Econometrica, Vol. 58, No. 4 (July, 1990) 873–900

EFFICIENCY DESPITE MUTUALLY PAYOFF-RELEVANT PRIVATE INFORMATION: THE FINITE CASE

BY SCOTT JOHNSON, JOHN W. PRATT, AND RICHARD J. ZECKHAUSER[1]

Individuals have or observe partly private information. They independently choose acts, possibly including messages. The center may also act. Individuals' utilities may depend on all acts and information, including others' private information. Are there incentives depending only on public information that make desired behavior a Bayesian equilibrium?

Assume incentive payments are separable and fully transferable. Appropriate incentives exist either if the center's information—perhaps solely messages—depends stochastically, however slightly, on all relevant private information, or if individuals' relative valuations of acts, however divergent, are not too dissimilarly affected by different states of nature. More generally, we give necessary and sufficient conditions for existence whenever the strategy profile asks agents to reveal all private knowledge relevant to their beliefs about the center's information. We also develop equivalences on the possible values of private information—concepts of similarity of agent types—that are key to resolving existence questions without such responsiveness or requiring budget balance.

KEYWORDS: Decentralization, incentives, private information, group decision, proper scoring, permutation dominance.

1. INTRODUCTION

THE PROBLEMS OF *eliciting honest information* and *inducing appropriate actions*, central challenges for most groups of individuals, are the subject of this paper. We make no welfare judgments beyond the desirability of increasing any individual's expected utility; Pareto optimality is our implicit criterion for group performance. When we seek to influence individual behavior, it is natural to introduce monetary incentives, given the substantial drawbacks and social inefficiencies of alternatives such as sodium pentothal, mind control, and torture. Our formulation is traditional: a central authority makes a monetary payment (possibly negative) to an individual depending on his actions and additional information available to the central authority, including the actions and reports of others. The central authority may be some overarching ruler, a hired professional outside the group, a trusted agent within the group, or the group members acting collectively on their own behalf, as they might in a partnership. The payments may or may not be constrained to some arbitrary budget, or to balance within the group.

We assume that individuals, whom we often call agents, are risk neutral. We derive surprisingly strong positive results. For incentive payments to exist it suffices that, at the desired equilibrium, the information available to the central authority—which may consist solely of reports from other individuals—be

[1] Scott Johnson died in a climbing accident after this paper was completed. We would like to thank Jerry Green, Aanund Hylland, participants in seminars at the Business School and Kennedy School, Harvard University, M.I.T., the Australian National University, and the NBER-NSF Decentralization Conference (New York), May 1987, and two referees for helpful comments. Green, Hylland, Pratt, and Zeckhauser (1984) provided the initial stimulus for this investigation. The Japanese Corporate Associates Program and the Business and Government Center, Kennedy School, provided research support.

stochastically dependent, however slightly, on all relevant private information of the individuals. Alternatively, even when there is independence, hence no possibility of monitoring, an effective payment scheme exists if no individual's evaluation of potential actions, however divergent from the group's evaluation, is too dissimilarly affected by different states of nature. To illustrate, if a couple is deciding what recreation to pursue, the tennis enthusiast and the movie maven must both believe rain to be more harmful to the pleasures of tennis.

We describe settings in which the issues mentioned above can be resolved either positively or negatively. Section 2 introduces the notation needed and presents a model encompassing all essential aspects of the problem. The infinite variety of nature is approximated by a finite set of states with positive probabilities. We assume throughout that each agent's utility has two separable components. The first component incorporates the effects of the actions and information of all agents, public information, and quite possibly actions by a central administration which depend upon these. The second component merely reflects that agent's financial payments from or to the central authority.

Section 2 also discusses the relationship of some of the results here to results in recent literature. We use the Bayesian equilibrium concept of Harsanyi (1967–68), coupled sometimes with Harsanyi's condition of consistent beliefs. This is well motivated by Myerson (1985). Although Myerson is not concerned with transfer payment schemes, this theme is taken up, for instance, by Arrow (1979), d'Aspremont and Gérard-Varet (1979), and in the special case of auctions, by Crémer and McLean (1985) and Wilson (1985). Of these, only Myerson and Crémer and McLean permit, as we do, *mutually payoff-relevant information*: the agents' utility functions may depend directly on the information and actions (including messages) of other members of the group.

One characteristic common to these papers has been that the center or mediator is free to choose any mechanism, which includes a choice of strategies to be implemented, in order to induce an efficient Bayesian equilibrium. Much of the literature invokes the revelation principle; this says that it suffices to consider only mechanisms that induce agents to reveal honestly all privately held information. By contrast, we find conditions under which a fixed strategy profile can be implemented using only transfer payments, whether or not the strategy profile happens to be efficient or completely or honestly revealing of private information. Thus we do not appeal to the revelation principle or require truth revelation, although we sometimes specialize our results to that important case. In our context, the revelation principle is intuitively obvious but confers no significant technical advantage. Moreover, routine transformation to an equivalent problem with truth revelation would obscure the implementability conditions, especially when there are real externalities. We do, however, get our best results under a responsiveness assumption that can be interpreted as a weakened version of truth revelation. It requires that the strategy ask agents to reveal any private knowledge relevant to their beliefs about information available to the central authority.

In Section 3 we present some results ignoring budget balance and assuming responsiveness. Necessary and sufficient conditions are given for the existence of a transfer payment scheme which can implement a given strategy profile.

Section 4 develops the notion of an equivalence on the possible values of private information, or *types* of agents. This notion can be used to address the implementation question when the assumption of responsiveness is dropped. In general the center cannot induce an agent to behave differently depending on differences of type between which the center cannot differentiate even stochastically. The more easily the center can differentiate stochastically between two types, the less "similar" those types are. Finding the correct *equivalence*, or notion of similarity between types, is the key to answering the strategy-inducement question.

For example, this enables us to show in Section 5 (assuming consistent beliefs and responsiveness) that appropriate behavior can always be induced with budget balancing exactly when no agent is asked to act differently as a function of types that are "similar" in a particular, constructively defined sense. This sense of similarity is, in general, strictly weaker than the sense of similarity relevant to the strategy-inducement question without budget balancing. In fact these two senses of similarity coincide (a condition which we call (LINK)) if and only if budget balancing always comes for free when strategy inducement is possible. In the case of truth revelation, and still assuming consistent beliefs, we show that (LINK) is equivalent to the discrete case of the compatibility condition of d'Aspremont and Gérard-Varet (1982),[2] thus offering an interpretation of that condition. A special case of (LINK) is that two types are similar (in either sense) only when they are equal. This condition is equivalent to d'Aspremont and Gérard-Varet's condition B, while their condition F is equivalent to the strong similarity of *all* of some agent's types, which also implies (LINK). It will be clear from our formulation that there is plenty of scope for (LINK) to hold aside from these two (mutually exclusive) cases, thus answering a question they raise.

When an agent is asked to act differently given different but similar types, results depend on the particular shape of that agent's utility function as well as on the beliefs of the agents. It turns out that work on what is called *announcement-proof equilibria* by Green, Hylland, Pratt, and Zeckhauser (1984), and *ex post Nash equilibria* by Crémer and McLean (1985) applies here. When certain results from these papers are translated and generalized to our context, they combine with the theory of equivalences to form a general picture of the responsive case.

This picture is blurred somewhat when responsiveness is not in force. Nevertheless, Section 4 shows how positive results can be obtained by weakening our notions of similarity to accommodate responsiveness in a weaker form.

Section 5 is devoted to the question of balancing the budget, assuming responsiveness. Johnson, Pratt, and Zeckhauser (1988) give further results without responsiveness, both with and without budget balance.

[2] We will show in the Appendix that this is strictly weaker than the compatibility condition given in d'Aspremont and Gérard-Varet (1979), even in the case of consistent beliefs.

There are n agents in the group. For each $i \in N = \{1,2,\ldots,n\}$, agent i privately observes his type r_i and chooses an act a_i depending only on r_i (and not, for instance on a_j for $j \neq i$). After the agents choose $\underline{a} = (a_1,\ldots, a_n)$, the central authority will observe $\underline{a}$ and the value of a random variable $\tilde{z}$ which includes all public information and possibly information possessed only by the center. The center then acts according to a known rule and pays each agent i an amount $t_i(z, \underline{a})$ (depending *only* on z and $\underline{a}$). Although r_i will not be observed by any but agent i, the *payment scheme* $t = (t_1, t_2,\ldots, t_n)$ is common knowledge. It may sometimes be convenient to let $-\infty$ be in the range of t_i, representing death or some other event, any positive probability of which is worse than losing one dollar. A budget balance constraint $\sum_{i=1}^{n} t_i(z, \underline{a}) = 0$ for all z and $\underline{a}$ will sometimes be imposed. Throughout the remainder of this paper we assume *finiteness*:

(FIN) $\tilde{z}$ and each $\tilde{r}_i$ can have only finitely many different values.

