SOCIAL INTERACTION UNDER TRULY PERFECT INFORMATION

By

Earl A. Thompson
University of California, Los Angeles

and

Roger L. Faith
Virginia Polytechnic Institute and State University

Discussion Paper Number 175
August 1980

SOCIAL INTERACTION UNDER TRULY PERFECT INFORMATION

Earl A. Thompson, UCLA, and

Roger L. Faith, VPI&SU*

INTRODUCTION

Classical sociologists have long been distinguished from political philosophers, economists, and other social thinkers by their insistence that orderly social interactions require certain individuals to be governed by "moral principles" or "social values" guiding their responses to others and thereby enforcing the norms of the group. While most social thinkers, including most modern exchange theorists, regard moral principles as merely an incidental part of social reality, the authors of the various classics in sociology -- esp. Durkheim, Weber, and Parsons -- regarded moral principles to be somehow necessary to the existence of social norms and the creation of order in a Hobbesian jungle. While their views emerged largely out of critical examinations of empirical phenomena, modern interaction theory provides theoretical support for their observation. The consideration leading to this theoretical support is simple: While rewards or punishments are generally necessary to induce certain individuals to behave in the social interest, providing these rewards or punishments -- i.e. "enforcing the social contract" -- requires responses that are ultimately irrational and therefore forthcoming only when some higher moral principle governs

---

*The authors benefited substantially from discussions with Armen Alchian, Ron Batchelder, Phil Bonacich, Jack Hirshleifer, David McFarland, Louis Makowski, Joe Ostroy, and Lloyd Shapley.

the response.[1]

While a moral principle or ethical code allows an individual to carry out behavior that, at the time of its occurrence, is irrational, there is a prior calculation determining the choice of a particular ethical code. At this level, the choice is not simply a choice among various possible actions. It is a choice among various possible principled reactions to the actions of others, each reaction valued in large part for its effect on the subsequently chosen actions of others. Yet we can find no precise statement of this optimality problem in the literature. Sociologists have apparently not developed their theories of moral principles sufficiently to endogenously derive moral principles and the corresponding role sets and norms from a theory of individually rational choice. As a result, they have not provided us with a theory from which we can derive the nature of social institutions and social behavior from basic technological data.

This paper provides a first step in closing this gap in the literature by constructing a "pure" theory of social institutions, a theory stripped of all informational imperfections in order to expose certain underlying

---

[1]The game-theoretic foundation of such arguments have become increasingly clear in recent years (see, e.g., Rapoport-Chammah, McFarland, Selten). Consider a world containing a known future date beyond which social interaction will not occur. Devoting resources to meting out rewards or punishments on this last possible date can not be narrowly rational because there would be no future behavior to influence and the past behavior, having already occurred, is water over the dam. It would be a waste of time and trouble to reward or punish anyone. Therefore, on the next-to-last conceivable date at which socially relevant actions are chosen, behavior is undisciplined and thoroughly noncooperative. But then it doesn't pay to devote resources to punishing or rewarding behavior in the second-to-last period either because behavior in subsequent dates will be noncooperative anyway. The argument thus proceeds backward in time to the present so that it never pays to cooperate when individuals are rational in the above, narrow sense.

tendencies characterizing such systems. The primary theoretical tendencies we shall derive are very similar to those found in the relatively empirical writings of Weber and, more recently, of numerous sociologists and game theorists. In particular, we find that under perfect information regarding the reaction functions of others: (1) there is a hierarchy of decision makers (Section I); (2) if everyone has strict preferences over the entire set of social alternatives, the decision hierarchy -- once formed -- will produce economically efficient roles, norms and social actions (Section II); but (3) substantial resource losses occur in the stratification process (Section III).

These primary efficiency results suggest that a leader, once established, will attempt to impose systems that reduce the economic wastes due to further stratification without sacrificing the efficiency of the institutions evolved by the lower members of the strata. Along these lines, we shall find that a private property system has this characteristic. That is, a private property system reduces the resource wastes involved in middle-level stratification while still allowing for the creation of efficient, middle-level roles and norms (Section IV). Private property institutions are thus seen to represent systems that reduce the amounts of resources devoted to stratification without sacrificing the natural ability of a "free," or unregulated, social system to develop efficient social institutions.

While private property systems do not deter the achievement of efficient roles and norms by subgroups in the system, when norm selectors are indifferent between more than one solution norm, a special problem arises that requires higher-level intervention in order to assure efficiency. After showing this, we apply the result to an actual historical institution --

slavery -- in order to test some of the detailed workings of the general model (Section V).

Our secondary efficiency results help explain the survival of hierarchal organizations within private property systems despite their extra stratification costs -- and despite the perennial attacks on their efficiency -- while our primary efficiency results will be seen to offer an explanation for the survival of hierarchial social systems in nature at large (Section VI).

In terms of game theory, our general approach is unusual in that it explicitly models strategic communication. Standard noncooperative game theory does not permit such communication; it allows only the communication of simple actions, leaving the communication of principled reaction functions, i.e., strategies, to be handled by something called "cooperative game theory." Unfortunately, however, cooperative game theory has never developed an explicit communication structure and consequently has not provided us with a consistent theory of how individuals cooperate (Section VII).

