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ be a GSS. Then $\sigma \in \triangle(A)$ is a properly mixed strategy Nash equilibrium (PMSNE) if the following hold:

1. $\forall i \in I$, $\forall a_i \in A_i$, $a_i \in supp(\sigma_i)$ only if $a_i \in BR_i(\sigma_{-i})$.

2. There exist at least two players i, $j \in I$ such that $|supp(\sigma_k)| > 1$, $k = i, j$.

---

[5]For all a_i', $a_i \in A_i$ such that $a_i' > a_i$, and all a_{-i}', $a_{-i} \in A_{-i}$ such that $a_{-i}' \geq (>)a_{-i}$, we have that $u_i(a_i, a_{-i}') - u_i(a_i', a_{-i}') \geq (>)u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})$.

[6]As is standard, $supp(\mu_i) = \cap \{E \,|\, E \, closed, \, \mu_i(E) = 1\}$

3.2.1 Repeated Games Framework

The model will consist of a fixed GSS Γ, which, starting at a given reference period t_0, is repeated for $t = t_0, t_0 + 1, t_0 + 2, ...,$ *etc.* To this end, each player i is endowed with some probability space $(\Omega_i, \mathscr{F}_i, m_i)$. In each period $t \geq t_0 + 1$, a payoff-irrelevant $\omega^t = (\omega_1^t, \omega_2^t, ..., \omega_N^t) \in \prod_{\forall i} \Omega_i$ is drawn, after which each player i privately observes ω_i^t and chooses an action a_i^t. This procedure produces for each t a **time$-t$ history profile** $h^t = (a^{t_0+1}, a^{t_0+2}..., a^t)$. We let H^t denote all possible time-t history profiles, where $H^{t_0} = \emptyset$, and denote $H = \overset{\infty}{\underset{t=t_0}{\cup}} H^t$. Each player i has **repeated game beliefs** $\mu_i : H \to \triangle(A_{-i})$ with the interpretation that, for each $h^t \in H$, $\mu_i(h^t)$ describes player i's beliefs about opponents' actions based on the current history of play. Letting $\mu(h^{t_0})$ denote arbitrary initial beliefs, the a^t are chosen in the following way:

Definition 19. A repeated games strategy for player i is a function $\varepsilon_i : \Omega_i \times H \to A_i$ such that $\forall \omega_i \in \Omega_i$, $\forall t \geq t_0$,

$$\varepsilon_i(\omega_i, h^t) \in BR_i(\mu_i(h^t))$$

where, for each h^t, $\varepsilon_i(\cdot, h^t) \to A_i$ is measurable. The interpretation is that, after viewing a history h^t and formed beliefs $\mu_i(h^t)$, a time$-t+1$ signal ω_i^{t+1} is realized, and player i chooses some best response $a_i^{t+1} = \varepsilon_i(\omega_i^{t+1}, h^t) \in BR_i(\mu_i(h^t))$. Then, after viewing a_{-i}^{t+1}, new period $t+1$ beliefs $\mu_i(h^{t+1})$ are formed, and so on. Note that each ω_i is payoff-irrelevant, so that $\varepsilon_i(\cdot, h^t) : \Omega_i \to A_i$ serves only as a randomization device. Thus, for any history h^t, we will define

$$m_i^{t+1} = m_i\left(\varepsilon_i^{-1}(\cdot, h^t)\right) : \mathscr{F}_{A_i} \to [0, 1]$$

as **time$-t+1$ intended play**, or mixed strategy best response, from which some a_i^{t+1} in the support is chosen after ω_i^{t+1} is realized. Finally, if $\varepsilon \equiv (\varepsilon_i)_{i=1}^N$ is a collection of repeated games strategies for each player, and $\mu \equiv (\mu_i)_{i=1}^N$ is a collection of repeated games beliefs, then we call $(\varepsilon, \mu, \mu(h^{t_0}))$ a **system of behavior and beliefs**.

3.2.2 The First-Order Stochastic Dominance Order

Our results will rely on lattice programming techniques applied to the set of mixed strategies. Suppose that $(X, \succeq)$ is a poset. A subset $E \subseteq X$ is called **increasing** if $\forall x \in E$, $y \in X$ and $y \succeq x$ implies that $y \in E$. If $\triangle(X)$ is the set of probability measures on X such that $\mu', \mu \in \triangle(X)$, then we say that μ' **first-order stochastically dominates** μ, written $\mu' \succeq_F \mu$, if, for every increasing set $E \subseteq X$, $\mu'(E) \geq \mu(E)$. $\mu' \succ_F \mu$ holds if $\mu' \succeq_F \mu$ and for some $E \subseteq X$ increasing, $\mu'(E) > \mu(E)$. Note that $\mu' \succeq_F \mu$ is equivalent to saying that for all $f : X \to \mathbb{R}$ increasing and integrable with respect to μ' and μ, we have

$$\int_X f d\mu' \geq \int_X f d\mu$$

It is easy to verify that for any $\triangle(X)$, $(\triangle(X), \succeq_F)$ forms a poset. Lastly, for two sets $A, B \subseteq X$, A is said to dominate B in the **strong set order**, written $A \succeq_S B$, if $\forall x \in A$, $\forall y \in B$, $x \vee y \in A$ and $x \wedge y \in B$.

As is shown in Roy and Sabarwal (2012), the process of iterated deletion of strictly dominated strategies in a GSS leads to the existence of upper and lower serially undominated strategies, denoted $\bar{a}$ and $\underline{a}$, respectively. We then have the following:

Proposition 6. *Let $\Gamma = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ be a GSS. Then for each player $i \in I$,*

1. *$\forall \mu \in \triangle(A_{-i})$, $\wedge BR_i(\mu) \in BR_i(\mu)$, and $\vee BR_i(\mu) \in BR_i(\mu)$.*

2. *If $\mu', \mu \in \triangle(A_{-i})$ are such that $\mu' \succeq_F \mu$, then*

$$BR_i(\mu) \geq_S BR_i(\mu')$$

3. *Let Γ be a strict GSS, and $\hat{\sigma} = (\hat{\sigma}_i, \hat{\sigma}_{-i})$ a PMSNE. For any $\varepsilon \in (0, 1]$, define*

$$\mu^\varepsilon = (1 - \varepsilon)\hat{\sigma}_{-i} + \varepsilon 1_{\{a_{-i}\}}$$

If $a_{-i} = \underline{a}_{-i}$, then

$$\wedge BR_i(\mu^{\varepsilon}) \succeq \vee BR_i(\hat{\sigma}_{-i})$$

Similarly, if $a_{-i} = \bar{a}_{-i}$,

$$\wedge BR_i(\hat{\sigma}_{-i}) \succeq \vee BR_i(\mu^{\varepsilon})$$

Proof. For the first claim, since each u_i is continuous, bounded, and supermodular in A_i (a property which is preserved under integration), then for each $\mu \in \triangle(A_{-i})$, $U_i(,\mu) : A_i \to \mathbb{R}$ satisfies supermodularity and upper semi-continuity. Since A_i is a complete lattice, then by Milgrom and Shannon (1994),

$$BR_i(\mu_i) = \underset{a_i \in A_i}{argmax}\,(U_i(a_i, \mu))$$

is a non-empty, complete lattice.

Because $(\triangle(A_{-i}), \succeq_F)$ is a poset, Claim 2 will follow from Sabarwal and Roy (2010) by showing that U_i satisfies decreasing differences in (a_i, μ_i). To that end, let $a_i' > a_i$ and $\mu' \succeq_F \mu$. Then $l : A_{-i} \to \mathbb{R}$ defined by

$$l_i(a_{-i}) = u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})$$

is increasing in a_{-i} by decreasing differences. Therefore

$$U_i(a_i, \mu) - U_i(a_i', \mu) = \int_{A_{-i}} l(a_{-i}) d\mu$$

$$\leq \int_{A_{-i}} l(a_{-i}) d\mu' = U(a_i, \mu') - U_i(a_i', \mu')$$

giving decreasing differences, where the third inequality follows from first order stochastic dominance.

The third claim is shown in the Appendix. $\square$

3.3 Belief Formation

In this section we describe how players update their beliefs from one round of play to another. To establish notation, if $\mu_i(h^{t_0})$ is any initial belief, then $\mu_i(h^{t_0}, h^t)$ denotes the result of $\mu_i(h^{t_0})$ being updated according to some subsequent history h^t. Likewise, if h^t is any history of play, and x^{t+1}_{-i} any time$-t+1$ profile of opponents' actions, $\mu_i(h^t, x^{t+1}_{-i})$ represents $\mu_i(h^t)$ being updated according to the subsequent play x^{t+1}_{-i}.

Definition 20. Let $(\varepsilon, \mu, \mu(h^{t_0}))$ be a system of behavior and beliefs. Define $\underline{M}_i = \wedge supp(\mu_i(h^{t_0}))$ and $\overline{M}_i = \vee supp(\mu_i(h^{t_0}))$. Then

1. Beliefs are **monotone** if $\forall i \in I, \forall t \geq t_0$

    (a) $\mu_i(\underline{h}^{t_0}) \leq_F \mu_i(h^{t_0}) \leq_F \mu_i(\overline{h}^{t_0})$ and $\underline{h}^t \preceq h^t \preceq \overline{h}^t$ implies

    $$\mu_i(\underline{h}^{t_0}, \underline{h}^t) \leq_F \mu_i(h^{t_0}, h^t) \leq_F \mu_i(\overline{h}^{t_0}, \overline{h}^t)$$

    (b) $y^{t_0+1}_{-i} \preceq \underline{M}_i \preceq \overline{M}_i \preceq z^{t_0+1}_{-i}$ implies

    $$\mu_i(h^{t_0}, y^{t_0+1}_{-i}) \leq_F \mu_i(h^{t_0}) \leq_F \mu_i(h^{t_0}, z^{t_0+1}_{-i})$$

    And for $t > t_0$, $y^{t+1}_{-i} \preceq (\wedge h^t_{-i}) \wedge \underline{M}_i \preceq (\vee h^t_{-i}) \vee \overline{M}_i \preceq z^{t+1}_{-i}$ implies

    $$\mu_i(h^t, y^{t+1}_{-i}) \leq_F \mu_i(h^t) \leq_F \mu_i(h^t, z^{t+1}_{-i})$$

2. Beliefs are **asymptotically empirical** if, whenever a sequence of play is convergent, or $a^t \to a$, then $\mu_i(h^t) \to 1_{\{a_{-i}\}}$.

Note that the above conditions are rather weak. 1 (a) requires that observing a higher (lower) history of actions results in higher (lower) beliefs in FOSD. 1 (b) requires that that at any time t, if next period play by opponents is higher (lower) than both anything played up to that point, as

68

well as the support of initial beliefs, then next period beliefs are higher (lower) in FOSD. These assumptions allow for a very wide range of updating rules. Consider, for example, the geometrically weighted beliefs of Benaïm, Hofbauer, and Hopkins (2009), which are updated according to, for each player i,

$$\mu_i(h^t) = (1 - \lambda_t)\mu_i(h^{t-1}) + \lambda_t 1_{\{a^t_{-i}\}}$$

where for each round t, $\lambda_t \in [0, 1]$. If $\lambda_t = 1$ for all t, then players exhibit "Cournot beliefs", where next period play is determined by best responding to the pure strategy play of the previous round. Alternatively, if we allow each player i's initial beliefs $\mu_i(h^{t_0})$ to be of the form $\mu_i(h^{t_0})(a_{-i}) = \frac{K_{a_{-i}}}{t_0}$, where the $(K_{a_{-i}})_{a_{-i} \in A_{-i}}$ describe player i's initial fictitious weight on opponents' actions, then setting $\lambda_t = \frac{1}{t+1}$ for each player gives us the fictitious play beliefs introduced by Brown (1951). That is, each player i's beliefs are updated according to the historical frequency, given by, for all $t \geq t_0$,

$$\mu_i(h^t) = \left(\frac{t}{t+1}\right)\mu_i(h^{t-1}) + \left(\frac{1}{t+1}\right)1_{\{a^t_{-i}\}}$$

Both of these models of learning have shown positive results in experimental settings.[7] Lemma 8 below states that geometrically weighted beliefs, which allow for any combination of these two models, satisfy the requirements of Definition 20.