A *strategy* for agent i is a function A_i of agent i's private information specifying the action agent i will take. We call $\underline{A} = (A_1,\ldots, A_n)$ a *strategy profile* and say agent i follows A_i or follows $\underline{A}$ if, whenever r_i is observed, $A_i(r_i)$ is agent i's chosen action. We let r_{-i} denote $(r_1,\ldots, r_{i-1}, r_{i+1},\ldots, r_n)$, and similarly for a_{-i} and A_{-i}.

For each i and r_i, agent i has a joint probability distribution over the other agents' types $\tilde{r}_{-i}$ and the center's information $\tilde{z}$. We write $P(z, r_{-i}|r_i)$ for the probability that $\tilde{z} = z$ and $\tilde{r}_{-i} = r_{-i}$ given type r_i. (The center is assumed to know the functions $P(-|-)$.) These beliefs are said to be *consistent* (see Harsanyi (1967–68)) if:

(CON) The probabilities $P(z, r_{-i}|r_i)$ are derived from a common joint prior distribution of $(\tilde{z}, \underline{\tilde{r}})$ using Bayes' rule.

Although (CON) is commonly assumed, only some results in Section 5 need it in their proofs and we will not invoke it until then.

The utility function of agent i is of the form $V_i(z, \underline{a}, \underline{r}) + W_i(t_i(z, \underline{a}))$ where V_i is what we call the *direct return function*, the utility agent i would receive in the absence of any transfer payments, and W_i is some increasing function from R to R such that $W_i(0) = 0$. V_i incorporates the effects on agent i of the center's action. Agent i is assumed to act so as to maximize his or her expected utility. Throughout the remainder of this paper we assume *risk neutrality*:

(RNEUT) For each i, W_i is linear.

For convenience, we take W_i to be the identity, so that agent i's utility is $V_i(z, \underline{a}, \underline{r}) + t_i(z, \underline{a})$ with t_i separable and fully transferable.

To simplify notation, let

$$U_i(a_i|r_i) = E\{V_i(\tilde{z}, A_{-i}(\tilde{r}_{-i}), a_i, \underline{\tilde{r}}) + t_i(\tilde{z}, A_{-i}(\tilde{r}_{-i}), a_i)|\tilde{r}_i = r_i\},$$

which is agent i's expected utility (after observing r_i) if everyone else uses the strategy profile $\underline{A}$ and the payment scheme is $\underline{t}$. This notation suppresses the dependence of U_i on $\underline{A}$, $\underline{V}$, and $\underline{t}$.

Clearly, given some designated strategy profile $\underline{A}$, agent i does not necessarily have an incentive to follow A_i even if it is assumed that the other agents are following A_{-i}. A pair $(\underline{t}, \underline{A})$ is called an *equilibrium* if $\underline{A}$ is a Bayes-Nash equilibrium given $\underline{t}$, that is, for all i and r_i,

$$U_i(A_i(r_i)|r_i) = \max_{a_i} U_i(a_i|r_i).$$

Then we say that each agent has an incentive to follow $\underline{A}$ and that $\underline{t}$ is *$\underline{A}$-inducing.*

Given a designated strategy profile $\underline{A}$, we seek conditions under which there is an $\underline{A}$-inducing payment scheme, and conditions under which such a payment scheme can be chosen to balance the budget.

A special case of this model, of additional interest because of its relationship to the revelation principle (see Myerson (1982)), is that of *truth revelation*:

(TR) $\quad A_i$ is a one-to-one function for all i.

This is equivalent for mathematical purposes to the requirement that A_i be the identity for all i, and corresponds to the case that agents are asked to reveal all their private information fully. When (TR) is in effect, "$\underline{A}$-inducing" will be replaced by "truth-inducing."

This model also includes the case that for some functions $f_1, \ldots, f_n$, $V_i(z, \underline{a}, \underline{r}) = f_i(C(z, \underline{a}), \underline{r})$ for all i, z, $\underline{a}$ and $\underline{r}$, where $C(z, \underline{a})$ is a choice of action by the central authority (usually a specified part of a mechanism). If the central authority is seeking to induce $\underline{A}$ so as to maximize group utility we would have:

(2.1) $\quad \sum V_i(z, \underline{A}(\underline{r}), \underline{r}) \geqslant \sum V_i(z, \underline{a}, \underline{r}) \quad$ for all z, $\underline{a}$, and $\underline{r}$.

We call $\underline{V}$ *standard* if it satisfies (2.1). Under natural definitions of efficiency, such as ex ante or ex post Pareto optimality, $\underline{A}$ will be efficient (for a given $\underline{V}$) whenever (2.1) holds, though the converse may fail for weak notions of efficiency.

Another reasonable restriction one could place on the utility profile is that no agent's private information can directly affect the utility of another agent:

(2.2) $\quad V_i(z, \underline{a}, \underline{r}) = V_i(z, \underline{a}, r_i) \quad$ for all z, $\underline{a}$, $\underline{r}$, and i.

Either of these assumptions by itself is innocuous:

PROPOSITION 2.1: *Given an arbitrary utility profile $\underline{V}$, there exist a standard $\underline{V}^s$ and a $\hat{\underline{V}}$ satisfying* (2.2) *such that any payment scheme $\underline{t}$ is $\underline{A}$-inducing for $\underline{V}$ iff it is $\underline{A}$-inducing for $\underline{V}^s$ and iff it is $\underline{A}$-inducing for $\hat{\underline{V}}$.*

PROOF: For all i, let

$$V_i^s(z, \underline{a}, \underline{r}) = \begin{cases} -L & \text{if } a_j \neq A_j(r_j) \text{ for some } j \neq i, \\ V_i(z, \underline{a}, \underline{r}) & \text{otherwise,} \end{cases}$$

let

$$\hat{V}_i(a_i, r_i) = E\{ V_i(z, A_{-i}(\tilde{r}_{-i}), a_i, \underline{\tilde{r}}) | \tilde{r}_i = r_i \}$$

and notice that the expected utilities $\underline{U}$ are the same whether they are derived from $\underline{V}$, $\underline{V}^s$, or $\underline{\hat{V}}$. It is clear that $\underline{\hat{V}}$ satisfies (2.2) and that $\underline{V}^s$ is standard provided L is sufficiently large. *Q.E.D.*

Thus, when considering the problem of finding $\underline{A}$-inducing payment schemes for arbitrary utility profiles, one could (though we do not do so here) restrict attention to standard utilities *or* to those utilities satisfying (2.2). However (2.1) and (2.2) taken together have strong consequences. We call a utility profile $\underline{V}$ satisfying both (2.1) and (2.2) *public*. Under (TR), d'Aspremont and Gérard-Varet (1979) have shown that a truth-inducing payment scheme $\underline{t}$ exists for *all* public $\underline{V}$ (by (2.2) one can construct an externality payment which is $\underline{A}$-inducing by (2.1)). On the other hand our Proposition 3.3 (for instance) shows that there are many general $\underline{V}$ for which there is no $\underline{A}$-inducing payment scheme. Thus we will allow arbitrary utility profiles $\underline{V}$, since neither (2.1) nor (2.2) greatly simplifies our analysis, and since there is real loss of generality in assuming both.

A case considered by Arrow (1979) and by d'Aspremont and Gérard-Varet (1979) in their work on collective decision problems is that of truth revelation, no $\tilde{z}$ and public $\underline{V}$. Assuming also that the types are mutually independent, they obtain analytic expressions for transfer payments that induce honest reporting (with a balanced budget). Pratt and Zeckhauser (1986) derive these payments directly as expected externalities. They also allow agents to take actions that directly affect other agents, to have information other than on preferences, possibly not independent, and to act and signal repeatedly, simultaneously or sequentially. However, when one agent's unsignalled private information affects others directly, his expected externality may be unknowable by others, which would prevent the approach from being implemented. Whether or not efficiency-inducing transfers nevertheless exist is the question addressed in this paper.

