1994

Three Essays On Bargaining

Murali K. Agastya

Three Essays on Bargaining

by
Murali K. Agastya

Department of Economics

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
August 1993

National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-315-90537-9

Canada

Three Essays on Bargaining

Murali K. Agastya. Ph.D

The University of Western Ontario, 1993

Department of Economics

Advisor: Philip J. Reny

Abstract

In the first two chapters, a non-negative function defined on the class of subsets of a finite set of players (or factors) describes the technology for produc ing a single good. Given this aggregate data, the problem of allocation of surplus among individual players (or factors) is studied in two different models.

In Chapter 1, an axiomatic approach is adopted to construct an allocation rule that is immune to *positive monotone transformations* of the players' utilities. Under this rule, each player is paid a weighted average of his (or her) marginal contributions to various coalitions. In fact, these weights coincide with the Shapley weights. The model also provides a proper framework for interpreting the Shapley value as the ex-ante evaluation of a conflict situation.

Chapter 2 studies an *evolutionary bargaining model* in which myopic players with limited memory make simultaneous demands, naively based on precedent. Necessary and sufficient conditions are provided under which the long-run equilibria coincide with the core allocations. Refining the set

of equilibria by allowing for the possibility of *mistakes*, it is shown that the unique limiting equilibrium allocation maximizes the product of players' utilities subject to being in the core of the technology.

Chapter 3 studies the effect of different communication possibilities on the coalitional stability of bargained outcomes. Formally, a graph with the set of players as its vertices describes the communication possibilities. A coalition can be a threat if and only if it is connected. It is then shown that, for a large class of environments, stability with respect to coalition formation ensues if each connected coalition is a tree. Conversely, if there is a connected coalition that is not a tree, there are bargaining situations in which no outcome is stable.

Three Essays on Bargaining

by
Murali K. Agastya

Department of Economics

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
August 1993

Abstract

In the first two chapters, a non-negative function defined on the class of subsets of a finite set of players (or factors) describes the technology for producing a single good. Given this aggregate data, the problem of allocation of surplus among individual players (or factors) is studied in two different models.

In Chapter 1, an axiomatic approach is adopted to construct an allocation rule that is immune to *positive monotone transformations* of the players' utilities. Under this rule, each player is paid a weighted average of his (or her) marginal contributions to various coalitions. In fact, these weights coincide with the Shapley weights. The model also provides a proper framework for interpreting the Shapley value as the ex-ante evaluation of a conflict situation.

Chapter 2 studies an *evolutionary bargaining model* in which myopic players with limited memory make simultaneous demands, naively based on precedent. Necessary and sufficient conditions are provided under which the long-run equilibria coincide with the core allocations. Refining the set of equilibria by allowing for the possibility of *mistakes*, it is shown that the unique limiting equilibrium allocation maximizes the product of players' utilities subject to being in the core of the technology.

Chapter 3 studies the effect of different communication possibilities on the coalitional stability of bargained outcomes. Formally, a graph with the set of players as its vertices describes the communication possibilities. A coalition can be a threat if and only if it is connected. It is then shown that, for a large class of environments, stability with respect to coalition formation ensues if each connected coalition is a tree.

Conversely, if there is a connected coalition that is not a tree, there are bargaining situations in which no outcome is stable.

Acknowledgements

Whatever may be the shortcomings of this work, they would have been greater (to an exponent) but for the sharp, incisive comments and valuable suggestions of both Phil Reny and Arthur Robson. Their constant encouragement combined with (perhaps an overdose of) scepticism of various results in earlier versions has, I believe, resulted in a substantial improvement. I express my deepest gratitude to them both. I am also indebted to Abhijit Sengupta and Motty Perry who have also contributed to this thesis through their several useful comments and encouraging remarks. I am grateful to Motty for a memorable course on Bargaining Theory that he taught in fall of 1989. A very special thanks is due to Ig Horstmann. He has been a source of constant encouragement, help and advice. Neils Anthoniesen and Brian Rivard were a kind audience to my several initially confused thoughts. Their patience has helped me clarify my ideas. I thank them both.

I would also like to extend my thanks to all the departmental staff, particularly Jane McAndrew at the Reference Center and Graduate Secretary Sue Brown, for their help in various matters. Amazingly, Jane manages to smile at me and is enormously helpful despite my phenomenal forgetfulness (and laziness) in returning library material. Sue has typed innumerable letters for me, some of which work like a charm with Immigration Canada and USIS.

And of course, I thank my friends Benoit, Brad, Brian, John, Dave Andolfatto, Dave Koop, Dimitri, Ian, Toni and Glen for creating and sharing the necessary social diversions.

To Mom, Dad, Gautami and Geethika

Contents

Chapter 1

Multiplayer Bargaining Situations: A Decision Theoretic Approach

1.1 Introduction

Since the pioneering work of John Nash in the fifties, the formal analysis of the bargaining problem has evolved in two different directions. One is axiomatic (e.g., Nash (1950); Kalai and Smordinsky (1975).). In this approach the bargaining process is only implicit. One tries to characterize the negotiated outcome through a set of axioms without formally modelling the process of settlement.

The other approach is to make explicit the strategic aspects of the bargaining problem. Here, the negotiation procedure is described explicitly as a non-cooperative game and the predictions are identified with its equilibrium points. For many years following Nash's seminal contributions, the axiomatic approach was the predominant one. In recent years, following the work of Rubinstein (1982), the perspective has shifted. The last decade has seen a flurry of research activity in non-cooperative bargaining theory.

A central insight obtained from this literature is the importance of the *exact spec-*

ification of the bargaining process in determining the outcomes. It was discovered early on that even slight changes in the bargaining procedure can yield quite different results. This insight questions the validity of the axiomatic approach which endeavours to find a fruitful solution concept that is valid for a variety of negotiation procedures. At the same time, the excruciating detail with which the rules of the game need to be specified casts doubt on the appropriateness of the predictions obtained as solutions to non-cooperative procedures.

If the rules of the game are well laid out for the players, a non- cooperative game-theoretic model is appropriate. Although one does get to know the rules of the game eventually, often a choice between various bargaining situations has to be made *before* the rules are known. If one subscribes to the basic tenet that agents make consistent choices even in these situations, one is left with (at least to my mind) no other alternative than to rely on a few basic rules of behaviour, i.e., resort to the axiomatic approach in predicting their choice.

In this chapter, I adopt the axiomatic approach and analyze the choice of a single individual amongst environments where side- payments are possible in a single good. These environments are referred to as *bargaining situations*. Loosely speaking, a bargaining situation consists a set of n players and a non- negative and super-additive function f defined over the class of all coalitions of this set. Here, $f(S)$ is the quantity of a certain infinitely divisible and transferable commodity available to a coalition S. I postulate a binary relation $\succeq_i$ on the class of all bargaining situations. A player anticipates how much of the good he will get should he find himself in any of the bargaining situations. Thus by $f \succeq_i g$, we mean that player i expects to obtain at least as much of the good in situation f as in g. The strategy of this paper is to impose restrictions on this binary relation. A representation of $\succeq_i$ is the *utility* of a certain bargaining situation for the player in terms of the underlying good itself. The main result of the paper is as follows. For a bargaining situation f, $[f(S) - f(S \setminus i)]$ is the marginal contribution of player i to the coalition S. Under intuitive axioms I show that in choosing between bargaining situations, an agent maximizes a weighted average of his marginal contributions to all the possible coalitions. Theorem 1.3 shows that these weights in fact coincide with the *Shapley* weights [1].

That the weights obtained in Theorem 1.3 coincide with the Shapley weights may lead one to interpret this representation theorem as a mere recharacterization of the Shapley value itself. Theorem 1.3 is in fact more general. The initial construction of the Shapley value is for games with transferable utility (TU). The assumption of TU requires the existence of a single composite good for making side-payments *and* the utility of this good be linear for all the players, with the same scales. While I do assume that side-payments are made in a single good, I do not assume TU. In fact, the exact utility for the good is irrelevant. The only thing that matters is the implicit assumption that the players prefer more of the good to less of it.

This irrelevance of the exact specification of the players' utilities for the underlying good is interesting. For, a common feature of both the strategic and the axiomatic approaches is the assumption that only *utilities* for the final outcome and not the outcomes themselves that matter. In the words of Nash: " What the actual courses of action are among which the individuals must choose is not regarded as essential

[1] The literature on Shapley value is enormous. See Roth (1988) for a list of references.

information ... Only the attitudes (like or dislike) of the two individuals towards the ultimate results are considered." This preoccupation with the utility and not the outcomes is especially important because the ultimate result in both the approaches is quite sensitive to the exact representation of preferences. In fact, the actual outcomes are invariant only up to an affine transformation of the players' utilities. For this reason, it is often asserted that bargaining theory must involve interpersonal comparisons of utility.

Restriction of players' preferences to those that are invariant only up to an affine transform is perhaps a reasonable assumption when bargaining over indivisible goods is carried out by means of lotteries. However, in many other instances this is not the case. Often, we do have access to an (almost) perfectly divisible unit (such as money), to compensate for the lack of divisibility in others. If one were to consider induced preferences for this good, it is not clear why one should not allow for a much larger class of preferences. Indeed, if we wish to construct a theory that is devoid of interpersonal comparison of utilities, one must ensure that the outcome is invariant under all *positive monotone transforms* of the utility for the underlying good.

The bargaining rule naturally implied by Theorem 1.3 is trivially immune to *ordinal transformations of the players utility for this good.* Thus the model here provides a resolution to the *Bargainers' Paradox.* (See Shapley (1969) or Shubik (1982).) The above bargaining solution also determines a Pareto-optimal allocation without actually imposing it as an axiom. Moreover, given that only side-payments are assumed, the model characterizes a non-transferable utility (NTU) value for the above physical environments. And finally, the approach adopted in this paper provides a framework for a proper interpretation of the value as the ex-ante evaluation of a game. To ease exposition, an elaborate discussion of various issues raised in this paragraph is deferred to Section 1.4. Section 1.4 also compares the present work to the relevant literature. The model and the axioms are laid out in Section 1.2. Section 1.3 contains all the results. Formal arguments, for the most part, appear in Appendix A.

1.2 The Model

Let $N = \{1, 2, \ldots n\}$ be a finite set denoting the set of players. A *coalition* is a non-empty subset of N. Given coalitions $S, T, U, \ldots$, let $s, t, u \ldots$ denote the number of players in the respective coalitions. A non-negative function, f, defined on the class of all coalitions is said to be *super-additive*, if for any two disjoint coalitions S and T, $f(S \cup T) \geq f(S) + f(T)$. As is usual, let 2^N denote the class of all subsets of N. Furthermore, let Π_i denote the class of all coalitions that contain player i and Π_{-i} denote the class of all coalitions that do not contain player i.

Definition 1.1 *A n-player bargaining situation is a non-negative function $f : 2^N \longrightarrow \Re_+$ such that*

1. *f is super-additive*

2. *$f(\phi) = 0$*

Let $\mathcal{G}$ denote the class of all bargaining situations. We regard $\mathcal{G}$ as a subset [2] of $\Re^{2^N - 1}$. By way of interpretation, given a bargaining situation f and a coalition S,

[2] Hence the addition of two bargaining situations or multiplication of a bargaining situation by a scalar are well defined.

$f(S)$ is the quantity of a single good, say money, that is available to the members of a coalition S if members of S^c leave [3]. I assume that side- payments are made in this good. It is worthwhile to re-emphasize that this is not the same as the assumption of transferable utility.

A typical player i, is completely characterized by a complete, transitive and reflexive relation $\succeq_i$ defined on $\mathcal{G}$. Such a relation is called a preference relation. Thus, given two bargaining situations f and g, by $f \succeq_i g$ should be read as " player i weakly prefers the prospect of being in situation f to the prospect of being in situation g ". The symbols $\succ_i$ and $\sim_i$ have the usual interpretation of strict preference and indifference. Certain situations play a key role in the analysis. Let f_ϕ denote the *null* situation, that is $f_\phi(S) = 0$, for all $S \subseteq N$. Formally, this situation corresponds to the origin of $\Re^{2^N-1}$ of which $\mathcal{G}$ is a subset. In f_ϕ, no coalition is productive. At the other extreme, there are situations in which the presence of a certain coalition is necessary and sufficient for production. Formally, given a coalition S, let δ_S the following situation:

$$\delta_S(T) = \begin{cases} 1 & \text{if } S \subseteq T \\ 0 & \text{otherwise} \end{cases}$$

In the situation δ_S, the members of S are solely responsible for any surplus that is available for a coalition T. For a real number $x \geq 0$, I will refer to $x\delta_S$ as a *pure bargaining situation* (PBS) of size x for S. A PBS is said to include player i if it is a PBS for a coalition S that contains i. A PBS is easy to analyze. Hence the strategy will be to impose intuitive axioms on the $\succeq_i$ between various PBS and then extend

[3] See Hart and Mas- Colell (1992) for a discussion on this issue.

this on to more complex bargaining situations.

The first axiom is a continuity assumption. It is well known that such an assumption is a minimal requirement to ensure the existence of any real representation.

Axiom 1.1 *(Continuity): The graph of* $\succeq_i$ *is closed, for all* $i = 1, 2, \ldots N$.

Consider δ_S and $x\delta_S$ where x is very close to one. Each is a PBS for S and they are approximately of the same size. If a player i who is *not* in S is asked to choose between the above two situations, it seems reasonable to assume that he will be indifferent between the two. This is the content of the next axiom.

Axiom 1.2 *(Nullity) Suppose* $S \in \Pi_{-i}$. *There exists an* $x \neq 1$ *such that* $\delta_S \sim_i x\delta_S$.

REMARK: A game with TU, say v, is said to be null for player i if for every coalition S, $v(S) = v(S \setminus i)$. In a related context, Roth (1988) requires that if any two games v and w are null for player i, then $v \sim_i w$. Since v_ϕ, the game analogous to f_ϕ is null for a player, this assumption basically assumes that in a null situation, a player obtains nothing. Although the implication of the *Nullity* axiom stated here is the same, i.e. players outside S in a pure- bargaining situation for S evaluate it as if they obtained nothing, I do not assume this outright. Technically the form of the assumption here is weaker.

Now consider $x\delta_S$ and $y\delta_{N\setminus S}$ with $x > y$. Suppose player i is in S. The next axiom requires that player i strictly prefer a larger PBS (that is $x\delta_S$) that includes him to a PBS (that is $y\delta_{N\setminus S}$) of smaller size that does not include him.

Axiom 1.3 *(Productivity) Suppose* $S \in \Pi_i$. *Then for any* $x > y \geq 0$, $x\delta_S \succ_i y\delta_{N\setminus S}$.

An alternative interpretation is as follows: Player i has to choose between an organization $S \setminus i$ and an organization $N \setminus S$. Player i's presence will enable $S \setminus i$ produce x units and nothing otherwise. On joining $N \setminus S$, the production will continue to be y units, which is strictly less than x. *Productivity* requires that player i choose $S \setminus i$.

Suppose player i prefers f to g. Now suppose player i is informed that in *both* f and g a coalition S is able to produce an extra unit of output. That is, he is now asked to choose between $f + \delta_S$ and $g + \delta_S$. The next axiom states that he does not change his preference.

Axiom 1.4 *(Strateqic Equivalence) Suppose* $f \succ_i g$. *Then for any* $x \geq 0$, $f + x\delta_S \succ_i g + x\delta_S$.

Axioms Axiom 1.5 and Axiom 1.6 below are of a different nature compared to axioms Axiom 1.1-Axiom 1.4. The first four axioms are concerned with the preferences of a single player. In the next two axioms I tie in the preferences of the various players. Axiom 1.5 is a simple *Symmetry* assumption. Such an assumption is natural in many instances.

In δ_i, one expects player i to obtain the 1 unit. Consider a PBS that includes players i and j. The Symmetry assumption requires that if the above PBS is large enough for a player i to prefer it to δ_i (and hence player i expects to obtain more than one unit), then the PBS is large enough for player j to at least weakly prefer it to δ_j.

Axiom 1.5 *(Symmetry): Suppose $i, j \in S$. For any real number x, if $x\delta_S \succ_i \delta_i$, then $x\delta_S \succeq_j \delta_j$*

REMARK: In value theory, typically a Symmetry assumption is made (See Shapley (1953)). The *Symmetry* assumption made here is specific to a particular coalition.

To introduce the next axiom, it is useful to introduce some definitions. To understand the motivation for the terminology, it is useful to recall that in δ_i, one expects player i to keep the entire surplus for himself.

Definition 1.2 *Suppose $i \in S$ and $x < |S|$. Player i is said to be an optimist if $x\delta_S \succ_i \delta_i$*

Similarly, one defines an pessimist as follows:

Definition 1.3 *Suppose $i \in S$ and $x > |S|$. Player i is said to be a pessimist if $\delta_i \succ_i x\delta_S$.*

Suppose there are only two players, say A and B. Furthermore, suppose that, in a loose and informal sense, that $(\succeq_A, \succeq_B)$ are common-knowledge between the two players. That is, both players know the preferences of the other player, both players know that both players knows the preferences of the other player, ... and so on *ad infinitum*. Now consider a PBS of size x for the coalition AB, where $x < 2$. Suppose that, at the chosen x if both A and B are optimists. So B expects to obtain more than one unit $x\delta_{AB}$ and so does A. Moreover both players know that each of them expects to get more than one unit, *and* both players know that both players know that both of them expects more than one unit in $x\delta_{AB}$. Since there is not enough surplus to go around, this cannot happen.