Lemma 8. *Let $(\varepsilon, \mu, \mu(h^{t_0}))$ be a system of behavior and beliefs. If $\forall i \in I$, $\forall t \geq t_0$, $\mu_i(h^t)$ is updated according to geometrically weighted beliefs and $\lambda_t \in [\frac{1}{t+1}, 1]$, then beliefs are monotone and asymptotically empirical.*

One immediate question arises from the above formulation. Suppose that $\mu_i(h^{t_0})$ are player i's initial beliefs representing opponents' play in a properly mixed strategy Nash equilibrium. How then is it possible that she may be a Cournot learner when Cournot learning never lends itself to properly mixed beliefs? Certainly the fact that an individual may be a Cournot learner does not prevent her from holding more complex beliefs about opponents' actions. The resolution of this issue

---

[7]See Boylan and El-Gamal (1993), Cheung and Friedman (1997), Huck, Normann, Oechssler (2002), and Gerber (2006), for example.

comes by way of making the distinction between "stated beliefs" and "latent beliefs". Following Rutström and Wilcox (2009), stated beliefs are those beliefs that a player would consciously hold if they were elicited from her by a third party. Latent beliefs refer to the true but unobserved "beliefs in the head", which may be driven by emotional or automatic, subconscious processes, and are those which are consistent with a player's behavior.

3.4   Bounds on Learning

We now study the limits of intended play in GSS starting at any initial beliefs $\mu(h^{t_0})$. To do this, we draw on a connection between iterated deletion of strictly dominated strategies (IESDS) and learning in a repeated GSS. This allows us to establish an interval for the support of any mixed play in the limit, and give sufficient conditions for when this interval converges to a singleton, guaranteeing a PSNE. In order to proceed, we introduce a necessary definition.

Definition 21. The **best response dynamic** starting from $\wedge A$ and $\vee A$ are given by the sequences $(y^t)_{t=0}^{\infty}$ and $(z^t)_{t=0}^{\infty}$ defined as

1. $y^0 = \wedge A$ and $z^0 = \vee A$

2. $y^t = \begin{cases} \wedge BR(y^{t-1}), & t \text{ even} \\ \vee BR(y^{t-1}), & t \text{ odd} \end{cases}, z^t = \begin{cases} \vee BR(z^{t-1}), & t \text{ even} \\ \wedge BR(z^{t-1}), & t \text{ odd} \end{cases}$

3. The **lower mixture and upper mixtures** of $(y^t)_{t=0}^{\infty}$ and $(z^t)_{t=0}^{\infty}$ are the sequences $(\underline{x}^t)_{t=0}^{\infty}$ and $(\bar{x}^t)_{t=0}^{\infty}$ defined as

$$\underline{x}^t = \begin{cases} y^t, & t \text{ even} \\ z^t, & t \text{ odd} \end{cases}, \bar{x}^t = \begin{cases} z^t, & t \text{ even} \\ y^t, & t \text{ odd} \end{cases}$$

Proposition 7. Let $\Gamma = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ be a GSS, $(\vec{x}^t)_{t=0}^{\infty}$ and $(\underline{x}^t)_{t=0}^{\infty}$ be the upper and lower mixtures of the best response dynamics. Then the following are true:

1. $\vec{x}^t \to \bar{a}$, and $\underline{x}^t \to \underline{a}$, where $\bar{a}$ and $\underline{a}$ are the largest and the smallest serially undominated strategies of Γ, respectively.

2. If $\wedge BR_i$ and $\vee BR_i$ are continuous, then $\underline{a}_i = \vee BR_i(\bar{a}_{-i})$ and $\bar{a}_i = \wedge BR_i(\underline{a}_{-i})$.

Proof. See Sabarwal and Roy (2012). $\qquad\square$

We now come to the first of the main results. It states that under weak assumptions on how beliefs are updated, the evolution of intended play starting at any initial beliefs will eventually be contained within the interval determined by the upper and lower serially undominated strategies.

Theorem 9. *Let $\Gamma = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ be a GSS, and $(\varepsilon, \mu, \mu(h^{t_0}))$ a system of behavior and beliefs. If, for each i, $\wedge BR_i$ and $\vee BR_i$ are continuous, and beliefs are monotone and asymptotically empirical, then*[8]

$$\left[liminf \left(supp \left(m_i^{t+1} \right) \right), limsup \left(supp \left(m_i^{t+1} \right) \right) \right] \subseteq [\underline{a}_i, \bar{a}_i]$$

Proof. First, we define the histories $\underline{h}^t, \overline{h}^t$ by the following:

- $\mu(\underline{h}^{t_0}) = 1_{\{\wedge A\}}, \mu(\overline{h}^{t_0}) = 1_{\{\vee A\}}, \underline{h}^0 = \{\wedge A\}, \overline{h}^0 = \{\vee A\}$

- $\forall t \geq t_0, \underline{s}^{t+1} = \wedge BR(\mu(\overline{h}^t)), \overline{s}^{t+1} = \vee BR(\mu(\underline{h}^t)), \underline{h}^{t+1} = \{\underline{h}^t, \underline{s}^{t+1}\}, \overline{h}^t = \{\overline{h}^t, \overline{s}^{t+1}\}$

Let $(\omega^t)_{t=t_0+1}^{\infty} \subseteq \Omega^{\infty}$ be arbitrary. Note that we have that

$$\mu(\underline{h}^{t_0}) \leq_F \mu(h^{t_0}) \leq_F \mu(\overline{h}^{t_0})$$

---

[8]Recall that if $\{Z^t\}_{t=1}^{\infty}$ is a sequence of sets, then $liminf(Z^t) = \bigvee_{t \geq 1} \bigwedge_{m \geq t} Z^m$ and $limsup(Z^t) = \bigwedge_{t \geq 1} \bigvee_{m \geq t} Z^m$.

Proposition 6 gives

$$\underline{s}^{t_0+1} = \wedge BR(\mu(\overline{h}^{t_0})) \preceq BR(\mu(h^{t_0})) \preceq \vee BR(\mu(\underline{h}^{t_0})) = \overline{s}^{t_0+1}$$

or, since $a^{t_0+1} = \varepsilon(\omega^{t_0+1}, h^{t_0})$

$$\underline{s}^{t_0+1} \preceq a^{t_0+1} \preceq \overline{s}^{t_0+1}$$

We proceed by induction by supposing that for $t_0 + 1 \le l \le t$,

$$\underline{s}^l \preceq a^l \preceq \overline{s}^l$$

Note that by $\mu(\underline{h}^{t_0}) \le_F \mu(h^{t_0}) \le_F \mu(\overline{h}^{t_0})$ and by the induction hypothesis $\underline{h}^t \preceq h^t \preceq \overline{h}^t$, monotonicity of beliefs gives

$$\mu(\underline{h}^{t_0}, \underline{h}^t) \le_F \mu(h^{t_0}, h^t) \le_F \mu(\overline{h}^{t_0}, \overline{h}^t)$$

By Proposition 6, since $a^{t+1} = \varepsilon(\omega^{t+1}, h^t)$,

$$\underline{s}^{t+1} \preceq a^{t+1} \preceq \overline{s}^{t+1}$$

completing the induction step. Because $(\omega^t)_{t=t_0+1}^{\infty} \subseteq \Omega^{\infty}$ was arbitrary, it follows that for each player i,

$$supp\left(m_i^{t+1}\right) \subseteq \left[\underline{s}_i^{t+1}, \overline{s}_i^{t+1}\right]$$

It is now shown that the sequences $(\underline{s}^t)_{t=t_0+1}^{\infty}$ and $(\overline{s}^t)_{t=t_0+1}^{\infty}$ are monotone increasing and decreasing, respectively. For each player i, define $\overline{M}_i^l \equiv \vee supp\left(\mu_i(\underline{h}^{t_0})\right) = \wedge A_{-i}$ and $\underline{M}_i^u = \wedge supp\left(\mu_i(\overline{h}^{t_0})\right) = \vee A_{-i}$. Since for each i, $\overline{s}_{-i}^{t_0+1} \preceq \underline{M}_i^u$, then by monotone beliefs we have that $\mu_i(\overline{h}^{t_0}, \overline{s}_{-i}^{t_0+1}) \le_F \mu_i(\overline{h}^{t_0})$. By Proposition 6,

$$\underline{s}^{t_0+2} = \wedge BR(\mu(\overline{h}^0, \overline{s}^{t_0+1})) \succeq \wedge BR(\mu(\overline{h}^{t_0})) = \underline{s}^{t_0+1}$$

A similar argument shows that $\bar{s}^{t_0+1} \succeq \bar{s}^{t_0+2}$. Suppose that for $t_0 + 1 \leq l \leq t$, $(\underline{s}^l)_{l=t_0+1}^t$ and $(\bar{s}^l)_{l=t_0+1}^t$ are increasing and decreasing sequences, respectively. Again, since for each i, $\bar{s}_{-i}^t \preceq \left(\wedge \bar{h}_{-i}^{t-1}\right) \wedge \underline{M}_i^u$, monotonicity of beliefs gives $\mu_i(\bar{h}^{t-1}, \bar{s}_{-i}^t) \leq_F \mu_i(\bar{h}^{t-1})$. Thus

$$\underline{s}^{t+1} = \wedge BR(\mu(\bar{h}^{t-1}, \bar{s}^t)) \succeq \wedge BR(\mu(\bar{h}^{t-1})) = \underline{s}^t$$

Similarly, $\bar{s}^t \succeq \bar{s}^{t+1}$, establishing monotonicity. Therefore, we have that $\underline{s}^t \to \underline{s}$ and $\bar{s}^t \to \bar{s}$ for some $\underline{s}, \bar{s} \in A$.

By the continuity of $\wedge BR$ and $\vee BR$, and the fact that beliefs are asymptotically empirical, we have that

$$\underline{s} = \lim_{t\to\infty} \left(\wedge BR(\mu(\bar{h}^t))\right) = \wedge BR\left(\lim_{t\to\infty}\left(\mu(\bar{h}^t)\right)\right) = \wedge BR(\bar{s}) \in BR(\bar{s})$$

Likewise, $\bar{s} \in BR(\underline{s})$. Therefore, $\underline{s}$ and $\bar{s}$ are rationalizable strategies, and hence serially undominated. Because $\underline{a}$ and $\bar{a}$ are the smallest and largest serially undominated strategies, respectively, then for each player i,

$$\underline{a}_i \preceq \underline{s}_i = \bigvee_{z>t_0} \bigwedge_{t\geq z} \underline{s}^t \preceq \bigvee_{z\geq t_0} \bigwedge_{t\geq z} supp\left(m_i^t\right)$$

Similarly,

$$\bar{a}_i \succeq \bigwedge_{z>t_0} \bigvee_{t\geq z} supp\left(m_i^t\right)$$

giving the result. □

Say that a game Γ is **globally stable** if it contains a (mixed strategy) Nash equilibrium $\hat{\sigma}$ such that intended play converges to it under any system of behavior and beliefs. We then have the following immediate Corollary.

Corollary 4. *Under the conditions of Theorem 9, if Γ is dominance solvable, then Γ is globally stable.*

Milgrom and Roberts (1991) show that if beliefs are "adaptive" in the sense that the probability assigned to actions which are played only a finitely amount of times eventually goes to 0, then any dominance solvable game will be globally stable. However, the set of beliefs which are asymptotically empirical is strictly larger than those which are adaptive. Consider the learning model of Moreno and Walker (1991)(MW), where players best respond to the entire sample average,[9] where for any history h^t,

$$\mu_i(h^t) = \frac{1}{t+1} \sum_{j=0}^{t+1} a_{-i}^j$$

It is straightforward to show that these beliefs are asymptotically empirical, but are not adaptive: If play were to cycle between 0 and 1, for example, then beliefs would always be strictly contained in $(0,1)$, although nothing within this set is ever actually played. This argument can be extended to include beliefs whose distribution is centered around the sample average along with a sample variance, among others.