Myerson (1982, 1985) allows for mutually payoff-relevant information, though there is nothing in his general formulation that corresponds to our restriction to separable utility functions that are linear in the transfer payments. Thus, for instance, Myerson's result that there is no loss of generality in restricting consideration to the case of consistent beliefs and independent types is not valid for our model. Myerson defines an incentive-compatible mechanism as one for which honest and obedient behavior by the agents is a Bayes-Nash equilibrium. He then characterizes those incentive compatible mechanisms which are also efficient. Here we focus, rather, on the question of the existence of transfer payment schemes for which obedience to a given strategy profile is an equilibrium.

A special problem of strategy or truth inducement studied extensively in the literature is optimal auction design. Crémer and McLean (1985), for example, give conditions under which a seller can extract full surplus from a group of buyers who have mutually payoff-relevant information. They also use a strong equilibrium concept—which they call *ex post Nash equilibrium*—and require as an individual rationality constraint that all buyers have nonnegative expected returns. Though the question they consider differs somewhat from the questions that concern us, two of their assumptions relate directly to conditions we employ. First, they use a condition of monotonicity in the derivative of the agents' utility functions to construct payments which induce honesty. The same condition always implies the existence of truth-inducing payment schemes in our context, though we give a weaker condition, based on Green, Hylland, Pratt, and Zeckhauser (1984), which does the same. Also, in Sections 4 and 5 we use a condition of linear independence of certain probability vectors of which Assumption 4 of Crémer and McLean (1985) is a special case used for a similar purpose.

Another treatment of auctions under incomplete information is given by Wilson (1984) who finds that a certain double auction is both optimal and individually rational, assuming independent information, no payoff-relevant information, truth revelation, and a restricted form of utility function.

Although it is true that, following Myerson (1982), one can essentially reduce the problem of finding a strategy-inducing transfer-payment scheme to a linear feasibility problem, we seek simpler or more insightful characterizations where possible. Our conditions involve only those parameters which are part of the data of the problem, i.e., the probability assessments, the strategy profile, and the direct returns. We do not ask for the consistency of inequalities in external parameters.

For example, as noted in the Introduction, we define a condition (LINK) which we show is equivalent to the discrete case of the compatibility condition of d'Aspremont and Gérard-Varet (1982) assuming consistent beliefs and truth revelation. Unlike the compatibility condition, however, (LINK) is defined in a way which makes its verification a straightforward (and polynomial-time) computational procedure. Our conditions on beliefs can be interpreted in terms of certain notions of similarity on types as discussed in the Introduction. However, computational complexity of necessary and sufficient conditions for strategy inducement involving the utility functions as well as the beliefs is (in a certain formal sense) unavoidable even in the responsive case.

3. RESPONSIVE STRATEGY INDUCEMENT

The strategy profile $\underline{A}$ will be fixed throughout. We will assume that every action lies in the range of $\underline{A}$ (without loss of generality, since the center can always punish heavily for actions chosen outside this range, even while balancing the budget). Hence we may write $P(z, a_{-i}|r_i)$ for the probability that $\tilde{z} = z$ and $A_{-i}(\tilde{r}_{-i}) = a_{-i}$ given the type r_i.

As mentioned in the Introduction, notions of similarity between different possible types of an agent will play a key role in our results. The first and most basic of these notions is *strong similarity*.

We say two types r_i and r_i' are *strongly similar*, and write $r_i \sim r_i'$, if they give agent i the same distribution of $(A_{-i}(\tilde{r}_{-i}), \tilde{z})$, that is, if $P(z, a_{-i}|r_i) = P(z, a_{-i}|r_i')$ for all z and a_{-i}. This says that the conditional distribution of all other agents' acts (under $\underline{A}$) and the center's information $\tilde{z}$ is the same given r_i as it is given r_i'. Equivalently, under consistent beliefs, whether $\tilde{r}_i = r_i$ or r_i' is independent of the other agents' acts and $\tilde{z}$ jointly. Hence the latter provide no information whatever distinguishing in any way between r_i and r_i' and, conversely, this distinction is useless to agent i, in the sense of sufficient statistics or partitions, for inference about the other agents' acts and $\tilde{z}$. Clearly $\sim$ is a vector of equivalence relations, one for each agent.

The next condition can be regarded as a weakened form of truth revelation, though it includes other cases of interest not covered by truth revelation. Throughout this section we assume *responsiveness*:

$$\text{(RESP)} \quad A_i(r_i) = A_i(r_i') \Rightarrow r_i \sim r_i'.$$

In words, (RESP) says that the designated strategy profile must ask each agent to reveal any information relevant to others' acts and z that he or she observes. Note that it is implied by truth revelation (TR) since $r_i = r_i' \Rightarrow r_i \sim r_i'$.

In this section we give necessary and sufficient conditions for the existence of a transfer payment scheme which is $\underline{A}$-inducing, assuming responsiveness. The next section examines the case when (RESP) may fail, and Section 5 examines the problem of balancing the budget.

Under the assumption (CON) of consistent beliefs, a weaker form of budget balancing used by d'Aspremont and Gérard-Varet (1982) is *expected budget balance*:

$$\text{(EBB)} \quad E\sum t_i(\tilde{z}, \underline{A}(\tilde{\underline{r}})) = 0.$$

This always comes for free when there exists an $\underline{A}$-inducing payment scheme $\underline{t}$, since we can set $t_i'(z, \underline{a}) = t_i(z, \underline{a}) - E\{t_i(\tilde{z}, \underline{A}(\tilde{\underline{r}}))\}$. The payment scheme $\underline{t}'$ will still be $\underline{A}$-inducing, since adding a constant to an agent's payment cannot alter incentives, and $\underline{t}'$ clearly satisfies (EBB).

To motivate the ideas behind the theorems we will look at two fundamental special cases of responsiveness. These two cases are primitives from which all other responsive cases can be derived. The first is the case of *belief announcement*:

$$\text{(BA)} \quad A_i(r_i) = A_i(r_i') \Leftrightarrow r_i \sim r_i'.$$

This means that A_i essentially asks agent i to announce his beliefs concerning $A_{-i}(\tilde{r}_{-i})$ and $\tilde{z}$ and no more. Clearly (BA) implies (RESP). Under (BA) the action desired of an agent strictly depends stochastically on the other information available to the center. This makes it possible for the center to induce the desired action. In an extreme case, if the central authority's information is the same as all of the private information, it can simply punish an agent heavily for any deviation from $\underline{A}$. (BA) guarantees only that obedience to $\underline{A}$ can be monitored probabilistically, but this suffices for an $\underline{A}$-inducing payment scheme to exist.

The second special case of responsiveness, in a sense the opposite of (BA), is the case that each agent's type is independent of all information obtained by the center not from that agent:

(IND) The distribution of $A_{-i}(\tilde{r}_{-i})$ and $\tilde{z}$ is independent of agent i's type.

Put another way, (IND) says $r_i \sim r_i'$ for *all* types r_i and r_i'. Thus (IND) implies (RESP) trivially. For this case we present simple linear inequalities on the direct returns only, inequalities that are satisfied if and only if an $\underline{A}$-inducing payment scheme exists.

(BA) and (IND) are incompatible conditions unless $\underline{A}$ is a degenerate, single-action strategy profile.

3.1. *Belief Announcement*

The special case of the problem under (BA) in which the direct returns are zero has been studied in the literature on proper scoring rules. It is well known (e.g., Good (1952), van Naerssen (1962), Mosteller and Wallace (1964), Winkler (1967)), that an expected-value maximizing agent i can be induced to state his true distribution of $(\tilde{z}, A_{-i}(\tilde{r}_{-i}))$ by paying him the logarithm of his stated probability $P'(\tilde{z}, A_{-i}(\tilde{r}_{-i}))$. Under (BA), therefore, the central authority can make agent i's payment such a large multiple of this log likelihood that on average the benefit to agent i from the increased payment for signalling the correct distribution exceeds any benefit through agent i's direct return from choosing to deceive.

LEMMA 3.1: *Let S be any finite set and P and P' be arbitrary probability mass functions on S. Let $g(P, P') = \Sigma_{s \in S} P(s) \ln P'(s)$. Then for each P, $g(P, P')$ is strictly maximized at $P' = P$.*

PROOF: By Jensen's inequality, $g(P, P') - g(P, P) = \Sigma P(s) \ln [P'(s)/P(s)] < \ln [\Sigma P(s) P'(s)/P(s)] = 0$ if $P' \neq P$. See also references above. *Q.E.D.*

If (BA) fails, the existence of an $\underline{A}$-inducing payment scheme depends on the utility profile (this is made precise in Theorem 3.8). To analyze this we turn to the case where (BA) fails maximally.

