EFFICIENCY AND STABILITY IN PARTNERSHIPS

Thesis by

Patrick Legros

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1989

(Submitted on May 9, 1989)

DEDICATION

To my parents

ACKNOWLEDGEMENTS

Financial support for the research and writing of this dissertation was provided by the John Randolph Haynes and Dora Haynes Foundation, the Alfred P. Sloan Foundation and Graduate Research Assistanships awarded by the Californian Institute of Technology.

I have been fortunate to work under the supervision of John Ledyard during my three years at Caltech. His availability, encouragements and comments on my work during these years were invaluable. He set an example for me that I will try to follow in my career as a researcher and as a teacher.

I wish to thank the other members of my committee, Kim Border, Richard McKelvey, and Thomas Palfrey for allowing me to steal part of their time, and for their comments on various portions of my work.

ABSTRACT

A partnership is an organization in which the owners of the firm provide inputs into the production process and in which they have, collectively, the power to make decisions. An *institution* defines how the output of the partnership is shared among the partners and also the collective decision process that will be used. An institution should have two desirable properties: efficiency and stability. Efficiency means that the partners have an incentive to provide efficient levels of inputs (the moral hazard problem) and that the decision process selects an efficient decision. Stability means that the partners do not want to modify the institution (renegotiation proofness).

When the inputs that the partners provide are not verifiable, there is a well established belief in the literature that efficiency cannot be sustained in partnerships. The first part of the dissertation establishes, contrary to this common belief, that the moral hazard problem can be almost eliminated in partnerships: there exists an allocation of the final output which induces each partner to almost always take an efficient action. It is in fact sometimes possible for the partners to attain full efficiency: necessary and sufficient conditions are established.

The second part of the thesis considers a situation in which renegotiation takes place through a mediator. It is shown that, under some sufficient conditions on the environment, there exist collective decision making processes which are (interim) efficient and which are renegotiation proof, i.e., stable.

TABLE OF CONTENTS

Page number

LIST OF FIGURES

CHAPTER 1

PARTNERSHIPS AS A FORM OF ORGANIZATION: AN OVERVIEW

1. INTRODUCTION

Institutions have occupied until recently a very modest place in economic modeling. They have been considered mainly as parameters of the environment, exogenous data that influence the behavior of the economic agents.[1] Even if the idea that institutions should adapt to the information of the agents and to their rationality was already in the work of Hayek (1945) and Knight (1921),[2] it is one of the great achievements of the mechanism design literature to have introduced into economic theory the idea that institutions can be chosen by economic agents.[3]

One of the most fecund applications of the mechanism design approach is in the field of industrial organization and especially in the (theoretical) literature on the theory of the firm. Economists are able to ask, and sometimes to answer, questions related to the optimal form of organization of a firm. Putterman (1986), Holmström and Tirole (1987) and Hart and Holmström (1986) offer a good panorama of the recent work in this area.

A theory of the firm should be able to predict which form of organization will appear and to explain why it exists. However, besides notable exceptions like Hart and Moore (1988a) or Beckman (1984), most of the recent work has been restricted to the study of the "classical" or "capitalist" firm. The question of the optimal form of organization is reduced to a question about the optimal form of contracts between owners, managers and workers.[4] In the classical firm, owners, managers and workers occupy different spheres: the owners own the equity of the firm, the managers have the control and the workers provide the labor inputs. In particular, the ownership sphere is separated from the productive sphere and from the control sphere. This approach can be called the "agency" approach to the firm (Fama (1980)) and uses as its main theoretical tool the principal-agent model (Grossman and Hart (1983)); or the principal-principal-agent as in the three tier hierarchy of Tirole (1986).

Consequently, this literature is of little use for predicting which form of organization will arise since only one basic form is studied (again, there are exceptions like Hart and Moore (1988a)). There exist some theoretical reasons—which mainly rest on efficiency grounds—for not studying other forms of organization. But it follows that the current state of the theory of the firm is not able to explain the persistence in our societies of alternative forms of organizations. This would be a moot issue if these alternative forms were rare. However, empirical evidence suggests that this is not the case. As Rosen (1988:58) argues, the "survivor principle suggests that efficiency losses from these schemes must have been kept at tolerable proportions."

This thesis will study one particular form of organization. A partnership is an organization in which the owners of the firm provide inputs into the production process and in which they have, collectively, the power to make all decisions. The aim of the thesis is not to explain why partnerships are formed but to understand how they can be stable. (Section 2 of this Introduction presents some arguments for the appearance of partnerships.)

Whereas ownership is generally defined with respect to possession of assets, Grossman and Hart (1986) propose to define ownership with respect to ex-post decision rights of control.[5] This definition of ownership can be applied to a partnership, as defined above. Joint ownership of the firm in a partnership implies, on one hand that the partners have a *right to the share of the output* and, on the other hand that there is *collective decision making* whenever a decision has to be made (including decisions concerning changes in the organization or in the decision making process). It might help to consider the following two stylized phases of the life of a partnership. The first phase is a productive phase in which each partner chooses an amount of input which is not observed by the other partners. The second phase is a decision phase, e.g., the partners make operating decisions. An institution defines, first, how the output of the production stage is shared among the partners and, second, the collective decision process that will be used in the second phase. At each stage, the utility payoff of a partner depends not only on his own action but also on the actions of the other partners and on the institution that was chosen.

In this framework, is it possible to find an institution which is efficient and stable? Efficiency means that the consequences of opportunistic behavior are minimized and that the collective decision process selects a Pareto optimal decision (efficiency must be defined here with respect to the information structure). Stability means that the partners will not want to modify the institution once it has been chosen. It is important to note that stability does not obligatorily mean that the partnership will not be modified but that the decision process will not change. For instance, the partners might decide to dissolve the partnership. What is important is that the process by which the decision is made does not change.

The remainder of this introductory chapter is organized as follows. In the next section I present the main issues that arise when there is joint ownership. I review, in Section 3, the state of the art concerning opportunistic behavior in firms. I analyze, in Section 4, the issues that are related to the problem of collective decision making with incomplete information when the decision process can be modified by the agents.

2. PARTNERSHIPS: A DEFINITION

From an historical perspective, neoclassical, capitalist firms have made only a recent appearance in our economies. Cooperative production within tribes preceded capitalist production. Even now, most primitive tribes are organized around principles based on cooperation and sharing (Posner (1980)).

The principles on which neoclassical firms are based are no more natural than a principle like profit sharing. For instance, Weitzman (1984) argues that profit-sharing has beneficial macroeconomic effects. (See the symposium at Yale University on the share-economy which was published in the Journal of Comparative Economics (1986) for a discussion and a criticism of his theory.) Even in modern capitalist firms, profit-sharing is present. Kruse (1987) reports that twenty percent (22 million employees) of the U.S. labor force participate in over 400,000 profit-sharing plans (cited in Baker et al. (1988)).

Partnerships require a type of sharing arrangement among the partners, either in the form of a share of profits or in the sharing of the decision power. Partnerships can take different forms and arise in many different spheres of activity. I distinguish between "equity" partnerships, "production" partnerships, and "full" partnerships.

In an equity partnership, the partners invest a certain amount in a project but do not actively participate in the production phase. Examples are some joint ventures, "limited partnerships" (energy or master limited partnerships for example), physician partnerships created in order to buy expensive medical equipment or "joint-venturing" in which physicians get a share in a medical laboratory.

In a production partnership the partners participate in the joint production but do not obligatorily have the right to make all decisions. Examples are some sharecropping systems (Adams and Rask (1968), Braveman and Stiglitz (1986)), collaboration between artists in a company, team production in a classical firm.[6]

Finally, in a full partnership the agents not only participate in the production phase but also have full power for making decisions. This is the definition that I will use from now on, and I will simply refer to partnerships to mean this type of organization. Collaboration between researchers in academia, international research and development projects (Jacquemin (1987), Picard and Rey (1987), Roberts and Weitzman (1981)), consulting firms, law firms (Gilson and Mnookin (1985)), medical group practices (Gaynor and Pauly (1987)), Yugoslav self-managed firms (Jones and Svejnar (1982)), and worker cooperatives are examples of full partnerships.

Why do we observe partnerships?[7] Because partnerships involve cooperation, an obvious answer to this question is that partnerships are formed in order to exploit the gains (e.g., economies of scope) from cooperation. This approach is taken, for instance, in a recent paper by Farrell and Scotchmer (1988). (They define a partnership in the following way: an organization in which the profits of the cooperation are *shared equally* among the partners.) In their model, there is no moral hazard problem even if there is joint production: the output is fully determined by the size of the partnership and the (known) distribution of abilities within the partnership. In this world, a partnership is similar to a club (Pauly (1970)) and the principal question that one can ask is about the relationship between efficiency and the size of the partnership. Access to a partnership is function of the ability of the agent.[8] The question of the optimal size of a partnership is important. As Gilson and Mnookin (1985) show, large law firms become the norm and are more profitable than small law firms, but large law firms are more unstable than small law firms.

Hart and Moore (1988a) develop a theory of ownership rights based on the level of investment that each agent makes. These ownership rights influence not only the share of the profit that each agent will receive but also the distribution of power that will be observed. The distribution of power is measured by the Shapley value (Shapley (1953)) of a certain characteristic function game. Hart and Moore show that if a player is "indispensable" then this player should have "more" power in the decision making. In Hart and Moore (1988a), partnerships are created because of complementarities between the assets (which are in human capital in their model, but they could as well be in nonhuman capital) that the agents own initially. The interest of their approach is that it leads to a theory of endogenous distribution of power.

The notion of asset specificity is also developed in Joskow (1988) and Williamson (1975), (1988) in order to explain vertical integration. Their analysis can be applied to the creation of partnerships as well. Asset specificity is defined with respect to the degree to which the asset can be redeployed to alternative uses and users without sacrificing its productive value. For instance, in most cooperative research and development projects, the investments that the partners make are highly specific to the project. Specificity has the beneficial effect of reinforcing cooperation—as in Hart and Moore (1988a), if an asset is specific to the project, the investment is sunk and the partner wants the project to be completed—but has the negative effect of creating incentives for opportunistic behavior—through a "hold-up" effect as in Klein et al. (1981).[9]

3. OPPORTUNISTIC BEHAVIOR

From a theoretical point of view, partnerships are of interest if the inputs of the partners are not verifiable. In a contractual world, contracts can be enforced by an outside authority, a court for instance. A court can enforce a contract only if the objects of the clauses of the contract are verifiable. For this reason, verifiability is more important than observability.[10]

Since the revenue of a single partner depends upon the actions of everyone else, a partner can be induced to take an efficient action only if his marginal revenue is equal to his marginal contribution. If the allocation of the output cannot be made contingent on the individual inputs (for lack of verifiability), it is *in general* impossible to induce the partners to be efficient.[11,12] Because of the "survivor principle" put forth by Rosen (1988:58), the theoretical prediction is at variance with the reality.

Opportunistic behavior has been extensively studied in the principal-agent literature (Grossman and Hart (1983), Holmström (1979), Shavell (1979), Gjesdal (1982)). By contrast, moral hazard in partnerships has received a scant attention. I will review four solutions that have been proposed in the literature. The first three solutions (bonding, budget breaking, and repeated game strategies) have received a formal treatment while the fourth solution (peer pressure, mutual monitoring) has been the object of informal arguments.[13]

<u>Bonding</u>: In a labor market framework, Kennan (1979) proposes that the workers post bonds at the signing of a contract, before the employer incurs expenses for their training.[14] Bonding has the effect to fully insure the employer against default by the workers. (Since training can be highly specific and is time consuming, one can also imagine that the employer posts a bond in order to insure the workers.) This solution applies when the actions that the agents can take are binary (the workers can

threaten to quit, but if they work, the output is well defined). Following the terminology of Klein et al. (1978), the workers and the employer face a "hold-up" effect (a worker can ask for a larger salary in return for his participation *after* his training, i.e., once the employer has incurred the training expenses). The bonding, or double-bonding solution is particular to this type of situation. In particular, it cannot solve the moral hazard problem in partnerships because, by assumption, it is not possible to identify who defaulted.[15] Moreover, to the best of my knowledge, this solution does not seem to be applied in actual partnerships.

Budget breaking: Alchian and Demsetz (1972) and Holmström (1982) (see also Groves (1973)) propose to introduce an additional agent in order to solve the moral hazard problem in partnerships. The agent is a monitor in the Alchian and Demsetz (1972) model. As Holmström (1982) points out, monitoring is not important once we allow the additional agent to become the owner of the firm and to acquire residual rights on the profit. All that is needed is a commitment of the partners to abandon part of the profit to an outsider (a charitable organization could be the additional agent). This theory has two weaknesses. First, we have to believe that the partners will indeed abandon their property rights. Second, it does not provide an explanation for the existence of partnerships since the empirical evidence which is available (e.g., Gilson and Mnookin (1985)) suggests that the partners generally share the total residual profit. I will return to this theory in Chapter 2.

Repeated games: The output obtained by joint production in a partnership has a character of semi-public good: whatever the contribution of a partner is, he is entitled to receive a given share of the final output. The situation is similar to a multi-person prisoner's dilemma game. It has been known for a long time that cooperation can be attained in the prisoner's dilemma game when the static game is repeated infinitely often (or finitely many times if there exist two Nash equilibria to the static game as Benoit and Krishna (1987) note). A small literature has applied this idea to solve the moral hazard problem in partnerships (Radner (1986), Radner et al. (1986), Abreu et al. (1987)). This literature rests on two assumptions: an infinite horizon and a static environment (the same game is repeated over and over). The first assumption eliminates from consideration short-lived partnerships, like limited partnerships or specific cooperative research and development projects. The second assumption is at variance with the life-cycle of many partnerships: partners are likely to leave, some associates will be promoted to partners. Nevertheless, the repeated game approach has definitive value. For instance, it can be a way to formally model the concepts of "peer pressure" (Kandel and Lazear (1989)) or "corporate culture" (Kreps (1988)).

Mutual monitoring: Finally, some authors have argued that profit-sharing systems encourage mutual monitoring. This joint monitoring (coupled with bonding in some cases) could enforce cooperation if the co-workers are able to "punish" a defecting worker. As in the repeated game literature, repetition is necessary in order to implement punishments. As Baker et al. (1988:606) point out, there is a risk of *over-monitoring*, i.e., the workers might spend more time monitoring each other than producing. It is not clear, then, if the benefits of monitoring will exceed the losses in productivity. This idea is reformulated by Kandel and Lazear (1989) under the name "peer pressure". (See also Kreps (1988) for the related and vaguely defined concept of "corporate culture.") Kandel and Lazear present a model in which the partners are able to use non-monetary punishments (not made explicit in their model) in order to punish their co-workers. In their model, it is crucial that the workers are able to observe, and *evaluate properly*, the performance of each other. In team production, a production chain at G.M. for instance, such an assumption is realistic. For other partnerships, this assumption is not straightforward. (See Gilson and Mnookin (1985) on the difficulty of evaluating the performance of lawyers, even if their activities are observable *and recorded.*)

4. DECISION MAKING AND STABILITY OF A PARTNERSHIP

As Grossman and Hart (1986) and Hart and Moore (1988a) have observed, the choices of a collective decision process and of an ex-post production distribution of power influence the levels of investment that will be made before the production phase. (Both papers are mainly concerned with efficient investment levels rather than with efficient production.) The choice of a collective decision process has also important consequences for the stability of a partnership when there is a possibility of renegotiation.[16] Renegotiations can be induced by a hold-up effect (as in the discussion of the bonding solution in Section 2) or by a change in the information structure (as in the stylized model of Section 1 in which the partners acquire some private information after the production phase).

A literature is developing on the situation where the renegotiation process is directed by a single agent. (See Myerson (1983), Tirole (1986), Green and Laffont (1987), Hart and Tirole (1987), Fudenberg and Tirole (1988), among others.) Very few studies look at situations in which more than one agent can renegotiate. Myerson (1984), Holmström and Myerson (1983), Crawford (1985), Cramton (1985), Hart and Moore (1988b) seem to exhaust the list of the relevant papers.

A decision process is a certain type of mechanism. In a world of incomplete information, when the agents renegotiate, they have to take into account the inferences that the other players will make during the renegotiation phase. It might be the case that the fact that a player does not renegotiate when he has the possibility to do so is by itself informative for the other players (Myerson (1983)). Modelling renegotiation is a difficult exercise. Ideally, we would like to give as much freedom as possible to the agents but at the same time we are limited by the possibilities of solving a model that is too general. As Crawford (1985) points out, it is necessary to anchor the game theoretical analysis. This means, in particular, to assume from the outset certain rules for renegotiating.

Holmström and Myerson (1983) suppose that the agents renegotiate through a mediator. That is to say, the mediator proposes to the agents to modify the initial mechanism (or institution) and lets them vote between this initial mechanism and an alternative mechanism. An alternative mechanism is selected whenever the agents unanimously vote for it. Stability in this framework means that the agents will never unanimously agree to play another mechanism. Cramton (1985) and Hart and Moore (1988) allow more freedom in the renegotiation; in particular, the agents renegotiate without a mediator by making proposals and counter-proposals.

Cramton (1985) looks at a buyer-seller sequential bargaining model when both agents are impatient (i.e., they discount the future). In his model, the mechanism specifies not only the price at which the good will be exchanged but also the time at which the exchange will happen. (When the players reveal their types, the mediator will not tell them *when* the exchange will happen, otherwise the agents would not want to wait.) In the framework of Holmström and Myerson (1983), ex-ante efficiency is defined with respect to the restricted set of incentive compatible mechanisms. Cramton shows that ex-ante efficient mechanisms *without renegotiation* are not robust to the renegotiation process, i.e., that they will violate sequential rationality (sequential rationality means that inefficiencies are never common knowledge). Since renegotiation imposes more restrictions on the set of mechanisms, it is not surprising that ex-ante efficiency might be incompatible with sequential rationality. For instance, Holmström and Myerson (1983:1806) give an example in which no classically efficient decision rule can be incentive compatible. The result of Cramton (1985) that some ex-ante efficient mechanisms are not sequential rational is akin to this example. In the static framework, the players understand the constraints imposed by strategic behavior and realize that even if a mechanism gives to every agent a larger ex-ante utility payoff than his ex-ante payoff in an ex-ante efficient mechanism, these payoffs are not realizable because incentive compatibility is violated. In the dynamic framework of Cramton, ex-ante efficiency should be defined with respect to the class of mechanisms which do not violate sequential rationality.

Hart and Moore (1988b) propose a model of renegotiation between a buyer and a seller who are *locked into* a relationship and can renegotiate the initial contract by sending messages following a certain technology. Their main result is the identification of sufficient conditions under which the first best can and cannot be attained. Without the fact that the agents are locked in their relationship, their work would be very similar to the work of Rubinstein (1982) and, to a lesser extent, Cramton (1985). The originality of their work is to model explicitly the communication technology and to allow the players to use the imperfections of this technology (for instance, if certified mail does not exist, an agent can send a message and the agent who receives it can deny that he received it).

It is not clear that it will be possible to extend easily the results of Cramton (1985) and Hart and Moore (1988b) to an environment involving more than two agents. In order to address the question of the stability of a mechanism in a more general environment, I will follow the approach of Holmström and Myerson (1983) and will consider a situation in which the agents renegotiate through a mediator.

4. OUTLINE OF THE RESULTS

The rest of this thesis will follow the order of the two preceding sections. Chapter 2 is devoted to the analysis of the moral hazard problem in partnerships. The main results of Chapter 2 are the following.

First, I show that by relaxing in a negligible way the efficiency requirement, it is possible for the partners to be almost always efficient. Thus, opportunistic behavior can be minimized and almost disappear if one allows a small loss in efficiency. This result differs from the other solutions that were reviewed in Section 3 in that it is compatible with observed compensation schemes used in partnerships. An interpretation of the result is the following. Following the Bayesian approach to game theory, one can consider that before providing an input into the production process, a partner has a certain "idea" of the inputs that will be provided by the other partners (or of the frequency with which the inputs will be provided). The problem for this partner is then similar to a problem of decision under uncertainty. The main result of Chapter 2 states that there exists a sharing rule and certain "ideas" that the partners can have which will induce them to be almost always efficient.[17]

Second, I offer a complete characterization of the situations in which the partners can sustain full efficiency. This characterization explains why it might be in the interest of the partners to choose a manager among themselves and to rotate the management function (as in Beckman (1984)).

Chapter 3 analyzes a renegotiation process in the spirit of the Holmström and Myerson (1983) model. I introduce the concept of strong durability and characterize the class of mechanisms which are strongly durable. When we consider renegotiation, it is not clear that the revelation principle can be applied, especially when there are potentially many equilibria in a direct mechanism. I will show that for strongly durable mechanisms, there is no loss of generality in considering status-quo which are incentive compatible direct mechanisms.

Strongly durable incentive compatible direct mechanisms are always interim efficient, and a mechanism is not strongly durable if, and only if, there exists an interim efficient mechanism which is not interim payoff equivalent to the initial mechanism and which is selected with probability one by all the types of all the players. Strongly durable mechanisms have also the "focal" property that all the equilibria of the mechanism are interim payoff dominated by the truthful equilibrium. In fact, it can be shown that, under some regularity conditions, a direct mechanism is strongly durable only if all its equilibria are interim payoff equivalent to the truthful equilibrium. This gives a stronger rationale for using the revelation principle in environments in which renegotiation through a mediator takes place.

NOTES

1 As Coase (1937), Arrow (1974) and Williamson (1975) have pointed out, organizations arise because of market failures (e.g., transaction costs), i.e., the inability of the price system to achieve satisfying collective action. It follows that the choice among different types of organizations depends upon the success of each type in reducing the underlying market failure(s). The goal that can be assigned to a theory of organization has been put forth by Arrow (1987) in a different context "... the New Institutional Economics movement ... [does] not consist primarily of giving new answers to the traditional questions of economics—resource allocation and degree of utilization. Rather it consists of answering new questions, why economic institutions have emerged the way they did and not otherwise; it merges into economic history, but brings sharper nanoeconomic reasoning to bear than has been customary" (1987:734).

2 For instance, Hayek (1945) had the idea that institutions should adapt to the information available in the society: "Which of these systems is likely to be more efficient depends mainly on the question under which of them we can expect that fuller use will be made of the existing knowledge. This, in turn, depends on whether we are more likely to succeed in putting at the disposal of a single authority all the knowledge which ought to be used but which initially dispersed among many different individuals, or in conveying to the individuals such additional knowledge as they need in order to enable them to dovetail their plans with those of others."

3 See Groves and Ledyard (1987) for a survey.

4 A firm can be defined by a nexus of contracts (e.g., Jensen and Meckling (1976), Hart (1988)).

5 In the model of Grossman and Hart (1986), ex-ante investment decisions are

followed by ex-post operating decisions once the state-of-the-world realizations obtain. (This realization is common knowledge to all the players. Thus, their model is not a model of decision making under incomplete information at the decision stage.) Grossman and Hart (1986) ask which assignment of ex-post decision rights is best, given that the ex-ante investment decisions are a function of the distribution of power at the decision phase.

6 The usage of the word "team" can be misleading. In the original work of Marschak and Radner (1972) (see also Arrow and Radner (1979), Groves and Radner (1972)), the members of the team share the same objectives but do not have the same information. Starting with Groves (1969), (1973), a second literature on teams continues to suppose that the agents have different information but it introduces the possibility for the agents to have different goals as well: this is the incentive problem. The team literature is concerned with the possibility for a center to achieve an efficient decision when the information is decentralized. There is a third literature, on *syndicates*, that considers the possibility for a group of agents with differing information and differing utility functions to achieve a collective decision (Wilson (1968), (1978), Kobayashi (1980)). A syndicate is a team without a center. The question is to construct a surrogate group utility function for the syndicate (see also Hylland and Zeckhauser (1979) for related work). These three literatures are concerned with the optimal use, or the manipulation, of information. Opportunistic behavior in these literatures is associated with the incentives to misrepresent one's information and is an adverse selection phenomenon while opportunistic behavior in partnerships is principally related to the moral hazard literature. (Clearly, it is possible to combine the two aspects, adverse selection and moral hazard, e.g., Picard and Rey (1987), McAfee and McMillan (1988).)

7 How do partnerships function? Empirical evidence on the functioning of partnerships is generally lacking. However, some evidence is available for law firms. The excellent article by Gilson ad Mnookin (1985) is a good reference for understanding the issues (see also Flood (1985)). Concerning the sharing system in law firms, historically compensation was a direct function of seniority. This type of compensation system continues to be widely applied. If anything else, there is a growing concern in the legal profession (Gilson and Mnookin (1985:318 ff.)) about the correct design of the compensation scheme.

8 Farrell and Scotchmer give the interesting example of the salmon fishermen in the Pacific Northwest who are organized in "partnerships," where fishermen in the same partnership share information (through a secret code) about the schools of salmon. Such sharing of information is akin to the sharing of information in oligopolies.

9 Admati and Perry (1987) analyze a model in which the parties avoid to become "completely sunk" into the relationship by making incremental investments instead of a one-shot investment.

10 Observability is by itself an ill-defined concept. It is possible to observe the input of a partner without being able to evaluate its quality. For instance, actors on stage can observe each other but they might disagree in their evaluation of their relative performances. (The fact that critics generally disagree in their evaluations reinforces the argument.) It is possible to observe the number of hours that a lawyer works on a case but it is difficult to assess his relative performance (see Gilson and Mnookin (1985)).

11 This observation is not recent. Smith (1776) made the early comment that the type of ownership and of distribution of risks will influence the risk taking behavior of the agents (e.g., his discussion about joint stock companies versus

copartneries [sic] pages 690-ff of *The Wealth of Nations*).

12 One of the principal results of Chapter 2 is that this result is not as general as it is believed to be.

13 Radner and Williams (1988) and Rasmusen (1987) take an approach that is similar to the one that I take in Section 3 of Chapter 2. I will postpone the discussion of their papers until Chapter 2.

14 Becker (1964) and Williamson (1975) note that when investment costs are shared, as in joint ventures or when an employer trains a worker, the future returns will be also shared and that this could lead to ex-post contract enforcement problems due to the opportunistic behavior of the parties. This is similar to the "hold-up" effect in Klein et al. (1978).

15 The bonding solution is also proposed by Holmström (1982:328). Specifically, Holmström suggests that the partners post a bond equal to the value of the optimal output level and that each partner receives, after production, the totality of what is produced. It is not clear who will be penalized and what happens to the bonds if the optimal output is not produced.

16 Cramton et al. (1987) show that if the decision process has to do with the dissolution of the partnership, then it is, in fact, easier to dissolve partnerships in which the initial shares are distributed among the partners than if only one partner has all the shares. This result suggests that if partnerships are unstable (as Gilson and Mnookin (1985) note for large law firms), it is easier to efficiently dissolve them than a capitalist firm. This is only a theoretical argument; I do not know of any evidence or empirical analysis of the process by which partnerships dissolve.

17 It is natural to question the origin of these "ideas", or beliefs, that the partners have about each other's plan of action. I do not provide an answer to this

question. However, one can imagine that the beliefs come from past experience, knowledge of the partners in other circumstances, or ex-ante communication.

REFERENCES

Abreu, Dilip, Milgrom, Paul and Pearce, David (1987). "Information and Timing in Repeated Partnerships." Mimeo.

Adams, Dale W. and Rask, Norman (1968). "Economies of Cost-Share Leases in Less-Developed Countries," *American Journal of Agricultural Economics*, 50, 935-42.

Admati, Anat R. and Perry, Motty (1987). "Joint Projects Without Commitment." Mimeo.

Alchian, Armen A. and Demsetz, Harold (1972). "Production, Information Costs and Economic Organization," *American Economic Review*, 62, 777-85.

Arrow, Kenneth J. (1974). *The Limits of Organization.* New York: W.W. Norton & Company, Inc.

Arrow, Kenneth J. (1987). "Reflections on the Essays," in George Feiwel, ed., *Arrow and the Foundations of the Theory of Economic Policy.* New York: New York University Press.

Arrow, Kenneth J. and Radner, Roy (1979). "Allocation of Resources in Large Teams," *Econometrica*, 47(2), 361-85.

Baker, George P., Jensen, Michael C. and Murphy, Kevin J. (1988). "Compensation and Incentives: Practice vs. Theory," *The Journal of Finance*, 43(3), 593-616.

Becker, Gary S. (1964). *Human Capital.* New York: Columbia University Press.

Beckman, Martin J. (1984). "Hierarchy vs. Partnership," *Journal of Economic Behavior and Organization*, 5, 237-45.

Benoit, Jean-Pierre and Krishna, Vijay (1985). "Finitely Repeated Games," *Econometrica*, 45, 905-22.

Braverman, Avishay and Stiglitz, Joseph (1986). "Cost-Sharing Arrangements under

Sharecropping: Moral Hazard, Incentive Flexibility, and Risk," *American Journal of Agricultural Economics*, August, 642-52.

Coase, Ronald (1937). "The Nature of the Firm," *Economica*, 4, 386-405.

Cramton, Peter (1985). "Sequential Bargaining Mechanisms," in *Game Theoretic Models of Bargaining*, edited by Alvin Roth, Cambridge: Cambridge University Press, chapter 8, pp. 149-79.

Cramton, Peter, Gibbons, Robert and Klemperer, Paul (1987). "Dissolving a Partnership Efficiently," *Econometrica*, 55(3), 615-32.

Crawford, V. (1985). "Efficient and durable decision rules: a reformulation." *Econometrica*, 53(4), 817-35.

Fama, Eugene F. (1980). "Agency Problems and the Theory of the Firm," *Journal of Political Economy*, 88(21), 288-307.

Farrell, Joseph and Scotchmer, Suzanne (1988). "Partnerships," *The Quarterly Journal of Economics*, 103(2), 279-97.

Flood, John (1985). *The Legal Profession in the United States.* Chicago: The American Bar Foundation.

Fudenberg, Drew and Tirole, Jean (1988). "Moral Hazard and Renegotiation in Agency Contracts." Mimeo, M.I.T..

Gaynor, Martin and Pauly, Mark (1987). "Alternative Compensation Arrangements and Productive Efficiency in Partnerships: Evidence from Medical Group Practice." Working Paper No. 2170, National Bureau of Economic Research.

Gilson, Ronald J. and Mnookin, Robert H. (1985). "Sharing Among the Human Capitalists: An Economic Inquiry into the Corporate Law and How Partners Split Profits," *Stanford Law Review*, 37(2), 313-92.

Gjesdal, Frøystein (1982). "Information and Incentives: The Agency Information

Problem," *Review of Economic Studies*, 49, 373-90.

Green, J. and Laffont, J.J. (1987). "Renegotiation and the Form of Efficient Contracts." D.P. 1338, Harvard University.

Grossman, Sanford J. and Hart, Oliver D. (1983). "An analysis of the principal-agent problem," *Econometrica*, 51(1), 7-46.

Grossman, Sanford J. and Hart, Oliver D. (1986). "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94(4), 691-719.

Groves, Theodore (1969). "The Allocation of Resources Under Uncertainty: The Informational and Incentive Roles of Prices and Demand in a Team." Technical Report No. 1, Center for Research in Management Science, University of California, Berkeley.

Groves, Theodore (1973). "Incentive in Teams," *Econometrica*, 41(4), 617-31.

Groves, Theodore and Ledyard, John O. (1987). "Incentive Compatibility Since 1972," in *Information, Incentives, and Economic Mechanisms. Essays in Honor of Leonid Hurwicz*, edited by Theodore Groves, Roy Radner and Stanley Fisher, Minneapolis: University of Minnesota Press, pp. 48-111.

Groves, Theodore and Radner, Roy (1982). "Allocation of Resources in a Team," *Journal of Economic Theory*, 4, 415-41.

Hart, Oliver (1988). "Incomplete Contracts and the Theory of the Firm," *Journal of Law, Economics, and Organization*, 4(1), 119-39.

Hart, Oliver and Holmström, Bengt (1986). "The Theory of Contracts." Mimeo. Presentation at the Fifth World Congress of the Econometric Society, Cambridge, Massachusetts.

Hart, Oliver and Moore, John (1988a). "Property Rights and the Nature of the Firm," Discussion Paper No. TE/88/174, London School of Economics.

Hart, Oliver and Moore, John (1988b). "Incomplete Contracts and Renegotiation," *Econometrica*, 56(4), 755-85.

Hart, O. and Tirole, J. (1987). "Contract Renegotiation and Coasian Dynamics." W.P. 442, M.I.T..

Hayek, Friedrich (1945). "The Use of Knowledge in Society," *American Economic Review*, 35(4), 519-30.

Holmström, Bengt (1979). "Moral Hazard and Observability," *The Bell Journal of Economics*, 10(1), 74-91.

Holmström, Bengt (1982). "Moral Hazard in Teams," *The Bell Journal of Economics*, 13, 324-40.

Holmström, B. and Myerson, R.B. (1983). "Efficient and durable decision rules with incomplete information." *Econometrica*, 51(6), 1799-819.

Holmström, Bengt and Tirole, Jean (1987). "The Theory of the Firm." Manuscript. Forthcoming in the Handbook of Industrial Organization, R. Schmalensee and R. Willig, eds.

Hylland, Aanund and Zeckhauser, Richard (1979). "The Impossibility of Bayesian Group Decision Making with Separate Aggregation of Beliefs and Values," *Econometrica*, 47(6), 1321-36.

Jacquemin, Alexis (1987). "Comportements collusifs et accords en recherche-développement," *Revue d'Economie Politique*, 1, 1-23.

Jensen, Michael C. and Meckling, William H (1976). "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3, 305-60.

Jones, Derek C. and Svejnar, Jan (1982). *Participatory and Self-Managed Firms.* Lexington: Lexington Books.

Joskow, Paul L. (1988). "Asset Specificity and the Structure of Vertical Relationships:

Empirical Evidence," *Journal of Law, Economics, and Organization*, 4(1), 95-117.

Journal of Comparative Economics (1986). "The Share Economy: A Symposium," edited by William Nordhaus and Andrew John, 10(4), 414-75.

Kandel, Eugene and Lazear, Edward P. (1989). "Peer Pressure and Partnership." Working Paper E-89-5, The Hoover Institution, Stanford University.

Kennan, John (1979). "Bonding and the Enforcement of Labor Contracts," *Economics Letters*, 3, 61-6.

Klein, Benjamin, Crawford, Robert G. and Alchian, Armen A. (1978). "Vertical Integration, Appropriable Rents and the Competitive Contracting Process," *Journal of Law and Economics*, 21, 297-326.

Knight, Frank (1921). *Risk, Uncertainty and Profit.* Haper Torchbooks, New York: Harper & Row, Publishers, Inc.

Kobayashi, Takao (1980). "Equilibrium Contracts for Syndicates with Differential Information," *Econometrica*, 48(7), 1635-65.

Kreps, David (1988). "Corporate Culture and Economic Theory." Mimeo, Stanford University.

Kruse, Douglas L. (1987). "Profit-Sharing and Employment Variability: Microeconomic Evidence." Mimeo, Department of Economics, Harvard University, November.

Marschak, Jacob and Radner, Roy (1972). *Economic Theory of Teams.* Cowles Foundation for Research in Economics at Yale University, Monograph 22. New Haven and London: Yale University Press.

McAfee, Preston R. and McMillan, John (1988). "Optimal Contracts for Teams." Mimeo.

Myerson, R.B. (1983). "Mechanism design by an informed principal." *Econometrica*, 51, 1767-97.

Myerson, R.B. (1984). "Two-person bargaining problems with incomplete information." *Econometrica*, 52(2), 461-87.

Pauly, Mark (1970). "Cores and Clubs," *Public Choice*, 9, 53-65.

Picard, Pierre and Rey, Patrick (1987). "Incentives in Cooperative Research and Development." Mimeo, C.E.P.R.E.M.A.P., France.

Posner, Richard A. (1980). "A Theory of Primitive Society, with Special Reference to Law," *Journal of Law and Economics*, 23, 1-53.

Putterman, Louis. *The Economic Nature of the Firm.* Cambridge: Cambridge University Press, 1986.

Radner, Roy (1986). "Repeated Partnership Games with Imperfect Monitoring and No Discounting," *Review of Economic Studies*, 53, 43-57.

Radner, Roy, Myerson, Roger and Maskin, Eric (1986). "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria," *Review of Economic Studies*, 53, 59-69.

Radner, Roy and Williams, Steven (1988). "Efficiency in Partnership when the Joint Output is Uncertain.," Discussion Paper No. 760, Center for Mathematical Studies in Economics and Management Science, Northwestern University, Evanston, Ill..

Rasmusen, E. (1987). "Moral hazard in risk-averse teams," *Rand Journal of Economics*, 18(3), 428-35.

Roberts, Kevin and Weitzman, Martin L. (1981). "Funding Criteria for Research, Development, and Exploration Projects," *Econometrica*, 49(5), 1261-88.

Rosen, Sherwin (1988). "Transaction Costs and Internal Labor Markets," *Journal of Law, Economics, and Organization*, 4(1), 49-64.

Rubinstein, Ariel (1982). "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97-110.

Shapley, Loyd S. (1953). "A Value for n-Person Games," in *Contributions to the Theory of Games*, H. Kuhn and A. Tucker eds. Princeton: Princeton University Press.

Shavell, Steven (1979). "Risk Sharing and Incentives in the Principal and Agent Relationship," *The Bell Journal of Economics*, 10(1), 55-73.

Smith, Adam (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by Edwin Cannan from the text of the 5th edition, New York: Modern Library.

Tirole, Jean (1986). "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations," *Journal of Law, Economics, and Organization*, 2(2), 181-214.

Tirole, J. (1986). "Procurement and renegotiation," *Journal of Political Economy*, 94(2), 235-59.

Weitzman, Martin (1984). *The Share Economy*. Cambridge: Harvard University Press.

Williamson, Oliver E. (1975). *Market and Hierarchies: Analysis and Antitrust Implications*. New York: The Free Press.

Williamson, Oliver E. (1988), "Corporate Finance and Corporate Governance," *The Journal of Finance*, 43(3), 567-91.

Wilson, Robert (1968). "The Theory of Syndicates," *Econometrica*, 36(1), 119-32.

Wilson, Robert (1978). "Information, Efficiency, and the Core of an Economy," *Econometrica*, 46(4), 807-16.

CHAPTER 2

SUSTAINABILITY IN PARTNERSHIPS

1. OVERVIEW

1.1. *INTRODUCTION*

It is well accepted that efficiency cannot be sustained in partnerships when the partners act noncooperatively and when actions are not observable (e.g., Holmström (1982)). This chapter shows that for any positive ϵ, it is possible for the partners to sustain ϵ-efficiency. It follows that monitoring or budget breaking is not necessary at all in partnerships if one accepts some efficiency loss, where this loss can be made as small as we want. This conclusion is in sharp contrast with the diverse theories of the firm which are based to a lesser or greater extent on the notion that monitoring or a separation of control and production is necessary in order to overcome problems of moral hazard.[1] In particular, the result of this chapter implies that it is necessary to reconsider the relative advantages of the "classical firm" versus partnerships.

A way to interpret the result is to imagine that the partners design a group mechanism which consists of a sharing rule and of a randomizing device. The randomizing device tells each partner what action to take without telling him what actions the other partners are told to take. The sharing rule always satisfies the condition that the output is fully allocated among the partners. A group mechanism can be contracted by the partners before production begins. Sustainability means in this framework that the partners will always take the action that they are told to take when they believe that the other partners also obey the recommendations of the device. This approach is very similar to the approach of Myerson (1986) and Forges (1986) on generalized mechanisms with communication.

Another way to interpret the result is that the partners can design a sharing rule for which there exists a mixed strategy equilibrium of the resulting game. Following a recent literature on the Bayesian foundations of game theory (Mertens and Zamir (1986), Aumann (1987), Brandeburger and Dekel (1987), Tan and Werlang (1988)) we can interpret these mixed strategies not as a lottery that each player designs for choosing his plan of action but as beliefs that each player has about the plans of actions of all the other players. Thus, if the partners are Bayesian players, at the signing of the contract, i.e., when the sharing rule is agreed upon, we can suppose that each player has in mind the possible plans of actions of the other partners. Following the signing of the contract, each partner will have to solve a problem of decision making under uncertainty, where the uncertainty is about the actions that the other partners will take. For consistency, one can require that the plans of actions of each player coincide with the beliefs that the other players have about his actions. (With this interpretation, a mixed strategy equilibrium has strong similarities with the concept of equilibrium introduced in Hayeck (1945)). The main result of this chapter is that there exist a structure of beliefs and a sharing rule for which each partner will find optimal to be almost always efficient.

With either of these two interpretations, the message of this chapter is that the partners can always shape their environment in such a way that moral hazard has "almost" no negative impact on efficiency. Thus, it is possible for partnerships to fulfill, without the presence of a monitor, the objective set forth in Alchian and Demsetz (1972)

> "*The economic organization through which input owners cooperate will make better use of their comparative advantage to the extent that it facilitates the payment of rewards in accord with productivity.*"

These results are important for the theory of the firm. Indeed, it has been argued (Alchian and Demsetz (1972) or Holmström (1982)) that, since efficiency cannot be sustained in partnerships, optimal forms of organizations must be characterized by a separation of ownership and active participation in firm decisions. This chapter shows the limits of this argument and proves that partnerships are not less efficient forms of organization. Because the existence of the classical firm can no longer be defended on simple efficiency arguments, it is necessary to develop alternative theories explaining its appearance. Similarly, it is not possible anymore to consider the existence of partnerships or of profit sharing firms in capitalist economies as pathological. There must be reasons, not based on efficiency grounds, for the simultaneous existence of different forms of organizations in our societies.

This chapter also gives necessary and sufficient conditions for attaining full efficiency. Somewhat surprisingly, the characterization of partnership problems which have a solution has never appeared, to my knowledge, in the literature. An example of sufficient condition is that there always exists a partner who cannot "mimic" the deviation of other partners. For instance, whenever it is possible to identify who *did not deviate*, full efficiency can be attained. The model of Holmström (1982) is a particular illustration of such a phenomenon. In the models of Alchian and Demsetz (1972) and of Holmström (1982), it is possible to attain efficiency once a manager is introduced because the manager cannot "mimic" the deviation (from efficiency) of a worker.

This characterization result suggests that, if some control or coordination is necessary, partnerships could rotate the management function among themselves. An advantage of such a system is that it enables the partner-manager to play the role of budget breaker. (See Beckman (1984) on this idea).

It can be argued that, concerning the design of incentive efficient sharing rules, there is no difference between partnerships and firms.[2] In the latter case, there is one more claimant to the profits. Thus, the manager can be seen as a member of a team whose actions do not influence directly the level of production, e.g., the manager is used only to police the workers' actions. However, the distinction between a partnership and a firm becomes important when we consider the process by which the reward scheme is chosen. In a firm, the manager alone chooses the reward scheme for the workers; in a partnership, the decision is made by all the members. The distinction is also important when there are decisions to be taken after production has taken effect; in the agency approach to the firm, workers have no rights in the decision process (e.g., choice of investment) while in a partnership, all the partners have a say in the decision process.