Our particular model of strategic communication is developed for the special case of an unrestricted communication of (and hence, perfect information regarding) the strategies of others. As conventional "perfect information" games have each player perfectly communicating only his actions while our game has perfect communication of both actions and strategies, we label our assumption, for lack of a better term, "truly perfect information." A basic feature distinguishing our game under truly perfect information from cooperative games and "supergames," both of which are designed to allow sufficient communication to prevent Pareto nonoptimal solutions, is that we do not assume Pareto optimality to be a characteristic of solutions or of points on rational reaction functions. We even find an exception (the case containing non-strict preferences discussed in Section V) in which Pareto nonoptima may result no matter how perfect the information structure. Also, our solution set, being based on an explicit model of communication of strategies, does not have the problem, chronic in cooperative game theory, of being either empty or too large to be of much practical interest.

Conventional game theory's treatment of the communication of strategies has also been noted and criticized, albeit indirectly, by Schelling [1960] and Howard [1971]. Our contributions relative to Schelling's seminal work on 2-person bargaining are: (1) To derive Schelling's game from a prior specification on the information structure; (2) To generalize it to n-players and (3) most important, to characterize the resulting general solutions as to their Pareto optimality and, correspondingly, to apply the model to Pareto optimality observed social institutions. The fact that these extensions of Professor Schelling's work on communicated strategies have not been heretofore developed is perhaps because Schelling did not properly contrast his implied model with more conventional games. In particular, he failed to note that he was merely applying the standard, von Neumann-Morgenstern, perfect information solution concept to strategies rather than actions (or "plays of a game"). While von Neumann and Morgenstern [1944, Sec. 11.3] explicitly recognized that games could be constructed in which strategies are communicated in the same way as the actions in their perfect information games, they saw nothing novel about such games. For such games posed no new problem in the development of solution concepts or the existence of solutions. We conjecture that had they been more interested in evaluating the Pareto optimality of solution actions, or in formally capturing the microsociology of social institutions, they would have devoted more intellectual resources to games with perfect information concerning strategies as well as actions. But von Neumann and Morgenstern also expressed rather serious doubts about having players rely on the rationality of others, a reliance required by their perfect information solution concept. Their argument supporting these doubts[2] is that it may pay a player to deviate from "rational" responses if he knows that another player's strategy depends on his responses. But it is precisely <u>these</u> deviations which are at the heart of any theory of perfect strategy communication. For example, in Schelling's two-person bargaining problem, the first strategy selector is that player who can first prevent himself from following his narrowly rational responses to the actions of the player and communicate the fact to the other player. Von Neumann and Morgenstern did not see that their justifiable skepticism with respect to their perfect information game leads towards the development of games with truly perfect information rather than towards the imperfect information games which they so elegantly explored.

Professor Howard's work, developed as a generalization of von Neumann-Morgenstern's majorant-minorant game, has strategies contingent on strategies, thus apparently implying perfect information of strategies. But Howard employs a conventional, Nash-type, solution set, where a prior strategy selector takes <u>as given</u> the strategies of subsequent strategy selectors. This is not generally consistent with the model we are presenting here in which rational prior strategy selectors can communicate their strategies to

---

[2]Von Neumann and Morgenstern [1944, Sec. 4.1.2].

subsequent strategy selectors. The result is a solution set which lacks the powerful optimality and distributional characteristics of the solution set of a game with truly perfect information (Section I).[3]

[3]The following two sections are almost identical to the first two sections of Thompson-Faith (1980). The reader familiar with this material are urged to move on to Section III.

I. THE BASIC MODEL AND ITS SOLUTIONS

A. The Physical Environment.

An individual is denoted i, i = 1,...,n. An action of individual i is denoted x_i, where $x_i \in X_i$, a finite set of feasible actions of individual i. A possible social action is defined by an n-dimensional set of individual actions, and is denoted $x = (x_1, x_2, \ldots, x_n)$, so that $x \in \prod_{i=1}^{n} X_i$. To describe individual preferences, each individual, i, is given a complete, transitive, irrelexive, antisymmetric, binary relation, $\succ_i$, defined over $\prod_{i=1}^{n} X_i$. This description, in effect, assumes away indifference between any pair among the finite set of possible social actions. The motivation for this assumption and the effects of indifference on our central results will be discussed later. A Pareto optimum is a social action, x' , $x' \in \prod_{i=1}^{n} X_i$, for which there is no alternative, social action x", $x'' \in \prod_{i=1}^{n} X_i$, such that $x'' \succ_i x'$ for all i. Several Pareto optima may exist.

B. Institutional Possibilities under Truly Perfect Information.

The institutions facing an individual can be completely described by the reactions of other individuals to his own actions. But institutions, or reactions, are not taken here as given; they are derived. This is done by allowing each individual to select, among all feasible reaction functions, a function which is maximal with respect to his preference relation. But we want individuals to *know* the institutions and thus the reaction functions of others. And for this to generally hold, the functions must be communicated in sequence. Thus, for the individuals to know the institutions, i.e., for truly perfect information, the first communicator, say individual 1, presents the reaction function,

$$x_1 = f_1(x_2, \ldots, x_n), \qquad (1)$$

to the other individuals; the second communicator, say individual 2, then presents

$$x_2 = f_2(x_3, \ldots, x_n) \qquad (2)$$

to individuals 3 through n; the third communicator then presents

$$x_3 = f_3(x_4, \ldots, x_n) \qquad (3)$$

to individuals 4 through n, and so on up to the n - $1^{\underline{st}}$ communicator, who presents

$$x_{n-1} = f_{n-1}(x_n) \qquad (4)$$

to the $n^{\underline{th}}$ individual, who has no need to communicate. Once the action of the $n^{\underline{th}}$ individual is taken, the action of the n - $1^{\underline{st}}$ individual is determined. Once this pair of actions is taken, the action of individual n - 3 is determined, and so on up until a social action is determined as a chain reaction from the $n^{\underline{th}}$ individual's action. The set $(f_1, f_2, \ldots, f_{n-1})$ is thus a complete institutional description. The feasible choice set, or strategy set, of individual 1 is the set of <u>all</u> functions from $\prod_{i=2}^{n} X_i$ to X_1. This can be represented by the functional variable, F_1. Similarly, $F_2, \ldots, F_{n-1}$ can be used to represent the respective strategy sets of individuals 2 to n-1. The product space, $\prod_{i=1}^{n-1} F_i$, thus represent the world's institutional possibilities. The strategy set of individual n is X_n.