3.2. *Independent Information*

To avoid cumbersome notation, we use the following abbreviation: for any i, and $\underline{A}$, and any function f depending on the information and the actions of the agents, let

$$Ef(a_i|r_i) = E\{f(\tilde{z}, \tilde{r}_{-i}, r_i, A_{-i}(\tilde{r}_{-i}), a_i)|r_i\}.$$

In the case of independent information the existence of an $\underline{A}$-inducing payment scheme $\underline{t}$ is equivalent to the existence, for each i, of quantities $c(a_i)$ such that $EV_i(a_i|r_i) + c(a_i)$ is maximized at $a_i = A_i(r_i)$. If there is such a function c, we can let $t_i(z, \underline{a}) = c(a_i)$ and $\underline{t}$ will be $\underline{A}$-inducing. Conversely, if $(\underline{t}, \underline{A})$ is an equilibrium pair we can let $c(a_i) = Et_i(a_i|r_i)$, which does not depend on the choice of r_i by the independence assumption.

If for a given i there is a function c as above, then we will call the pair $\langle EV_i, A_i \rangle$ a *transfer maximum*. More generally, if B is a set of types for agent i, we say $\langle EV_i, A_i \rangle$ is a *transfer maximum on B* if there exists a real-valued function c of agent i's actions such that $EV_i(a_i|r_i) + c(a_i)$ is maximized at $a_i = A_i(r_i)$ for all $r_i \in B$. One would like a straightforward way of checking this property. For example, suppose A_i is the identity and $B = \{1,2\}$. Then $\langle EV_i, A_i \rangle$ is a transfer maximum on B if and only if there exist two numbers c_1 and c_2 such that

$$EV_i(1|1) + c_1 \geqslant EV_i(2|1) + c_2, \quad \text{and}$$
$$EV_i(2|2) + c_2 \geqslant EV_i(1|2) + c_1.$$

Adding the inequalities together yields

$$(3.1) \qquad EV_i(1|1) + EV_i(2|2) \geqslant EV_i(2|1) + EV_i(1|2)$$

which is therefore a necessary condition for $\langle EV_i, A_i \rangle$ to be a transfer maximum on B. Setting $c_1 = EV_i(2|2)$ and $c_2 = EV_i(1|2)$, we see that (3.1) is also sufficient.

For a more general example, we get from Crémer and McLean (1985) that in the case where $B = \{1, \dots, k\}$ and A_i is the identity (i.e., (TR) holds), the following condition is sufficient for $\langle EV_i, A_i \rangle$ to be a transfer maximum on B:

$$(3.2) \qquad EV_i(m+1|l+1) - EV_i(m|l+1) \geqslant EV_i(m+1|l) - EV_i(m|l) \quad \text{for all } l < k \quad \text{and } m < k.$$

This requires the cross-difference of EV_i with respect to l and m to be nonnegative. However, this condition is far from necessary for $k > 2$.

The appropriate generalization of (3.1) is straightforward except possibly in the absence of truth revelation when for some action a_i there may be several types r_i such that $A_i(r_i) = a_i$. A *representative function* is a function r from the actions in $A_i(B) = \{A_i(r_i)\colon r_i \in B\}$ to the types in B such that $A_i(r(a_i)) = a_i$ for all $a_i \in A_i(B)$. Under (TR) (where A_i is one to one) there is only one such function, namely the inverse of A_i. (In general r is sometimes called a partial inverse to A_i.)

We call A_i *permutation dominant for EV_i on B* if:

(PD) For all permutations π of $A_i(B)$ and all representative functions r,

$$\sum_{a_i \in A_i(B)} EV_i(a_i|r(a_i)) \geqslant \sum_{a_i \in A_i(B)} EV_i(\pi(a_i)|r(a_i)).$$

Here $A_i(B)$ is the range of A_i on B, so that A_i is a function *onto* $A_i(B)$. If B consists of all possible types for agent i we simply say that A_i is permutation dominant for EV_i.

If $A_i(B)$ consists of only one action, then permutation dominance is trivially satisfied as (PD) becomes an equality in a single summand. If A_i is the identity and if $B = \{1,2\}$, (PD) is precisely the condition (3.1).

It is easiest to interpret permutation dominance when $B = \{1, 2, \dots, k\}$ and A_i is the identity. In this case (PD) becomes

$$\sum_{l=1}^{k} EV_i(l|l) \geqslant \sum_{l=1}^{k} EV_i(\pi(l)|l) \quad \text{for all permutations } \pi \text{ of } \{1, \dots, k\}.$$

This says that in the $k \times k$ matrix $\{EV_i(l, j)\colon 1 \leqslant l \leqslant k,\ 1 \leqslant j \leqslant k\}$, the sum of the elements along the main diagonal is at least as big as the sum of any other k elements picked with one from each row and one from each column. This gives $k! - 1$ inequalities, though efficient algorithms can reduce the number of calculations to the order of 2^k.

In general, note that if $B \subseteq B'$ then permutation dominance on B' implies permutation dominance on B, that is, permutation dominance is more easily satisfied on smaller sets of types.

LEMMA 3.2: *$\langle EV_i, A_i \rangle$ is a transfer maximum on B iff A_i is permutation dominant for EV_i on B. In this case we may take*

$$c(a_i) = \min\Big\{ \sum_{k=1}^{m} \big[EV_i\big(A_i(r_i^k) | r_i^k\big) - EV_i\big(A_i(r_i^{k+1}) | r_i^k\big)\big] : m \geqslant 1,\ r_i^k \in B \text{ for } 1 \leqslant k \leqslant m+1, \text{ and } A_i(r_i^1), \ldots, A_i(r_i^m) \text{ are distinct} \Big\}$$

for $a_i \in A_i(B)$ (provided we take $c(a_i)$ sufficiently small for $a_i \notin A_i(B)$).

This is proved in the Appendix. The choice of c in Lemma 3.2 is by no means unique even up to an additive constant. The case of this lemma when A_i is the identity is indicated in Green, Hylland, Pratt, and Zeckhauser (1984).

If B is not too large it should be much easier to check permutation dominance than to try to verify directly whether there is a c such that $EV_i(a_i | r_i) + c(a_i)$ is maximized at $a_i = A_i(r_i)$.[3]

From the remarks above we get:

PROPOSITION 3.3: *Under (IND), there exists an $\underline{A}$-inducing payment scheme $\underline{t}$ iff for all i, A_i is permutation dominant for $EV_i(a_i | r_i)$. In this case we can choose $\underline{t}$ to be budget balancing.*

PROOF: All except budget balancing follows immediately from Lemma 3.2 since we are assuming finiteness (FIN). To get budget balancing, let

$$t_i(z, \underline{a}) = c(a_i) - \frac{1}{n-1} \sum_{j \neq i} c(a_j)$$

with c as in the Lemma. Clearly $\underline{t}$ is budget balancing, and since agent i has no control over a_j for $j \neq i$, his incentives are as if $t_i = c$. *Q.E.D.*

[3] If B is large, however, permutation dominance is nontrivial to check. It is equivalent under (TR) to a form of traveling salesman problem which is known to be NP-complete (for instance, TSP-SUBOPTIMALITY in Lawler, Lenstra, Rinnooy Kan, and Shmoys (1985)). Any efficient algorithm for the usual travelling salesman problem gives an efficient algorithm for checking permutation dominance. However, by the next proposition the $\underline{A}$-inducement problem is itself NP-complete, so there is probably no polynomial-time algorithm for checking $\underline{A}$-inducement.

3.3. *The General Responsive Case*

Although the Lemmas above seem to apply only to the two fundamental cases mentioned so far, they can be used to resolve the general responsive case.

Given any r_i we define $\sim[r_i]$ to be $\{r_i': r_i' \sim r_i\}$, the equivalence class of r_i under $\sim$. We call any such class an *alikeness-class* for agent i.

The first main result follows easily from Lemma 3.2.