---

[9]Strategy spaces are assumed to be a convex subset of $\mathbb{R}^n$, and a_{-i}^0 can be represented by initial beliefs $1_{\{a_{-i}^0\}}$. Healy (2006) generalizes this model by considering truncated histories and a $k-$periods average.

Example 8. Consider the Type II duopoly of Cox and Walker (1998)(CW), where the marginal costs of each firm are sufficiently large so that the best response of firm 2 cuts firm 1's from above,[10] as shown below

Type II Duopoly
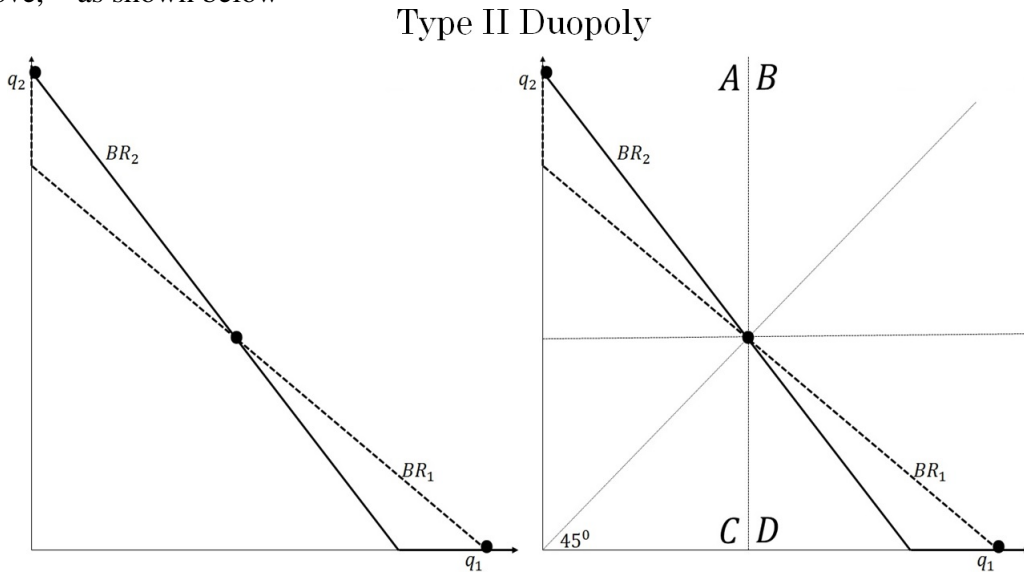
CW test in an experimental setting whether human subjects can learn to play the stable border equilibria in a repeated game. Intuition for the stability of such equilibria is clear: Small perturbations away from a boundary equilibrium must begin either in region A or D, in which case it is easy to verify that Cournot dynamics quickly lead back to the respective equilibrium. In order to obtain such baseline theoretical predictions as to when experimental behavior can be expected to converge under more complex learning rules, CW simulate various models of learning and conclude that fictitious play and MW learning seem to converge as well.

However, Theorem 9 and Corollary 4 immediately offer a much more general conclusion, which does not require ex-ante computer simulations: Since $BR(A) \subseteq A$ and $BR(D) \subseteq D$, regions A and D can themselves be viewed as sub-GSS. Ignoring the interior equilibrium, we see that the boundary equilibria are the dominance solvable solutions in their respective regions, and hence any beliefs which are monotone and asymptotically empirical, including Cournot, fictitious play, and MW, can be guaranteed to converge.

---

[10]We analyze a simplified version with continuous action spaces and symmetric payoffs.

3.5 Instability of PMSNE

In this section we study whether mixed behavior is a valid equilibrium prediction when intended play does not converge to a singleton as in Corollary 4. Section 3.5.1 deals with the case when players put sufficient weight on the play of the previous round, which, as is supported empirically by Cheung and Friedman (1997), is likely a reasonable assumption in more informative environments. This assumption will also always incorporate Cournot learning.

3.5.1 p-Instability

In this subsection we will assume that each A_i is finite and linearly ordered. Recall from Proposition 6 that $\bar{a}$ and $\underline{a}$ are the upper and lower serial undominated strategies, respectively.

Definition 22. For each player i, and $a_i \in A_i$, let $p_i^{\bar{a}_i}(a_i) \in [0, 1]$ be the smallest value such that for any $\lambda \in \triangle(A_{-i})$ satisfying $\lambda(\underline{a}_{-i}) \geq p_i^{\bar{a}_i}(a_i)$,

$$\sum_{a_{-i}} u_i(\bar{a}_i, a_{-i})\lambda(a_{-i}) \geq \sum_{a_{-i}} u_i(a_i, a_{-i})\lambda(a_{-i})$$

or

$$l^{\bar{a}_i}(a_i, \lambda) \equiv \sum_{a_{-i}} \triangle u_i(\bar{a}_i, a_i, a_{-i})\lambda(a_{-i}) \geq 0$$

We define $p_i^{\underline{a}_i}(a_i)$ similarly as the smallest probability that player i must see $\bar{a}_{-i}$ being played in order that $\underline{a}_i$ does better than a given a_i. We then have the following characterization of $p_i^{\bar{a}_i}(a_i)$ and $p_i^{\underline{a}_i}(a_i)$:

76

Proposition 8. *Let $\Gamma = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ be a strict GSS. Suppose there exists a PMSNE $\hat{\sigma}$ such that for each player $i \in I$, $\underline{a}_i, \overline{a}_i \in BR_i(\hat{\sigma}_{-i})$. Then, for each player i, and $a_i \neq \overline{a}_i, \underline{a}_i$, we have that*

$$
1. \ p_i^{\overline{a}_i}(a_i) =
\begin{cases}
0, & \overline{a}_i \ strictly \ dominates \ a_i \\[2ex]
\max\limits_{\substack{\lambda \in \triangle(A_{-i}) \\ l^{\overline{a}_i}(a_i, \lambda) = 0}} \left(\lambda(\underline{a}_{-i}) \right), & otherwise
\end{cases}
$$

$$
2. \ p_i^{\underline{a}_i}(a_i) =
\begin{cases}
0, & \underline{a}_i \ strictly \ dominates \ a_i \\[2ex]
\max\limits_{\substack{\lambda \in \triangle(A_{-i}) \\ l^{\underline{a}_i}(a_i, \lambda) = 0}} \left(\lambda(\overline{a}_{-i}) \right), & otherwise
\end{cases}
$$

3. $p^{\overline{a}_i} \equiv \max\limits_{a_i} \left(p^{\overline{a}_i}(a_i) \right), \ p^{\underline{a}_i} \equiv \max\limits_{a_i} \left(p^{\underline{a}_i}(a_i) \right) \in [0, 1)$

4. For all $\lambda \in \triangle(A_{-i})$,

$$
\mu(\underline{a}_{-i}) > p_i^{\overline{a}_i} \left(resp. \ \mu(\overline{a}_{-i}) > p_i^{\underline{a}_i} \right) \Rightarrow
$$

$$
BR_i(\mu) = \overline{a}_i \ (resp. = \underline{a}_i)
$$

Proof. Appendix. □

Proposition 8 generalizes the extent to which $\underline{a} \in BR(\overline{a})$ and $\overline{a} \in BR(\underline{a})$: If, for player i, $p_i^{\overline{a}_i} = 1$, then $\overline{a}_i$ is simply a best response to $\underline{a}_{-i}$. If $p_i^{\overline{a}_i} = 0$, then $\overline{a}_i$ is a strictly dominant action, and if $p_i^{\overline{a}_i} \in (0, 1)$, then $\overline{a}_i$ is a best response to any λ so long as $\lambda(\underline{a}_{-i}) \geq p_i^{\overline{a}_i}$ and strictly so if the inequality is strict.

The next definition formalizes a lower bound for the weight associated to previous action played for beliefs in a repeated game.

Definition 23. Let $(\varepsilon, \mu, \mu(h^{t_0}))$ be a system of behavior and beliefs, and $p = (p_i)_{i \in I} \in [0, 1]^N$. Then $(\varepsilon, \mu, \mu(h^{t_0}))$ is $p-$**consistent** if $\forall i \in I, \forall t \geq t_0 + 1, \mu_i(h^t)(a_{-i}^t) \geq p_i$.

It is easy to check that geometric beliefs satisfy $p-$consistency as long as $\lambda_t \geq p_i$ for each player $i \in I$ and $t \geq t_0 + 1$. In fact, Cournot beliefs are always $p-$consistent for any p. This notion,

together with Proposition 8, will be the driving force behind the instability of MSNE, which is defined below. For any $\hat{\mu} \in \triangle(A) \equiv \prod_{i \in I} \triangle(A_i)$, we will denote by $\hat{\mu}_{-i}$ the product measure of the $\hat{\mu}_j, j \neq i$.

Definition 24. Let $\Gamma = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ be a GSS, and $\hat{\mu} \in \triangle(A)$. Then $\hat{\mu}$ is $p-$**unstable** if, for each player i, and each open neighborhood V_i of $\hat{\mu}_{-i}$, there exists a $\mu' \in V_i$ such that for any system of p-consistent behavior and beliefs $(\varepsilon, \mu, \mu(h^{t_0}))$ such that $\mu_i(h^{t_0}) = \mu'$, we have that intended play remains outside of some neighborhood W of $\hat{\mu}$. That is, for each $t \geq t_0$, $(m_i^{t+1})_{i \in I} \notin W$.

We are now ready to state the first instability result.

Theorem 10. *Let $\Gamma = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ be a strict GSS, and $\hat{\sigma} \in \triangle(A)$ be a PMSNE. If $\forall i \in I$, $\underline{a}_i, \overline{a}_i \in BR_i(\hat{\sigma}_{-i})$, then there exists a $p \in [0, 1)^N$ such that $\hat{\sigma}$ is $p-$unstable.*

Proof. Let $i \in I$, and V_i be an open set containing $\hat{\sigma}_{-i}$. Since

$$\left\| \left((1 - \varepsilon)\hat{\sigma}_{-i} + \varepsilon 1_{\{\overline{a}_{-i}\}} \right) - \hat{\sigma}_{-i} \right\| \to 0$$

as $\varepsilon \to 0$, then we can follow the requirements of instability by choosing an $\varepsilon_i \in (0, 1)$ such that

$$\mu_i\left(h^{t_0}\right) \equiv (1 - \varepsilon_i)\hat{\sigma}_{-i} + \varepsilon_i 1_{\{\overline{a}_{-i}\}} \in V_i$$

We will now choose the set W according to instability in the following way: For each player i, let W_i be any open set containing $\hat{\sigma}_i$ if $\hat{\sigma}_i$ is degenerate. If $\hat{\sigma}_i$ is purely mixed, let d_i be such that

$$0 < d_i < min\left(\frac{\left\| \hat{\sigma}_i - 1_{\{\overline{a}_i\}} \right\|}{2}, \frac{\left\| \hat{\sigma}_i - 1_{\{\underline{a}_i\}} \right\|}{2} \right)$$

By letting W_i be the open ball $B(\hat{\sigma}_i, d_i)$, we have that $W = \prod_{i \in I} W_i$ is open set containing $\hat{\sigma}$. Finally, by Proposition 3, both $p^{\underline{a}_i}$ and $p^{\overline{a}_i}$ lie in $[0, 1)$. Let $p_i \in (p^{\underline{a}_i}, 1] \cap (p^{\overline{a}_i}, 1]$ for each player i, and define $p = (p_i)_{i \in I}$. We now show that intended play starting at $\mu(h^{t_0})$ and updated according to $p-$consistency remains outside of the open set W containing $\hat{\sigma}$.

