Discussion Paper No. 731

COOPERATION THROUGH DELEGATION*

by

Chaim Fershtam**

Kenneth L. Judd***
and

Ehud Kalai***

February 1987
Revised December 1987

---

*This research was partly supported by NSF Economics Grants Nos. SES-8409798 and SES-8606581.

**Department of Economics, Tel Aviv University, Tel Aviv 69967, Israel.

***Department of Managerial Economics and Decision Sciences, J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60208.

[3-5]

COOPERATION THROUGH DELEGATION

Abstract

Most economics and political games are played by agents of the real players. This paper analyzes the outcomes of general noncooperative games in which the players can strategically use agents to play on their behalf. It is shown that when players in a noncooperative game delegate strategy choices to their agents every Pareto optimal outcome of the game becomes a Nash equilibrium of the delegation game. A folk theorem using delegation games is presented. The result casts serious doubt about the properness of many models applying game theory to the analysis of economic activity.

COOPERATION THROUGH DELGATION

1. Introduction

Game theory is extensively used in describing and analyzing conflict situations in economics and social science. Such an analysis contributes to our understanding of market functions, how firms compete, the effect of government policies on markets, and may lead to policy recommendations. The extensive use of game theory only emphasizes that we should be careful in such applications. A major question that must be addressed is the sensitivity of the outcome of the analysis to the specification of the game used in modeling the problem. In the presence of such a huge wave of game theory applications, this problem of robustness should be a main concern of game theorists as well as social scientists.

It is a common observation in conflict situations that players are represented by agents who play the game on their behalf. Lawyers often represent clients in negotiation, agents represent actors and sports players, managers represent owners and elected officials represent voters. Thus, in discussing conflict situations in social science, specific attention must be paid to the possibility of hiring agents who participate in the game on behalf of the real players.

Besides considering the implications of such delegations, a fundamental question is the explanation of this phenomenon. Why does a player hire someone who will represent him in a game? Clearly one possible explanation is that there are games in which having special skills is essential. But can we argue that the only purpose in hiring a lawyer is always his superior knowledge of the law? Or can we argue that players can gain strategic advantage just by having someone with different incentives play the game on their behalf? For example, consider a bargaining game in which the two parties have to agree on a partition of a pie. If we adopt the axiomatic approach, it is straightforward to demonstrate (see Kihlstrom, Roth and Schmeidler (1981)) that a player can benefit by hiring an agent who is less risk averse than he is. If we consider the strategic approach (see Rubinstein, 1982), it is clear that a player can benefit by sending a representative who is less impatient than he is. Thus, delegating the power to play can be advantageous to players. However, before rushing to quick conclusions, one should understand that the risk aversion, the degree of impatience, as well as other characteristics of the agent depend crucially on the contract that is signed between the principal and his agent and on the compensation scheme designated by such a contract.

The main purpose of this paper is to analyze the effects of strategic delegation on the outcome of games. That is, to what extent does the set of equilibria of a strategic game change when agents are allowed to strategically represent the main players? It turns out that delegation substantially alters the set of outcomes. Even in highly noncooperative games, cooperative outcomes emerge as equilibria in the game with delegation.

In game theoretic language, a full folk theorem is obtained. That is to say, every individually rational feasible outcome of the underlying game is an equilibrium of the game with delegation. Obtaining such folk theorems has been a major thrust of research in game theory. For example: (i) cooperation through repetition (see, for example, Aumann (1981), Rubinstein (1979), Friedman (1985), and Benoit-Krishna (1985)); (ii) cooperation through correlations and binding agreements (see, for example, Aumann

(1974), Cass-Shell (1983), Forges (1986), Myerson (1986), and Kalai-Samet (1985)); (iii) cooperation through reputation building (see, for example, Kreps-Milgrom-Roberts-Wilson (1982), Fudenberg-Maskin (1986)); (iv) cooperation due to bounded rationality (see, for example, Radner (1980), Aumann-Sorin (1985), and Neyman (1985)). This paper may also be viewed as an addition to this literature, as it suggests another way of obtaining the folk theorem.