A question may arise as to why some individuals do not present reaction functions to other individuals who are higher up in the communication hierarchy. Consider individual n. Facing the prior strategies of the other n - 1 individuals, he sees that the eventual social action must be consistent with the chosen reaction functions of each of the n - 1 prior selectors. Hence, if individual n responds to the prior selectors with a simple action, he will have a free choice over all social actions consistent with the prior reaction functions. But if n responds with a function of prior actions, thus giving further choices to the prior strategy selectors, he can only reduce his original choice out of the same set of possible social actions. He cannot expand the set of possible outcomes because any eventual outcome must be consistent with the given n-1 reaction functions.

Similarly, if the n-1st strategy selector presents a reaction function rather than an action to his prior strategy selectors for a given action of individual n, he is giving them the choice of actions consistent with the set of reaction functions he faces and thus can be no better off. This also applies, in like fashion, to individuals n-2 to 2, so that it is in no individual's interest to present a reaction function to a prior strategy selector.

The above world, which can now be viewed as a "game," differs from the standard, von Neumann-Morgenstern, "perfect information" games in that some individuals are allowed to communicate their strategies to others before the latter select their own strategies. Thus, in the von Neumann-Morgenstern world, a player will not adopt a special strategy in order to influence the subsequent strategies (and actions) of others simply because he cannot communicate it and therefore cannot use it to influence the subsequently chosen strategies. In contrast, in the above world, each of the first n-1 players communicates his strategy to all subsequent strategy selectors. And response strategies of the subsequent selectors are known a priori by the prior strategy selector because they are the rational responses to the given strategy of the prior selector.

While Howard (4) has produced a general class of games (called "jk-metagames") containing strategies contingent on the strategies of other strategy selectors, he does not assume truly perfect information. Correspondingly, he does not adopt a perfect information solution concept. Rather, he adopts, without substantive justification, the von Neumann-Morgenstern-Nash "no-regret" solution concept in which each strategy selector accepts <u>as given</u> the strategies of all other strategy selectors. This amounts, as Howard recognizes, to assuming uniformly zero information regarding the strategies of others at the time of strategy selection. For if the choice of strategy selector were perceived by subsequent strategy selectors, it would, in general, influence the latter's selections. Such games, besides being theoretically unsatisfying in that they typically generate a multiplicity of solution points, some of which are optimal and others nonoptimal (Howard, p. 58), are empirically unsatisfying in that observed commitments are, as pointed out in the Introduction, typically communicated to others in order to influence their strategy selections.

It may be convenient to think of this problem as one in classical sociology: The individuals are prestratified according to their "power," or ability to influence others. The individuals above the lowest stratum can "punish" certain "deviant" behavior of others, thereby serving as members of "reference groups" who induce others to conform to certain "norms," i.e., to adopt particular actions. The middle-status individuals are in turn punished by higher-ups in ways that induce them to adopt certain "roles," i.e., to select certain reaction functions. After defining solution norms and roles, we shall look for certain properties of these solutions. Due largely to our previous work in economics, we look for economic efficiency, or Pareto optimality, properties. After that we inquire into the process of stratification and the efficiency problems that arise therefrom. The last few sections of the paper deal with the efficiency properties of various extensions of the basic model. Because we wish to avoid terminological squabbles, and also because we are hoping for an audience beyond sociology, we shall maintain our rather neutral terminology instead of adopting that of classical sociology.

C. Equilibrium Institutions, or "Solutions", under Truly Perfect Information.

A solution, $(f^*_1, \ldots, f^*_{n-1}, x^*_n)$, is a set in which the ith variable is maximal with respect to $\succ_i$ for given values of $f_1, \ldots, f_{i-1}$. A solution can be constructed as follows: First, we find, for individual n, x_n^*, the point in X_n such that, for all $x_n \neq x_n^*$,

$$\{f_1(f_2,\ldots,f_{n-1},x_n^*),\ f_2(f_3,\ldots,f_{n-1},x_n^*),\ldots,x_n^*\} \succ_n$$
$$\{f_1(f_2,\ldots,f_{n-1},x_n),\ f_2(f_3,\ldots,f_{n-1},x_n),\ldots,x_n\}.$$

This solution determines a dependency of x_n^* on $f_1,f_2,\ldots,f_{n-1}$, which we write $x_n^*[f_1,\ldots,f_{n-1}]$. Then, for individual n-1, we find a reaction function, f_{n-1}^*, such that for all $f_{n-1} \in F_{n-1}$, $f_{n-1} \neq f_{n-1}^*$,

$$\{f_1(f_2,\ldots,f_{n-2},f_{n-1}^*,x_n^*[f_1,\ldots,f_{n-2},f_{n-1}^*]),\ldots,f_{n-2},f_{n-1}^*,x_n^*[f_1,\ldots,f_{n-2},f_{n-1}^*]\} \succ_{n-1}$$
$$\{f_1(f_2,\ldots,f_{n-2},f_{n-1},x_n^*[f_1,\ldots,f_{n-2},f_{n-1}]),\ldots,f_{n-2},f_{n-1},x_n^*[f_1,\ldots,f_{n-2},f_{n-1}]\}$$