THEOREM 3.4: *Given direct return functions* $\underline{V}$ *and a strategy profile* $\underline{A}$, *if there is an* $\underline{A}$*-inducing payment scheme, then for all i and all alikeness classes B for agent i,* A_i *is permutation dominant for* EV_i *on B.*

PROOF: Suppose $(\underline{t}, \underline{A})$ is an equilibrium pair. For $a_i \in A_i(B)$ let $c(a_i) = Et_i(a_i|r_i)$ where $r_i \in B$. As $(\underline{t}, \underline{A})$ is an equilibrium, $EV_i(a_i|r_i) + Et_i(a_i|r_i)$ must be maximized at $a_i = A_i(r_i)$ for all $r_i \in B$. But this means $EV_i(A_i(r_i)|r_i) + c(A_i(r_i)) \geqslant EV_i(a_i|r_i) + c(a_i)$ for all $a_i \in A_i(B)$. So $\langle EV_i, A_i \rangle$ is a transfer maximum on B and therefore, by Lemma 3.2, A_i is permutation dominant for EV_i on B. *Q.E.D.*

Thus $\underline{A}$-inducement is impossible if permutation dominance fails on any alikeness class of any agent. Notice that the above proof made no use of (RESP). Responsiveness is essential only for our positive results in this section.

Given $\underline{A}$, each r_i corresponds to an action, namely $A_i(r_i)$. Without truth revelation one action may correspond to more than one type. Under responsiveness, however, each action a_i for agent i corresponds to a unique probability distribution over $\tilde{z}$ and $A_{-i}(\tilde{r}_{-i})$ (which we write as $P(-|a_i)$). If two actions a_i and a_i' correspond to the same distribution, then we say a_i is strongly similar to a_i' and write $a_i \sim a_i'$. That is, $a_i \sim a_i'$ iff there exist $r_i \sim r_i'$ with $a_i = A_i(r_i)$ and $a_i' = A_i(r_i')$. With this notation (BA) is equivalent to $a_i \sim a_i'$ iff $a_i = a_i'$.

THEOREM 3.5: *Under (RESP), given* $\underline{V}$ *and* $\underline{A}$, *there is an* $\underline{A}$*-inducing payment scheme* $\underline{t}$ *(not necessarily balancing the budget) iff, for all i and all alikeness classes B for agent i,* A_i *is permutation dominant for* EV_i *on B.*

PROOF: Theorem 3.4 gives us the necessity, so we need only prove sufficiency.

We will assume that $-\infty$ is a valid payment. If finite payments are required we can always find an M large enough so that $-\infty$ can be replaced by $-M$ throughout without affecting the validity of the result.

Since for each i and each alikeness class B, A_i is permutation dominant for EV_i on B, we have from Lemma 3.2 that $\langle EV_i, A_i \rangle$ is a transfer maximum on B for each B. Let c_i be a real-valued function such that $EV_i(a_i|r_i) + c_i(a_i)$ is maximized among the $a_i \in A_i(B)$ at $a_i = A_i(r_i)$, for each alikeness class B and $r_i \in B$. Now let $\hat{t}_i(z, \underline{a}) = \ln P(z, a_{-i}|a_i)$. Here $\hat{t}_i(z, \underline{a}) = -\infty$ is possible (if $\tilde{z} = z$ and $A_{-i}(\tilde{r}_{-i}) = a_{-i}$ is impossible given a_i). Invoking Lemma 3.1 we see that $E\hat{t}_i(A_i(r_i)|r_i) \geqslant E\hat{t}_i(A_i(r_i')|r_i)$, with equality holding iff $r_i \sim r_i'$.

Let $t_i(z, \underset{\sim}{a}) = L\hat{t}_i(z, \underset{\sim}{a}) + c_i(a_i)$ for all i, z and $\underset{\sim}{a}$. If L is sufficiently large, then a best action for any agent i with any type r_i must maximize $E\hat{t}_i(a_i|r_i)$ and must therefore be strongly similar to $A_i(r_i)$. Among the actions strongly similar to $A_i(r_i)$, $E\hat{t}_i(a_i|r_i)$ is constant and $EV_i(a_i|r_i) + c_i(a_i)$ is maximized at $a_i = A_i(r_i)$ (by definition of c). Hence $A_i(r_i)$ is a best action for agent i with type r_i and $(\underset{\sim}{t}, \underset{\sim}{A})$ is an equilibrium. *Q.E.D.*

COROLLARY 3.6: *Under (TR), given direct returns $\underset{\sim}{V}$, there is a truth-inducing payment scheme $\underset{\sim}{t}$ (not necessarily balancing the budget) iff, for all i and all alikeness classes B for agent i, truth is permutation dominant for EV_i on B.*

3.4. *Conditions on Beliefs Only*

The conditions in Theorem 3.5 involve both the beliefs of each agent and their direct returns. As d'Aspremont and Gérard-Varet (1982) do in their situation, we now give conditions involving only the beliefs of the agents (and the strategy profile $\underset{\sim}{A}$).

First we observe that it does not matter whether or not we require $\underset{\sim}{V}$ to be standard, even without assuming responsiveness, and even with budget balancing.

PROPOSITION 3.7: *For any strategy profile $\underset{\sim}{A}$, the following are equivalent. (a) For all $\underset{\sim}{V}$ there is an $\underset{\sim}{A}$-inducing payment scheme. (b) For all standard $\underset{\sim}{V}$ there is an $\underset{\sim}{A}$-inducing payment scheme. (c) There is a payment scheme $\underset{\sim}{t}$ such that:*

$$Et_i(A_i(r_i)|r_i) > Et_i(a_i|r_i) \quad \text{for all } i,\ r_i, \text{ and } a_i \neq A_i(r_i).$$

These remain equivalent if "payment scheme" is replaced throughout by "budget-balancing payment scheme."

PROOF: (a) ⇔ (b) by Proposition 2.1.

(a) ⇒ (c): Define $\underset{\sim}{V}$ by

$$V_i(z, \underset{\sim}{a}, \underset{\sim}{r}) = \begin{cases} 1 & \text{if } a_i \neq A_i(r_i) \text{ and } a_{-i} = A_{-i}(r_{-i}) \\ 0 & \text{otherwise.} \end{cases}$$

Under (a) there is a $\underset{\sim}{t}$ such that:

$$\begin{aligned} Et_i(A_i(r_i)|r_i) &= EV_i(A_i(r_i)|r_i) + Et_i(A_i(r_i)|r_i) \\ &\geqslant EV_i(a_i|r_i) + Et_i(a_i|r_i) \\ &= 1 + Et_i(a_i|r_i) > Et_i(a_i|r_i) \end{aligned}$$

$$\text{for all } i,\ r_i, \text{ and } a_i \neq A_i(r_i).$$

Hence (c) holds.

(c) ⇒ (a): Given $\underset{\sim}{t}$ as in (c) and given any direct returns $\underset{\sim}{V}$, by finiteness there is an L large enough so that $L\underset{\sim}{t}$ is $\underset{\sim}{A}$-inducing for $\underset{\sim}{V}$. Clearly this proof works in the budget-balancing case as well. *Q.E.D.*

In the case of budget-balancing, truth revelation, and no $\tilde{z}$, condition (c) above is the same as Condition B of d'Aspremont and Gérard-Varet (1982).

THEOREM 3.8: *Under (RESP), (BA) holds for all* $i \Leftrightarrow$ *for all* $\underline{V}$ *there is an* $\underline{A}$*-inducing payment scheme.*

PROOF: $\Rightarrow$: This follows from Theorem 3.5 since under (BA) an alikeness class B for agent i corresponds to a particular action a_i such that $r_i \in B$ iff $A_i(r_i) = a_i$. Permutation dominance holds trivially on such an alikeness class.

$\Leftarrow$: If (BA) fails for some i, then there are types $r_i \sim r_i'$ such that $A_i(r_i) \neq A_i(r_i')$ (by (RESP)). But then for any $\underline{t}$ satisfying (c) of Proposition 3.7,

$$Et_i(A_i(r_i)|r_i) > Et_i(A_i(r_i')|r_i) = Et_i(A_i(r_i')|r_i') > Et_i(A_i(r_i)|r_i')$$
$$= Et_i(A_i(r_i)|r_i),$$

a contradiction. *Q.E.D.*

As indicated in Section 2, we cannot replace "standard" by "public" in part (b) of Proposition 3.7. The result proved by d'Aspremont and Gérard-Varet (1979) that there always exists a transfer payment scheme for public $\underline{V}$ in the truth revelation case generalizes to the case of responsiveness.

PROPOSITION 3.9: *If* $\underline{V}$ *is public and if (RESP) holds, then there is always an* $\underline{A}$*-inducing payment scheme.*

PROOF: Let

$$_i(z, \underline{a}) = \sum_{j \neq i} E\{V_j(z, \underline{a}, \tilde{r}_j)|a_i\} + L \ln P(z, a_{-i}|a_i) \quad \text{for all } i.$$

If L is large enough, a best act for agent i must be strongly similar to $A_i(r_i)$. Among such acts, the total expected utility for agent i is equal to his expectation of the group's total direct return, which is maximized at $a_i = A_i(r_i)$ since $\underline{V}$ is public. *Q.E.D.*

Assuming consistent beliefs, d'Aspremont and Gérard-Varet (1982) show that this payment scheme can be chosen to balance the budget on expected value, a fact that also follows from the remarks at the beginning of this section.