Our delegation game is described as follows. To every two person strategic game we associate a delegation game in which agents play the original game on behalf of the original players--their principals. More specifically we formulate a two-stage game. In the first stage each principal provides his agent with a compensation scheme. These compensation schemes determine each agent's final reward as a function of the principals' payoffs, i.e., the payoffs that are obtained in the original game. In the second stage, each agent, after learning both compensation schemes,[1] plays the original game, choosing his principal's strategy so as to maximize his own final payoff. The principals then receive their payoffs in the original game net of promised compensation to their agents.

This paper is an extension of the principal-agent theory. Except for few papers (see, for example, Fershtman-Judd (1986, 1987), Myerson (1982), Katz (1987), and Vickers (1985)), this literature concentrates on the strategic game played between the principal and his agent; see, for example, Ross (1973) and Holmstrom (1973). However, we will find that the strategic use of agents competing against other players is an important phenomenon

---

[1]See Myerson (1982) and Katz (1987) for a discussion on delegation games in which compensation schemes are private knowledge.

that should be analyzed.

As an example we consider a Cournot type duopolistic game. We allow the owners of firms to hire managers who run the firms on their behalf. We show that even when we consider a one shot game such that the two firms meet only once in the market the existence of the owners-managers strucutre affect dramatically the set of possible equilibrium outputs. In particular we show that every collusive outcome can be obtained in such an oligopolistic game with delegation. We feel that since most firms are not managed by the original owners our analysis, along with Fershtman-Judd (1986, 1987) and Vickers (1985), casts serious doubts about the properness of the usual analyses of such oligopolistic markets.

2. The Delegation Game

Our analysis of delegation begins with a game representing the basic conflict. The underlying game is a 2-player strategic form game with the set of players $P = (p_1, p_2)$. We let $S = S_1 \times S_2$ be the set of strategy combinations in this game. The payoffs of the players are described by a utility function $u = (u_1, u_2): S \to \mathbb{R}^2$. We will use $G = (P, S, u)$ to denote this underlying game. A Nash equilibrium of this game is a strategy combination (s_1^*, s_2^*) with $u_1(s_1^*, s_2^*) \geq u_1(s_1, s_2^*)$ for every $s_1 \in S_1$ and similarly for player 2.

For such a game we define the associated delegation game, D, as follows: the set of players $N = (p_1, p_2, a_1, a_2)$ where p_1 and p_2 are called principals and a_i is called the agent of p_i. The set of strategies of p_i is

$$C_i = \{c_i: \mathbb{R} \to \mathbb{R}_+ : c_i \text{ is weakly monotonically increasing}\}$$

We refer to an element of C_i as a <u>compensation function</u> of agent a_i and we let $C = C_1 \times C_2$ denote the set of compensation pairs. Note that agent a_i's compensation depends only on his principal's gross payoff. Also, we restrict the compensation schemes to be weakly monotonic. Beside being intuitively appealing, there is a technical need for such a restriction. As will be discussed later subgame perfection cannot be obtained without such a restriction. The strategy set of every agent a_i is $R_i = \{r_i : C \to S_i\}$ and we let $R = R_1 \times R_2$. We call an element of R_i a <u>response function</u> of agent i.

Given a 4-tuple of strategies $g = (c_1,c_2,r_1,r_2)$ we define the utilities of the four players in D as follows:

$$U_i^P(c_1,c_2,r_1,r_2) = u_i(r(c)) - c_i(u_i(r(c))),$$

and

$$U_i^a(c_1,c_2,r_1,r_2) = c_i(u_i(r(c))).$$

Note that u is used for payoffs in the underlying game while U is used for payoffs in the delegation game.

<u>Definition 1</u>: $(r_1^*,r_2^*) \in R$ is a <u>Nash equilibrium</u> in the <u>agents' game</u> induced by $(c_1,c_2) \in C$ if

$$U_i^a(c_1,c_2,r_1^*,r_2^*) \geq U_i^a(c_1,c_2,r_1,r_2^*)$$

for every $r_1 \in R_1$ and similarly for agent 2. We let

$$EA(c) = \{(r_1^*(c), r_2^*(c)) \in S: \quad (r_1^*, r_2^*)$$
$$\text{is an equilibrium in the agents game induced by } c\}.$$

Notice that the weak monotonicity of compensation functions guarantees that every Nash equilibrium s^* of the underlying game G is in EA(c) for every c. However EA(c) will typically contain other equilibria as well.