Now suppose that $x > 2$ and player B is a pessimist. So B expects to obtain less than one unit. Now, both players prefer more of the good to less of it. If B is a pessimist, it must be because he expects player A to obtain more than one unit. If fact, given our assumption of common-knowledge, not only does B expect A to obtain more than one unit, he *knows* that A obtains more than one unit. Hence when A has to choose between $x\delta_{AB}$ and δ_A, he will opt for the former. Axiom 6 formalizes this intuition.

Axiom 1.6 *(Consistency) For a PBS of size x for a coalition S,*

1. *If $x < |S|$. there exists a player $i \in S$ who is* not *an optimist.*

2. *If $x > |S|$, there exists a player $i \in S$ who is* not *a pessimist.*

1.3 Representation Theorems

In this section, I state all Theorems. I provide a sketch of the proof for some of them. The sketch also highlights the role of the various axioms. All formal arguments are in the Appendix A.

Theorem 1.1 *Consider a preference relation $\succeq_i$ on $\mathcal{G}$. The following are equivalent:*

1. *The preference relation $\succeq_i$ satisfies Axiom 1.1-Axiom 1.4.*

2. *There is a unique set of positive weights* $\{p(i,S) \ : \ S \in \Pi_i\}$ *, such that* $\sum_{S \in \Pi_i} p(i,S) = 1$ *and*

$$f \succ_i g \iff \sum_{S \in \Pi_i} p(i,S)[f(S) - f(S \setminus i)] > \sum_{S \in \Pi_i} p(i,S)[g(S) - g(S \setminus i)].$$

PROOF: (Sketch) It is easy to see that the second statement implies the first. Suppose now that $\succeq_i$ satisfies 2. We need to show that it satisfies 1. Towards this end we show that there is an extension $>_i^*$ of $\succ_i$ to the whole of $\Re^{2^N-1}$ such that for any three vectors $f, g,$ and$h \in \Re^{2^N-1}$, $f >_i^* g$ if and only if $f + h >_i^* g + h$. This is Lemma A.2 in the Appendix A. In particular, Lemma A.2 implies that for any two bargaining situations f and g, $f \succ_i g$ if and only if $f - g >_i^* 0$. Now, let E_i be the set of all vectors that are strictly better than 0 under $>_i^*$. Then clearly, $f \succ_i g$ if and only if $(f - g) \in E_i$. It can be shown (Lemma A.5) that E_i is an open and convex subset of $\Re^{2^N-1}$ that does not contain the origin. Hence there is a hyperplane that strictly separates E_i and the origin. It is now easy to obtain the following lemma:

Lemma A.6 *There is a set of weights* $\{\lambda(i,S) \ : \ S \subseteq N\}$ *such that for any* $f, g \in \mathcal{G}$*,*

$$f \succ_i g \iff \sum_{S \subseteq N} \lambda(i,S) f(S) > \sum_{S \subseteq N} \lambda(i,S) g(S).$$

Now, it may be verified that for any f,

$$\sum_{S \subseteq N} \lambda(i,S) f(S) = \sum_{S \in \Pi_i} \lambda(i,S)[f(S) - f(S \setminus i)] + \sum_{S \in \Pi_{-i}} [\lambda(i,S) + \lambda(i, S \cup i)] f(S).$$

I use Axiom 1.2 to conclude that $\lambda(i,S) + \lambda(i, S \cup i) = 0$, if $S \in \Pi_{-i}$. This is Lemma A.7. Finally Lemma A.8 shows that $\lambda(i,S) > 0$, if $S \in \Pi_i$. Now, set $p(i,S) = \frac{\lambda(i,S)}{\sum_{T \in \Pi_i} \lambda(i,T)}$.

The uniqueness of the set of weights follows from the fact that the ordering is complete and the weights are normalized to add up to one. This completes the proof.

Armed with Theorem 1.1, I now proceed to characterize the weights obtained therein more completely by imposing other axioms. Theorem 1.2 below, states the implication of the *Symmetry* axiom in addition to those stated in Theorem 1.1.

Theorem 1.2 *Consider n preference relations $\succ_i, i = 1, 2, \ldots n$, defined on $\mathcal{G}$ such that each of them satisfies Axiom 1.1-Axiom 1.5. Then the weights obtained in Theorem 1.1 satisfy the following:*

If $S \in \Pi_i \cap \Pi_j$, then $p(i, S) = p(j, S)$.

PROOF: See Appendix A.

Note that our assumption of *Symmetry* is specific to a given coalition. That is, even under the *Symmetry* axiom, player i's evaluation of a coalition with player k (and not containing player j) may be different from player j's evaluation of a coalition with player [4] k (and not containing player i).

The assumption of *Consistency* (along with the assumptions of Theorem 1.2) is sufficient to pin down weights exactly. Theorem 1.3 details the claim:

[4]Consider the following axiom: AXIOM* 5: (Anonymity) Let S be a coalition that does not contain the three distinct players i, j and k. For any real number x, if $x\delta_{S\cup i,j} \succ_i \delta_i$, then $x\delta_{S\cup i,k} \succ_i \delta_i$. Replacing the *Symmetry* axiom with the *Anonymity* axiom, one can show that, for each player, the weights obtained in Theorem 1.1 must depend only on the size of various coalitions. However, different players may yet have different *sets* of weights. If one imposes *Anonymity* along with the *Symmetry*, the weights coincide for all the players and depend only on the size of the various coalitions.

Theorem 1.3 *Consider n preference relations, $\{\succeq_i : i = 1, 2, \ldots n\}$ defined on $\mathcal{G}$. Then the following two statements are equivalent:*

1. *Each $\succeq_i$ satisfies Axiom 1.1-Axiom 1.6.*

2. *For any two bargaining situations f and g, $f \succ_i g$ if and only if*

$$\sum_{S \in \Pi_i} \frac{(n-s)!(s-1)!}{n!}[f(S) - f(S \setminus i)] > \sum_{S \in \Pi_i} \frac{(n-s)!(s-1)!}{n!}[g(S) - g(S \setminus i)]$$

PROOF: See Appendix A.

1.4 Discussion

This section elaborates on the issues raised in the last paragraph of the Introduction.

There have been several formulations of the Shapley value for TU games besides Shapley's own seminal contribution. (See for e.g. Young (1988), Hart and Mas-Colell (1987), Chun (1989) and Roth (1977).) A typical characterization involves imposing a set of axioms on a cardinal function and showing that the Shapley value is the unique function that satisfies these axioms. The different axioms in these characterizations yieid alternative interpretations to the value. Young (1988) is motivated by ideas of distributive justice while Hart and Mas-Colell (1987) is motivated by notions of cost-allocation. However, the original motivation for the value (see Roth (1977)) is to regard it as the utility of a game.

Every one of the formulations mentioned above, except Roth (1977), imposes *Efficiency* as an axiom. Efficiency requires that the sum of the values of a game for various players add up to the worth of the grand coalition. If one were to interpret the value as the utility of a game, efficiency [5] requires summing up the utilities of different players. It is then not clear as to what interpersonal comparisons of utility are being made [6]. For the present model, Theorem 1.3 provides the utility of a game situation. Nowhere in my construction do I add the utilities or the representations of different players. In fact, axioms Axiom 1.1-Axiom 1.4 are restrictions on the preferences of a single player. It is only Axiom 1.5 and Axiom 1.6 that make a direct interpersonal comparison. While Axiom 1.5 is a symmetry assumption, Axiom 1.6 is a statement on the knowledge of the model among various players rather than a direct comparison of utility. Hence, Theorem 1.3 provides a proper basis for interpreting the value as the utility of a conflict situation.

Roth (1977) also provides a framework in which he obtains the Shapley value without using efficiency as an axiom. Roth (1977) is concerned with the preferences of a single player who is uncertain about the TU game to be played and the position or the role that the player may be in. Allowing for lotteries over this space of uncertainty, he assumes that the player satisfies all the standard axioms of expected utility theory. Thus he obtains a von Neumann- Morgenstern (VNM) utility function, say θ. If the realized outcome of the lottery is a game v with the position i, then $\theta(v,i)$, is the VNM utility. Neutrality to "strategic" and "ordinary" risk is shown to imply that $\theta(v,i)$ is in fact the Shapley value.

[5]Another important axiom used in most of the constructions is *Additivity*, that is the value of the sum of two different games add up to the sum of the values of the games. This assumption too is not meaningful in the present context.

[6]Nonetheless, some authors have associated the Shapley value with classical utilitarianism (See Moulin (1988)), as it seems to maximize the sum of utilities of the various players.

The present work differs from that of Roth (1977) in two important ways. In the model presented in Section 1.2, there is no uncertainty about the games being played. Players are involved in a conceptual exercise of ranking one bargaining situation over another but never lotteries over bargaining situations. Consequently attitudes to risk are irrelevant. Second, since the Shapley value is obtained only as a VNM utility functional in Roth (1977), the representation is unique only up to a positive linear transformation. On the other hand, the representation obtained in Theorem 1.3 is ordinal; it is unique up to any positive monotone transform.

Theorem 1.3 also determines a bargaining solution when side- payments are made in a single good. Players obtain a weighted average of their marginal contributions with the weights being the Shapley weights. There are two interesting features of this bargaining solution. First, we obtain a Pareto-optimal allocation without resorting to efficiency as an axiom.

Second, for two player games, the Shapley value coincides with another widely studied sharing rule; the Nash Bargaining solution (see Nash (1950)). One of the axioms in the Nash bargaining solution requires that the physical outcome corresponding to the agreed utility payoff be invariant to positive linear transforms of the players' utilities over the underlying set of alternatives. This invariance of the physical outcome with respect to positive linear transforms is a natural requirement if one were bargaining over the exchange of goods by means of lotteries. Decision- theoretically equivalent utility functions must yield the same physical outcome. However, in the present context where the players are bargaining over a single infinitely divisible good, rather than lotteries over (perhaps indivisible) goods, one might naturally require that the physical outcome be invariant to ***ordinal transforms*** of the players' utility for the underlying good.

However, it is shown in Shapley (1969) [7], that at least for two person games it is impossible to construct a theory where the physical outcome is invariant to ordinal transforms of utility. This is sometimes referred to as the ***Bargainers" Paradox***. In fact, in the above paper, Shapley goes on to assert that:

> Interpersonal comparability of utility ... enters naturally - and, I believe, properly - as a nonbasic, derivative concept playing an important if sometimes hidden role in theories of bargaining...

The problem pointed out in Shapley (1969) appears since axiomatic bargaining theory focuses on the shapes of the utility possibility frontiers and other such geometric criteria rather than the underlying physical or economic environment. This paper studies a restricted yet an interesting class of physical environments. The procedure adopted here yields a bargaining solution that is trivially immune to ordinal transforms of utility of the underlying good. In fact, this is perhaps the central contribution of the paper.

A by-product of this paper is an NTU value when side-payments are possible only in a single good. The general NTU value was formulated in Harsanyi (1963) followed by Shapley (1969). (See also Aumann (1985b) and Hart (1985).) The NTU value has been a subject of considerable controversy in the last decade or so. This controversy follows certain puzzling and counter-intuitive predictions the value offers for examples

[7] See also Chapter 5, Shubik (1982)

presented in Roth (1980). (See also Aumann (1986) and Roth (1987)). In particular, concern has also been expressed regarding two related issues: the interpersonal comparisons of utility implicit in the NTU value and how one might accommodate the various interpretations that are accorded to the value with transferable utility in the case of NTU. It is clear from the preceding discussion that the model presented here answers these concerns for the environments considered therein.

1.5 Conclusion

The primary aim of the paper has been to see how a player evaluates certain conflict situations in the presence of other rational players. In the conflict situations that we consider the evaluation is a weighted average of his marginal contributions to the various coalitions. The interpersonal comparisons made are minimal and in our opinion very natural. Three central insights have been obtained by studying the preference ordering over bargaining situations. First, we have a proper framework for interpreting the "Shapley value" as an ex-ante evaluation of a conflict situation. Second, an extension of the original interpretation of the value to the NTU case. And most significantly, a bargaining solution invariant to ordinal transforms of utility for the underlying good.

Several extens· ·ns can be suggested. All the results in this paper hinge very critically on *Strategic Equivalence.* A weakening of this axiom would be of considerable interest. One expects that any such weakening would lead to a loss of separability between the weighting measure and the bargaining situation.

Another important extension would be to allow for side-payments in several goods.

This extension seems rather complicated. This is because with several goods, how the strategic position of a player changes when endowed with more of some of the goods is not clear.

Chapter 2

An Evolutionary Bargaining Model

2.1 Introduction

Analysis of strategic interaction normally requires either the assumption of common-knowledge of preferences or at the very least that the agents have a prior on the opponents preferences. These assumptions are certainly suspect when analysing markets with several anonymous participants. One might argue that the Walrasian model is better suited to analyze such situations. Nonetheless in some markets, such as the housing and the used car markets, although the players are relatively anonymous, there is some scope for strategic interaction. In these situations, the prior a player may have regarding the utilities of the other participants could be quite diffuse to be of any value.

Individuals however, either through their friends who have been in such situations before or through various magazines and newspapers, do have access to information on how their opponents may have acted in similar situations in the past. Since there is a lack of knowledge about the exact preferences of the opponents, it is not an unreasonable hypothesis that agents rely on history to devise an optimal strategy.

In a recent paper, Young (1993b) studies a model in which two players sample a fraction of history and demand a share of the surplus simultaneously. The split is enforced if and only if their demands are compatible. Different agents, drawn from a finite population, play this game repeatedly. It is then shown that any efficient split can be sustained as a long-run equilibrium. However, if one allows for the possibility of mistakes in players' responses, the set of divisions that one observes is considerably smaller. In fact, when the probability of making mistakes is small, a single division will be observed most of the time. This division is close to a generalized Nash bargaining solution and the shares depend on the amount of information that each player has.

An interesting feature of the above model is the following: Suppose that a players' type is completely specified by his utility for the underlying good and the extent of his information of the past. If both agents are drawn from a set of common types, the unique division of the surplus that is observed most often is a fifty-fifty split, regardless of which two types are matched.

While the fifty-fifty split obtained by Young is interesting, it is perhaps not very surprising in the following sense. If the common set of types is a singleton, then all players are symmetric in terms of their preferences over the surplus (and the information regarding the past). Since any solution concept that is unique treats symmetric players exactly the same, the outcome follows.

In general, the final allocation in most allocation mechanisms depends on two factors. First it depends on the preferences of the players over the allocations. Second, it depends on the strategic position of the players, usually manifest in their endowments of their contribution to the various coalitions. In the model of Young, while the players are allowed to be different in terms of their preferences, they are symmetric with respect to how much they contribute in physical terms; Two players have a pie to share if they strike a bargain or they both get zero otherwise. It is then of interest to study the allocation of scarce resources when players make demands optimally, based only on precedent.

Section 2.2 studies a variant of the model due to Young (1993b). There is a technology for producing an infinitely divisible good using the services of at least two factors of production. At each date a representative for each of the factors is drawn at random from a finite pool of representatives. One representative for each of the factors are jointly matched. They simultaneously make wage demands. The representatives are referred to as players.

The technology indicates the contribution of a player to various coalitions. Players have varying preferences over this good. Thus, players are allowed to be different in terms of their physical contributions to the various coalitions (or equivalently endowments) as well as their preferences for various goods.

The main results are as follows. If the technology is convex (i.e., displays increasing returns to the factors,) and one allows for the possibility of mistakes, the model predicts a unique outcome (in the limit) that maximizes the product of players' utilities for the surplus, subject to being in the core of the technology. Of course, if an all round equal split is in the core, this is the outcome provided the players are drawn from a common set of types.

Consideration of examples without a convex technology generates some interesting results. For instance, consider the technology in which the first factor can produce a hundred units of surplus with the help of one or both of the other factors. Then, while the competitive solution would predict that the wages for the first factor must be a hundred and the other two zero, the prediction here is different. If all three players have the same information and linear utilities the first factor obtains two thirds of the surplus while the other two factors obtain a sixth each, in expectation. If the utilities display constant absolute risk aversion, the return to player one is higher. Thus, unlike in Young (1993), the outcome is quite sensitive to the exact specification of the utilities of the players. The problem of existence of equilibria without convex technologies is also discussed by means of an example.

Like in most models that study naive behaviour (See for e.g., Foster and Young (1990), Kandori *et.al* (1993), Young (1993a and 1993b)), I resort to the possibility of players making mistakes in refining the set of equilibria. The model studied here differs from all these models in that players are restricted to make only small trembles. Intuitively, a small tremble is an action that is close to a best response in the physical sense. This restriction creates a technical problem. The stochastic process induced by the play of the game is not strongly ergodic, even after allowing for mistakes. Since the refinements are typically based on the convergence properties of the invariant distributions of the perturbed process, they are not immediately applicable.