There is some relation between the approach taken here and the standard mechanism design problem in principal and agents models (e.g. Myerson (1983)). Indeed, one can think of a fictitious principal whose utility function corresponds to a social welfare function (e.g., the sum of the individual players' utilities) and who always gets a zero share of the output. The problem of the partners (attaining efficiency) is equivalent to the problem in which the benevolent principal maximizes his utility. In the terminology of Myerson (1983), the partnership problem has a solution if and only if the (benevolent) principal can design an incentive-efficient mechanism (efficiency refers to the principal maximizing his utility, incentive refers to the fact that it is an equilibrium strategy for the players to act as the principal wants them to behave). When the necessary and sufficient conditions exhibited in the second part of this chapter fail to be satisfied, there is no incentive-efficient mechanism; this non-existence result is in sharp contrast with existence results which are obtained in the principal and agents literature. This chapter shows that there is always ϵ-incentive-efficient mechanisms (which I call ϵ-equilibrium group mechanisms).

The techniques used in this chapter rely on very mild assumptions about the production and the disutility functions. In particular, assumptions of concavity and differentiability are not necessary. All the results are obtained under the assumption of separability of the utility functions and of risk neutrality.

This chapter is organized as follows. In next part of this section I analyze an example and give an intuition for the results. I introduce in Section 2 the concept of group mechanism and presents the main result of this chapter, i.e., that ϵ-efficiency can always be sustained by group mechanisms when ϵ is positive. I offer in Section 3 a characterization of the partnership problems for which full efficiency can be implemented in Nash equilibrium (this is the "classical problem" that is addressed in the literature). An alternative proof of the negative result of Holmström (1982) appears in Section 3.1; this proof will allow a simple explanation of why a positive result can be obtained. The characterization result appears in Section 3.2. In Section 4, I propose an illustration of the previous results and analyze the question of limited liability in two examples. The proofs of the results which did not appear in the text appear in Section 5. Concluding reamarks an extensions for future research appear in Section 6.

1.2. *AN EXAMPLE*

There are two partners, Mr. 1 and Ms. 2. The action space of each partner is the interval [0,2]. The output function is linear, i.e., $y(a_1,a_2)=a_1+a_2$, for all i and $a_i \in [0,2]$. For each output level y, $s(y)$ denotes the share of Mr. 1. $s(y)$ can be nonpositive. The utility function of each partner is given by

$$u_1(y(a),a_1)=s(y(a))-v_1(a_1)$$

$$u_2(y(a),a_2)=y(a)-s(y(a))-v_2(a_2),$$

where $v_i(a_i)=a_i^2/2$, for all $i=1,2$.

Efficiency is attained in this example when $a^*=(1,1)$, the sum of the utilities being $U^*=1$. For any sharing rule s, (1,1) cannot be an equilibrium strategy since for any $x\neq 1$ either $s(2)-s(1+x)<(1-x^2)/2$ or $s(2)-s(1+x)>(1-x^2)/2$. In the first case, Mr. 1 wants to take the action $x\neq 1$, in the other case, Ms. 2 wants to take the action $x\neq 1$. Consequently, efficiency is incompatible with strategic behavior.

Suppose now that we relax the efficiency requirement, and that we impose that at the signing of the contract the sum of the expected utility of the agents is ϵ-close to the level attained at efficiency. With this simple modification, it is no longer necessary to force the agents to be always efficient; they can be inefficient with some positive probability. The main result of this chapter will show that this is enough to reconcile strategic behavior and efficiency. For a given $\epsilon>0$, there are in general many solutions which will lead to ϵ-efficient in equilibrium. In the example of this section, there is an especially simple solution that I describe now.

Let $\tilde{A}_1=[0,1]$, $\tilde{A}_2=[1,2]$ denote two subsets of the original action sets. Let $\Gamma=\langle A_1,A_2,y,u_1,u_2\rangle$ be the original partnership game and let $\tilde{\Gamma}=\langle\tilde{A}_1,\tilde{A}_2,\tilde{y},\tilde{u}_1,\tilde{u}_2\rangle$ be the reduced partnerhip game (where the functions $\tilde{y}$, $\tilde{u}_1$, $\tilde{u}_2$ are the restrictions of the functions y, u_1, u_2 to the spaces $\tilde{A}_1\times\tilde{A}_2$, $\tilde{A}_1$, $\tilde{A}_2$). The vector (1,1) is still optimal in the reduced game $\tilde{\Gamma}$.

If we impose only an ϵ-efficiency requirement, Mr. 1 can take action 0 with positive probability and Ms. 2 can take action 2 with positive probability. Suppose that a group mechanism selects only pairs of actions in the set $\{(0,1),(0,2),(1,1),(1,2)\}$

and suppose that a level of output strictly greater than 3 is observed after the production phase. Then, it must be true that Mr. 1 deviated: the only possibility for obtaining an output strictly greater than 3 is that Mr. 1 takes an action strictly greater than 1. (It is possible that both partners deviate but this is irrelevant here.) Mr. 1 can credibly commit to be "punished" when an output strictly greater than 3 is observed because he can be sure not to be punished if he does not take an action strictly greater than 1. Similary, if an output level strictly lower than 1 is observed, Ms. 2 must have deviated and can credibly commit to be punished in this case.

Consequently, if one can find a solution to the partnership problem in the reduced game $\tilde{\Gamma}$, it will be possible to find a solution to the initial partership problem by using such punishments. It is crucial that each partner believes that the other partner will be inefficient with some probability for this scheme to function. In particular, even if (1,1) is an equilibrium in the reduced game for some sharing rule, a group mechanism cannot put all the probability mass on this vector of actions because then no partner will believe that an output greater than 3 or smaller than 1 will be realized when he is the only one to deviate.

I provide in Appendix A a formalization of these arguments and I show that there exists group mechanisms in which Mr. 1 is asked to take only one of the two actions {0,1} and Ms. 2 is asked to take only one of the two actions {1,2}. Figure 2.1 gives an illustration of two possible sharing rules which are compatible with such group mechanisms. When the output is smaller than 1, $s(y)$ is very large positively, i.e., Ms. 2 is punished and when the output is larger than 3, $s(y)$ is very large negatively, i.e., Mr. 1 is punished. Thus, in these group mechanisms, each partner will behave as if he or she could take *only one of two actions* (*low-high for* Mr. 1, *high-very high for* Ms. 2) and the probability that each partner takes an efficient action can be made as close as we want from 1.

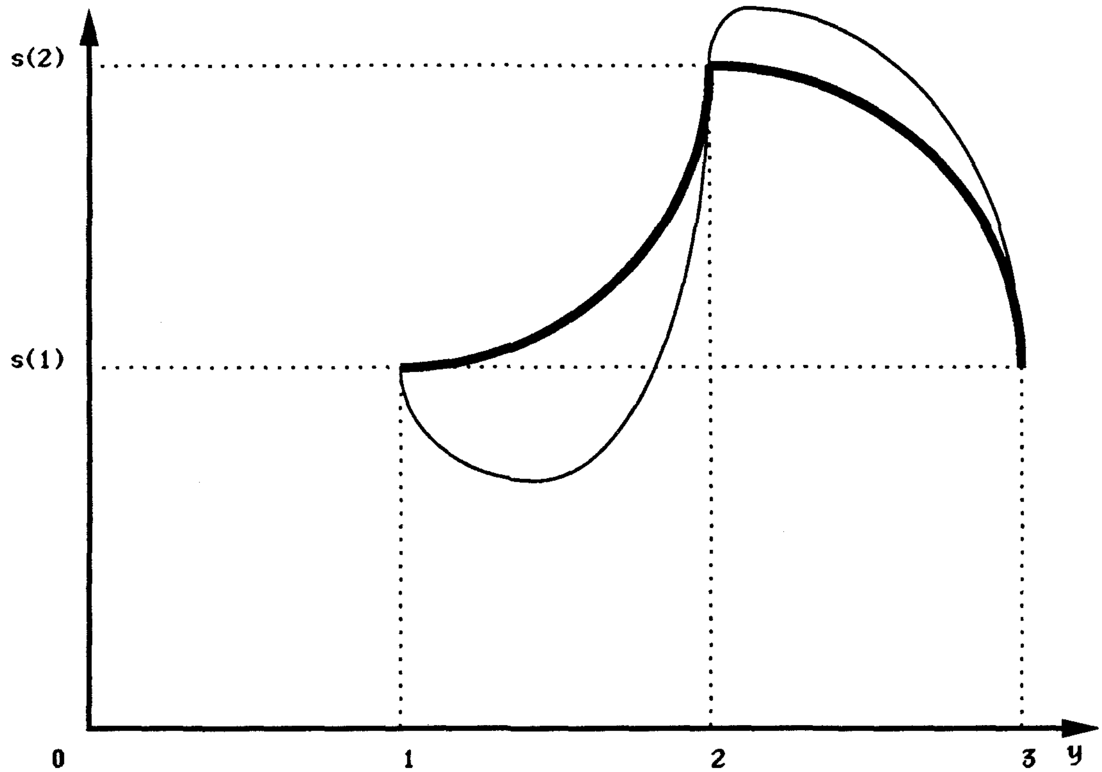

Figure 2.1

Examples of sharing rules in the

reduced game of the example of Section 1.2

2. ϵ-EFFICIENCY AND GROUP MECHANISMS

2.1. *THE MODEL*

There are n partners indexed by $i=1,\ldots,n$. Each partner can take an action in a set A_i. Let A denote the space of joint actions, i.e., $A=\otimes_i A_i$ is the Cartesian product of the individual sets of actions. There exists a production function $y\colon A\rightarrow\mathbb{R}$, which can also be considered as a profit function, which maps each vector of joint actions $a\in A$ to an output level. A *sharing rule* is a scheme to allocate any output level in the image set $y(A)$ among the n partners. Shares are possibly nonpositive. I suppose that there is an outside court which can enforce this sharing rule. Because the utility of each partner will be increasing in money, the partners will never agree on a sharing rule which does not satisfy budget balance. Thus, if s is a sharing rule, it must be true that for any y, $\sum_i s_i(y)=y$.

There is moral hazard because the partners cannot observe each other's actions and because they enjoy some disutility from taking an action. Specifically, I will suppose that the utility function of each partner is separable in money and action and that the partners are risk-neutral. Thus, if a is the vector of joint actions and if s is the sharing rule, then the utility of partner i is equal to $s_i(y)-v_i(a_i)$, where $v_i(\cdot)$ is the disutility function of partner i.

Because of risk neutrality, a vector of actions is efficient when the sum of the utilities, which is equal to $y(a)-\sum_i v_i(a_i)$, is maximized. The partnership problem is to design a sharing rule s such that each partner has the incentive to take an efficient action when he believes that the other partners will take an efficient action. I.e., s must generate a game $\Gamma=\left\langle\{A_i\},\{s_i(y(a)-v_i(a_i)\},i=1,\ldots,n\right\rangle$ for which the vector of efficient actions is a Nash equilibrium. If we require full efficiency, and if the functions y and v_i are not trivial, then, the partnership problem has a solution if and only if the vector of efficient actions a^* is a pure strategy Nash equilibrium of the game Γ.

It is well accepted in the literature that the partnership problem *does not* have a solution, i.e., that full efficiency cannot be sustained under the Nash assumption and under some conditions on the output and utility functions. Does this mean that a partnership will not form? Consider the solution proposed in the literature. Efficiency can be attained when property rights are transferred to a "manager" who will pay each worker a certain wage depending on the final output. A typical contract will be: each worker receives $s_i \cdot y^*$ if $y^* = y(a^*)$ is produced and receives 0 otherwise. Here the s_i are not constrained to sum to 1. A typical contract penalizes the partners if the efficient output is not produced. The analysis in Section 3 will make clear why efficiency can be enforced in this framework. The contracting process is not bounded by a budget constraint, i.e., the manager is entitled to receive the residual profit $y - \sum_i s_i(y)$, where $s_i(y)$ is the wage paid to partner i if output y is observed. In this framework, the manager is the owner (or the representative of the owner(s)) of the firm.

What role does a mechanism designed by a manager play? It is essentially a way to tell the workers "take the optimal actions because if any one of you deviates, he will be worse off." Such a mechanism can consequently be implemented by having the workers sign, before taking an action, a contract in which they commit themselves to taking an efficient action and to receiving certain wages depending on the final output. This commitment on the part of the workers is credible because, by accepting to be paid a certain amount in case of deviations, they indeed have no incentives to deviate.

Thus, a mechanism is a way for the workers and the manager to correlate their actions.[3] One can wonder at this point why the partners, keeping their property rights over the final output, cannot design such a mechanism by themselves. We know from Holmström (1982) the answer: efficiency conflicts with budget balancing. By introducing the manager as a residual claimant, it is credible (they have no choice!) for the workers to accept to be paid less than the value of the output that they produce and consequently to attain efficiency.

If the story is that "we must attain efficiency at any cost to us (the potential partners)," then there is no way to escape the conclusion of the literature when partnerships are inefficient. However, if one does not believe that economic agents will abandon their property rights for the sake of efficiency, one cannot accept the solution of the literature. I would like to argue here that there is a solution to the partnership problem which is consistent with the partners' keeping their property rights.

This solution extends the idea of a mechanism in the simple principal-agents framework and corresponds to the concept of correlated strategy equilibrium as formalized in Aumann (1974). Another related reference is Myerson (1983). Correlation of strategies can be attained when the players base their strategies on the realization of a random event, whose distribution may or may not be "objectively" known. For instance, the profile of strategy in which each partner plays his efficient action corresponds to correlated strategies in which each partner plays his efficient action when an event, which is known to happen with probability one, is realized. In particular, a mixed strategy is a special correlated strategy in which there is no correlation; the probability that a partner takes a given action is independent of the probability that another partner takes a given action.[4]

A *group mechanism* defines each partner's share for each possible realization of output, and the probabilities of a random device that the partners will build. A strategy for each partner is then a probability measure over his strategy set conditional on the realization of the random device. This is equivalent to having a random device selecting a vector of actions with a certain probability and telling each partner only the action that he must take, but not the actions which are selected for the other partners. For convenience, I will use the second interpretation.

A way to think of a group mechanism is the following. The partners sign, before taking any action, a contract which is enforceable. The contract specifies, on one hand, that a randomizing machine is built before production begins, and, on the other hand, a sharing rule. Once the machine is built, each partner receives a message telling him which action he should take, but he is not able to observe the messages that the other partners receive. After all the partners have received their messages, production begins, i.e., each partner takes an action and an output results. It is not necessarily true that each partner will take an action which corresponds to his message. An equilibrium mechanism is a correlated device and a sharing rule such that if partner i believes that the other partners will take actions similar to the messages that they have received from the machine, then i has no incentive not to follow the instructions of the machine. Figure 2.2 provides an illustration of the time sequence.

I will show that for any $\epsilon>0$, there exists a correlated device (i.e., a probability measure over the joint actions of the partners) for which no partner has an incentive to deviate from what he is told to do and for which the ex-ante sum of the utilities is ϵ-close to the sum of the utilities corresponding to the efficient vector of actions. This device can be chosen in such a way that the actions of the partners are independent.

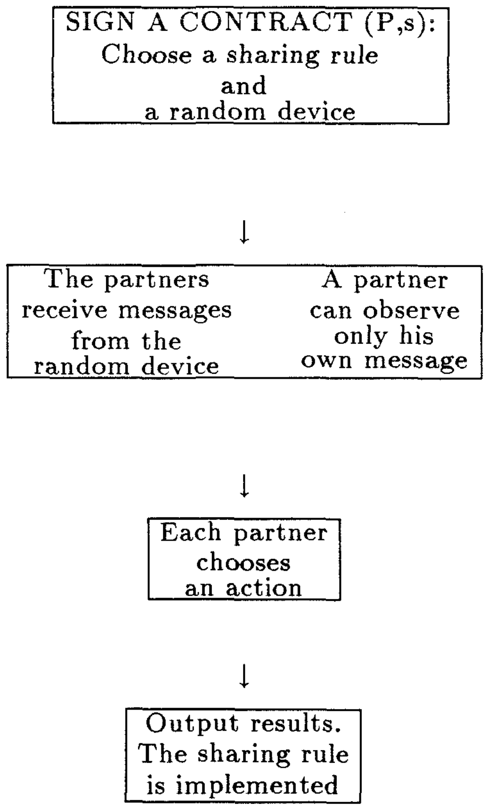
SIGN A CONTRACT (P,s):
Choose a sharing rule
and
a random device
The partners
receive messages
from the
random device
A partner
can observe
only his
own message
Each partner
chooses
an action
Output results.
The sharing rule
is implemented

Figure 2.2

I make the following assumptions.

H1. $\forall i$, $A_i \subset \mathbb{R}$, A_i is complete.[5]

H2. $\forall i$, $\forall a_{-i} \in A_{-i}$, $[a_i > a_i'] \Rightarrow [y(a_{-i}, a_i) > y(a_{-i}, a_i')]$.

H3. y and v_i are continuous.

I make no further assumption on y or on the v_i functions. The set of joint actions A is endowed with the product topology. Below, A will be considered either as a normed linear space or as a metric space. This should create no confusion.

$\mathcal{A}$ will denote the set of Borel subsets of the metric space A, i.e., the σ-field generated by the open sets of A. A probability measure on $\mathcal{A}$ is a nonnegative, countably additive set function P with $P(A)=1$. For each i, I will say that an action a_i is *supported by P* whenever $a_i \in \pi_i(supp(P))$, where $supp(P)$ is the support of P. If a_i is supported by P, I denote the conditional probability on $(A_{-i}, \mathcal{A}_{-i})$ by $P(D|a_i)$. I will write $a_i \in supp_i(P)$ to indicate that a_i is supported by P.[6]

Because of the budget balancing condition, for any output y, it is enough to define the shares of $n-1$ partners. The share of the *n-th* partner is then the residual $s_n(y) \equiv y - \sum_{i \neq n} s_i(y)$. If X is a normed linear space, the dual of X is the space of all continuous linear functionals defined on X. Let $Y \equiv y(A)$ be the range of the output function. It is convenient to think of a sharing rule as an element of the dual of the space $\left(\mathbb{R}^{n-1}\right)^Y$, i.e., the dual of the space of all bounded functions from Y to $\mathbb{R}^{n-1}$. For instance, if Y is finite, the dual coincides with the predual and a sharing rule is a vectorial function from Y to $\mathbb{R}^{n-1}$.

DEFINITION 1: *An ϵ-equilibrium group mechanism (ϵ-EGM) is a pair (P,s) where P is a probability measure on $(A,\mathcal{A})$, s is an element of the dual of $\left(\mathbb{R}^{n-1}\right)^Y$, which satisfy the following conditions,*

*E*1. $\forall i$,

$$[a_i \in supp_i(P)] \Rightarrow [a_i \in \underset{\hat{a}_i}{argmax} \int \left(s_i(y(a\backslash \hat{a}_i)) - v_i(\hat{a}_i)\right) P(da_{-i}|a_i)].$$

*E*2. $$\int \left(y(a) - \sum_{i=1}^{n} v_i(a_i)\right) P(da) = y(a^*) - \sum_{i=1}^{n} v_i(a_i^*) - \epsilon.$$

E1 is the equilibrium condition. Once partner i receives the message a_i, i.e., if he is told to do action a_i, he will believe that there is a probability $P(D|a_i)$ that the other partners have received messages a_{-i} in the set D. Under the Nash assumption, i will choose an action which maximizes his expected utility, given that the other agents do not deviate from what they are told to do. E2 is the efficiency requirement: the ex-ante sum of the utilities must be ϵ close to the sum of the utilities which is attained at the efficient vector of actions.

It is clear that once the sharing rule s is known, the utility functions are well defined. Consequently, (P,s) satisfies E1 if and only if P is a correlated equilibrium of the game whose payoff function is generated by s. If the set of actions is finite, for any s, the set of correlated equilibria is nonempty since it includes the convex hull of the set of Nash equilibria, which is nonempty by the Nash Theorem. The main result of the chapter is that for any ϵ, there exists a random device P for which the actions taken by the players are independent events and for which there exists a sharing rule s such that (P,s) is an ϵ-EGM.[7]

THEOREM 1: *For any partnership problem which satisfies H1-H3, there exists an ϵ-equilibrium group mechanism (P,s). It is always* *possible to choose P in such a way that (i) the actions taken by one partner are independent of the actions taken by the other agents, (ii) each partner takes in equilibrium only countably many actions with positive probability, and (iii) the probability that an agent takes his efficient action is greater than* $1-\delta(\epsilon)$, *where* $\delta(\epsilon)$ *is increasing in* ϵ.

The proof of this theorem is given in Section 5.1. Independence means that the partners follow mixed strategies in the game induced by the sharing rule s.

Remark 1: Theorem 1 begs the question of how and why the partners will ever agree on an ϵ-equilibrium mechanism since by choosing a smaller ϵ there is always a possibility of gain. Disagreement delays production and revenues for the partners. If one acknowledges the fact that communication between individuals uses media which are time consuming (the voice, handwriting), it can be shown that if the partners have a non-zero discount rate, they would rather settle on a given ϵ rather than continue to disagree forever. A formal proof of this claim necessitates the definition of the extensive form game that the partners play for choosing the group mechanism. Appendix C provides such an extensive form in the framework of the examples of Section 4. If the partners do not discount their future revenues they will disagree forever, each announcing a smaller ϵ than the other partner.

Remark 2: It is well known that the set of correlated equilibria of a game is compact under certain assumptions on the strategy spaces. There is no reason to expect that the set of *all* ϵ-EGM sets is compact (the problem is with closedness; see the first example in Section 4 in which a sequence of ϵ-EGM converges to a pair (P,s) which is not an $\epsilon-$EGM). However, one can prove weaker results. Compactness results are important when we consider the problem of the selection of a group mechanism. For instance, if the partners agree ex-ante to choose a group mechanism which gives them their Nash (cooperative) levels of utility, compactness is a sufficient condition for the existence of such a selection (e.g., see Appendix C). Two results are established below. First, if the partners agree on a level of efficiency loss ϵ, then the set of ϵ-EGM is (weak*) compact. Second, if they agree on a sharing rule s, the set of probability measures P for which (P,s) is an EGM is compact.

Denote by $\Delta(A)$ the set of regular probability measures on $(A,\mathcal{A})$. $\Delta(A)$ is a subset of the unit sphere. By the Banach-Alaoglu Theorem, $\Delta(A)$ is (weak*) compact since it is closed. Let Y be the set of outputs which can be generated by actions in A. A sharing rule is an element of the dual of $L=\left(\mathbb{R}^{n-1}\right)^Y$ (see Section 5.1 for more details). The first result states that if the partners decide to sacrifice a given amount of efficiency, i.e., they fix an ϵ, then the set of ϵ-EGM from which they can choose is a compact set.

COROLLARY 1: *For each $\epsilon>0$, for each $\rho>0$, $\rho<\infty$, if the set of ϵ-EGM (P,s) for which $\|s\|\leq\rho$ is nonempty, then it is compact in the product topology on $\Delta(A)\otimes L^*$.*

Proof: By Alaoglu's Theorem, the set $K=\{s\in L^*;\ \|s\|\leq\rho\}$ is (weak*) compact. Consider a sequence $\{(P^m,s^m)\}$ of ϵ-EGM, ϵ being given, for which $s^m\rightarrow s$ and $P^m\rightarrow P$. By compactness, $P\in\Delta(A)$ and $s\in K$. (P,s) is also an ϵ-EGM since by taking limits one does not change the direction of the inequalities defining the equilibrium conditions (see also Section 5.1). □

The second result states that to each sharing rule corresponds a compact set of probability measures on A which are compatible with ϵ-efficiency, ϵ being arbitrary.

COROLLARY 2: *The set of probability measures P' for which (P',s) is an ϵ-EGM is compact in the product topology on $\Delta(A) \otimes L^*$.*

Proof: For s given, the utility functions of the players are well defined. It is clear that the set of correlated equilibria of the resulting game is closed in $\Delta(A)$. Compactness of the set of ϵ-EGM follows the compactness of $\Delta(A)$. □

Remark 3: Theorem 1 holds for any *positive* ϵ. This triggers the question of the existence of group mechanisms when ϵ *is equal* to zero. The literature argues that in this case, there does not exist 0-equilibrium group mechanisms since full efficiency cannot be sustained. We will see in the next section that this negative conclusion must be qualified: sometimes full efficiency can be sustained. Nevertheless, there are many situations in which nonsustainability is the rule. There is consequently a discontinuity in the equilibrium properties of a partnership at $\epsilon=0$. The reason is the following. At full efficiency, it is *necessary* to consider only pure strategy equilibria. Indeed, if a partner puts a nonzero probability on an inefficient action, then, by definition, an inefficient level of output has a positive probability of occurring and this contradicts full efficiency. The literature tells us that, in general, there is no sharing rule which generates a game $\Gamma=\left\langle\{A_i\},\{s_i(y(a))-v_i(a_i)\},i=1,\ldots,n\right\rangle$ where the vector of efficient actions is a pure strategy Nash equilibrium. Theorem 1 establishes that if we "sacrifice" some ϵ, then there exists a sharing rule for which the game Γ has a mixed strategy equilibrium and for which the probability that a partner takes an efficient action is close to one. The discontinuity at $\epsilon=0$ comes from the fact that at $\epsilon\neq0$, we are not restricted anymore to consider only pure strategies.

3. EFFICIENCY AND PURE STRATEGY NASH EQUILIBRIA

3.1. *WHY DO THINGS GO WRONG WHEN ϵ GOES TO ZERO?*

Consider the environment of Section 2. I make the following assumptions.[8] The sets of actions are identical, $\forall i=1,\ldots,n,\ A_i=[0,\infty)$. The function y is strictly concave, the functions v_i are strictly convex and y and v_i are strictly increasing in their respective arguments and are of class C^2. This environment applies only to Section 3.1.

Let $\Sigma(y)\equiv\{x\in\mathbb{R}^n \mid \sum_i x_i=y\}$. The partners must agree, before taking an action, on a sharing rule s, where $\forall y,\ s(y)\in\Sigma(y)$. Let a^* be the (unique) efficient vector of actions, where $a^*\neq 0$ and a^* is interior. The partnership problem is to find a sharing rule for which each partner is compelled to take action a_i^* when he believes that the other partners will take action a_j^*, $j\neq i$. Suppose that there exists such a sharing rule s. Let $\tilde{a}$ be a vector of actions such that $\forall i,\ y(a^*\backslash\tilde{a}_i)\equiv\tilde{y}$, where $\tilde{y}$ is a given output level, and where $a^*\backslash\tilde{a}_i\equiv\ (a_1^*,\ldots,a_{i-1}^*,\tilde{a}_i,a_{i+1}^*,\ldots,a_n^*)$. Thus if partner i takes an action $\tilde{a}_i$, while the other partners take their efficient actions, a certain output level $\tilde{y}=y(a^*\backslash\tilde{a}_i)$ will be produced. Thus, the $\tilde{a}_i$'s are *observationally equivalent* unilateral deviations from the optimal vector of actions a^*. If s is a solution to the partnership problem, then

$$\forall i,\ s_i(y(a^*))-v_i(a_i^*)\geq s_i(\tilde{y})-v_i(\tilde{a}_i).$$

Summing the above inequalities over i and using the fact that $s(y)\in\Sigma(y)$,

$$y(a^*)-\sum_i v_i(a_i^*)\geq\tilde{y}-\sum_i v_i(\tilde{a}_i). \tag{3.1}$$

It is now easy to remark that s is a solution if and only if (3.1) holds when the right hand side is maximized. Since $\tilde{a}=a^*$ is a possible case, the right hand side in (3.1) is always at least as great as the left hand side. Thus, we must have an equality, i.e., $\tilde{a}=a^*$, since the right hand side has a unique maximum. Now, to find this maximum, one solves the program,

$$(3.2) \qquad \underset{y,\tilde{a}}{\mathrm{MAX}} \; [y - \textstyle\sum_i v_i(\tilde{a}_i)]$$

$$\text{s.t. } \forall i,\; y(a^*\backslash \tilde{a}_i) = y.$$

Let λ_i be the Lagrangian coefficient associated with the constraint of partner i. By concavity, the optimal $\tilde{a}$ must solve the equations

$$(3.3) \qquad \textstyle\sum_i \lambda_i = 1,\; \lambda_i \geq 0$$

$$\forall i,\; \lambda_i \cdot \frac{\partial y(a^*\backslash \tilde{a}_i)}{\partial a_i} = v_i'(\tilde{a}_i)$$

If we suppose that $\forall i,\; v_i'(0) > 0$, it follows that $\lambda_i \neq 0$. Then, from the first equality,

$$(3.4) \qquad \forall i,\; \lambda_i < 1.$$

The efficient vector of actions satisfies the conditions

$$(3.5) \qquad \forall i,\; \frac{\partial y(a^*)}{\partial a_i} = v_i'(a_i^*).$$

Using (3.4) and the assumptions on y and v_i, it follows, by comparing (3.3) and (3.5), that $\forall i,\; \tilde{a}_i < a_i^*$. Thus, there does not exist a solution to the partnership problem.

This proof makes clear that in order to check the existence of a solution for the partnership problem, it is enough to check that one inequality is satisfied (replace the right hand side of (3.1) by the value for the solution of the program (3.2)). This suggests that the partnership problem might have a solution. In particular, if there always exists a partner who cannot "mimic" the other partners, then the partnership will have a solution. (Because in this case, there is no vector $\tilde{a}$ for which $y(a^*\backslash \tilde{a}_i)$ is a constant for all i.)

3.2. *MIMIC AND PERISH*

3.2.1. *The finite case*

There are n partners indexed by $i=1,\ldots,n$. The action space of partner i is the set A_i with generic element a_i. This set is finite and is written $A_i=\{a_i(1),\ldots,a_i(T_i)\}$. With an abuse of notation, I will denote $a_i<a_i'$ whenever $a_i=a_i(j)$, $a_i'=a_i(j')$, and $j<j'$. Note that, contrary to Sections 2 and 3.1, there is no restriction on the underlying set in which the actions lie.[9] Moreover, it is possible that the underlying spaces are different for all agents, e.g., $\forall i$, $A_i\subset\mathbb{R}^{k_i}$, where the k_i are all different. For the purpose of this section, I will suppose that the sets of actions have the same cardinality, i.e., that for each partner i, $T_i=T$. This is only for convenience; the arguments below follow if two partners have sets of actions with a different cardinality. For a given vector of actions $a\in A$, the corresponding (unique) output is $y(a)$. Clearly, the set Y of possible outputs is finite and has a cardinality which is bounded above by T^n. I denote $Y=\{y_1,\ldots,y_l\}$, where $t>t'\Rightarrow y_t>y_{t'}$. By finiteness of the set A, an efficient vector of actions, a^* exists. Below, a_{-i} denotes the vector $(a_1,\ldots a_{i-1},a_{i+1},\ldots,a_n)$. With an abuse of notation, I will sometimes denote $a^*\backslash a_i=(a^*_{-i},a_i)$. I make the following assumptions.

H4. $a_i>a_i'\Rightarrow y(a^*\backslash a_i)>y(a^*\backslash a_i')$.

H5. $\forall i$, $\exists j$, such that $v_i(a_i(j))<\infty$.

H6. $y^*-\sum_i v_i(a_i^*)<\infty$.

H4 is only for convenience.[10] (Note that H4 does not mean that y is increasing.) If H5 does not hold, then the partnership is trivial since any vector of actions is efficient. If H6 does not hold, the partnership problem is also trivial since any sharing rule is a solution. (Moreover, in this case, since the efficient vector of actions is well defined, it is not possible that $y^*=\infty$ and that $\exists i$, such that $v_i(a_i^*)=\infty$. Thus, if H6 is violated, it is possible for the partners to attain an infinite output with any of them enjoying only a finite disutility.)

A *sharing rule* defines for each level of output y a vector $s(y) \in \mathbb{R}^n$ such that $\sum_i s_i(y) = y$, i.e., $s \in \Sigma(y)$. Since the set of outputs is finite, it is enough to define l sharing rules (the sharing rule can depend on the output only since the actions are not observable).

From now on, I will use either $y(a^*)$ or y^* or y_t to denote the efficient level of output. From H4, by deviating from a_i^*, partner i can generate exactly $T-1$ different outputs. If partner i anticipates that all the other partners will play action a_j^*, where $j \neq i$, and if partner i deviates to actions $a_i(k) \neq a_i^*$, then $\tau(i,k)$ denotes the index of the level of output that results from the joint actions $(a_{-i}^*, a_i(k))$, i.e., $y((a_{-i}^*, a_i(k))) = y_{\tau(i,k)}$. From H4, $k \neq k' \Rightarrow \tau(i,k) \neq \tau(i,k')$.

The partnership problem is to find sharing rules $s(\cdot)$ such that no partner wants to deviate when he believes that the other partners will use their efficient actions. If this is the case, a^* is a Nash equilibrium of the game with strategy space A_i for each partner and with associated payoff functions of $s_i(y_h) - v_i(a_i(k))$ when partner i takes action $a_i(k)$ and the other partners take actions which result in output y_h. a^* will be an equilibrium if and only if the following system has a solution ($s(y_h)$, where $h=1,\ldots,l$, and $s(y_h) \in \Sigma(y_h)$)

(3.6) $\qquad \forall i,\ \forall k,\ a_i(k) \neq a_i^*,\ s_i(y^*) - v_i(a_i^*) \geq s_i(y_{\tau(i,k)}) - v_i(a_i(k)).$

(3.7) $\qquad \forall r,\ \sum_i s_i(y_r) = y_r.$

Inequalities in (3.6) are the equilibrium constraints, (3.7) are the budget balancing conditions. Clearly, (3.6) and (3.7) are equivalent to

$$(3.8)\qquad \forall i\neq n,\ \forall k\neq a_i^*,\ s_i(y^*)-v_i(a_i^*)\geq s_i(y_{\tau(i,k)})-v_i(a_i(k)).$$

$$\forall k\neq a_n^*,\ y^*-\sum_{i\neq n} s_i(y^*)-v_n(a_n^*)\ \geq\ y_{\tau(n,k)}-\sum_{i\neq n} s_i(y_{\tau(n,k)})-v_n(a_n(k)).$$

Note that (3.8) can be written in the form $B\cdot s=\beta$, where

s is the $l\times(n-1)$ vector

$$s=(s_1(y_1),\ldots,s_{n-1}(y_1),s_1(y_2),\ldots,s_{n-1}(y_2),\ldots,s_1(y_l),\ldots,s_{n-1}(y_l))$$

β is the $n\cdot(T-1)$ vector defined by

$\beta=(\beta 1,\ldots,\beta n)$ where each βi consists of $T-1$ values such that

$$\forall i<n,\ \beta i(k)=\begin{cases} v_i(a_i^*)-v_i(a_i(k)) & \text{if } a_i(k)<a_i^* \\ v_i(a_i^*)-v_i(a_i(k+1)) & \text{if } a_i(k)>a_i^*. \end{cases}$$

$$\beta n(k)=\begin{cases} y_{\tau(n,k)}-y^*+v_n(a_n^*)-v_n(a_n(k)) & \text{if } a_n(k)<a_n^* \\ y_{\tau(n,k+1)}-y^*+v_n(a_n^*)-v_n(a_n(k+1)) & \text{if } a_n(k)>a_n^*. \end{cases}$$

Note that, by efficiency, $\beta n(k)\leq 0$, for all k. B is a $(n\times(T-1),l\times(n-1))$ matrix consisting of n vertical $(T-1,l\times(n-1))$ blocks $B(i)$. $B(i)(r,c)$ denotes the element in the r-th row and c-th column of the matrix $B(i)$, $\forall i<n$,

$$(3.9)\qquad B(i)(r,c)=\begin{cases} 1 & \text{if } c=(t-1)\cdot(n-1)+i \\ -1 & \text{if } c=(\tau(i,r)-1)\cdot(n-1)+i \\ 0 & \text{otherwise.} \end{cases}$$

$$B(n)(r,c)=\begin{cases} -1 & \text{if } (t-1)\cdot(n-1)+1\leq c\leq t\cdot(n-1) \\ 1 & \text{if } (\tau(n,r)-1)\cdot(n-1)+1\leq c\leq \tau(n,r)\cdot(n-1) \\ 0 & \text{otherwise.} \end{cases}$$

It is useful to compute the rank of the matrix B. Let $\mathfrak{F}\equiv\{f\colon\{1,\ldots,n-1\}\rightarrow\{1,\ldots,T\}\}$ and $\forall r$, $\mathfrak{F}_0(r)\equiv\{f\in\mathfrak{F}\mid\forall i\in\{1,\ldots,n-1\},\ \tau(i,f(i))=\tau(n,r)\}$ be the subset of $\mathfrak{F}$ such that each partner i different from n deviates to an action $f(i)$

for which the resulting output index coincides with the index of the output obtained when partner n deviates to $a_n(r)$ (if $a_n(r)<a_n^*$) or $a_n(r+1)$ (if $a_n(r)>a_n^*$). Finally, let χ be the indicator function, $\chi(r)=1$ if $\mathfrak{F}_0(r)\neq\emptyset$ and $\chi(r)=0$ otherwise. Lemma 1 characterizes the rank of B in terms of χ. Theorem 2 gives the necessary and sufficient conditions for having a solution to partnership problems with finite sets of actions. The proof of Theorem 2 relies on a separating hyperplane argument.

LEMMA 1: $rank(B)=n\cdot(T-1)-\sum_{r=1}^{T-1}\chi(r)$.

THEOREM 2: (3.8) *has a solution if and only if either one of the two following conditions hold*

(i) $\forall r,\ \chi(r)=0$.

(ii) $[\exists r,\ \ \chi(r)=1]\Rightarrow[y^*-\sum_i v_i(a_i^*)\geq y_{\tau(n,r)}-\sum_i v_i(a_i(f(i)))$, *where* $\forall i,\ \tau(i,f(i))=\tau(n,r)]$.

When either (i) *or* (ii) *hold, the set of solutions is a linear manifold of dimension* $l\cdot(n-1)-rank(B)$.

Let me interpret these necessary and sufficient conditions. I interpret (i) as saying that all partners cannot "mimic" partner n. Thus, when partner n deviates unilaterally from his efficient action, there is at least one partner who cannot, by deviating *unilaterally from his efficient action*, generate the same output as n was able to generate. Note that in this case, by observing the result of a deviation, the partners know who did not unilaterally deviate; thus deviations are always informative (i.e., restrict the set of the possible deviants). *(i) is not satisfied if the partnership is* symmetric, i.e., if the output function is symmetric and the disutility functions are symmetric. (ii) is equivalent to condition (3.1). As noted at the end of Section 3.1, there is a simple reason why (ii) is not vacuous: it is always possible to eliminate the actions which are not consistent with (ii) and to have a new partnership problem with smaller sets of actions that has a solution.

I rewrite the inequality in (ii) as $y_{\tau(n,r)} - y^* \leq \sum_i \Big(v_i(a_i(f(i))) - v_i(a_i^*)\Big)$. This tells us that the variation in output when one partner deviates must be smaller than the sum of the variations in disutilities for all the partners, when the partners use observationally equivalent deviations. Suppose that the reverse (strict) inequality holds. Then, for any sharing rule, the partners are "collectively" better off; there is more output to share among them than they lose in terms of disutilities. With nonobservable actions, such a "collective" improvement implies that, for any sharing rule, at least one partner will have an incentive to deviate.

Remark 4 : A partnership problem is defined by the cardinality of the set of actions, the function y and the functions v_i. Let P be the class of partnerships with n players and T actions. A partnership is consequently a point in the space $P \equiv \mathbb{R}^{T^n} \times \mathbb{R}^{n \cdot T}$. I will denote $\Gamma \equiv \langle y(a), v_i(a_i); a \in A \rangle$. Define the product topology in this space. Let $P^1 \subset P$ be the subclass of partnership problems which can sustain full efficiency. I claim that P^1 is dense in P. To prove this claim, it is enough to show that the closure of P^1 coincides with P. I will in fact prove a stronger result.

Let $P^0 \subset P^1$ be the set of partnership problems for which condition (i) of Theorem 2 holds. That is to say, for any deviation from efficiency of partner n, it is possible to identify a partner $i \neq n$ who could not have possibly deviated. Let $\Gamma \in P \setminus P^1$, i.e., suppose that Γ does not have a solution. Then, (i) and (ii) of Theorem 2 do not hold. Let $R \equiv \{r_1, \ldots, r_m\}$ be the set of indices for which $\chi(r) = 1 \Leftrightarrow r \in R$. Since (i) does not hold, $R \neq \emptyset$.

I show that there exists a sequence of partnerships problems $\{\Gamma^m\} \subset P^0$ with the property that $\Gamma^m \rightarrow \Gamma$. Let y and v_i be the output and disutility functions in Γ. Consider the following quantities,

$$\delta_i^1 = \min_{\substack{a_i, a_i' \\ a_i \neq a_i'}} \left| y(a^* \backslash a_i) - y(a^* \backslash a_i') \right|$$

$$\delta^1 = \min_i \delta_i^1$$

$$\delta_i^2 = \min_{r \notin R} \min_{\substack{a_i: \\ y(a^* \backslash a_i) \neq y_{\tau(n,r)}}} \left| y_{\tau(n,r)} - y(a^* \backslash a_i) \right|$$

$$\delta^2 = \min_i \delta_i^2.$$

By assumption, $\delta^1 > 0$. $\delta^2 > 0$ only if $\#R < T$. Let $\delta = \min(\delta^1, \delta^2)$ if $\delta^2 > 0$, let $\delta = \delta^1$ if $\delta^2 = 0$. Choose $0 < \epsilon < \delta$, and let $\epsilon^m = \epsilon/m$. Define the following pertubed functions

$$y^m(a) = \begin{cases} y(a) + \epsilon^m & \text{if } a = (a^* \backslash a_n) \\ y(a) & \text{otherwise.} \end{cases}$$

$$v_i^m(a_i) = \begin{cases} v_i(a_i) + \epsilon^m & \text{if } i = n \\ v_i(a_i) & \text{otherwise.} \end{cases}$$

Observe that the function y^m is well defined since there is only one vector of actions of the form $a^* \backslash a_n$. If a is not of the form $a^* \backslash a_n$, then $y^m(a) - \sum_i v_i^m(a_i) = y(a) - \sum_i v_i(a_i) - \epsilon^m$ since the term in ϵ^m appears only in $v_n^m(a_n)$. If a is of the form $a^* \backslash a_n$, then, $y^m(a) - \sum_i v_i^m(a_i) = y(a) - \sum_i v_i(a_i)$ since the term ϵ^m disappears. It follows that a^* is still efficient in the game Γ^m. The difference $y^m_{\tau(n,r)} - y^m(a^* \backslash a_i)$ is equal to $y_{\tau(n,r)} - y(a^* \backslash a_i) - \epsilon^m$. This expression is equal to zero only if $y_{\tau(n,r)}-y(a^*\setminus a_i)=\epsilon^m$. Since $\epsilon^m>0$, it is necessary that $y_{\tau(n,r)}-y(a^*\setminus a_i)\neq 0$. The choice of δ insures that whenever $y_{\tau(n,r)}-y(a^*\setminus a_i)\neq 0$, this difference is greater than δ. Since $\delta>\epsilon^m$, it is not possible to have $y_{\tau(n,r)}-y(a^*\setminus a_i)=\epsilon^m$. It follows that in each game Γ^m, condition (i) of Theorem 2 is satisfied, i.e., $\forall m$, $\Gamma^m\in P^0$. Clearly, $\Gamma^m\rightarrow\Gamma$. •

Remark 5: It is useful to rewrite Theorem 2 when the cardinalities of the sets of actions of the partners are different. The reader can easily verify that the following is true.