<u>Definition 2</u>: $(c_1^*, c_2^*, r_1^*, r_2^*)$ is a <u>subgame perfect Nash equilibrium</u> of D if

i. $U_1^P(c_1^*, c_2^*, r_1^*, r_2^*) \geq U_1^P(c_1, c_2^*, r_1^*, r_2^*)$ for every $c_1 \in C_1$ and the same for principal 2, and

ii. For every pair of compensation functions $c \in C$, (r_1^*, r_2^*) is a Nash equilibrium of the agents' game induced by c, i.e., $r^*(c) \in EA(c)$.

<u>Definition 3</u>: c <u>implements</u> $u \in \mathbb{R}^2$ via $r \in \mathbb{R}$ if (c,r) is a subgame perfect equilibrium of D with $u(r(c)) = u$.

Under this definition when u is implemented, u_i is shared by principal i and his agent.

<u>Theorem 0</u>: If s* is a Nash equilibrium of the underlying game, every feasible payoff pair $u \geq u(s^*)$ is implementable in the delegation game.

<u>Proof</u>: Let s* be a Nash equilibrium and let $(\hat{s}_1, \hat{s}_2)$ be a pair of actions such that $u(\hat{s}_1, \hat{s}_2) = \hat{u} \geq u(s^*)$. Let f_i be the flat compensation scheme such that $f_i(u_i) = 0$ for every u_i and $i = 1,2$. Let r_i be defined by $r_i(f_1, f_2) = \hat{s}_i$ and for every $c \in C$ with $c \neq f$ let $r_i(c) = s_i^*$. The monotonicity of the

compensation function guarantees that the above construction is a subgame perfect Nash equilibrium of the delegation game. Q.E.D.

Note that without monotonicity or other similar restrictions obtaining subgame perfection is impossible. It is always possible to construct "wild" compensation schemes for which there is no equilibrium in the induced agents' game. The weak monotonicity guarantees the existence of an equilibrium in every conceivable agents' game since it preserves the equilibria of the original game. We could have foregone this assumption at the cost of losing subgame perfection.

The above theorem shows that without any additional requirements on the choice of equilibrium in the agents' game any feasible individual rational payoffs can be trivially implemented. However this implementation although formally correct, is intuitive unattractive. It suffers from the usual difficulty of dealing with multiple equilibrium as well as the choice of Pareto dominated equilibrium. Under the flat compensation scheme the agents are completely indifferent regarding their choice of strategy. Yet the principals have to count on them choosing the particular $\hat{s}$ that yields $\hat{u}$. But it would be naive on the part of the principals to expect the agents to bring about the exact $\hat{u}$ under such a great indifference since every feasible u can be an equilibrium response to a flat compensation scheme. Moreover off the equilibrium path the agents are choosing s* which may be Pareto dominated by other equilibria of the agents game. Game theory however does not have yet the proper equilibrium concept to deal with such conceptual difficulties. We thus reconstruct our formulation of the delegation game and the solution concept in order to exclude some unattractive phenomena of this type and still retain the folk theorem flavor of this result.

Multiplying all the payoffs of a player in a strategic game by a positive constant results in a new game, isomorphic to the original one. For our delegation game an implication of this observation is important. It is reasonable to assume that if the principals change a compensation scheme (c_1, c_2) say, to $(.5c_1, c_2)$ the strategies of the agents will stay the same yielding principal one (and two) the same payoff. Yet principal one would have to spend half as much on compensating his agents. Since this argument can be applied repeatedly, the only equilibrium for the principals is to come back to the flat compensation scheme.

In order to deal with the above difficulty we modify the delegation game in the following way: for a fixed but arbitrary $\varepsilon > 0$ we define the ε-associated delegation game D_ε as follows. The set of players, the payoff functions and the agents' strategies remain as in our former definitions. However the set of possible compensation scheme that the principals can provide is restricted to be $C_i = \{c_i : \mathbb{R} \to \varepsilon Z \mid c_i$ is weakly monotonically increasing$\}$ where Z is the set of all nonnegative integers.