Binmore (1987) and Carlsson (1991) analyze perturbations of the original Nash demand game and study the properties of non- cooperative equilibria of the perturbed game as the perturbations become very small. They show that the Pareto optimal Nash equilibria of the perturbed game converge to the original Nash Bargaining solution. The model presented below (as does Young (1993b)) differs from the above models in that they presuppose the equilibrium of both the original as well as the unperturbed game. In this model, players reach a unique long-run equilibrium as the likelihood of mistakes becomes very small, without any knowledge of the utilities of other players.

For the most part, attention is restricted only to the case of three players. The proofs, for the most part, can be generalized in a natural way for arbitrary (but finitely many) players. This is not done with rigour for all the results because the notation conceals more than what the increased generality reveals. However, brief pointers that help towards this generalization are provided.

The rest of the paper is organized as follows. Section 2.2 presents the basic model. The notion of a convention is presented in Section 2.3. Section 2.4 onwards, attention is restricted only to the case of three players. In Section 2.4 presents conditions for the convergence in probability to a convention. Section 2.5 then embarks on the problem of refining the set of conventions. Section 2.6 compares the results obtained here to those of Young (1993b) and the competitive outcomes. A short discussion in Section 2.7 concludes the chapter.

2.2 The Model

Time is discrete, $t = 1, 2, \ldots$. At each date t, a single good y can be produced using the services of at least two of n factors of production denoted by the set $N = \{1, 2, \ldots n\}$. The technological possibilities are described by a non-negative function f defined on the class of all coalitions of N. By way of interpretation, $f(S)$ is the total quantity of a single good that produced using the services of the factors in S alone. I will assume the each $f(S)$ is a rational number and $f(i) = 0$. Since f takes only finitely many values, for a given f, we can redefine the units in which the good is being measured such that $f(S)$ is a positive integer for each $S \subseteq N$.

At each date, a representative for each of the factors is drawn at random from a finite class of representatives for each factor. We will refer to them as players. At a date t, player i demands a wage w_i for the services of the factor i. To steer clear of the complications that arise from infinite dimensional strategy spaces, I assume that only finitely many demands are feasible. For a positive integer p, let D denote the set of all $p-$decimal fractions that are positive and less than the maximum value that f can take. D is the (finite) set of strategies for each player and $\delta = 10^{-p}$ is the *precision* of the demands. A central concern of the paper is the behaviour of the model for small values of δ. In particular, we will be interested in the set of payoffs that emerge as $\delta \rightarrow 0$.

The factors are identified with the players. In particular, when we say that a coalition S has formed at date t, we mean that only the services of the factors in S have been used for production. Wages are paid in kind. Given a vector of wage demands W, let $W(S)$ denote the total wages demanded by the coalition S. The demands of a coalition S are said to be compatible, if $W(S) \leq f(S)$.

Let w_{it} and W^t denote a typical demand of a player i and the vector of wage demands at date t respectively. The complete *history* up to and including period t is a sequence of demands $W^1, W^2, \ldots W^t$.

Consider a typical player i who has been chosen at date t. Players have no knowledge or a prior regarding the utility of the other players. To determine an optimal response, they have to rely on the historical records. Formally, player i samples at random, a subset of size k_i of the last m records, s= $(W^{t-m+1}, W^{t-m+2}, \dots W^t)$. It is important for this model that every sample of size k_i is sampled with a positive probability. However one need make no assumptions on the relative probabilities with which different parts of the history are sampled. The variable k_i/m is a measure of player i's *information.*

Recall that the history up to date t consisted of only past demands. In particular, the history was silent as to which coalition has formed when similar demands were in place. To proceed, then we will need to assume that the players have some subjective beliefs regarding the likelihood of the formation of coalitions when the demands are compatible in several coalitions. Consider a player who has picked the sample $\sigma_i = (W^1, W^2, \dots W^{k_i})$. The probability with which the demand of player i is met is

$$F_i(w|\sigma_i) = \frac{1}{k_i}[p_i(w|W^1_{-i}) + p_i(w|W^2_{-i}) + \dots + p_i(w|W^{k_i}_{-i})],$$

where $p_i(w|W_{-i})$ is player i's subjective belief that his demand w is met when the others have demanded W_{-i}. Assuming that a player obtains zero if his demand is not met, the expected payoff of player i on sampling σ_i is

$$U_i(w)F_i(w|\sigma_i) + U_i(0)[1 - F_i(w|\sigma_i)$$

where $U_i(w)$ is the utility derived from consuming w. U_i is assumed to be a strictly increasing, differentiable and concave function. Normalize U_i so that $U_i(0) = 0$.

Players maximize their expected utility. Hence if the state at time t is s, the wage demand at $t+1$ is simply

$$w_{it} = \arg\max U_i(w)F(w|\sigma_{it}) \tag{2.1}$$

for some sample σ_{it} of size k_i from the state s. If there are several values of w that solve 2.1 above, then each of them is played with a strictly positive probability.

The above model is similar to fictitious play in the sense that players make their demands naively based on empirical distributions. Unlike in fictitious play, where a player samples the entire history, in the above process, a player samples only a fraction of the most recent history. This process has been termed *adaptive play* by Young (1993a). Since players are sophisticated enough to actually play a best response at each date, it seems that the above behavioural process, as is fictitious play, makes sense only if one assumes that players do not know each other's utility functions.

Fix m. The response rules of the players (determined by equation 2.1), determine a stationary Markov chain. The state space Ω consists of all sequences s of length m. Each entry of s, is a vector of wage demands by the agents. Let $p_i^*(w_i|s)$ denote the probability with which player i demands w_i in the state s. For each i, p_i^* is a best response distribution, i.e., $p^*(w_i|s) > 0$ if and only if w_i solves equation 2.1 for some sample σ_i in s.

For a state $s = (W^1, W^2, \ldots W^m)$, a state $s' = (W^2, W^3, \ldots, W^m, W)$ is said to be its successor. The probability of transition from s to its successor s' is

$$P^0_{ss'} = \prod_{i \in N} p_i^*(w_i|s)$$

and $P^0_{ss'} = 0$, if s' is not a successor of s.

Let P^0 denote the matrix of the above transition probabilities. As in Young (1993b), the above Markov process, with the state space Ω, and transition probability matrix P^0 is said to be an *evolutionary bargaining process* (abbreviated EBP) with memory m, precision δ, information parameters $\{k_i/m\}$ and best reply distributions $\{p_i^*\}$.

2.3 Evolutionary Bargaining Process and Conventions

CONVENTION: A state $\mathbf{s}$ is said to be a *convention* if and only if it is an absorbing state of the evolutionary bargaining process, *i.e.*, $P^0_{\mathbf{ss}} = 1$.

Since there is a positive probability of reaching only an immediate successor from any state, it is clear that a convention must involve of a sequence composed of m repetetions of a single demand. Hence, once the EBP locks into a convention, it becomes the unique way of playing a game. This is the primary concept of an equilibrium that will be employed.

Let $\mathbf{w}$ denote the state in which each entry is the vector of demands W. Now suppose that the process in state $\mathbf{w}$ at time t. Then regardless of which sample player i samples, the probability with which his demand wil. ˙ e met is simply given by

$$F_i(w|\sigma) = p_i(w_i|W_{-i}).$$

Hence whether m repetitions of a single vector of demands constitutes a convention or not, depends (besides the utility function) on what we assume of p_i. The following assumption will be maintained throughout this chapter.

Assumption 2.1 *The likelihood of the formation of a coalition depends only on the set of all coalitions in which the demands are compatible and not on the exact demands themselves.*

Given that the formation of a coalition depends only on the compatibility of the demands and not on the demands themselves is indeed a strong assumption. However, it is partially justified in the present context because the agents make their demands simultaneously and are committed to whatever demands they make.

Definition 2.1 *A vector of wage demands W is said to be in the core of the technology f, if for every coalition S, $W(S) \geq f(S)$ and $W(N) = f(N)$*

At a first glance, it seems somewhat intuitive that conventions should consist of elements of the core alone. Indeed, elements of the core are usually strong equilibria and immune to even coalitional deviations. But this intuition is misguided in the present context. Note that in our model, players make simultaneous demands, and not proposals. When a player makes a proposal, not only does he seek a payoff for himself, but also specifies a payoff for the a subset of other players as well. Hence, he has not only specified a payoff, but also suggested which coalition is to form. In the present model, players independently and simultaneously make their demands. The demands are made based only on precedent. Hence, it is possible that individual players may demand wages much higher than the grand coalition can afford but they expect to obtain these smaller payoffs in smaller coalitions.

Changing the payoffs at this stage may not be optimal because it may change the number of coalitions in which the demands are compatible. For instance, if by demanding less, the number of coalitions in which the demands can be met does not change, it is clearly sub-optimal because of Assumption 2.1. But by making a smaller demand, it is possible that the number of coalitions in which the demand is compatible does increase. Whether demanding less is optimal or not, depends on the exact specification of the beliefs. Similarly for increasing the demands.

For further analysis, we need to make further assumptions on the beliefs of the players regarding the formation of coalitions. An initially appealing assumption that partly compensates the for the admittedly strong Assumption 2.1 is that each of the coalitions in which the demands can be met forms with a strictly positive probability. Such a hypothesis leads to fairly unintuitive results as the following example shows.

Example 2.1 Suppose the $f(ij) = 1$ for all i, j and $f(123) = 2$, and $U_i(w) = w$ for all i. Consider the demands $(1/2, 1/2, 1/2)$ repeated m times. These demands are compatible in each of the two player coalitions as well as the grand coalition. By carefully choosing the beliefs, we wish to show that these demands can be supported in a convention if we assume that each of the coalitions in which the demands are feasible has a positive probability of forming.

Let us assume that each of these coalitions are equally likely to form. The expected utility of demanding 1/2 is then 3/8 for every player. Now suppose that in this state player i raises his demand. Then the demands cannot be compatible in either of the two player coalitions involving player i. The demands will still continue to be compatible in the $\{jk\}$ coalition and perhaps in the grand coalition depending on the player i's demand. The best player i can do, given that players j and k continue to ask for 1/2 each, is to demand for 1. This will be met in the grand coalition with some probability.

Now suppose that player i believes that the jk coalition forms with probability 3/4 while grand coalition forms with probability 1/4. Then the expected utility of asking for more than 1/2 is 1/4 which is less than what he would have obtained by asking for 1/2.

It is clear that demanding less than half is dominated by asking for half. Hence, with the above set of beliefs, each player demanding 1/2 can constitute a convention.

In the above example, player i believes that a smaller coalition will form with a higher probability although the demands of all three players are compatible within a larger coalition leading to the non-intuitive, Pareto sub-optimal outcome. In order to rule out such non-intuitive behaviour, we make the following assumption:

Assumption 2.2 *Suppose that there is a set S^* such that*

$$\textit{For every } S \textit{ such that } W(S) \leq f(S) \textit{ implies } W(S \cup S^*) \leq f(S \cup S^*).$$

Then $p_i(w_i|W_{-i}) = 1$ for all $i \in S^*$.

The above assumption is equivalent to the assumption that the largest coalition forms, if there is a single largest coalition in which the demands are compatible. But, more generally this is not the case as the following example shows:

Example 2.2 Let $n = 4$ and let the demands W be compatible in the coalitions $\{123, 12, 13, 234\}$. Assumption 2.2 only requires that player 2's demand must be met with probability 1. Indeed, this merely implies that $\{13\}$ does not form.

Both Assumptions 2.1 and 2.2 are sufficient to ensure that for example 2.1 the set of conventions coincide with the core. This is however not always the case.

Example 2.3 Let $f(12) = f(13) = 300$, $f(123) = 300 + 2\epsilon$ where $0 < \epsilon < 75$, $U_i(w) = w$ and $k_i = k$ for all i. Here, player 1 is essential for production. The set of demands that players 2 and 3 make in conventions is shown in the figure below:

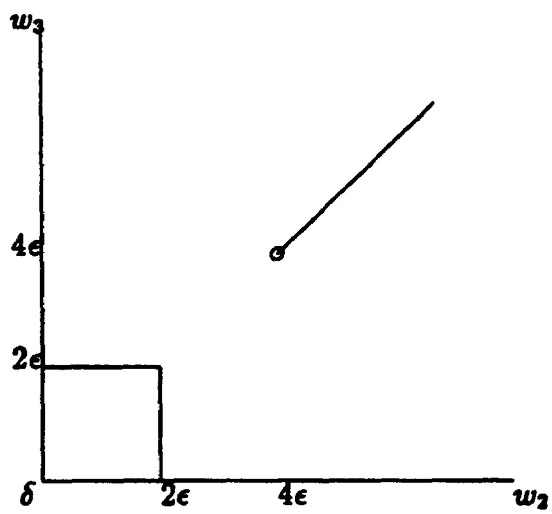

Figure 2.1: Demands in Conventions

The demands made in the core correspond to the square formed with (δ, δ), $(\delta, 2\epsilon), (2\epsilon, \delta)$ and $(2\epsilon, 2\epsilon)$ as its vertices. The set of conventions must include these demands. The set of points on the line of slope 1 starting from $(4\epsilon, 4\epsilon)$, correspond to the demands of the form, $(300 - w, w, w)$ for all $w > 4\epsilon$. I now claim that these are conventions as well. Indeed, let **w** denote the state in which one of these demands has been repeated m times. Regardless of which sample of size k player 1 samples from w, a demand of w is compatible in both the two player coalitions whereas a higher demand cannot be met in any coalition. Hence, demanding more is not optimal. Demanding less is of course not optimal either.

For player 2, demanding more than w is not optimal. On the other hand demanding less than w may be optimal, if the likelihood of obtaining a lower wage for sure in the grand coalition exceeds getting w with some probability. But the most that he can get from the grand coalition, given players 1 and 3's demands is 2ϵ. Let us take his beliefs to be such that all the coalitions in which the demands are compatible are equally likely to form. In the present context this does not seem unreasonable. Then, player 2 will weigh the chances of obtaining w with a probability half against the sure option of 2ϵ. Indeed, for[1] all $w > 4\epsilon$, he must continue with the above strategy and demand w. The analysis is similar for player 3.

It is also useful to point out that conventions with demands such as $(300-w, w, w)$ are *semi-stable* (See Selten (1972))

Proposition 2.1 below identifies sufficient conditions for the set of conventions to coincide with the core.

Proposition 2.1 *Suppose that the technology is convex, i.e.,*

$$f(S) + f(T) \leq f(S \cup T) + f(S \cap T) \qquad \textit{for all } S, T$$

Then w *is a convention if and only if* w *is in the core of the technology* f.

[1]However, note that $(300-4\epsilon, 4\epsilon, 4\epsilon)$ cannot be part of any convention. In a state involving made m repetitions of these demands, the best response for player 2 (and 3) is not unique. Demanding 4ϵ and obtaining it with probability half or demanding 2ϵ and obtaining it for sure are both best responses. By our earlier assumption that all best responses are played with a strictly positive probability, the above will not be an absorbing state.

Proof: See Appendix B.

2.4 Convergence of the EBP with three players

From this section onwards, rigorous proofs are provided only for the case of three players. It will be evident from the proofs that generalization for arbitrary number of players is fairly straightforward, except for Theorem 2.1.

In a Markov Process the set of all aperiodic states are either transient or persistent. If the set of all persistent states are absorbing, then the EBP converges from an arbitrary state to some convention. Theorem 2.1 below, uses some of the ideas from Young (1993a), the only persistent states are the conventions, thereby ruling out cycles and establishing global convergence in probability.

Theorem 2.1 *Suppose that $k_i/m \leq 1/3$ for all i. Then the EBP converges with probability one to a convention.*

PROOF: Assume, for ease of exposition, that $k_i = k$ for all i. Let the process be in state s (t) at date t. Let σ denote the last k elements of this state. Let W^1 denote a best-response vector to the sample σ. Since every sample of size k has a positive probability of being sampled, there is a positive probability of observing W^1 at date $t+1$. In fact, there is a positive probability of observing a run of W^1 between $t+1$ and $t+k$. If $\mathbf{w}^1$ is a convention, we are done. For, between $t+1$ and $t+2k$ all the players will sample the records containing W^1 alone and respond with w_i^1. Hence there is a positive probability p of reaching a convention in $m+2k$ periods. So the probability of not reaching a convention in $r(m+2k)$ periods at most $(1-p)^r$, which goes to zero as $r \to \infty$. Suppose that W^1 is not in the core but is compatible in at least one of the three two player coalitions say $\{12\}$. Then, between $t+k+1$ and $t+2k$, there is a positive probability that players 1 and 2 will continue to sample σ, while player three will sample the records consisting of W^1 alone. His best response[2], by virtue of Assumption 2.2 is $f(123) - w_1 - w_2$. Hence, between period $t+k+1$ and $t+2k$, there is a positive probability of seeing a run of $W^2 = (w_1, w_2, f(123) - w_1 - w_2)$.

It is useful to note that until now we needed a history of length at most $3k$. From now on will make use of samples of size k that appear from dates $t+k+1$ onwards.