It is an easy fact to check that

$$\hat{\sigma}_{-i} <_F \mu_i(h^{t_0}) \leq_F 1_{\{\overline{a}_{-i}\}}$$

for each player i. Thus,

$$\underline{a}_i \overset{(1)}{=} \wedge BR_i(1_{\{\overline{a}_{-i}\}}) \overset{(2)}{\preceq} \wedge BR_i\left(\mu_i(h^{t_0})\right) \preceq \vee BR_i\left(\mu_i(h^{t_0})\right) \overset{(3)}{\preceq} \wedge BR_i(\hat{\sigma}_{-i}) \overset{(4)}{=} \underline{a}_i$$

where (1) follows from Proposition 7 (2), (2) and (3) from Proposition 6, and (4) from the fact that

so that $\underline{a}_i \in BR_i(\hat{\sigma}_{-i})$. Therefore, for each player i,

$$\left(m_i^{t_0+1} = 1_{\{\underline{a}_i\}}\right)_{i \in I} \notin W$$

To proceed by induction, suppose that for some $t \geq t_0 + 1$, $m_i^{t+1} = 1_{\{\underline{a}_i\}}$ for each i, so that $a^{t+1} = \underline{a}$. By $p-$consistent beliefs, since $\mu_i(h^{t+1})\left(\underline{a}_{-i}\right) \geq p_i > p_i^{\overline{a}_i}$, we have by Proposition 8 that $\overline{a}_i$ is a strict best response, so that

$$\left(m_i^{t+2} = 1_{\{\overline{a}_i\}}\right)_{i \in I} \notin W$$

Because a similar argument holds for any $t \geq t_0 + 1$ such that $m_i^{t+1} = 1_{\{\underline{a}_i\}}$ for each i, we have the result. $\qquad \square$

Recall that if $\hat{\sigma} = (\hat{\sigma}_i)_{i \in I}$ is any MSNE, then for each i, it must be that $supp(\hat{\sigma}_i) \subseteq [\underline{a}_i, \overline{a}_i]$. The condition $\underline{a}_i, \overline{a}_i \in BR_i\left(\hat{\sigma}_{-i}\right)$ for each player i simply requires that the extremal strategies that define these bounds are themselves best responses to opponents' equilibrium strategy, and does not require them to lie in the support of $\hat{\sigma}_i$ itself. Notice that this is always satisfied for binary-action games with equilibria in which each player purely mixes over both actions.

Example. Again consider the Dove-Hawk-Chicken example in the introduction. It is easily verified that this is a strict GSS for $\varepsilon \in (0, \frac{1}{2})$, and that $\hat{\sigma}$ defined by $\forall i$, $\hat{\sigma}_i = (\frac{1}{2}, \frac{1}{2})$ is a MSNE satisfying the conditions of Theorem 10. Also, since for each player i, $p_i^{\overline{a}_i} = p_i^{\underline{a}_i} = \frac{1}{2}$, then re-

gardless of how players update their beliefs, as long as any probability larger than $\frac{1}{2}$ is ascribed to the previous round's play, intended play will never return to $\hat{\sigma}$ after a perturbation. That is, $\hat{\sigma}$ if $p \in [0, 1)^N$ is any vector such that $p_i > \frac{1}{2}$ for each i, then $\hat{\sigma}$ is $p-$unstable.

References

Aliprantis, C. D. & Border, K. C. (1999). *Infinite dimensional analysis*, volume 32006. Berlin: Springer.

Athey, S. (1998). *Characterizing Properties of Stochastic Objective Functions*. MIT Working Paper 96-1R.

Baliga, S. & Sjostrom, T. (2009). Conflict games with payoff uncertainty (no. 2009, 05). Working Papers, Department of Economics, Rutgers, The State University of New Jersey.

Basteck, C., Daniëls, T. R., & Heinemann, F. (2013). Characterising equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory*, 148(6), 2620–2637.

Benaïm, M. & Hirsch, M. W. (1999). Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Games and Economic Behavior*, 29(1), 36–72.

Benaïm, M., Hofbauer, J., & Hopkins, E. (2009). Learning in games with unstable equilibria. *Journal of Economic Theory*, 144(4), 1694–1709.

Boylan, R. T. & El-Gamal, M. A. (1993). Fictitious play: A statistical study of multiple economic experiments. *Games and Economic Behavior*, 5(2), 205–222.

Brown, G. W. (1951). Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1), 374–376.

Carlsson, H. & Van Damme, E. (1993). Global games and equilibrium selection. *Econometrica: Journal of the Econometric Society*, (pp. 989–1018).

Cheung, Y. W. & Friedman, D. (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior*, 19(1), 46–76.

Cox, J. C. & Walker, M. (1998). Learning to play Cournot duopoly strategies. *Journal of economic behavior and organization*, 36(2), 141–161.

Crawford, V. P. (1985). Learning behavior and mixed-strategy Nash equilibria. *Journal of Economic Behavior and Organization*, 6(1), 69–78.

Dubey, P., Haimanko, O., & Zapechelnyuk, A. (2006). Strategic complements and substitutes, and potential games. *Games and Economic Behavior*, 54(1), 77–94.

Echenique, F. (2002). Comparative statics by adaptive dynamics and the correspondence principle. *Econometrica*, 70, 833–844.

Echenique, F. & Edlin, A. (2004). Mixed equilibria are unstable in games of strategic complements. *Journal of Economic Theory*, 118(1), 61–79.

Ellison, G. & Fudenberg, D. (2000). Learning purified mixed-equilibria. *Journal of Economic Theory*, 90(1), 84–115.

Frankel, D. M., Morris, S., & Pauzner, A. (2003). Equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory*, 108(1), 1–44.

Fudenberg, D. (1998). *The theory of learning in games (Vol. 2)*. Chicago: MIT press.

Fudenberg, D. & Kreps, D. M. (1993). Learning mixed equilibria. *Games and Economic Behavior*, 5(3), 320–367.

Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F., & Yariv, L. (2010). Network games. *The review of economic studies*, 77(1), 218–244.

Gerber, A. (2006). *Learning in and about Games*. University of Zurich: Institute for Empirical Research in Economics.

Goldmann, K. (1994). *The logic of internationalism: coercion and accommodation*. Psychology Press.

Govindan, S., Reny, P. J., & Robson, A. J. (2003). A short proof of harsanyi's purification theorem. *Games and Economic Behavior*, 45(2), 369–374.

Harrison, R. (2003). *Equilibrium selection in global games with strategic substitutes (No. 0306003)*. EconWPA.

Healy, P. J. (2006). Learning dynamics for mechanism design: An experimental comparison of public goods mechanisms. *Journal of Economic Theory*, 129(1), 114–149.

Healy, P. J. (2007). Group reputations, stereotypes, and cooperation in a repeated labor market. *The American Economic Review*, 97(5), 1751–1773.

Hofbauer, J. & Hopkins, E. (2005). Learning in perturbed asymmetric games. *Games and Economic Behavior*, 52(1), 133–152.

Hoffmann, E. (2014). *Bayesian Games of Strategic Substitutes*. University of Kansas.

Huck, S., Normann, H. T., & Oechssler, J. (2002). Stability of the Cournot process-experimental evidence. *International Journal of Game Theory*, 31(1), 123–136.

Jordan, J. S. (1993). Three problems in learning mixed strategy Nash equilibria. *Games and Economic Behavior*, 5(3), 368–386.

Kajii, A. & Morris, S. (1997). The robustness of equilibria to incomplete information. *Econometrica: Journal of the Econometric Society*, (pp. 1283–1309).

Kaniovski, Y. M. & Young, H. P. (1995). Learning dynamics in games with stochastic perturbations. *Games and Economic Behavior*, 11(2), 330–363.

Keser, C., Suleymanova, I., & Wey, C. (2012). Technology adoption in markets with network effects: Theory and experimental evidence. *Information Economics and Policy*, 24(3), 262–276.

Kilgour, D. M. & Zagare, F. C. (1991). Credibility, uncertainty, and deterrence. *American Journal of Political Science*, (pp. 305–334).

Macrae, C. N. & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual review of psychology*, 51(1), 93–120.

Milgrom, P. & Roberts, J. (1990). Rationalizability, learning and equilibrium in games with strategic complementarites. *Econometrica*, 58, 1255–1277.

Milgrom, P. & Roberts, J. (1991). Adaptive and sophisticated learning in normal form games. *Games and economic Behavior*, 3(1), 82–100.

Milgrom, P. & Shannon, C. (1994). Monotone comparative statics. *Econometrica*, 62, 157–180.

Moreno, D. & Walker, M. (1991). On recursive learning in noncooperative games (pp. 91-20). Discussion paper.

Morris, S. & Shin, H. S. (2009). Coordinating expectations: Global games with strategic substitutes. Working paper, Department of Economics, Princeton University.

Morris, S. E. & Shin, H. S. (2000). *Global games: theory and applications*. Cowles Foundation for Research in Economics.

Reny, P. (2011). On the existence of monotone pure-strategy equilibria in Bayesian games. *Econometrica*, 79, 499–553.

Roy, S. & Sabarwal, T. (2008). On the (non-) lattice structure of the equilibrium set in games with strategic substitutes. *Economic Theory*, 37(1), 161–169.

Roy, S. & Sabarwal, T. (2010). Monotone comparative statics for games with strategic substitutes. *Journal of Mathematical Economics*, 46(5), 793–806.

Roy, S. & Sabarwal, T. (2012). Characterizing stability properties in games with strategic substitutes. *Games and Economic Behavior*, 75(1), 337–353.

Rutström, E. E. & Wilcox, N. T. (2009). Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test. *Games and Economic Behavior*, 67(2), 616–632.

Sabarwal, T. & Roy, S. (2010). Monotone comparative statics for games with strategic substitutes. *Mathematical Economics*, 46, 793–806.

Shannon, C. (1995). Weak and strong monotone comparative statics. *Economic Theory*, 5(2), 209–227.

Topkis, D. M. (1998). *Supermodularity and complementarity*. Princeton University Press.

Van Zandt, T. (2010). Interim Bayesian Nash equilibrium on universal type spaces for supermodular games. *Journal of Economic Theory*, 145(1), 249–263.

Van Zandt, T. & Vives, X. (2007). Monotone equilibria in Bayesian games of strategic complementarities. *Journal of Economic Theory*, 134, 339–360.

Vives, X. (1990). Nash equilibria with strategic complementarities. *Journal of Mathematical Economics*, 9, 305–321.

Weinstein, J. & Yildiz, M. (2007). A structure theorem for rationalizability with application to robust predictions of refinements. *Econometrica*, 75(2), 365–400.

Appendix A

Additional Proofs, Chapter 2

The proof of Proposition 4 is given below. In order to do so, we first establish Lemmas $9-12$:

Lemma 9. *Suppose $\tilde{a}$ is a strict, $p-$dominant Nash equilibrium and for some player $i \in \mathscr{I}$, and $a_i \in A_i$, $x \notin D^{\tilde{a}_i,a_i}$. Then $\exists \lambda_i \in \triangle(A_{-i})$ such that $l_i(a_i, \lambda_i, x) = 0$.*

Proof. Since $\tilde{a}$ is a strict Nash equilibrium, we have that $l_i(a_i, 1_{\tilde{a}_{-i}}, x) > 0$. Suppose for all $\lambda_i \in \triangle(A_{-i})$, $l_i(a_i, \lambda_i, x) > 0$. Then for all a_{-i}, $\triangle u_i(\tilde{a}_i, a_i, a_{-i}, x) > 0$, contradicting the fact that $x \notin D^{\tilde{a}_i,a_i}$. Thus for some $\bar{\lambda}_i$, $l_i(a_i, \bar{\lambda}_i, x) \leq 0$. Consider the set of probability measures

$$\mathscr{Z} = \left\{ \lambda_i^{\alpha} = \begin{cases} \alpha + (1-\alpha)\bar{\lambda}_i(\tilde{a}_{-i}) & if\, a_{-i} = \tilde{a}_{-i} \\ (1-\alpha)\bar{\lambda}_i(a_{-i}) & if\, a_{-i} \neq \tilde{a}_{-i} \end{cases} \mid \alpha \in [0, 1] \right\}$$

Note that $\lambda_i^1 = 1_{\tilde{a}_{-i}}$ and $\lambda_i^0 = \bar{\lambda}_i$, giving us $l_i(a_i, \lambda_i^1, x) > 0$ and $l_i(a_i, \lambda_i^0, x) \leq 0$. Since $l_i(a_i, \lambda_i^{\alpha}, x)$ is a continuous function of α, by the intermediate value theorem $\exists \bar{\alpha} \in [0, 1)$ such that $l_i(a_i, \lambda_i^{\bar{\alpha}}, x) = 0$, giving the result. □ □

Lemma 11 will show that the set $\{\lambda_i \in \triangle(A_{-i}) \mid l_i(a_i, \lambda_i, x) = 0\}$ is compact, and since it is non-empty by Lemma 9, the value

$$\hat{p}_i(a_i, x) \equiv \max_{\substack{\lambda_i \in \triangle(A_{-i}) \\ l_i(a_i, \lambda_i, x) = 0}} (\lambda_i(\tilde{a}_{-i}))$$

is well-defined. We now show that for all such $x \in X$, $p_i(a_i, x) = \hat{p}_i(a_i, x)$.