THEOREM 2′: *Suppose that* $\#A_i=T_i$, *where* $i=1,\ldots,n$. *Then,*

$$rank(B)=\sum_{i=1}^{n} T_i-n-\sum_{r=1}^{T_n}\chi(r).$$

(3.8) has a solution if and only if either one of the two following conditions hold

$(i)'$ $\forall r$, $\chi(r)=0$.

$(ii)'$ $[\exists r,\ \chi(r)=1]\Rightarrow[y^*-\sum_i v_i(a_i^*)\geq y_{\tau(n,r)}-\sum_i v_i(a_i(f(i)))$, *where* $\forall i$, $\tau(i,f(i))=\tau(n,r)]$.

When either $(i)'$ *or* $(ii)'$ *hold, the set of solutions is a linear manifold of dimension* $l\cdot(n-1)-rank(B)$.

3.2.2. *General sets of actions*

I consider general sets of actions. A_i is an abstract set of actions for partner i. (For instance, A_i can be taken to be a subset of the Euclidian space E^{k_i}.) The production function y: $A\rightarrow\mathbb{R}$ induces a set of attainable outputs Y (observe that Y does not have obligatorily a lowest bound here). I will suppose that there exists an efficient vector of actions. Assumptions H4-H6 hold.[11] The notation of Section 3.2.1 can be extended in a straightforward way. In particular the blocks $B(i)$ in the matrix B have an infinite (possibly nondenumerable) number of rows and of columns; the elements $B(i)(r,c)$ continue to satisfy (3.9). The partnership problem is to find an infinite dimensional vector s such that the system $B \cdot s \geq \beta$ has a solution. Theorem 3 below is the extension of Theorem 2 and states that there exists a solution with finite norm to the partnership problem if conditions (i) and (ii) of Theorem 2 hold for any finite subsystem.

Let $L=\left(\mathbb{R}^{n-1}\right)^{Y}$ be the space of all real bounded functions $x(y)=(x_1(y),\ldots,x_{n-1}(y))$ on the set Y. For instance, if each set of actions is countable, Y itself is countable and $L=\mathbb{R}^{\infty}$. I will consider L as a real normed linear space (L is always normable, e.g., the sup norm defines L as a Banach space). With this notation, each row of the matrix B is a vector in L. Note that by construction, a row corresponding to a deviation of partner $i<n$ has exactly two nonzero components, while the rows corresponding to deviations of partner n each have exactly $2\cdot(n-1)$ nonzero components. A solution to the partnership problem is an element s of the dual L^* such that the inequalities $s(B_\alpha) \geq \beta_\alpha$ hold (where $s(B_\alpha)$ denotes the image by s of the α-*th* row of B and β_α denotes the α-th element of β).

A *finite subsystem* of $B \cdot s \geq \beta$ is a system $\tilde{B} \cdot s \geq \tilde{\beta}$, where $\tilde{B}$ and $\tilde{\beta}$ have a finite number of rows and where $\forall \alpha, \rho, \tilde{B}_\alpha = B_\rho \Leftrightarrow \tilde{\beta}_\alpha = \beta_\rho$. Each row of $\tilde{B}$ corresponds to a deviation of a given player. For each partner i, let $\tilde{A}_i$ be the union of $\{a_i^*\}$ and of the set of actions to which i deviates in the submatrix $\tilde{B}$. Let $\tilde{Y}$ be the image by y of $\tilde{A}$ (where $\tilde{A}$ is the Cartesian product of the $\tilde{A}_i$). By assumption, $\tilde{Y}$ has finite cardinality. For $x \in L$, let $\pi(x)$ denote the projection of x on $\tilde{L}=\left(\mathbb{R}^{n-1}\right)^{\tilde{Y}}$. It is immediate that a component of $\tilde{B}_\alpha$ is nonzero if and only if this component is in

$\pi(\tilde{B}_\alpha)$. By finiteness of the dimension, the dual of $\tilde{L}$ coincides with $\tilde{L}$. Consequently, it is possible to associate to the subsystem $\tilde{B}\cdot s \geq \tilde{\beta}$ its finite dimensional version $\hat{B}\cdot\hat{s} \geq \hat{\beta}$, where $\hat{B} \equiv \pi(\tilde{B})$, $\hat{\beta} \equiv \pi(\tilde{\beta})$ and where $\hat{s} \in \tilde{L}$. Doing so, a finite subsystem "looks like" a partnership problem in which each partner has a set of actions with finite cardinality. I will say that the resulting partnership problem is *generated* by the finite subsystem $\tilde{B}\cdot s \geq \tilde{\beta}$. This observation enables me to generalize the result of the preceding section in a straightforward way.

> THEOREM 3: *Consider the real normed linear space L, where the norm of s is denoted by* $\|s\|$. *Then, there exist* $0 \leq \rho < \infty$, $s \in L^*$ *with* $\|s\| \leq \rho$ *such that the system* $B\cdot s \geq \beta$ *admits s as a solution if and only if conditions* $(i)'$-$(ii)'$ *of Theorem* $2'$ *are satisfied for any partnership problem generated by a finite subsystem* $\tilde{B}\cdot s \geq \tilde{\beta}$.

The "only if" part of this theorem is obvious. The "if" part uses the facts that the unit sphere is weak*-compact, and that compact sets have the finite intersection property.

Remark 6: The interpretation of the conditions of Theorem 3 is identical to the interpretation given in the previous section. Note that the previous condition $y_{T(n,r)} - y^* \leq \sum_i \left(v_i(a_i(f(i))) - v_i(a_i^*) \right)$ implies, when the functions y and v_i are differentiable, that the marginal product is less than the sum of the marginal disutilities when the functions are evaluated at observationally equivalent deviations. We saw in the example of Section 1.2 that it is in general necessary for at least one of the sets of actions to be *unconnected* in order to have a solution. From the previous section, unconnectedness is not a sufficient condition. It is now clear why there exist efficient mechanisms when a manager is introduced. If the manager's actions do not influence in a "continuous" way the output, or if these actions are publicly observable (for instance, a reorganization of the production process is observable and could be part of a contract), then condition (i)′ of Theorem 2′ will always be satisfied. The next section illustrates these points in a two person framework.

3.2.3. *AN ILLUSTRATION IN THE TWO PERSON CASE*

With only two partners, the set of attainable outputs can be represented in $\mathbb{R}^2$ as a map of isoquants. I will suppose that the assumptions on y and v_i of Section 3.1 hold. The partnership problem in this case is illustrated in Figure 2.3. S is the locus of observationally equivalent deviations,

$$S \equiv \left\{ a;\ \forall i=1,\ldots,n,\ \forall j=1,\ldots,n,\ y(a^*\backslash a_i)=y(a^*\backslash a_j) \right\}.$$

$\underline{a}$ is the point of S such that, (these points are well defined by concavity)

$$\underline{a} \equiv min\left\{ a \in S \,\middle|\, y(a^*\backslash a_i)-\sum_{j=1}^{2} v_i(a_i)=y(a^*)-\sum_{j=1}^{2} v_i(a_i^*) \right\}.$$

The conditions of Theorem 3 are equivalent to saying that the curve S does not "enter" the interior of the box delineated by the four points a^*, $\underline{a}$, $(a_1^*,\underline{a}_2)$, $(\underline{a}_2,a_1^*)$ (the dotted area in Figure 2.3). Clearly, if $A_i=[0,\infty)$, this will not be the case. Let $A_1=[0,\alpha_1)\cup(\alpha_2,\alpha_3]\cup[a_4,\infty)$ and let $A_2=[0,\beta_2]\cup(\beta_3,\infty)$ (see Figure 2.3). In this case, the curve S does not enter the dotted area and a solution exists. This example emphasizes that one needs disconnectedness of the set of actions for at least one player in order to fulfill the conditions of Theorem 3.

By introducing a manager with an action space $A_3=\{0,1\}$, where 0 means "do not produce" and 1 means "authorize production", i.e., the manager's actions do not have any productive role once production is begun, then efficiency is attained at the vector $(a_1^*, a_2^*, 1)$ and condition (i) of Theorem 3 is trivially satisfied. Geometrically, this amounts to saying that the interior of the box delineated by the four points $(a_1^*, a_2^*, 1)$, $(\underline{a}_1, \underline{a}_2, 1)$, $(a_1^*, \underline{a}_2, 1)$, $(\underline{a}_2, a_1^*, 1)$ is empty in the $\mathbb{R}^3$ space (see Figure 2.4). The fact that the relative interior of this box (the dotted area in Figure 2.4) is not empty exemplifies the role of budget breaker of the manager. These remarks extend naturally to n partners.

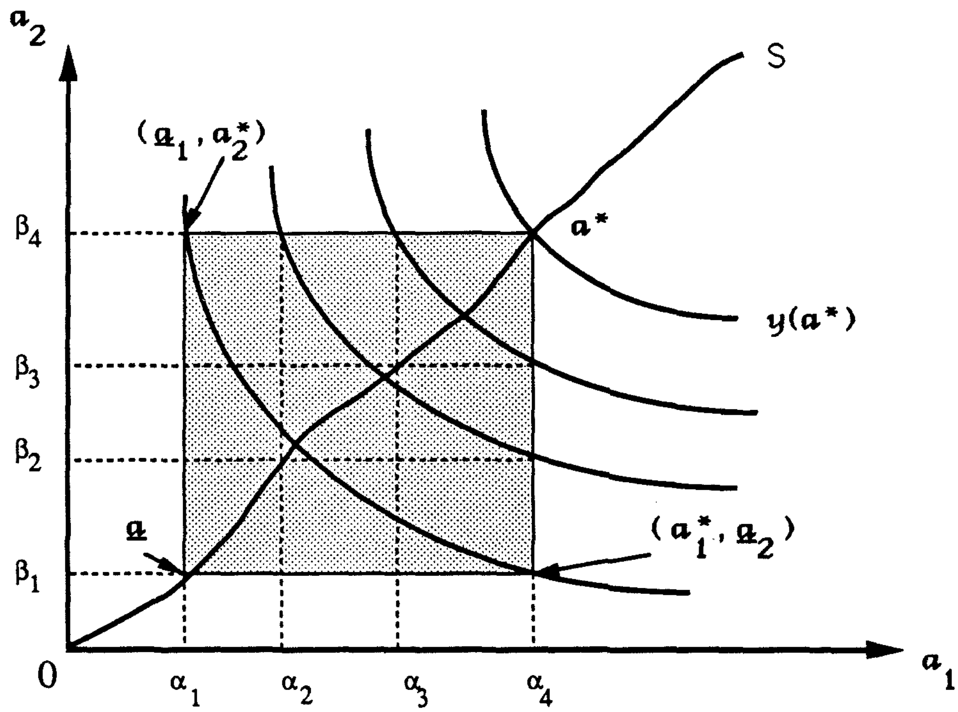

Figure 2.3

A two person example (without a budget breaker)

[S=set of observationally equivalent deviations]

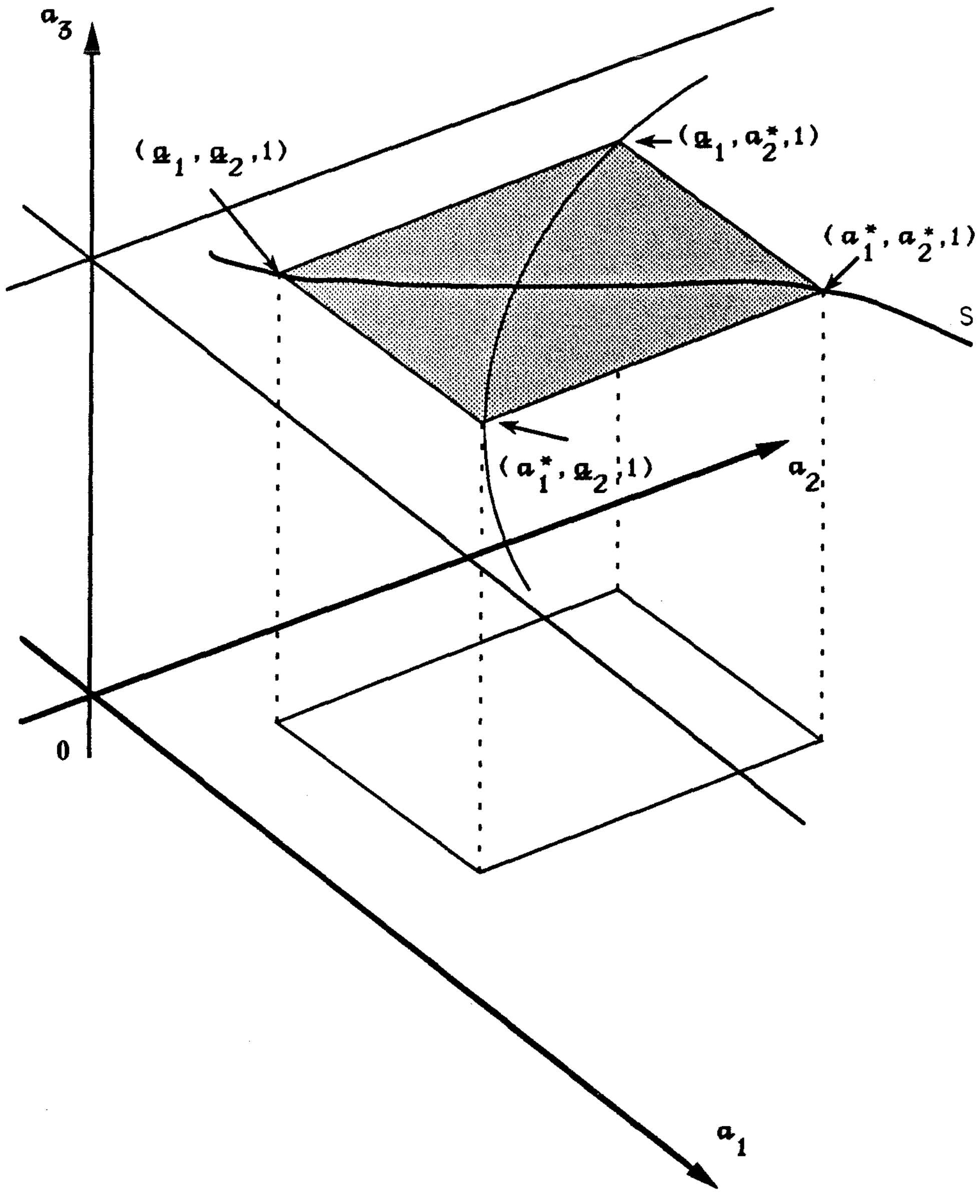

Figure 2.4

A two person example (with a budget breaker)

[S=set of observationally equivalent deviations]

4. LIMITED LIABILITY: TWO EXAMPLES[12]

Theorem 1 tells us that we can always find a sharing rule for which the partnership is sustainable. Because I made no restriction on the sharing rule, it is possible that the share of one partner is negative. If the wealth of the partners is finite, or if the partners cannot borrow money from a financial institution, it is necessary to address the problem of limited liability of the partners. Limited liability means that a partner cannot be asked to pay, after production ends, an amount which is superior to his initial endowment, or to the maximum that he could borrow to pay his debts.[13] This is equivalent to saying that a partner cannot go bankrupt. (Note that this is a weaker requirement than individual rationality.) It is still an open question whether ϵ-EGM which satisfy a no-bankruptcy condition exist in general. However, the following examples should shed some light on this problem.

In the first example, the *only* ϵ-EGM are in mixed strategy and imply, as ϵ goes to zero, that one partner will have to pay a very large amount. I will argue that if the partners can sign a contract which does not satisfy the no-bankruptcy condition, then it is possible for them to sustain ϵ-efficiency with *none* of them going bankrupt. The second example is an example in which no partner goes bankrupt *in equilibrium.* I conjecture that in general, when the sets of actions are "large enough," no partner goes bankrupt in some equilibrium.

4.1. *AN EXAMPLE WITH BANKRUPTCY*

There are two partners 1 and 2 who can take one of two actions 1 or 0 (e.g., work hard or shirk). The output function is defined by $y(0,0)=y_0=0$, $y(0,1)=y(1,0)=y_1=1$, $y(1,1)=y_2=\pi$, $\pi>1$. The partners have the same disutility function which is given by $v(0)=0$, $v(1)=1$. The efficient vector of actions is $(1,1)$ when $\pi\geq 2$. Let α_i be the share of the first partner when the output is y_i, where $i=0,1,2$. Thus $y_i-\alpha_i$ is the share of the second partner. There will exist a sharing rule α for which (1,1) is a noncooperative equilibrium if and only if $1 \leq \alpha_2 - \alpha_1 \leq \pi - 2$, i.e., if $\pi \geq 3$. In this case, there is an infinite number of solutions. Note that only y_1 is attainable by a unilateral deviation.

When $\pi \in [2,3)$, efficiency cannot be sustained. In the context of the example, a random device is a 4-tuple

$$P = (P(0,0), P(1,0), P(0,1), P(1,1)),\ \forall i,j \in \{0,1\},\ P(i,j) \geq 0,\ \sum_{i,j} P(i,j) = 1,$$

where $P(i,j)$ is the probability that the first partner is told to do action i and the second partner is told to do action j. These probabilities are common knowledge. A group mechanism consists of a pair (P,α) where α is defined as above. This mechanism is an equilibrium mechanism if, when a player is told to take action $a_i \in \{0,1\}$, he has no incentive to make action $a \neq a_i$ when he believes that the other partner will obey the mechanism. I.e., if 1 is asked to take action a_1, he will believe that the other partner has been asked to take action a_2 with probability $P(a_1,a_2)/[P(a_1,0)+P(a_1,1)]$. This gives us four incentive compatibility constraints.

(4.1a) $\quad -b \cdot P(0,0) - a \cdot P(0,1) \geq -P(0,0) - P(0,1)$

(4.1b) $\quad b \cdot P(1,0) + a \cdot P(1,1) \geq P(1,0) + P(1,1)$

(4.1c) $\quad b \cdot P(0,0) + a \cdot P(1,0) \geq (\pi - 2) \cdot P(1,0)$

(4.1d) $\quad -b \cdot P(0,1) - a \cdot P(1,1) \geq -(\pi - 2) \cdot P(1,1).$

where $a \equiv \alpha_2 - \alpha_1$ and $b \equiv \alpha_1 - \alpha_0$. Assume that $\epsilon > 0$. A mechanism is an ϵ-equilibrium mechanism if it is an equilibrium mechanism and if

(4.1e) $\quad \sum_{i,j} P(i,j) \cdot [y(i,j) - v(i) - v(j)] = \pi - 2 - \epsilon,$

where $\pi - 2$ is the sum of the utilities of the two partners when they take efficient actions. Thus, an ϵ-equilibrium mechanism solves (4.1a)-(4.1e). When $\pi \geq 3$, the mechanism in which $P(1,1)=1$ and $1 \leq \alpha_2 - \alpha_1 \leq \pi - 2$ is a 0-equilibrium mechanism. When $\pi \in [2,3)$, there does not exist a 0-equilibrium mechanism. We know from Theorem 1 that, for any positive ϵ, there exists an ϵ-equilibrium mechanism. (4.1e) is satisfied if and only if $P(1,1) = 1 - \frac{\epsilon}{\pi - 2}$. Thus, there is no loss of generality in supposing that $P(1,1)=1-\epsilon$. I will restrict myself to mechanisms in which $P(1,1)=1-\epsilon$. In this case, it is only necessary to verify that (4.1a)-(4.1d) and the balance condition $P(0,0)+P(1,0)+P(0,1)=\epsilon$ are satisfied.

Let $\gamma > \frac{1-\epsilon}{\epsilon}$ be a scalar. Let (P,α) be the following mechanism,

(4.2) $$P(0,0)=\frac{\epsilon\cdot(\gamma+1)-1}{\gamma\cdot(\gamma+1)};\ P(1,0)=\frac{\epsilon\cdot(\gamma+1)-1}{\gamma+1};\ P(0,1)=\frac{1-\epsilon}{\gamma};\ P(1,1)=1-\epsilon;$$

$$a=\frac{\gamma\cdot K\cdot(\pi-2)-K-1}{\gamma\cdot K-1},\ b=\gamma\cdot(\pi-2-a)$$

where $$K \equiv \frac{\epsilon\cdot(\gamma+1)-1}{(\gamma+1)\cdot(1-\epsilon)}.$$

This mechanism is well defined if $K \neq 1/\gamma$ which is always possible by an appropriate choice of γ (there are infinitely many γ for which this is possible). To verify that this mechanism is an $\epsilon\cdot(\pi-2)$-equilibrium mechanism, observe that $\frac{P(1,1)}{P(0,1)}=\frac{P(1,0)}{P(0,0)}=\gamma$. Consequently, (4.1a)-(4.1d) are consistent if and only if the following system has a solution

(4.3a) $$b+\gamma\cdot a=(\pi-2)\cdot\gamma$$

(4.3b) $$P(0,0)\cdot(b-1)=(1-a)\cdot P(0,1).$$

The reader can easily verify that the above values of a and b are the unique solutions of this system. Let $P_1(0)=\frac{1}{\gamma+1}$, $P_1(1)=\frac{\gamma}{\gamma+1}$, $P_2(0)=\frac{\epsilon\cdot(\gamma+1)-1}{\gamma}$, $P_2(1)=\frac{(\gamma+1)\cdot(1-\epsilon)}{\gamma}$. Then, $\forall a_1 \in \{0,1\}$, $\forall a_2 \in \{0,1\}$ $P(a_1,a_2)=P_1(a_1)\cdot P_2(a_2)$. I.e., the partners follow a mixed strategy. There exist an infinite number of values for α_i, where $i=0,1,2$, corresponding to such a and b. In fact, it can be shown that ϵ-equilibrium mechanisms *must* be of the above form (see Appendix B). Consequently, for this example, for any ϵ, there exist an infinite number of ϵ-equilibrium mechanisms, all parametrized by a scalar γ.

The details in Appendix B explain why there is no solution when $P(1,1)=1$, i.e., when we try to enforce the efficient vector of actions. By introducing nonzero (but negligible) probabilities that the partners will take inefficient actions one can induce them to almost always take an efficient action.

Remark 7: A mechanism is here a point in the space $E\equiv\Delta^3\times\Sigma(Y)$, where Δ^3 is the 3-dimensional simplex and $\Sigma(Y)\equiv\cup\Sigma(y)$, where the union is taken over the possible outputs. We can define the product topology in E and we have as a consequence a natural measure of distance between mechanisms. If we define a correspondence ϕ which maps each ϵ in the set of ϵ-equilibrium mechanisms $\left\langle P(\epsilon),s(\epsilon)\right\rangle$ of E, it is clear that $\forall\epsilon>0$, $\phi(\epsilon)\neq\emptyset$ and that ϕ is continuous on $(0,1]$ but is discontinuous at 0. In fact, $\lim_{\epsilon\downarrow 0}\phi(\epsilon)=\emptyset$, and ϕ is not upper-hemicontinuous at 0 (note that ϕ is trivially lower-hemicontinuous). •

After straightforward, but tedious, computations, it can be shown that as $\epsilon\downarrow 0$, i.e., $\gamma\uparrow\infty$, $a\rightarrow\pi-2$, $b\rightarrow\infty$. Since $b\equiv\alpha_1-\alpha_0$, either $\alpha_1\rightarrow\infty$ or $\alpha_0\rightarrow-\infty$. Thus, one partner will have to pay a very large amount when y_1 or y_0 is produced. If both partners have finite initial endowments or do not have enough assets to borrow the necessary amount of money, they will know that for small ϵ, one of them will not be able to meet his or her financial obligations. Does this mean that such mechanisms are not feasible for small ϵ? Suppose that partner 1 has to pay a very large amount to partner 2 when output y_0 is produced. Suppose that when 1 is not able to pay, 2 can seek compensation before the legal authorities (there exist criminal courts and financial commitments can be enforced) and lead to a situation that is comparable to "ruin", "lifetime in jail", "death" or other "worst outcome" for 1. With the proposed sharing rules, 1 will avoid taking action 0 whenever there is a nonzero probability that 2 takes action 0, since there is a small probability that he will be ruined (note that (4.1a)-(4.1b) are incentive compatible conditions *only if* partner 1 believes that the sharing rule α is feasible). Knowing this, 2 will take action 1 or 0 *given that* 1 takes action 1 if he expects that 2 will take action 0 with positive probability, and given the initial sharing rules, but ignoring the correlated device. It can be shown that for small ϵ, a is never equal to $\pi-2$. (Precisely, it can be shown that $a>(<)\pi-2$ as $\delta\leq(>)2\cdot(1-\epsilon)/\epsilon$.) Consequently, 2 will want to put all the probability mass on action 0 (1) when $1-\alpha_1>(<)\pi-\alpha_2-1$. (1,1) cannot be sustained as a Nash equilibrium; consequently, only (1,0) can be an equilibrium. But (1,0) could have been enforced as a Nash equilibrium (choose $\alpha_1\geq\alpha_0+1$ and $\alpha_2-\alpha_1\geq\pi-2$). Thus there was no need to design an ϵ-mechanism in which 1 commits to be ruined if y_0 is realized!

The previous paragraph points out the consequences of designing an ϵ-equilibrium mechanism and "forgetting" the feasibility of the mechanism. In that case nothing is gained for the partners since their expected utility is zero. However, the idea that players can commit to infeasible sharing rules is potentially interesting. By committing to a sharing rule in which he will be ruined if output y_0 is realized, partner 1 commits credibly to take action 1 whenever there is a positive probability that partner 2 takes action 0. With mechanisms defined by (4.2), partner 2 will never be indifferent between the two actions and will always choose action 0 in equilibrium (because the incentive compatibility constraints which correspond to the ϵ-mechanism when the feasibility constraint is "forgotten" implies that such indifference is not possible). There is consequently a simple solution to the partnership problem if we accept the possibility of a player to commit to infeasible sharing rules: let α_0 be so large (negatively) that partner 1 will not be able to pay such amount and will have a utility of $-\infty$ if y_0 is realized. Let $\alpha_2-\alpha_1=\pi-2$ and let, for $\epsilon>0$, $P(1,0)=\epsilon$, $P(1,1)=1-\epsilon$, $P(0,0)=P(0,1)=0$. Thus, the correlated device P corresponds to partner 1 using a pure strategy and partner 2 using a mixed strategy. Because $\epsilon>0$, partner 1 wants to play 1 and partner 2 is indifferent (ex-ante) between his two pure strategies, so he can as well obey the device P (he cannot use $\epsilon=0$ since this will not be an equilibrium).

This reasoning is correct as long as we believe in the possibility for players to commit themselves to schemes which are not feasible. Observe that such schemes are not trembling hand perfect, since if partner 1 believes that there is a small probability that he will tremble, he will never agree to sign contracts in which he can be ruined). These mechanisms have some appeal. They correspond, to a certain extent, to an attenuation of Schelling's (1963) ideas, later developed by Crawford (1982), that impasses in bargaining situations arise because the bargainers try to convince their opponent(s) that they will not retreat from certain strategies, i.e., try to attain the best bargaining positions. With partnerships, impasses (to attain efficiency) arise because the players might not be able to commit credibly to actions which are best for both of them. According to Schelling, players can destroy a unit of production or go on strike in order to commit to a line of action. Here a player commits to taking an efficient action by signing a contract in which he can be ruined for certain contingencies.

4.2. *AN EXAMPLE WITH NO BANKRUPTCY*

There are two partners with the same set of actions {0,1,2}. The output and the disutility functions are symmetric, i.e., $y(a_1,a_2)=y_{a_1+a_2}$, $v_1(a)=v_2(a)=v(a)$. The parameters of the example are

$y_0=0;\ y_1=2;\ y_2=3.5;\ y_3=5;\ y_4=6$

$v(0)=0;\ v(1)=1;\ v(2)=3.$

Efficiency is attained at (1,1), the output is $y_2=3.5$ and the sum of the utilities is 1.5. I will use here reasoning which is close in spirit to the observations made in Section 3 and in the example of Section 1.2. Consider the following probability measures

$P_1(0)=\epsilon_1,\ P_1(1)=1-\epsilon_1,\ P_1(2)=0,$
$P_2(0)=0,\ P_2(1)=1-\epsilon_2,\ P_2(2)=\epsilon_2,$

where $\epsilon_i\in(0,1)$. The resulting random device has the property that the messages that each partner receives are not correlated. With these probability measures, there is a zero probability that, when the partners obey the device, output levels y_4 and y_0 are observed. Moreover, if either one of these output levels is observed, then it is possible to know precisely *who* deviated. Indeed, to obtain output y_4 it is necessary that partner 1 uses action $a_1=2$ and to obtain output y_0 it is necessary that partner 2 uses action $a_2=0$. Let $A_1=\{0,1\}$, $A_2=\{1,2\}$, and suppose that it is possible to find a sharing rule s such that $(\tilde{P},s)$ is an ϵ-EGM of the game in which A_1 and A_2 are the sets of actions and in which $\tilde{P}$ coincides with $P=P_1\cdot P_2$ over $A_1\otimes A_2$. Then it is easy to obtain an ϵ-EGM for the original game. Indeed, if y_4 is observed, give a large negative amount to partner 1 and if y_0 is observed, give a large negative amount to partner 2. Because, deviations are informative when y_4 or y_0 are observed, it is indeed possible to punish the player who deviates. Figure 2.5 illustrates this reasoning.

The gray area is the "reduced" game that I consider, and it corresponds to the actions which are supported by P_1 and P_2. It is clear from the figure why punishments are possible. Let α_k be the share of the first partner when the output is y_k. It is simple algebra to show that for $\tilde{P}$ as given above, the sharing rule s which solves the following system is such that $(\tilde{P},s)$ is an ϵ-EGM of the reduced game (where ϵ is a function of ϵ_1, ϵ_2)

$$\alpha_1=\frac{1-2\cdot\epsilon_2}{1-\epsilon_2}\cdot\alpha_2+\frac{\epsilon_2}{1-\epsilon_2}\cdot\alpha_3-\frac{1}{1-\epsilon_2}$$

$$\alpha_2=\left(\epsilon_1\cdot\frac{1-2\cdot\epsilon_2}{1-\epsilon_2}+1-2\cdot\epsilon_2\right)^{-1}\cdot\left(\left(1-\epsilon_1-\frac{\epsilon_1\cdot\epsilon_2}{1-\epsilon_2}\right)\cdot\alpha_3-3.5\right).$$

When ϵ_1 and ϵ_2 are close to 0, i.e., when we get close to full efficiency, it is possible to choose α_3 in such a way that $\alpha_3\in(0,5)$, $\alpha_2\in(0,3.5)$ and $\alpha_1\in(0,2)$, i.e., such that no partner goes bankrupt even when they have a zero initial endowment. Out of the equilibrium, e.g., when partner 1 uses action $a_1=2$ or partner 2 uses action $a_2=0$, for any pair (ϵ_1,ϵ_2), there exists a large enough punishment when y_4 or y_0 are observed which will deter the partners from deviating.

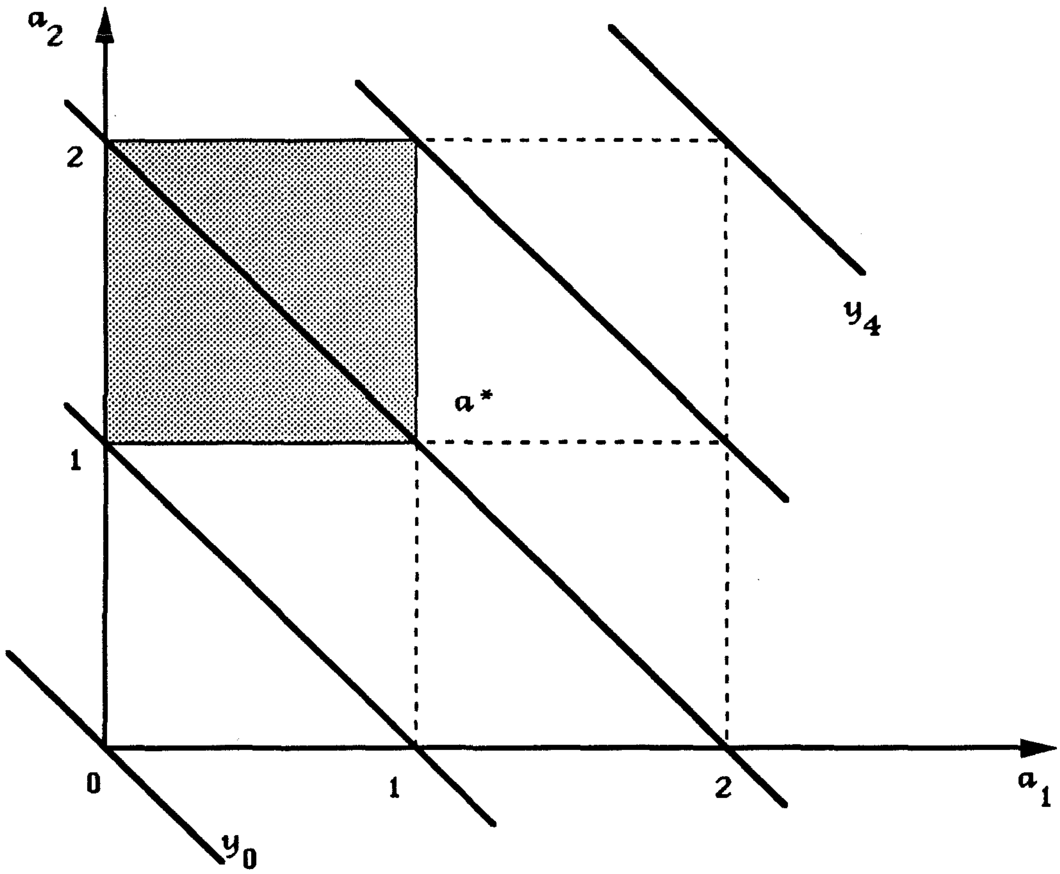

Figure 2.5

An ϵ-group mechanism

5. PROOFS

I will use below the following result of Fan that I present here as a lemma.

LEMMA 2 (Fan): *Let* $\{x_v\}_{v\in I}$ *be a family of elements in a real normed linear space* X*, and let* $\{\beta_v\}_{v\in I}$ *be a corresponding family of real numbers. Then for any* $\rho\geq 0$*, the following two conditions are equivalent:*

(i) There exists a continuous linear functional f *on* X *with* $\|f\|\leq\rho$ *such that* $f(x_v)\geq\beta_v$*, where* $v\in I$.

(ii) For any number m *of indices* $v_1, v_2,\ldots,v_m$ *of* I *and for any* m *positive numbers* $\lambda_1,\lambda_2,\ldots,\lambda_m$*, the inequality*

$$(5.1)\quad \rho\cdot\left\|\sum_{i=1}^{m}\lambda_i\cdot x_{v_i}\right\|\geq\sum_{i=1}^{m}\lambda_i\cdot\beta_{v_i}$$

holds.

Proof: Theorem 12, pages 123-125 in Fan (1956). It follows from Theorem 1, the compactness (in the weak*-topology on X^*) of the set $\{f\in X^* \mid \|f\|\leq\rho\}$ (Alaoglu's Theorem) and a separating hyperplane argument. □

5.1. *PROOFS OF THE RESULTS OF SECTION* 2

Sketch of the proof: Let $\Gamma=\left\langle\{A_i\},\{v_i\},y\right\rangle$ be a partnership problem. I suppose that H1-H3 hold. Let a^* be an efficient vector of actions (existence is assumed). I construct a sequence of games $\Gamma^r=\left\langle\{A_i^r\},\{v_i^r\},y^r\right\rangle$ as follows. For each r ($r\in\mathbb{Z}_+$), $A_i^r\subset A_i$ is a finite set of actions such that for all i, $a_i^*\in A_i^r$. The sequence $\{A_i^r\}$ is nested, i.e., $r>r'\Rightarrow A_i^r\supset A_i^{r'}$. Define $A_i^\infty\equiv\bigcup_{r\in\mathbb{Z}_+}A_i^r$, $A^r\equiv\times_i A_i^r$, $A^\infty\equiv\times_i A_i^\infty$. It is always possible to find sequences $\{A_i^r\}$ such that A^∞ is dense in A (e.g., choose A_i^r such that the Hausdorff distance between A_i^r and A_i is less than $1/r$). Once A^r is defined, y^r and v_i^r are the restrictions of y and of v_i to A^r and A_i^r (i.e., y^r and y agree on A^r, v_i^r and v_i agree on A_i^r).

I will show that for any subgame Γ^r, there exists an ϵ-equilibrium group mechanism (P^r,s^r). The proof will be constructive; the probability measure that I will use has the property that the conditional beliefs of a player about the messages received by the other players are independent of his own message. In other words, the event a_i is independent of the event a_j, for any partners i and j and any actions in A_i and A_j. It can easily be shown that the sequence $\{P^r\}$ converges weakly to a probability measure P^∞ on $(A^\infty,\mathcal{A}^\infty)$. I will suppose that there is no sharing rule s^∞ such that the pair (P^∞,s^∞) is an ϵ-EGM of the game Γ^∞ and reach a contradiction since this would imply that there exists r large enough for which there is no sharing rule s^r such that (P^r,s^r) is an ϵ-EGM of the game Γ^r. To finish the proof, I will show that there is an extension s of s^∞ such that (P^∞,s) is an ϵ-EGM of the game Γ (this will follow the continuity of y and v_i and the denseness of A^∞ in A). H3 is used only in the last part of the proof. Existence of ϵ-EGM when the sets of actions are discrete or countable does not depend on the continuity of the functions y or v_i.

5.1.1. *Finite games*

Otherwise stated, the next results are related to the finite games Γ^r. To save notation, I will suppress the superscript r. The first result shows that assumption H2 implies a lower bound on the cardinality of the set of output levels. T_i is the cardinality of the set of actions of partner i and l is the cardinality of the set of outputs.

LEMMA 3: *Under H2,* $l=\#Y\geq\sum_{i=1}^{n}T_i-n+1$.

Proof: Start with the vector $(a_1(1),\ldots,a_n(1))$ of A. By varying a_1 in $A_1\backslash\{a_1(1)\}$, H2 implies that one can obtain T_1-1 different levels of output. Consider the vector $(a_1(T_1),a_2(1),\ldots,a_n(1))$. By varying a_2 in $A_2\backslash\{a_2(1)\}$, one can obtain T_2-1 different levels of output. From H2, each output $y((a_1(T_1),a_2,\ldots,a_n(1))$ is greater than $y(a_1(T_1),a_2(1),\ldots,a_n(1))$ for $a_2\in A_2\backslash\{a_2(1)\}$.

Consider the vector $\left(a_1(T_1),a_2(T_2),\ldots,a_{i-1}(T_{i-1}),a_i(1),a_{i+1}(1),\ldots,a_n(1)\right)$. With the same argument as before, by varying a_i in $A_i\backslash\{a_i(1)\}$, one can obtain T_i-1 levels of outputs, all different from the previous levels of outputs. By doing the same reasoning for all the partners, the result follows. □

In the following, I will consider the case in which the probability measure P is such that there exist probability measures $P_1,\ldots,P_n$ such that

$$\forall a\in A,\ P(a)=P_1(a_1)\cdot P_2(a_2)\cdot\ldots\cdot P_n(a_n),$$

which is equivalent to supposing that the partners follow mixed strategies. In such a case, the conditional probability measure is given by $P(\hat{a}_{-i}|a_i)=\prod_{j\neq i}P_j(\hat{a}_j)$. I will write $(a_i\rightarrow\tilde{a}_i)$ to denote that partner i takes (deviates to) action $\tilde{a}_i$ when he is told to do action a_i and I will denote by $E1(a_i\rightarrow\tilde{a}_i)$ the corresponding equilibrium condition. We have,

$$\sum_{\hat{a}_{-i}\in A_{-i}}\prod_{j\neq i}P_j(\hat{a}_j)\cdot\left(s_i(y(\hat{a}\backslash a_i))-s_i(y(\hat{a}\backslash\tilde{a}_i))\right)\geq v_i(a_i)-v_i(\tilde{a}_i),$$

$$\sum_{\hat{a}_{-i}\in A_{-i}}\prod_{j\neq i}P_j(\hat{a}_j)\cdot\left(s_i(y(\hat{a}\backslash\tilde{a}_i))-s_i(y(\hat{a}\backslash a_i))\right)\geq v_i(\tilde{a}_i)-v_i(a_i).$$

The first line is the content of $\text{E1}(a_i\rightarrow\tilde{a}_i)$, the second line refers to $\text{E1}(\tilde{a}_i\rightarrow a_i)$. Let

$supp(P_i)$ be the set of a_i such that $P_i(a_i)>0$. If a_i and $\tilde{a}_i$ are in the support of P_i, then there must be an equality in $E1(a_i \rightarrow \tilde{a}_i)$. This is the well known fact that, in a mixed strategy equilibrium, every player must be indifferent between the pure strategies that he plays with positive probability. If $a_i \in supp(P_i)$ and $\tilde{a}_i \notin supp(P_i)$, partner i must (at least weakly) prefer the first action, i.e., $E1(a_i \rightarrow \tilde{a}_i)$ must be satisfied with an inequality.