Now between $t+2k+1$ and $t+3k$, there is a positive probability all the three players will sample demand W^2. The best response of player 3 continues to be $f(123)-w_1-w_2$ while player 1 and player 2 must demand $f(12) - w_2$ and $f(12) - w_1$ respectively. Hence we will see a run of $W^3 = (f(12) - w_2, f(12) - w_1, f(123) - w_1 - w_2)$ for k periods with a positive probability.

Between $t+3k+1$ and $t+4k$, there is a positive probability of players 2 and 3 continuing to sample demands of W^2 alone while player 1 samples W^3. The best responses of players 2 and 3 remain unchanged while that of player 1 is now w_1. Hence, there is a positive probability of seeing a run of $W^4 = (w_1, f(12)-w_1, f(123)-f(12))$.

Finally, there is a positive probability of all three players sampling the most recent k records consisting of W^4 alone. The best response of players 1 and 2 continue to be w_1 and $f(12) - w_1$ respectively while player three must now demand $f(123) - f(12)$. It may be verified that the demands $W^5 = (w_1, f(12) - w_1, f(123) - f(12))$ is an

[2]It may be verified that since the technology is convex, $f(123) - w_i$ is not feasible in either of the smaller coalitions that contain player 3.

element of the core. Now we repeat arguments similar to those found in the first paragraph of this proof to conclude that we the EBP converges with probability one to a convention if the information is less than or equal to 1/3.

The only other case to consider is when W^1 is feasible is strictly compatible in the grand coalition but is not compatible in any of the smaller coalitions. In this case, it is clear that W^2 above is in the core.

Remark 2.1 Although Theorem 1 requires an upper bound of only 1/3 on the information parameters of the players, I will in fact assume that $k_i/m \leq 1/4$. This bound considerably simplifies later analysis. The tighter bound that n players require would make this assumption unnecessary in general.

Remark 2.2 With n-players, the bound depends on the size of smallest coalition in which positive demands can be met. For example, for a pure bargaining game for N, it can be proved that a bound of $1/n$ is sufficient. However, for more complicated games, a sufficient condition would be $\frac{k_i}{m} \leq 2/n(n-1)$. This number is simply the reciprocal of the number of two player coalitions that can be formed with n players. The rather cumbersome proof is omitted. The idea of the proof, however, will be evident from that of Theorem 2.1 and Proposition 2.1.

Remark 2.3 Note that for Example 2.3, although the set of conventions correspond to those depicted in Figure 2.1, it does not mean that all other states are transient. Cycles are a possibility.

2.5 Stochastically Stable Conventions

The core of a typical game can be quite large. Hence, the set of conventions is very large as well. In order to select among the set of conventions, in the spirit of the models in Kandori, et. al (1993), Young (1993a) and Young (1993b), I consider the possibility of players making mistakes and study the conventions that are stochastically stable. To illustrate the notion of stochastic stability, certain definitions are in order.

Fix the sample sizes k_i and the memory m.

Definition 2.2 *Mistake:* Suppose that the EBP is in state $s = (W^{t-m+1}, W^{t-m+2,\dots,W^t})$ at time t and $s' = (W^{t-m+2}, W^{t-m+3}, \dots W^t, W)$. The transition s to s' is said to involve a mistake on th part of player i if there is no sample in s of size k_i for which w_i is a best-response, i.e., $p_i^*(w_i|s) = 0$. Clearly the number of mistakes involved in a transition from a state to its successor can only be zero one, two or three depending on the number of players that have made a mistake.

Suppose that the probability with which player i makes a mistake is given by $\epsilon\lambda_i > 0$. Conditional on the fact that player i has made a mistake, let $q_i(w_i|s)$ be the probability with which he demands the wage w_i in state s. Clearly, q_i is different from p_i^*. The parameter ϵ is the absolute probability with which players make mistakes and λ_i/λ_j is the relative probability of players i and j making mistakes. The event that player i makes a mistake is assumed to be independent of the that event j makes a mistake.

Now suppose that the process is in state s at time t. The probability that exactly the members in the coalition S make mistakes is $\epsilon^s(\prod_{i\in S}\lambda_i)(\prod_{i\notin S}(1-\epsilon\lambda_i))$. Conditional on this event, the transition probability of moving from a state s to a state s' is

$$Q^S_{ss'} = \begin{cases} \prod_{i\in S} q_i(w_i|s) \prod_{i\notin S} p_i^*(w_i|s) & \text{if } s' \text{ is a successor of } s \\ & \text{and the demands to the far right are } w \\ 0 & \text{otherwise.} \end{cases}$$

If none of the players make mistakes, then the transition probability of moving from state s to a state s' is given by the earlier transition probabilities $P^0_{ss'}$. This event has the probability $\prod_{i=1,2,3}(1-\epsilon\lambda_i)$.

Allowing for the possibility of mistakes, we now obtain a new Markov process with the same state space Ω as before but with the transition function:

$$P^\epsilon_{ss'} = \left(\prod_{i=1,2,3}(1-\epsilon\lambda_i)\right) P^0_{ss'} + \sum_{S\subsetneq N} \epsilon^{|S|}(\prod_{i\in S}\lambda_i)(\prod_{i\notin S}(1-\epsilon\lambda_i))Q^S_{ss'}.$$

Let P^ϵ denote the above matrix of transition probabilities. In most models similar to the one presented here, including Kandori, et. al. (1993) and Young (1993a) and Young (1993b), it is assumed that when players make mistakes, every feasible strategy is played with a strictly positive probability. Mistakes then, constantly perturb the process away from a convention. Now there are no absorbing states. However, since the transition probabilities of the perturbed process converge to those of the unperturbed process as ϵ converges to zero, for small values of ϵ, the perturbed process continues to be attracted to conventions, without actually settling down. Which of the conventions that the process stays at for the most part depends on the number of mistakes that are required to move it far enough to a state from which it would gravitate toward a different another convention. Hence, in the long run, if when the probability of mistakes is very small, the convention that is observed most of the time will be the one that requires the largest number of mistakes to displace.

The asymptotic (or long run) behaviour of a Markov Process is captured completely by its invariant distributions. When one assumes that there is a positive probability of *every* strategy being played, the perturbed process is irreducible. It is easy to show that each of the states is aperiodic as well. Hence there is a unique invariant distribution μ^ϵ for the perturbed process, for each $\epsilon > 0$. For a state s, μ^ϵ_s is the relative frequency with which it will be observed in the first t periods as $t \rightarrow \infty$. Since the invariant distributions (perhaps along a subsequence), converge to the invariant distribution of the unperturbed process, the convention that will be observed most often when the probability of mistakes is small in the one corresponding to the invariant distribution that the invariant distributions of the perturbed process converge to.

This motivates the following refinement of the set of conventions, first introduced by Foster and Young (1990).

Definition 2.3 STOCHASTICALLY STABLE CONVENTION: A convention s is stochastically *stable* if $\lim_{\epsilon\rightarrow} \mu^\epsilon_s$ exists and is positive. A state is strongly stable, if $\lim_{\epsilon\rightarrow} \mu^\epsilon_s = 1$.

While the assumption that players may play any strategy, when they make mistakes is an alternative, I will assume that the players' make small trembles.

Definition 2.4 *Small Trembles:* Let s be the state at date t. A mistake w_i on the part player i is said to be a small tremble if there is a $\hat{w}$ such that (i). $p^*_i(\hat{w}|s) > 0$, and (*ii*). $|\hat{w} - w_i| \leq \delta$.

In other words, a player is deemed to have made a small tremble, if his demand differs from a best-response by at most δ. The support of $q_i(.|s)$ consists of best-responses and demands that are a δ distance away from them.

While the results in the sequel will probably hold under more general assumptions with regard to the trembles or mistakes, such a characterisation is left open as a possibility for future research.

It must be pointed out that we are assuming that the trembles are with regard to the actions rather than payoffs (and hence the terminology trembles rather than mistakes). The set of stochastically stable outcomes can be characterized with trembles in payoffs rather than actions. This will be elaborated in a later section.

When players make only small trembles the matrix of transition probabilities P^ϵ is no longer irreducible nor is it clear that the states are aperiodic. Consequently, it is now not immediate that a unique invariant distribution exists and the notion of stochastic stability may not be a proper refinement concept. However, we have the following theorem:

Theorem 2.2 *Suppose that players only make small trembles. Then for each* $\epsilon > 0$, P^ϵ *admits a unique invariant distribution* μ^ϵ. *Moreover, the support of* μ^ϵ *contains the set of all conventions.*

Sketch of the Proof: It can be shown that even allowing for only small trembles, starting in an arbitrary convention **w**, there is a positive probability of reaching another arbitrary convention (See Corollary B.2). Now, let Ω_1 denote the set of all states that can be reached from some convention with a positive probability when one allows for small trembles. Let s be an arbitrary state in Ω_1. From any other state s', there is positive probability of reaching s. For, by definition s can be reached from some convention, say w. But from Theorem 2.1, we already know that some convention w' can be reached with positive probability from state s'. Since we have already proved that we can reach any two conventions with a positive probability when allowing for small trembles, it follows that one can reach w from w', thereby reaching the state s. Hence, the matrix of transition probabilities associated with the states in Ω_1 is irreducible.

Recall that as long as players have sufficiently small information, it is possible to reach any convention with a positive probability from an arbitrary state. Since Ω_1 is a closed set, each state $\Omega_2 = \Omega_1^c$ is transient.

Hence, the matrix P^ϵ can be decomposed as follows:

$$P^\epsilon = \begin{bmatrix} Q & 0 \\ U & V \end{bmatrix}$$

where Q is an irreducible square matrix corresponding to the states in Ω_1 while the matrix V is a square matrix correspond to the states in Ω_2.

There[3] is a unique invariant distribution corresponding to Q since the matrix is irreducible and the states are aperiodic. By extending this vector with zeroes, we obtain an invariant distribution for P^0. Furthermore, since the support of any other invariant distribution for the matrix P^ϵ must have Ω_1 as its support, it follows that there is the unique distribution. □

[3]All the following assertions that follow are applications of well known results from Stochastic Processes. See for e.g. Feller (1957)

Following Theorem 2.2, the use of stochastic stability as a refinement concept is justified. In fact, we will be interested in something more than stochastic stability. Let α_i be a positive rational number. Suppose that player i samples at most a fraction α_i of the records. We will say that m is admissible if and only if $m\alpha_i$ is an integer. We are interested in conventions that are stochastically stable for every admissible m.

Theorem 2.3 *There exists a level of precision $\delta^* > 0$ such that for every level of precision $\delta \leq \delta^*$,*

1. *There are at most six conventions that are stochastically stable for all admissible values of m. Furthermore, for every if $\mathbf{w}^1$ and $\mathbf{w}^2$ are any stochastically stable conventions,*

$$\| \mathbf{w}^1 - \mathbf{w}^2 \| \leq 2\sqrt{2}\delta.$$

2. *A convention $\mathbf{w}_\delta$ is stochastically stable for every admissible m if and only if it maximizes the function r_δ over the set of all conventions, where*

$$r_\delta(\mathbf{w}) = \min_{i=1,2,3} \left\{ \alpha_i [1 - \frac{U_i(w_i - \delta)}{U_i(w_i)}] \right\}.$$

The above result is formally weaker than a corresponding result obtained by Young (1993b) in the case of two players. The difference is that an upper bound δ^* is required on the level of precision. Although the theorem holds true from somewhat higher values of δ, $f(123)/8$ is a sufficient upper bound in Theorem 2.3. This bound however, depends on the number of players.

The intuition for the theorem is as follows. Starting in a convention, the integer[4] $\lceil mr_\delta(\mathbf{w})\rceil$ is the minimum number of trembles required for some player to have a best-response different from the conventional demand. It turns out that for a convex technology, this minimum number of trembles is sufficient to lead one away from the current convention to another, with no further trembles. Now when r_δ is maximized at a point such as $\mathbf{w}_\delta^*$, the number of trembles required to displace it is the largest. Hence, when the likelihood of trembling is very small, this is the one convention that is the hardest to displace and consequently is observed most often.

Even when the technology is not convex, it is still relatively easy to compute the minimum number of trembles that are required before some player has a best response different from the conventional one. These turn out to be functions that look like r_δ. But now, this minimum number of trembles is no longer sufficient for the players .ɔ continue with demands (as best response,) that would lead them to a different convention. Consider the following example:

Example 2.4 In Example 2.3, set $\alpha_i = 1/4$ for all i. The minimum number of trembles $\hat{r}_\delta(\mathbf{w})$ before which some player has a best response different from the conventional demand in a convention $\mathbf{w}$ is

$$\frac{4}{\delta}\hat{r}_\delta(\mathbf{w}) = \begin{cases} \min_{i=1,2,3} \frac{1}{w_i} & \text{if } W \text{ is in the core} \\ \min\{\frac{2}{w}, \frac{1}{300-w}\} & \text{for a convention such as } W = (300-w, w, w) \end{cases}$$

where $w > 4\epsilon$. This bound may be obtained by constructing arguments similar to those found in Lemma B.1 and Lemma 𝔅.3 in the Appendix B. For ek ments

[4]For a real number x, $\lceil x\rceil$ is the smallest integer larger than x.

in the core, $\hat{r}_\delta$ coincides with r_δ that appears in Theorem 2.3. Among the set of conventions that are not in the core, the fewest number of trembles required before player 1 demands a δ less is $2/w_1$, whereas for players 2 and 3 the minimum number of trembles continues to be $1/w_i$. This is because for player 1 to demand δ less as a best response in a convention such as $(300 - w, w, w)$, we require *both* player 2 and player 3 to make the mistake of asking for δ more than w. Hence, the correction of 2 for player 1.

The least player 1 obtains in the core is in the allocation $(300 - 2\epsilon, 2\epsilon, 2\epsilon)$. In the convention involving these demands, the minimum number of trembles may no longer be sufficient to displace it from the convention. To see this let us, for concreteness, first suppose that $\epsilon = 1$. Then, for a large enough m, it is player 1 who has a best response for the first time. This corresponds to the case where both players 2 and 3 have made the mistake of asking for $2\epsilon + \delta$. Let 1^* be the number of entries in which players 2 and 3 have demanded $2 + \delta$. To this, the best response for player 1 is to demand $298 - \delta$. Suppose that from this point in time, say T, onwards no further trembles are made. We may assume, without loss of generality, that the trembles actually occurred in the last $T - 1^*$ periods.

Now, for every sample of size $m/4$ that player 1 chooses that does not include the last $T - 1^*$ records, he must demand 298. If the sample includes that last $T - 1^*$ records, he demands $298 - \epsilon$. Now consider a sample of size $m/4$ chosen by player 2. Furthermore suppose that in i entries of sample, player 1 has asked for $298 - \delta$. For both players 2 and 3, demanding 2 dominates (strictly) a demand smaller than 2. Furthermore, since both of them (players 2 and 3) are demanding at least 2 and player 1 is demanding either 298 or $298 - \delta$, asking for more than $2 + \delta$ is a strictly dominated strategy.

Now, suppose that in the sample drawn by player 2, every demand of player 1 is $298 - \delta$ while all but one demand of player 3 is 2. The other demand is of course, $2 + \delta$. Then, if player 2 demands 2, he gets it for sure, while if he demands $2 + \delta$, he would obtain $2 + \delta$ on $m/4 - 1$ occasions but obtain only $(2 + \delta)/2$ on the one occasion. Hence, demanding 2 is a unique best-response if

$$2 > \frac{(m/4-1)}{m/4}(2+\delta) + \frac{1}{m/4}(2+\delta)/2 \quad \text{or}$$

$$2 > (1 - 2/m)(2+\delta)$$

For an appropriate δ the last inequality will hold. Since two and three are symmetric, for such levels of precision, both players 2 and 3, will continue to demand 2 units. In due course, the process returns to the convention involving the original set of demands namely, $(298, 2, 2)$.

The problem in this example appears due to the fact that the set of conventions is not connected. In fact, a similar analysis can be carried out for ever $0 < \epsilon < 50$. For $\epsilon \geq 50$, player 1 is no longer the one who deviates with the minimum number of trembles.

When δ is small, the number of trembles required to reach a convention involving demands of the form $(300 - w, w, w)$, starting from a demand in the core, turns out to be very large. This is because we require a series of trembles on the part of players 2 and 3 each of them demanding a δ higher at each instant. On the other hand, the number of trembles required to reach an element of the core from the convention with demands $(296, 4, 4)$ is at least $298m/[4(298+\delta)]$. For a fixed δ, this is a large number. Hence, one cannot obtain a bound independent of δ and m. Questions of existence and characterization of stochastically stable conventions for examples such as these are left open as possibilities for future research.

COROLLARY TO THEOREM 2.3: Suppose that $\delta \to 0$. Then the set of conventions obtained in Theorem 2.3 converge to $\mathbf{w}^*$, where

$$W^* = \arg \max_{W \in \text{Core of } f} \prod_{i=1,2,3} U_i^{\alpha_i}(w_i)$$

The proof of this corollary is a straightforward application of Lemma 3 in Young (1993b).