This solution determines the dependency of f_{n-1}^* on $f_1,f_2,\ldots,$ and f_{n-2}, which we describe as $f_{n-1}^*[f_1,\ldots,f_{n-2}]$. Then, for individual n-2, we find a reaction function, f_{n-2}^*, such that, for all $f_{n-2} \in F_{n-2}$, $f_{n-2} \neq f_{n-2}^*$,

$$\{f_1(f_2,\ldots,f_{n-2}^*,f_{n-1}^*[f_1,\ldots,f_{n-2}^*],x_n^*[f_1,\ldots,f_{n-2}^*,f_{n-1}^*[f_1,\ldots,f_{n-2}^*]]),$$
$$\ldots,f_{n-2}^*,\ f_{n-1}^*[f_1,\ldots,f_{n-2}^*],x_n^*[f_1,\ldots,f_{n-2}^*,f_{n-1}^*[f_1,\ldots,f_{n-2}^*]]\} \succ_{n-2}$$
$$\{f_1(f_2,\ldots,f_{n-2},f_{n-1}^*[f_1,\ldots,f_{n-2}],x_n^*[f_1,\ldots,f_{n-2},f_{n-1}^*[f_1,\ldots,f_{n-2}]]),$$
$$\ldots,f_{n-2},\ f_{n-1}^*[f_1,\ldots,f_{n-2}],x_n^*[f_1,\ldots,f_{n-2},f_{n-1}^*[f_1,\ldots,f_{n-2}]]\}.$$

This solution thus determines the dependency of f_{n-2}^* on $f_1,f_2,\ldots,$ and f_{n-3}, which we write as $f_{n-2}^*[f_1,\ldots,f_{n-3}]$. The process continues until we have determined f_1^*. Since f_1^* does not depend on any prior functions, we can use it to determine the succeeding reaction functions by successively substituting starred values into $f_2^*[f_1]$, $f_3^*[f_1,f_2],\ldots,$ and $f_{n-1}^*[f_1,f_2,\ldots,f_{n-2}]$.

In this way, a solution, $(f_1^*,f_2^*,\ldots,x_n^*)$, which implies a solution social choice, $(x_1^*,x_2^*,\ldots,x_n^*)$, is determined.

The finite structure of the successive maximization problems, along with the completeness and transitivity of $\succ_i$, assures us that a solution <u>always</u> exists.

II. PARETO OPTIMALITY

Besides unqualified existence, the solution has the important property of Pareto optimality. That is, institutions formed under truly perfect information always imply Pareto optimal allocations.

To prove this, suppose the solution allocation, $(x_1^*,..,x_{n-1}^*,x_n^*) = x^*$, is not Pareto optimal. Then there is a point, $x^o = (x_1^o,...,x_n^o) \in \prod_{i=1}^n X_i$ such that $x^o \succ_i x^*$ for all i. A set of reaction functions generating x^o as social action is given by $(f_1^o,...,f_{n-1}^o)$. Of course, $(f_1^*,...,f_{n-1}^*,x_n^o) \neq (f_1^o,...,f_{n-1}^o,x_n^o)$; otherwise x^o would be the solution. Now let individual 1 consider:

$$\text{(A)} \quad f_1(f_2,..,f_{n-1},x_n) = \begin{cases} f_1^o \text{ if } (f_2,..,f_{n-1},x_n) = (f_2^o,..,f_{n-1}^o,x_n^o) \\ f_1^* \text{ otherwise.} \end{cases}$$

This may induce each subsequent strategy selector to reorder his strategy in $f_2^o,...,f_{n-1}^o,x_n^o$ relative to $f_2^*,...,f_{n-1}^*,x_n^*$. However, as it does not alter the social actions resulting from non-solution strategies <u>other than</u> $f_2^o,...,f_{n-1}^o,x_n^o$, it does not alter anyone's ordering of these other strategies relative to $f_2^*,...,f_{n-1}^*,x_n^*$. Therefore, because $x^o \succ_1 x^*$, individual 1 is no worse off under (A) than under his original strategy.

We next let individual 2 consider, in view of (A),

$$\text{(B)} \quad f_2(f_3,...,f_{n-1},x_n) = \begin{cases} f_2^o \text{ if } (f_3,...,x_n) = (f_3^o,...,x_n^o) \\ f_2^* \text{ otherwise.} \end{cases}$$

This similarly cannot hurt individual 2. We continue on to individual n, who now faces (A), (B),..., Thus, $(f_1^o,...,f_{n-1}^o,x_n^o) \Rightarrow x^o$ will result if he picks $x_n = x_n^o$; and $(f_1^*,...,f_{n-1}^*,x_n^*)$ if he picks his solution action. Since $x^o \succ_n x^*$, he picks the former. The supposition that there is a Pareto nonoptimal solution is thus immediately contradicted: For the supposition implies that the players individually prefer a non-solution set of strategies $(f_1^o,\ldots,f_{n-1}^o,x_n^o)$ to the solution set, $(f_1^*,\ldots,f_{n-1}^*,x_n^*)$.