It follows that condition E1 can be written

$\forall i$, $\forall a_i \in supp(P_i)$, $\forall \tilde{a}_i \in supp(P_i)$,

$$\sum_{\hat{a}_{-i} \in A_{-i}} \prod_{j \neq i} P_j(\hat{a}_j) \cdot \Big(s_i(y(\hat{a} \backslash a_i)) - s_i(y(\hat{a} \backslash \tilde{a}_i)) \Big) = v_i(a_i) - v_i(\tilde{a}_i).$$

$\forall i$, $\forall a_i \in supp(P_i)$, $\forall \tilde{a}_i \in A_i$,

$$\sum_{\hat{a}_{-i} \in A_{-i}} \prod_{j \neq i} P_j(\hat{a}_j) \cdot \Big(s_i(y(\hat{a} \backslash a_i)) - s_i(y(\hat{a} \backslash \tilde{a}_i)) \Big) \geq v_i(a_i) - v_i(\tilde{a}_i).$$

These conditions can be represented in matrix form by $B \cdot s \geq \beta$, where the rows of the matrix B correspond to deviations $\big(a_i(h) \rightarrow a_i(p)\big)$, if $a_i(h) \in supp(P_i)$. Precisely,

(5.2)
$$B = \begin{bmatrix} B(1) & 0 & \cdots & 0 \\ 0 & B(2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & B(n-1) \\ B(n) & B(n) & \ldots & B(n) \end{bmatrix}$$

Each matrix $B(i)$ is a $\big[\#supp(P_i) \cdot (T_i - 1)\big] \times l$ matrix, with row elements $B(i)\big(a_i(h) \rightarrow a_i(p)\big)$, where $a_i(h) \in supp(P_i)$ and where the k-th element of this vector is,

(5.3) For $i<n$,

$$B(i)\big(a_i(h) \rightarrow a_i(p)\big)(y_k) = \sum_{\hat{a}_{-i} \in A_{-i}} \prod_{j \neq i} P_j(\hat{a}_j) \cdot \Big[\chi_{E(y_k; a_i(h))}(a_{-i}) - \chi_{E(y_k; a_i(p))}(a_{-i}) \Big]$$

$$B(n)\Big(a_n(h)\rightarrow a_n(p)\Big)(y_k)=\sum_{\hat{a}_{-n}\in A_{-n}}\prod_{j\neq n}P_j(\hat{a}_j)\cdot\Big[\chi_{E(y_k;a_n(p))}(a_{-n})-\chi_{E(y_k;a_n(h))}(a_{-n})\Big]$$

where, $E(y_k;a_i)\equiv\{a_{-i}\in A_{-i};\ y(a_{-i},a_i)=y_k\}$ and χ_C is the indicator function of the set C ($\chi_C(\alpha)=1$ if $\alpha\in C$, $\chi_C(\alpha)=0$ if $\alpha\notin C$). $E(y_k;a_i)$ is the set of actions that the other partners can take and which together with a_i generate output y_k.

The reason for the expressions in (5.3) is the following. If i is told to do a_i, he will compare his expected utility of obeying the machine and playing a_i versus his expected utility of disobeying it and playing $\tilde{a}_i$. I.e., partner i will compare the probability of obtaining a given output y_k under each possibility. The term in (5.3) is the difference between these two probabilities. If $s_i(y_k)$ is the share of partner i ($i<n$) if output y_k is observed, then the variation in his expected revenues when output y_k is observed is equal to $B(i)(a_i\rightarrow\tilde{a}_i)(y_k)\cdot s_i(y_k)$. It follows that the total variation in expected revenues for partner i is

$$\sum_{y_k\in Y}B(i)(a_i\rightarrow\tilde{a}_i)(y_k)\cdot s_i(y_k)=\sum_{a_{-i}\in A_{-i}}\prod_{j\neq i}P_j(a_j)\cdot\Big(s_i(y(a))-s_i(y(a\backslash\tilde{a}_i))\Big).$$

Partner i will compare this value to his variation in disutility which is equal to

$$v_i(a_i)-v_i(\tilde{a}_i)=\beta(a_i\rightarrow\tilde{a}_i),\text{ if } i<n.$$

The conditions corresponding to partner n include the balance condition $s_n(y)=y-\sum_{i\neq n}s_i(y)$. For this reason, the expression in (5.3) is not the variation in expected revenue of partner n when output y_k is observed. Similarly, the term $\beta(a_n\rightarrow\tilde{a}_n)$ is not his variation in disutility.

s is a vector in $\left(\mathbb{R}^{n-1}\right)^l$ with the interpretation that the first $n-1$ components are the shares of partners 1, 2,..., $n-1$ of output y_1, etc. β is a $\sum_i T_i-n$ vector where the first T_1-1 components are equal to

$$\beta\left(a_1(1)\rightarrow a_1(h)\right)=v_1\left(a_1(1)\right)-v_1\left(a_1(h)\right),\text{ for } h=2,3,\ldots,T_1,$$

and similarly for the first $n-1$ first partners. For the last partner, the last T_n-1 components are equal to

$$\beta\left(a_n(1)\rightarrow a_n(h)\right)=$$

$$\sum_{\hat{a}_{-n}\in A_{-n}}\ \prod_{j\neq n} P_j(\hat{a}_j)\cdot\left(y(\hat{a}_{-n},a_n(h))-y(\hat{a}_{-n},a_n(1))\right)+v_n\left(a_n(1)\right)-v_n\left(a_n(h)\right).$$

I will say that the probability measure P_i is *complete* when $supp(P_i)=A_i$, i.e., when each action can be played with positive probability (this corresponds to the notion of completely mixed strategy). If the probability measure is complete for a partner, then the incentive conditions for this partner are equalities. Indeed, if the incentive conditions $E1\left(a_i(1)\rightarrow a_i(h)\right)$ are satisfied when h varies in $\{1,\ldots,T_i\}$, then all incentive conditions $E1\left(a_i(h)\rightarrow a_i(k)\right)$ are also satisfied (by taking the difference between $E1\left(a_i(1)\rightarrow a_i(h)\right)$ and $E1\left(a_i(1)\rightarrow a_i(k)\right)$. It follows that it is enough to consider deviations from $a_i(1)$ for any partner i when all the probability measures P_i are complete.

LEMMA 4: *For any i, the l-vectors $B(i)\left(a_i(1)\rightarrow a_i(h))\right)$ are independent whenever the probability measures P_i satisfy $P_i(a_i(T_i))>0$.*

Proof of Lemma 4: Suppose that for each i, $P_i(a_i(T_i))>0$. For each partner i, denote by $m_i(a_i)$ the maximum index of the output that it is possible to attain when i takes action a_i. I.e., $y_{m_i(a_i)}=\max\{y_k\colon\ \exists a_{-i}\in A_{-i},\ y_k=y(a_{-i},a_i)\}$. From H2, $y_{m_i(a_i)}$ is always equal to $y(a_{-i}(T_{-i}),a_i)$, where $a_{-i}(T_{-i})$ is a short hand notation for $\forall j \neq i$, $a_j = a_j(T_j)$. Also from H2 it follows that $\forall i \in \{1,\ldots,n\}$, $\forall j \in \{1,\ldots,T_i\}$,

$$B(i)\Big(a_i(1) \to a_i(h)\Big)\Big(y_{m_i(a_i)}\Big) = 0, \text{ if } a_i(h) < a_i,$$

(5.4)
$$B(i)\Big(a_i(1) \to a_i(h)\Big)(y) = 0, \text{ if } y > y_{m_i(a_i(h))},$$

$$B(i)\Big(a_i(1) \to a_i(h)\Big)\Big(y_{m_i(a_i(h))}\Big) = \prod_{k \neq i} P_k\Big(a_k(T_k)\Big), \text{ if } i < n,$$

$$B(n)\Big(a_n(1) \to a_n(h)\Big)\Big(y_{m_n(a_n(h))}\Big) = \prod_{k \neq n} P_k\Big(a_k(T_k)\Big).$$

Suppose that the vectors $B(i)\Big(a_i(1) \to a_i(h)\Big)$ are dependent. Then there exist $T_i - 1$ scalars $\lambda(h)$ such that for all y_k,

$$\sum_{h=1}^{T_i} \lambda(h) \cdot B(i)\Big(a_i(1) \to a_i(h)\Big)(y_k) = 0.$$

From (5.4), $\lambda(T_i) = 0$, otherwise,

$$\sum_{h=1}^{T_i} \lambda(h) \cdot B(i)\Big(a_i(1) \to a_i(h)\Big)(y_l) = \prod_{k \neq i} P_k(a_k(T_k))$$

since the terms $B(i)\Big(a_i(1) \to a_i(h)\Big)(y_l)$ are all equal to zero for $h < T_i$. Given that $\lambda(T_i) = 0$, it must be true that $\lambda(T_i - 1) = 0$ for the same reason as above. Repeating the same reasoning finishes the proof since there are $T_i - 1$ rows in $B(i)$. □

LEMMA 5: *Suppose that H1 and H2 are satisfied. For any* $\epsilon > 0$, *there exists probability measures* $P_1, \ldots, P_n$ *such that for all* i, $P_i(a_i^*) > 1 - \delta$, *such that the probability measures* $P_1, \ldots, P_{n-1}$ *are complete and such that there exists a sharing rule for which* (P, s) *is an* ϵ*-EGM.*

Proof of Lemma 5: Since $P_i(a_i^*)>1-\delta$, it is possible to choose $\epsilon>1-(1-\delta)^n$. Thus, it is enough to check the equilibrium conditions. The proof is by induction. Consider probability measures $P_1,\ldots,P_{n-1}$ such that for all $i<n$, $P_i(a_i^*)>1-\delta$, and P_i is complete. From Lemma 4, the rows of each submatrix $B(i)$, $i\geq 1$, are independent since for all i, $P_i(a_i(T_i))>0$. I fix the sets of actions of partners 1 to $n-1$. I want to show that the assertion of the lemma is true for any set of actions of partner n.

Induction hypothesis: *Given sets $A_1,\ldots,A_{n-1}$, the statement of Lemma 5 is true for any partnership problem satisfying assumptions H1-H2 in which the set of actions of the n-th partner has less than T_n actions.*

The induction hypothesis is true if $T_n=1$. In this case, $A_n=\{a_n^*\}$ and the matrix B in (5.2) can be written in the form

$$B=\begin{bmatrix} B(1) & 0 & \cdots & 0 \\ 0 & B(2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & B(n-1) \end{bmatrix}$$

since partner n cannot deviate. By completeness of the probability measures P_i, for $i<n$, the rows of each submatrix $B(i)$ consist of the deviations $\left(a_i(1)\rightarrow a_i(h)\right)$, $h\geq 2$. From Lemma 4, each submatrix is of full rank. It follows that the matrix B itself is of full rank. Note that when A_n is a singleton, partner n plays the role of budget breaker. We will see in Theorem 2 that, in this case, a^* is sustainable in pure strategies. Here, I have proved that there are infinitely many ϵ-EGM in which the probability measures of at least $n-1$ partners are complete.

I suppose that the induction hypothesis is true for partnership problems in which the set A_n has T_n actions and I show that it is true at T_n+1.

Case 1: $a_n^* < a_n(T_n+1)$. In this case, it is possible to set $P_n(a_n(T_n+1))=0$ without contradicting the choice of $P_n(a_n^*)>1-\delta$. The partnership problem under consideration is defined by $\Gamma=\left\langle\{A_i\},\{v_i\},y\right\rangle$. Consider the *reduced partnership problem* $\hat{\Gamma}=\left\langle\{\hat{A}_i\},\{\hat{v}_i\},\hat{y}\right\rangle$ where,

(5.5a) $\forall i<n,\ \hat{A}_i=A_i$.

(5.5b) $\hat{A}_n=A_n\setminus\left\{a_n(T_n+1)\right\}$.

(5.5c) $\hat{y}$ agrees with y on $\hat{A}$ (i.e., $\hat{y}=y$ on $\hat{A}$).

(5.5d) $\forall i$, $\hat{v}_i$ agrees with v_i on $\hat{A}_i$ (e.g., $\hat{v}_i=v_i$ on A_i when $i<n$, $\hat{v}_n=v_n$ on $A_n\setminus\{a_n^*\}$).

By the induction hypothesis and by (5.5a), there exist probability measures $P_1,\ldots,P_{n-1}$ on $A_1,\ldots,A_{n-1}$ which are complete such that $P_i(a_i^*)>1-\delta$, and a probability measure $\hat{P}_n$ on $\hat{A}_n$ such that $\hat{P}_n(a_n^*)>1-\delta$ for which the resulting system $\hat{B}\cdot\hat{s}\geq\hat{\beta}$ has a solution. From (5.5a)-(5.5c), the range of $\hat{y}$ is

$$\hat{Y}=Y\setminus\left\{y_k: y(a)=y_k\Rightarrow a_n=a_n(T_n+1)\right\}. \tag{5.6}$$

In particular, the maximal output that can be attained in the partnership game Γ cannot be attained in the partnership game $\hat{\Gamma}$. Indeed, from H2,

$$\left[y_l=y(a)\right] \Rightarrow \left[a=\left(a_1(T_1),\ldots,a_{n-1}(T_{n-1}),a_n(T_n+1)\right)\right].$$

Because the probability measures P_i are complete for $i<n$, if partner n uses the probability measure $\hat{P}_n$ in the original partnership game Γ, it is known that the

output y_l *cannot* be attained when every partner follows the instructions of the device. Moreover, if y_l is observed while the device $\mathcal{P}=\left\langle P_i, i<n, \hat{P}_n\right\rangle$ is used, then it must be true that partner n deviated.

From above, in the game $\hat{\Gamma}$, there exists an ϵ-EGM $(\mathcal{P},\hat{s})$ which satisfies the conditions of the lemma. From (5.6), the sharing rule $\hat{s}$ does not specify the shares when y_l is observed since y_l does not belong to $\hat{Y}$, the range of $\hat{y}$. It follows that in order to prove that there exists a sharing rule s defined on Y such that the pair $(\mathcal{P},s)$ is and ϵ-EGM of Γ, it is necessary and sufficient to show that partner n does not want to deviate to $a_n(T_n+1)$. Consider the sharing rules s on Y which satisfy, $\forall y \in \hat{Y}$, $s(y)=\hat{s}(y)$. Let $U_n(\mathcal{P},\hat{s})$ and $U_n(\mathcal{P},s)$ be the expected utility function of partner n in the games $\hat{\Gamma}$ and Γ respectively. We have

$$
\begin{aligned}
U_n(\mathcal{P},s)\Big(a_n(T_n+1)\Big) &= \sum_{a_{-n}\in A_{-n}} \prod_{i\neq n} P_i(a_i)\cdot s_n\Big(y(a_{-n},a_n(T_n+1))\Big)-v_n\Big(a_n(T_n+1)\Big) \\
&= \sum_{a_{-n}\in \tilde{A}_{-n}} \prod_{i\neq n} P_i(a_i)\cdot s_n\Big(y(a_{-n},a_n(T_n+1))\Big) \\
&\qquad + \prod_{i\neq n} P_i(a_i(T_i))\cdot s_n(y_l)-v_n\Big(a_n(T_n+1)\Big)
\end{aligned}
$$

where $\tilde{A}_{-n}\equiv A_{-n}\setminus\Big\{\Big(a_1(T_1),\ldots,a_{n-1}(T_{n-1})\Big)\Big\}$.

Define arbitrary values for $s_i(y)$, $i=1,2,\ldots,n$ when $y\in Y\setminus\Big(\hat{Y}\cup\{y_l\}\Big)$. Define $s(y_l)$ in such a way that

$$
s_n(y_l) < \frac{U_n(\mathcal{P},\hat{s})(a_n^*)-\displaystyle\sum_{a_{-n}\in \tilde{A}_{-n}} \prod_{i\neq n} P_i(a_i)\cdot s_n\Big(y(a_{-n},a_n(T_n+1))\Big)-v_n\Big(a_n(T_n+1)\Big)}{\displaystyle\prod_{i\neq n} P_i(a_i(T_i))}.
$$

This expression is well defined by completeness of the measures P_i. The new sharing rule s is well defined since the shares of the partners for any output level in Y are well defined. From the value of $s_n(y_l)$, it is immediate that the difference $U_n(\mathcal{P},s)\Big(a_n(T_n+1)\Big)-U_n(\mathcal{P},\hat{s})(a_n^*)$ is negative. Recall that $\hat{P}_n(a_n^*)>1-\delta$, and that $U_n(\mathcal{P},\hat{s})(a_n^*)=U_n(\mathcal{P},s)(a_n^*)$ since outputs in $Y\backslash\hat{Y}$ are not attainable when every partner obeys the device $\mathcal{P}$. It follows that in the game Γ, when partner n believes that the other partners obey the device $\mathcal{P}$, he does not want to deviate to action $a_n(T_n+1)$. Partner n does not want to deviate to actions $a_n\notin supp(\hat{P}_n)\backslash\Big\{a_n(T_n+1)\Big\}$ since he did not want to deviate to these actions in the reduced game $\hat{\Gamma}$ and since by deviating to such actions partner n can generate only outputs in $\hat{Y}$. Thus, it is an equilibrium strategy for partner n to follow the instructions of the device $\mathcal{P}$ in the game Γ. It is obviously an equilibrium strategy for the other partners to follow the instructions of the device $\mathcal{P}$ since the games Γ and $\hat{\Gamma}$ "look" the same for these partners in terms of incentives when the device assigns a zero probability to action $a_n(T_n+1)$.

Case 2: $a_n^*=a_n(T_n+1)$. It is possible to replicate exactly the reasoning of case 1 by considering the reduced game $\hat{\Gamma}$, where (5.5b) is replaced by

(5.5b$'$) $$\hat{A}_n=A_n\backslash\Big\{a_n(1)\Big\}. \quad \square$$

Lemma 5 proves the following theorem.

THEOREM 4: *For any partnership game with finite sets of actions which satisfies H1-H2, for any positive ϵ, there exists an ϵ-equilibrium group mechanism. Moreover, for any $\epsilon>0$, there exists $\delta(\epsilon)>0$, $\delta(\epsilon)$ an increasing function of ϵ, such that there exists an ϵ-EGM (P,s) with the property that the game induced by s has a mixed strategy* *equilibrium in which each partner assigns a probability greater than $1-\delta(\epsilon)$ to his efficient action.*

Proof: Lemma 5 gives an example of ϵ-EGM which satisfies the conditions of the theorem. □

5.1.2. *General sets of actions–A convergence result*

From Theorem 4, for any $\epsilon>0$, each game Γ^r has an ϵ-EGM. Let $\left\{(P^r,s^r)\right\}$ be a sequence of ϵ-EGM. Such a sequence exists by Theorem 4 and by the construction in Lemma 4. Each probability measure P^r can be looked at as a probability measure on A^r or as a probability measure on A whose support is A^r. A first lemma establishes that the sequence $\{P^r\}$ is weakly convergent.

LEMMA 6: *The sequence $\{P^r\}$ converges weakly to a probability measure on A.*

Proof: By Prohorov Theorem (Billingsley (1968)), the family of probability measures $\{P^r\}$ is relatively compact if it is tight. Tightness follows completeness and separability of A. □

Since A^∞ is the limit of the sets A^r (the supports of the measures P^r,) there exists a measure of probability on $(A^\infty,\mathcal{A}^\infty)$ which is the limit of some subsequence of $\{P^r\}$. I call P^∞ this limit. Thus, the set $A\backslash A^\infty$ has P^∞-measure zero. s^∞ is a solution to the partnership problem Γ^∞ when the correlated device is P^∞ if each condition $\mathrm{E1}(a_i\rightarrow\tilde{a}_i)$ is satisfied. These conditions can be written in matrix form $B\cdot s^\infty\geq\beta$, where B, s^∞ and β have the same interpretation as in Section 5.1.1. I will consider the vector $B(i)(a_i \to \tilde{a}_i)$ to be a vector in the space $L \equiv \left(\mathbb{R}^{n-1}\right)^{Y^{\infty}}$ corresponding to the row $(a_i \to \tilde{a}_i)$ of the matrix B (Y^{∞} is the set of possible output levels when the set of joint actions is A^{∞}). With an abuse of notation, I will denote by $B(i)(a_i \to \tilde{a}_i)(y_k)$ the element in the column corresponding to output y_k and in the row corresponding to the deviation $(a_i \to \tilde{a}_i)$ of the submatrix $B(i)$. Note that because P^{∞} has countable support, the set of outputs that can be attained when i deviates to $\tilde{a}_i$ (in A_i^{∞}) while he expects the other partners to take actions in A_{-i}^{∞} is also countable. As in Section 5.1.1, it can be shown that (recall that $a_i \in supp_i(P^{\infty})$, $a_n \in supp_i(P^{\infty})$.)

$$(5.7) \qquad B(i)(a_i \to \tilde{a}_i)(y_k) = \sum_{a_{-i} \in A_{-i}^{\infty}} \left[\chi_{E(y_k;a_i)}(a_{-i}) - \chi_{E(y_k;\tilde{a}_i)}(a_{-i})\right] \cdot P^{\infty}(a_{-i}|a_i), \quad \text{if } i<n$$

$$B(n)(a_n \to \tilde{a}_n)(y_k) = \sum_{a_{-n} \in A_{-n}^{\infty}} \left[\chi_{E(y_k;\tilde{a}_n)}(a_{-n}) - \chi_{E(y_k;a_n)}(a_{-n})\right] \cdot P^{\infty}(a_{-n}|a_n)$$

where the notation has the same interpretation as in Section 5.1.

LEMMA 7: *There exists a sharing rule s^{∞} such that (P^{∞}, s^{∞}) is an ϵ-EGM of the game Γ^{∞}. Moreover, s^{∞} is bounded ($\|s^{\infty}\| \le \rho$, some ρ).*

Proof: Suppose that there does not exist a sharing s^{∞} rule such that (P^{∞}, s^{∞}) is an ϵ-EGM of Γ^{∞}. Clearly, because E2 is independent of the sharing rule, E2 is satisfied at P^{∞} for some ϵ since E2 is satisfied at any P^r. Consequently, it is only necessary to check E1. I apply Lemma 2.

From Lemma 2, for any $\rho>0$, there exist a finite set of indices $J=\{(a_i \to \tilde{a}_i)\}$ and positive numbers $\lambda(a_i \to \tilde{a}_i)$ such that

$$(5.8)\qquad \rho\cdot\left\|\sum_J \lambda(a_i\rightarrow\tilde{a}_i)\cdot B(i)(a_i\rightarrow\tilde{a}_i)\right\| < \sum_J \lambda(a_i\rightarrow\tilde{a}_i)\cdot\beta(a_i\rightarrow\tilde{a}_i).$$

I will argue that this leads to a contradiction. By construction of the sequence $\{A_i^r\}$, there exists $\tilde{r}$ large enough such that all the actions in the deviations of J are in A^r for $r\geq\tilde{r}$. Since (5.8) is true for any ρ, it is true for any ρ^r, where ρ^r is such that $\|s^r\|\leq\rho^r$.[15]

Case 1: $\left\|\sum_J \lambda(a_i\rightarrow\tilde{a}_i)\cdot B(i)(a_i\rightarrow\tilde{a}_i)\right\|\neq 0$. Let $\delta(\rho,\lambda)$ be the difference between the right hand side and the left hand side in (5.8). To obtain a contradiction, it is enough to show that there is r large enough such that the absolute value (where $\rho>\rho^r$)

$$(5.9)\qquad \rho\cdot\left|\left\|\sum_J \lambda(a_i\rightarrow\tilde{a}_i)\cdot B(i)(a_i\rightarrow\tilde{a}_i)\right\| - \left\|\sum_J \lambda(a_i\rightarrow\tilde{a}_i)\cdot B^r(i)(a_i\rightarrow\tilde{a}_i)\right\|\right|$$

is strictly less than $\delta(\rho,\lambda)$, where B^r refers to the (finite dimensional) matrix (5.9) induced by the game Γ^r. Precisely,

(5.9')

$$B^r(i)(a_i\rightarrow\tilde{a}_i)(y^k)=\sum_{a_{-i}\in A_{-i}^{\infty}} P^r(a_{-i}|a_i)\cdot\left[\chi_{E(y_k;a_i)}(a_{-i})-\chi_{E(y_k;\tilde{a}_i)}(a_{-i})\right],$$

where the summation is taken over A_{-i}^{∞} instead of A_{-i}^r because for any $a_{-i}\in A_{-i}^{\infty}\setminus A_{-i}^r$, $P^r(a_{-i}|a_i)=0$. For any v, there exists r large enough, such that for all $a_{-i}\in A_{-i}^{\infty}$, $\left|P^{\infty}(a_{-i}|a_i)-P^r(a_{-i}|a_i)\right|<v$ (since P^{∞} has countable support and since $P^r\rightarrow P^{\infty}$).

The right hand side of equality (5.9') is linear in $P^r(\cdot|a_i)$. Thus, for any $\gamma>0$, there exists r large enough (i.e., v small enough,) such that,

$$\left|B(i)(a_i\rightarrow\tilde{a}_i)(y_k)-B^r(i)(a_i\rightarrow\tilde{a}_i)(y_k)\right|<\gamma.$$

Thus, the difference $\left|\sum_J \lambda(a_i\rightarrow\tilde{a}_i)\cdot B(i)(a_i\rightarrow\tilde{a}_i)(y_k) - \sum_J \lambda(a_i\rightarrow\tilde{a}_i)\cdot B^r(i)(a_i\rightarrow\tilde{a}_i)(y_k)\right|$

can be made as small as we want (since J is finite) by choosing r large enough. It follows that the difference between the norms in (5.9) can be made less than $\delta(\rho,\lambda)$ since the norm is a continuous operator. Thus, inequality (5.8) holds in the game Γ^r for $\rho > \rho^r$. This contradicts the fact that (P^r, s^r) is an ϵ-EGM of Γ^r.

Case 2: $\left\| \sum_J \lambda(a_i \to \tilde{a}_i) \cdot B(i)(a_i \to \tilde{a}_i) \right\| = 0$. This implies that for all level of output $y_k \in Y^\infty$, $\sum_J \lambda(a_i \to \tilde{a}_i) \cdot B(i)(a_i \to \tilde{a}_i)(y_k) = 0$. In particular, this is true for any $y_k \in Y^r$. As in case 1, we can deduce that there exists r large enough such that the difference $\left| \sum_J \lambda(a_i \to \tilde{a}_i) \cdot B^r(i)(a_i \to \tilde{a}_i)(y_k) \right|$ is as small as we want (this expression is well defined for $y_k \in Y^r$). But then, we conclude that there exists r for which, inequality (5.8) hold. This is a contradiction.

Combining cases 1 and 2, the lemma is proved. □

To finish the proof of Theorem 3, I will show that there exists a sharing rule defined on Y such that the pair (P^∞, s) is an ϵ-EGM of the game Γ. This is where continuity of the functions y and v_i becomes important.

LEMMA 8: *Suppose that assumptions H1-H3 are satisfied. Then, there exists an extension s of s^∞ to Y such that the pair (P^∞, s) is an ϵ-EGM of Γ.*

Proof: Y is a subset of $\mathbb{R}$ and can be considered as a metric space, with the usual metric. By continuity of the output function y and by denseness of the set A^∞ in A, the set Y^∞ is dense in Y. Precisely, the closure of Y^∞ is Y, i.e., for any output $y \in Y \setminus Y^\infty$, there exists a sequence $\{y^k\} \subset Y^\infty$ such that $y = \lim y^k$. Define for $i \neq n$,

$$s_i(y)=\begin{cases} s_i^{\infty}(y) & \text{if } y\in Y^{\infty} \\ \lim\limits_{\delta\downarrow 0} \inf\limits_{y^k\in Y^{\infty}\cap N(y,\delta)} s_i^{\infty}(y^k) & \text{if } y\notin Y^{\infty}, \end{cases}$$

where $N(y,\delta)$ is the ball around y of radius δ. By Lemma 7, $s_i^{\infty}(y^k)$ is bounded for any $y^k\in Y^{\infty}$. Moreover, the sequence $\left\{\inf\left\{s_i^{\infty}(y^k)\colon\ y^k\in N(y,\delta)\right\},\ \delta>0\right\}$ is monotone increasing in δ. Thus, the limit exists and $s_i(y)$ is well defined for $i\neq n$. For partner n, define $s_n(y)$ as the residual $y-\sum_{i\neq n} s_i(y)$. s is consequently a well defined sharing rule. The claim is that (P^{∞},s) is an ϵ-EGM of the game Γ. Clearly, it is only necessary to verify that the equilibrium conditions are satisfied since the sum of the expected utilities under the mechanism (P^{∞},s) coincide with the sum that is obtained under the group mechanism (P^{∞},s^{∞}) in the game Γ^{∞} (in equilibrium, no actions in $A\backslash A^{\infty}$ are taken).

Suppose that (P^{∞},s) is not an ϵ-EGM. Then, for some partner $i\neq n$, there exist actions $a_i\in A_i^{\infty}$ and $\tilde{a}_i\in A_i\backslash A_i^{\infty}$ such that,

$$\sum_{\hat{a}_{-i}\in A_{-i}^{\infty}} P^{\infty}(\hat{a}_{-i}|a_i)\cdot s_i(y(\hat{a}_{-i},a_i))-v_i(a_i) < \sum_{\hat{a}_{-i}\in A_{-i}^{\infty}} P^{\infty}(\hat{a}_{-i}|a_i)\cdot s_i(y(\hat{a}_{-i},\tilde{a}_i))-v_i(\tilde{a}_i).$$

I.e., i wants to deviate to an action which does not belong to $supp_i(P^{\infty})$ when he is told to do an action a_i which is in $supp_i(P^{\infty})$. By denseness of the set A_i^{∞}, there exists a sequence $\{\tilde{a}_i^k\}\subset A_i^{\infty}$ such that $\tilde{a}_i=\lim \tilde{a}_i^k$. By definition of the sharing rule s,

$$\sum_{\hat{a}_{-i}\in A_{-i}^{\infty}} P^{\infty}(a_{-i}|a_i)\cdot s_i(y(\hat{a}_{-i},\tilde{a}_i))-v_i(\tilde{a}_i)$$

$$= \sum_{\hat{a}_{-i}\in A_{-i}^{\infty}} P^{\infty}(\hat{a}_{-i}|a_i)\cdot\left[\lim_{\delta\downarrow 0} \inf_{y^k\in Y^{\infty}\cap N(y(\hat{a}_{-i},\tilde{a}_i),\delta)} s_i(y^k)\right]-v_i(\tilde{a}_i)$$

$$\leq \sum_{\hat{a}_{-i}\in A_{-i}^{\infty}} P^{\infty}(\hat{a}_{-i}|a_i)\cdot\left[\lim_{\delta\downarrow 0} \inf_{\tilde{a}_i^k\in A_i^{\infty}\cap N'(\tilde{a}_i,\delta)} s_i^{\infty}(y(\hat{a}_{-i},\tilde{a}_i^k))\right]-v_i(\tilde{a}_i)$$

$$\leq \lim_{\delta \downarrow 0} \inf_{\tilde{a}_i^k \in A_i^{\infty} \cap N'(\tilde{a}_i,\delta)} \left[\sum_{\hat{a}_{-i} \in A_{-i}^{\infty}} P^{\infty}(\hat{a}_{-i}|a_i) \cdot s_i^{\infty}(y(\hat{a}_{-i},\tilde{a}_i^k)) - v_i(\tilde{a}_i^k) \right].$$

The first equality is by definition of s. The first inequality follows since we take the infimum over a subset. The second inequality follows the continuity of v_i (lim inf $v_i(\tilde{a}_i^k) = v_i(\tilde{a}_i)$) and the fact that $\sum(\inf) \leq \inf(\sum)$. Consequently, there exists k large enough such that

$$\sum_{\hat{a}_{-i} \in A_{-i}^{\infty}} P^{\infty}(\hat{a}_{-i}|a_i) \cdot s_i^{\infty}(y(\hat{a}_{-i},a_i)) - v_i(a_i) <$$

$$\sum_{\hat{a}_{-i} \in A_{-i}^{\infty}} P^{\infty}(\hat{a}_{-i}|a_i) \cdot s_i^{\infty}(y(\hat{a}_{-i},\tilde{a}_i^k)) - v_i(\tilde{a}_i^k),$$

but this contradicts the fact that (P^{∞},s^{∞}) is an ϵ-EGM of the game Γ^{∞}.

From the previous arguments, for each $\tilde{a}_n \in A_n \setminus A_n^{\infty}$, there exists $\tilde{a}_n^k$ such that the difference

$$\sum_{i \neq n} \left| \sum_{\hat{a}_{-n} \in A_{-n}^{\infty}} P^{\infty}(\hat{a}_{-n}|a_n) \cdot s_i^{\infty}(y(\hat{a}_{-n},\tilde{a}_n^k)) - \sum_{\hat{a}_{-n} \in A_{-n}^{\infty}} P^{\infty}(\hat{a}_{-n}|a_n) \cdot s_i(y(\hat{a}_{-n},\tilde{a}_n)) \right|$$

is as small as we want. By continuity of y and of v_n, it follows that partner n has no incentive to deviate to an action in $A_n \setminus A_n^{\infty}$ (otherwise, one obtains a contradiction by using the same reasoning as for a partner $i \neq n$). □

Finally, all the reasoning could have been made when the probability device is such that the actions of the players are independent random variables, i.e., when we consider a sequence of $\epsilon-$EGM for which the device corresponds to a mixed strategy equilibrium of the game induced by the sharing rule. Theorem 1 is consequently proved. □

5.2. *PROOFS OF THE RESULTS OF SECTION* 3

Proof of Lemma 1: From H4, $l \geq T$. It follows that there are at least $(n-1)\cdot T$ columns in the matrix B. By construction, there are $n\cdot(T-1)$ rows in B. First, I show that the $(n-1)\cdot(T-1)$ first rows of B are independent. From (3.9), if "-1" appears in row r and column c of the block $B(i)$, then it must be true that $c=\tau(i,r)$. Consider $j \neq i$ and suppose that there exists r' such that -1 appears in row r' and column c of the block B_j. By construction, it must be true that

$$[\tau(i,r)-\tau(j,r')]\cdot(n-1)+i-j=0.$$

Since $i \neq j$, $|i-j|<n-1$ and $|i-j|/(n-1)<1$, which is a contradiction. Independence follows. Consider now the block $B(n)$. In row r of this block $n-1$ ones appear between columns $(\tau(n,r)-1)\cdot(n-1)+1$ and $\tau(n,r)\cdot(n-1)$. Since $r \neq r' \Rightarrow \tau(n,r) \neq \tau(n,r')$, it follows that the other elements of the above columns of the matrix $B(n)$ consist of zeros. Independence follows.

It is now enough to consider the case when one of the rows of $B(n)$ is dependent of the rows of the matrix obtained by deleting the rows of $B(n)$. Suppose that there exist scalars $\lambda(i,k)$, not all zero, such that

$$\exists r,\ \forall c,\ B(n)(r,c)=\sum_{i \neq n} \sum_{k=1}^{T-1} \lambda(i,k)\cdot B(i)(k,c).$$

From (3.9), for $(\tau(n,r)-1)\cdot(n-1)+1 \leq c \leq \tau(n,r)\cdot(n-1)$, $B(n)(r,c)=1$. For these values of c, $B(i)(k,c)=-1$ if $c=\tau(i,k)$ and $B(i)(k,c)=0$ otherwise. Since for $(t-1)\cdot(n-1)+1 \leq c \leq t\cdot(n-1)$, $\sum_{i \neq n} \sum_{k=1}^{T-1} \lambda(i,k)\cdot B(i)(k,c)=1$, it must be true that for each i there exists *exactly one* k such that $\lambda(i,k) \neq 0$, i.e., such that $\lambda(i,k)=-1$, *and* such that $\tau(i,k)=\tau(n,r)$. Reciprocally, if the condition $\forall i \neq n$, $\exists k=f(i)$, $\tau(i,k)=\tau(n,r)$, then the *r-th* row of the block $B(n)$ is equal to the opposite of the sum of the rows $f(i)$ of each block $B(i)$, $i \neq n$. Thus, the rank of B is indeed determined by the condition of the lemma. □

Proof of Theorem 2: I use a basic result of linear algebra, (e.g., Fan (1956:Theorem 1)) which tells us that the system $B \cdot s \geq \beta$ has a solution in s if and only if when there exist non-negative scalars $\lambda(i,k)$ such that

(5.10) $$\forall c=1,\ldots,(n-1)\cdot(T-1),\ \sum_{i=1}^{n}\sum_{k=1}^{T-1}\lambda(i,k)\cdot B(i)(k,c)=0$$

then

$$\sum_{i=1}^{n}\sum_{k=1}^{T-1}\lambda(i,k)\cdot\beta i(k)\leq 0.$$

If B is of full rank, the result is immediate since the scalars are all equal to zero. From Lemma 1, this is equivalent to (i) of Theorem 2. Suppose that the rank of B is one less than its number of rows. From Lemma 1, it follows that there exists a unique r such that $\chi(r)=1$ and that $\lambda(i,k)\neq 0$ if and only if $a_i(k)=f(i)$, $f\in\mathfrak{F}_0(r)$ and $\lambda(i,f(i))/\lambda(n,r)=+1$. Thus, (5.10) implies, in this case,

(5.11) $$\forall c,\ B(n)(r,c)=-\sum_{i\neq n}B(i)(f(i),c).$$

Observe that (5.11) is the only possibility. In particular, there cannot exist $g\in\mathfrak{F}_0(r)$, $g\neq f$ such that $\forall c$, $B(n)(r,c)=-\frac{1}{2}\cdot[\sum_{i\neq n}B(i)(f(i),c))+\sum_{i\neq n}B(i)(g(i),c))]$. Indeed, if this were true, then $\forall i\neq n$, $\tau(i,f(i))=\tau(i,g(i))$ where $f(i)\neq g(i)$, which is impossible by H4. Existence of a solution to (3.8) is now equivalent to the fact that the following inequality holds,

$$\beta n(r)+\sum_{i\neq n}\beta i(f(i))\leq 0.$$

This can be rewritten as

$$y_{\tau(n,r)}-y^{*}+v_n(a_n^{*})-v_n(r)+\sum_{i\neq n}v_i(a_i^{*})-\sum_{i\neq n}v_i(a_i(f(i)))\ \leq\ 0,$$

since $\forall i\neq n$, $\tau(i,f(i))=\tau(n,r)$, it follows that

(5.12) $$y^* - \sum_i v_i(a_i^*) \geq y_{\tau(n,r)} - \sum_i v_i(a_i(f(i))).$$

From Lemma 1 and the above observations, it is clear that (5.12) holds if and only if $\exists r_1,\ldots,\exists r_h$, $h \leq \sum_{r=1}^{T-1} \chi(r)$, such that $\forall k$, $\lambda(n,r_k) \neq 0$, it must be true that $\lambda(i,r) \neq 0$ if and only if $\exists k$, such that $r = f_k(i)$. In this case $\lambda(i,f_k(i)) = \lambda(n,r_k)$ and

$$\forall k=1,\ldots,h,\ \forall c,\ B(n)(r_k,c) = -\sum_{i \neq n} B(i)(f_k(i),c)).$$

Applying (5.12) h times finishes the proof of (ii) of Theorem 2. The maximum dimension of a linear manifold contained in the solution set is equal to the dimension of the original space minus the maximum of independent linear functionals, i.e., is $l \cdot (n-1) - rank(B)$, (Fan (1956:Theorem 7)). □

Proof of Theorem 2: The proof of this theorem uses Lemma 2. An immediate corollary of Lemma 2 is that (5.1) holds for some ρ when the left hand side is nonzero if and only if the expression on the right hand side of (5.1) is finite for any λ and any m indices $(i_1,r_1)\ldots(i_m,r_m)$. In that case, choose $\rho = \sup[\sum_{j=1}^{m} \lambda(i_j,r_j) \cdot \beta i_j(r_j)]$ where $\lambda(i_j,r_j) > 0$ and $\left\| \sum_{j=1}^{m} \lambda(i_j,r_j) \cdot B(i_j)(r_j) \right\| = 1$; see also Theorem 13 in Fan (1956). Define,

$$\left\| \sum_{j=1}^{m} \lambda(i_j,r_{i_j}) \cdot B(i_j)(r_{i_j}) \right\| = \sup_c \left| \sum_{j=1}^{m} \lambda(i_j,r_{i_j}) \cdot B(i_j)(r_j,c) \right|.$$

Clearly, this norm is equal to zero if and only if the vectors $B(i_j)(r_{i_j})$ are positively dependent. Moreover, it is easy to see that this norm is finite since the matrix B is bounded (from the discussion in the text, the element of L which consists of "1" is an upper bound for the row elements of B).

Consider first the case of independence. By setting the above norm equal to 1 implies that the scalars are bounded. It will follow that ρ is finite.

If $\forall j=1,\ldots m$, $i_j=n$, then the sum $\sum_{j=1}^{m} \lambda(i_j,r_j)\cdot\beta i_j(r_j)\leq 0$ since $\beta n(r_j)\leq 0$ for all r_j. Thus, (5.1) holds for such indices. Reorder the indices $i_1,\ldots,i_m$ such that if $\exists k$ s.t. $i_k=n$, then $i_j=n$, for all $j'\geq k$. If such a k does not exist, let $k\equiv m+1$. By definition, $\sum_{j=1}^{m} \lambda(i_j,r_j)\cdot\beta i_j(r_j)$ is equal to

$$(5.13) \qquad \sum_{j\leq k-1} \lambda(i_j,r_j)\cdot[v_{i_j}(a^*_{i_j})-v_{i_j}(a_{i_j}(\tilde{r}_j))]$$

$$+\sum_{j\geq k} \lambda(n,r_{i_j})\cdot[y_{\tau(n,\tilde{r}_{i_j})}-y^*+v_n(a^*_n)-v_n(a_n(\tilde{r}_{i_j}))]$$

where $\forall i_j$, $\tilde{r}_{i_j}=r_{i_j}$ if $a_{i_j}(r_{i_j})<a^*_{i_j}$ and $\tilde{r}_{i_j}=r_{i_j}+1$ if $a_{i_j}(r_{i_j})>a^*_{i_j}$. The second term in the sum is finite and nonpositive. Suppose that the set of indices is such that there exists j for which $v_{i_j}(a_{i_j}(r_j))=\infty$. Then, the sum in (5.13) is equal to $-\infty$. Thus, the supremum ρ must be finite.

Consider now the case of dependence. With the previous notation, and recalling the arguments of Lemma 1, since (5.1) must be true for any dependent system, it must be true for $\{(1,f(1)),(2,f(2)),\ldots,(n-1,f(n-1)),(n,r))\}$, where, $\forall i$, $\tau(i,f(i))=\tau(n,r)$. In this case, $\forall i_j$, $\lambda(i_j,r_{i_j})=1$ and thus, (ii) of Theorem 1 must hold. If this is true for any r for which $\chi(r)=1$, it follows that (5.1) is satisfied for any dependent finite set of rows of B. □

6. CONCLUSION

6.1. *GROUP MECHANISMS AND THE THEORY OF THE FIRM*

A firm and a partnership both fulfill a contracting and an informative role. They define a contract between the owners of the outputs and the owners of the inputs of production (who coincide for partnerships) and define networks by which the information is transmitted, analyzed, and redistributed. With a firm, where ownership and production are separated, the owners will try to maximize their rent. With a partnership, the partners try to maximize (under risk neutrality) the sum of their utilities. In each case, moral hazard can lead to inefficiencies.

Whereas the literature has pointed out for a long time that a firm is compatible with efficiency, the analysis of this chapter suggests that partnerships can also have this property. Baker et al. (1988:606) point out that "The productive effects and popularity of profit-sharing plans are poorly understood by economists." This chapter provided a possible explanation which does not rely on monitoring or a separation of production and control. If the partners believe that the other partners do not always take an efficient action, then it might be optimal for them to almost always take an efficient action.

If one acknowledges the fact that partnerships can be efficient forms of organization, one is led to ask new questions. In particular, it is necessary to go beyond the classical principal-agent framework and to look at mechanisms which are designed by a *group* of individuals. Moreover, it is necessary to give reasons for the existence of hierarchical forms of organization which are not based on efficiency; efficiency does not seem any longer to be a sufficient criterion for comparing the merits of partnerships and classical firms.