2.6 Competition and Evolution

As in Young (1993b), let us define the type of a player by the his utility function and the extent of his information. To assert that players are drawn from a common set of types, is in the present context equivalent to saying that $(U_i, \alpha_i) = (U_j, \alpha_j)$ for all i, j. If the technology is convex, by virtue of the above corollary, the stochastically stable outcomes converge to a three-way even split of the surplus generated by all three factors of production, i.e., $(f(123/3, f(123)/3, f(123)/3))$. The outcome is independent of the specification of the utilities, as articulated in Young (1993b).

This outcome can be quite different from the competitive outcome in the following sense. Assume that the utilities of the three agents are linear. Then, the technology f can be thought of as being the representation of market game induced by an economy with quasi-linear preferences. Several different economies can induce the same market game. Now, suppose that only two of the three players are symmetric and the induced technology is convex. Then, the competitive outcome does not treat all three players symmetrically, in contrast to the evolutionary model.

When the technology is not convex, the deviation from the competitive norm is much more conspicuous. This is shown by means of the following examples.

Example 2.5 Consider the limiting case of Example 2.3 when $\epsilon \to 0$. Here, $f(12) = f(13) = f(123) = 300$. The remaining variables are as in Example 2.4. This technology corresponds to the well-known representation of a game with two sellers and one buyer. When one does not allow for zero demands, the core of this technology is empty. But the set of absorbing states corresponds to m repetitions of the form $(300 - w, w, w)$, where $\delta \leq w \leq 300 - \delta$. Player 1 obtains $300 - w$ in one of the two smaller coalitions while 2 and 3 obtain w with some probability. It can also be shown that, with the same bound as in Theorem 2.1, all other states are transient. Of course, it is being assumed that both the two player coalitions form with positive probabilities.

As mentioned before, the problem in Example 2.4 appears because the set of absorbing states is not a connected set. However, when the set of absorbing states is a connected set, as it is here, the techniques in the proof of Theorem 2.3 can still be applied. A stochastically stable convention maximizes the function $\hat{r}_\delta$ below:

$$\hat{r}_\delta(w) = \frac{\delta}{4} \min\left\{\frac{2}{300 - w}, \frac{1}{w}\right\}.$$

The competitive outcome on the other hand, is one in which player 1 yields the least amount to either player 2 or 3. In the present case this corresponds to $(300 -$

$\delta, \delta, \delta)$.

When $\delta \to 0$, the stochastically stable outcomes converge to $(200, 100, 100)$ whereas the competitive outcome is the one in which player 1 gets the whole surplus, i.e., $(300, 0, 0)$. In expected terms, players 2 and 3 obtain 50 in the bargained outcome while the competitive outcome gives them 0.

Example 2.6 In Example 2.5 let $U_i(w) = 1 - e^{-w}$. Then the minimum number of trembles that are required before some player has a best response different from the conventional one is given by

$$\hat{r}_\delta(\mathrm{w}\ ) = \frac{e^\delta - 1}{4} \min\left\{\frac{2}{\mathrm{e}^{300-w} - 1}, \frac{1}{\mathrm{e}^w - 1}\right\}.$$

The stochastically stable outcome is the convention that maximizes the above function. When $\delta \to 0$, the stochastically stable outcome converge to $(300-w^*, w^*, w^*)$ where w^* solves

$$2(e^{w^*} - 1) = (e^{300-w^*} - 1).$$

It may be checked that $w^* > 100$.

The competitive outcome on the other hand, converges to be $(300, 0, 0)$. Relative to the competitive outcome, player 1 continues to fare worse in the stochastically stable outcome. In fact, he does worse in the stochastically stable outcomes with constant absolute risk-aversion than he did with under similar outcomes when all the players had linear utilities in Example 2.5. This is in contrast to Young (1993) where the outcomes were invariant under the specification of utilities.

2.7 Discussion

The preceding analysis has been admittedly vague for the general n player case. But it will be evident from the proofs of various results, that an extension to the general n player case is fairly straight-forward, except for perhaps Theorem 2.1. As indicated in a remark following the proof of that theorem, it is possible to prove it albeit with a much less interesting bound on the information parameters than is available for the general 3 player technology. See Remark 2.2 for further discussion.

Another criticism is perhaps that one should be looking at small mistakes in terms of payoffs rather than small trembles that concentrate on actions. Note that in a convention $\mathbf{w}$, if player i demands $w_i + \delta$, his payoff is zero. But a demand of $w_i - \delta$ will be met for sure, and hence the difference in payoff is only $U_i(w_i) - U_i(w_i - \delta)$. Thus a small tremble $w_i + \delta$ is "big" mistake whereas a tremble $w_i - \delta$ is a small mistake as well. In fact since any demand less than or equal to the conventional demand will be met with probability one and no demand greater than the conventional demand is compatible in any coalition, the smallest mistake in a convention (in terms of payoffs) is a tremble that is a δ less than the conventional demand.

The entire model can be redone with the assumption of small mistakes rather than small trembles. Following the observations from the preceding paragraph, a careful reinterpretation of the various lemmas that appear in Appendix B, yield the following result:

Theorem 2.4 *If one permits only small mistakes, then*

1. *There are at most six conventions that are stochastically stable for all admissible*

values of m. Furthermore, for if $\mathbf{w}^1$ *and* $\mathbf{w}^2$ *are any two stochastically stable conventions,*

$$\| \mathbf{w}^1 - \mathbf{w}^2 \| \leq 2\sqrt{2}\delta.$$

2. *A convention* $\mathbf{w}_\delta$ *is stochastically stable for every admissible m if and only if it maximizes the function* r_δ^* *over the set of all conventions, where*

$$r_\delta^*(\mathbf{w}) = \min_{i=1,2,3} \left\{ \alpha_i \frac{U_i(w_i)}{U_i(w_i + \delta)}] \right\}.$$

A limiting result that corresponds to the Corollary to Theorem 2.3 is a subject of current research by the author.

Chapter 3

Negotiation Schemes and Stability of Bargained Outcomes

3.1 Introduction

The concept of an equilibrium is inextricably linked to some notion of stability. While in non-cooperative games stability of strategies is demanded against individual defections, it assumes a much more stringent form in cooperative and semi-cooperative solution concepts such as the strong equilibrium and coalition-proof Nash equilibrium. In such solution concepts the suggested equilibrium payoffs (and hence the strategies that induce them) are required to be stable against coalitional deviations besides being individually rational.

Stability of equilibrium payoffs against coalitional deviations is an attempt to extend the concept of individual rationality to groups. However, as Luce and Raiffa point out (See Luce and Raiffa(1954)), "the notion of group rationality is neither a basic postulate of the model nor does it appear to follow as a logical consequence of individual rationality." Moreover, the nature of a coalitional deviation is quite different from that of individual defections. To see this, suppose that the current status-quo lies in the interior of the set of all utility payoffs that a coalition can guarantee its members. Should we expect a deviation? The answer obviously depends on what we assume of the bargaining process. If we assume that bargaining follows the coalitional deviation, then it is not clear what the threat points for the members are. In particular, the status-quo may not continue to be the threat point as this may not be attained once the coalitional deviation is effected. Hence, if a coalitional deviation is effected before an agreement is reached, some of the members of the deviating coalition could be worse off relative to the status-quo. Hence, we may not observe a deviation in the first place.

Considerations such as above suggest that coalitional deviations (and group rationality) must at a minimum involve considerable communication between various players regarding the payoffs that will be obtained following a deviation. Furthermore, this agreement must be reached while the status-quo is still a valid threat point. Now, if we take as axiomatic that a coalition cannot be a threat unless its members can communicate between themselves at the status-quo, a question of interest is the maximum extent of communication a designer can permit among a set of players and still hope for stability against coalitional deviations, regardless of the economic possibilities.

In this Chapter, I consider a large class of NTU games and study the effect of different communication possibilities on the stability of bargained outcomes against coalitional deviations. More precisely, I take as given which pairs of agents can communicate between themselves. A coalition involving some two players, say i and j, can be a potential threat to a suggested payoff only if either i can directly communicate with j or there is a sequence of players *in the threatening coalition* through which player i can transmit (receive) a message to (from) player j. If neither of these possibilities exist, a coalition involving i and j cannot be a threat regardless of its economic possibilities described by the characteristic function.

The main result in this chapter is as follows. Suppose the communication possibilities are such that for any two players there is a *unique* set of players who are involved in transmitting a message between them. Then, a stable coalition structure exists. That is, there is a vector of payoffs x^* and a partition of the set of players such that each player obtains x_i^* in a single coalition. Furthermore, this configuration of coalitions is stable against further deviations by coalitions in which players can communicate between themselves. Of course, when one assumes that the characteristic form is super-additive, the payoff vector x^* can be achieved in the grand coalition. Then, the above is nothing but an assertion that the core is not-empty.

Communication possibilities offer a natural restriction on coalition formation. At the same time, in several economic instances certain coalitions do not have any power even if they did form. For example, in labour markets, coalitions consisting of only employers or only employees cannot achieve anything on their own. For these situations, we might as well assume that these coalitions cannot form. Motivated in the above manner, Kaneko and Wooders(1982) take as exogenous a given set of coalitions that can form. Such coalitions were described as *effective* coalitions. Then, they provide a list of necessary and sufficient conditions for a partitioning game to have a non-empty core regardless of the payoff functions of the effective coalitions. Their main conclusion is that as long as the set of effective coalitions is *strongly balanced*, the core of a "partitioning game" non-empty regardless of the economic possibilities described by the characteristic form. Hence, the results in this chapter provide a natural geometric (graph theoretic) interpretation of their results.

In a recent paper Le Breton et. al(1992) independently provide a slightly different proof of the results obtained in this paper. They also discuss the notion of *strong balancedness* for various kinds of games including communication games.

The model is laid out in Section 3.2. It also contains a formal statement of the main result. Proofs of formal statements appear in Section 3.3.

3.2 The Model

Let $N = \{1, 2, \ldots n\}$ be a finite set of players. A coalition S is a non-empty subset of N. As usual, let 2^N denote the class of all coalitions. A *communication structure* (CS) is a subset of $N \times N$. Given a typical CS R, player i_0 can communicate with i_n, if there is a sequence of players $i_1, i_2, \ldots i_k$, such that $(i_{l-1}, i_l) \in R$, for all $l = 1, 2, \ldots k$. The above sequence of players between i_j, $j = 1, 2, n-1$, is said to be a path between i_0 and i_n. A coalition S is said to be *connected* if and only if there is a path between any two of its members that is completely contained in S.

Definition 3.1 An N-player game V with *non-transferable utility* (NTU) in characteristic function form associates with each coalition S, a set $V(S) \subseteq R^{|S|}$

$V(S)$ is the set of all utility levels that a coalition S can guarantee its members if it should form, regardless of the activities of S^c. However, for a coalition S to form, as articulated in the introduction, the players in S must be able to communicate with each other. A connected coalition formalises the notion of communication between players. We take as a given which pairs of players can directly communicate with each other. Player i can transmit (receive) a certain proposal for deviation to (from) player j only if he there is a path between i and j. Hence, if a coalitional deviation containing i and j is to take place, the coalition must involve every player on some path between i and j.

Definition 3.2 Given a CS R, a partition $\{S_i\}$ of N is said to be a ***stable coalition structure*** if and only if there exists an $x^* \in R^n$ such that,

1. Each S_k is connected.

2. $x^*_{S_k} \in V(S_k)$, for all $k = 1, 2 \ldots K$.

3. If $y \in V(S)$ and $y_i \geq x^*_i$ for all $i \in S$ with a strict inequality for at least one player, then S is not connected.

The above definition is a direct generalisation of the notion of a stable coalition structure due to Aumann and Dreze(1972), to the case where there is an associated CS. In fact, if one sets $R = N \times N$, thereby allowing every coalition to form, our definition coincides with that of Aumann and Dreze. If one further requires that the players obtain the payoffs only in the grand coalition, then stable payoffs coincide with that of the core.

Note that the existence of a stable coalition structure is in general not an issue when one is at a liberty to specify the communication structure. For example, if $R = \{(i,i) : \ i \in N\}$, then one only need to check the stability of an allocation against individual defections. Indeed, in this case we are preventing all communication between agents. Consequently, no coalition would ever form. In fact, any partition of N that awards at least the individually rational payoffs constitutes a stable coalition structure. The main question of interest is the largest extent of communication that one can permit and yet ensure existence. The Theorem 3.1 below provides a complete characterization.

Definition 3.3 A connected coalition S is said to be a *tree* if there is a *unique* path between any two of its members.

Theorem 3.1 *Let V be an n-player NTU game in characteristic form such that*

1. $V(\{i\}) = \Re_-$
2. $V(S)$ *is closed and bounded from above,*
3. *Each $V(S)$ is comprehensive, i.e. if $x \in V(S)$ and $y \in R^{|S|}$ such that $y_i \leq x_i$, for all $i \in S$, then $y \in V(S)$.*

Given a CS R, a stable coalition structure of V exists if every connected coalition S is a tree.

Furthermore, if the CS R is not a tree, there is an NTU game statisfying 1, 2 *and* 3 *above that does not admit a stable coalition structure.*

3.3 Proof of Theorem 3.1

Several concepts from the theory of cooperative games are required to prove the above theorem. Many of the definitions are taken straight from Border(1985).

Definition 3.4 A family β of subsets of N is said to be *balanced* if for each $S \in \beta$, there is a positive real number λ_S (called a *balancing weight*) such that for each $i \in N$,

$$\sum_{\beta(i)} \lambda_S = 1,$$

where $\beta(i) = \{S \in \beta : i \in S\}$.

Note that in particular, any partition of N constitutes a balanced family with a unit weight assigned to each of its elements.

The following lemma relates balancedness to connectedness.

Lemma 3.1 *Let R be a CS such that every connected coalition is a tree. Then a balanced collection of connected coalitions contains a partition of N.*

PROOF: Assume without loss of generality that N is connected. Let $\{S_i\}_{i=1}^{i=n}$ be a balanced collection of connected coalitions such that each S_i is a tree. If $S_1 = N$, we are done. Suppose $S_1 \neq N$. Since N is connected, there is a player $i_1 \in S_1$ and an $i_2 \notin S_1$ such that $(i_1, i_2) \in R$.

I now claim that there is a set, say S_2 such that $i_2 \in S_2$ but $i_1 \notin S_2$. Suppose this claim were false. Then i_2 appears in every coalition in the collection that i_1 appears *except* in S_1. Hence, the following holds:

$$\sum_{i_1 \in S} \lambda_S = \sum_{i_2 \in S} \lambda_S + \lambda_{S_1}.$$

where the λ's are the balancing weights. Using the above relation and balancedness, we conclude that $\lambda_{S_1} = 0$. But this contradicts the fact that the above collection is balanced.

Note that $S_1 \cup S_2$ is a connected coalition. Furthermore, $S_1 \cap S_2$ is empty. For, suppose by way of contradiction, that there is an $i_3 \in S_1 \cap S_2$. Clearly, i_3 is distinct from i_1 and i_2. Since S_1 is a connected coalition, there is a path between i_1 and i_3 that does not involve i_2. Furthermore, since S_2 is a connected coalition, there is a path between i_2 and i_3 that does not involve i_1. Since i_1 can directly communicate with i_2, it follows that there are two paths between i_1 and i_2, one that involves i_3 and another that does not. Hence, $S_1 \cup S_2$ is not a tree. But this contradicts our hypothesis that every connected coalition is a tree. Hence, it must be the case that $S_1 \cap S_2 = \phi$.

If $S_1 \cup S_2 = N$, we are done. Suppose this were not the case. we repeat the above procedure to obtain a connected coalition S_3 such that S_1, S_2 and S_3 and mutually disjoint, and $S_1 \cup S_2 \cup S_3$ is connected. The proof is complete on observing that N is finite (and hence the above process stops in finitely many steps).

Lemma 3.2 *K.K.M.S (Shapley(1973)) Let $\{a^i : i \in N\} \subset R^m$ and let $\{\Gamma(S) : S \subset N\}$ be a family of closed subsets of R^k such that for each nonempty $A \subset N$,*

$$co\ \{a^i : i \in A\} \subset \cup_{S \subset A} \Gamma(S).$$

Then there is a balanced family β of subsets of N such that

$$\cap_{S \in \beta} \Gamma(S) \quad \textit{is nonempty and conpact.}$$

PROOF: See Border(1985).

PROOF OF THEOREM 3.1: The proof of necessity is by means of a counter-example. Let R be a CS such that there is a coalition S that is not a tree. Hence there is a sequence of players $i_1, i_2, \ldots i_k$ such that $(i_j, i_{j+1}) \in R$ for all $j = 1, 2, \ldots k-1$. Note that $k \geq 3$. Let $k = 2l+1$ or $k = 2l$. Now consider the following TU game:

$$V(T) = \begin{cases} 1 & \text{if } T \subset S \text{ is connected and } |T| \geq |S| - 1 \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to check that the above game does not have a stable payoff vector. Clearly partition involving S is not stable if all the players in S receive strictly positive payoff. A strictly smaller subset can share the same amount of surplus. If at one of the players in S receive a zero payoff, he can ask for a surplus slightly smaller than one of the players receiving a positive surplus and a coalition of size $|S| - 1$ that is connected dominates the status-quo.