The discretization of the set of compensation schemes is natural and not unrealistic since any known currency has a minimal indivisible unit. An alternative interpretation of ε is to be the minimal noticeable unit required to move the agent from one level of effort to another.

For the modified delegation game D_ε we define the notions of Nash equilibrium, subgame perfect equilibrium and implementability as before. We are now able to make the following additional definitions.

<u>Definition 4</u>: c <u>implements</u> u <u>via</u> r <u>with individually rational agents</u> if in

addition to implementing u, r satisfies the following property: for every $\bar{c} \in C$, $\bar{c}(u(r(\bar{c}))) \geq (\epsilon,\epsilon)$ whenever there is $s \in EA(\bar{c})$ with $\bar{c}(u(s)) \geq (\epsilon,\epsilon)$.

Intuitively the above definition requires that if there is an equilibrium in the agent game in which both agents get at least ε they will choose such an equilibrium. Moreover, such a choice is done at every subgame, i.e., for every pair of compensation schemes. Clearly the unattractive implementation discussed in Theorem 0 does not satisfy this requirements since in some subgames the agents choose to play s* yielding them the payoff (0,0) while other agents' equilibria would yield them $(\varepsilon,\varepsilon)$ or better.

Notice that the individual rationality condition is an assumption on the agents selection among multiple Nash equilibria. It assumes that the agents will coordinate their actions in order to avoid 0-payoffs if possible and nothing else.

<u>Definition 5</u>: c <u>uniquely implements</u> u <u>with individual rational agents</u> if there is an $r \in R$ such that c implements u via r with individual rational agents and if for some $\bar{u}$ and $\bar{r} \in R$, c implements $\bar{u}$ via $\bar{r}$ with individual rational agents then $\bar{u} = u$.

Unique implementations are attractive since they guarantee the principals the vector u without depending on a specific choice of equilibrium by their agents. Thus we exclude counterintuitive and unattractive implementation like the one presented before with the flat compensation schemes.

We say that T_i is a <u>target compensation function</u> if for every u_i, $T_i(u_i) \in \{0,\epsilon\}$. Target compensation functions pay nothing unless a minimal level of utility is obtained for the principal and pay ε if that target level is obtained or exceeded.

Theorem 1 (folk theorem in delegation games): If u is (strictly) Pareto optimal in G and for i = 1,2, $u_i - u_i(s^*) > \epsilon$ for some Nash equilibrium s* of G, then u is uniquely implementable with individual rational agents in the game D_ϵ. Moreover the implementation can be done by target compensation functions.

Proof: Let $T_{u_i} \in C_i$ be a target compensation function defined by:

$$T_{u_i}(\bar{u}_i) = \begin{cases} \epsilon, \text{ if } \bar{u}_i \geq u_i \\ 0, \text{ otherwise} \end{cases}$$

and define $r(\bar{c})$ as follows:

i. If there is an $s \in EA(\bar{c})$ with $\bar{c}(u(s)) \geq (\epsilon,\epsilon)$ then let $r(\bar{c})$ be such an s.

ii. Otherwise, let $r(\bar{c}) = s^*$.

We will show that (T_u,r) uniquely implements u with individually rational agents.

It is clear that $u(r(T_u)) = u$. By playing the above strategies the agents' payoffs is (ϵ,ϵ) and by the fact that u is strictly Pareto optimal in G it is the only outcome such that $T_u(u) \geq (\epsilon,\epsilon)$. By construction $r(c) \in EA(c)$ for every $c \in C$ and it is individually rational. Now we check that T_{u_1} is indeed a best response strategy to (T_{u_2}, r_1, r_2).

Suppose that principal one deviates and plays $c_1 \in C_1$. One of the following two cases must hold:

a. There exists $s \in EA(c_1, T_{u_2})$ such that $(c_1(u_1(s)), T_{u_2}(u_2(s))) \geq (\epsilon,\epsilon)$. By the Pareto optimality of u it follows that principal one cannot be better off from this deviation.

b. Otherwise, the outcome of the game is s^* and $u_1(s^*) < u_1 - \epsilon$.