6.2. *RELATION WITH THE PRINCIPAL-AGENTS APPROACH*

As noted in the introduction, there is an obvious relationship between the approach of this chapter and the approach taken in the principal-agents literature. In particular, the approach of Myerson (1983) which transforms problems of moral hazard (and adverse selection) into games of communication with a mediator can be used to point out the main differences between the two approaches. In the principal-agents approach, the principal has his own utility function that he wants to maximize.

A partnership problem can be transformed into a principal-agents problem if the principal has a utility function of the form, $U_0(a)=y(a)-\sum_{i=1}^{n} v_i(a_i)$ and if all the profit is distributed to the partners (or for that matter, all the profit minus an amount which is independent of the output,) i.e., if the principal must choose a sharing rule in $\Sigma(y)$. In this case, there does not exist an incentive-efficient compatible mechanism. In this respect, partnership problems are a very special case of principal-agents problems in which there is no incentive-efficient mechanism.

6.3. *GROUP MECHANISMS AND CORRELATED EQUILIBRIA*

While correlated equilibria have received much attention recently, it is not known how much better the players can be with correlation than without. Partial answers to this question have been provided in Rosenthal (1974) for some examples and in Moulin and Vial (1978) in a more general framework. Moulin and Vial (1978) give some necessary and some sufficient conditions for having correlated equilibria improving upon Nash equilibria but they do not tell us the extent of the improvement.

The present chapter tells us the following. If, for a given sharing rule, there exists a mixed strategy equilibrium of the resulting game (i.e., a group mechanism where the correlated device is independent), then, by definition, it is possible that there exists a correlated equilibrium at which efficiency is at least as large as with the mixed strategy Nash equilibria. However, from the main result of this chapter, Nash equilibria (in mixed strategies) are compatible with ϵ-efficiency, for any $\epsilon>0$. Thus, as ϵ goes to zero, it must be true that the correlated equilibria improve only in a negligible way on Nash equilibria.

6.4. *STATICS VERSUS DYNAMICS*

The approach taken in this chapter is static. Radner (1986) shows that if the partners do not discount the future and if the partnership game is repeated infinitely many times, efficiency can be attained. Radner, Myerson and Maskin (1986) study an example with discounting in which it is not possible to attain full efficiency, and where every supergame equilibrium leads to a discounted sum of utilities which is uniformly bounded below full efficiency. These two papers are to my knowledge the only works on repeated partnerships and they emphasize the role played by discounting. Because discounting is a very natural assumption, the result of Radner et al. (1986) proves that inefficiency in partnerships can persist in a repeated framework. However, as Radner and Williams (1988) show, this negative result is due to the special assumptions that are made in Radner et al. (1986).

The relative merits of a static versus a dynamic approach must be evaluated with respect to the economic environment that one desires to explain. In some situations, the assumption of infinite repetition is not tenable. This is the case for most cooperative projects of research and development where the project is highly specific (e.g., the space station). If there is repetition of the cooperation in later periods, it will be in a completely different economic environment; thus, the standard assumption of the repeated game literature that the environment is fixed is not adequate. In other cases, the repeated game approach might be a good approximation of reality.

6.5. *EXTENSIONS AND RELATED WORK*

The most obvious extensions of the results of this chapter concern uncertain output and risk aversion. I conjecture that the results of Section 2 extend to these two cases. There are already some results concerning full efficiency in the literature. When the output is a random variable, Radner and Williams (1987) show that it is possible for the partners to attain efficiency when the density function of the output has some (strong) properties. With risk-aversion, Rasmusen (1987) observes that if the sharing rule can be made random, then it is possible to attain the first best. When the sharing rule is not stochastic, partnerships with risk averse partners cannot in general attain full efficiency.

APPENDIX A

This Appendix provides the proof of the claims made in the example of Section 1.2.

1. *An equilibrium in the reduced game*

Consider the following distribution function. $p=Prob(a_1=1)$, $1-p=Prob(a_1=0)$, $p=Prob(a_2=1)$ and $1-p=Prob(a_1=0)$. Suppose that $1>p>1/2$. Thus, all actions in the interiors of $\tilde{A}_1$ and $\tilde{A}_2$ have a zero probability. Given these distribution functions, the expected utility functions of the two players are,

$$U_1(a_1)=p\cdot s(a_1+1)+(1-p)\cdot s(a_1+2)-\frac{a_1^2}{2}, \text{ for } a_1\in\tilde{A}_1,$$

$$U_2(a_2)=(1-p)\cdot(a_2-s(a_2))+p\cdot(a_2+1-s(a_2+1))-\frac{a_2^2}{2}, \text{ for } a_2\in\tilde{A}_2.$$

Define s such that

$$\text{(A.1)}\qquad \begin{aligned} &s(1)=s(3)\\ &0<s(1)<p-\frac{1-p}{2\cdot(2\cdot p-1)}\\ &s(2)=s(1)+\frac{1}{2\cdot(2\cdot p-1)}. \end{aligned}$$

It follows that for these values,

$$\begin{aligned} U_1(0)&=U_1(1)\\ &=s(1)+\frac{1-p}{2\cdot(2\cdot p-1)}\\ &=EU_1\\ U_2(1)&=U_2(2)\\ &=p-s(1)-\frac{1-p}{2\cdot(2\cdot p-1)}\\ &=EU_2\\ &=p-EU_1. \end{aligned}$$

With the assumptions on s, it follows that $EU_i > 0$, for $i=1,2$.

It is now necessary to consider deviations of the players in the interior of their reduced set of actions $\tilde{A}_i$, i.e., on (0,1) for Mr. 1 and on (1,2) for Ms. 2. Until now, the sharing rule has been defined for output levels in the set {1,2,3}. I will show that it is possible to extend continuously this sharing rule to the interval [1,3], i.e., the range of the output function on $\tilde{A}_1 \times \tilde{A}_2$, in such a way that each player is strictly worse off by deviating in the interior of his set of actions.

Let $Q(\alpha,\beta,\gamma)$ denote the quadratic function $Q(\alpha,\beta,\gamma)(y)=\alpha\cdot y^2+\beta\cdot y+\gamma$. I define the following function on the interior of $\tilde{A}_1 \times \tilde{A}_2$,

(A.2) $\quad s(y)=Q(\alpha/2,\beta,\gamma)(y)$ if $y\in(1,2)$

$\quad =Q(-\rho/2,\sigma,\xi)(y)$ if $y\in(2,3)$,

where

(A.3) $\quad \rho > \mathrm{Max}\left\{\frac{1}{2\cdot p-1},\frac{1}{p}\right\}=\frac{1}{2\cdot p-1}$ for $p>\frac{1}{2}$

(A.4) $\quad \frac{1+(1-p)\cdot\rho}{p} < \alpha < \frac{p\cdot\rho-1}{1-p}$.

Note that the inequalities in (A.4) are compatible since ρ obeys (A.3). In order to have continuity at $y=1$, $y=2$ and $y=3$, the parameters of the quadratic functions must satisfy,

(A.5a) $\quad s(1)=\alpha/2+\beta+\gamma \quad (Q(\alpha/2,\beta,\gamma)(1)=s(1))$

(A.5b) $\quad s(3)=-\rho/2+\sigma+\xi \quad (Q(-\rho/2,\sigma,\xi)(3)=s(3))$

(A.5c) $\quad s(2)=2\cdot\alpha+2\cdot\beta+\gamma \quad (Q(\alpha/2,\beta,\gamma)(2)=s(2))$

(A.5d) $\quad s(2)=-2\cdot\rho+2\cdot\sigma+\xi \quad (Q(-\rho/2,\sigma,\xi)(2)=s(2))$

where $s(1)$, $s(2)$ and $s(3)$ satisfy (A.1) and (A.2).

One can show that the system given by (A.1)-(A.5d) admits a solution. For instance, by taking

$$s(1)=s(3)=\frac{1}{2\cdot(2\cdot p-1)},$$

$$s(2)=2\cdot s(1),$$

$$\alpha=\frac{1-2\cdot p+(1+2\cdot p^2-2\cdot p)\cdot\rho}{2\cdot p\cdot(1-p)},$$

$$\rho>\frac{1}{2\cdot p-1},$$

conditions (A.1)-(A.4) are satisfied (α is chosen as the average of the two bounds in (A.3)). To solve (A.5), it is enough to solve the system

$$\begin{cases} \beta+\gamma=\dfrac{1}{2\cdot(2\cdot \mathrm{p}-1)}-\dfrac{\alpha}{2} \\ \sigma+\xi=\dfrac{1}{2\cdot(2\cdot \mathrm{p}-1)}+\dfrac{\rho}{2}. \end{cases}$$

This system has clearly a solution in $(\beta,\gamma,\sigma,\xi)$.

(Note that by adding the condition $2\cdot\alpha+\beta=-2\cdot\rho+\sigma$, one could ensure that the first order derivatives of the two quadratic functions coincide at $y=2$.)

It follows that the expected utility functions $U_1(a_1)$ and $U_2(a_2)$ are convex on the intervals (respectively) (0,1) and (1,2). Indeed,

$$\begin{aligned} U_1''(a_1)&=p\cdot s''(a_1+1)+(1-p)\cdot s''(a_1+2)-1 \\ &=p\cdot\alpha-(1-p)\cdot\rho-1 \\ &>0 \\ U_2''(a_2)&=-(1-p)\cdot s''(a_2)-p\cdot s''(a_2+1)-1 \\ &=-(1-p)\cdot\alpha+p\cdot\rho-1 \\ &>0, \end{aligned}$$

where the inequalities follow from (A.3) and (A.4). By continuity of s and by convexity, it must be true that

$$\forall a_1 \in (0,1),\ U_1(a_1) < EU_1$$

$$\forall a_2 \in (1,2),\ U_2(a_2) < EU_2.$$

Consequently, the sharing rule defined by (A.1), (A.2), and (A.3) is together with the probability distribution defined by p an $(1-p)$-efficient group mechanism of the reduced game.

2. *An equilibrium of the original game*

If Ms. 2 takes an action a_2 in $[0,1)$, there is a positive probability that an output $y = a_2 < 1$ will result. If Mr. 1 takes an action a_1 in $(1,2]$, there is a positive probability that an output $y = a_1 + 2 > 3$ will result. When an ouput level less than 1 is observed, the partners know for sure that Ms. 2 deviated. If an output level greater than 3 is observed, the partners know for sure that Mr. 1 deviated. To insure that Ms. 2 has no incentive to deviate to $a_2 < 1$, it is enough to extend the sharing rule to $[0,1)$ in such a way that $U_2(a_2) < EU_2$, for all $a_2 < 1$. This implies that,

$$\text{(A.7)} \qquad (1-p)\cdot s(a_2) > (1-p)\cdot a_2 + p\cdot\left[a_2 + 1 - s(a_2+1)\right] - \frac{a_2^2}{2} - EU_2$$

$$= h(a_2, p) - EU_2.$$

For any p, the function $h(a_2,p)$ is bounded since the function $s(y)$ is bounded on $[1,2)$ and since $a_2 \in [0,1)$. EU_2 is an implicit function of p, and it is easy to verify that this function in concave, increasing in p and bounded by $1-s(1)$. Let $s(a_2) = \frac{K}{1-p}$, for $a_2 < 1$, where K is a positive scalar which is chosen large enough. By uniform boundeness of the right hand side of (7), K can be chosen finite. However, as $p \uparrow 1$, $s(a_2) \uparrow \infty$, i.e., very large are necessary. Note that the expected utility payoffs of Mr. 1 and of Ms. 2 are finite.

3. *Final comments*

The class of ϵ-efficient mechanisms constructed in this example have the nice feature that the partners will only take one of two actions with positive probability. Observe that the efficient vector of actions (1,1) can be supported as a pure strategy Nash equilibrium in the reduced game of Section 2. (Indeed, the conditions of Theorem 3 are satisfied.) If $a_1 < 1$, then a level of output inferior to 2 will be generated, if $a_2 > 1$, a level of output larger than 2 will be observed; thus, deviations are informative in the reduced game, i.e., Mr. 1 cannot mimic the deviations of Ms. 2. However, it is necessary that each partner is inefficient with positive probability in order to being able to use punishments for deviations outside the reduced game.

APPENDIX B

I want to show that the mechanisms described in (4.2) of Section 3 are the only ϵ-equilibrium mechanism. I rewrite the system (4.1a)-(4.1d) in matrix form $A\cdot\alpha\geq\beta$

$$\begin{bmatrix} -P(0,0) & -P(0,1) \\ P(1,0) & P(1,1) \\ P(0,0) & P(1,0) \\ -P(0,1) & -P(1,1) \end{bmatrix} \cdot \begin{bmatrix} b \\ a \end{bmatrix} \geq \begin{bmatrix} -P(0,0)-P(0,1) \\ P(1,0)+P(1,1) \\ (\pi-2)\cdot P(1,0) \\ -(\pi-2)\cdot P(1,1) \end{bmatrix}$$

I will use the same argument as in the proof of Theorem 1. Let λ_i, for i=1,2,3,4, be nonnegative scalars such that $\forall r=1,2$, $\sum_i \lambda_i\cdot B(i,r)=0$, where $B(i,r)$ is the element in the i-th row and r-th column of the matrix B. There is a solution α to $B\cdot\alpha\geq\beta$ if and only if

(B.1) $$[\forall i=1,2,3,4,\ \exists\lambda_i,\ \sum_i \lambda_i\cdot B(i,r)=0] \Rightarrow [\sum_i \lambda_i\cdot\beta_i\leq 0].$$

The left hand side of (a) implies that

(B.2) $$-\lambda_1\cdot P(0,0)+\lambda_2\cdot P(1,0)+\lambda_3\cdot P(0,0)-\lambda_4\cdot P(0,1)=0$$

(B.3) $$-\lambda_1\cdot P(0,1)+\lambda_2\cdot P(1,1)+\lambda_3\cdot P(1,0)-\lambda_4\cdot P(1,1)=0.$$

The right hand side of (a) implies that

(B.4)

$$\lambda_1\cdot(-P(0,0)-P(0,1)+\lambda_2\cdot(P(1,0)+P(1,1))+\lambda_3\cdot(\pi-2)\cdot P(1,0)-\lambda_4\cdot(\pi-2)\cdot P(1,1)\leq 0.$$

Combining (B.2) and (B.3) with (B.4) implies that

(B.5) $$\lambda_3\cdot[(\pi-3)\cdot P(1,0)-P(0,0)]+\lambda_4\cdot[P(0,1)+(3-\pi)\cdot P(1,1)]\leq 0.$$

Since (B.1) must be true for any λ_i satisfying $\sum_i \lambda_i \cdot B(i,r)=0$, it must be true that it is not possible that $\lambda_4 \neq 0$ and $\lambda_3=0$ (recall that $\pi<3$) or $(\pi-3)\cdot P(1,0)-P(0,0)=0$. When $\lambda_4=0$, (B.5) is always satisfied since $\pi-3<0$. If $\lambda_3=0$, then $B(4)$ is a linear combination of only $B(1)$ and $B(2)$. This is always impossible if and only if $B(1)$ and $B(2)$ are dependent and $B(4)$ is not dependent of either $B(1)$ or $B(2)$. It follows that

(B.6) $$\frac{P(1,1)}{P(0,1)}=\frac{P(1,0)}{P(0,0)}=\gamma$$

(B.7) $$P(0,1)\neq P(1,0).$$

(B.6) insures dependence between $B(1)$ and $B(2)$ and (B7) independence between $B(4)$ and $B(2)$. Observe that (B.6) implies that $P(i,j)\neq 0$ whenever $(i,j)\neq(1,1)$ and that $(\pi-3)\cdot P(1,0)-P(0,0)\neq 0$. Because $P(1,1)=1-\epsilon$ and $P(0,1)<\epsilon$, it follows that $\gamma>\frac{1-\epsilon}{\epsilon}$. From (B.6), it follows that (4.1a)-(4.1d) reduce to (4.3a)-(4.3b). (B.6) and (B.7) are necessary conditions for the existence of a solution. They clearly imply the other conditions of (4.2). Sufficiency is obvious.

If $P(1,1)=1$, (B.5) reduces to $\lambda_4\cdot(3-\pi)$ which is positive. We already know that there does not exist a 0-equilibrium mechanism when $3-\pi>0$. Introducing an ϵ in the analysis allows us to render $B(2)$ and $B(4)$ (the only equilibrium constraints which appear when $P(1,1)=1$) independent. □

APPENDIX C

This appendix offers an example of a process by which the partners might select a group mechanism in the environment of the example of Section 4.1. Since the main question is why the partners will ever settle on a particular ϵ, I will suppose that the partners agree to use sharing rules which correspond to the Nash cooperative solution. Nash (1950), (1953) give the axiomatic and bargaining basis for such a solution. Alternatively, one can suppose that the partners first bargain on which P (i.e., which pair (ϵ,γ)) to choose since such a pair uniquely defines the correlated device) to choose and then bargain on which shares to choose.

At the beginning of the second phase of this two stage mechanism, the expected utility of each partner is,

$$u_1(\alpha)\equiv\sum_{k,l}P(k,l)\cdot[\alpha_{k+l}-v(k)]$$

$$u_2(\alpha)\equiv\sum_{k,l}P(k,l)\cdot[y_{k+l}-\alpha_{k+l}-v(l)].$$

It can easily be shown that for a given ϵ, the Nash solution (I suppose that the disagreement point is (0,0)) implies that α must be chosen in such a way that $u_1(\alpha)=u_2(\alpha)=\frac{1-\epsilon}{2}\cdot(\pi-2)$, i.e., the utility is independent of γ. There are clearly many possible such α. However, there is a unique sharing rule which satisfies the above condition and which is compatible with an ϵ-equilibrium mechanism, i.e., satisfies (4.2). Such a sharing rule is dependent on ϵ and on γ and is equal to

(C.1) $$\alpha(\epsilon,\gamma)=B(\epsilon,\gamma)^{-1}\cdot\beta(\epsilon,\gamma)$$

where,

$$\alpha(\epsilon,\gamma)\equiv[\alpha_o(\epsilon,\gamma),\alpha_1(\epsilon,\gamma),\alpha_2(\epsilon,\gamma)]$$

$$\beta(\epsilon,\gamma)\equiv[(1-\epsilon)\cdot(\pi-1)+P(1,0)\,,a(\epsilon,\gamma)\,,b(\epsilon,\gamma)]$$

with $a(\epsilon,\gamma)$, $b(\epsilon,\gamma)$,$P(1,0)$ as in (4.2)

$$B(\epsilon,\gamma)\equiv\begin{bmatrix} P(0,0) & P(1,0)+P(0,1) & P(1,1) \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

where $P(0,0)$, $P(1,0)$, $P(1,1)$ are as in (4.2).

Observe that the matrix $B(\epsilon,\gamma)$ has a nonzero determinant and that $\alpha(\epsilon,\gamma)$ in (C.1) is well defined. Because $\forall\, i,\ u_i(\alpha(\epsilon,\gamma))=\frac{1-\epsilon}{2}\cdot(\pi-2)$, the partners are indifferent (ex-ante) between all γ once ϵ is fixed. Let

$$\xi(\epsilon)\equiv\Big\{(\gamma,\alpha(\epsilon,\gamma))|\,\alpha(\epsilon,\gamma)=B(\epsilon,\gamma)^{-1}\cdot\beta(\epsilon,\gamma)\Big\}$$

be the set of possible ϵ-equilibrium Nash mechanisms, i.e., pairs (p,α), which lead to Nash utility payoffs.

At the first step of the negotiation process, the partners will bargain on ϵ since they are indifferent between all elements of $\xi(\epsilon)$. I suppose that if the partners agree on an ϵ, they randomly select an element of $\xi(\epsilon)$ (with the interpretation that they sign the corresponding contract). Without loss of generality, suppose that the mechanism is implemented, and that the partners take their actions, immediately after the contract is signed.

Suppose that the partners settle on a given ϵ. Since before taking their actions, the expected utility of each partner is $\frac{1-\epsilon}{2}\cdot(\pi-2)$, both partners have an incentive to tear the ϵ-contract up and to agree on an ϵ'-contract in which $\epsilon'<\epsilon$. Clearly, this process might continue forever. Thus, if the partners do not discount future revenues (i.e., if the negotiation is costless) then there is no possibility to have an equilibrium at the negotiation phase.

I suppose that the negotiation goes as follows. A partner proposes an ϵ, and the other partner can agree or disagree. If the other partner disagrees, he can propose a new ϵ, while if he agrees, a contract in $\xi(\epsilon)$ is signed and implemented. I make the following assumptions.[15]

C1. Time is perceived as a discrete variable and $\delta<1$ is the discount rate per unit of time.

C2. $\exists x: \mathbb{R}_+ \to \mathbb{R}_+$, where $\forall T$, $x(T)=\text{argmin}\,\{x|\, x$ can be announced in a period of length $T\}$. x is decreasing and satisfies $\lim_{T\to\infty} x(T)=0$.

C3. Politeness prevails: one does not interrupt a person who is talking.

C1 is made only for convenience. The analysis can be made when time is perceived as a continuous variable. C2 illustrates the limitations put on the players by the language that they are using.[16] C3 simplifies the analysis and is natural in the present context since the interests of the two parties coincide.

The problem for the partner who starts to speak first (his identity does not matter, the partners will be willing to decide at random) is to find T which maximizes the function

$$f(T)=\delta^{T-1}\cdot\frac{1-x(T)}{2}\cdot(\pi-2).$$

By assumption, as $T\to\infty$, $f(T)\to 0$. Consequently, there exists $T<\infty$ which maximizes f (use the monotonicity of x). The other partner will accept this proposal since there is coincidence of interests.

NOTES

1 Current theories of the firm find the classical works of Coase (1937), Alchian and Demsetz (1972), Jensen and Meckling (1976), and Williamson (1975). The reader compiled in Putterman (1986) and the survey of Holmström and Tirole (1987) are good references for an understanding of the state of the art in the literature. The existing theoretical work transforms the partnership problem in an agency problem, i.e., a game between a principal and many agents. Work in this area looks at partnerships from an informational point of view: the question is to define an information system which is Pareto superior in terms of risk sharing and incentives. Given an information structure, the manager is able to observe some signal which is, or is not, correlated with the workers' actions and that he might use as a basis for his reward scheme. It is necessary and sufficient that the signal is a "sufficient statistics" of the workers' actions in order to lead to a Pareto improvement. Holmström (1982) is the classic example of this approach in a multi-agent framework. Holmström (1979), Shavell (1979) and Gjesdal (1982) present the same qualitative results in the one-agent situation.

2 I wish to thank Kim Border and Preston McAfee for a discussion which clarified these issues.

3 The manager's action is to pay wages depending on the output; this action is binding by the contract and is not a strategy in the game theoretical sense. A large part of the existing literature on the principal-agent allows the principal to define random reward schemes (which makes sense only if the output is uncertain, if the principal is risk-neutral and the agent is risk-averse). Once the agent has taken his action, a random scheme is not (ex-post) optimal for the principal. Thus, the agent will credibly believe that the principal will use the random scheme only if a random device is built and if the contract stipulates

that the agent's reward is function of the observable realization of the random device. The same is true with multiple agents. In that model the agents move before the random device. With the model of this chapter, the agents move after the random device.

4 Some care must be taken here. As shown in Aumann (1974), mixed strategies are always uncorrelated, but strategies can be uncorrelated even though they are pegged on events regarding which all players are informed. I thank Tom Palfrey for making this remark.

5 Completeness is not crucial. If A_i is not complete, enlarge A_i to include the limits of all Cauchy sequences and define the output function in such a way that $y(a)$ is very large negatively if there exists i for which a_i did not belong to the original set.

6 There are some technicalities involved in defining mixed or correlated strategies when the set of pure strategies and when the set of signals that the players can receive are not denumerable. Aumann (1964) and Milgrom and Weber (1985) offer a good discussion of the issues. I avoid these technicalities in this section since the probability measure P that is constructed in the existence theorem has a countable support ($supp\, P^{\infty} \subset A^{\infty}$ which is countable). Countable support means here that only countably many signals can be received with positive probability by each agent.

7 Following Aumann (1974:75), independence means that for any player i, and for any vector of actions a, $P(a) = P(a_1) \cdot \cdots \cdot P(a_n)$.

8 It seems that every paper on partnerships starts with this example (Groves (1973), Holmström (1982), Radner (1986)). My excuse for following this trend is that I do not intend to show that partnerships are inefficient in this example; rather, I want to argue that the reason that they are inefficient gives intuition

for why some partnerships can be efficient in other environments.

9 I distinguish between the set of actions for a player, which is the list of the actions that he can take, from the space in which the actions lie. For instance, $\{(0,\sqrt{2}),(1,1)\}$ can be the set of actions, while $\mathbb{R}^2$ is the underlying space.

10 It is always possible to change the indices in such a way that assumption H4 holds with a weak inequality. I want to show that there is no loss of generality in considering a strict inequality. Suppose that there exists a partner i and actions $a_i(1),\ldots,\ a_i(t)$ for which $\forall j\in\{1,\ldots,t\},\ y(a^*_{-i}|a_i(j))=\hat{y}$. For any sharing rule s, the share that partner i will receive by deviating unilaterally to any action $a_i(j)$ is the same since the output is the same for any deviation. If s is a solution to the partnership problem, then,

(a) $\forall j\in\{1,\ldots,t\},\ s_i(\hat{y})-v_i(a_i(j))\leq s_i(y(a^*))-v_i(a^*_i)$.

Clearly, the left hand side is maximum for $v_i(a_i(j))$ minimum. Consequently, in order to verify that s is a solution, it is enough to verify that (a) holds for one of the actions which minimize $v_i(a_i(j))$, let say $a_i(1)$. Since $\forall j,\ a_i(j)\neq a^*_i$, efficiency is not affected if one eliminates the other actions $a_i(j)$, where $j\in\{2,\ldots,t\}$. This process can be repeated for each player. By successive eliminations, one obtains a reduced partnership problem which satisfies H4. It is now clear that there exists a solution to the initial partnership problem if and only if there exists a solution to the reduced partnership problem.

11 H4 must be slightly modified. Write A_i as $A_i\equiv\{a_i(\alpha);\ \alpha\in I_i\}$, where I_i is an index set. H4 reads now: for any i, there exists I_i such that $\alpha>\beta\Rightarrow y(a^*\backslash a_i(\alpha))>y(a^*\backslash a_i(\beta))$. The statement is always true with a weak inequality (from the Well Ordering Theorem). The fact that the statement (with a strict inequality) does not entail any loss of generality follows the same

arguments as in the previous note.

12 I thank John Ledyard for raising this question. Note that the problem of limited liability also arises in the principal-agents literature. For instance, if the workers have limited initial endowments, Holmström (1982) must suppose that the manager has infinite resources in order to implement the optimal mechanism.

13 In corporations, stockholders have a different sort of limited liability. They are liable up to the amount invested in the firm. A contrario, in partnerships, the partners are liable up to their total wealth. What is important for the discussion in the text is that the partners have finite wealth.

14 I abuse the notation here. To be exact, I should subscript each norm with r since the norm $\|s^r\|$ applies to the space $\mathbb{R}^{Y^r}$, where Y^r has finite cardinality and is in general different when r varies. Similarly, since the vectors $B(i)(a_i \to \tilde{a}_i)$ and $B^r(i)(a_i \to \tilde{a}_i)$ lie in different spaces (L for the first and $\left(\mathbb{R}^{n-1}\right)^{Y^r}$ for the second), the norms are different operators. I do not distinguish the norms in the text in order to minimize notation. I hope that this does not create any confusion.

15 I thank Richard McKelvey for a helpful discussion on this topic.

16 The brain might work at the speed of light, the voice or the hand do not...

REFERENCES

Alchian, Armen A. and Demsetz, Harold (1972). "Production, Information Costs and Economic Organization," *American Economic Review*, 62, 777-85.

Aumann, Robert, J (1964). "Mixed and Behavior Strategies in Infinite Extensive Games," in *Advances in Game Theory*, Annals of Mathematics Studies, No.52, pp.627-50, Princeton: Princeton University Press.

Aumann, Robert, J (1974). "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67-96.

Aumann, Robert J. (1987). "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55(1), 1-18.

Baker, George P., Jensen, Michael C. and Murphy, Kevin, J. (1988). "Compensation and Incentives: Practice vs. Theory," *The Journal of Finance*, 43(3), 593-616.

Beckman, Martin J. (1984). "Hierarchy vs. Partnership," *Journal of Economic Behavior and Organization*, 5, 237-45.

Billingsley, Patrick (1968). *Convergence of Probability Measures.* New York: John Wiley and Sons.

Brandenburger, Adam and Dekel, Eddie (1987). "Rationalizability and Correlated Equilibria," *Econometrica*, 55(6), 1391-402.

Coase, Ronald H (1937). "The Nature of the Firm," *Economica*, IV, 386-405.

Crawford, Vincent F (1982). "A Theory of Disagreement in Bargaining," *Econometrica*, 50(3), 607-37.

Fan, Ky (1956). "On Systems of Linear Inequalities," in *Linear Inequalities and Related Systems*, H.W. Kuhn and A.W. Tucker, eds., Princeton: Princeton University Press, pp. 99-156.

Forges, Françoise (1986). "An Approach to Communication Equilibria," *Econometrica*, 54(6), 1375-85.

Gjesdal, Frøystein (1982). "Information and Incentives: The Agency Information Problem," *Review of Economic Studies*, 49, 373-90.

Groves, Theodore (1973). "Incentive in Teams," *Econometrica*, 41(4), 617-31.

Hayek, Friedrich (1945). "The Use of Knowledge in Society," *American Economic Review*, 35(4), 519-30.

Holmström, Bengt (1979). "Moral Hazard and Observability," *The Bell Journal of Economics*, 10(1), 74-91.

Holmström, Bengt (1982). "Moral Hazard in Teams," *The Bell Journal of Economics*, 13, 324-40.

Holmström, Bengt and Tirole, Jean (1987). "The Theory of the Firm." Manuscript. Forthcoming in the Handbook of Industrial Organization, R. Schmalensee and R. Willig, eds..

Jensen, Michael C. and Meckling, William H (1976). "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3, 305-60.

Mertens, Jean-Francois and Zamir, Samuel (1986). "Formulation of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory*, 14, 1-29.

Milgrom, Paul R. and Weber, Robert J. (1985). "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operations Research*, 10(4), 619-32.

Moulin, H. and Vial, J.-P (1978). "Strategically Zero Sum Games: The Class of Games Whose Completely Mixed Equilibria Cannot be Improved Upon," *International Journal of Game Theory*, 7 (3/4), 1978, 201-21.

Myerson, Roger B (1983). "Bayesian Equilibrium and Incentive-Compatibility: an

Introduction." Discussion Paper No. 548, Center for Mathematical Studies in Economics and Management Science, Northwestern University.

Myerson, Roger B. (1986). "Multistage Games with Communication," *Econometrica*, 54(2), 67-81.

Nash, John F (1950). "The Bargaining Problem," *Econometrica*, 18, 155-62.

Nash, John F (1953). "Two-Person Cooperative Games," *Econometrica*, 21, 128-40.

Putterman, Louis (1986). *The Economic Nature of the Firm.* Cambridge: Cambridge University Press.

Radner, Roy (1986). "Repeated Partnership Games with Imperfect Monitoring and No Discounting," *Review of Economic Studies*, 53, 43-57.

Radner, Roy, Myerson, Roger and Maskin, Eric (1986). "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria," *Review of Economic Studies*, 53, 59-69.

Radner, Roy and Williams, Steven (1988). "Efficiency in Partnership when the Joint Output is Uncertain." Discussion Pqaper No. 760, Center for Mathematical Studies in Economics and Management Science, Northwestern University.

Rosenthal, Robert W. (1974). "Correlated Equilibria in Some Classes of Two-Person Games," *International Journal of Game Theory*, 3(3), 119-28.

Rubinstein, Ariel (1982). "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97-109.

Schelling, Thomas C. (1963). *The Strategy of Conflict.* London and New York: Oxford University Press.

Shavell, Steven (1979). "Risk Sharing and Incentives in the Principal and Agent Relationship," *The Bell Journal of Economics*, 10(1), 55-73.

Tan, Tommy Chin-Chiu and Werlang, Sérgio Ribeiro da Costa (1988). "The Bayesian Foundations of Solution Concepts of Games," *Journal of Economic Theory*, 45(2), 370-91.

CHAPTER 3

STRONGLY DURABLE MECHANISMS

1. INTRODUCTION

In economic models institutions are generally taken as parameters of the environment. By institution, I mean any set of rules (legislative, cultural, contract laws, etc.) which govern the behavior of the agents. The mechanism design literature addresses the question of institution creation when a unique entity (called government, central planner, principal, etc.) has the power to define the institution which will serve best his or her objectives.

Little work exists when the institution must be chosen by a group of individuals with conflicting interests. Allowing more than one agent to have some decision power introduces some interesting but difficult questions. This paper considers a situation in which the agents have to make a decision in the future and in which there exists a "default" decision process which will be implemented if the players decide not to modify it. The decision to change the decision process is typically due to a better information of the agents about their environment or about their type. I will study the "default" decision processes which are stable against such changes.

It is now well-understood that there is in general a tradeoff between incentive-compatibility and efficiency. Many authors have argued that when inefficiency of an incentive compatible direct mechanism is common knowledge, it is not reasonable to assume that the players will stick to the recommendations of the initial mechanism instead of designing a new mechanism for which all of the agents will be better off. (See Holmström-Myerson (1983), Cramton (1985).) However, even if efficiency is common knowledge, the mere fact that agents refuse to renegotiate can leak information. A special case is when the agents have to vote between the status-quo and an alternative mechanism. The outcome of the vote can leak information about the types of the agents.

The first paper to take into account the possibility of redesigning mechanisms is Holmström-Myerson (1983). They introduce a notion of stability called durability. A mechanism is durable if for any alternative mechanism, there is an equilibrium in which the status-quo is voted with probability one. This notion of stability can be criticized on two basis. First, a mechanism can be durable even if there is an equilibrium of the voting game in which all the types of all the players are better off than in the original mechanism. Second, Holmström and Myerson suppose that the alternative mechanism is proposed by an outside agent (e.g., a mediator). In a truly decentralized environment, the agents should be able to make proposals. (This is also true when there is a mediator: an agent could stand up and make a proposal if he or she does not like the proposal of the mediator.) The present paper is devoted to the analysis of a new concept of stability when the proposals are made by a mediator. Consequently, it does not escape the criticism that the agents should be able to make proposals and counterproposals.

I define a new concept of stability which I call *strong durability.* A mechanism *strongly endures* another mechanism if there exists a Bayesian Nash equilibrium of this mechanism such that for any alternative mechanism and for any sequential equilibrium of the voting game between the two mechanisms, the interim utility payoffs of *all* the players in the sequential equilibrium are less than or equal to those obtained in the Bayesian equilibrium in the initial mechanism. I will say that a mechanism HM-endures or strongly endures another one to refer to one or the other definition of durability.

The following example shows that an incentive compatible direct mechanism can be HM-durable while being interim dominated by another mechanism.

<u>Example 1</u> (Holmström-Myerson (1983:1817).) This is a coordination game. There are two players 1 and 2 with two possible types a and b. Players have common prior beliefs and types have the same probability of occurrence. There are two possible decisions A and B. The payoffs to the players are as in the table below.[1]

		$2a$	$2b$
$1a$	A	2 2	2 2
	B	3 3	0 0
$1b$	A	2 2	2 2
	B	0 0	3 3

Consider the two mechanisms $\mu^0=\begin{bmatrix} A & A \\ A & A \end{bmatrix}$, $\mu^1=\begin{bmatrix} B & A \\ A & B \end{bmatrix}$. μ^0 HM-endures μ^1 but μ^0 does not strongly endure μ^1.

To see this, let each type of each player vote for μ^1 with probability one. Let each type be truthful in each mechanism. This form a sequential equilibrium once we define the beliefs of the players when μ^0 is played by $q_i(t_{-i}|t_i,\mu^0)=0.5$. Here $q_i(t_{-i}|t_i,\mu^0)$ denote the interim beliefs of type t_i if μ^0 is played. The interim utility payoffs of each type of each player are 2.5 in this sequential equilibrium while they are equal to 2 if μ^0 is played. Thus, μ^0 does not strongly endure μ^1.

To show that μ^0 HM-endures μ^1, let each type of each player vetoes μ^1 with probability one. Let each type be truthful in μ^0 and let each type play a mixed strategy $[0.5,ia\oplus 0.5,ib]$ in μ^1. (I.e., type t_i of player i sends message ia with probability 0.5 and sends message ib with probability 0.5.) This forms an equilibrium rejection of μ^1 if the out of the equilibrium beliefs (i.e., when μ^1 is played) are such that each player believes that there is an equal probability of facing one of the two types of the other player. •

Holmström and Myerson argue that because the voting game is played noncooperatively, "*Individuals cannot be forced to communicate effectively in a noncooperative game with incomplete information.*" This argument is correct but is a weak one. In fact, if one considers a situation in which the players choose the status-quo *before* knowing their own type, it is difficult to defend the idea that they will choose an interim dominated mechanism.

Observe that μ^1 is HM-durable. Thus, ex-ante, the players would be indifferent between two mechanisms which are HM-durable, i.e., which are compatible with the notion of stability that is used, while each type of each player has a greater interim expected utility with one mechanism than with the other.[2] We will see later that two strongly durable incentive compatible direct mechanisms cannot be Pareto ranked at the interim stage. For this reason, strong durability is more attractive a concept than durability.

In Example 1, the status-quo is an incentive compatible mechanism. Without further explanations, such a choice is a restriction. The revelation principle is of no use here as long as there are multiple equilibria in the status-quo mechanism. (The question of the uniqueness of equilibria in direct mechanisms is linked to the problem of full implementation. See Palfrey and Srivastava (1987), (1989) and Jackson (1988) and the references therein for a recent discussion.) In the model of Holmström and Myerson (1983), the problem is assumed away since the authors *force* the players to report truthfully when the status-quo is played. Here, it is necessary to provide a formal proof that there is no loss of generality in restricting attention to incentive compatible mechanisms.

If there are multiple equilibria, a definition of stability for a mechanism should incorporate not only the idea that the players will not renegotiate to *another contract*, but also the idea that they will not renegotiate to *another equilibrium.*[3] The concept of strongly durable mechanism satisfies these two requirements. In particular, if a mechanism is strongly durable, then there exists an equilibrium of this mechanism such that there does not exist another equilibrium for which one type is strictly better off (Lemma 1). In fact, under a regularity condition, *all* equilibria of a strongly durable mechanism must be either interim payoff equivalent or if they are not, then at most one type has a different interim payoff (Theorem 3). In particular, if the status-quo is an incentive compatible direct mechanism and is strongly durable, then the truthful equilibrium becomes a "focal" equilibrium (Schelling (1963)).

The main results of this paper are the characterization of the strongly durable mechanisms and a sufficient condition for their existence. I prove (Lemma 2) that there is no loss of generality in restricting attention to incentive compatible status-quo. This can be considered as an "extended revelation principle" when renegotiation takes place through a mediator. Strongly durable incentive compatible direct mechanisms are always HM-durable whenever each profile of types has a positive probability of occurring. Furthermore, an incentive compatible direct mechanism is strongly durable if, and only if, it is interim incentive efficient and if there does not exist an interim incentive efficient direct mechanism which is not (interim) payoff equivalent to the status-quo with the property that there is a sequential equilibrium of the voting game between this mechanism and the status-quo in which this alternative mechanism is chosen with probability one by all types. In other words, if a mechanism is *not* strongly durable, *then* there is an interim incentive efficient direct mechanism which is selected with probability one by *all types* in a sequential equilibrium of the voting game and the players will announce their true type once this alternative mechanism is implemented. I show existence of the strongly durable mechanisms when at least one player has a utility function which is independent of the types of the other players and when any decision is utility improving.

Finally, there is an interesting relationship between the question of multiplicity of equilibria and the stability of a mechanism. In particular, an incentive compatible direct mechanism is strongly durable only if the truthful equilibrium is "focal."

The rest of the paper is organized as follows. In Section 2, I present the main assumptions of the model. In Section 3, I introduce the definitions and I point out the main differences between the concepts of durability and strong durability. I show, at the end of Section 3, that there is no loss of generality in considering incentive compatible status-quo. In Section 4, I characterize the strongly durable incentive compatible direct mechanisms. I propose in Section 5 a geometric characterization of the strongly durable mechanisms and an existence theorem when at least one player has a utility function that is a function of her own type only. Some proofs are relegated to Appendix A. I present some final comments in Section 6.

2. THE MODEL

The model is one of Bayesian collective decision making (see Myerson (1983) for more details). There is a set N of n agents. For each agent i, there exists a finite set of types T_i. The set of decisions D is finite with generic element d. The prior beliefs of partner i are described by the probability measure p_i over the set $T_1 \times \cdots \times T_n$, where $\times$ denotes the Cartesian product. An *admissible mechanism* is a pair $\mu = (\{M_i\}_{i \in N}, g)$ where g: $M_1 \times \ldots \times M_n \rightarrow \Delta(D)$ is measurable, $\Delta(D)$ is the set of probability measures over the set D and M_i is the, finite, set of messages that player i can send in the mechanism.[4] A mechanism is *direct* if $M_i = T_i$, for all player i. Each player i has a bounded and measurable utility function and $u_i(d,t)$ denotes the utility of player i for the decision d when the profile of types is t.

I will denote by Ξ the set of admissible mechanisms. Ξ^{DI} denotes the set of direct mechanisms. $IC \subset \Xi^{DI}$ is the set of *incentive compatible* direct mechanisms. E_A and E_I denote, respectively, the sets of ex-ante and interim incentive efficient mechanisms. I will restrict myself to the status-quo mechanisms which satisfy ex-ante individual rationality, i.e., such that the ex-ante expected utility of each player exceeds his reservation value (taken to be zero for simplification). This is a natural condition to impose when the partners are free *not to* sign an ex-ante contract.

I will suppose that the players must agree on a mechanism *before* they learn their type. In my partnership model, this assumption follows the fact that the players will learn their type after the production phase and that a contract stipulating the sharing rule and the decision process must be signed before production begins. The initial mechanism can be considered as a "standard" contract which is applied as long as the players do not decide to change it.

The timing is the following. Initially, before the private information is revealed, the partners sign a contract μ^0. Then, the types are privately revealed to the players. There is then a possibility of renegotiation through a voting procedure in which the alternative mechanism is proposed by a mediator.