III. STRATIFICATION

There is, in general, a significant advantage to being the first to establish a reaction function. If, for example, any individual can adopt actions that punish any single other individual sufficiently that the other individual would be better-off serving as a slave than suffering the punishment, then the equilibrium social action is that most desired by the first strategy selector (Thompson-Faith, 1980, Part IV). Correspondingly, there is a game, preceding our own, representing a competition to be the first strategy selector. This higher-order game is a generally inefficient, war-like, Nash-VNM noncooperative affair because strategic communication is -- by definition -- not yet established. The obvious qualities conducive to winning this game are the abilities to (a) inflict physical punishment on potential competitors while withstanding their attacks, (b) flexibly adopt moral principles assuring broadly rational (but narrowly irrational) responses and (c) act benevolent toward potential competitors under normal conditions in order to reduce their net profit from challanging the leader. The resource wastes involved in this competition, while of some descriptive-historical significance (Section VI), are socially unavoidable and of no direct relevance to social policy. Regarding the potential resource wastes involved in acquiring subsequent hierarchial positions, several theoretical reasons exist to believe that these possible resource wastes are insignificant. For one, the first strategy selector, aware of the possibility of such waste, may use his prior commitment ability to assign the remaining hierarchal positions, punishing individuals who attempt to deviate from their assigned status. For another, since the return to hierarchal positions below the first is plausibly insignificant in our world (Thompson-Faith, 1980, Part IV), significant resources would not be devoted to acquire such positions.

However, the analysis, to apply beyond family and small village societies, must admit the possibility that several individuals have substantial informational advantages over the first strategy selector. Once this is admitted, certain, systematic changes appear in the subsequently chosen reaction functions. In particular, the subsequent strategy selectors, being sometimes able to escape the disipline of a reaction by prior selectors and also being generally less wealthy and therefore less generous towards the lower strata than the prior selectors, will adopt reaction functions generating solution norms that are systematically less beneficial to the lower strata (and, of course, more beneficial to themselves) than the reaction functions under truly perfect information. In other words, the members in the middle of the decision hierarchy will tend to "exploit their power positions" over the lower members. This not only produces solutions that differ from our own, optimal solution, it gives these decision makers an incentive to devote resources to hiding their true reaction functions from higher authorities in an attempt to gain personal redistributions from those lower in the social hierarchy. This latter effect generates pure economic waste. Similarly, efficient middle-level decision makers would tend to be competed out of their position by less efficient decision makers who are, nonetheless, relatively talented at deception.

The first strategy selector, understanding this problem, should then look for special institutions or reaction functions in order to reduce these economically wasteful and, to him, distributionally undesireable tendencies. If, for example, he imposes a common-law private-property system, a system with tort and contract law provisions serving to compensate those whose centrally determined "endowments" are somehow reduced by the action of others, he can limit the ability of middle-level members of the hierarchy to punish lower level members in order to obtain redistributional benefits from them and, at the same time, assure certain minimum benefit levels for the lower strata of the social decision hierarchy. While institutions other than common-law private-property -- e.g., central direction of final consumption activities or enforced egalitarianism -- could also achieve these goals, private property systems free the central authority from the job of figuring out the various preferences of the others while allowing more informed individuals to cooperate in a less centralized fashion. The key question that arises, however, is: Does a private property constraint distort reaction functions in ways that induce inefficient allocative solutions? The following section shows that the answer is negative -- that perfect cooperation achieves a Pareto optimum despite a common-law private property constraint. In economics, this theorem, although heretofore more of a conjecture than a theorem, is called the "Coase Theorem."

When communication between various, higher-level decision makers is costly, serious inefficiency problems arise that are alleviated -- both theoretically and empirically -- by other institutional selections of the first strategy selector (Thompson-Faith, 1980, Part IV). A discussion of this would take us into the history of social organization and thus beyond the scope of this paper.

IV. THE PARETO OPTIMALITY OF INTRAGROUP INTERACTION UNDER PRIVATE PROPERTY AND TRULY PERFECT INFORMATION

We now let n be a subset of a larger group that has imposed a private property constraint on the interactions between the n individuals. The constraint has the effect of limiting the set of feasible social actions to $\prod_{i=1}^{n} \overline{X}$, the subset of $\prod_{i=1}^{n} X$ such that no one can be made worse-off than he would be under a certain, "endowed" set of benchmark actions, x^B, $x^B \in \prod_{i=1}^{n} X_i$.

That is, private property restricts social actions to the set,

$$\prod_{i=1}^{n} \overline{X}_i = \{x \in \prod_{i=1}^{n} X_i\text{: either } x = x^B \text{ or } x \succ_i x^B \text{ for all } i\}.$$

While this constraint may dramatically alter the solution social choice, the alteration being obtained by replacing the original X-constraint with an $\overline{X}$-constraint, the altered solution is still Pareto optimal.

Our proof of this is a simple variant of the above proof for the unconstrained case: Suppose the new solution $\overline{x}*$ is not Pareto optimal. Then there is an $x^0 \in \prod_{i=1}^{n} X_i$ such that $x^0 \succ_i \overline{x}*$ for all i. Individual 1 could, as in the Section II proof, be no worse off with an altered reaction function that delivered x_1^0 if everyone else chose his part of x^0 and $\overline{f}_1^*$ otherwise. Since $x^0 \succ_i \overline{x}*$ for all i and $\succ_i$ is transitive, $x^0 \succ_i x^B$ for all i so the private property constraint cannot be violated by 1's altered reaction function. This is the key. Individual 2 similarly has a feasible reaction function that he could be no worse-off choosing in which he selects x_2^0 if the subsequent strategy selectors choose their parts of x^0 but $\overline{f}_2^*$ otherwise. This continues on to individual n, who would rationally pick x^0 instead of $\bar{x}*$, thereby contradicting the supposition that $\bar{x}*$ is a Pareto nonoptimal solution.