Lemma 10. *Suppose that $x \notin D^{\tilde{a}_i, a_i}$. Then $p_i(a_i, x) = \hat{p}_i(a_i, x)$.*

Proof. It is first shown that $p_i(a_i, x) \leq \hat{p}_i(a_i, x)$. In order to do so, we show that for all $\lambda_i' \in \triangle(A_{-i})$ such that $\lambda_i'(\tilde{a}_{-i}) \geq \hat{p}_i(a_i, x)$, $l_i(a_i, \lambda_i', x) \geq 0$. Because $p_i(a_i, x)$ is defined as the lowest value that satisfies this property, the conclusion follows. Suppose $\lambda_i' \in \triangle(A_{-i})$ is such that $\lambda_i'(\tilde{a}_{-i}) \geq \hat{p}_i(a_i, x)$, but $l_i(a_i, \lambda_i', x) < 0$. Since $\tilde{a}$ is a strict Nash equilibrium, we have that $l_i(a_i, 1_{\tilde{a}_{-i}}, x) > 0$. Consider the set of probability measures

$$\mathcal{L} = \left\{ \lambda_i^\alpha = \begin{cases} \alpha + (1-\alpha)\lambda_i'(\tilde{a}_{-i}) & if\, a_{-i} = \tilde{a}_{-i} \\ (1-\alpha)\lambda_i'(a_{-i}) & if\, a_{-i} \neq \tilde{a}_{-i} \end{cases} \Bigg| \, \alpha \in [0, 1] \right\}$$

Note that $\lambda_i^1 = 1_{\tilde{a}_{-i}}$ and $\lambda_i^0 = \lambda_i'$, giving us $l_i(a_i, \lambda_i^1, x) > 0$ and $l_i(a_i, \lambda_i^0, x) < 0$. Since $l_i(a_i, \lambda_i^\alpha, x)$ is a continuous function of α, by the intermediate value theorem $\exists \bar{\alpha} \in (0, 1)$ such that $l_i(a_i, \lambda_i^{\bar{\alpha}}, x) = 0$. Since $\lambda_i^{\bar{\alpha}}(\tilde{a}_{-i}) = \bar{\alpha} + (1 - \bar{\alpha})\lambda_i'(\tilde{a}_{-i}) > \lambda_i'(\tilde{a}_{-i})$, we have found a $\lambda_i^{\bar{\alpha}}$ such that $l_i(a_i, \lambda_i^{\bar{\alpha}}, x) = 0$ but $\lambda_i^{\bar{\alpha}}(\tilde{a}_{-i}) > \lambda_i'(\tilde{a}_{-i}) \geq \hat{p}_i(a_i, x)$, contradicting the definition of $\hat{p}_i(a_i, x)$. Hence $p_i(a_i, x) \leq \hat{p}_i(a_i, x)$.

To show equality, suppose that $p_i(a_i, x) < \hat{p}_i(a_i, x)$. We will show that for any $\lambda_i'' \in \triangle(A_{-i})$ such that $\lambda_i''(\tilde{a}_{-i}) > p_i(a_i, x)$, we must have $l_i(a_i, \lambda_i'', x) > 0$. Hence if $p_i(a_i, x) < \hat{p}_i(a_i, x)$ is true, any λ_i'' such that $\lambda_i''(\tilde{a}_{-i}) = \hat{p}_i(a_i, x)$ must be such that $l_i(a_i, a_{-i}, \lambda_i'', x) > 0$, a contradiction to the existence of $\hat{p}_i(a_i, x)$. Let $\mathscr{A} = \underset{a_{-i}}{argmin}(\triangle u_i(\tilde{a}_i, a_i, a_{-i}, x))$. Note that $\tilde{a}_{-i}$ is not part of this set, or else $x \notin D^{\tilde{a}_i, a_i}$ would be contradicted. Let λ_i be such that $\lambda_i(\tilde{a}_{-i}) = p_i(a_i, x)$, and assign to all a_{-i} in $\mathscr{A}$ the probability $\frac{1 - \lambda_i(\tilde{a}_{-i})}{|\mathscr{A}|}$. By the definition of $\tilde{a}_i$ being $p_i(a_i, x)$ dominant, we must have that $l_i(a_i, \lambda_i, x) \geq 0$. Now let λ_i' be such that $\lambda_i'(\tilde{a}_{-i}) > p_i(a_i, x)$, and assign to all a_{-i} in $\mathscr{A}$ the

87

probability $\frac{1-\lambda_i'(\tilde{a}_{-i})}{|\mathscr{A}|}$. It follows that[1]

$$l_i(a_i, \lambda_i', x) = \triangle u_i(\tilde{a}_i, a_i, \tilde{a}_{-i}, x)\lambda_i'(\tilde{a}_{-i}) + \left(\frac{1-\lambda_i'(\tilde{a}_{-i})}{|\mathscr{A}|}\right)\sum_{\mathscr{A}}\triangle u_i(\tilde{a}_i, a_i, a_{-i}x) >$$

$$\triangle u_i(\tilde{a}_i, a_i, \tilde{a}_{-i}, x)\lambda_i(\tilde{a}_{-i}) + \left(\frac{1-\lambda_i(\tilde{a}_{-i})}{|\mathscr{A}|}\right)\sum_{\mathscr{A}}\triangle u_i(\tilde{a}_i, a_i, a_{-i}x) = l_i(a_i, \lambda_i, x) \geq 0$$

Finally, let λ_i'' be arbitrary but satisfying $\lambda_i''(\tilde{a}_i) = \lambda_i'(\tilde{a}_{-i})$. By the construction of λ_i', $l_i(a_i, \lambda_i', x)$ gives the smallest value for such probability measures, and therefore $l_i(a_i, a_{-i}, \lambda_i'', x) \geq l_i(a_i, a_{-i}, \lambda_i', x) > l_i(a_i, a_{-i}, \lambda_i, x) \geq 0$, or $l_i(a_i, \lambda_i'', x) > 0$, giving the result. $\qquad\square\quad\square$

The upper semi-continuity of $p_i(a_i, x)$ is now established in two steps.

Lemma 11. $p_i(a_i, x)$ *is an upper semi-continuous function on* $X \cap (D^{\tilde{a}_i, a_i})^C$.

Proof. Recall the Maximum Theorem:[2] If $p_i(a_i, x) = \max\limits_{\substack{\lambda_i \in \triangle(A_{-i}) \\ l_i(a_i, \lambda_i, x) = 0}} (\lambda_i(\tilde{a}_{-i}))$ is such that $\varphi : X \cap (D^{\tilde{a}_i, a_i})^C \rightrightarrows \triangle(A_{-i})$, $\varphi(x) = \{\lambda_i \in \triangle(A_{-i}) \mid l_i(a_i, \lambda_i, x) = 0\}$ is upper hemi-continuous with non-empty, compact values, and $f : gf(\varphi) \to \mathbb{R}$ defined as $f(x, \lambda_i) = \lambda_i(\tilde{a}_{-i})$ is upper semi-continuous, then $p_i(a_i, \cdot)$ is upper semi-continuous. We show one-by-one that these conditions are met:

$\square$

Let $(x^n, \lambda_i^n)_{n=1}^{\infty} \subseteq gf(\varphi)$ be such that $(x^n, \lambda_i^n) \to (x, \lambda)$. Then $\lim\limits_{n\to\infty} (f(x^n, \lambda_i^n)) = \lim\limits_{n\to\infty} (\lambda_i^n(\tilde{a}_{-i})) = \lambda_i(\tilde{a}_{-i}) = f(x, \lambda_i)$, showing that f is continuous, and therefore upper semi-continuous.

By Lemma 9, $\varphi : X \cap (D^{\tilde{a}_i, a_i})^C \rightrightarrows \triangle(A_{-i})$ is non-empty valued. To see that it is compact valued, let x be given and suppose $(\lambda_i^n)_{n=1}^{\infty} \subseteq \varphi(x)$ is such that $\lambda_i^n \to \lambda_i$. Because $\triangle(A_{-i})$ is closed, $\lambda_i \in \triangle(A_{-i})$. Because $l_i(\cdot, x) : \triangle(A_{-i}) \to \mathbb{R}$ is continuous, $l_i(a_i, \lambda_i, x) = 0$, and hence $\varphi(x)$

---

[1]Note that $\sum_{\mathscr{A}}\triangle u_i(\tilde{a}_i, a_i, a_{-i}, x) \leq 0$: If $\triangle u_i(\tilde{a}_i, a_i, a_{-i}, x) > 0$ for each a_{-i}, $x \notin D^{\tilde{a}_i, a_i}$ is contradicted. Thus for some a_{-i}, $\triangle u_i(\tilde{a}_i, a_i, a_{-i}, x) \leq 0$, and the conclusion follows by the definition of $\mathscr{A}$.

[2]Aliprantis, Lemma 17.30

is closed-valued. Since $\varphi(x) \subseteq \triangle(A_{-i})$, it is therefore compact valued.

Finally, we see that φ is upper hemi-continuous. Recall that a correspondence with a compact and Hausdorff range space has a closed graph if and only if it is upper hemi-continuous and closed valued. It therefore suffices to show that φ has a closed graph. To that end, suppose $(x^n, \lambda_i^n)_{n=1}^\infty \subseteq gf(\varphi)$ is such that $(x^n, \lambda_i^n) \to (x, \lambda_i)$. Because $X \cap (D^{\tilde{a}_i, a_i})^C$ is closed, $x \in X$. Because $\triangle(A_{-i})$ is closed, $\lambda_i \in \triangle(A_{-i})$. Lastly, because $l_i : X \times \triangle(A_{-i}) \to \mathbb{R}$ is continuous, $l_i(a_i, \lambda_i, x) = \lim\limits_{n \to \infty} (l_i(a_i, \lambda_i^n, x^n)) = 0$, and thus $(x, \lambda_i) \in gf(\varphi)$, so that $gf(\varphi)$ is closed, completing the Lemma. $\square$

Lemma 12. $p_i(a_i, x)$ *is an upper semi-continuous function on all of X.*

Proof. Let $(x_n)_{n=1}^\infty \subseteq X$ be such that $x_n \to x$. It's shown that $limsup\limits_{n}(p_i(a_i, x_n)) \leq p_i(a_i, x)$ by considering two cases:

$\square$

Case 1: Suppose $\exists K > 0, \forall n \geq K, x_n \in D^{\tilde{a}_i, a_i}$. If $x \in D^{\tilde{a}_i, a_i}$, then

$$limsup\limits_{n}(p_i(a_i, x_n)) = \inf\limits_{k \geq 1}\sup\limits_{n \geq k}(p_i(a_i, x_n)) \leq \sup\limits_{n \geq K}(p_i(a_i, x_n)) = 0 \leq p_i(a_i, x)$$

by definition.

If $x \notin D^{\tilde{a}_i, a_i}$, then $\sup\limits_{n \geq K}(p_i(a_i, x_n)) \leq p_i(a_i, x)$. Thus

$$limsup\limits_{n}(p_i(a_i, x_n))) = \inf\limits_{k \geq 1}\sup\limits_{n \geq k}(p_i(a_i, x_n)) \leq \sup\limits_{n \geq K}(p_i(a_i, x_n)) \leq p_i(a_i, x)$$

giving the result.