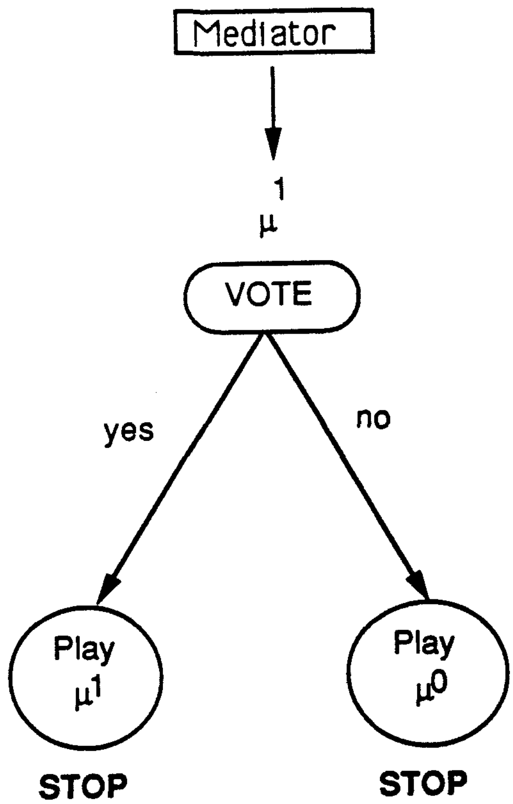

Figure 3.1

Figure 3.1 corresponds to the situation analyzed by Holmström and Myerson (1983) and to the environment of this paper. A mediator proposes an alternative mechanism to the status-quo and after the vote, the players play the winning mechanism and the game stops.

3. DEFINITIONS AND INTERMEDIATE RESULTS

3.1. *DEFINITIONS AND NOTATION*

Let $\mu^0=(M,g^0)\in\Xi$ be the status-quo and let $\mu^1=(S,g^1)$ be an alternative mechanism, where the sets M_i and S_i are finite. Here I suppose that the status-quo can be any mechanism, direct or indirect. The use of a status-quo which is an incentive compatible direct mechanism is, without a formal argument, a restriction. I will show in Lemma 2 that there is in fact no loss of generality in considering incentive compatible direct mechanisms. This will enable me to pursue the analysis by using only incentive compatible direct mechanisms as status-quo.

The players play the following voting game. The mechanism μ^1 is proposed by a mediator. Each player votes for or against this new mechanism (mixed strategies are allowed). Then the mediator announces which mechanism has to be played: μ^1 if there is unanimity and μ^0 if there is at least one player who cast a negative vote for μ^1. The players then choose a strategy, i.e., send messages in M if μ^0 is played or messages in S if μ^1 is played, and the final decision is implemented.

A *sequential equilibrium* of this voting game is an n-tuple of triples $((r_i,\sigma_i,q_i);$ $i\in N)$, where r_i is the voting strategy of player i, σ_i is his play strategy and q_i are his beliefs at his different information sets. I will denote by $\Phi(\mu^0,\mu^1)$ the set of sequential equilibria of the voting game when μ^0 is the status-quo and μ^1 is the alternative mechanism.

The extensive form of the game is sensitive to the hypothesis on how the players vote, secretly or publicly. If votes are secret, then, when the status-quo is played, player i of type t_i will have, in general, different beliefs about the types of the other players depending on whether t_i voted for or against μ^1. It he voted for μ^1, then his beliefs will be a function of the voting strategies of the other players. If he voted against μ^1, his posterior beliefs will always coincide with his prior beliefs (because once μ^1 is vetoed by one player μ^0 is played). If the votes are public, the beliefs of a player depend only on the observation of the votes of the other players, not on his own vote. Following Holmström and Myerson (1983), I will suppose that the votes are secret. Figures 3.2 and 3.3 give a representation of the information sets if there are only two players, if player 1 has only one type and if player two has two types. In the public voting case (Figure 3.3) one of the information sets of player 1 in the extensive form when the votes are secret (Figure 3.2, see the information set $(0,\mu^0)$) has been "split" in two (more information is always available in the public voting case).

There are four types of information sets for type t_i of player i. (See Figure 3.2)) First, when t_i must vote. Second, when t_i must play the game μ^1. Third, when t_i must play the game μ^0 while he voted against μ^1. Fourth, when t_i must play μ^0 while he voted for μ^1. Let h denote one of the three last types of information sets. To simplify, I will write $h=\mu^1$ if μ^1 must be played, $h=(1,\mu^0)$ if μ^0 must be played while the player voted for μ^1, and $h=(0,\mu^0)$ if μ^0 must be played and the player voted against μ^1. Let $H=\{\mu^1,(1,\mu^0),(0,\mu^0)\}$.

I use the following notation. If X is a set, then $\Delta(X)$ denotes the set of probability measures over the set X. If $M=M_1\times\cdots\times M_n$, then M_{-i} denotes the set $\times_{j\neq i}M_j$.

The behavioral strategies are defined as follows.[5]

$$\forall i,\ r_i\colon\ T_i\to[0,1]$$

$$\forall i,\ \sigma_i\colon\ H\times T_i\to\Delta(N_i)$$

where $N_i=M_i$ if $h\in H\setminus\{\mu^1\}$ and $N_i=S_i$ otherwise.

Thus, $\sigma_i(n_i|t_i,h)$ is the probability that player i of type t_i will send the message n_i at his information set h. A belief system for player i is defined by,

$$\forall i, \quad q_i\colon H \times T_i \to \Delta(T_{-i}).$$

Thus, $q_i(t_{-i}|t_i,h)$ are the beliefs of player i of type t_i about the types of the other players when his information set h is reached.

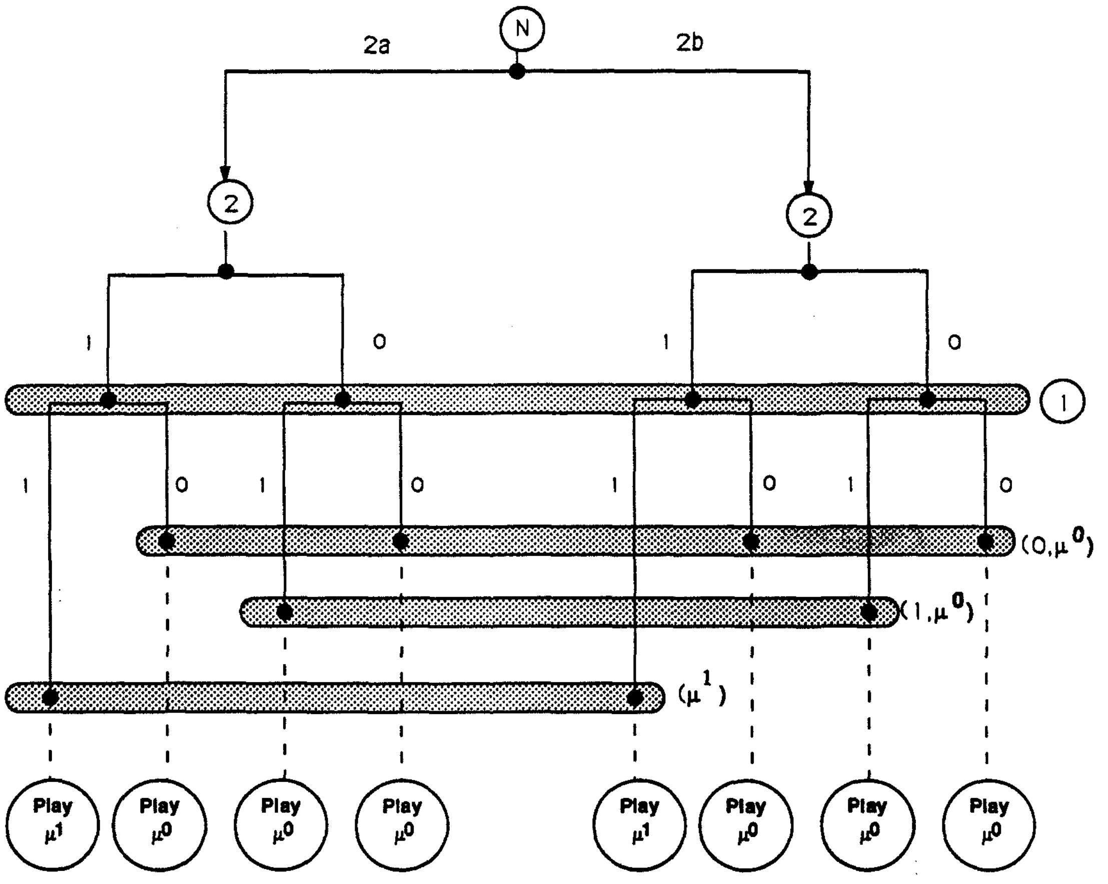

Figure 3.2

Extensive form when the votes are secret

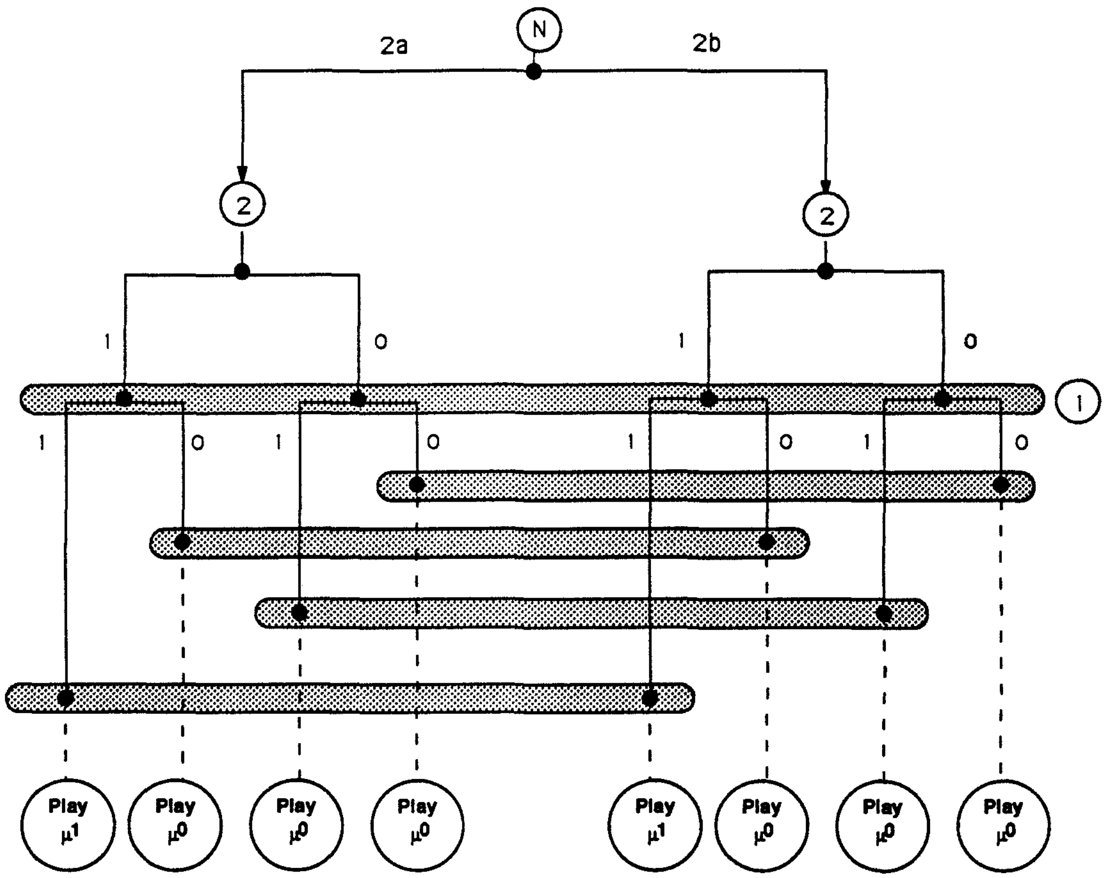

Figure 3.3

Extensive form when the votes are public

3.2 *EQUILIBRIA OF THE VOTING GAME*

The concept of sequential equilibrium has been defined in Kreps and Wilson (1982) and is a weakening of the concept of perfect equilibrium of Selten (1976). A sequential equilibrium of the voting game between μ^0 and μ^1 is a triple $(r,\sigma,q)=\{(r_i,\sigma_i,q_i);\, i\in N\}\in\Phi(\mu^0,\mu^1)$ which satisfies the following four conditions.[6]

(S1) (Best response in μ^1)

$\forall i,\ \forall t_i,$

$$\sigma_i(\cdot|t_i,\mu^1)\in \underset{\tilde\sigma_i\in\Delta(S_i)}{Argmax} \sum_d \sum_{t_{-i}} \sum_s q_i(t_{-i}|t_i,\mu^1)\cdot\sigma_{-i}(s_{-i}|t_{-i},\mu^1).\tilde\sigma_i(s_i)\cdot g^1(d|s)\cdot u_i(d,t).$$

(S2) (Best response in μ^0)

$\forall i,\ \forall t_i,\ \forall h\in H\backslash\{\mu^1\},$

$$\sigma_i(\cdot|t_i,h)\in \underset{\tilde\sigma_i\in\Delta(M_i)}{Argmax} \sum_d \sum_{t_{-i}} q_i(t_{-i}|t_i,h)\cdot\sum_m K(\tilde\sigma_i)(m)\cdot g^0(d|m)\cdot u_i(d,t)$$

where,

$$K(\tilde\sigma_i)(m)\equiv\prod_{j\neq i}\Big[r_j(t_j)\cdot\sigma_j(m_j|t_j,(1,\mu^0))+(1-r_j(t_j))\cdot\sigma_j(m_j|t_j,(0,\mu^0))\Big]\cdot\tilde\sigma_i(m_i)\cdot$$

(S3) (Best response at the voting stage)

Let,

$$V_i(t_i|0)\equiv\sum_d \sum_{t_{-i}} p_i(t_{-i}|t_i)\cdot\sum_m K\Big(\sigma_i((0,\mu^0),t_i)\Big)(m)\cdot g^0(d|m)\cdot u_i(d,t)$$

$$V_i(t_i|1)\equiv\sum_d \sum_{t_{-i}} p_i(t_{-i}|t_i)\cdot\Big\{r_{-i}(t_{-i})\cdot\sum_s \sigma(s|t,\mu^1)\cdot g^1(d|s)$$

$$+\Big[1-r_{-i}(t_{-i})\Big]\cdot\sum_m K\Big(\sigma_i(\cdot|t_i,(1,\mu^0))\Big)(m)\cdot g^0(d|m)\Big\}\cdot u_i(d,t).$$

where $K(\cdot)$ has been defined in S(2). $V_i(t_i|0)$ and $V_i(t_i|1)$ are, respectively, the expected utility levels of type t_i of player i if he votes against or for μ^1 at the voting stage. In a sequential equilibrium, t_i will vote for the alternative which leads to the greatest level of utility, i.e.,

$$\forall i,\ \forall t_i,\ r_i(t_i) \begin{cases} =1 \\ =0 \\ \in(0,1) \end{cases} \text{as } V_i(t_i|0) \begin{matrix} < \\ > \\ = \end{matrix} V_i(t_i|1).$$

(S4) (Consistent belief structure)

$\forall i,\ \forall t_i,$

(S41) $$q_i(t_{-i}|t_i,\mu^1)=\frac{p_i(t_{-i}|t_i)\cdot r_{-i}(t_{-i})}{\sum_{\hat{t}_{-i}} p_i(\hat{t}_{-i}|t_i)\cdot r_{-i}(\hat{t}_{-i})}.$$

(S42) $$q_i(t_{-i}|t_i,(1,\mu^0))=\frac{p_i(t_{-i}|t_i)\cdot(1-r_{-i}(t_{-i}))}{\sum_{\hat{t}_{-i}} p_i(\hat{t}_{-i}|t_i)\cdot(1-r_{-i}(\hat{t}_{-i}))}.$$

(S43) $$q_i(t_{-i}|t_i,(0,\mu^0))=p_i(t_{-i}|t_i).$$

The expressions in (S41) and (S42) are well defined only if their denominators are nonzero. If this is not the case, let $\{r^k\}$ be a sequence of strategies with the property that $\forall i,\ \forall t_i,\ r_i^k(t_i)\in[0,1]$ and $r_i^k(t_i)\rightarrow r_i(t_i)$. (It follows that all the players form the same conjecture on events of probability zero.) Define the beliefs in (S41) and (S42) as the limit as $k\uparrow\infty$ when the right hand side is evaluated using the sequence $\{r^k\}$.

3.3. *HM-DURABLE MECHANISMS*

Holmström and Myerson (1983) consider only status-quo μ^0 which are incentive compatible direct mechanisms, i.e., $\mu^0\in IC$. I will denote by HMD the set of HM-durable mechanisms. A mechanism μ^0 is HM-durable, i.e., $\mu^0\in HMD$, if for any alternative mechanism μ^1, there exists a sequential equilibrium

$\{(r_i,\sigma_i,q_i);i\in N\}\in\Phi(\mu^0,\mu^1)$ with the following properties.

(HMD1) **(Equilibrium rejection)**

$\forall t,\ r(t)=0.$

(HMD2) **(Truthful reporting in μ^0)**

$\forall i,\ \forall t_i,\ \forall h\in H\backslash\{\mu^1\},\ \sigma_i(t_i|t_i,h)=1.$

(HMD3) **("Strong best response")**

$\forall i,\ \forall t_i$, if

$$\sum_{t_{-i}} q_i(t_{-i}|t_i,\mu^1)\cdot\sum_{s}\sigma(s|t,\mu^1)\cdot g^1(d|s)\cdot u_i(d,t) > \sum_{t_{-i}} q_i(t_{-i}|t_i,\mu^1)\cdot g^0(d|t)\cdot u_i(d,t)$$

then, $r_i(t_i)=1$.

(HMD2) requires that the players tell the truth in μ^0, given the beliefs that they have when μ^0 is played. (Recall that μ^0 is incentive compatible when the beliefs of the players are given by $p_i(t_{-i}|t_i)$; μ^0 is not necessarily incentive compatible for other interim beliefs.)

Remark 1: The condition (HMD3) *is not* part of the definition of a sequential equilibrium, nor of the definition of a perfect equilibrium. Note that (HMD3) is always part of the definition of a sequential equilibrium *only if the votes are public* and if (HMD2) is satisfied. In this case, the beliefs of t_i are independent of his own vote when μ^0 has to be played. When player t_i considers voting for or against μ^1, he will only consider the two expressions in (HMD3). (All the other terms in (S3) cancel out.) Because of (HMD2), the players are always truthful if μ^0 is played, the expression in (HMD3) follows. Holmström and Myerson (1983) introduce this condition in order to eliminate the trivial equilibrium rejection in which all the types vote against μ^1 with probability one. •

3.4. *STRONGLY DURABLE MECHANISMS*

A mechanism HM-endures another mechanism whenever *there exists one* sequential equiliibrium in which the alternative mechanism is rejected with probability one. We observed in Example 1 that the existence of such an equilibrium rejection does not preclude the existence of another sequential equilibrium in which the alternative mechanism is selected *with probability one* and in which the interim payoffs of all the types are larger than in the equilibrium rejection. The concept of durability ignores the other sequential equilibria of the voting game and presupposes that the players will automatically select the equilibrium rejection. Without a formal theory of selection of equilibria, there is no reason to suppose that the players will choose, as in Example 1, a sequential equilibrium in which all of them are worse off than in another sequential equilibrium. By taking into account all the possible sequential equilibria of the voting game, one obtains a more natural concept of stability for mechanisms.

To prove that a mechanism HM-endures another mechanism, it is enough to find *one* sequential equilibrium satisfying the conditions (HMD1)-(HMD3). For strong durability, it is necessary to verify that *no* player can be made better off in *any* sequential equilibrium of the voting game. This difference between the two concepts has important consequences. In particular, strong durability takes into account the possibility a mechanism might have multiple equilibria and that some "undesirable" equilibria can be used in a sequential equilibria in order to destabilize the status-quo. Lemma 1 below shows if a mechanism is strongly durable, then any "undesirable" equilibrium is interim Pareto dominated by the "desirable" equilibrium. Theorem 1 of Section 4 will establish that strongly durable mechanisms are HM-durable and induce interim efficient payoffs whenever each profile of types has a nonzero probability of occurrence.

The concept of strong durability incorporates the idea that the players will not renegotiate to another mechanism *or* to another equilibrium. A mechanism together with a Bayesian Nash equilibrium is strongly durable if for any alternative mechanism there is no sequential equilibrium of the voting game between these two mechanisms for which the interim utility payoff, in the sequential equilibrium, of one type is strictly larger than if the initial equilibrium of the status-quo mechanism had been played (without voting). The following two examples will give some intuition for the necessity of requiring that one type is made better off. The formal definition of a strongly durable mechanism will follow these examples.

<u>Example 2</u>. Consider the following environment in which the types of each player are equally probable. (The notation is the same as in Example 1.)

A
B
C

	2a		2b	
1a	2	0	2	−2
	1	2	1	0
	0	1	0	1
1b	0	0	0	−2
	4	2	4	0
	9	1	9	1

$\mu^0=\begin{bmatrix} B & B \\ C & C \end{bmatrix}$, $\mu^1=\begin{bmatrix} B & C \\ C & C \end{bmatrix}$. Both μ^0 and μ^1 are incentive compatible and ex-ante efficient. Consider the following strategies and beliefs. ("&" is the logician symbol for "and".)

$r_1(1a)=0,\ r_1(1b)=1,\ r_2(2a)=0,\ r_2(2b)=1.$

$\forall i,\ \forall t_i,\ \forall h,$

$\sigma_i(t_i|t_i,h)=1,$

$\&\ q_1(2b|t_1,\mu^1)=1\ \&\ q_1(2a|t_1,(1,\mu^0))=1\ \&\ q_1(2a|t_1,(0,\mu^0))=0.5,$

& $q_2(1b|t_2,\mu^1)=1$ & $q_2(1a|t_2,(1,\mu^0))=1$ & $q_2(1a|t_2,(0,\mu^0))=0.5$.

It can be verified that $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$. Moreover, in this equilibrium, the interim expected utility payoffs coincide with those obtained when μ^0 is played (without vote). Note that in this equilibrium, the resulting mechanism is $\mu=\mu^0=\begin{bmatrix} B & B \\ C & C \end{bmatrix}$. Thus, the players are in fact indifferent between playing μ^0 immediately versus voting first between μ^0 and μ^1 and then playing the mechanism which has been voted. •

Example 3:

A
B
C

	$2a$		$2b$	
$1a$	2	0	2	−2
	1	2	1	0
	0	1	0	1
$1b$	0	0	0	−2
	4	2	4	0
	9	1	9	1

Consider the two mechanisms $\mu^0=\begin{bmatrix} B & B \\ B & C \end{bmatrix}$, $\mu^1=\begin{bmatrix} B & B \\ B & B \end{bmatrix}$. It can be verified that $\mu^0\in E_I$. Observe that truth is a (strictly) best response strategy in μ^0 even if t_i knows the type of t_{-i}. (This is what Myerson (1983) calls a "safe" mechanism.) In μ^0, there is only another equilibrium involving both players pooling at a (i.e., $t_1\in\{1a,1b\}$ announces $1a$ and $t_2\in\{2a,2b\}$ announces $2a$). Consider the following strategies,

$\forall i$, $\forall t_i$, $\forall h\in H\backslash\{\mu^1\}$

$r_i(t_i)=0.5$

& $\sigma_i(ia|t_i,h)=1$

& $\sigma_i(t_i|t_i,\mu^1)=1$.

Since each information set is reached with positive probability, the beliefs q_i are computed by Bayes' law. It is easy to check that $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$. The mechanism resulting in equilibrium coincides with μ^1. The expected interim payoffs are

$U_1(\mu^1|1a)=1=U_1(\mu^0|1a)$; $U_1(\mu^1|1b)=4<6.5=U_1(\mu^0|1b)$

$U_2(\mu^1|2a)=2=U_2(\mu^0|2a)$; $U_2(\mu^1|2b)=0<0.5=U_2(\mu^0|2b)$.

This sequential equilibrium is possible because there exists another equilibrium in μ^0 which is dominated by the truthful equilibrium. •

These examples suggest that it would be unecessarily severe to say that a mechanism strongly endures another one if there is no equilibrium of the voting game in which the alternative mechanism is selected with positive probability. A mechanism together with a Bayesian equilibrium induces certain interim payoffs. It seems reasonable to exclude situations in which the voting game leads to the same (as in Example 2) interim expected payoffs for all the types of players or to payoffs that are interim dominated by the initial payoffs (like in Example 3). This suggests the following definition. For $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$, let $\mu=\psi(r,\sigma)$ be the *resulting equivalent direct mechanism.* I.e., $\mu=(T,g)$, where

(3.1) $\forall t,\ g(d|t)\equiv r(t)\cdot\beta(d|t)+(1-r(t))\cdot\alpha(d|t;r),$

with

(3.2) $\beta(d|t)\equiv\sum_{d}\sum_{s}\sigma(s,t)\cdot g^1(d|s),$

(3.3)

$$\alpha(d|t;r)\equiv\sum_{d}\sum_{m}\prod_{i}\Big\{r_i(t_i)\cdot\sigma_i(m_i|t_i,(1,\mu^0))+(1-r_i(t_i))\cdot\sigma_i(m_i|t_i,(0,\mu^0))\Big\}\cdot g^0(d|m).^7$$

Define,

$$U_i(\psi(r,\sigma)|t_i)\equiv\sum_{t_{-i}}p_i(t_{-i}|t_i)\cdot g(d|t)\cdot u_i(d,t)$$

to be the interim payoff of type t_i in the sequential equilibrium (r,σ,q) of the voting game. If the prior beliefs are given by p, then I will denote the set of Bayesian Nash equilibria of a mechanism μ by $E(\mu,p)$.

DEFINITION 1: *Suppose that the prior beliefs are given by p. A pair* (μ^0,σ^0), *where* $\mu^0 \in \Xi$, $\sigma^0 \in E(\mu^0,p)$, *strongly endures an alternative mechanism* μ^1 *if the set of sequential equilibria* $\Phi(\mu^0,\mu^1)$ *satisfies*

$$\forall i,\ \forall t_i,\ \forall (r,\sigma,q) \in \Phi(\mu^0,\mu^1),\ U_i(\psi(r,\sigma)|t_i) \leq U_i(\mu^0,\sigma^0|t_i).$$

A mechanism $\mu^0 \in \Xi$ *is strongly durable if there exists an equilibrium* $\sigma^0 \in E(\mu^0,p)$ *for which the pair* (μ^0,σ^0) *strongly endures any alternative mechanism.*

Let $U_i(\mu,\sigma|t_i)$ denote the expected utility of type t_i if the equilibrium $\sigma \in E(\mu,p)$ is played. I.e.,

$$U_i(\mu,\sigma|t_i) \equiv \sum_d \sum_{t_{-i}} p_i(t_{-i}|t_i) \cdot \sum_m \prod_j \sigma_j(m_j|t_j) \cdot g(d|m) \cdot u_i(d,t). \tag{3.4}$$

The next result is immediate from the definition of a strongly durable mechanism.

LEMMA 1: *Suppose that* μ^0 *is strongly durable, then the following is true*

$$\exists \sigma \in E(\mu^0,p), \forall \sigma' \in E(\mu^0,p),\ \forall i,\ \forall t_i,\ U_i(\mu,\sigma^0|t_i) \geq U_i(\mu^0,\sigma'|t_i).$$

Proof:[8] Suppose not. Consider the voting game between μ^0 and μ^0. The voting strategies are irrelevant here because $\mu^1 = \mu^0$. If there exists $\sigma' \in E(\mu,p)$ violating the condition in the lemma, then there exists a sequential equilibrium of the voting game between μ^0 and μ^0 for which μ^0 does not strongly endure μ^0. □

The revelation principle tells us that if σ is a Bayesian Nash equilibrium of the mechanism $\mu=(M,g)$, then there is an incentive compatible mechanism $\hat{\mu}=(T,\hat{g})$, where $\hat{g}=g\circ\sigma$, whose truthful equilibrium leads to the same interim payoffs than the pair (μ,σ). Because an incentive compatible mechanism can have other equilibria than the truthful equilibrium, it is not immediately obvious that there is no loss of generality in using an incentive compatible status-quo. Precisely, it is necessary to show that if the pair (μ,σ) is strongly durable, then the pair $(\hat{\mu},\beta)$, where $\hat{\mu}$ is the direct mechanism induced by (μ,σ) and β is the truthful equilibrium of $\hat{\mu}$ is also strongly durable. This "extended" revelation principle is established in Lemma 2. As a consequence of lemmas 1 and 2, it follows that the incentive compatible direct mechanism which is induced by a strongly durable pair (μ,σ) has the property that its truthful equilibrium interim Pareto-dominates all the other equilibria of this mechanism.

LEMMA 2: *Suppose that the pair (μ,σ), where $\mu\in\Xi$ and $\sigma\in E(\mu,p)$, is strongly durable. Then there exists an incentive compatible direct mechanism $\hat{\mu}$ such that the pair $(\hat{\mu},\beta)$, where β is the truthful equilibrium, is strongly durable and such that the pairs (μ,σ) and $(\hat{\mu},\beta)$ induce the same interim payoffs.*

Proof: See Appendix A. □

COROLLARY 1: *Let $\mu=(S,g)\in\Xi$ and $\sigma\in E(\mu,p)$. Let μ^0 be the incentive mechanism which is induced by (μ,σ). If (μ,σ) is strongly durable then the equilibrium in which all the types report truthfully in μ^0 dominates all the other equilibria of μ^0.*

Proof: Combine Lemmas 1 and 2. □

If $\mu \in IC$ is an incentive compatible mechanism, I will write $U_i(\mu|t_i) \equiv U_i((\mu,\sigma)|t_i)$ where β is the equilibrium in which every type t_i has the degenerate strategy $\beta_i(t_i|t_i)=1$. That is, $U_i(\mu|t_i)$ is the interim payoff to type t_i of player i corresponding to the truthful equilibrium of the incentive compatible mechanism μ. By a little abuse of definition, I will say that the incentive compatible direct mechanism μ^0 is strongly durable if the pair (μ^0,β), where β is the truthful equilibrium, is strongly durable. From Lemma 2, there is no loss of generality in considering status-quo that are incentive compatible direct mechanisms. For this reason, I will suppose from now on that the status-quo is an incentive compatible mechanism and I will use the following definition of strong durability.

DEFINITION 2: *An incentive compatible direct mechanism $\mu^0 \in IC$ is strongly durable if for any alternative mechanism $\mu^1 \in \Xi$ the set of sequential equilibria $\Phi(\mu^0,\mu^1)$ satisfies*

$$\forall i,\ \forall t_i,\ \forall (r,\sigma,q) \in \Phi(\mu^0,\mu^1),\ U_i(\psi(r,\sigma)|t_i) \leq U_i(\mu^0|t_i),$$

where $\psi(r,\sigma)$ is given by (3.1).

I will denote by SD the set of incentive compatible direct mechanisms which satisfy Definition 2. From Lemma 2, if a pair (μ,σ) is strongly durable in the sense of Definition 1, there exists an incentive compatible direct mechanism $\hat{\mu}$ which is strongly durable in the sense of Definition 2 and which leads to the same interim payoffs than (μ,σ).

In the definition of $U_i(\psi(r,\sigma)|t_i)$, no claim was made about the fact that the players will indeed be truthful if they have to play the mechanism (T,g). The following lemma proves that, in fact, $\psi(r,\sigma)$ is incentive compatible. I will say that a direct mechanism μ is *incentive compatible given certain interim beliefs* if the interim expected utility of each type of each player (using the interim beliefs in question) is maximized at t_i when all the other players are expected to be truthful.

LEMMA 3: *Consider a sequential equilibrium* $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$, *where* $\mu^0\in IC$, $\mu^1\in\Xi$. *Let* $\psi(r,\sigma)=(T,g)$, *where* g *is as in* (3.1).

(i) $\psi(r,\sigma)$, (T,α) *and* (T,β) *are admissible direct mechanisms.*

(ii) $\psi(r,\sigma)\in IC$, *given the interim beliefs* $p_i(t_{-i}|t_i)$.

(iii) $(T,\beta)\in IC$, *given the interim beliefs* $q_i(t_{-i}|t_i,\mu^1)$.

(iv) $(T,\alpha)\in IC$, *given the interim beliefs*

$$\tilde{q}_i(t_{-i}|t_i)=r_i(t_i)\cdot q_i(t_{-i}|t_i,(1,\mu^0))+(1-r_i(t_i))\cdot p_i(t_{-i}|t_i).$$

Proof: Appendix A. □

One can prove a stronger result than (ii) of Lemma 3. Suppose that μ^0 is any admissible mechanism, i.e., $\mu^0=(S,g^0)\in\Xi$. Let $\mu^1=(R,g^1)$ be another admissible mechanism and suppose that the players play the voting game as before. Let $\mathfrak{C}$ be the set of collective decision making problems $C=(D,T_j,p_i,u_i)$, where the sets D and T_i are finite. Let Λ be the correspondence Λ: $\mathfrak{C}\twoheadrightarrow\Xi^{DI}$ which maps each collective decision making problem to the set of incentive-compatible mechanisms ($\forall C\in\mathfrak{C}$, $\Lambda(C)=IC$). Let $\Phi(\mu^0,\mu^1)$ be the equilibrium correspondence. Consider the set $\theta(\mu^0,\mu^1)$ of all pairs (r,σ) which can be part of a sequential equilibrium. (I.e., $(r,\sigma)\in\theta(\mu^0,\mu^1)\Leftrightarrow\exists q$, $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$.) Let Θ be the correspondence defined by

$$\forall C\in\mathfrak{C},\ \Theta(C)=\bigcup_{(\mu^0,\mu^1)\in\Xi\times\Xi}\theta(\mu^0,\mu^1)\times\{(\mu^0,\mu^1)\}.$$

Let $\chi \equiv \Theta(\mathcal{C})$ be the image by Θ of $\mathcal{C}$. Finally, let ψ be the function $\psi: \Theta(\mathcal{C}) \rightarrow \Xi^{DI}$ defined by $\psi((r,\sigma),(\mu^0,\mu^1)) = \mu = (T,g)$ where g is given by (3.1) (keeping in mind that $\mu^0 = (S,g^0)$ and $\mu^1 = (R,g^1)$).

The following result states that the set of incentive compatible mechanisms coincides with the set of mechanisms induced by the sequential equilibria of the voting game when all possible pairs of mechanisms are considered. In other words, the diagram in Figure 3.4 commutes.

LEMMA 4: $\forall C \in \mathcal{C}$, $\Lambda(C) = \psi(\Theta(C))$, *where the equality is the equality of sets.*

Proof: The proof of $\psi(\Theta(C)) \subset \Lambda(C)$ follows the same lines than the proof of Lemma 3 (ii). To prove $\psi(\Theta(C)) \supset \Lambda(C)$, let $\mu^0 \in IC$ and consider $\mu^1 = \mu^0$.[9] □

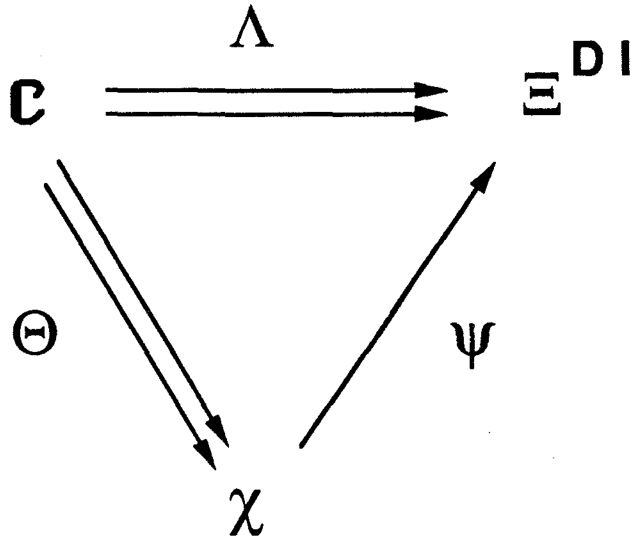

Figure 3.4

4. CHARACTERIZATION OF STRONG DURABILITY

Following Lemma 2, I will consider for now on only incentive compatible status-quo (i.e., I will use Definition 2). We saw in Example 1 that HM-durable incentive compatible direct mechanisms can be interim, and ex-ante, dominated by another incentive compatible direct mechanism. I show in the next lemma that a direct mechanism is strongly durable only if it is interim efficient. I also show that some ex-ante efficient mechanisms are not strongly durable while some interim (but not ex-ante) efficient mechanisms are strongly durable. Lemma 6 will establish the relationship between strong durability and durability. The relationship between HM-durability and strong durability is established in Theorem 1. A strongly durable mechanism is HM-durable and interim efficient whenever each type has a positive probability of occurrence.

Let E_A and E_I be the sets of ex-ante efficient and interim efficient incentive compatible mechanisms. (We know from Holmström-Myerson (1983) that $E_A \subset E_I$.)

LEMMA 5: (*i*) $SD \subset E_I$,

(*ii*) $E_A \setminus SD \neq \emptyset$,

(*iii*) $SD \cap (E_I \setminus E_A) \neq \emptyset$,

Proof: (i) Let $\mu^0 \in IC \setminus E_I$. Let $\mu^1 \in E_I$ such that $\forall i$, $\forall t_i$, $U_i(\mu^1|t_i) \geq U_i(\mu^0|t_i)$, with a strict inequality for some t_i. Such a mechanism exists since $\mu^0 \notin E_I$. Consider the following strategies and beliefs.

$\forall i$, $\forall t_i$,

$$r_i(t_i)=1;\ \sigma_i(t_i|t_i,\mu^1)=1;\ \sigma_i(t_i|t_i,(1,\mu^0))=1;\ \sigma_i(t_i|t_i,(0,\mu^0))=1.$$

$q_i(t_{-i}|t_i,\mu^1)=p_i(t_{-i}|t_i)$; $q_i(t_{-i}|t_i,(0,\mu^0))=p_i(t_{-i}|t_i)$.

Choose a sequence $\{r_i^k\}$ such that $\forall i,\forall t_i$, $r_i^k(t_i)=1-\frac{1}{k}$, $k>0$. Then define,

$$q_i(t_{-i}|t_i,(1,\mu^0)) = \lim_{k\uparrow\infty} \frac{p_i(t_{-i}|t_i)\cdot\left[1-\left(1-\frac{1}{k}\right)^{n-1}\right]}{\sum_{\hat{t}_{-i}} p_i(\hat{t}_{-i}|t_i)\cdot\left[1-\left(1-\frac{1}{k}\right)^{n-1}\right]}$$

$$= p_i(t_{-i}|t_i).$$

By incentive compatibility of μ^0 and of μ^1, $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$. Since μ^1 interim dominates μ^0, $\mu^0\notin SD$.

(ii) (An ex-ante efficient mechanism which is not strongly durable: $E_A\setminus SD\neq\emptyset$.) This example is due to Holmström-Myerson (1983).

A
B
C

	2a		2b	
1a	2	2	2	2
	1	1	1	1
	0	0	0	−8
1b	0	2	0	2
	4	1	4	1
	9	0	9	−8

Consider the two mechanisms $\mu^0=\begin{bmatrix} A & B \\ C & B \end{bmatrix}$, $\mu^1=\begin{bmatrix} A & A \\ A & A \end{bmatrix}$. The reader can verify that the following assessment is indeed a sequential equilibrium.

$r_1(1a)=r_2(2a)=r_2(2b)=1$ & $r_1(1b)=0$;

$\forall t_1,\forall t_2$, $\sigma_1(t_1|t_1,\mu^1)=1$ & $\sigma_2(t_2|t_2,\mu^1)=1$;

$\forall h\in H\setminus\{\mu^1\}$, $\forall t_1,\forall t_2$, $\sigma_1(t_1|t_1,h)=1$ & $\sigma_2(2b|t_2,h)=1$,

$\forall t_1$, $\forall t_2$, $q_1(t_2|t_1,\mu^1)=0.5$ & $q_1(t_2|t_1,(1,\mu^1))=0.5$ & $q_1(t_2|t_1,(0,\mu^0))=0.5$,

$\forall t_1,\ \forall t_2,\ q_2(1a|t_2,\mu^1)=1\ \&\ q_2(1a|t_2,(1,\mu^0))=0\ \&\ q_2(t_1|t_2,(0,\mu^0))=0.5.$

Thus, all types are truthful if μ^1 is played, while the two types of player 2 pool at $2b$ if μ^0 is played. The resulting mechanism is $\begin{bmatrix} A & A \\ B & B \end{bmatrix}$. In this equilibrium, both types of player 2 are made better off, type $1a$ of player 1 is made better off and type $1b$ of player 1 is made worse off.

(iii) (An interim efficient mechanism which is not ex-ante efficient but which is strongly durable: $SD \cap (E_I \setminus E_A) \neq \emptyset$.) Consider the environment of Example 3. Let $\mu^0 = \begin{bmatrix} B & B \\ B & C \end{bmatrix}$. Observe that $\mu^0 \in E_I$ but that $\mu^0 \notin E_A$ since the mechanism $\mu = \begin{bmatrix} B & C \\ C & C \end{bmatrix}$ ex-ante dominates μ^0. ($U_1(\mu^0)=3.75$, $U_1(\mu)=4.75$, $U_2(\mu^0)=U_2(\mu)=1.25$.) Consider another admissible mechanism $\mu^1=(S,g^1)$. Consider any sequential equilibrium $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$. First, we observe that whenever μ^0 is played, it is possible for types $1a$ and $2a$ of players 1 and 2 to enforce the decision B. Thus, it must be true that

$$U_1(\psi(r,\sigma)|1a) \geq U_1(\mu^0|1a),$$

$$U_2(\psi(r,\sigma)|2a) = U_2(\mu^0|2a).$$

Where the equality follows the fact that B is the preferred decision of $2a$.

<u>Case 1</u>: $U_1(\psi(r,\sigma)|1a) > U_1(\mu^0|1a)$. This is true only if decision A is obtained with positive probability in μ^1. If A is obtained with positive probability in μ^1, type $2b$ of player 2 will always prefer to vote against μ^1, since $2b$ can insure that his second most preferred decision (B) is implemented in μ^0. (For instance by pooling at $2a$.) Thus, μ^1 can be voted only if type $2a$ votes for μ^1 and if the decision B is implemented with probability one when $2a$ uses his strategy $\sigma_2(2a|2a,\mu^1)$ in μ^1. Otherwise, $2a$ will vote against μ^1 and μ^0 endures μ^1 with respect to the equilibrium (r,σ,q). If decision B is obtained when μ^1 is played, then, it follows that type $1a$ cannot be made better off

than in μ^0. This contradicts the assumption that $U_1(\psi(r,\sigma)|1a) > U_1(\mu^0|1a)$.

Case 2: $U_1(\psi(r,\sigma)|1a) = U_1(\mu^0|1a)$. Suppose that μ^0 does not strongly endure μ^1. Then, by definition, it must be true that either $U_1(\psi(r,\sigma)|1b) > U_1(\mu^0|1b)$ or that $U_2(\psi(r,\sigma)|2b) > U_2(\mu^0|2b)$. It is not possible to have *both* inequalities since this would imply that $\psi(r,\sigma)$ interim dominates μ^0, which contradicts the fact that $\mu^0 \in E_I$. In fact, if $1b$ is made (strictly) better off, then $2b$ is made (strictly) worse off, and vice-versa, in $\psi(r,\sigma)$ than in the truthful equilibrium of μ^0. Since we have seen that decision B must be implemented in $\psi(r,\sigma)$ when $2a$ follows his equilibrium strategy, $1b$ *cannot* be made better off. Indeed, for any strategy that player 2 follows in μ^1, when 2 is of type $2b$, the best decision that $1b$ can expect is C, which is exactly what μ^0 will select. Suppose that $1b$ is made worse off and that $2b$ is made better off. Then a decision different from C is selected when $1b$ and $2b$ follow their equilibrium strategies. Since $1a$ must be indifferent between $\psi(r,\sigma)$ and μ^0, and because B is implemented when player 2 is of type $2a$, it must be true that decision B is also implemented when 1 is of type $1a$. But this implies that the decision B is implemented when 1 is of type $1a$ and 2 is of type $2b$. This implies that type $2b$ is strictly worse off.