The most immediate application of this secondary optimality result is apparently to hierarchial organizations such as are found in firms and small societies within a private property system. The results suggest that, except for deadweight resource costs in the process of establishing hierarchial position, such organizations produce Pareto optimal decisions. This, perhaps, explains the persistent survival of bureaucratic, hierarchial forms of organization despite the perennial intellectual attacks on their efficiency.

Since, however, strategic communication costs are obviously not trivial, we should consider an alternative form of organization suggested by the use of the polar alternative, game-theoretic assumption of zero information with respect to the strategies of others. Here, a Nash solution is appropriate. The incentive system inducing actions that are Pareto optimal from the standpoint of the firm under this assumption is equivalent to one which pays each decision maker in the firm the entire total profits of the firm and has these decision makers pay lump sums for the right to join the firm. (This is essentially the incentive system of Groves. It has the important advantage of solving the "shirking" problem emphasized by Alchian-Demsetz.) But we do not observe these "optimal" incentive systems! The reason is presumably that the Nash assumptions which underly their "optimality" are completely untenable. That is, under such incentive systems, there would be an irresistible incentive to members of the organization to communicate strategies and thus cooperate so as to induce substantial overworking by the group. The fact that these "optimal" incentive systems have been rejected in favor of hierarchial systems in real world organization therefore provides strong empirical evidence for the applicability of our strategic communication assumptions relative to Nash assumptions to analyze behavior within organizations.

When communication between subgroups is sufficiently costly and hence rare, these groups, empirically speaking, become labelled separate "organizations" and, at the same time, Nash assumptions become appropriate in describing certain interactions between these organizations. Correspondingly, the first strategy selector may find different kinds of institutions desirable in responding to inter-organizational interactions. An analysis of such institutions is beyond the scope of this paper (see Thompson-Faith, 1980, Part IV, for such an analysis).

V. THE DIFFICULTIES PRESENTED BY INDIFFERENCE

If someone other than the first strategy selector were indifferent between two or more possible solution strategies, a prior selector, who would otherwise have no way of knowing what the indifferent one would do, would -- assuming that he could perform a reaction that would leave this later selector uniformly worse off than in a solution -- simply adjust his reactions to all but one of the later-selected strategies so as to make these strategies suboptimal for the later selector. The resulting solution in this case is also Pareto optimal, as can be seen by noting that our above optimality proof also applies here as long as the Pareto dominating strategy used in the proof is still feasible, which is the case because no prior strategy selector, in inducing a specific choice of a later selector between strategies about which the later selector would otherwise be indifferent, would eliminate the Pareto superior strategy choice.

However, when the <u>first</u> strategy selector is indifferent, Pareto nonoptima may easily arise. Consider the following, "slave master's insensitivity," payoff matrix illustrated in Figure 1. The standard, VNM-Nash, no-regret solution has the slave resting while the master insults the slave; this is both nonoptimal <u>and</u> empirically unrealistic! The solution set under perfect strategic communication, with the master as the first strategy selector, contains the Pareto optimum (10, 0) where the master will beat the slave if he rests and leave him alone if he works. But the set also contains (10, -4), as the master may also insult the slave, lowering the slave's benefit to -4 without altering either the master's payoff or the slave's optimal decision. The point, (10, -4), is obviously Pareto inferior to (10, 0). To see all this, set up the majorant of the payoff matrix (Fig. 2), which defines, in the

MASTER \ SLAVE	WORK x_2'	REST x_2''
BEAT THE SLAVE x_1'	5, -10	0, -6
INSULT THE SLAVE x_1''	10, -4	1, 0
LEAVE THE SLAVE ALONE x_1'''	10, 0	0, 4

FIGURE 1. THE SLAVE MASTER'S INSENSITIVITY-
ACTIONS AND PAYOFFS

F_1	MASTER \ SLAVE	x_2'	x_2''
f_1^{I}	x_1'	5, -10	0, -6
f_1^{II}	x_1''	10, -4	1, 0
f_1^{III}	x_1'''	10, 0	0, 4
f_1^{IV}	$x_1' \mid x_2',\ x_1'' \mid x_2''$	5, -10	1, 0
f_1^{V}	$x_1' \mid x_2',\ x_1''' \mid x_2''$	5, -10	0, 4
f_1^{VI}	$x_1'' \mid x_2',\ x_1' \mid x_2''$	10, -4	0, -6
f_1^{VII}	$x_1'' \mid x_2',\ x_1''' \mid x_2''$	10, -4	0, 4
f_1^{VIII}	$x_1''' \mid x_2',\ x_1'' \mid x_2''$	10, 0	1, 0
f_1^{IX}	$x_1''' \mid x_2',\ x_1' \mid x_2''$	10, 0	0, -6

note: "$x_1^{\phi} \mid x_2^{\psi}$" means "player 1 chooses x_1^{ϕ} if player 2 chooses x_2^{ψ}."

FIGURE 2. NORMAL FORM OF THE SLAVE MASTER'S INSENSITIVITY

the master's column, the F_1 functional. A standard, VNM, perfect information solution -- which has the player with the expanded strategy set waiting for the other player to move first -- has the slave choosing between his two actions on the basis of what the master's best response would be. This would obviously lead to the (1, 0) solution where the slave doesn't work and the master insults the slave, a quite unsatisfactory solution. Our game -- which differs in that the player with the expanded strategy set is the first to move -- has the master picking f_1^*, that f such that the slave's best response yields the maximum payoff to the master. It is easy to see using Figure 2, that only f^{VI} and f^{IX} will assure the master of his maximal payoff, 10. Since the master is indifferent between these two strategies, he may pick either. Since the slave is clearly worse off at f_1^{VI}, where he gets -4, than he is at f_1^{IX}, where he gets 0, the solution set obviously contains a Pareto nonoptimal as well as a Pareto optimal point.