Proof. Case 2: Suppose $\forall K > 0, \exists k_K \geq K, x_{k_K} \notin D^{\tilde{a}_i, a_i}$. Let $\tilde{N} \subseteq \mathbb{N}$ be those indices such that $n \in \tilde{N} \Rightarrow x_n \notin D^{\tilde{a}_i, a_i}$, and define $m : \mathbb{N} \to \mathbb{N}$ by $m(n) = \min\limits_{\substack{j \in \tilde{N} \\ j \geq n}}(j)$. Define the sequence $(x_n')_{n=1}^\infty$ by the

89

formula $x_n' = x_{m(n)}$. Then we have the following:

$\square$

1. $\forall n,\, p_i(a_i, x_n') \geq p_i(a_i, x_n)$: Let n be given. If $x_n \in D^{\tilde{a}_i, a_i}$, then $p_i(a_i, x_n') \geq 0 = p_i(a_i, x_n)$. If $x_n \notin D^{\tilde{a}_i, a_i}$, then $x_n' = x_{m(n)} = x_n$, so the inequality follows.

2. $x_n' \to x$: Let $\varepsilon > 0$ be given. Since $x_n \to x$, $\exists K,\, \forall n \geq K,\, |x_n - x| < \varepsilon$. Since for all n, $m(n) = \min_{\substack{j \in \tilde{N} \\ j \geq n}}(j) \geq n$, then for all $n \geq K$, $|x_n' - x| = |x_{m(n)} - x| < \varepsilon$, giving convergence.

Finally, $\limsup_n(p_i(a_i, x_n)) \leq \limsup_n(p_i(a_i, x_n')) \leq p_i(a_i, x)$, where the first inequality follows from 1. The second follows from 2., the fact that $(x_n')_{n=1}^\infty \subseteq (D^{\tilde{a}_i, a_i})^C$, and that by Lemma 11, $p_i(a_i, \cdot)$ is upper semi-continuous on $X \cap (D^{\tilde{a}_i, a_i})^C$. $\square$

We are now in a position to complete the proof of Proposition 4 below:

Proof. (of Proposition 4) The first claim was established after Lemma 10. For the second claim, let x be given. Let $\lambda_i \in \triangle(A_{-i})$ be such that $\lambda_i(\tilde{a}_{-i}) \geq \max_{a_i}(p_i(a_i, x))$. Choose any a_i', giving $\lambda_i(\tilde{a}_{-i}) \geq p_i(a_i', x)$. If $x \in D^{\tilde{a}_i, a_i'}$, we trivially have that $l_i(a_i', \lambda_i, x) \geq 0$. If $x \notin D^{\tilde{a}_i, a_i'}$, then by the first part of Lemma 10, $\lambda_i(\tilde{a}_{-i}) \geq p_i(a_i', x) \Rightarrow l_i(a_i', \lambda_i, x) \geq 0$. Thus $\lambda_i(\tilde{a}_{-i}) \geq \max_{a_i}(p_i(a_i, x))$ implies that for any a_i', $l_i(a_i', \lambda_i, x) \geq 0$. Since $p_i(x)$ is defined as the smallest value satisfying this property, we must have that $\max_{a_i}(p_i(a_i, x)) \geq p_i(x)$. For a contradiction, suppose that this inequality is strict. Since for all $x \in D^{\tilde{a}_i}$, $\max_{a_i}(p_i(a_i, x)) = 0$, we must have that $x \notin D^{\tilde{a}_i}$. In particular, let a_i' be such that $x \notin D^{\tilde{a}_i, a_i'}$ and $\max_{a_i}(p_i(a_i, x)) = p_i(a_i', x)$, so that $x \notin D^{\tilde{a}_i, a_i'}$ and $p_i(a_i', x) > p_i(x)$. By repeating the second part of Lemma 10 with $p_i(x)$ in place of $p_i(a_i, x)$, we must have that for any $\lambda_i \in \triangle(A_{-i})$ such that $\lambda_i(\tilde{a}_{-i}) > p_i(a_i', x)$, $l_i(a_i', \lambda_i, x) > 0$, contradicting $p_i(a_i', x) = \max_{\substack{\lambda_i \in \triangle(A_{-i}) \\ l_i(a_i', \lambda_i, x) = 0}}(\lambda_i(\tilde{a}_{-i}))$.

Therefore, $p_i(x) = \max_{a_i}(p_i(a_i, x))$. The fact that $p_i(x)$ is upper semi-continuous on X follows from the fact that each $\hat{p}_i(a_i, x)$ is. $\square$ $\square$

Lemma 5 is now proven:

Proof. (Of Lemma 5) The first assertion is given in CvD (1993) and is thus omitted.

For the second assertion, for each player i, let φ_i^v denote the pdf of $v\varepsilon_i$, which is distributed uniformly on $[-v, v]$. After receiving x_i, we have that

$$f_i^*(x_{-i}|x_i, v) = \frac{Pr(x_i \& x_{-i})}{Pr(x_i)} = \frac{\int_{x_i-v}^{x_i+v} f^*(x)\varphi_i^v(x_i - x)\left(\prod_{j\neq i}\varphi_j^v(x_j - x)\right)dx}{\int_{\mathbb{R}^{N-1}}\int_{x_i-v}^{x_i+v} f^*(x)\varphi_i^v(x_i - x)\left(\prod_{j\neq i}\varphi_j^v(x_j - x)\right)\left(\prod_{j\neq i}dx_j\right)dx}$$

Because $x_i \in I$, $f^*(x)$ will be constant on the entire range of integration, and thus after factoring this out and noticing that the denominator now equals 1, we have

$$f_i^*(x_{-i}|x_i, v) = \int_{x_i-v}^{x_i+v} \varphi_i^v(x_i - x)\left(\prod_{j\neq i}\varphi_j^v(x_j - x)\right)dx$$

Since each $\varphi_j^v(x_j - x) = \frac{1}{2v}$ on $[x_j - v, x_j + v]$ and zero elsewhere, we then have that

$$f_i^*(x_{-i}|x_i, v) = \left(\frac{1}{2v}\right)^N m\left(\bigcap_{j\in I}[x_j - v, x_j + v]\right)$$

where $m(I)$ is the length of any interval I.

To make the calculation $F_i^*(\vec{x}_{-i}|x_i, v)$, consider vectors of the form $\vec{x}_i - a = (x_i - a_j)_{j\neq i}$, where each $a_j \in [0, 2v)$. By the above equation we have that whenever $a_m = \max_{j\neq i}(a_j)$,

$$f_i^*(\vec{x}_i - a|x_i, v) = \left(\frac{1}{2v}\right)^N (2v - a_m), \ \ 0 \leq a_m \leq 2v$$

If we let $\mathscr{I}/\{i\} = \{1.2, ..., N - 1\}$ denote the set of player i's opponents, then for any $k \in \mathscr{I}/\{i\}$ and corresponding a_k, the above equation then gives

91

$$Pr\left(\{0 \leq a_1 < a_2 < \cdots < a_k \leq 2v\} \mid x_i, v\right) = \int_0^{2v} \int_0^{a_k} \cdots \int_0^{a_2} \left(\frac{1}{2v}\right)^N (2v - a_k) \prod_{j=1}^k da_j$$

For any such k, there are $(N-2)!$ orderings of the $(a_j)_{j \neq i,k}$ with a_k as the largest. Thus if we denote the set

$$A_k = \{\vec{x}_i - a \mid 0 \leq a_1 < a_2 < \cdots < a_k \leq 2v\}$$

and notice that[3] $\{x_i - 2v \leq x_k \leq \cdots \leq x_2 \leq x_1 \leq x_i\} \supseteq A_k$, we then have that

$$Pr\left(\{x_i - 2v \leq x_k \leq \cdots \leq x_2 \leq x_1 \leq x_i\} \mid x_i, v\right) \geq \int_{A_k} f_i^*(\vec{x}_i - a \mid x_i, v) da =$$

$$(N-2)! \int_0^{2v} \int_0^{a_k} \cdots \int_0^{a_2} \left(\frac{1}{2v}\right)^N (2v - a_k) \prod_{j=1}^k da_j$$

Finally, notice that $\{x_{-i} \leq \vec{x}_i\} \supseteq \bigcup_{j \neq i} A_j$, so that

$$F_i^*\left(\vec{x}_{-i} \mid x_i, v\right) \geq \sum_{j \neq i} \left(\int_{A_j} f_i^*(\vec{x}_i - a \mid x_i, v) da\right) =$$

$$(N-1)(N-2)! \int_0^{2v} \int_0^{a_k} \cdots \int_0^{a_2} \left(\frac{1}{2v}\right)^N (2v - a_j) \prod_{j=1}^k da_j =$$

$$(N-1)! \int_0^{2v} \int_0^{a_k} \cdots \int_0^{a_2} \left(\frac{1}{2v}\right)^N (2v - a_k) \prod_{j=1}^k da_j = \frac{1}{N}$$

giving the result. $\qquad \square \quad \square$

Theorem 8 is proven below:

Proof. (of Theorem 8) Choose $\varepsilon > 0$ and let $v(\varepsilon) \in (0, \frac{\varepsilon}{K}]$, where K is the Lipschitz constant associated with players' utility. Choosing any $v \in (0, v(\varepsilon)]$, we proceed by induction by showing that for each $n \geq 0$, $S_n(0) \subseteq S_n^{**}(\varepsilon)$.

For $n = 0$, $S_0(0) = S = S_0^{**}(\varepsilon)$. Suppose that for some $n \geq 0$, $S_n(0) \subseteq S_n^{**}(\varepsilon)$. Let $s \in S_{n+1}(0) =$

---

[3]This is true on $supp\left(f_i^*(x_{-i} \mid x_i, v)\right)$.

$UR(S_n(0), \varepsilon)$, and fix x_i for some player $i \in I$. Then $\forall a_i \in A_i$, $\exists s_{-i} \in S_n(0)$, $\triangle \pi_{x,x_{-i}}(s_i(x_i), a_i, s_{-i}, x_i) \geq$

0. This implies that[4]

$$0 \leq \triangle \pi_{x,z_{-i}}(s_i(x_i), a_i s_{-i}, x_i) = \int\limits_{Z_{x_i}} \int\limits_{\mathbb{R}} \triangle u_i(s_i(x_i), a_i, s_{-i}, x) d\tilde{\mu}_i(z_{-i}, x | x_i, v)$$

$$\leq \int\limits_{Z_{x_i}} \int\limits_{\mathbb{R}} \triangle u_i(s_i(x_i), a_i, s_{-i}, x_i) d\tilde{\mu}_i(z_{-i}, x | x_i, v) + Kv$$

$$= \int\limits_{Z_{x_i}} \triangle u_i(s_i(x_i), a_i, s_{-i}, x_i) d\tilde{\mu}_i^{**}(z_{-i} | x_i, v) + Kv$$

That is,

$$\int\limits_{Z_{x_i}} \triangle u_i(s_i(x_i), a_i, s_{-i}, x_i) d\tilde{\mu}_i^{**}(z_{-i} | x_i, v) \geq -Kv \geq -\varepsilon$$

Since $s_{-i} \in S_n(0) \subseteq S_n^{**}(\varepsilon)$, this shows that s_i survives ε−iterated deletion at round n in the game $G^{**}(v)$, and thus

$$S_{n+1}(0) = UR(S_n(0), \varepsilon) \subseteq UR(S_n^{**}(\varepsilon), 0) = S_{n+1}^{**}(\varepsilon)$$

completing the induction step. Therefore,

$$SU(G(v), 0) = \bigcap_{n \geq 0} S_n(\varepsilon) \subseteq \bigcap_{n \geq 0} S_n^{**}(\varepsilon) = SU(G^{**}(v), \varepsilon)$$

concluding the proof.   □   □

---

[4]The second line follows from the fact that after receiving signal x_i, the support of player i's beliefs about x are contained in $(x_i - v, x_i + v)$ and that by Lipschitz continuity, $x \in B(x_i, v)$ implies $|u_i(a, x_i) - u_i(a, x)| \leq Kv$, $\forall a \in A$.

Appendix B

Additional Proofs, Chapter 3

First, Lemma 8 is proven, showing that when beliefs are geometrically weighted, they satisfy monotonicity and are asymptotically empirical.