Consequently, for any μ^1, for any equilibrium $(r,\sigma,q) \in \Phi(\mu^0,\mu^1)$, no type can be made better off. This implies that μ^0 is strongly durable. □

The next result shows that the set of strongly durable mechanisms is a subset of the set of interim efficient HM-durable mechanisms whenever each type has a positive probability of occurrence. An example shows that the inclusion is strict.

THEOREM 1: *Suppose that* $\forall i$, $\forall t_i$, $\forall t_{-i}$, $p_i(t_{-i}|t_i) > 0$. *Then, (i)* $SD \subset HMD \cap E_I$. *(ii) The inclusion can be strict.*

Proof: (i) I show that $\mu^0 \in SD \Rightarrow \mu^0 \in HMD$. Since $SD \subset E_I$, (i) will follow. Denote by $\Phi^0(\mu^0,\mu^1)$ the set of sequential equilibria for which the players are truthful when μ^0 is played. Clearly, $\Phi^0(\mu^0,\mu^1) \subset \Phi(\mu^0,\mu^1)$. Suppose by way of contradiction that $\mu^0 \notin HMD$. Then, it must be true that there exists a mechanism μ^1 and that for any $(r,\sigma,q) \in \Phi^0(\mu^0,\mu^1)$, if (HMD3) is satisfied then there exists a profile of types $\tilde{t}$ for which $r(\tilde{t})>0$. It follows that for each type t_i, the quantity $\sum_{t_{-i}} p_i(t_{-i}|t_i) \cdot r_{-i}(t_{-i})$ is nonzero.

Consider the following assessment $(\hat{r},\hat{\sigma},\hat{q})$.

$\forall i,\ \forall t_i,\ \forall t_{-i},\ \forall h \in H \backslash \{\mu^1\}$,

(4.1) $\hat{r}_i(t_i)=0$.

(4.2) $\hat{\sigma}_i(s_i|t_i,\mu^1)=\sigma_i(s_i|t_i,\mu^1)$.

(4.3) $\hat{\sigma}_i(t_i|t_i,h)=1$.

(4.4) $\hat{q}_i(t_{-i}|t_i,\mu^1)=q_i(t_{-i}|t_i,\mu^1)$.

(4.5) $\hat{q}_i(t_{-i}|t_i,h)=p_i(t_{-i}|t_i)$.

With this assessment, the information set μ^1 is reached with a zero probability. Let $\{\hat{r}_i^k(t_i);\ k \in \mathbb{N} \backslash \{0\}\}$ be the sequence defined by

$\forall k,\ \forall i,\ \forall t_i,\ \hat{r}_i^k(t_i)=r_i(t_i)/k$.

With this sequence, the beliefs in (4.4) are consistent since the beliefs q_i were consistent. Since (r,σ,q) is a sequential equilibrium, by (S1), $(\hat{\sigma}_i(s_i|t_i,\mu^1);i=1,\cdots,n)$ as defined in (4.2) is an equilibrium of the mechanism μ^1 when the interim beliefs are given by (4.4). From (4.1), the beliefs (4.5) are consistent. (4.3) is obviously an equilibrium of the mechanism μ^0 since μ^0 is incentive compatible. Since every type votes against μ^1, it is a (possible weak) best response for each type to vote against μ^1.

Suppose now that (HMD3) is not satisfied for the assessment $(\hat{r},\hat{\sigma},\hat{q})$. Then, there exists a type t_i for which,

$$\sum_d \sum_{t_{-i}} \hat{q}(t_{-i}|t_i,\mu^1)\cdot g^0(d|t)\cdot u_i(d,t) <$$

$$\sum_d \sum_{t_{-i}} \hat{q}(t_{-i}|t_i,\mu^1)\cdot \sum_s \sigma(s|t)\cdot g^1(d|s)\cdot u_i(d,t)$$

From (4.4) and the fact that $\sum_{t_{-i}} p_i(t_{-i}|t_i)\cdot r_{-i}(t_{-i})\neq 0$, this inequality can be rewritten as,

$$\sum_d \sum_{t_{-i}} p_i(t_{-i}|t_i)\cdot r_{-i}(t_{-i})\cdot g^0(d|t)\cdot u_i(d,t) <$$

$$\sum_d \sum_{t_{-i}} p_i(t_{-i}|t_i)\cdot r_{-i}(t_{-i})\cdot \sum_s \sigma(s|t)\cdot g^1(d|s)\cdot u_i(d,t)$$

This inequality cannot be true for t_i such that $r_i(t_i)=0$ since by assumption (r,σ,q) satisfies (HMD3). Consequently, this inequality is true only if $r_i(t_i)>0$. Multiplying both sides of this inequality by $r_i(t_i)$ and rearranging yields to,

$$\sum_d \sum_{t_{-i}} p_i(t_{-i}|t_i)\cdot g^0(d|t)\cdot u_i(d,t) <$$

$$\sum_d \sum_{t_{-i}} p_i(t_{-i}|t_i)\cdot \Big[r(t)\cdot \sum_s \sigma(s|t)\cdot g^1(d|s) + (1-r(t))\cdot g^0(d|t) \Big]\cdot u_i(d,t),$$

which is equivalent to $U_i(\mu^0|t_i) < U_i(\psi(r,\sigma)|t_i)$. This implies that (r,σ,ψ) is a sequential equilibrium in which one type is strictly better off than in μ^0. This contradicts $\mu^0 \in SD$. Thus, $(\hat{r},\hat{\sigma},\hat{q})\in\Phi^0(\mu^0,\mu^1)$ satisfies (HMD3) and for any t, $\hat{r}(t)=0$. This proves that μ^0 is HM-durable after all.

(ii) I exhibit an example for which the inclusion is strict. Consider the following modification of the example which was used in the proof of Lemma 5 (ii).

		$2a$		$2b$	
$1a$	A	2	2	2	2
	B	1	1	2	1
	C	0	0	0	−8
$1b$	A	0	2	0	2
	B	4	1	4	1
	C	9	0	9	−8

Consider the two mechanisms $\mu^0=\begin{bmatrix} A & B \\ C & B \end{bmatrix}$, $\mu^1=\begin{bmatrix} A & A \\ A & A \end{bmatrix}$. Then, $\mu^0 \in HMD \cap E_I$ but μ^0 does not strongly endure μ^1. (Observe that μ^0 is HM-durable now because type $1a$ is indifferent between decisons A and B when player 2 is of type $2b$.) To show that μ^0 does not strongly endure μ^1, it is possible to use the same sequential equilibrium than in Lemma 4 (ii). (Note that there exists a sequential equilibrium satisfying the conditions of Theorem 1 in the voting game between μ^0 and the mechanism $\hat{\mu}=\begin{bmatrix} A & A \\ B & B \end{bmatrix}$. This equilibrium involves both types of player 2 pooling at $2b$ in μ^0.) □

It is clear from Lemma 5 (iii) that it is more difficult to show that a mechanism is strongly durable than to show that a mechanism is not strongly durable. This is because it is necessary to show that the status-quo strongly endures *any* admissible mechanism. The following results show that there is a nice characterization of the strongly durable mechanisms. I show in Theorem 2 that a mechanism μ^0 *is not* strongly durable if, and only if, there exists an interim efficient mechanism μ^1 and a sequential equilibrium in $\Phi(\mu^0,\mu^1)$ in which each type votes for μ^1 with probability one and for which the interim utility of at least one type is greater with μ^1 than with μ^0. It follows (Corollary 2) that a mechanism is strongly durable if, and only if, it is interim efficient and any other interim efficient mechanism which is not payoff equivalent is *not* voted with probability one.

In order to prove this result, I will need two intermediate results. The first one is technical in nature and is Lemma 6. The second one is summarized in Lemma 7 below, and has the following interpretation. Recall that for a sequential equilibrium $(r,\sigma,q) \in \Phi(\mu^0,\mu^1)$, $\psi(r,\sigma)$ denotes the equivalent direct mechanism which is obtained in the sequential equilibrium. (See (3.1)-(3.3).) I prove in Lemma 7 the intuitive result that if μ^0 does not strongly endure μ^1, (i.e., $\exists(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$ such that $\exists i$, $\exists t_i$, $U_i(\mu^0|t_i)< U_i(\psi(r,\sigma)|t_i)$), then μ^0 does not strongly endure $\psi(r,\sigma)$. It is relatively easy to show that there exists a Bayesian Nash equilibrium for which $\psi(r,\sigma)$ is selected with probability one. To show that such an equilibrium can be made sequential requires the technical result that is presented in Lemma 6. The proofs of Lemma 6 and 7 are presented in Appendix A.

LEMMA 6:[10] *Let* $I=[0,1]$. *Let* $\forall i\in\{1,\cdots,n\}$, $x_i\in X_i=I^{c_i}$, $c_i\in\mathbb{N}\setminus\{0\}$. *Then,* $\exists\,\bar{d}\in\mathbb{R}$, $\bar{d}>>0$, $\forall d<<\bar{d}$, $\forall i$, $\exists\, y_i\in X_i$ *such that* $\forall i$, $\forall t_{-i}\in \times_{j\neq i}\{1,\cdots,c_j\}$,

$$\prod_{j\neq i} y_j(t_j)=1-\Big(1-\prod_{j\neq i} x_j(t_j)\Big)\cdot d_i.$$

Lemma 6 is useful because it tells us that for any initial voting strategies, there exist other voting strategies that lead to the same belief structure and that have the property that in the limit (as $d_i\downarrow 0$), each type will vote for the alternative mechanism with probability one.

LEMMA 7: *If* μ^0 *does not strongly endur* μ^1 *then there exists* $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$ *such that* μ^0 *does not strongly endure* $\psi(r,\sigma)$. *Moreover, there exists* $(\hat{r},\hat{\sigma},\hat{q})\in\Phi(\mu^0,\psi(r,\sigma))$ *such that* $\psi(\hat{r},\hat{\sigma})=\psi(r,\sigma)$ *and such that* $\forall i,\forall t_i$, $\hat{r}_i(t_i)=1$.

Lemma 7 states that, if μ^0 is not strongly durable. then, the incentive compatible mechanism resulting from a sequential equilibrium (in which one type is made better off) has the property that if the players had had to play the voting game between μ^0 and this resulting mechanism, there would have existed a sequential equilibrium in which the status-quo is rejected with probability one and in which the players are truthful in the alternative mechanism. From this lemma, the main theorem of this section follows.

> THEOREM 2: *An incentive compatible direct mechanism μ^0 is not strongly durable if and only if there exists an interim efficient mechanism $\hat{\mu}$ and a sequential equilibrium of the voting game between μ^0 and $\hat{\mu}$ in which $\hat{\mu}$ is chosen with probability one, in which the players are truthful in $\hat{\mu}$ and in which at least one type is made better off.*

Proof of Theorem 2: Formally, Theorem 2 can be rewritten as,

$$\mu^0 \notin SD \Leftrightarrow \exists \hat{\mu} \in E_I,\ \exists (r,\sigma,q) \in \Phi(\mu^0,\hat{\mu}),\ \left[\forall t,\ r(t)=1\ \&\ \exists i,\ \exists t_i,\ U_i(\mu^0|t_i) < U_i(\hat{\mu}|t_i)\right].$$

Sufficiency ($\Leftarrow$) is obvious by the definition of a strongly durable mechanism. To prove necessity ($\Rightarrow$), suppose that $\mu^0 \notin SD$. Then, there exist μ^1, $(r,\sigma,q) \in \Phi(\mu^0,\mu^1)$ and a type t_i such that $U_i(\mu^0|t_i) < U_i(\psi(r,\sigma)|t_i)$, where $\psi(r,\sigma)$ is the direct mechanism resulting from the equilibrium (r,σ,q). From Lemma 7, there exists $(\hat{r},\hat{\sigma},\hat{q}) \in \Phi(\mu^0,\psi(r,\sigma))$ such that, $\forall i$, $\hat{r}_i(t_i)=1$. If $\hat{\mu}=\psi(r,\sigma) \in E_I$, then the theorem is proved. Otherwise, let $\hat{\mu} \in E_I$, such that $\hat{\mu}$ interim dominates $\psi(\hat{r},\hat{\sigma})$. Such a mechanism exists by assumption. It is immediate that $(\hat{r},\tilde{\sigma},\hat{q}) \in \Phi(\mu^0,\hat{\mu})$, where $\tilde{\sigma}$ is defined by: $\forall i$, $\forall t_i$, $\tilde{\sigma}_i(t_i|t_i,\hat{\mu})=\hat{\sigma}_i(t_i|t_i,\mu^1)=1$ and that $\forall h \in \{(0,\mu^0),(1,\mu^1)\}$, $\tilde{\sigma}_i(\hat{t}_i|t_i,h)=\hat{\sigma}_i(\hat{t}_i|t_i,h)$. (The difference in notation between $\hat{\sigma}$ and $\tilde{\sigma}$ takes into account the fact that a different alternative mechanism is played.) Because $\hat{\mu}$ interim dominates μ^1, $\hat{\mu}$ and $(\hat{r},\tilde{\sigma},\hat{q})$ satisfy the conditions of Theorem 2. □

Example 1 is an obvious illustration of Theorem 2. Another illustration of this theorem is given by the example used in the proof of Lemma 5 (ii). Consider the mechanism $\mu^1=\begin{bmatrix} A & A \\ B & B \end{bmatrix}$. Let, for each type t_i, $r_i(t_i)=1$, and let the players report their true type in μ^1. (This is possible since μ^1 is in E_I.) If μ^0 must be played, suppose that each type does not change his beliefs and that both types of player 2 pool at $2b$ and that both types of player 1 report truthfully. It can be verified that these strategies and beliefs constitute a sequential equilibrium. (The argument seems to rely on the fact that μ^0 has two equilibria. However, the argument could have been made by supposing that each type of player 2 believe that, if μ^0 is played while he voted for μ^1, 1 is of type $1b$ with probability one.)

Theorem 2 has for corollary that it is possible to find a necessary and sufficient condition for strong durability.

COROLLARY 2: *An incentive compatible direct mechanism μ^0 is strongly durable if and only if it is interim incentive efficient and for any other interim incentive efficient mechanism μ^1 either this mechanism is chosen with probability strictly less than one in any equilibrium of the game between μ^0 and μ^1 or is payoff equivalent to μ^0.*

Proof: Formally, the statement of the corollary can be written as ("$\vee$" is the logician symbol for "or")

$$\mu^0 \in SD \Leftrightarrow \forall \mu^1 \in E_I,\ \forall (r,\sigma,q) \in \Phi(\mu^0,\mu^1),\ \left[\exists t,\ r(t)<1\ \vee\ \forall i,\ \forall t_i,\ U_i(\mu^0|t_i) \geq U_i(\mu^1|t_i)\right].$$

This is the negation of Theorem 2. Because $\mu^0 \in E_I$ and $\mu^1 \in E_I$, all the inequalities are in fact equalities, i.e., μ^1 is payoff equivalent to μ^0. □

We will see in the next section that whenever the set of interim payoffs corresponding to the set of interim efficient mechanism is not a singleton, then all the equilibria of a strongly durable direct mechanism are, under a regularity condition, either interim payoff equivalent or *at most one type* has an interim payoff which is different, hence lower, than the payoff in the truthful equilibrium.

5. A GEOMETRIC CHARACTERIZATION AND AN EXISTENCE RESULT

5.1. *A GEOMETRIC CHARACTERIZATION*

Theorem 2 has the following geometric interpretation.[11] Consider a given mechanism μ^0 and a system of prior beliefs p. Consider the set of beliefs q which are obtained in a consistent way from p when each type votes for the alternative mechanism with probability one. That is to say,

$$(5.1)\quad \exists\{r_i^k\}, \forall i,\ \forall t_i,\ \lim_{k\uparrow\infty} r_i^k(t_i)=1,\ q_i(t_{-i}|t_i)=\lim_{k\uparrow\infty}\frac{p_i(t_{-i}|t_i)\cdot(1-r_{-i}^k(t_{-i}))}{\sum_{\hat{t}_{-i}} p_i(\hat{t}_{-i}|t_i)\cdot(1-r_{-i}^k(\hat{t}_{-i}))}.$$

Let us map each possible system of beliefs q to the set of corresponding Bayesian Nash equilibria σ of the mechanism μ^0. Given the beliefs q and the strategy vector σ, each type can consider what strategy will maximize his interim utility given that the other players obey the strategy σ and given that this type has beliefs p. Doing so, we can define for μ^0 the set $A(\mu^0)$ of interim payoffs which are obtained by following the previous construction.

Theorem 2 can be rephrased as follows. μ^0 is strongly durable if and only if there does not exist a vector of interim payoffs in the set $A(\mu^0)$ which is (possibly weakly) dominated by a vector of interim payoffs corresponding to an interim efficient mechanism μ^1 (given the prior beliefs p). The reason is that if all the types vote with probability one for μ^1, then the players are free to form their beliefs in any way which is consistent with (5.1) and to play a certain equilibrium in μ^0 if μ^0 is played. However, in a sequential equilibrium of the voting game in which each type votes for the alternative mechanism with probability one, a type should not have an incentive to change his vote and then to change his strategy when μ^0 is played. Thus, the best that a type can obtain by deceiving the other players must be dominated by what he gets by playing the alternative mechanism. Because this must be true for every type, the result follows.

Consider a Bayesian collective decision making problem $C=(D,T_i,u_i,p_i)\in\mathcal{C}$. Denote by $U(\mu)\equiv\Big(U_i(\mu|t_i);\ i=1,\cdots,n,\ t_i\in T_i\Big)$ the vector of interim payoffs when all the players have beliefs p and when all the players are truthful. Let $\mathcal{U}$ be the set of interim payoffs corresponding to the set of incentive compatible mechanisms and let $\overline{\mathcal{U}}$ be the set of interim payoffs corresponding to the interim efficient mechanisms. Formally,

$$\mathcal{U}\equiv\bigcup_{\mu\in IC}\Big\{U(\mu)\Big\}$$

$$\overline{\mathcal{U}}\equiv\Big\{u\in\mathcal{U}\,|\ \forall w\in\mathcal{U},\ \sim(w>u)\Big\}.$$

$\mathcal{U}$ and $\overline{\mathcal{U}}$ are subsets of the Euclidian space $E^{\sum_i|T_i|}$, considered as a metric space (with the usual metric). It is immediate that $\mathcal{U}$ is a bounded polyhedral set.[12] (For an earlier recognition of this fact, see, e.g., Myerson (1983) or Ledyard (1986).)

Recall that if μ is a mechanism, then $E(\mu,q)$ denotes the set of Bayesian Nash equilibria of μ given the interim beliefs q. It is a matter of routine to check that $E(\mu,q)$ is a upper hemi-continuous correspondence. (This follows closeness and the fact that the interim payoff function is continuous in (μ,q).) I will write $U_i((\mu,q,\sigma)|t_i)$ to denote the expected utility of type t_i of player i in the equilibrium σ, i.e.,

$$U_i((\mu,q,\sigma)|t_i)\equiv\sum_d\sum_{t_{-i}}q_i(t_{-i}|t_i)\cdot\sum_{\hat{t}}\prod_j\sigma_j(\hat{t}_j|t_j)\cdot g(d|\hat{t})\cdot u_i(d,t).$$

With this notation, $U_i(\mu|t_i)=U_i((\mu,p,\sigma)|t_i)$ where σ is the equilibrium in which every type t_i has the degenerate strategy $\sigma_i(t_i|t_i)=1$. Observe that this notation makes sense since μ is supposed to be incentive compatible. If $\sigma\in E(\mu,q)$, $\forall i$, $\forall t_i$, $U_i((\mu,q,(\sigma_{-i},\tilde{\sigma}_i))|t_i)\leq U_i((\mu,q,\sigma)|t_i)$ by definition of the set $E(\mu,q)$. This does not mean however that $\sigma_i(\cdot|t_i)$ is a best response for type t_i when the other players are using the strategy σ_{-i} and when t_i has beliefs $p_i(\cdot|t_i)$. Let $u_i(\mu,\sigma|t_i)$ be the maximum interim payoff that type t_i can obtain when he has beliefs p_i and when the other players are using the strategies σ_{-i},

$$(5.2) \qquad u_i(\mu,\sigma|t_i) \equiv \underset{\tilde{\sigma}_i}{Max} \sum_{t_{-i}} p_i(t_{-i}|t_i) \cdot \sum_{\hat{t}} \sigma_{-i}(\hat{t}_{-i}|t_{-i}) \cdot \tilde{\sigma}_i(\hat{t}_i) \cdot g(d|\hat{t}) \cdot u_i(d,t).$$

This expression is well defined since the right hand side is a continuous function of $\tilde{\sigma}_i$ and the maximization is performed on a compact domain. Let $u(\mu,\sigma) \equiv \left(u_i(\mu,\sigma|t_i);\ i=1,\cdots,n,\ t_i \in T_i\right)$ denote the resulting vector in $E^{\sum_i |T_i|}$. The set $A(\mu)$ is defined as the union of all possible vectors $u(\mu,\sigma)$,

$$(5.3) \qquad A(\mu) \equiv \bigcup_{\substack{q \in \times_i \Delta(T_{-i})^{|T_i|-1} \\ q \text{ satisfies } (5.1)}} \ \bigcup_{\sigma \in E(\mu,q)} \left\{u(\mu,q)\right\}.$$

For each $u \in \mathcal{U}$, there are possibly many incentive compatible mechanisms μ such that $u = U(\mu)$. Consider the correspondence $F: \mathcal{U} \twoheadrightarrow IC$ which maps each vector $u \in \mathcal{U}$ to the set of incentive compatible mechanisms μ such that $u = U(\mu)$. $f: \mathcal{U} \rightarrow IC$ is a *selection* of F if for each $u \in \mathcal{U}$, $f(u) \in F(u)$. In the following I will consider a particular selection f and I will identify the vector u with the corresponding mechanism $f(u)$.

The following lemma summarizes the previous discussion. Theorem 3 which follows gives a necessary condition for an interim efficient mechanism to be strongly durable. From Theorem 3, the truthful equilibrium of a strongly durable direct mechanism has in general the property that all the equilibria of this mechanism are interim payoff equivalent.

LEMMA 7: *Let $\mu \in IC$. Then μ is strongly durable if and only if μ is interim efficiency and if for all q which satisfies* (5.1), *for all $\sigma \in E(\mu,q)$, it is not true that there exists $u' \in \overline{\mathfrak{U}} \setminus \{u^0\}$ such that $u' > u(\mu,\sigma)$.*

Proof: This is a restatement of Theorem 2. □

We observe that if $\overline{\mathfrak{U}} = \{U(\mu^0)\}$ then μ^0 is obligatorily strongly durable since any other interim efficient mechanism is interim payoff equivalent to μ^0. When $\overline{\mathfrak{U}}$ is not a singleton, we have the following result. Let $N_\epsilon(u)$ denote an ϵ-neighborhood of u. If will say that $u \in \overline{\mathfrak{U}}$ is *irregular with respect to* $\overline{\mathfrak{U}}$ if for any $\epsilon > 0$, $N_\epsilon(u) \cap (relbd\,\mathfrak{U} \setminus \overline{\mathfrak{U}}) \neq \emptyset$. Otherwise, u is *regular with respect to* $\overline{\mathfrak{U}}$. By extension, an interim efficient mechanism is regular if the interim payoff corresponding to the truthful equilibrium is regular with respect to $\overline{\mathfrak{U}}$.

THEOREM 3: *Suppose that $\overline{\mathfrak{U}}$ is not a singleton. Let $\mu^0 \in E_I$, and suppose that $U(\mu^0)$ is regular with respect to $\overline{\mathfrak{U}}$. Then μ^0 is strongly durable only if (i) for all $\sigma \in E(\mu^0,p)$, $u(\mu^0,\sigma) \leq U(\mu^0)$ and (ii) whenever there exists $\sigma \in E(\mu^0,p)$ for which $u(\mu^0,\sigma) \neq U(\mu^0)$, $\overline{\mathfrak{U}}$ is included in a hyperplane H and there exists a type t_i such that $u((\mu^0,\sigma)|t_i) < U(\mu^0|t_i)$ and such that for all types $t_j \neq t_i$, $u((\mu^0,\sigma)|t_j) = U(\mu^0|t_j)$.*

Proof: See Appendix A. □

Thus, a regular interim efficient mechanism is strongly durable only if the truthful equilibrium interim dominates all the other equilibria and *at most one type is made strictly worse off* in a non-truthful equilibrium than in the truthful equilibrium. The following example illustrates the necessity of imposing the regularity of $U(\mu^0)$ with respect to $\overline{\mathfrak{U}}$ in Theorem 3.

<u>Example 4</u>:

A
B

	2a		2b	
1a	2	1	2	1
	1	2	1	2
1b	4	1	4	1
	4	2	4	2

Let $\mu^0=\begin{bmatrix} A & A \\ B & B \end{bmatrix}$ and $\mu^1=\begin{bmatrix} A & A \\ A & A \end{bmatrix}$. Then, $U(\mu^0)=(2,4,1.5,1.5)$ and $U(\mu^1)=(2,4,1,1)$, where the order in the vectors $U(\cdot)$ corresponds to the order $(1a,1b,2a,2b)$. $U(\mu^1)$ is also the utility payoff vector that is obtained when player 1 pools at $1a$ in the mechanism μ^0. $U(\mu^0)$ is irregular with respect to $\overline{\mathfrak{U}}$ and $U(\mu^1)<U(\mu^0)$, but μ^0 is strongly durable. •

5.2. *AN EXISTENCE RESULT*

I will prove the existence of strongly durable mechanisms under two assumptions. First, there is at least one partner whose utility function is a function only of his own type

(5.4) $\quad \forall d,\ \forall t,\ u_1(d,t)=u_1(d,t_1).$

Second, the decision set is such that each partner has a utility payoff at any decision which is greater than his reservation value.

(5.5) $\forall i,\ \forall t,\ \forall d,\ u_i(d,t) \geq 0.$

Condition (5.5) implies that the any incentive compatible mechanism is ex-ante individual rational. Consider the set:

$$G = \{\mu \in \Xi^{DI} | \mu = (T,g),\ g\colon T_1 \to \Delta(D)\}.$$

G is the set of direct mechanisms which depend only of Mr. 1's message.

(IC_1) $\forall t_1,\ \forall \tilde{t}_1 \neq t_1,\ \sum_d [g(d|t_1) - g(d|\tilde{t}_1)] \cdot u_1(d,t_1) \geq 0.$

(IC_1) is an incentive compatibility condition for Mr. 1 since t_1's utility payoff is independent of the types of the other partners if μ belongs to the set G. (Indeed, $U_1(\mu|t_1) = u_1(g(t_1),t_1)$.) Let G_1 be the set of mechanisms in G which satisfy (IC_1). Clearly, G_1 is a compact and convex subset of IC. [13]

Consider $g^* \in G_1$ such that g^* is nonrandom, i.e., $g^*\colon T_1 \to D$, and such that $g^*(t_1)$ maximizes $u_1(d,t_1)$ over $d \in D$.

<u>Case 1</u>: Suppose that Mr. 1 is never indifferent between two decisions,

(5.6) $\forall t_1,\ \forall d,\ \forall d',\ d \neq d' \Rightarrow u_1(d,t_1) \neq u_1(d',t_1).$

Then, any other incentive compatible mechanism μ such that $g(g^*(t_1)|t_1) \neq 1$ for some type t_1 will lead to a lower utility payoff for t_1. Thus, $\mu^* = (T,g^*) \in E_I$. Because of (5.5), $\mu^* = (T,g^*)$ is an admissible ex-ante contract. From the definition of G_1 and (5.6), for *any* belief structure satisfying (5.3) there exists a unique equilibrium to the game μ^* since the interim payoff of Mr. 1 is independent of the types of the other players. From Lemma 7, μ^* is strongly durable.

<u>Case 2</u>: Suppose that Mr. 1 is not indifferent between two decisions. Let $\mu \in G_1$ be a

nonrandom mechanism which has the property that $u_1(d,t_1)$ is maximized at $g(t_1)$ for any t_1. If for any t_1 and $\tilde{t}_1 \neq t_1$, $u_1(g(t_1),t_1) \neq u_1(g(\tilde{t}_1),t_1)$, then the preceding arguments apply (all the equilibria of μ are interim payoff equivalent). Otherwise, it is possible to construct a new mechanism $\mu' = (T,g')$ with the property

$$(5.7) \qquad \forall t_1, \forall \tilde{t}_1,\ \tilde{t}_1 \neq t_1,\ u_1(g'(t_1),t_1) = u_1(g'(\tilde{t}_1),t_1) \Rightarrow g'(t_1) = g'(\tilde{t}_1),$$

and such that the interim utility payoffs of player 1 are the same.[14] Consider the set $\tilde{G}_1$ of mechanisms in G_1 which maximize the interim utility payoffs of Mr. 1 and which have property (5.7). For all these mechanisms, the interim utility payoffs of Mr. 1 are the same. Let μ^0 be a mechanism in $\tilde{G}_1$ with the additional property that no other mechanism in G_1 interim dominates μ^0, i.e., μ^0 maximizes a certain welfare function $\sum_i \sum_{t_i} \lambda(t_i) \cdot U_i(\mu|t_i)$ where the weights are such that $\forall i$, $\forall t_i$, $\lambda(t_i) > 0$. μ^0 is interim efficient since any incentive compatible mechanism which is interim payoff equivalent to μ^0 for Mr. 1 must be a convex combination of mechanisms in the set $\tilde{G}_1$. The arguments of case 1 can be applied to show that μ^0 satisfies the conditions of Lemma 7.

This proves the following theorem.

THEOREM 4: *Under the assumptions* (5.4) *and* (5.5), *the set of strongly durable mechanisms is nonempty.*

Example 5 illustrates the role of assumptions (5.5). Without, assumption (5.5) it is possible that an incentive compatible mechanism violates the ex-ante individual rationality condition of one player but maximizes the interim utility of all the other types of the other players. In such a case, any ex-ante contract which satisfies individual rationality will not strongly endure that mechanism.

Example 5. There are three players. Players 1 and 2 can have two types and player 3 has only one type. Types are independent and have the same probability of occurrence. Below, α is a real number strictly greater than 1.

A
B

		2a			2b	
1a	2	1	4	2	1	4
	1	2	0	1	2	0
1b	4	$-\alpha$	4	4	$-\alpha$	4
	4	2	0	4	2	0

3

Consider any ex-ante contract μ^0 which satisfies ex-ante individual rationality. Such a contract cannot select decision A when 1 is of type $1b$ since otherwise the individual rationality constraint for player 2 would be violated. For instance, consider the mechanism $\mu^0 = \begin{bmatrix} A & A \\ C & C \end{bmatrix}$, where C is the lottery $[x, A \oplus (1-x), B]$, $x = 3/(2+\alpha)$. In μ^0, both types of player 2 get an ex-ante utility payoff of 0. Consider the alternative mechanism $\mu^1 = \begin{bmatrix} A & A \\ A & A \end{bmatrix}$. Both players 1 and 3 prefer μ^1 to μ^0. Moreover, if player 1 pools at $1a$ in the contract μ^0, there is a sequential equilibrium of the voting game between μ^0 and μ^1 in which μ^1 is selected with probability one. Thus μ^0 is not strongly durable. •

In the following example, Mr. 1 is indifferent between two decisions and an ex-ante contract which maximizes his interim utility payoffs is not strongly durable. This is the reason for which it was necessary to consider separately the case in which some type of Mr. 1 is not indifferent between all decisions in the proof of Theorem 4.

Example 6. The assumptions are the same as in Example 5 except that the utility payoffs are now given by the following matrix:

		2a	2b
1a	A	2 1 4	2 1 4
	B	1 2 0	1 2 0
1b	A	4 1 4	4 1 4
	B	4 2 0	4 2 0

3

Consider the mechanisms $\mu^0=\begin{bmatrix} A & A \\ B & B \end{bmatrix}$ and $\mu^1=\begin{bmatrix} A & A \\ A & A \end{bmatrix}$. Then, $U(\mu^1)=(2,4,1,1,4)$ and $U(\mu^0)=(2,4,1.5,1.5,2)$, where the order in each vector corresponds to the order of the types $(1a,1b,2a,2b,3)$. μ^0 and μ^1 are interim efficient and maximize the interim utility of both types of partner 1 but only μ^1 is strongly durable. Indeed, if μ^0 is the status-quo, then, in the voting game between μ^0 and μ^1, there exists a sequential equilibrium in which all the players vote for μ^1, and in which both types of player 1 pool at $1a$. In this equilibrium, player 3 is strictly better off. μ^1 is strongly durable by Lemma 7. •

6. CONCLUSION

Modeling renegotiation is a difficult exercise. Ideally, one would like to give as much freedom as possible to the players for renegotiating and for choosing the process of renegotiation. Pragmatically, there is a tradeoff between the level of endogeneity that is allowed and the resulting complexity of the model as well as the possible predictions that can be obtained. The analysis made in this chapter should be considered as a first step toward more general models of renegotiation.

I have introduced a new concept of stability when ex-ante contracts can be renegotiated through a mediator, i.e., when the partners are not able to communicate directly. Strongly durable mechanisms are interim efficient and have the property that the truthful equilibrium of an incentive compatible mechanism has the "focal" property that it interim dominates all the other possible equilibria. Under a regularity condition, it can be showed that at most one type can be made worse off in any non-truthful equilibrium of a strongly durable mechanism; all the other types obtain the same interim payoff. In this respect, strong durability highlights the relationship between the question of multiplicity of equilibria in a mechanism and its stability. Multiplicity of equilibria is not so much a problem as long as the truthful equilibrium is "focal." The fact that this is also a necessary condition for stability is a nice outcome of the model.

There are two major weaknesses of this new concept. First, like the concept of durability of Holmström and Myerson (1983), it assumes a communication process which is restrictive. In particular, the fact that the alternative mechanism is proposed exogenously is an extreme assumption. Second, it lacks a general existence theorem. The correction of both problems is the topic of future research.

APPENDIX A

Proof of Lemma 2

Suppose that the status-quo $\mu^0=(M,g^0)$ is not a direct mechanism. Let σ^0 be a Bayesian equilibrium of the game μ^0. The pair (μ^0,σ^0), $\sigma^0 \in E(\mu^0,p)$, is strongly durable if for any μ^1, and any $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$, the interim payoffs of all the types are larger with (μ^0,σ^0), i.e., when μ^0 is played with the strategies σ^0, than in the mechanism induced by the sequential equilibrium (r,σ,q). From Lemma 1, this implies that *all the other* Bayesian equilibria (if they exist) of μ^0 are interim dominated by σ^0. By the revelation principle, there is an incentive compatible mechanism $\hat{\mu}^0=(T,\hat{g}^0)$ that is (interim) payoff equivalent to (μ^0,σ^0). Specifically, $\forall d$, $\forall t$, $\hat{g}^0(d|t)=\sum_m \sigma^0(m|t)\cdot g^0(d|m)$. I need to show that the pair $(\hat{\mu}^0,\beta^0)$ is strongly durable, where β^0 is the truthful strategy $(\forall t_i,\ \beta^0(t_i|t_i)=1)$.

Suppose not. Then, there exist a mechanism $\mu^1=(S,g^1)$ and a sequential equilibrium $(\hat{r},\hat{\sigma},\hat{q})\in\Phi(\hat{\mu}^0,\mu^1)$ such that for some type t_i, $U_i(\hat{\mu}^0|t_i)< U_i(\psi(\hat{r},\hat{\sigma})|t_i)$.

Let $\forall i$, $\forall t_i$, $\forall h\in H$, $\forall h'\in H\backslash\{\mu^1\}$,

$r_i(t_i)=\hat{r}_i(t_i)$,

& $\sigma_i(s_i|t_i,\mu^1)=\hat{\sigma}_i(s_i|t_i,\mu^1)$,

& $q_i(t_{-i}|t_i,h)=\hat{q}_i(t_{-i}|t_i,h)$,

& $\sigma_i(m_i|t_i,h')=\sum_{\hat{t}_i}\hat{\sigma}_i(\hat{t}_i|t_i,h')\cdot\sigma_i^0(m_i|\hat{t}_i)$.

Observe that $\forall i$, $\forall t_i$, $\forall h'\in H\backslash\{\mu^1\}$, $\sigma_i(\cdot|t_i,h')\in\Delta(M_i)$ is an acceptable strategy for t_i. I claim that $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$. By construction, it is clear that it is enough to show that when $h'\in H\backslash\{u^1\}$, the profile of strategies $(\sigma_i(m_i|t_i,h'))$ belongs to $E(\mu^0,q(\cdot|\cdot,h'))$. $\forall h\in H\backslash\{\mu^1\}$,

$$\sum_d \sum_{t_{-i}} q_i(t_{-i}|t_i,h)\cdot\sum_{\hat t}\hat\sigma_{-i}(\hat t_{-i}|t_{-i},h)\cdot\sigma_i(\hat t_i)\cdot\hat g^0(d|\hat t)\cdot u_i(d,t)$$

$$=\sum_d \sum_{t_{-i}} q_i(t_{-i}|t_i,h)\cdot\sum_{\hat t}\Big[\hat\sigma_{-i}(\hat t_{-i}|t_{-i},h)\cdot\sigma_i(\hat t_i)\cdot\Big[\sum_m \sigma^0(m|\hat t)\cdot g^0(d|m)\Big]\Big]\cdot u_i(d,t)$$

(by definition of $\hat g^0$)

$$=\sum_d \sum_{t_{-i}} q_i(t_{-i}|t_i,h)\cdot\sum_{\hat t}\Big[\sum_m\Big[\hat\sigma_{-i}(\hat t_{-i}|t_{-i},h)\cdot\sigma_i(\hat t_i)\cdot\sigma^0(m|\hat t)\cdot g^0(d|m)\Big]\Big]\cdot u_i(d,t)$$

(since the coefficient of the inner bracket was independent of m)

$$=\sum_d \sum_{t_{-i}} q_i(t_{-i}|t_i,h)\cdot\sum_m\Big[\sum_{\hat t}\Big[\Big(\prod_{j\neq i}\hat\sigma_j(\hat t_j|t_j,h)\cdot\sigma_j^0(m_j|\hat t_j)\Big)\cdot\sigma_i(\hat t_i)\cdot\sigma_i^0(m_i|\hat t_i)\Big]\Big]\cdot g^0(d|m)\cdot u_i(d,t)$$

(by rearranging)

$$=\sum_d \sum_{t_{-i}} q_i(t_{-i}|t_i,h)\cdot\sum_m\Big[\prod_{j\neq i}\sum_{\hat t_j}\Big(\hat\sigma_j(\hat t_j|t_j,h)\cdot\sigma_j^0(m_j|\hat t_j)\Big)\cdot\sum_{\hat t_i}\sigma_i(\hat t_i)\cdot\sigma_i^0(m_i|\hat t_i)\Big]\cdot g^0(d|m)\cdot u_i(d,t)$$

(because each $\hat\sigma_j$ is independent of $\hat t_{-j}$.)

If $\hat\sigma$ is an equilibrium in $\hat\mu^0$ when the interim beliefs are given by q_i, it is true that $\sigma_i(\hat t_i)=\hat\sigma_i(\hat t_i|t_i,h)$ maximizes the above expression. Consequently, $\sigma_i(m_i|t_i)=\sum_{\hat t_i}\hat\sigma_i(\hat t_i|t_i,h)\cdot\sigma_i^0(m_i|\hat t_i)$ is a best response to the strategies $\sigma_j(m_j|t_j)=\sum_{\hat t_j}\hat\sigma_j(\hat t_j|t_j,h)\cdot\sigma_j^0(m_j|\hat t_j)$ of the other players in the mechanism μ^0. Thus, (r,σ,q) is indeed a sequential equilibrium of the game between μ^0 and μ^1. Since $(\hat\mu^0,\beta^0)$ and (μ^0,σ^0) lead to the same interim payoffs, $\exists i$, $\exists t_i$, $U_i(\mu^0,\sigma^0|t_i)< U_i(\psi(r,\sigma)|t_i)$, and (μ^0,σ^0) is not strongly durable, contradicting our assumption. □

Proof of Lemma 3

(i) (T,α) and (T,β) as given in (3.2) and (3.3) are admissible direct mechanisms. Indeed, $\forall t$, $\forall d$, $\alpha(d|t)\geq 0$, $\beta(d|t)\geq 0$. Moreover, $\forall t$,

$$\sum_d \alpha(d|t;r)=\sum_{\hat{t}}\prod_i\left\{r_i(t_i)\cdot\sigma_i(\hat{t}_i|t_i,(1,\mu^0))+(1\text{-}r_i(t_i))\cdot\sigma_i(\hat{t}_i|t_i,(0,\mu^0))\right\}\cdot\left[\sum_d g^0(d|\hat{t})\right]$$

$$=\sum_{\hat{t}}\prod_i\left\{r_i(t_i)\cdot\sigma_i(\hat{t}_i|t_i,(1,\mu^0))+(1\text{-}r_i(t_i))\cdot\sigma_i(\hat{t}_i|t_i,(0,\mu^0))\right\}$$

(because $\forall \hat{t}$, $g^0(\cdot,\hat{t})\in\Delta(D)$)

$$=\prod_i\left\{\sum_{\hat{t}_i} r_i(t_i)\cdot\sigma_i(\hat{t}_i|t_i,(1,\mu^0))+(1\text{-}r_i(t_i))\cdot\sigma_i(\hat{t}_i|t_i,(0,\mu^0))\right\}$$

(because the term in braces depended only on $\hat{t}_i$)

$$=1$$

(because $\forall i$, $\forall t_i$, $\forall h\in H$, $\sigma_i(\cdot|t_i,h)\in\Delta(T_i)$).

$$\sum_d \beta(d|t)=\sum_s\prod_i\sigma_i(s_i|t_i,\mu^1)\cdot\sum_d g^1(d|s)$$

$$=\prod_i\sum_{s_i}\sigma_i(s_i|t_i,\mu^1)\cdot\sum_d g^1(d|s)$$

$$=1,$$

since $g^1(\cdot,s)$ is an element of $\Delta(D)$ and $\sigma_i(s_i|t_i,\mu^1)$ is an element of $\Delta(S_i)$.

Finally, $\psi(r,\sigma)$ is admissible since for each profile t, $g(d|t)$ is a convex combination of $\beta(d,t)$ and of $\alpha(d,t)$ which are admissible mechanisms.