To prove the converse, let R be a CS. Fix an $\epsilon > 0$. Consider the correspondence $W : 2^N \longmapsto \Re^n$ as follows:

$$W(S) = \begin{cases} \{x \in R^n : \pi_S(x) \in V(S)\} & \text{if } S \text{ is connected} \\ \{x \in R^n : x_i \leq -\epsilon, \forall i \in N\} & \text{otherwise} \end{cases}$$

Since each $V(S)$ is bounded from above, there is a uniform bound, say M such that $x \in W(S)$, implies $x_i \leq M$. Now, for each unit coordinate vector e^i of $\Re^n$, set $g^i = -nMe^i$. Let $K = con\{g^i : i = 1, 2 \ldots n\}$, where $con\{X\}$ denotes the convex hull of the set X. Define $\tau : K \longrightarrow \Re$

$$\tau(x) = \max\{\lambda \in \Re : x + \lambda u \in \cup_{S \subseteq N} W(S)\},$$

where u is the unit vector in $\Re^n$.

Finally, for each coalition S, define

$$\Gamma(S) = \{x \in K : x + \tau(x)u \in W(S)\}.$$

Since each $V(S)$ is closed and bounded from above, τ is finite and continuous. Consequently, $\{\Gamma(S) : S \subseteq N\}$ is a family of closed sets. We will now show that the above collection satisfies the hypothesis of the K.K.M.S lemma.

Suppose that $x \in \Gamma(S) \cap con\{g^i : i \in A\}$. We need to show that $S \subseteq A$. Without loss of generality, assume that $A \neq N$. If individual rationality is to hold,

$$x_i + \tau(x) \geq 0 \qquad \forall i \in N \tag{3.1}$$

$$x_i + \tau(x) \leq M \qquad \forall i \in S \tag{3.2}$$

Equation 3.1 holds due to individual rationality whereas equation 3.2 because $W(S)$ is bounded above by M.Since $x \in con\{g^i : i \in A\}$, it follows that $\sum_{i \in A} x_i = -nM$. Using this fact and equation 1 above,

$$\begin{aligned} \textstyle\sum_{i \in A} x_i + |A|\tau(x) &\geq 0 \\ \implies \quad -nM + |A|\tau(x) &\geq 0 \\ \implies \quad \tau(x) &> M \end{aligned}$$

Using equation 3.2 above and the fact that $\tau(x) > M$ it follows that $x_i < 0$ for all $i \in S$. But since $x_i = 0$ for all $i \notin A$, it follows that $S \subset A$.

Hence we have proved that the family of sets $\{\Gamma(S) : S \subseteq N\}$ satisfies the hypothesis of K.K.M.S. lemma. Hence, there is a balanced family β, such that

$$\bigcap_{S \in \beta} \Gamma(S) \text{ is non-empty and compact.}$$

Now by virtue of equation 3.1 and the construction of $W(S)$, each $\Gamma(S)$ is empty if S is not a connected coalition. Thus the balanced family obtained above consists of connected coalitions alone. By virtue of the Lemma 3.1, this balanced family contains a partition. □

Appendix A
Proofs of Chapter 1

The set $\delta_S \ : \ S \subseteq N$ constitutes a basis for $\Re^{2^N-1}$. Hence every bargaining situation $f \in \mathcal{G}$, can be expressed a a unique linear combination of the elements of the above set. That is to say

$$\begin{aligned} f &= \sum_{S \subseteq N} \gamma_f(S)\delta_S \\ &= \sum_{\gamma_f(S) \geq 0} \gamma_f(S)\delta_S + \sum_{\gamma_f(S)<0} \gamma_f(S)\delta_S \\ &\equiv f^+ - f^- \end{aligned}$$

Note that both f^+ and f^- are both made up of *non-negative* linear combinations of the situations δ_S. I will assume that $\succ_i$ satisfies Axioms 1-4, although some of them are superflous for certain lemmas.

Lemma A.3 *For any two bargaining situations f and g,*

1. *Suppose $h \equiv h^+$. Then $f \succ_i g$ if and only if $f + h \succ_i g + h$.*

2. *$f \succ_i g$ if and only if $f^+ + g^- \succ_i g^+ + f^-$.*

Proof: Part 1 of the Lemma follows follows from a repeated application of *Strategic Equivalence*. The proof of Part 2 is complete by taking $h = f^- + g^-$.

Now consider the extension $>_i^*$ of $\succ_i$ from $\mathcal{G}$ to $\Re^{2^N-1}$.

For any f and g in $\Re^{2^N-1}$, $f >_i^* g$ if and only if $f^+ + g^- \succ_i g^+ + f^-$.

By virtue of the fact that δ_S constitutes a basis for the space $\Re^{2^N-1}$, $>_i^*$ is indeed an extension of $\succ_i$.

Lemma A.4 *The relation $>_i^*$ is complete and for any three vectors f, g and h in $\Re^{2^N-1}$, $f >_i^* g$ if and only if $f + h >_i^* g + h$.*

Proof: Completeness of $>_i^*$ follows from completeness of $\succ_i$. To prove the latter part of the lemma, I first prove that for any two vectors $f, g \in \Re^{2^N-1}$, $f >_i^* g$ if and only if $f - g >_i^* 0$. Note that the following equation holds because both sides represent the same vector, $(f - g)$:

$$(f-g)^+ - (f-g)^- = (f^+ - f^-) - (g^+ - g^-),$$

or,

$$(f-g)^+ = (f^+ + g^-) - (g^+ + f^-) + (f-g)^- \tag{A.3}$$

Now consider the following sequence obtained on using equation A.3:

$$\begin{aligned} & f - g >_i^* 0 \\ \Leftrightarrow\ & (f-g)^+ \succ_i (f-g)^- \\ \Leftrightarrow\ & f^+ + g^-) - (g^+ + f^-) + (f-g)^- \succ_i (f-g)^- \end{aligned}$$

Adding $(g^+ + f^-)$ to both sides of the last relation , and using part 1 of Lemma A.1,

$$\begin{aligned} & (f-g) >_i^* 0 \\ \Leftrightarrow \; & (f^+ + g^-) + (f-g)^- \succ_i (g^+ + f^-) + (f-g)^- \\ \Leftrightarrow \; & (f^+ + g^-) \succ_i (g^+ + f^-) \\ \Leftrightarrow \; & f >_i^* g \end{aligned}$$

Now since f and g are arbitrary vectors, we have the following sequence:

$$\begin{aligned} f >_i^* g \;\; & \Leftrightarrow \;\; f - g >_i^* 0 \\ & \Leftrightarrow \;\; (f+h) - (g+h) >_i^* 0 \\ & \Leftrightarrow \;\; (f+h) >_i^* (g+h), \end{aligned}$$

□

REMARK: Lemma A.2 shows that if $\succ_i$ satisfies *Strategic Equivalence* on $\mathcal{G}$, there is an extension $>_i^*$ to the entire space $\Re^{2^N-1}$ that satisfies an analouge Part 1 of Lemma A.1. Although the relation $>_i^*$ is complete, it need not be transitive. The following lemma, however isolates the property sufficient for our purposes.

Lemma A.5 *Suppose* $f >_i^* g$ *and* $g >_i^* 0$. *Then* $f >_i^* 0$.

Proof: Suppose that, by way of contradiction, $0 \geq_i^* f$. Then using Lemma 1, $f^- + g^+ \succeq_i f^+ + g^+$. Now, $g^+ \succ_i g^-$ by hypothesis. Hence using Lemma A.1, I get $g^+ + f^+ \succ_i g^- + f^+$. The previous two expressions and transitivity imply, $f^- + g^+ \succ_i f^+ + g^-$, or $g >_f^*$. This is not compatible with the hypothesis $f >_i^* g$. Hence, $f >_i^* g$. □.

Lemma A.6 *Suppose* $f \geq_i^* 0$. *Then for any* $\alpha \geq 0$, $\alpha f \geq_i^* 0$.

Proof: I will first prove the lemma when α is an integer. Suppose $f \geq_i^* 0$. Then $f^+ \succeq_i f^-$. The proof is by induction. Suppose the lemma is true for any integer

$\alpha \le k$. Then $kf^+ \succeq_i kf^-$. Now let $\alpha = k+1$. By *Strategic Equivalence*

$$f^+ + kf^+ \succeq_i f^- + kf^+$$

$$kf^+ + f^- \succeq_i kf^- + f^-$$

The above two equations together with transitivity yield $(k+1)f^+ \succeq_i (k+1)f^-$.

I now claim that for any integer k, $(\frac{1}{k})f^+ \succeq_i (\frac{1}{k})f^-$. Suppose not. Then $(\frac{1}{k})f^- \succ_i (\frac{1}{k})f^+$. But then multiplying both sides by k, $f^- \succ_i f^+$, a contradiction.

Using these two facts, it is easy to see that the lemma holds when α is a rational number. For a general α, choose a seqence of rational numbers α_n such that $\lim_{n\to\infty} \alpha_n = \alpha$. Then $\alpha_n f^+ \succeq_i \alpha_n f^-$. Since $\succeq_i$ is continuous, it follows that $\alpha f^+ \succeq_i \alpha f^-$, or $\alpha f \ge_i^* 0$. This proves the lemma.

Let E_i denote the set of all f that are strictly better than 0, i.e.,

$$E_i = \{ f \in \Re^{2^N-1} : f >_i^* 0\}$$

Lemma A.7 *The set E_i is an open and convex subset of $\Re^{2^N-1}$.*

Proof: I will first show that E_i is open. Let f_* be a limit point of E_i^c. Then there is an infinite sequence of points f_n in E_i^c, such that $\lim_{n\to\infty} f_n = f_*$. For each n, $f_n^- \succeq_i f_n^+$. So the pair (f_n^-, f_n+) is a point in the graph of $\succeq_i$. Since f_n converges to f_*, it follows that f_n^+ converges to f_*^+ and f_n^- converges to f_*^-. By continuity, it follows that (f_*^-, f_*^+) is in the graph of $\succeq_i$. That is to say $f_*^- \succeq_i f_*^+$, or $f_* \in E_i^c$. Hence E_i^c is closed, or E_i is open.

Let $f, g \in E_i$. By virtue of Lemma A.4, for any $\alpha \in (0,1)$,

$$\alpha f >_i^* 0 \qquad \text{and} \qquad (1-\alpha)g >^* 0.$$

Adding αf to both sides of the expression on the far right and using Lemma A.2

$$(1-\alpha)g + \alpha f >_i^* \alpha f.$$

Now using Lemma A.3 I conclude that $(1-\alpha)g + \alpha f \in E_i$. Hence E_i is convex. Q.E.D.

Lemma A.8 *There is a set of weights* $\{\lambda(i,S) \ : \ S \subseteq N\}$*, such that for any* $f, g \in \mathcal{G}$*,*

$$f \succ_i g \iff \sum_{S \subseteq N} \lambda(i,S) f(S) > \sum_{S \subseteq N} \lambda(i,S) g(S).$$

Proof: By Lemma 5, the set E_i is an open and convex subset of $\Re^{2^N-1}$ that does not contain the origin. Hence the separating hyperplane theorem applies. There is a hyperplane that strictly separates the origin and the set E_i. Let $\{\lambda(i,S) \ : \ S \subseteq N\}$ denote this hyperplane. Hence $f \in E_i$ if and only if $\sum_{S \subseteq N} \lambda(i,S) f(S) > 0$.

Now consider any two bargaining situations f and g such that $f \succ_i g$. By virtue of Lemma A.2, this is true if and only if $f - g >_i^* 0$, i.e. $f - g \in E_i$, or

$$\sum_{S \subseteq N} \lambda(i,S)[f(S) - g(S)] > 0,$$

This completes the proof. □.

Lemma A.9 *Suppose that* $S \in \Pi_{-i}$*. Then* $\lambda(i,S) + \lambda(i, S \cup i) = 0$*.*

Proof: The proof is by induction on the size of the coalition. By *Nullity,* there is an $x \neq 1$ such that $\delta_{N\setminus i} \sim_i \alpha\delta_{N\setminus i}$. Using the representation in the previous lemma,

$$\sum_S \lambda(i,S)\delta_{N\setminus i} = x \sum_S \lambda(i,S)\delta_{N\setminus i}$$

Or,

$$\lambda(i,N)+\lambda(i,N\setminus i)=0.$$

Now suppose that we have proved the lemma for every coalition $S\in\Pi_{-i}$ of size k, $N-1\geq k\geq 2$. It remains to prove for the lemma for a coalition of size $k-1$. Let $S\in\Pi_{-i}$ be a coalition of size $k-1$. By *Nullity*, it follows that there is an $x\neq 1$ such that $\delta_S\sim_i x\delta_S$. Hence,

$$\sum_{T\subset N}\lambda(i,T)\delta_S(T)=x\sum_{T\subset N}\lambda(i,T)\delta_S.$$

Or,

$$\lambda(i,S)+\lambda(i,S\cup i)+\sum_{T\in\Pi_{-i},T\supset S,t\geq k}\lambda(i,T)+\lambda(i,T\cup i)=0.$$

But by the induction hypothesis, second expression on the left hand side is zero. Hence,

$$\lambda(i,S)+\lambda(i,S\cup i)=0.$$

This completes the proof. □

Corollary A.1 *For any $S,T\in\Pi_{-i}$ and $x\geq 0$, $x\delta_S\sim_i\delta_T$.*

PROOF: For any bargaining situation f, the following is true:

$$\begin{aligned}\sum_{S\subsetneq N}\lambda(i,S)f(S) &= \sum_{S\in\Pi_i}\lambda(i,S)f(S)+\sum_{S\in\Phi_i}\lambda(i,S)f(S)\\ &= \sum_{S\in\Pi_i}\lambda(i,S)f(S)-\sum_{S\in\Pi_{-i}}\lambda(i,S\cup i)f(S)+\\ &\quad \sum_{S\in\Pi_{-i}}\lambda(i,S\cup i)f(S)+\sum_{S\in\Pi_{-i}}\lambda(i,S)f(S)\\ &= \sum_{S\in\Pi_i}\lambda(i,S)[f(S)-f(S\setminus i)]+\\ &\quad \sum_{S\in\Pi_{-i}}[\lambda(i,S)+\lambda(i,S\cup i)]f(S)\end{aligned}$$

By virtue of Lemma A.7, it follows that the very last expression above is zero. Hence

$$\sum_{S\in\Pi_i} \lambda(i,S)f(S) = \sum_{S\in\Pi_i} \lambda(i,S)[f(S) - f(S\setminus i)]$$

The corollary is now an easy implication of the above expression. □

Lemma A.10 *For every coalition $S \in \Pi_i$, $\lambda(i,S) > 0$.*

PROOF: We will first prove the statment when $S \neq i$. By virtue of Lemma A.7, it is sufficient to show that $\lambda(i, S\cup i) > \lambda(i,S)$ for each $S \in \Pi_{-i}$. Suppose, by way of contradiction, that $\lambda(i,S) \geq \lambda(i, S\cup i)$. Now by virtue of the previous lemma,

$$\sum_{S\subset T, T\in\Pi_{-i}} \lambda(i,T) + \lambda(i, T\cup i) = 0$$

Hence,

$$\lambda(i,S) \geq \lambda(i, S\cup i),$$

Or,

$$\lambda(i,S) + \sum_{S\subset T, T\in\Pi_{-i}} [\lambda(i,T) + \lambda(i,T\cup i)] \geq \lambda(i, S\cup i) + 2\sum_{S\subset T, T\in\Pi_{-i}} [\lambda(i,T) + \lambda(i,T\cup i)]$$

Or,

$$\lambda(i,S)+\lambda(i,S\cup i)+ \sum_{S\subset T, T\in\Pi_{-i}} [\lambda(i,T)+\lambda(i,T\cup i)] \geq 2\lambda(i,S\cup i)+2\sum_{S\subset T, T\in\Pi_{-i}} [\lambda(i,T)+\lambda(i,T\cup i)]$$

Or,

$$\sum_{T\supseteq S} \lambda(i,T) \geq 2\sum_{T\supseteq S\cup i} \lambda(i,T)$$

But this is the same as saying,

$$\delta_S \succeq_i 2\delta_{S\cup i}$$

By virtue Corollary A.1, the above is equivalent to

$$\delta_{N\setminus S\cup i} \succeq_i 2\delta_{S\cup i}.$$

This, contradicts *Productivity*. □

We are now all set to prove Theorem 1.