4. GENERAL STRATEGY INDUCEMENT

This section presents an approach to the implementation question when responsiveness is not assumed. As noted in the last section, the permutation-dominance requirement in Theorem 3.4 does not depend on responsiveness. However, the converse of Theorem 3.4 is false, as the next result shows. For this we consider an assumption used in Crémer and McLean (1985).

For each agent i we define the *belief matrix* Q_i with rows indexed by the alikeness classes for agent i and columns indexed by all possible values of $(z, A_{-i}(r_{-i}))$. If B is an alikeness class for agent i and $r_i \in B$, let $Q_i(B, z, a_{-i})$ be $P(z, a_{-i}|r_i)$. By the definition of alikeness class, this value depends only on B and not on the particular representative r_i in B. The rows of Q_i are precisely the different possible beliefs agent i could have about $\tilde{z}$ and $A_{-i}(\tilde{r}_{-i})$. We call the alikeness classes *completely distinguishable* if:

(CD) For every i, the rows of Q_i are linearly independent.

This condition is used in Crémer and McLean (1985) (their Assumption 4) to construct a payment scheme where a seller can extract full surplus from a group of buyers in an auction. An indication that something like this condition is needed for $\underline{A}$-inducement is given by the following theorem.

THEOREM 4.1: *Given the designated strategies A_{-i} of the other agents, if the alikeness classes for agent i are not completely distinguishable (i.e., if (CD) fails), then there is a standard $\underline{V}$ and an A_i such that A_i is permutation dominant for EV_i on B for all alikeness classes B for agent i, but there is no $\underline{A}$-inducing payment scheme.*

PROOF: Suppose $Q_i(B_0) = \Sigma_{B \neq B_0} m(B) Q_i(B)$. Pick $a_i^0 \neq a_i^1$ and let

$$A_i(r_i) = \begin{cases} a_i^0 & \text{if } r_i \in B,\ m(B) > 0, \text{ and } B \neq B_0, \\ a_i^1 & \text{otherwise,} \end{cases}$$

and let

$$\overline{V}(\underline{a}, \underline{r}) = \begin{cases} 1 & \text{if } \underline{a} = \underline{A}(\underline{r}), \\ 0 & \text{otherwise.} \end{cases}$$

Let $V_i = -\overline{V}$ and $V_j = 2\overline{V}$ for $j \neq i$. Clearly A_i is permutation dominant for EV_i on B since there is only one element of $A_i(B)$ for each B. Suppose $(\underline{t}, \underline{A})$ is an equilibrium and thus gives the agents an incentive to act according to the group optimal strategy profile $\underline{A}$. Since agent i's interest is directly opposed to that of the group, we must have

$$Et_i(a_i^0|B) > Et_i(a_i^1|B) \quad \text{if } m(B) > 0 \text{ and } B \neq B_0,$$

$$Et_i(a_i^0|B) < Et_i(a_i^1|B) \quad \text{otherwise.}$$

But then $Et_i(a_i^0|B_0) = \Sigma_{B \neq B_0} m(B) Et_i(a_i^0|B) > \Sigma_{B \neq B_0} m(B) Et_i(a_i^1|B) = Et_i(a_i^1|B_0)$, which is a contradiction. *Q.E.D.*

This theorem suggests that somewhat different conditions from those of Section 3 will be needed when responsiveness is not assumed. One way to extend the results of Section 3 is to force a condition like responsiveness to hold by enlarging the alikeness classes of an agent. This method is explored in 4.2 using the general notion of an equivalence. Alternatively we can ask for some form of linear independence in the matrix Q_i such as complete distinguishability and use this (in a way similar to Crémer and McLean) to construct an $\underline{A}$-inducing payment scheme. These two approaches are combined to give more general positive results in Johnson, Pratt, and Zeckhauser (1988, Section IV.3).

4.2. *Equivalences*

An *equivalence* ρ is an n-tuple $(\rho_1, \dots, \rho_n)$ where for each i, ρ_i is an equivalence relation on the possible types of agent i. (We will write $r_i \rho r_i'$ in place of $r_i \rho_i r_i'$ since the added subscript is unnecessary.) An equivalence ρ partitions the set of possible types of each agent i into equivalence classes, which we will call *ρ-alikeness classes for agent i* or ρ_i-alikeness classes. An *equivalence on actions* is defined similarly.

Examples of equivalences include the two trivial ones, Δ (equality) and ∇, where $r_i \Delta r_i'$ iff $r_i = r_i'$ and $r_i \nabla r_i'$ for all r_i and r_i'. Of course $\sim$ is an equivalence. There is also an equivalence α naturally associated with the strategy profile $\underline{A}$ defined by $r_i \alpha r_i'$ iff $A_i(r_i) = A_i(r_i')$.

Given two equivalences ρ and σ, we call ρ *finer* than σ and write $\rho \leqslant \sigma$ if every ρ-alikeness class is contained in a σ-alikeness class (i.e., if $r_i \rho r_i' \Rightarrow r_i \sigma r_i'$). Thus Δ is the finest equivalence while ∇ is the coarsest. In general $\leqslant$ is only a *partial* order on the set of equivalences since two equivalences may be incomparable. We also define the *intersection* $\rho \cap \sigma$ by $r_i(\rho \cap \sigma) r_i'$ iff $r_i \rho r_i'$ and $r_i \sigma r_i'$. Clearly this always gives another equivalence which is the coarsest equivalence finer than both ρ and σ.

An equivalence ρ is called *responsive* if $\alpha \leqslant \rho$, i.e., if $A_i(r_i) = A_i(r_i') \Rightarrow r_i \rho r_i'$. If ρ is responsive, then it induces an equivalence on actions (which we also denote by ρ) by letting $A_i(r_i) \rho A_i(r_i')$ iff $r_i \rho r_i'$. In fact this gives a one-to-one correspondence between responsive equivalences and equivalences on actions.

Many of our assumptions from Section 3 can be expressed in terms of these equivalences. Specifically,

$$\text{(RESP)} \quad \Leftrightarrow \alpha \leqslant \sim,$$

$$\text{(BA)} \quad \Leftrightarrow \alpha = \sim,$$

$$\text{(IND)} \quad \Leftrightarrow \sim = \nabla,$$

$$\text{(TR)} \quad \Leftrightarrow \alpha = \Delta.$$

It is clear how any one of (BA), (IND), and (TR) implies (RESP).

Proper scoring works on the basis that an agent's action can be interpreted as announcing a belief about other information available to the center. We therefore focus on equivalences which can be defined in terms of such beliefs. An *i-external event* is any subset e of the possible values of (z, a_{-i}), and we let $P(e \mid r_i)$ denote the probability that $(\tilde{z}, A_{-i}(\tilde{r}_{-i})) \in e$ given type r_i. Given a set S_i of i-external events for each i we may define an equivalence $\text{ext}(\underline{S})$ by

(4.1) $\quad r_i \,\text{ext}(\underline{S})\, r_i'$ iff $P(e|r_i) = P(e|r_i')$ for all $e \in S_i$.

Informally, (4.1) says that r_i and r_i' are $\text{ext}(\underline{S})$-alike just when they correspond to the same beliefs about those events concerning the information available to the center which lie in S_i. We call any equivalence which can be defined in this way an *external equivalence*.

For example if S_i consists of all i-external events for each i, then $\text{ext}(\underline{S})$ is just $\sim$ (of course, it suffices here to take only all singleton events $\{(z, a_{-i})\}$ for each i). At the other extreme, if S_i is empty for each i, $\text{ext}(\underline{S})$ is the trivial equivalence ∇. Thus $\sim$ is the finest external equivalence while ∇ is the coarsest.