(ii) Suppose that $\psi(r,\sigma)\notin IC$. Then, there exists a type t_i such that for $\hat{t}_i\neq t_i$,

$$\sum_d\sum_{t_{-i}}p_i(t_{-i}|t_i)\cdot g(d|(t_{-i},t_i))\cdot u_i(d,t)<\sum_d\sum_{t_{-i}}p_i(t_{-i}|t_i)\cdot g(d|(t_{-i},\hat{t}_i))\cdot u_i(d,t),$$

but this means that type t_i prefers to use the strategy $(r_i(\hat{t}_i),\sigma_i(\cdot|\hat{t}_i,h))_{h\in H}$ instead of the strategy $(r_i(t_i),\sigma_i(\cdot|t_i,h))_{h\in H}$. This contradicts the fact that (r,σ) is a Nash equilibrium. (Recall that every sequential equilibrium is a Nash equilibrium.)

(iii) If (T,β) is not incentive compatible given beliefs $q_i(t_{-i}|t_i,\mu^1)$, then there exists a type t_i and $\hat{t}_i \neq t_i$, such that

$$\sum_d \sum_{t_{-i}} q_i(t_{-i}|t_i,\mu^1)\cdot\beta(d|t_{-i},t_i)\cdot u_i(d,t) < \sum_d \sum_{t_{-i}} q_i(t_{-i}|t_i,\mu^1)\cdot\beta(d|t_{-i},\hat{t}_i)\cdot u_i(d,t),$$

and this contradicts (S3).

(iv) Using (S2) and the fact that $q_i(t_{-i}|t_i,(0,\mu^0)) = p_i(t_{-i}|t_i)$, we have,

$$\text{(B.1)}\quad \sigma_i(\cdot|t_i,(1,\mu^0)) \in \underset{\tilde{\sigma}_i\in\Delta(T_i)}{Argmax} \sum_d \sum_{t_{-i}} q_i(t_{-i}|t_i,(1,\mu^0))\cdot\sum_{\hat{t}} K(\tilde{\sigma}_i)(\hat{t})\cdot g^0(d|\hat{t})\cdot u_i(d,t).$$

$$(B.2)\quad \sigma_i(\cdot|t_i,(0,\mu^0)) \in \underset{\tilde{\sigma}_i\in\Delta(T_i)}{Argmax} \sum_d \sum_{t_{-i}} p_i(t_{-i}|t_i)\cdot\sum_{\hat{t}} K(\tilde{\sigma}_i)(\hat{t})\cdot g^0(d|\hat{t})\cdot u_i(d,t).$$

Suppose that (T,α) is not incentive compatible given the interim beliefs $\tilde{q}_i(t_{-i}|t_i)$ as given in (iv) of the lemma. Then, there exists i and t_i, $\hat{t}_i \neq t_i$ for which

$$\sum_d \sum_{t_{-i}} \tilde{q}_i(t_{-i}|t_i)\cdot\alpha(d|t)\cdot u_i(d,t) < \sum_d \sum_{t_{-i}} \tilde{q}_i(t_{-i}|t_i)\cdot\alpha(d|t_i,\hat{t}_i)\cdot u_i(d,t).$$

For $r_i(t_i)\in[0,1]$, (B.1) is still true if both sides are multiplied by $r_i(t_i)$ and (B.2) is still true if both sides are multiplied by $(1-r_i(t_i))$. (If $r_i(t_i)=0$, by multiplying (B.1) by $r_i(t_i)$, we get a tautology; if $r_i(t_i)=1$, by multiplying (B.2) by $(1-r_i(t_i))$ one also get a tautology; otherwise, for $r_i(t_i)\in(0,1)$, the fact that the set Argmax is independent if the right hand side is multiplied by a positive constant leads to the result.) It follows that

$$r_i(t_i)\cdot\sigma_i(\cdot|t_i,(1,\mu^0)) + (1-r_i(t_i))\cdot\sigma_i(\cdot|t_i,(0,\mu^0))$$

is a best response to the strategies

$$r_j(t_j)\cdot\sigma_j(\cdot|t_j,(1,\mu^0)) + (1-r_j(t_j))\cdot\sigma_j(\cdot|t_j,(0,\mu^0))$$

of the other players given the interim beliefs

$$r_i(t_i)\cdot q_i(t_{-i}|t_i,(1,\mu^0)) + (1-r_i(t_i))\cdot q_i(t_{-i}|t_i,(0,\mu^0)).\quad\square$$

Proof of Lemma 6

Recall that $x_{-i}(t_{-i})$ denotes $\prod_{j\neq i}x_j(t_j)$. (Observe that if $x_{-i}(t_{-i})=1$, then $x_j(t_j)=1$ for the respective $j\neq i$, and it is possible to choose $y_j(t_j)=1$. It follows that there would be no loss of generality in restricting the attention to $X_i=[0,1)^{c_i}$.) Let,

(B.1) $$\bar{d}_i = \max_{t_{-i}:\ x_{-i}(t_{-i})\neq 1}\Big(1-x_{-i}(t_{-i})\Big)^{-1}.$$

For $x\in X=X_1\times\cdots\times X_n$ given, $\bar{d}$ is well defined. Choose any $d\in \times_i(0,\bar{d}_i)$. (If $\bar{d}=0$, the conclusion is immediate since $y=1$ satisfies the condition in the lemma.) For any i, I denote the set $\times_{j\neq i}\{1,\cdots,c_j\}$ by A^{-i} and I write,

(B.2) $$A^{-i}=\{A_1^{-i},A_2^{-i},\cdots A^{-i}_{\prod_{j\neq i}c_j}\}.$$

where,

$$\forall k,k',\ k\neq k',\ A_k^{-i}\neq A_{k'}^{-i}.$$

For $z\in X$, denote $z=(z_1(1),\cdots,z_1(c_1),z_2^2(1),\cdots,z_2(c_2),\cdots,z_n(1),\cdots,z_n(c_n))$. Consider the following functions. (Recall that $x_{-i}(A_l^{-i})=\prod_{j\neq i}x_j(\hat{t}_j)$, where $A_l^{-i}=\{\hat{t}_j,j\neq i\}$.)

(B.3) $\forall i,\ F_i: X\to \mathbb{R}^{\prod_{j\neq i}c_j}$ where for any $y\in X$, the l-th component of $F_i(y)$ is given by

$$F_i(y)(l)=y_{-i}(A_l^{-i})+\Big[1-x_{-i}(A_l^{-i})\Big]\cdot d_i-1.$$

(B.4) $$F: X\to \mathbb{R}^{\sum_i\prod_{j\neq i}c_j}$$

where $\forall y\in X,\ F(y)=(F_1(y),F_2(y),\cdots,F_n(y))$.

Note that F is a well defined function as soon as we define the ordered sets A^{-i} and once x and d are specified. Let, $\forall i$,

$$a_i=\left(\left(1-x_{-i}(A_l^{-i})\right)\cdot d_i-1;\ i=1,\cdots,n,\ l=1,\cdots,\prod_{j\neq i} c_j,\ A_l^{-i}\in A^{-i}\right).$$

$$b_i=\left(\left(1-x_{-i}(A_l^{-i})\right)\cdot d_i;\ i=1,\cdots,n,\ l=1,\cdots,\prod_{j\neq i} c_j,\ A_l^{-i}\in A^{-i}\right).$$

Clearly, , $a_i=F_i(0)$, $b_i=F_i(1)$ (where $0\in X$, $1\in X$). It follows from the continuity of F (which is immediate) that $F(X)=\times_i[a_i,b_i]$, i.e., F maps the compact, convex set X to a compact and convex set, i.e., the rectangle $\times_i[a_i,b_i]$. For $d<<\bar{d}$, where $\bar{d}$ is given by (B.1), $a_i<0$ and $b_i>0$. Consequently, when $d<<\bar{d}$, there exists $y\in X$ such that $F(y)=0$. But $F(y)=0$ proves the lemma. □

Proof of Lemma 7

Let $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$ and $\psi\equiv\psi(r,\sigma)$. Recall from (S3) that $V_i(t_i|0)$ and $V_i(t_i|1)$ denote, respectively, the expected utility levels of type t_i if he votes against or for μ^1 at the voting stage. From (S3), it is true that $\forall i$, $\forall t_i$,

(C.1) $$U_i(\psi|t_i)=r_i(t_i)\cdot V_i(t_i|1)+(1-r_i(t_i))\cdot V_i(t_i|0)$$

$$=max\left\{V_i(t_i|0),V_i(t_i|1)\right\}.$$

By assumption (μ^0 does not strongly endure μ^1) there exist an equilibrium and a type t_i for which

(C.2) $$U_i(\psi|t_i)>U_i(\mu^0|t_i).$$

Consider now the voting game between μ^0 and ψ. Consider the following assessment $(\hat{r},\hat{\sigma},\hat{q})$. $\forall i$, $\forall t_i$, $\forall \hat{t}_i$, $\forall t_{-i}$,

(C.3) $$\hat{r}_i(t_i)=1.$$

(C.4.1) $$\hat{\sigma}_i(t_i|t_i,\psi)=1.$$

(C.4.2) $$\hat{\sigma}_i(\hat{t}_i|t_i,(1,\mu^0))=r_i(t_i)\cdot\sigma_i(\hat{t}_i|t_i,(1,\mu^0))+(1-r_i(t_i))\cdot\sigma_i(\hat{t}_i|t_i,(0,\mu^0)).$$

(C.4.3) $\hat{\sigma}_i(\hat{t}_i|t_i,(0,\mu^0))=\sigma_i(\hat{t}_i|t_i,(0,\mu^0))$.

(C.5.1) $\hat{q}_i(t_{-i}|t_i,\mu^1)=p_i(t_{-i}|t_i)$.

(C.5.2) $\hat{q}_i(t_{-i}|t_i,(1,\mu^0))=q_i(t_{-i}|t_i,(1,\mu^0))$.

(C.5.3) $\hat{q}_i(t_{-i}|t_i,(0,\mu^0))=q_i(t_{-i}|t_i,(0,\mu^0))$.

where r, σ and q always refer to the assessment $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$.

I claim that $(\hat{r},\hat{\sigma},\hat{q})\in\Phi(\mu^0,\psi)$. From Lemma 2 (ii), $\psi=(T,g)$ where g is given by (3.1)-(3.3) is incentive compatible given the interim beliefs $p_i(t_{-i}|t_i)$. Consequently, by (C.5.1), telling the truth is an equilibrium strategy when ψ is played. This is the content of (C.4.1). Because of (C.3), the beliefs when the history is $h=\psi$ in (C.4.1) are compatible with Bayes' rule.

Suppose that type t_i deviates at the voting stage and votes against ψ. In the resulting subform, if the players who did not deviate use their strategies as given by (C.4.2) and the player who deviated uses the strategy given by (C.4.3), then, because $(r,\sigma,q)\in\Phi(\mu^0,\mu^1)$, it follows that $\hat{\sigma}(\cdot|\cdot,(\cdot,\mu^0))$ is an equilibrium because it was an equilibrium in the game between μ^0 and μ^1.

When t_i deviates and vetoes ψ, then his expected utility is equal to $V_i(t_i|0)$. (Indeed, with beliefs and strategies given by (C.3)-(C.5.3), if t_i vetoes ψ in the present game, his expected utility is the same as if he vetoes μ^1 in the first game.) If t_i does not deviate, his expected utility is equal to $U_i(\psi|t_i)$, which by (C.1) is greater than $V_i(t_i|0)$. Consequently, it is indeed a (possibly weak) best response for each type of each player to vote for ψ if all the other players are expected to vote for ψ with probability one. Thus, $(\hat{r},\hat{\sigma},\hat{q})$ is a Bayesian Nash equilibrium of the voting game between μ^0 and ψ.

In order to prove that $(\hat{r},\hat{\sigma},\hat{q})\in\Phi(\mu^0,\psi)$, it is necessary to show that the beliefs $\hat{q}$ are consistent. The only problem is to show consistency of the beliefs in (C.5.2).

((C.5.3) is trivially satisfied since the posterior beliefs correspond to the prior beliefs if the player votes for the status-quo.)

First, suppose that $\sum_{\hat{t}_i} p_i(\hat{t}_{-i}|t_i)\cdot(1-r_{-i}(\hat{t}_{-i}))\neq 0$. From Lemma 6, for any integer k, there exists n positive real numbers d_i^k and a belief structure $\{r_i^k(t_i)\}$ satisfying,

(C.6.1) $\quad \forall i,\ \forall t_{-i},\ \prod_{j\neq i} \hat{r}_j^k(t_j)\equiv 1-\left(1-\prod_{j\neq i} r_j(t_j)\right)\cdot d_i^k.$

(C.6.2) $\quad \forall j,\ \forall t_j,\ r_j^k(t_j)\in[0,1].$

For any i, choose a decreasing positive sequence $\{d_i^k\}$ such that $\forall i$, $lim_{k\uparrow\infty} d_i^k=0$. Then, from Lemma 6, there exist corresponding sequences $\{\hat{r}_i^k(t_i)\}$ satisfying (C.6.1)-(C.6.2). By assumption, the right hand side of (C.6.1) has limit 1. Consequently, each term $\hat{r}_i^k(t_i)$ must also have 1 for limit. (Otherwise, the left hand side of (C.6.1) has a limit strictly bounded below 1, but then, for some d_i^k, (C.6.1) is violated, which contradicts Lemma 6.) Thus, $\forall t_i$, $lim_{k\uparrow\infty}\hat{r}_i^k(t_i)=1$. Finally, observe that for any k, the posterior beliefs of player t_i are

$$\hat{q}_i^k(t_{-i}|t_i,(0,\mu^0))=p_i(t_{-i}|t_i)$$

$$\hat{q}_i^k(t_{-i}|t_i,(1,\mu^0))=\frac{p_i(t_{-i}|t_i)\cdot(1-\hat{r}_{-i}^k(t_{-i}))}{\sum_{\hat{t}_{-i}} p_i(\hat{t}_{-i}|t_i)\cdot(1-\hat{r}_{-i}^k(\hat{t}_{-i}))}$$

(Because of (C.6.1))

$$=\frac{p_i(t_{-i}|t_i)\cdot(1-r_{-i}(t_{-i}))}{\sum_{\hat{t}_{-i}} p_i(\hat{t}_{-i}|t_i)\cdot(1-r_{-i}(\hat{t}_{-i}))}$$

(since the term in d_i^k disappears)

$$=q_i(t_{-i}|t_i,(1,\mu^0)).$$

Finally, consider the situation in which $\sum_{\hat{t}_i} p_i(\hat{t}_{-i}|t_i)\cdot(1-r_{-i}(\hat{t}_{-i}))=0$ for some t_i. Then, in the sequential equilibrium (r,σ,q), the beliefs $q_i(t_{-i}|t_i,(1,\mu^0))$ must have been computed by taking the limit of the posteriors obtained by perturbing the initial voting strategies. (See the paragraph after (S43).) In this case, it is possible to obtain a sequence of perturbed beliefs for the new game by applying Lemma 6 to the initial perturbed beliefs. (I.e., if $\{r_i^k\}$ is the initial sequence of perturbed beliefs that is used to compute $q_i(t_{-i}|t_i,(1,\mu^0))$, use Lemma 6 in order to find a sequence $\{\hat{r}_i^k\}$; this last sequence will have the desired property that (C.5.2) is satisfied and that $\hat{r}_i^k \uparrow 1$. This argument has already being used in the proof of Theorem 1 (i).)

This concludes the proof. □

Proof of Theorem 3

The proof is simple but long. Figure 3.5 might be helpful (in this figure, $\overline{\mathfrak{U}}$ is the heavy line, the vertices v and v' are irregular with respect to $\overline{\mathfrak{U}}$). I introduce the following sets.

$$(u,w) \equiv \{v | \exists \lambda \in (0,1),\ v = \lambda \cdot w + (1-\lambda) \cdot u\}.$$

$$ray(u,w) \equiv \{v | \exists \gamma > 0,\ v = \gamma \cdot w + (1-\gamma) \cdot u\}.$$

Let $\mu^0 \in E_I$ and let $u^0 = U(\mu^0)$. (i) of Theorem 3 is true by Lemma 1. Suppose that μ^0 is strongly durable and that there exists $\sigma \in E(\mu^0,p)$ such that $u(\mu^0,\sigma) \neq U(\mu^0)$. From (i), it must be true that $u(\mu^0,\sigma) < U(\mu^0)$. Suppose, contrary to the assumption, that there exist two indices i and j such that $\forall k \in \{i,j\}$, $u(\mu^0,\sigma)_k < U(\mu^0)_k$. I denote an element of the open segment $(u(\mu^0,\sigma),u^0)$ by u^λ, where $\lambda \in (0,1)$ is such that $u^\lambda = \lambda \cdot u^0 + (1-\lambda) \cdot u(\mu^0,\sigma)$. To simplify the notation, I will denote the component of an interim payoff vector u by u_i instead of $u(t)$.

Case A: $u(\mu^0,\sigma) << u^0$. By assumption, $\overline{\mathfrak{U}} \neq \{u^0\}$, and it is possible to choose $\epsilon > 0$

small enough such that there exists $u^1 \in N_\epsilon(u^0)$, $u^1 \neq u^0$, $u^1 >> u(\mu^0,\sigma)$. By Lemma 7, μ^0 is not strongly durable.

Case B: $\sim[u(\mu^0,\sigma) << u^0]$. That is to say, there exists at least one index i for which $u(\mu^0,\sigma)_i = u_i^0$.

I define the dimension of the polytope $\mathfrak{U}$, and I write $dim\,\mathfrak{U}$, to be the dimension of the flat of smallest dimension which contains $\mathfrak{U}$. I will say that $\mathfrak{U}$ is of full dimension if $dim\,\mathfrak{U}$ is equal to the dimension of the underlying space ($\sum_i |T_i|$ here) I will need the following result.

CLAIM: (*Under the assumptions of Theorem* 3.) *Let* $u \in relint\,\mathfrak{U}$. *Let* $u' \in ray(u,u^0)$ *such that* $u' > u$. *Then there exists* $w \in relint\,\mathfrak{U}$ *such that* $w > u$ *and* $w \notin ray(u,u^0)$.

Proof of the claim: Let $\{H^k = [h^k,\alpha^k]\}$ be a collection of hyperplanes that define the proper—i.e., different from $\emptyset$ or $\mathfrak{U}$—faces of $\mathfrak{U}$. I.e., if F is a face of $\mathfrak{U}$, then there exists a unique hyperplane H^k in the collection for which $F = H^k \cap \mathfrak{U}$. Two cases are of interest.

(i) $\mathfrak{U}$ is of full dimension. Then $\mathfrak{U}$ cannot be a subset of a hyperplane and $u \in relint\,\mathfrak{U} \Leftrightarrow \forall k, h^k(u) < \alpha^k$. Consider a vector $\beta >> 0$ such that $max_k\, h^k(\beta) < min_k\left(\alpha^k - h^k(u)\right)$. Such a choice is possible since each function h^k is continuous and since $h^k(0) = 0$. It follows that $w = u + \beta$ has the property that $\forall k$, $h^k(w) = h^k(u) + h^k(\beta) < \alpha^k$. Thus, $w \in relint\,\mathfrak{U}$. Observe that here $w >> u$.

(ii) $\mathfrak{U}$ is not of full dimension. Then, there exists a hyperplane $H = [h,\alpha]$ such that $\mathfrak{U} = H \cap \mathfrak{U}$. Consequently, $\forall u \in \mathfrak{U}$, $h(u) = \alpha$ and $\forall u \in relint\,\mathfrak{U}$, $\forall k$, $h^k(u) < \alpha^k$. By assumption, there exists $u'>u$ where $u'\in ray(u,u^0)$ such that the set $I=\{i|\, u_i=u'_i\}$ is nonempty. Observe that

(D.1) $\qquad v\in ray(u,u^0)\Rightarrow \forall\, i\in I,\ v_i=u_i=u'_i.$

The functional h can be written as $h(\beta)=\sum_i a_i\cdot\beta_i$, where the a_i are scalars. If $\exists i\in I$, $a_i=0$, consider β such that $\beta_i>0$ and $\forall j\neq i$, $\beta_j=0$. Then, $w=u'+\beta>u'>u$ and from (D.1), $w\notin ray(u,u')$. As in (i), it is possible to choose β such that $u'+\beta\in\mathcal{U}$. Suppose now that $\forall i\in I$, $a_i\neq 0$. If there exists $j\notin I$ such that $a_j\neq 0$, then choose β such that $\beta_j>u_j-u'_j$ and $\beta_j=-a_i\cdot\beta_i/a_j$, where $\beta_i>0$, for some $i\in I$ and let $\beta_k=0$ otherwise. Such a choice is clearly possible since $u_j-u'_j<0$. It follows that $h(\beta)=0$ and by choosing β_i small enough one can have $\forall k$, $h^k(u'+\beta)<\alpha^k$, i.e., $u'+\beta\in relint\,\mathcal{U}$.

Finally, suppose that for any $i\notin I$, $a_i=0$. (i1) If there exist $i,j\in I$ such that $a_i>0$ and $a_j<0$, then by choosing $\beta_i>0$, $\beta_j>0$ such that $a_i\cdot\beta_i+a_j\cdot\beta_j=0$, and $\beta_k=0$ for $k\notin\{i,j\}$, we have $h(\beta)=0$. β can always be chosen in such a way that $u'+\beta\in relint\,\mathcal{U}$. (i2) Either $\forall i\in I$, $a_i>0$ or $\forall i\in I$, $a_i<0$. By assumption, $\#\{i|\, i\notin I\}\geq 2$. Let for some $i\notin I$, $\beta_i>0$, and $\beta_k=0$ if $k\neq i$. In this case, $u'_i+\beta_i>u'_i$ and $u'_j+\beta_j=u'_j$. By definition of u' there exists a unique $\lambda\in(0,1)$ for which $\forall j$, $u'_j=\lambda\cdot u_j+(1-\lambda)\cdot u^0_j$, and consequently, $u'_i+\beta_i>\lambda\cdot u_i+(1-\lambda)\cdot u^0_i$ while $\forall j\neq i$, $u'_j+\beta_j=\lambda\cdot u_j+(1-\lambda)\cdot u^0_j$. Thus, $u'+\beta\notin ray(u,u^0)$.

Observe that if $\#\{i|\, i\notin I\}=1$, then in (i2) of the proof, for any $w>u$, it must be true that $\forall i\in I$, $w_i=u_i$. Since there is only one index j such that $j\notin I$, whenever $w_j>u_j$, there exists $\lambda\in(0,1)$ such that $w_j=\lambda\cdot u_j+(1-\lambda)\cdot u^0_i$. Thus, the only vector in $\overline{\mathcal{U}}$ which dominates u is u^0. □

Figure (3.6) illustrates a situation in which condition (ii) of Theorem 3 is violated: there, the only point of $\overline{\mathcal{U}}$ which dominates u^{λ} is u^0. Theorem 3 is proved by a sequence of steps.

Step 1: There exists $\lambda \in (0,1)$ such that $u^{\lambda} \in relint\,\mathcal{U}$.

Let $\lambda(\epsilon)$ such that $d(u^{\lambda(\epsilon)}, u^0) = \epsilon$, where $\epsilon > 0$ is such that $N_{\epsilon}(u^0) \cap (relbd\,\mathcal{U} \setminus \overline{\mathcal{U}}) = \emptyset$. Because u^0 is regular with respect to $\overline{\mathcal{U}}$, such an ϵ exists. Suppose that for $\lambda > \lambda(\epsilon)$, $u^{\lambda} \notin relint\,\mathcal{U}$. Then, $u^{\lambda} \in relbd\,\mathcal{U}$ since $u^{\lambda} \in \mathcal{U}$ (by convexity of $\mathcal{U}$). By definition of ϵ and of $\lambda(\epsilon)$, $(u^{\lambda}, u^0) \subset N_{\epsilon}(u^0)$. But this contradicts the assumption that $N_{\epsilon}(u^0) \cap (relbd\,\mathcal{U} \setminus \overline{\mathcal{U}}) = \emptyset$. Observe that for any λ, $u^{\lambda} < u^0$ since $u(\mu^0, \sigma) < u^0$.

Step 2: Suppose that $dim\,\mathcal{U} = 1$. Then, either $\mathcal{U} = \overline{\mathcal{U}}$, in which case $u(\mu^0, \sigma) < u^0$ is impossible, or $\mathcal{U} \neq \overline{\mathcal{U}}$, in which case $\overline{\mathcal{U}}$ is a singleton, which contradicts the assumption of the theorem. Note that when $dim\,\mathcal{U} = 1$, the theorem follows immediately from Lemma 1.

Step 3: Suppose that $dim\,\mathcal{U} \geq 2$. From Step 1, there exists a neighborhood $N_{\delta}(u^{\lambda})$ such that $N_{\delta}(u^{\lambda}) \cap \mathcal{U}$ is relatively open in $\mathcal{U}$. From the Claim, it is possible to find $w \in \mathcal{U}$ such that $w \in N_{\delta}(u^{\lambda}) \setminus ray(u^{\lambda}, u^0)$ and such that $u^{\lambda} < w$. For any such w, $w > u(\mu^0, \sigma)$ since $w > \lambda \cdot u^0 + (1-\lambda) \cdot u(\mu^0, \sigma)$, and since $u^0 > u(\mu^0, \sigma)$ and $\lambda \in (0,1)$.

Step 4: Let w be one of the elements of $N_{\delta}(u^{\lambda})$ as defined in Step 3. Since a polytope is bounded, the ray $ray(u^{\lambda}, w)$ must intersect the relative boundary of $\mathcal{U}$. From Step 3, u^0 does not belong to this intersection. Let $u(\lambda)$ denote a possible element of this intersection for a particular choice of w, i.e., $u(\lambda) \equiv ray(u^{\lambda}, w) \cap relbd\,\mathcal{U}$, for some $w \in N_{\delta}(u^{\lambda}) \setminus ray(u^{\lambda}, u^0)$, $u^{\lambda} < w$. As λ goes to one, $\delta \to 0$, $u^{\lambda} \to u^0$ and $u(\lambda) \to u^{\lambda}$, for any choice of w. Consequently, there must exist $\lambda < 1$ such that $d(u^0, u^{\lambda}) < \epsilon/2$ and $d(u^{\lambda}, u(\lambda)) < \epsilon/2$, for any possible choice of w. From the triangle inequality,

$d(u^0,u(\lambda)) \leq d(u^0,u^\lambda)+d(u^\lambda,u(\lambda))<\epsilon$. Thus, $u(\lambda)\in N_\epsilon(u^0)$, for $\lambda<1$ large enough. By assumption, $N_\epsilon(u^0)$ contains no element of $relbd\,\mathfrak{U}\backslash\overline{\mathfrak{U}}$. Thus, $u(\lambda)\in\overline{\mathfrak{U}}$ for λ large enough. Consequently, there exists $u^1\in\overline{\mathfrak{U}}\backslash\{u^0\}$ such that $u(\mu^0,\sigma)<u^1$. By Lemma 7, μ^0 is not strongly durable since it does not strongly endure $f(u^1)$. This contradicts our assumption. □

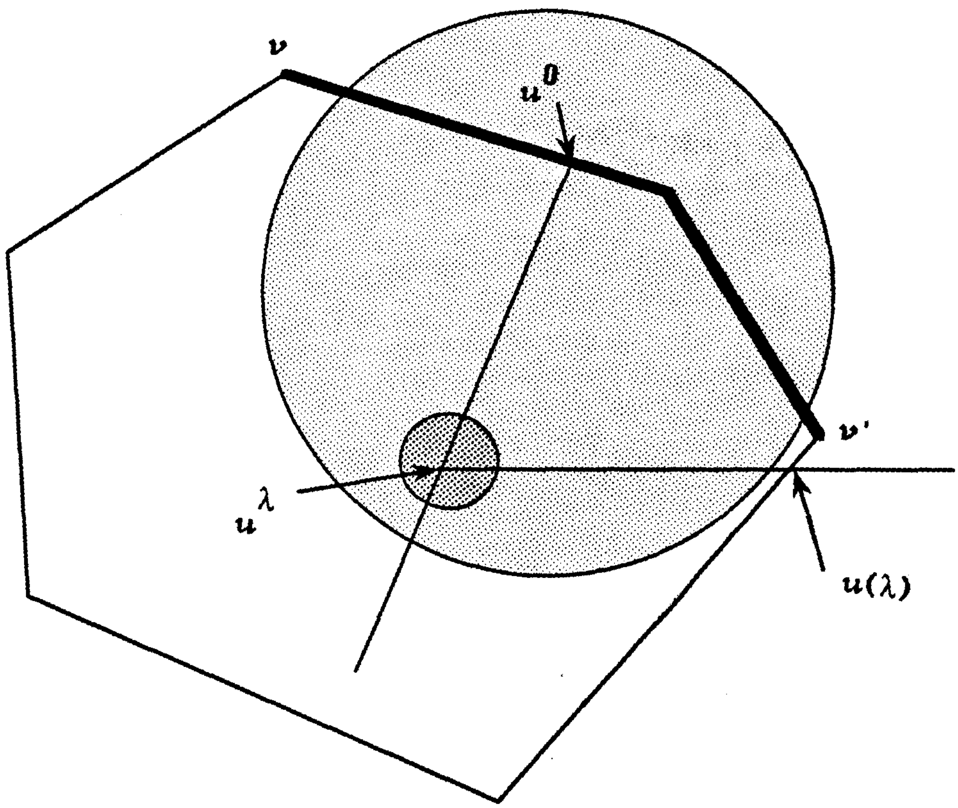

Figure 3.5

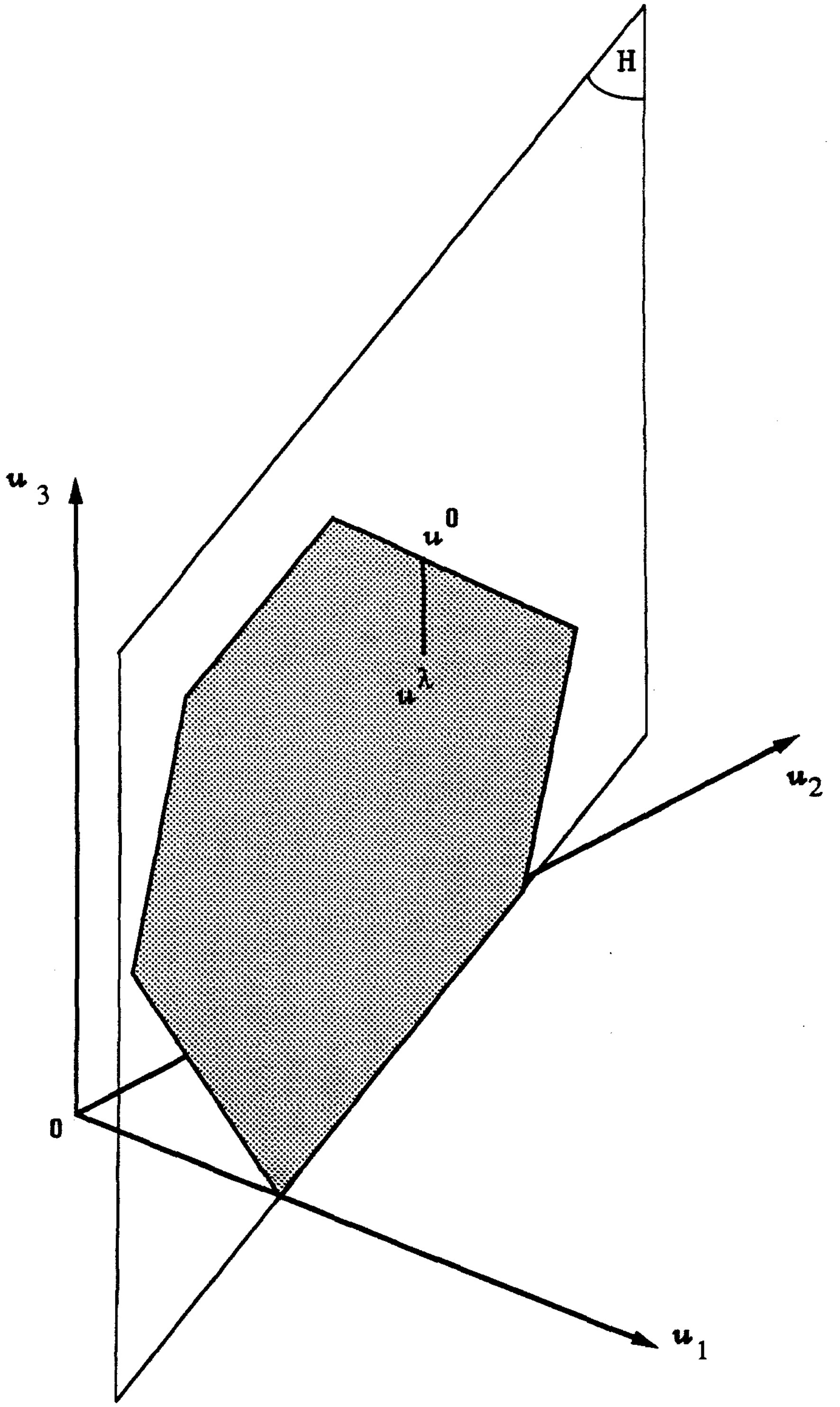

Figure 3.6

$[u \in H \Leftrightarrow u_1 + u_2 = \alpha]$

APPENDIX B

Convexity: Some definitions

Most of these definition are taken from Lay (1982).

A *polyhedral set* is the intersection of a finite number of closed half-spaces. A *polytope* is the convex hull of a finite set of points. The fact that a bounded polyhedral set is a polytope is immediate (e.g., see Lay (1982, Theorem 20.9)).

A translate of a linear subspace of an Euclidian space E^m is called a *flat.* A flat of dimension $m-1$ is called a hyperplane. The relative interior of S, *relint* S, is the interior of the set S relative to the minimal flat containing it. *relbd* S and *relcl* S, the relative boundary and the relative closure of S have similar definitions. For instance, in E^3, $I=[(0,0,0),(1,0,0))$ has an empty interior and the closure of I in E^3 is equal to its boundary, i.e., $cl\,I=bd\,I=[(0,0,0),(1,0,0)]$. The minimal flat containing I has dimension 1 and $relint\,I=\Big((0,0,0),(1,0,0)\Big)$, $relbd\,I=\Big\{(0,0,0),(1,0,0)\Big\}$ and $relcl\,I=cl\,I$. (It is always the case that the closure and the relative closure coincide.)

If H is a hyperplane in E^m, then there exists a linear functional h and a real number α such that $H=[h{:}\alpha] \equiv \{x\in E^m | h(x)=\alpha\}$. The notation $h(S)\leq\alpha$ stands for $\forall x\in S$, $h(x)\leq\alpha$ (similarly for $\geq$, $<$, $>$). If $S\subset E^m$ is a convex set, the hyperplane $H=[h,\alpha]$ *bounds* S if either $h(S)\geq\alpha$ or $h(S)\leq\alpha$ holds. A hyperplane H *supports* S at x if $x\in S$ and H bounds S.

If S is a compact and convex subset of some Euclidian space, a subset F of S is called a *face* if either $F=\emptyset$ or $F=S$, or if there exists a supporting hyperplane H of S such that $F=S\cap H$.

NOTES

1 This representation of a collective decision making problem is due to Matthews et al. (1989). The framed column on the left is the set of decisions that can be used. The order of the decisions in this column is the order that is used in the matrix on the right. For instance, in Example 1, if 1 is of type $1a$ and 2 is of type $2a$, then if decision A is chosen, $1a$ has a utility payoff of 2 while if the decision B is chosen, $1a$ has a utility payoff of 3. The notation for the (direct) mechanisms is consistent with this notation. For instance, μ^1 in Example 1 means that decision A is implemented if players 1 and 2 make announcements ($1a$,$2b$) or ($1b$,$2a$) and decision B is implemented otherwise.

2 The argument is similar to the consistency argument developed in the literature on renegotiation proof equilibria in repeated games. See, Pearce (1987), Farrell and Maskin (1987), Bernheim and Ray (1988).

3 Renegotiation of equilibria is used in the literature cited in note 2. It is worth noting that in that literature as well as in the present work, the *process* by which the players renegotiate to another equilibria is not modeled. But this is "consistent" with the fact that the process by which the players *choose* an equilibrium is not modeled.

4 Whether the restriction to mechanisms with finite message sets is or is not without loss of generality is an open question at this point.

5 It is sufficient to consider behavioral strategies. This was proven by Kuhn (1953) for extensive form games with finite length and finite number of moves at each node. The "Kuhn Theorem" was later extended by Aumann (1964) to infinite extensive form games with (possibly) a continuum of moves at each node.

6 For convenience, I use the following notation. If θ is a n-tuple of strategies then θ_{-i} denote the product $\prod_{j\neq i}\theta_j$. For instance, $\sigma_{-i}(s_{-i}|t_{-i},\mu^1)$ denotes the

product $\prod_{j\neq i}\sigma_j(s_j|t_j,\mu^1)$, $r_{-i}(t_{-i})$ denotes the product $\prod_{j\neq i} r_j(t_j)$, etc.

7 Observe that, $\forall i$, $\alpha(d|t;r)=r_i(t_i)\cdot\alpha(d|t;r\backslash r_i=1)+(1-r_i(t_i))\cdot\alpha(d|t;r\backslash r_i=0)$, where $r\backslash r_i=0$ ($r\backslash r_i=1$) denotes the vector of voting strategies in which player i votes against (for) μ^1 and all the other players play their original strategies. With the notation of (S2),

$$\alpha(d|t;r\backslash r_i=0)=\sum_{\hat{t}} K\Big(\sigma_i(\cdot|t_i,(0,\mu^0))\Big)(m)\cdot g^0(d|m)$$

and, $\alpha(d|t;r\backslash r_i=1)= \sum_{\hat{t}} K\Big(\sigma_i(\cdot|t_i,(1,\mu^0))\Big)(m)\cdot g^0(d|m)$.

8 The reader might have some difficulties to imagine that a mediator would propose the same mechanism as the status-quo. This is only for convenience as the following shows. If $\sigma\in E(\mu^0,p)$ and if there exists a type t_i such that $U_i(\mu^0|t_i)< U_i((\mu^0,\sigma)|t_i)$, let $\mu^1=(T,g^1)$, where $\forall d$, $\forall t$, $g^1(d|t)=\sum_{\hat{t}}\sigma(\hat{t}|t)\cdot g^0(d|t)$. Then, μ^1 is a different mechanism from μ^0. Suppose that the mediator proposes to vote between μ^0 and μ^1. Consider the sequential equilibrium in which each type votes for μ^1, each type reports truthfully in μ^1 and each type plays the strategy $\sigma_j(\hat{t}_j|t_j)$ in the mechanism μ^0. When μ^0 must be played, suppose that the players do not modify their prior beliefs (e.g., consider the sequence $\{r_j^k(t_j)=1/k\}$). This assessment forms a sequential equilibrium since each type is indifferent between voting for μ^1 and μ^0. The result now follows.

9 See note 7.

10 For $z,z'\in\mathbb{R}^m$, $z>>z'$ denotes $\forall j$, $z_j>z'_j$ while $z>z'$ denotes $\forall j$, $z_j\geq z'_j$ & $z\neq z'$.

11 Most of the terms related to convexity are defined in Appendix B.

12 It is immediate that $\overline{\mathfrak{U}}\subset relbd\,\mathfrak{U}$. In fact, $\overline{\mathfrak{U}}$ is the union of faces of $\mathfrak{U}$. This is a direct consequence of the definition of $\overline{\mathfrak{U}}$ and of the observation, e.g., Rockafellar (1970:162), that the set of points of $\mathfrak{U}$ where a certain linear function achieves its maximum over $\mathfrak{U}$ is a face of $\mathfrak{U}$.

13 G_1 is a uniformly incentive compatible mechanism (Holmström and Myerson

(1983:1814)). In fact, for every player $i \neq 1$, truth is a (weakly) dominant strategy.

14 Label the elements of the set T_1 as $\{t_1(1),\cdots,t_1(T_1)\}$. Find the set $T_1(1)$ of all the types t_1 such that $u_1(g(t_1),t_1)=u_1(g(t_1(1)),t_1)$ and define $g^1(t_1)=g(t_1(1))$ for such types in $T_1(1)$ and $g^1(t_1)=g(t_1)$ otherwise. Next, find the minimum index k for which $t_1(k_1)\notin T_1(1)$ and call this index k_1. Find the set $T_1(2)$ of all $t_1 \in T_1 \setminus T_1(1)$ such that $u_1(g(t_1),t_1)=u_1(g(t_1(k_1)),t_1)$ and define $g^2(t_1)=g^1(t_1(k_1))$ if $t_1 \in T_1(2)$ and $g^2(t_1)=g^1(t_1)$ otherwise. Define in the same manner the indices $k_2,\cdots$ and the sets $T_1(2),\cdots$ by following the previous construction. Clearly, by finiteness of the set T_1, this process must end and the resulting decision rule g' will have the property in the text.

REFERENCES

Aumann, Robert J. (1964). "Mixed and Behavior Strategies in Infinite Extensive Games," pp. 627-50 in *Advances in Game Theory*, Annals of Mathematics Studies No. 52. Princeton: Princeton University Press.

Bernheim, B. Douglas and Ray, Debraj (1987). "Collective Dynamic Consistency in Repeated Games." Mimeo.

Crawford, Vincent (1985). "Efficient and durable decision rules: a reformulation." *Econometrica*, 53(4), 817-835.

Farrell, Joseph and Maskin, Eric (1987). "Renegotiation in Repeated Games." D.P. 1335, Harvard University.

Hart, Oliver and Moore, John (1988). "Incomplete Contracts and Renegotiation." Mimeo, London School of Economics," *Econometrica*, 56(4), 755-85.

Holmström, Bengt and Myerson, Roger B. (1983). "Efficient and durable decision rules with incomplete information." *Econometrica*, 51(6), 1799-819.

Lay, Steven R. (1982). *Convex Sets and their Application.* New York, NY: John Wiley & Sons.

Kreps, David M. and Wilson, Robert (1982). "Sequential equilibrium," *Econometrica*, 50, 863-74.

Ledyard, John O. (1986). "The Scope of the Hypothesis of Bayesian Equilibrium," *Journal of Economic Theory*, 39, 59-82.

Matthews, Steven, Okuno-Fujiwara, Masahiro and Postlewaite, Andrew (1989). "Communication in Bayesian Games: Issues and Problems." Paper presented at the Caltech Theory Workshop.

Myerson, Roger B. (1983). "Mechanism design by an informed principal." *Econometrica*, 51, 1767-97.

Palfrey, Thomas R. and Srivastava, Sanjay (1987). "On Bayesian Implementable

Allocations,," *Review of Economic Studies*, 54, 193-208.

Palfrey, Thomas R. and Srivastava, Sanjay (1989). "Implementation with Incomplete Information," *Econometrica*, 57(1), 115-34.

Pearce, Douglas G. (1987). "Renegotiation-Proof Equilibria: Collective Rationality and Intertemporal Cooperation." D.P. 855, Cowles Foundation.

Rockafellar, R. Tyrrell (1970). *Convex Analysis.* Princeton: Princeton University Press.

Schelling, Thomas C. (1960). *The Strategy of Conflict.* Cambridge, MA: Harvard University Press.

Selten, Reinhart (1976). "A Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4(1), 25-55.