Proof. (Lemma 8) Concentrating on the right hand side, to establish monotonicity, assume that $\mu_i(h^{t_0}) \leq_F \mu_i(\bar{h}^{t_0})$ and $h^t \preceq \bar{h}^t$ for some $t \geq t_0$, where $h^t = \left(a^{t_0+1}, a^{t_0+2}, ..., a^t\right)$ and $\bar{h}^t = \left(\bar{a}^{t_0+1}, \bar{a}^{t_0+2}, ..., \bar{a}^t\right)$. The result holds trivially for $t = t_0$ by hypothesis. Suppose for $t_0 \leq l \leq t-1$, $\mu_i(h^{t_0}, ..., a^l) \leq_F \mu_i(\bar{h}^{t_0}, ..., \bar{a}^l)$, and let $E \subseteq A_{-i}$ be increasing. Since $\bar{a}^{l+1} \succeq a^{l+1}$, by the increasingness of E, $1_{\left\{a_{-i}^{l+1}\right\}}(E) \leq 1_{\left\{\bar{a}_{-i}^{l+1}\right\}}(E)$. Thus,

$$\mu_i(h^{t_0}, ..., a^l, a^{l+1})(E) = (1 - \lambda_t)\mu_i(h^{t_0}, ..., a^l)(E) + \lambda_t 1_{\left\{a_{-i}^{l+1}\right\}}(E) \leq$$

$$\mu_i(h^t, y^{t+1})) \leq_F \mu_i(h^t) \leq_F \mu_i(h^t, z^{t+1}))$$

Therefore, $\mu_i(h^{t_0}, ..., a^l, a^{l+1}) \leq_F \mu_i(\bar{h}^{t_0}, ..., \bar{a}^l, \bar{a}^{l+1})$. Because the left hand side can be done similarly, we see that the first condition is satisfied.

For the next condition, suppose that $\left(\vee h^t_{-i}\right) \vee \bar{M}_i \preceq z^{t+1}_{-i}$ for $t \geq t_0$, and let $E \subseteq A_{-i}$ be increasing. Note that we immediately have that

$$z^{t+1}_{-i} \succeq \left(\vee h^t_{-i}\right) \vee \bar{M}_i \succeq supp\left(\mu_i(h^t)\right)$$

If $z^{t+1}_{-i} \notin E$, then by increasingness of E and the inequality above, $supp(\mu_i(h^t)) \cap E = \emptyset$. Thus

$$\mu_i(h^{t+1})(E) \geq 0 = \mu_i(h^t)$$

If $z^{t+1}_{-i} \in E$, then

$$\mu_i(h^{t+1})(E) = (1 - \lambda_{t+1})\mu_i(h^t)(E) + \lambda_{t+1} = \mu_i(h^t)(E) + \lambda_{t+1}(1 - \mu_i(h^t)(E)) \geq \mu_i(h^t)(E)$$

Therefore $\mu_i(h^t) \leq_F \mu_i(h^{t+1})$. Since the left hand side can be done similarly, we see that beliefs satisfy monotonicity.

Lastly, we show that beliefs are asymptotically empirical, which must be done according to weak convergence because of arbitrary action spaces. Suppose that $a^t \to \hat{a}$, consider $i \in I$, and let $f : A_{-i} \to \mathbb{R}$ be continuous. For each $\varepsilon > 0$, let U^ε be the open set in A_{-i} containing $\hat{a}_{-i}$ such that $\forall a_{-i} \in U^\varepsilon, |f(\bar{a}_{-i}) - f(a_{-i})| < \frac{\varepsilon}{2}$. By compactness, let M be the absolute maximum value attained on A_{-i} by f. Since $a^t_{-i} \to \hat{a}_{-i}$, by geometric updating we have that for each $\varepsilon > 0$, there exists a $T^\varepsilon \geq 0$ such that for each $t \geq T^\varepsilon$, $\mu_i(h^t)\left((U^\varepsilon)^C\right) < \frac{\varepsilon}{4M}$.

Let $\varepsilon > 0$ be given, and choose the corresponding U^ε and T^ε. Then for each $t \geq T^\varepsilon$,

$$\left| \int_{A_{-i}} f d1_{\{\hat{a}_{-i}\}} - \int_{A_{-i}} f d\mu_i(h^t) \right| = \left| \int_{A_{-i}} f(\hat{a}_{-i}) d\mu_i(h^t) - \int_{A_{-i}} f d\mu_i(h^t) \right|$$

$$= \left| \int_{(U^\varepsilon)^C} (f(\hat{a}_{-i}) - f)\, d\mu_i(h^t) + \int_{U^\varepsilon} (f(\hat{a}_{-i}) - f)\, d\mu_i(h^t) \right|$$

$$\leq \left| \int_{(U^\varepsilon)^C} (f(\hat{a}_{-i}) - f)\, d\mu_i(h^t) \right| + \left| \int_{U^\varepsilon} (f(\hat{a}_{-i}) - f)\, d\mu_i(h^t) \right|$$

$$\leq (2M)\,\mu_i(h^t)\left((U^\varepsilon)^C\right) + \left(\frac{\varepsilon}{2}\right)\mu_i(h^t)(U^\varepsilon) < \left(\frac{\varepsilon}{2}\right) + \left(\frac{\varepsilon}{2}\right) = \varepsilon$$

Therefore $a^t \to \hat{a}$ implies that for each i, $\mu_i(h^t) \to 1_{\{\hat{a}_{-i}\}}$, giving the result. $\qquad\square$

Claim 3 of Proposition 6 is given below, which follows along the lines of Theorem 1 in EE. First, one more definition is introduced.

Definition 25. $\forall i \in I$, U_i satisfies the strict single-crossing property on $A_i \times \triangle(A_{-i})$ if for all $a_i', a_i \in A_i$ such that $a_i' > a_i$ and $\mu', \mu \in \triangle(A_{-i})$ such that $\mu' \succ_F \mu$,

$$U_i(a_i', \mu) - U_i(a_i, \mu) \leq 0 \Rightarrow U_i(a_i', \mu') - U_i(a_i, \mu') < 0$$

Lemma 13. *Define the poset $T = \left\{ \mu', \mu \right\} \subseteq \triangle(A_{-i})$ by supposing that $\mu' >_F \mu$. If U_i satisfies the strict single-crossing property on $A_i \times T$, then*

$$\wedge BR_i(\mu) \geq \vee BR_i(\mu')$$

Proof. Suppose that $\vee BR_i(\mu') > \wedge BR_i(\mu)$. By Claim 1 of Proposition 6, we have that $\wedge BR_i(\mu) \in BR_i(\mu)$, so that $U_i(\vee BR_i(\mu'), \mu) - U_i(\wedge BR_i(\mu), \mu) \leq 0$. Thus by the strict single-crossing property,

$$U_i(\vee BR_i(\mu'), \mu') - U_i(\wedge BR_i(\mu), \mu') < U_i(\vee BR_i(\mu'), \mu) - U_i(\wedge BR_i(\mu), \mu) \leq 0$$

so that $U_i(\wedge BR_i(\mu), \mu') > U_i(\vee BR_i(\mu'), \mu')$, which, again by Claim 1, contradicts $\vee BR_i(\mu') \in BR_i(\mu')$. $\square$

Proof. (Claim 3, Proposition 6) The case of when $a_{-i} = \bar{a}_{-i}$ is shown, the other case follows similarly. The result follows from Lemma 13 by showing that U_i satisfies the strict single-crossing property on $A_i \times T$, where $T = \{\hat{\sigma}_{-i}, \mu^{\varepsilon}\}$. It is easy to verify that for any increasing set $A \subseteq A_{-i}$, $\mu^{\varepsilon}(A) \geq \hat{\sigma}_{-i}(A)$, and that by considering the set $E \equiv [\bar{a}_{-i}, \vee A_{-i}] \subseteq A_{-i}$, $\mu^{\varepsilon}(E) > \hat{\sigma}_{-i}(E)$. Thus $\mu^{\varepsilon} >_F \hat{\sigma}_{-i}$.

Let $a_i', a_i \in A_i$ be such that $a_i' > a_i$, and suppose that $U_i(a_i, \hat{\sigma}) - U_i(a_i', \hat{\sigma}) \geq 0$. We must show that $U_i(a_i, \mu_i^{\varepsilon}) - U_i(a_i', \mu_i^{\varepsilon}) > 0$. Define the function $l : A \to \mathbb{R}$ as in Claim 2. Since $\mu_i^{\varepsilon} =$

$(1-\varepsilon)\hat{\sigma}_{-i}+\varepsilon 1_{\{\overline{a}_{-i}\}}$, we have that

$$U_i(a_i,\mu^\varepsilon)-U_i(a_i',\mu^\varepsilon)=(1-\varepsilon)\int_{A_{-i}} l(a_{-i})d\hat{\sigma}_{-i}+\int_{A_{-i}} l(a_{-i})d\left(1_{\overline{a}_{-i}}\right)$$

$$=(1-\varepsilon)\left(U_i(a_i',\hat{\sigma}_{-i})-U_i(a_i,\hat{\sigma}_{-i})\right)+\varepsilon l(\overline{a}_{-i})$$

Suppose by way of contradiction that $U_i(a_i,\mu^\varepsilon)-U_i(a_i',\mu^\varepsilon)\le 0$. By hypothesis, $U_i(a_i,\hat{\sigma}_{-i})-U_i(a_i',\hat{\sigma}_{-i})\ge 0$, so that we must have $l(\overline{a}_{-i})\le 0$. By strict decreasing differences, l is strictly increasing, so that $\forall a_{-i}\in BR_{-i}(\hat{\sigma})/\{\overline{a}_{-i}\}, l(a_{-i})<0$.

We now show that $\hat{\sigma}_{-i}(BR_{-i}(\hat{\sigma})/\{\overline{a}_{-i}\})=0$. To see this, note if $\hat{\sigma}_{-i}(BR_{-i}(\hat{\sigma})/\{\overline{a}_{-i}\})>0$, then

$$0>\int_{BR_{-i}(\hat{\sigma})/\{\overline{a}_{-i}\}} l(a_{-i})d\hat{\sigma}_{-i}\ge\int_{A_{-i}} l(a_{-i})d\hat{\sigma}_{-i}=U_i(a_i,\hat{\sigma}_{-i})-U_i(a_i',\hat{\sigma}_{-i})$$

where the first inequality comes from the fact that $l(a_{-i})<0$ for all $a_{-i}\in BR_{-i}(\hat{\sigma})/\{\overline{a}_{-i}\}$, and the second from the fact that since $\hat{\sigma}$ is a MSNE, $supp(\hat{\sigma}_{-i})\subseteq BR_{-i}(\hat{\sigma})$. This contradicts $U_i(a_i',\hat{\sigma}_{-i})-U_i(a_i,\hat{\sigma}_{-i})\ge 0$. Thus, $\hat{\sigma}_{-i}(BR_{-i}(\hat{\sigma})/\{\overline{a}_{-i}\})=0$. However, since $supp(\hat{\sigma}_{-i})\subseteq BR_{-i}(\hat{\sigma})$, this implies that $\hat{\sigma}_{-i}(\overline{a}_{-i})=1$, contradicting the fact that $\hat{\sigma}$ is properly mixed. $\square$

We now provide the proof for Proposition 8 in terms of $p^{\overline{a}_i}$, the case for $p^{\underline{a}_i}$ follows similarly. First, we begin with two necessary Lemmas.