PROOF OF THEOREM 1.1: It is easy to verify that 2 implies 1. Now suppose 1 holds. I will show that 2 must hold as well. Lemma A.6 and the expression obtained in the Corollary A.1 imply the following:

$$f \succ_i g \iff \sum_{S\in\Pi_i} \lambda(i,S)[f(S) - f(S\setminus i)] > \sum_{S\in\Pi_i} \lambda(i,S)[g(S) - g(S\setminus i)]$$

By virtue of Lemma A.8, each of the weights in the above expression is positive. Now we may obtain the weights statement of theorem 1 by defining $p(i,S)$ as follows:

$$p(i,S) = \frac{\lambda(i,S)}{\sum_{S\in\Pi_i} \lambda(i,S)}$$

The set of weights $p(i) = \{p(i,S) \ : \ S \in \Pi_i\}$ is said to represent $\succ_i$. It remains to demonstrate the uniqueness of these weights. To do this, it is useful to introduce some notation. Given a set of weights $p(i)$, such the $p(i,S) > 0$ and $\sum_{S\in\Pi_i} p(i,S) = 1$, define:

$$\alpha_{p(i)}(S) = \sum_{T\in\Pi_i} p(i,T)[\delta_S(T) - \delta_S(T\setminus i).$$

For any $S \in \Pi_i$, $\alpha_{p(i)}(S)$ denotes the evaluation of the situation δ_S by player i under the weights $p(i)$. Note that $\alpha_{p(i)}(i) = 1$ and $\alpha_{p(i)}(S) > 0$. Now label the sets $S \in \Pi_i$ and fix the order. Let $\alpha_{p(i)}$ denote the vector in $\Re^{2^{N-1}}$ whose entries correspond to

$\alpha_{p(i)}(S)$ and the order is as per the earlier labelling of the coalitions. Assume without loss of generality that $\alpha_{p(i)}(i)$ is the first coordinate.

Now for each $S \in \Pi_i$, let δ_S^i be the vector in $\Re^{2^{N-1}}$ obtained from each $\delta_S \in \Pi_i$ by restricting to the coalitions that contain player[1] i. Let $\mathbf{M}_i$ denote the matrix[2] whose rows are composed of the vectors $\delta^i(S)$. The k^{th} row correponds to the k^{th} set $S \in \Pi_i$ as per the earlier labelling.

Then, by construction,

$$\mathbf{M}_i p(i) = \alpha_{p(i)}.$$

Now let if possible $q(i)$ be another set of positive weights that sum to one and represent $\succ_i$. It is easy to see that $\mathbf{M}_i$ has full rank. Hence, $p(i) \neq q(i)$, implies that there is a coalition S such that $\alpha_{p(i)}(S) \neq \alpha_{q(i)}(S)$. Assume, without loss of generality, that $\alpha_{p(i)}(S) > \alpha_{q(i)}(S)$. Now consider the situation f^* defined as follows:

$$f^* = \frac{1}{\sqrt{\alpha_{p(i)}(S)\alpha_{q(i)}(S)}}\delta_S.$$

Now, using $p(i)$,

$$\sum_{S \in \Pi_i} p(i,S)[f^*(S) - f^*(S \setminus i)] = \sqrt{\frac{\alpha_{p(i)}(S)}{\alpha_{q(i)}(S)}} > 1,$$

Or,

$$f^* \succ_i \delta_i.$$

Now using $q(i)$ instead of $p(i$, we obtain

$$\delta_i \succ_i f^*,$$

[1]For e.g. with two players A and B, $(\delta_A(A), \delta_A(B), \delta_A(AB)) = (1,0,1)$ and $(\delta_A^A(A), \delta_A^A(AB)) = (1,1)$; $(\delta_{AB}(A), \delta_{AB}(B), \delta_{AB}(AB)) = (0,0,1)$ and $(\delta_{AB}^A(A), \delta_{AB}^A(A)) = (0,1)$.

[2]For the two player case, $\mathbf{M}_A = \begin{pmatrix} \delta_A^A \\ \delta_{AB}^A \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

which is a contradiction. It is now easy to see that[3] $\alpha_{p(i)} = \alpha_{q(i)}$. This concludes the proof. □

PROOF OF THEOREM 1.2: Let $i, j \in S$. We will first show that $\alpha_{p(i)}(S) = \alpha_{p(j)}(S)$. Assume, by way of contradiction, that $\alpha_{p(i)}(S) > \alpha_{p(j)}(S)$. Set $\alpha = \frac{1}{\sqrt{\alpha_{p(i)}(S)\alpha_{p(j)}(S)}}$. Now,

$$\sum_{T \in \Pi_i} p(i,S)[\delta_S(T) - \delta_S(T \setminus i)] = \sqrt{\frac{\alpha_{p(i)}(S)}{\alpha_{q(i)}(S)}} > 1.$$

Hence $\alpha\delta_S \succ_i \delta_i$. A similar exercise as above yields $\delta_j \succ_j \alpha\delta_S$. This violates *Symmetry*.

Hence $\alpha_{p(i)}(S) \leq \alpha_{q(i)}(S)$. Since a strict inequality yields a contradiction as the one above, it follows that

$$\alpha_{p(i)}(S) = \alpha_{p(j)}(S)$$

To show that $p(i,S) = p(j,S)$, first note that when $S = N$,

$$p(i,N) = \alpha_{p(i)}(N) = \alpha_{p(j)}(N) = p(j,N).$$

Now suppose that we have proven $p(i,S) = p(j,S)$ for every coalition S of size $k+1$, $1 \leq k \leq N-1$. If we show that the equality must hold for a coaltion of size k, then the proof is complete by induction on the size of the coalitions.

Let S be a coalition of size k. Then by the induction hypothesis,

$$\sum_{T \supset S} p(i,T)[\delta_S(T) - \delta_S(T \setminus i) = \sum_{T \supset S} p(j,T)[\delta_S(T) - \delta_S(T \setminus j)$$

[3]Note that we have critically used the fact that the set of weights sum up to one, or $\alpha_{p(i)}(i) = \alpha_{q(i)}(i) = 1$.

Assume by way of contradiction that $p(i,S) > p(j,S)$. Then,

$$\sum_{T \supseteq S} p(i,T)[\delta_S(T) - \delta_S(T \setminus i) > \sum_{T \supseteq S} p(j,T)[\delta_S(T) - \delta_S(T \setminus j)$$

Or, $\alpha_{p(i)}(S) > \alpha_{p(j)}(S)$, a contradiction. Hence, $p(i,S) \leq p(j,S)$. Since a strict inequality yields a similar contradiction, if follows that $p(i,S) = p(j,S)$.

PROOF OF THEOREM 1.3: From Theorem 2, it is clear that for every coalition S, there is a unique number $\alpha(S)$, such that $i \in S$, $\alpha(S)\delta_S \sim \delta_i$. I now claim that *Consistency* implies that $\alpha(S) = |S|$.

Suppose that, by way of contradiction, that $\alpha(S) > |S|$. Hence,

$$\delta_i \sim \alpha(S)\delta_S \succ_i |S|\delta_S \qquad \text{for all } i \in S,$$

for some $\alpha > |S|$. But clearly, this implies all the players in S are pessimists. This violates Part 2 of *Consistency*. Hence $\alpha(S) \leq |S|$. Since a strict inequality leads to a contradiction of Part 1 of the *Consistency*, one concludes that $\alpha(S) = |S|$.

Hence $|S|\delta_S \sim_i \delta_i$, or

$$|S| \sum_{T \in \Pi_i} p(i,T)[\delta_S(T) - \delta_S(T \setminus i)] = \sum_{T \in \Pi_i} p(i,T)[\delta_i(T) - \delta_i(T)]$$

Or,

$$\sum_{T \in \Pi_i} p(i,S)\delta_S(T) = \frac{1}{|S|} \qquad S \in \Pi_i$$

The Shapley weights solve the above system of equations. Since the solution must be unique, it follows that:

$$p(i,S) = \frac{(n-s)!(s-1)!}{n!}.$$

This completes the proof. □.

Appendix B
Proofs of Chapter 2

Lemma B.11 *Let the technology be convex. Fix W. Suppose that for some two coalitions S and T,*

1. $W(S) < f(S)$,
2. $W(T) \leq f(T)$ *and*
3. $W(S \cup T) \geq f(S \cup T)$.

Then, $S \cap T \neq \phi$ *and* $W(S \cap T) < f(S \cap T)$.

PROOF: That $S \cap T \neq \phi$ is clear. Furthermore, using convextity,

$$\begin{aligned} W(S \cap T) &= W(S) + W(T) - W(S \cup T) \\ &< f(S) + f(T) - W(S_1 \cup T) \\ &\leq f(S) + f(S \cup T) - W(S \cup T) \\ &\leq f(S \cap T) \end{aligned}$$

Let P be the following statement defined on the class of all coalitions.

$P(S)$: $W(T) \leq f(T)$ implies $W(S \cup T) < f(S \cup T)$.

Lemma B.12 *Fix W and let the technology be convex. If there is a coalition S_1 such that $W(S_1) < f(S_1)$, then there is a coalition S^* such that $W(S^*) < f(S^*)$ and $P(S^*)$ holds.*

PROOF: By hypothesis, $W(S_1) < f(S_1)$. If $P(S_1)$ holds, we are done. If not, there is a coalition T_1 such that S_1 and T_1 satisfy the hypothesis of Lemma B.11. Hence $S_2 = S_1 \cap T_1$ is not empty and $W(S_2) < f(S_2)$. If $P(S_2)$ holds set $S^* = S_2$ and we are done. Or else we repeat the above procedure. In finitely many steps, we obtain a set S^* that proves the lemma or arrive at a set $S_k = \{ij\}$ such that $w_i + w_j < f(ij)$. I now claim that $P(\{ij\})$ must hold.

Suppose not. Then, there is a set T such that $W(T) \leq f(T)$ but $W(T \cup \{ij\}) \geq f(T \cup \{ij\})$. Clearly, this cannot be true if T does not contain both i and j. Suppose that $j \in T$. Then the preceeding two inequalities imply

$$\begin{aligned} W(T \cup i) - W(T) &\geq [f(T \cup i) - f(T)] \\ w_i &\geq [f(T \cup i) - f(T)] \\ &\geq f(ij) \\ &> w_i \end{aligned}$$

which is a contradiction. Hence $P(\{ij\})$ holds. This proves the lemma. □

Proof of Proposition 2.1: Suppose that the EBP is in state w at time t and let W be an element of the core. Then, by taking $S^* = N$ in Assumption 2.2, we note that every players' demand is met with probability one. Hence, there is no incentive to lower a demand. Since raising the demand would lead to an incomptibility in every coalition, it follows that responding with the the same demand is the unique best-response. Hence **w** is a convention.

Conversly, let **w** denote a convention. Now, by virtue of Lemma B.12, if there is a coalition S for which $W(S) < f(S)$, then by Assumption 2.2, there is player who can increase his demand by a small amount, and still obtain it with probability one. This contradicts the fact that **w** is a convention. Hence, $W(S) \geq f(S)$ for all S.

The proof is now complete if we show that $W(N) = f(N)$. Let $\beta(W)$ denote the set of all largest coalitions in which the demands are compatible. It is clear that for **w** to be a convention, for every player there is at least one coalition in which the demands are compatible. Or else, his current expected payoff is zero and any other demand is weakly dominates the one in the candidate convention.

Let $S, T \in \beta(W)$ be two distinct coalitions. Then, we have $W(S) \leq f(S)$ and $W(T) \leq f(T)$ but $W(S \cup T) > f(S\cup)$. Proceeding as in Lemma B.11, we can conclude that this implies $W(S \cap T) < f(S \cap T)$, which, by the arguments of the preceeding paragraph, is not true. Now setting $S_1 = S \cap T$ in Lemma B.12, we arrive at a contradiction. Hence, there is a single largest coalition in which the demands are compatible. □

Lemma B.13 *The number of small trembles required by others in a convention* **w** *such that player i is the first to have a best response different from w_i is at least*

$$\min\{k_i \frac{U_i(w_i)}{U_i(w_i + 2\delta)}, k_i[1 - \frac{U_i(w_i - \delta)}{U_i(w_i)}]\}$$

PROOF: Starting from a convention **w**, let s be the first state in which player 1 is the first player to have a best-response different from w_1. Hence, there is a sample of size k_1 in which for $i \leq k_1$ of the periods player 2 and/or player 3 have demanded something different from w_2 or w_3 respectively. By the definition of s, it follows that these i entries must constitute trembles on the part of players 2 and(or) 3. Of these i entries, let i_0 of them be the ones in which both players 2 and 3 have made different demands, i_2 of them in which player 2 alone has made a different demand and i_3 where only player 3 has made a different demand. Of course, $i = i_0 + i_2 + i_3$. The total number of trembles however, is $2i_0 + i_2 + i{+}_3$.

Since we are allowing for only small trembles, the best response of player 1, w_1' must be in the set $\{w_1 - 2\delta, w_1 - \delta, w_1 + \delta, w_1 + 2\delta\}$.

Case 1a: $w_1' = w_1 + \delta$: Consider the following sample σ of size k_1, constructed from the original sample drawn by player 1. Replace each of the i demands of player 2 with $w_2 - \delta$ and those of player 3 with w_3. This sample σ has only i (and hence fewer) trembles than the original sample. Furthermore, if $w_1 + \delta$ was a best- response in the original sample, it continues to be a best-response to be a best-response in the sample σ. Since, player 1 would obtain $w_1 + \delta$ with probability i/k_1 while he would obtain w_1 for sure, it follows that

$$\frac{i}{k_1}U_1(w_1 + \delta) \geq U_1(w_1),$$

or

$$i \geq k_1 \frac{U_1(w_1)}{U_1(w_1 + \delta)} \tag{B.4}$$

Case 1b: $w' = w_1 + 2\delta$: Note that the only instances in which the demand of player 1 is met are those inst .nces where both players have made a mistake and in the same direction. So construct a sample σ of size k_1 from the original sample as follows. Replace each the i_2 demands of player 2 and i_3 demands of the player 3 among the i entries with w_2 and w_3 respectively. Furthermore, replace the i_0 demands of players 2 and 3 with $w_i - \delta$, $i = 2, 3$. Now, note that if $w_1 + 2\delta$ was a best response to the original sample, it continues to be a best response to σ. Hence, it must be the case that

$$\frac{i_0}{k_1} U_1(w_1 + 2\delta) \geq U_1(w_1),$$

or

$$i \geq k_1 U_1(w_1)/U_1(w_1 + 2\delta) \tag{B.5}$$

Case 2a $w_1' = w_1 - \delta$: This case may be analyzed like we did the earlier two cases. Ingore the trembles of player 3 and let player 2 tremble by demanding $w_2 + \delta$ in each of the i entries. Then, if $w_1 - \delta$ were a best-response for player 1 in the original sample, it continues to be a best response now. Then, player 1 gets $w_1 - \delta$ for sure in this sample, but obtains w_1 only $k_1 - i$ of the times. Hence,

$$U_1(w_1 - \delta) \geq (1 - \frac{i}{k_1}) U_1(w_1),$$

or

$$i \geq k_1 [1 - \frac{U_1(w_1 - \delta)}{U_1(w_1)}]. \tag{B.6}$$

Case 2b: $w_1' = w_1 - 2\delta$: Following in a simlar vein for this case, we get

$$i \geq k_1 [1 - \frac{U_1(w_1 - 2\delta)}{U_1(w_1)}]. \tag{B.7}$$

Note that the RHS of equation B.5 is less than the RHS of equation B.4. Similarly, the RHS of equation B.6 is less than the RHS of equation B.7. Since we can carry out this construction for every player, the bound stated in the statement of the lemma holds. □.

Le ıma B.14 Let **w** be a convention. There is a level of precision $\delta^* > 0$ such that for a level of precision $\delta \le \delta^*$,

$$k_i U_i(w_i)/U_i(w_i + 2\delta) \ge \min_{j=1,2,3} k_j[1 - U_j(w_j - \delta)/U_j(w_j)]$$

PROOF: The following two inequalitities are a straightforward implication of concavity of the utility functions and the fact that $U_i(0) = 0$.

$$\begin{aligned} \alpha_i U_i(w_i)/U_(w_i + 2\delta) &\ge \alpha_i w_i/(w_i + \delta) \\ \alpha_i[1 - U_i(w_i - \delta)/U_i(w_i)] &\le \alpha_i \delta/w_i \end{aligned}$$

Assume, without loss of generality that $k_1 \ge k_2 \ge k_3$. Recall that the minimum that can be demanded is δ. Then, the RHS of the first inequality above, is bounded below by $k_i/3$. Suppose that, in this convention, $w_3 \ge 3\delta$. Then, the RHS of the second inequality is bounded above by $k_3/3$. It is obvious that in this case, the assertion of the lemma holds for every $\delta > 0$.

Now consider the case when $w_3 \le 2\delta$. Then $w_i = \max\{w_1, w_2\} \ge f(123) - 2\delta$. Now, choose δ^* so that for all $\delta \le \delta^*$ the following inequality holds:

$$k_1 2\delta/[f(123) - 2\delta] \le k_3/3.$$

For this δ^*, it may be verified that the assertion of the lemma holds. □

COROLLARY: In a convention **w**, if $\delta \leq \delta^*$, the number of trembles that are required before some player has a best-response cannot fall below $[mr_\delta(\mathbf{w})]$, where m is admissible and

$$r_\delta(\mathbf{w}) = \min_{i=1,2,3} \alpha_i[1 - \frac{U_i(w_i - \delta)}{U_i(w_i)}].$$

The proof of the above corollary is immediate.