If ρ is external it may be definable from many different $\underline{S}$. Even when ρ is not external, however, there is a canonical vector of event sets $\underline{S}(\rho) = (S_1(\rho), \ldots, S_n(\rho))$ given by

$$S_i(\rho) = \{ e\colon r_i \rho r_i' \Rightarrow P(e|r_i) = P(e|r_i') \} = \bigcup_{\underline{S}\colon \rho \leqslant \text{ext}(\underline{S})} S_i.$$

This is the set of all external events which are stochastically independent of the types within every ρ-alikeness class. We call $\text{ext}(\underline{S}(\rho))$ the *external closure* of ρ (since it is the finest external equivalence coarser than ρ), and we denote it by $\bar{\rho}$. Of course, ρ is external iff $\rho = \bar{\rho}$. If ρ is responsive and if $e \in S_i(\rho)$, then we let $P(e|a_i)$ be $P(e|r_i)$ where $a_i = A_i(r_i)$ and note that this is well defined. An important equivalence when responsiveness is not assumed is $\bar{\alpha}$, the smallest equivalence which is both responsive and external.

Thus we have the following picture of equivalences:

$$\Delta \leqslant \left\{ \begin{array}{l} \sim \leqslant \\ \\ \alpha \leqslant \end{array} \begin{array}{l} \left\{ \begin{array}{l} \epsilon \text{ defines external} \\ \end{array} \right. \\ \bar{\alpha} \leqslant \text{ every external responsive } \sigma \\ \left\{ \begin{array}{l} \rho \text{ defines responsive} \end{array} \right. \end{array} \right\} \leqslant \nabla.$$

If b is a ρ-alikeness class for agent i and if $e \in S_i(\rho)$, then we define $P(e|b)$ to be $P(e|r_i)$ for some $r_i \in b$. By definition of $S_i(\rho)$, this definition is independent of the particular representative r_i of the alikeness class b.

If we now replace $\sim$ by $\bar{\alpha}$, the positive half of Theorem 3.5 still holds:

THEOREM 4.2: *If, for all $\bar{\alpha}$-alikeness classes B, A_i is permutation dominant for EV_i on B, then there exists an $\underline{A}$-inducing payment scheme $\underline{t}$.*

This is proved as a corollary of a more general result in Johnson, Pratt, and Zeckhauser (1988, Corollary IV.3).

5. BUDGET BALANCING

Until now we have ignored the constraint of budget-balancing:

(BB) $\quad \sum_{i=1}^{n} t_i(z, \underline{a}) = 0$ for all z and $\underline{a}$.

This has been a prime concern for many authors. In the case of public $\underline{V}$ and responsiveness it is the only case of difficulty, in view of Proposition 3.9, and d'Aspremont and Gérard-Varet in (1979, 1982) have studied the budget balance problem for the more general case of compact distributions.

We will not require $\underline{V}$ to be public, though some of our assumptions are related to those in d'Aspremont and Gérard-Varet (1979, 1982). As before, the situation is simpler under responsiveness (RESP), and we assume it here. Johnson, Pratt, and Zeckhauser (1988) give comparable results without responsiveness. Section 5.1 studies the budget balance problem for given $\underline{V}$, while 5.2 presents results involving conditions only on the agents' beliefs.

5.1. *Budget Balancing Under Responsiveness*

Just as $\sim$ was the important equivalence in Section 3, there will be a key equivalence β for the budget-balancing problem. In the responsive case without budget balancing we constructed payments depending on $\ln P(z, a_{-i}|a_i)$ which induced A_i to within $\sim$-alikeness. The difficulty with this if we are trying to balance the budget is that for any $j \neq i$, agent j cannot pay any part of $\ln P(z, a_{-i}|a_i)$ without (possibly) changing his incentives. We can solve this problem if we replace $P(z, a_{-i}|a_i)$ by $P(z, a_{-i-j}|a_i)$. If agent j pays a function of this to agent i, it will not affect the incentives of agent j. With an appropriate informational assumption $\ln P(z, a_{-i-j}|a_i)$ might also induce A_i up to $\sim$-alikeness, though in general it will only work up to a weaker alikeness, requiring permutation dominance on larger classes for full A_i-inducement. If some property corresponding to (BA) held with respect to (z, a_{-i-j}) then, using the techniques of Section 3 we could always balance the budget while inducing $\underline{A}$.

We can get by with weaker assumptions however. The remarks above motivate the first stage in the construction of the equivalence β which will play a crucial role in budget balancing. Before presenting this construction, some more notation is useful.

Let $\rho[r_i]$ denote the ρ-alikeness class of r_i, i.e., the set of types r_i' such that $r_i'\rho r_i$. Extending the notation introduced in Sections 3 and 4 for expressing probabilities, we write $\rho[r_i]$ within a probability expression to indicate the event $\tilde{r}_i \rho r_i$. Thus, for example, $P(z, r_{-i-j}, \rho[r_j]|r_i)$ is the probability that $\tilde{z} = z$, $\tilde{r}_{-i-j} = r_{-i-j}$ and $\tilde{r}_j \rho r_j$ given type r_i for agent i.

If ρ is responsive (as any external equivalence is if (RESP) is assumed), let $\rho[a_i]$ denote the ρ-alikeness class of a_i. We will use abbreviations such as $P(z, a_{-i-j}, \rho[a_j]|r_i)$ for agent i's probability that $\tilde{z} = z$, $A_{-i-j}(\tilde{r}_{-i-j}) = a_{-i-j}$, and $A_j(\tilde{r}_j)\rho a_j$ given type r_i.

We construct a sequence $\rho^1, \rho^2, \ldots$ of external equivalences inductively as follows (always assuming (RESP)).

Let $\rho^1 = \nabla$, so that $r_i \rho^1 r_i'$ for all i, r_i and r_i'. Given ρ^k, define ρ^{k+1} by $r_i \rho^{k+1} r_i'$ iff, for all $j \neq i$, all a_{-i} and all z,

$$(1_k) \qquad P\left(z, a_{-i-j}, \rho^k[a_j]|r_i\right) = P\left(z, a_{-i-j}, \rho^k[a_j]|r_i'\right).$$

That is, r_i and r_i^1 are ρ^{k+1}-equivalent if, for each $j \neq i$, they correspond to the same joint beliefs about the actions of all agents but agent j, and the action of agent j up to ρ^k-equivalence. The point of this construction is, once we are guaranteed that $\underline{A}$ can always be induced to within ρ^k-alikeness, we are free to let agent j make payments which only depend on the ρ^k-alikeness class for his action. Doing so may allow us to induce $\underline{A}$ to within a stronger alikeness, and so on.

This sequence has the following properties. (i) ρ^k is external, and therefore responsive for all $k \geqslant 1$. (ii) $\rho^{k+1} \leqslant \rho^k$ for all $k \geqslant 1$. (iii) There exists a K with $\rho^k = \rho^K$ iff $k \geqslant K$. (K is the stage at which the inductive construction stabilizes.)

PROOF: (i) $\rho^1 = \nabla$, which is external as already noted. Assuming ρ^k is external, we have $\sim \leqslant \rho^k$ and therefore $\alpha \leqslant \rho^k$ by responsiveness. Thus $\rho^k[a_j]$ is defined and ρ^{k+1} is external.

(ii) Clearly $\rho^2 \leqslant \rho^1 = \nabla$. Suppose inductively that $\rho^k \leqslant \rho^{k-1}$. If $r_i \rho^{k+1} r_i'$ then for all $j \neq i$, a_{-i} and z, (1_k) holds. Summing (1_k) over all ρ^k-alikeness classes for agent j which lie in the ρ^{k-1}-alikeness class of a_j gives (1_{k-1}) and hence $r_i \rho^k r_i'$. Thus $\rho^{k+1} \leqslant \rho^k$.

(iii) This follows from (ii) and finiteness. *Q.E.D.*

We define β to be ρ^K where K is as in (iii). Thus β satisfies:

(5.1) For all i, r_i and r_i': $r_i \beta r_i'$ if and only if for all $j \neq i$, all a_{-i} and all z, $P(z, a_{-i-j}, \beta[a_j] | r_i) = P(z, a_{-i-j}, \beta[a_j] | r_i')$.

(In fact, β is the coarsest such external equivalence.) If there are only two agents and no $\tilde{z}$, this process gives $\rho^k = \nabla$ for all k and hence $\beta = \nabla$. Otherwise it is not hard to show that $K \leqslant 3 + \sum_{i=1}^n (T_i - 2)$ where T_i is the number of possible types for agent i. In particular, if there are only two possible types for each agent, $K \leqslant 3$. However, given any k and $n \geqslant 3$, there exist consistent beliefs of n agents such that the construction above does not stop for at least k stages (i.e., $K \geqslant k$). In most practical cases a calculation of β should take only a few stages.

Since β is external, we have $\sim \leqslant \beta$.