Lemma 14. *Suppose that $\Gamma=(I,(A_i)_{i\in I},(u_i)_{i\in I})$ be a strict GSS and that there exsits a PMSNE $\hat{\sigma}$ such that for each player $i\in I$, $\underline{a}_i,\overline{a}_i\in BR_i(\hat{\sigma}_{-i})$. Then $BR(\overline{a})=\underline{a}$ and $BR(\underline{a})=\overline{a}$.*

Proof. Since each A_i is assumed finite, $\wedge BR_i$ and $\vee BR_i$ are continuous functions on A_{-i}, so that by Proposition 7, $\wedge BR_i(\overline{a}_{-i})=\underline{a}_i$ and $\vee BR_i(\underline{a}_{-i})=\overline{a}_i$. By Proposition 6 (3), since

$$1_{\{\overline{a}_{-i}\}}=(1-\varepsilon)\hat{\sigma}_{-i}+\varepsilon 1_{\{\overline{a}_{-i}\}}$$

with $\varepsilon = 1$, we then have

$$\wedge BR_i \left(1_{\{\bar{a}_{-i}\}}\right) = \underline{a}_i \in BR_i\left(\hat{\sigma}_i\right) \succeq \wedge BR_i\left(\hat{\sigma}_i\right) \succeq \vee BR_i\left(1_{\{\bar{a}_{-i}\}}\right) = \underline{a}_i$$

so that $BR_i\left(\bar{a}_{-i}\right) = \underline{a}_i$, the other case following similarly. $\qquad\square$

For $a_i \neq \bar{a}_i$, define $\Lambda^{\bar{a}_i}(a_i) = \left\{\lambda \in \triangle\left(A_{-i}\right) \mid l^{\bar{a}_i}(a_i, \lambda) = 0\right\}$, and

$$\bar{p}_i(a_i) = \begin{cases} 0, & \bar{a}_i \; strictly \; dominates \; a_i \\ \max_{\lambda \in \Lambda^{\bar{a}_i}(a_i)} \left(\lambda_i(\underline{a}_i)\right) & otherwise \end{cases}$$

Lemma 15. *If $a_i \in A_i$ is such that $\bar{a}_i$ does not strictly dominate a_i, then $\Lambda^{\bar{a}_i}$ is non-empty and compact. Therefore, $\bar{p}_i(a_i)$ is well-defined.*

Proof. Since $\bar{a}_i = BR_i(\underline{a}_{-i})$, $l^{\bar{a}_i}(a_i, 1_{\{\underline{a}_{-i}\}}) > 0$. Suppose for all $\mu \in \triangle(A_{-i})$, $l^{\bar{a}_i}(a_i, \mu) > 0$. Then for all a_{-i}, $l^{\bar{a}_i}(a_i, 1_{\{a_{-i}\}}) > 0$, contradicting the fact that $\bar{a}_i$ does not strictly dominate a_i. Thus for some $\bar{\mu}_i$, $l^{\bar{a}_i}(a_i, \bar{\mu}) \leq 0$. Consider the set of probability measures

$$\mathscr{Z} = \left\{\mu^\alpha = \begin{cases} \alpha + (1-\alpha)\bar{\mu}(\underline{a}_{-i}) & if \, a_{-i} = \underline{a}_{-i} \\ (1-\alpha)\bar{\mu}(a_{-i}) & if \, a_{-i} \neq \underline{a}_{-i} \end{cases} \mid \alpha \in [0, 1] \right\}$$

Note that $\mu^1 = 1_{\{\underline{a}_{-i}\}}$ and $\mu^0 = \bar{\mu}$, giving us $l^{\bar{a}_i}(a_i, \mu^1) > 0$, $l^{\bar{a}_i}(a_i, \mu^0) \leq 0$, and that $l^{\bar{a}_i}(a_i, \mu^\alpha)$ is a continuous function of α. By the intermediate value theorem, $\exists \bar{\alpha} \in [0, 1)$ such that $l^{\bar{a}_i}(a_i, \mu^{\bar{\alpha}}) = 0$. Therefore, $\Lambda^{\bar{a}_i}$ is non-empty.

To show compactness, suppose $(\mu^n)_{i=1}^\infty \subseteq \Lambda^{\bar{a}_i}$ is such that $\mu^n \to \mu$. By continuity, $l^{\bar{a}_i}(a_i, \mu) = \lim_{n\to\infty}\left(l^{\bar{a}_i}(a_i, \mu^n)\right) = 0$, thus $\mu \in \Lambda^{\bar{a}_i}$, so that $\Lambda^{\bar{a}_i}$ is closed. Since $\Lambda^{\bar{a}_i} \subseteq \triangle(A_{-i})$, which is compact, we have that $\Lambda^{\bar{a}_i}$ is compact also. $\qquad\square$

Proof. (Proposition 8) If $\bar{a}_i$ strictly dominates a_i, then by the definition of $p_i^{\bar{a}_i}(a_i)$, we have that $\bar{p}_i(a_i) = 0 = p_i^{\bar{a}_i}(a_i)$. Now suppose that $\bar{a}_i$ does not strictly dominate a_i.

It is first shown that $p_i^{\bar{a}_i}(a_i) \le \overline{p}_i(a_i)$. In order to do this, we show that for all $\mu' \in \triangle(A_{-i})$ such that $\mu'(\underline{a}_{-i}) \ge \overline{p}_i(a_i)$, $l^{\bar{a}_i}(a_i, \mu') \ge 0$. Because $p_i^{\bar{a}_i}(a_i)$ is defined as the lowest value that satisfies this property, the conclusion follows. Suppose $\mu' \in \triangle(A_{-i})$ is such that $\mu'(\underline{a}_{-i}) \ge \overline{p}_i(a_i)$, but $l^{\bar{a}_i}(a_i, \mu') < 0$. Since $\bar{a}_i = BR_i(\underline{a}_{-i})$, $l^{\bar{a}_i}(a_i, 1_{\{\underline{a}_{-i}\}}) > 0$. Consider the set of probability measures

$$\mathscr{Z} = \left\{ \mu_i^\alpha = \begin{cases} \alpha + (1-\alpha)\mu_i'(\underline{a}_{-i}) & if\, a_{-i} = \underline{a}_{-i} \\ (1-\alpha)\mu_i'(a_{-i}) & if\, a_{-i} \ne \underline{a}_{-i} \end{cases} \middle| \alpha \in [0,1] \right\}$$

and note that $\mu^1 = 1_{\{\underline{a}_{-i}\}}$ and $\mu^0 = \mu_i'$, giving us $l^{\bar{a}_i}(a_i, \mu^1) > 0$, and $l^{\bar{a}_i}(a_i, \mu^0) < 0$. Since $l^{\bar{a}_i}(a_i, \mu^\alpha)$ is a continuous function of α, by the intermediate value theorem $\exists \bar{\alpha} \in (0,1)$ such that $l^{\bar{a}_i}(a_i, \mu^{\bar{\alpha}}) = 0$. Since $\mu^{\bar{\alpha}}(\underline{a}_{-i}) = \bar{\alpha} + (1-\bar{\alpha})\mu'(\underline{a}_{-i}) > \mu'(\underline{a}_{-i})$, we have found a $\mu^{\bar{\alpha}}$ such that $l^{\bar{a}_i}(a_i, \mu^{\bar{\alpha}}) = 0$ but $\mu^{\bar{\alpha}}(\underline{a}_{-i}) > \mu'(\underline{a}_{-i}) \ge \overline{p}_i(a_i)$, contradicting the definition of $\overline{p}_i(a_i)$. Hence $p_i^{\bar{a}_i}(a_i) \le \overline{p}_i(a_i)$.

To show equality, suppose that $p_i^{\bar{a}_i}(a_i) < \overline{p}_i(a_i)$. We will show that for any μ'' such that $\mu''(\underline{a}_{-i}) > p_i^{\bar{a}_i}(a_i)$, we must have that $l^{\bar{a}_i}(a_i, \mu'') > 0$. Hence if $p_i^{\bar{a}_i}(a_i) < \overline{p}_i(a_i)$ is true, any μ_i' such that $\mu_i'(\underline{a}_{-i}) = \overline{p}_i(a_i)$ must be such that $l^{\bar{a}_i}(a_i, \mu'') > 0$, a direct contradiction to the existence of $\overline{p}_i(a_i)$. Note that this will also establish part 4 of Proposition 8, since A_i is finite. Let $\mathscr{A} = \underset{a_{-i}}{argmin}(\triangle u_i(\bar{a}_i, a_i, a_{-i}))$. Note that $\underline{a}_{-i}$ is not part of this set, for if it were, then for all a_{-i} we'd have by the fact that $\bar{a}_i = BR_i(\underline{a}_{-i})$ by Lemma 2, $\triangle u_i(\bar{a}_i, a_i, a_{-i}) \ge \triangle u_i(\bar{a}_i, a_i, \underline{a}_{-i}) > 0$, contradicting the fact that $\bar{a}_i$ does not strictly dominate a_i. Let μ be such that $\mu_i(\underline{a}_{-i}) = p_i^{\bar{a}_i}(a_i)$, and assign to all a_{-i} in $\mathscr{A}$ the probability $\frac{1-\mu_i(\underline{a}_{-i})}{|\mathscr{A}|}$. By the definition of $p_i^{\bar{a}_i}(a_i)$, we must have that $l^{\bar{a}_i}(a_i, \mu) \ge 0$. Now let μ' be such that $\mu'(\underline{a}_{-i}) > p_i^{\bar{a}_i}(a_i)$, and assign to all a_{-i} in $\mathscr{A}$ the probability

$\frac{1-\mu'_i(\underline{a}_{-i})}{|\mathscr{A}|}$. Note that we have[1]

$$l^{\bar{a}_i}(a_i, \mu') = \triangle u_i(\bar{a}_i, a_i, \underline{a}_{-i})\mu'_i(\underline{a}_{-i}) + \left(\frac{1 - \mu'_i(\underline{a}_{-i})}{|\mathscr{A}|}\right)\sum_{\mathscr{A}}\triangle u_i(\bar{a}_i, a_i, a_{-i}) >$$

$$\triangle u_i(\bar{a}_i, a_i, \underline{a}_{-i})\mu_i(\underline{a}_{-i}) + \left(\frac{1 - \mu_i(\underline{a}_{-i})}{|\mathscr{A}|}\right)\sum_{\mathscr{A}}\triangle u_i(\bar{a}_i, a_i, a_{-i}) = l^{\bar{a}_i}(a_i, \mu) \geq 0.$$

Finally, let μ'' be arbitrary but satisfying $\mu''(\underline{a}_{-i}) = \mu'(\underline{a}_{-i})$. Because $l^{\bar{a}_i}(a_i, \mu') \geq 0$ gives the smallest value for such probability measures, we have that $l^{\bar{a}_i}(a_i, \mu'') \geq l^{\bar{a}_i}(a_i, \mu') > l^{\bar{a}_i}(a_i, \mu) \geq 0$, or $l^{\bar{a}_i}(a_i, \mu'') > 0$. This establishes parts 1, 2, and 4 of Proposition 8.

For part 3 of Proposition 8, let $i \in I$ and consider any a_i not strictly dominated by $\bar{a}_i$. By above,

$$p_i^{\bar{a}_i}(a_i) = \max_{\substack{\lambda \in \triangle(A_{-i}) \\ l^{\bar{a}_i}(a_i, \lambda) = 0}} \left(\lambda(\underline{a}_{-i})\right)$$

Notice that for all $\lambda \in \triangle(A_{-i})$ such that $l^{\bar{a}_i}(a_i, \lambda) = 0$, we must have that $\lambda(\underline{a}_{-i}) < 1$. If not, then $\lambda = 1_{\{\underline{a}_{-i}\}}$. But by Lemma 14, since $\bar{a}_i = BR_i(\underline{a}_{-i})$ we must have that $l^{\bar{a}_i}(a_i, \lambda) > 0$. This contradicts $l^{\bar{a}_i}(a_i, \lambda) = 0$, and thus we must have that $p_i^{\bar{a}_i}(a_i) < 1$. Since each A_i is finite, it follows that $p_i^{\bar{a}_i} < 1$. The case for $p^{\underline{a}_i}$ following similarly. $\qquad\square$

---

[1] Note that $\sum_{\mathscr{A}}\triangle u_i(\bar{a}_i, a_i, a_{-i}) \leq 0$: If $\triangle u_i(\bar{a}_i, a_i, a_{-i}) > 0$ for each a_{-i}, the fact that $\bar{a}_i$ does not strictly dominate a_i is violated. Thus for some a_{-i}, $\triangle u_i(\bar{a}_i, a_i, a_{-i}) \leq 0$, and the conclusion follows by the definition of $\mathscr{A}$.