The exercise of Section II can be repeated for a weak preference relation to show that if x*, a Pareto nonoptimum, is a solution, so is x°. Hence, although the solution set with weak preference relations may sometimes contain a Pareto nonoptimal point, it must always also contain a Pareto optimum.

Because all standard competitive equilibria are Pareto optimal, economists have grown accustomed to the thought that individual indifference between various possible equilibria is unimportant. But, as individual indifference between the possible equilibria of a master-slave relationship can induce Pareto nonoptima, we should guard against the habit of ignoring solution indifference when examining decentralized slave economies. Apparently, the real world has not ignored the problem. As our model would predict, observed decentralized slavery systems have arisen only through the capture of social "outsiders" toward which initial benevolence could hardly have been widespread among insiders (Finlay) and have dissolved not by slave unrisings or voluntary manumissions but, at least in modern times, by the intervention of politically powerful humanitarians armed with a "moral argument" (Finlay) based on examples in which slaves were torn from their families, worked to death, tortured, or broken of spirit for the minor conveniences of their only mildly benevolent, and therefore largely indifferent, masters. Thus, as predicted by the model, solution indifference created inefficiencies only when the leaders were also indifferent; when the leaders lost their indifference, the institution permitting the inefficiencies was dissolved. While abolition has sometimes also served to redistribute away from the masters, as it apparently did in the American South in view of the slow pace of Southern Reconstruction, in most cases freed slaves have become serfs or debt-peons who provide about the same benefit as do slaves to the capitalist class (Finlay). The social advantage of serfdom and debt-peonage is that they prevent local slave master's insensitivity problems, the former having central authorities rigidly controlling the taxation of the immobile serfs and the latter, by granting a choice of creditor-employers to the peon, inducing the prospective employers to compete away payment systems which harm the worker without benefiting the employer.

VI. BIOLOGICAL APPLICATION

The optimality results also have some descriptive, biological relevance to the entire population. To illustrate this, we add the following biological assumptions: (1) each player's payoff is an increasing function of only the survival probabilities of himself and, perhaps, some others in his n-player group; (2) the "action", x, which determines survival probabilities include the physical characteristics of the players (i.e., their size, shape, mobility, etc.). In other words, each player is considered an abstract, amorphous unit that selects strategies defined over physical characteristics and behavior patterns so as to maximize its survival probability.

If preferences (i.e., survival probabilities) over x were identical for all players, there would be no conflict as all players would share the same set of optimal actions, x*. Each player, i, would rationally select x_i^*, and there would be no need for reaction functions (i.e., social institutions). Indeed, the "players" representing the vital parts of living creatures have such preferences. The heart and lungs of a given person share a common survival probability so that their abstract representations need no reaction function to achieve a joint optimum. To avoid these situations, we combine such "players" into units, thereby shortening our list of relevant players to include only those with unique preferences over x.

Applying our central optimality theorem, with exogenously fixed hierarchal positions, the social equilibrium has the property that there is no alternative set of physical characteristics or behavior patterns such that the survival probability is greater for at least one player and no lower for the others. This helps explain the marked tendency toward group efficiency and apparent self-sacrifice in nature (e.g., Wynne-Edwards) without resorting to implausible group-selection arguments. (See Ghiselin for a critical evaluation of such arguments.)

As indicated above in Section III, the mere existence of our process induces a higher-order competition for hierarchical position favoring those with characteristics of battle superiority, benevolence, and abilities to carry out reactions despite their narrow irrationality. Hence, the general theory also helps explain the socially <u>inefficient</u>, competitive evolution of size, speed and cunning observed by Darwin and others, the concomitant survival of extra-familial benevolence among possible leaders (e.g., Lorenz), and the survival of socially instinctive behavior towards carrying out one's promise, such as emotion-induced, punative reactions (e.g., Dawkins).

An example indicating how joint efficiency tends to emerge once hierarchal positions are fixed may be helpful. Many modern biologists (esp. Maynard Smith, and Dawkins) consider the phenomenon of sex (or Meiosis) somewhat of a paradox in the absence of group selection because, put in our own, crude terms, the abstract player-female appears to have a much higher chance of survival if <u>all</u>, rather than half, of the offspring which she nurtures have her own, female, characteristics. While the greater biological flexibility of the offspring in the world with sex produces a somewhat higher survival probability for each of the offspring (e.g., Ghiselin), there are many examples in nature in which sexual and asexual units live stable existences with almost indistinguishable physical characterics other than their differing reproductive apparatus (Dawkins, Ghiselin). Why wouldn't the asexual organisms in these circumstances compete out the sexual organisms; the female has only insignificant use for the male. But what makes the survival and predominance of sexual reproduction a paradox is a Nash assumption. Once we allow males to commit themselves to physically punish shirking or asexual females (in patriarchal societies) or (in matriarchies) allow females to commit themselves to punish males who do not participate in child rearing (or otherwise compensate the females), then it is easy to see that jointly efficient, sexual, individual will not be driven out by the self-interested calculation of those who specialize in child-rearing.

Similar arguments apply to explain the supposedly paradoxical existence of "warning coloration" and other such self-sacrificial characteristics (Wynne-Edwards), characteristics that cannot be easily explained with conventional Nash assumptions.