T_u uniquely implements u with individual rational agents since every $s \in EA(T_u)$ with $c(u(s)) \geq (\epsilon,\epsilon)$ has $u(s) = u$. Q.E.D.

One could conceive of defining an agent compensation schemes on other parameters of the game. For example an agent's compensation could depend on the difference in the payoff of his principal minus the payoff of his opponent's principal, or even on combinations of payoffs and actions in the underlying game. These modifications are possible but they only serve to complicate a simple idea. They may also come at a cost of losing the subgame perfection property of the implementations, since they may create a large number of unreasonable subgames.

The reason that the delegation game "works" is that through their agents (and their compensation schemes) the principals can commit to playing cooperative strategies (with enough flexibility to allow them to retaliate against deviations). When the principals cannot commit directly to each other (such as in a one shot game and when binding agreements are not permissible) the ability to commit to their agents is very beneficial as was illustrated in the theorem. One could question the ability of the principals to commit to their agents and actually pay them according to the promised compensation schemes. This commitment however is credible when binding agreements between them are possible and legally enforceable. Also, a long term relationship between the principal and an agent would enhance these commitment possibilities even when the principal is using the agent in a sequence of games with different opponents.

3. Discussion and Conclusion

To highlight the main feature of our analysis we next discuss a particular example, the Cournot duopoly, and its relation to previous similar analyses. Consider a duopolistic market in which the inverse demand function is given by $p = a - b(q_1 + q_2)$ where p is the market price and q_i is the i'th firm output. Assume that the cost function is given by $TC_i(q_i) = mq_i$ where $a > m \geq 0$ is the constant marginal cost. The i'th firm profit is thus $\pi_i(q_1,q_2) = q_i(a - b(q_1 + q_2)) - mq_i$. The unique equilibrium of such one shot game is $q_1^* = q_2^* = (a - m)/3b$ and the equilibrium payoffs are $\pi_i(q_1^*,q_2^*) = (a - m)^2/9b$.

Now let us consider the delegation game in which the owner of each firm hires a manager that will make the choice of q_i on his behalf. The owner and the manager sign a contract that specify the manager compensation as a function of the firm's performance.

Let $\hat{q}_1 = \hat{q}_2 = (a - m)/4b$ and $\hat{\pi}_i = (a - m)^2/8b$ be the maximum symmetric collusive output levels and the collusive payoffs, respectively.

Proposition 1: In the above duopolistic game the collusive outcome, $(\hat{q}_1,\hat{q}_2)$, can be uniquely implemented with individual rational agents.

Proposition 1 is an immediate corollary of Theorem 1. However to get some intuition on the type of strategies and contracts that give rise to such an equilibrium we will specify the players' strategies.

Let $T_{\hat{\pi}_i}$ be the following target compensation function for $i = 1,2$:

$$T_{\hat{\pi}_i}(\pi_i) = \begin{cases} \varepsilon, & \text{if } \pi_i \geq \hat{\pi}_i \\ 0, & \text{otherwise} \end{cases}$$

Let the managers' strategies for a specific pair of compensation schemes c be as follows: when possible, the managers will choose a pair of output rates q which are a Nash equilibrium in the managers' game induced by c and $c(\pi(q)) \geq (\varepsilon,\varepsilon)$; otherwise they will choose the Cournot production rate q_i^*.

Clearly the above construction implies that once the compensation scheme $T_{\hat{\pi}_i}$ are given the managers will choose the only production rate that yields the payoffs $(\varepsilon,\varepsilon)$ which is the collusive production rate $(\hat{q}_1,\hat{q}_2)$. Also given the response of the managers to various incentive schemes the owners can do no better than using these target schemes. Subgame perfection of the above equilibria means that the owners can be convinced that if they deviate from their compensation scheme the managers will indeed respond as stated.

It is interesting to observe that the principals' commitment to cooperate through their agents is done by sending out agents that are less "hungry" than themselves. This is done by flattening the agents' compensation from some critical level of the principals' utility and on.