Definition B.5 *Resistance:* Let $\mathbf{s}'$ be a successor of $\mathbf{s}$. The resistance between these two states, denoted by $\mathbf{r}(\mathbf{s},\mathbf{s}')$ is the minimum number of small trembles required in the one period transition $\mathbf{s} \longrightarrow \mathbf{s}'$. If $\mathbf{s}'$ can be obtained from a small mistake, then $\mathbf{r}(\mathbf{s},\mathbf{s}')$ takes the value 0,1,2 or 3. Otherwise, $\mathbf{r}(\mathbf{s},\mathbf{s}')$ is ∞. Similarly, for any two states $\mathbf{s}^1$ and $\mathbf{s}^2$, $\mathbf{r}(\mathbf{s}^1,\mathbf{s}^2)$ is the minimum number of small trembles required to reach $\mathbf{s}^2$ from $\mathbf{s}^1$ through a sequence of one period transitions involving small trembles.

Now, let e_{ij} denote the vector in R^3_+ such that the ith component is δ, the jth component is $-\delta$ and the kth component is zero. For a convention **w**, if $w + e_{ij}$ is in the core, $\mathbf{w}+e_{ij}$ is the convention in which player i demands a δ less and j demands a δ more than their respective demands in **w**. The following lemma shows that conventions such as those can be reached with the minimum number of small trembles. This is one place where the upper bound of 1/4 rather than 1/3 is of the players information parameters.

Lemma B.15 Let **w** and $\mathbf{w}'$ be any two conventions such that $w' = w + e_{ij}$ for some $i \neq j$. If $\delta \leq \delta^*$ then

$$\mathbf{r}(\mathbf{w},\mathbf{w}') = k_i[1 - \frac{U_i(w_i - \delta)}{U_i(w_i)}].$$

PROOF: Suppose that the process is in the convention **w** at time t. Between date $t+1$ and date $t+k_i$, there is a positive probability that player j demands a δ more than his conventional demand, w_j exactly j^* times while the other two players continue to demand w_{-j}. This is a best response for them because they can sample the earlier records where player j has not made trembles.

At time $t+k_i$, the players have a memory in which at least $3k$ of the records consist of conventional demands and k_i demands some of which involve trembles on part of player j. There is a positive probability the between period $t+k_i$ and period $t+k_i+k_j$, player i samples these k_i most recent records while player j and k sample only records consisting of the conventional demands only. So between $t+k_i+k_j$, there is positive probability the player i will respond with $w_i-\delta$ while players j and k continue with their conventional demands.

But between $t+k_i+k_j+1$ and $t+k_i+k_j+k$, there is a positive probability that player i will continue to sample the records consisting of trembles due to player j, while player j now samples the most recent records consisting of player i lower demand. Conditional on the above event, there is a positive probability that player k will continue to play the conventional demands alone, at least $2k$ of which are left at $t+k_i+k_j$. This leads to a run of k demands of the form $w+e_{ij}$.

Now, there is a positive probability the players will sample the k most demands and the convention w+ e_{ij} will soon be established.

All this is conditional on the fact that the j^* trembles in the initial k_i imply the the best response for player i is in fact $w_i-\delta$. But the point of indifference is given by

$$j^*/k_i = [1 - U_i(w_i - \delta)/U_i(w_i)]$$

as was derived in Lemma B.13, equation B.6.

Similarly, if one were to reach the convention $\mathbf{w}+e_{ij}$ with player i making the mistake of asking for a δ less to start with, then the number of trembles is given by considering equation B.4. The lemma follows by virtue of.

Now, it can be shown in the spirit of Lemma B.13 that any other path must involve at least as many trembles as this path. The proof is complete by resorting to Lemma B.14. □

Corollary B.2 Suppose that $\mathbf{w}^1$ and $\mathbf{w}^2$ are two conventions. Then, $\mathbf{r}(\mathbf{w}^1,\mathbf{w}^2)$ *is finite, i.e., with smal' trembles it is possible to reach one convention from another.*

PROOF: Straightforward verification shows that if $w_i \geq 2\delta$ can be met in the core, then the demand $w_i - \delta$ can also be met.

Definition B.6 w-tree: A w-tree is a graph with the set of conventions as vertices such that from every vertex $\mathbf{w}' \neq \mathbf{w}$ there is a unique path directed to $\mathbf{w}$, and there are no cycle. Let $\mathcal{T}_{\mathbf{w}}$ denote the set of all w-trees.

Definition B.7 *Stochastic Potential:* The *stochastic potential* of a convention $\mathbf{w}$ is the least resistance among all w-trees:

$$\gamma(\mathbf{w}) = \min_{T \in \mathcal{T}_{\mathbf{w}}} \sum_{(\mathbf{w}^1,\mathbf{w}^2) \in T} \mathbf{r}(\mathbf{w}^1,\mathbf{w}^2).$$

The following theorem is proved in Young(1993a).

THEOREM A: *The sequence of stationary distributions μ^ϵ converges to a stationary distribution μ^0 of P^0 as $\epsilon \to 0$. Moreover a state* **s** *is stochastically stable if and only if* **s**=**w** *is a convention and $\gamma(\mathbf{w})$ has the minimum stochastic potential among all conventions.*

We now use this theorem and the information gathered about the number of trembles required to go from one convention to another to prove Theorem 3.

For each convention **w**, let $R_\delta(\mathbf{w})$ denote the player who requires the minimum number of trembles before he demands δ less than in the convention **w** as a best-response. That is,

$$R_\delta(\mathbf{w}) = \arg\min_{i=1,2,3} \left\{ \alpha_i[1 - \frac{U_i(w_i - \delta)}{U_i(w_i)} \right\}.$$

Lemma B.16 *Suppose* $\mathbf{w}^*$ *maximizes* r_δ. *Then if in a convention* $\mathbf{w}$, $w_i < w_i^*$, *then* $R_\delta(\mathbf{w}) \neq i$.

PROOF: Suppose not. Then

$$\begin{aligned} r_\delta(\mathbf{w}) &= \alpha_i[1 - \frac{U_i(w_i - \delta)}{w_i} \\ &> \alpha_i[1 - \frac{U_i(w_i^* - \delta)}{w_i^*} \\ &\geq r_\delta(\mathbf{w}^*). \end{aligned}$$

This contradicts the fact that $\mathbf{w}^*$ maximizes r_δ. □

Lemma B.17 *Suppose* $\mathbf{w}^*$ *maximizes* r_δ *and* $R_\delta(\mathbf{w}^*) = i$. *Then, if* $\mathbf{w}$ *is another maximum of* r_δ, *then* $w_i \leq w_i^*$ *and* $w_j \leq w_j^* + 2\delta$, $j \neq i$.

PROOF: Assume without loss of generality that $R_\delta(\mathbf{w}) = 1$. If $w_1^* < w_1$, since $R_\delta(\mathbf{w}) = 1$, we contradict Lemma B.16.

Observe that at $W(2) = (w_1^* - \delta, w_2^* + 2\delta, w_3^* - \delta)$ $R_\delta(\mathbf{w}(2)) = 2$ and at $W(3) = (w_1^* - \delta, w_2^* - \delta, w_3^* + 2\delta)$, $R_\delta(\mathbf{w}(3)) = 3$ to conclude as above the other two inequalities.

□

PROOF OF THEOREM 2.3:

We first prove 1. Let M be the set of maximizers of r_δ. If $R_\delta(\mathbf{w}) = i$ for all $\mathbf{w} \in M$, then player i makes the same demand in each of those conventions. Hence, by virtue of Lemma B.17, there are at most five conventions.

Suppose that for some two conventions $\mathbf{w}^1$ and $\mathbf{w}^2 \in M$, $R_\delta(\mathbf{w}) = 1$ and $R_\delta(\mathbf{w}^2) = 2$. From Lemma B.16 and Lemma B.17, it follows that $w_2^1 \leq w_2^2 \leq w_2^1 + 2\delta$. Hence there are three cases to analyze.

Suppose that $w_2^2 = w_2^1$. Since $w_3 \leq w_3^1 + 2\delta$ and $w_1 \leq w_1^1$ in every convention, there are at most six conventions.

Let $W(2) = (w_1^1 - \delta, w_2^1 + 2\delta, w_3^1 - \delta)$. Note that $R_\delta(\mathbf{w}(2)) = 2$. Hence if $w_2^2 = w_2^1 + 2\delta$, it must be because $R_\delta(\mathbf{w}(3)) = 3$ where $W(3) = (w_1^1 - \delta, w_2^1 + \delta, w_3^1)$. Hence, again from Lemma B.16 it follows that $w_3 \geq w_3^1$ in any convention that maximizes r_δ. Hence, there are again at most six conventions. The final case may be analyzed similarly.

The bound is obtained on the distance is clear from the above arguments.

Now let w^* be a maximum of r_δ. To prove 2 we need to exhibit a $\mathbf{w}^*$ that has the lowest stochastic potential among all $\mathbf{w}$ -trees. Let A_i denote the set of all conventions in which player i demands at least a δ less than in the convention $\mathbf{w}^*$. By Lemma B.16, if $W \in A_i$, then $R_\delta(\mathbf{w}) \neq i$.

Fix a demand w_k on the part of player k. Let $w_\delta^{ij}(w_k)$ be such that

$$\alpha_i[1 - \frac{U_i(x-\delta)}{U_i(x)}] \leq \alpha_j[1 - \frac{U_j(f(123) - w_k - x - \delta)}{U_j(f(123) - w_k - x)}]$$

if and only if $x \geq w_\delta^k$, $i < j$.

Then, $w_\delta^{ij}(w_k)$ is largest demand of player i for which the number of trembles required for him to demand δ less as a best response is *lower* than the number of trembles required before player j has a best response that involves a demand of δ less than in W. That is, given in W and $i < j$,

$$\arg\min \left\{\alpha_i[1 - \frac{U_i(w_i - \delta)}{U_i(w_i)}, \alpha_j[1 - \frac{U_i(f(123) - w_k - w_i - \delta)}{U_i(f(123) - w_k - w_i)}\right\} = \begin{cases} i & \text{if } w_i \geq w_\delta^{ij}(w_k) \\ j & \text{otherwise} \end{cases}$$

Now construct the $\mathbf{w}^*$-tree $\Im$ of minimum reisistance as follows:

(i). Let $W \in A_i \cap A_j$, $i < j$. In this region $R_\delta(\mathbf{w}) = k$.

1. Suppose that $w_i < w_i^* - \delta$. Then

$$\mathbf{w} \longrightarrow \mathbf{w} + e_{ki}.$$

2. If $w_i = w_i^* - \delta$, then

$$\mathbf{w} \longrightarrow \mathbf{w} + e_{kj}.$$

The resistance of each edge is $\lceil m(r_\delta(\mathbf{w})\rceil$. Note well that this is the minimum number of trembles required to reach one convention to *any other* state and not just the convention that it is directed towards.

By construction, there is now a unique path from every convention in this region to $\mathbf{w}^*+e_{ik}$.

(ii). Let $W \in A_k \cap (A_i \cup A_j)^c$, $i < j$. In this region $R_\delta(\mathbf{w})$ is either i or j.

1. If $w_i > w_\delta^{ij}(w_k)$, then $R_\delta(\mathbf{w}) = i$. In this case,

$$\mathbf{w} \longrightarrow \mathbf{w} + e_{ij}$$

2. If $w_j > w_\delta^{ij}(w_k)$, then $R_\delta(\mathbf{w}) = j$. For these conventions,

$$\mathbf{w} \longrightarrow \mathbf{w} + e_{jk}.$$

3. If $w_i = w_\delta^{ij}$, then

$$\mathbf{w} \longrightarrow \mathbf{w} + e_{ik}.$$

The above construction yields a unique path that ends in either $\mathbf{w}^*+e_{ki}$ or $\mathbf{w}^*+e_{kj}$.

(iii). For each of the six conventions of the form $\mathbf{w}^*+e_{ij}$ for some i, j, place a unique outgoing edge directed towards $\mathbf{w}^*$.

Now, by construction, there is a unique edge directed away from each convention except $\mathbf{w}^*$. Moreover, the construction yields a unique path from every convention to $\mathbf{w}^*$. The resistance between any two edges is given by the function r_δ, which is a lower bound on the resistance along any edge. Hence stochastic potential of this tree is

$$\gamma(\Im) = \sum_{\mathbf{W}} r_\delta(\mathbf{w}\ ).$$

Now, consider any other w-tree. Let e denote the edge directed away from $\mathbf{w}^*$ and r its resistance. By virtue of the corollary to Lemma B.13, $r \geq [mr_\delta(\mathbf{w}^*)]$. But on the other hand, in the construction of the graph $\Im$, the resistance of the edge directed away from the convention $\mathbf{w}$ it $m[r_\delta(\mathbf{w})]$. Since, r_δ is maximized at $\mathbf{w}^*$, it follows that the resistance of this w-tree must exceed that of the tree $\Im$ by at least $r - m[r_\delta(\mathbf{w})]$.

For large enough m, the above difference will be positive if $\mathbf{w}^*$ is the unique maximum. In that case $\Im$ will have the least stochastic potential and by virtue of Theorem A, the unique stochastically stable convention for all admissible m. If there are several conventions, that maximize r_δ, each of them will be a stochastically stable convention for large enough m. This proves 2.

Bibliography

[1] AUMANN, R.J. (1985a): On the Non-Transferable Utility Value: A comment on the Roth Shafer Examples, *Econometrica*, 54, 667-77.

[2] AUMANN, R.J. (1985b): An Axiomatization of the Non- Transferable Utility Value, *Econometrica*, 53, 599-612.

[3] AUMANN, R.J. (1986): Rejoinder, *Econometrica*, 54, 985- 989.

[4] AUMANN, R.J. AND DREZE, J.H. (1972): "Cooperative Games with Coalition Structures," *International Journal of Game Theory*, 3, 217-37.

[5] BORDER, K.(1982):*Fixed point theorems with applications to economics and game theory*, Cambridge University Press, New York, 1989

[6] BINMORE, K.G: "Nash Bargaining Theory II, and Nash bargaining theory and incomplete information," in *The Economics of Bargaining* (K.G.Binmore and P. Dasgupta, Eds.), Basil Blackwell, Oxford, 1987

[7] CARLSSON, H. (1991): A bargaining model where parties make errors, *Econometrica*, 59, 1487-1496

[8] CHUN, Y. (1989): "A New Axiomatization of the Shapley Value," *Games and Economic Behaviour*, vol 1, no:2, June, 119-130.

[9] FOSTER, D. AND YOUNG, H.P.(1990): "Stochastic Evolutionary game dynamics," *Theoretical Population Biol.*, 38, 219-232.

[10] FRIEDLIN, M. AND WENTZELL, A. *Random Perturbations of Dynamical Systems*, Springer Verlag, New York,1984.

[11] GUL, F. (1985): "Bargaining foundations of the Shapley Value," *Econometrica*, 57, 81-95.

[12] HART, S. (1985): An Axiomatization of Harsanyi's Non- transferable utility Solutions, *Econometrica*, 53, 1295-1313.

[13] HART, S. and Mas-COLELL, A. (1992): "A Model of n- Person Non- Cooperative Bargaining", Working Paper Series, Centre for Rationality and Interactive Decision Theory, Jerusalem, ISRAEL.

[14] HART, S. and Mas-COLELL, A. (1987): "Potential, Value and Consistency," *Econometrica*, 57, 589-614.

[15] KALAI, E. AND SMORDINSKY, M. (1975): "Other solutions to Nash's bargaining problem," *Econometrica*, 45, 513-518

[16] KANDORI, M., MALAITH, G. AND ROB, R. (1993): "Learning, mutation and long-run equilibrium in games," *Econometrica*, 61, 29-56

[17] KANEKO, M. AND WOODERS, M.H.(1978): "Cores of Partitioning Games," *Mathematical Social Sciences*

[18] Le BRETON, M., OWEN, G. AND WEBER, S.(1992): "Strongly Balanced Games," *International Journal of Game Theory*,,20,419-427

[19] MOULIN, H. (1988): *Axioms of Cooperative Decision Making*, New York: Cambridge University Press.

[20] NASH, J. (1950): " The Bargaining Problem," *Econometrica*, 28, 155-162.

[21] ROTH A.E. (1977): "The Shapley Value as a von Neumann- Morgenstern Utility," *Econometrica*, 45, 657-664.

[22] ROTH A.E. (1980): "Values for games without Side- Payments: Some Difficulties with Current Concepts," *Econometrica*, 48, 457-65.

[23] ROTH A.E. (1986): "On the Non-Transferable Utility Value: A Reply to Aumann," *Econometrica*, 981-984.

[24] ROTH A.E. (1988): *The Shapley Value*, New York: Cambridge University Press.

[25] SHAPLEY, L.S (1953): "A value for n-Person Games," reprinted in *The Shapley Value*, Alvin Roth ed., New York: Cambridge University Press.

[26] SHAPLEY, L.S. (1969): "Utility Comparison and the theory of games," reprinted in *The Shapley Value*, Alvin Roth ed., New York: Cambridge University Press.

[27] YOUNG, H.P. (1993a): "The evolution of conventions," *Econometrica*, 61, 57-84.

[28] YOUNG, H.P (1993b): "An evolutionary model of bargaining," *Journal of Economic Theory*, 59,1, 145-168

[29] YOUNG H.P. (1988): "Individual Contribution and Just Compensation," in *The Shapley Value*, Alvin Roth ed., New York, Cambridge University Press.