VII. COOPERATIVE GAMES

Cooperative game theory is founded on the assumption that any subset of n can form a "blocking coalition," a group of players which can, presumably by a given set of actions, achieve a given payoff for themselves and thereby prevent certain outcomes. The excluded outcomes are the "non-solution" outcomes to the game. The assumption guarantees each player a payoff at least equal to the minimum of what he can achieve in a one-man coalition. On the additional assumption that any Pareto optimum can be achieved by a coalition of all n players, the theory guarantees that any solution must be Pareto optimal. For any Pareto nonoptimal solution would be blocked by an n-person coalition. The theory is then devoted to the search for a solution out of the resulting set of "imputations," i.e., Pareto optimal points which give each player at least the minimum of what he would receive in a one-man coalition. The standard solution set indicated above, the core, is the set of unblocked outcomes. A chronic problem with this theory is that its solution sets are often empty. Other solution sets, such as VNM's "stable set" and the "bargaining set" are less frequently empty but have the chronic problem of admitting a superabundance of outcomes in their solution sets (see, for example, Owen).

We object to cooperative game theory because of its inexplicit communication process and related absence of committed strategies. These weaknesses result in insufficient constraints on the set of blocking coalitions. This point requires some elaboration.

Blocking coalitions exist in a general form as a by-product of interaction under truly perfect information. For any subset of reaction functions effectively blocks all outcomes which do not simultaneously satisfy these functions; and the players in the subset may be thought of as a blocking coalition. However, in our model, the players may be worse-off under their blocking behavior but still engage in it because they recognize the effect of their strategies on the strategies of others. The commitment of the players to these strategies simultaneously prevents them from forming blocking coalitions with subsequent strategy selectors merely because they would be better off in these coalitions under the narrowly rational, uncommitted, reactions characteristic of standard game theory.

Consider, for example, a three-person, zero-sum, "majority" game, in which, say, a dime and a nickel are to be shared by the three players. If players 1 and 2 each select certain actions implying that they "get together," 2 gets a dime and 1 gets a nickel. If 1 and 3 each select certain actions, where the action is different for 1 than in the former case, then 1 gets a dime and 3 gets a nickel. If 2 and 3 each select new actions implying that they "get together," then 3 gets a dime and 2 gets a nickel. Cooperative game theory offers no meaningful solution to this game because, for any distribution of coins, there is a blocking coalition.[4] Under truly perfect information, where the order of strategy selection is, say, 1, 2, 3, player 1 will adopt the following strategy: "I will get together with 2 if he gets together with me; otherwise, I will perform my part of getting together with 3." Player 2 then selects: "I will perform my part of getting together with 1 regardless of the action of player 3. Player 3 gets nothing no matter what he does. It is easy to verify that there is no other solution. In sharp contrast, under cooperative game theory, 3 would offer to get

---

[4] While the core and VNM's stable set are empty, the bargaining set contains all possible allocations.

together with 1, who -- being unable to commit himself to a fixed strategy -- would be unable to refuse the offer. And we would be off on the never-ending cycle of coalition formation characteristic of existing cooperative game theory.

REFERENCES

Alchian, Armen A. and Demsetz, Harold M. [1972]. "Production, Information Costs, and Economic Organization." American Economic Review, Vol. 62 (5).

Dawkins, R., The Selfish Gene, Oxford University Press, N.Y., 1976, pp. 224.

Durkheim, Emile, Professional Ethics and Civic Morals, London, Routledge.

Finlay, M. I., "The Origins of Slavery," Ch. 1 of Slavery: A Comparative Perspective, R. W. Winks, Ed., New York University Press.

Ghiselin, M. T., The Economy of Nature and the Evolution of Sex, Univ.

Groves, Theodore [1973]. "Incentives in Teams," Econometrica, Vol. 41, pp. 617-663.

Howard, Nigel, Paradoxes of Rationality, MIT Press, Cambridge, 1971.

Lorenz, V. Z., "Ritualized Fighting," in The Natural History of Aggression, J. D. Carthy and F. J. Ebling, Eds., Academic Press, N.Y., 1964, pp. 39-50.

Maynard, Smith, J., "What Use is Sex," J. Theo. Biol., 1971, pp. 317-35.

McFarland, David D., "Notes on the Hobbesian Problem of Order: Individual Rationality, Collective Rationality, and Social Control in Prisoner's Dilemma Type Situations," Presented to the Seminar on Mathematical Theory of Collective Decisions, Harbour Town, Hilton Head Island, S. Carolina, August 8-14, 1971.

Nash, John, "Noncooperative Games," Annals of Mathematics, 54 (1951), pp. 286-295.

Owen, G., Game Theory, Saunders, Philadelphia, 1969.

Parsons, Talcott, The Structure of Social Action, N.Y., Free Press, 1937.

Rapoport, A. and A. M. Chammah, Prisoner's Dilemma, University of Michigan Press, Ann Arbor, pp. 26-29.

Schelling, Thomas C., The Strategy of Conflict, Harvard University Press, Cambridge, 1963.

Selten, Reinhart [1974], "The Chain-Store Paradox," Working Paper No. 18, Institute of Mathematical Economics, University of Bielefeld.

Thompson, Earl A. and R. L. Faith, "A Pure Theory of Strategic Behavior and Social Institutions," forthcoming, American Economic Review, 1980.

von Neumann, J., and O. Morgenstern, Theory of Games and Economic Behavior, 3rd Ed., Princeton University Press, Princeton, 1953.

Weber, Max, Economy and Society, Berkeley and Los Angeles, University of California Press, 1978 (originally 1972).

Wynne-Edwards, V. C., Animal Dispersion in Relation to Social Behavior, Hofner, N.Y., 1963, pp. 603.