This Cournot analysis should be compared to previous analyses of competing principal-agent pairs. Fershtman and Judd (1987) examined the same game but restricted contracts to be linear in profits and sales; in this case the owners' profits were less than in the underlying game's Nash equilibrium. In Fershtman and Judd (1986) it was assumed that the manager's effort was unobserved, creating a moral hazard problem; again the outcome was not a cooperative one for the principals. From the latter analysis it is clear that the folk theorem of our analysis may break down when contracts must deal with moral hazard problems with managers as well as coordinate cooperation with opposing principals. The value of delegation in incomplete information games is an open and interesting question.

The result reported in the paper can be viewed as a negative one. Delegation of games is very common. Rarely are serious large games played by the original owners. Therefore, analyzing the original game is naive. Given the large multiplicity of equilibria in the delegation game, the Nash equilibrium concept can predict very little unless we add some more assumptions about its selection. For example, one could consider various consistency conditions on the Nash equilibrium selection as we vary the games. These types of conditions turn out to be very useful in cooperative game theory.

References

Aumann, R. J. and S. Sorin (1985), "Bounded Rationality Forces Cooperation," unpublished manuscript.

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

Aumann, Robert J. (1981), "Survey of Repeated Games," in Essays in Game Theory and Mathematical Economics, Bibliographsiches Institut, Mannheim, Wein, Zurich, 11-42.

Benoit, Jean-Pierre and Vijay Krishna (1985), "Finitely Repeated Games," Econometrica, 53, 905-922.

Cass, David and Karl Shell (1983), "Do Sunspots Matter?", Journal of Political Economy, 91 (2), 193-227.

Fershtman, Chaim and Kenneth Judd (1987), "Incentive Equilibrium in Oligopoly," American Economic Review, forthcoming.

Fershtman, Chaim and Kenneth Judd (1986), "Strategic Incentive in Manipulation in a Rivalrous Agency," Technical Report No. 496, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Forges, Francoise (1986), "An Approach to Communication Equilibria," Econometrica, 54 (6), 1375-1386.

Friedman, James W. (1985), "Cooperative Equilibria in Finite Horizon Noncooperative Supergames," Journal of Economic Theory, 35, 390-398.

Fudenberg, Drew and Eric Maskin (1986), "The Folk Theorem in Repeated Games with Discounting and With Incomplete Information," Econometrica, 54 (3), 533-554.

Holmstrom, B. (1979), "Moral Hazard and Observability," Bell Journal of Economics, 10, 74-91.

Kalai, Ehud and Dov Samet (1985), "Unanimity Games and Pareto Optimality," International Journal of Game Theory, 14 (1), 41-50.

Katz, Michael L. (1987), "Game-Playing Agents: Contracts as Precommitments," unpublished draft, University of California, Berkeley.

Kihlstrom, R. E., A. E. Roth and D. Schmeidler (1981), "Risk Aversion and Nash's Solution to the Bargaining Problem," in Game Theory and Mathematical Economics, edited by O. Moeschlin and D. Pallaschke, Amsterdam, North-Holland.

Kreps, David, Paul Milgrom, John Roberts, and Robert Wilson (1982), "Rational Cooperation in the Finitely-Repeated Prisoners' Dilemma," Journal of Economic Theory, 27, 245-252.

Myerson, Roger (1982), "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," Journal of Mathematical Economics, 10, 67-81.

Myerson, Roger (1986), "Multistage Games with Communication," Econometrica, 54 (2), 323-358.

Neyman, Abraham (1985), "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoners' Dilemma," Economic Letters, 19, 227-229.

Radner, Roy (1980), "Collusive Behavior in Noncooperative Epsilon-Equilibrium of Oligopolies with Long but Finite Lives," Journal of Economic Theory, 22, 134-154.

Ross, S. (1973), "The Economic Theory of Agency: The Principal's Problem," American Economic Review, 63, 134-139.

Rubinstein, Ariel (1979), "Equilibrium in Supergames with the Overtaking

Criterion," Journal of Economic Theory, 21, 1-9.

Vickers, J. (1985), "Delegation and the Theory of the Firm," Economic Journal, conference supplement, 138-147.