THEOREM 5.1: *Under (RESP), if for all i and all β-alikeness classes B for agent i, A_i is permutation dominant for EV_i on B, then there exists an $\underline{A}$-inducing, budget-balancing payment scheme $\underline{t}$.*

PROOF: We define a sequence of payment schemes $\underline{t}^K, \underline{t}^{K-1}, \ldots, \underline{t}^1$ such that for all k:

$$(2_k) \qquad EV_i(A_i(r_i)|r_i) + Et_i^k(A_i(r_i)|r_i) \geqslant EV_i(a_i|r_i) + Et_i^k(a_i|r_i)$$

for all i, r_i, and a_i such that $a_i \rho^k A_i(r_i)$;

$$(3_k) \qquad \sum_{i=1}^n t_i^k(z, \underline{a}) = 0 \quad \text{for all } z \text{ and } \underline{a}.$$

This construction proceeds by backward induction. As A_i is permutation dominant for EV_i on B for any β-alikeness class B, we have from Lemma 3.2

that $\langle EV_i, A_i \rangle$ is a transfer maximum on B for each such B. Let c_i be a real-valued function such that for every r_i, the maximum of $EV_i(a_i|r_i) + c_i(a_i)$ over the $a_i \in \beta[A_i(r_i)]$ is attained at $a_i = A_i(r_i)$. Let

$$t_i^K(z, \underline{a}) = c_i(a_i) - \frac{1}{n-1} \sum_{j \neq i} c_j(a_j).$$

Then (2_K) holds, since the second term in the definition of t_i^K is independent of agent i's action, and (3_K) holds easily.

Suppose now that we are given a $\underline{t}^{k+1}$ satisfying (2_{k+1}) and (3_{k+1}). Let

$$\ln_M(x) = \begin{cases} \ln x & \text{if } x > 0, \\ -M & \text{if } x = 0, \end{cases}$$

and let

$$\hat{t}_i^k(z, \underline{a}) = \sum_{j \neq i} \ln_M P\left(z, a_{-i-j}, \rho^k[a_j] \| a_i\right) - \sum_{j \neq i} \ln_M P\left(z, a_{-i-j}, \rho^k[a_i] \| a_j\right).$$

By Lemma 3.1, if we take M to be ∞ so that $\ln_M x$ is just $\ln x$, then $E\hat{t}_i^k(A_i(r_i)|r_i) \geqslant E\hat{t}_i^k(a_i'|r_i)$ for all a_i' such that $a_i' \rho^k A_i(r_i)$ (notice that the second term in the definition of t_i^k is constant on ρ^k-alikeness classes), with equality iff $a_i' \rho^{k+1} A_i(r_i)$. By (FIN) there is a finite M large enough so that the above still remains true. Let $t_i^k(z, \underline{a}) = t_i^{k+1}(z, \underline{a}) + L\hat{t}_i^k(z, \underline{a})$ for all i, z and $\underline{a}$. If L is sufficiently large, then for any agent i and any type r_i, a best action from the ρ^k-alikeness class of $A_i(r_i)$ for agent i must maximize $E\hat{t}_i(a_i, r_i)$ and must therefore be ρ^{k+1}-alike to $A_i(r_i)$. But among the actions which are ρ^{k+1}-alike to $A_i(r_i)$, $E\hat{t}_i(a_i|r_i)$ is constant and $Et_i^{k+1}(a_i|r_i)$ is maximized at $A_i(r_i)$ (by the induction hypothesis). Thus $\underline{t}^k$ satisfies (2_k), and by its definition (and the induction hypothesis) it satisfies (3_k). So the induction argument is complete.

As $\rho^1 = \nabla$, $\underline{t}^1$ is budget balancing and $\underline{A}$-inducing, since we are assuming that each agent i must choose an action in the range of A_i. *Q.E.D.*

Comparing this result with Theorem 3.5 we see that the hypotheses of Theorem 5.1 are necessary as well as sufficient under the assumption $\beta = \sim$. In fact, we get necessary and sufficient conditions under the following weaker assumption:

(LINK) $\alpha \leqslant \sim$ and for some i, $\beta_i = \sim_i$.

This condition says that the inductive construction of β eventually collapses to give the equivalence $\sim$ for some agent i. Later (Theorem 5.7) we will show that (CON) and (LINK) together imply $\beta = \sim$ (i.e., $\beta_j = \sim_j$ for *all* j). A special case of (LINK) is Condition F of d'Aspremont and Gérard-Varet (1982) which states that some agent's type is independent of the other agents' types (and of $\tilde{z}$) so that $\sim_i = \nabla_i = \beta_i$. Theorem 3.5 and Theorem 5.1 combine to give the following corollary.

COROLLARY 5.2: *Under (LINK) (and therefore (RESP)), there is an $\underline{A}$-inducing payment scheme* $\Leftrightarrow$ *there is a budget-balancing, $\underline{A}$-inducing payment scheme* $\Leftrightarrow$ *for all j and all $\sim$-alikeness classes B for agent j, A_j is permutation dominant for EV_j on B.*

PROOF: The only thing left to prove is sufficiency of (LINK) for budget balancing. So suppose $\beta_i = \sim_i$. By the same argument as used in the proof of Proposition 3.3, we can find a budget-balancing $\underline{t}^{K+1}$ such that

$$EV_j(A_j(r_j)|r_j) + Et_j^{K+1}(A_j(r_j)|r_j) \geqslant EV_j(a_j|r_j) + Et_j^{K+1}(a_j|r_j)$$

for all j, r_j, and a_j such that $a_j \sim A_j(r_j)$.

For $j \neq i$, let $\hat{t}_j^K(z, \underline{a}) = \ln P(z, a_{-j}|a_j)$ and let $t_j^K(z, \underline{a}) = t_j^{K+1}(z, \underline{a}) + L(\hat{t}_j^K(z, \underline{a}) - E\hat{t}_j^K(a_i|a_i))$ where L is large. Let $t_i^K(z, \underline{a}) = t_i^{K+1}(z, \underline{a}) + L\Sigma_{j\neq i}(E\hat{t}_j^K(a_i|a_i) - \hat{t}_j^K(z, \underline{a}))$ so that $\underline{t}^K$ is budget-balancing. If L is sufficiently large, $EV_j(a_j|r_j) + Et_j^K(a_j|r_j)$ is maximized at $a_j = A_j(r_j)$ for all $j \neq i$ and all r_j.

For agent i, if $a_i \beta a_i'$, then

$$Et_i^K(a_i'|a_i) = Et_i^{K+1}(a_i'|a_i) + L\left(\sum_{j\neq i} E\hat{t}_j^K(a_i'|a_i') - E\hat{t}_j^K(a_i'|a_i)\right)$$

$$= Et_i^{K+1}(a_i'|a_i),$$

since we actually have $a_i \sim a_i'$ by (LINK). Thus (2_K) and (3_K) of the proof of Theorem 5.1 hold and the rest of the induction argument of that proof now applies. *Q.E.D.*

COROLLARY 5.3: *Under (LINK) there is a budget-balancing, $\underline{A}$-inducing payment scheme for all public $\underline{V}$.*

PROOF: This follows immediately from Proposition 3.9 and Corollary 5.2. *Q.E.D.*

It can be shown that under the assumption of consistent beliefs, (LINK) is equivalent to a compatibility condition which we denote by (C). It is the discrete case of the compatibility condition of d'Aspremont and Gérard-Varet (1982). In the Appendix we define (C) formally and prove the following proposition.

PROPOSITION 5.4: *Assuming (CON), (TR), and no $\tilde{z}$, (LINK)* $\Leftrightarrow$ *(C).*

As d'Aspremont and Gérard-Varet showed in (1982), the converse of Corollary 5.3 is not true. Their counterexample can be generalized[4] to show that given (CON), no $\tilde{z}$, and only two agents each with two states of the world, there is always a budget-balancing, truth-inducing payment scheme for any public $\underline{V}$. Of

[4] The details of this demonstration are rather lengthy and are not included here.

course, in this case (LINK) is equivalent to (IND) and is therefore far from necessary for the conclusion of Corollary 5.3.

Our next result is an impossibility result for budget-balancing and relies on the assumption (CON) that agents have consistent beliefs.

Assuming (CON), let $P(z, \underline{a})$ denote the probability that $\tilde{z} = z$ and $\underline{A}(\tilde{r}) = \underline{a}$ and let $P(a_i' | \beta[a_i])$ denote the conditional probability that $A_i(\tilde{r}_i) = a_i'$ given that $A_i(\tilde{r}_i) \beta a_i$. We will also use similar abbreviations following the same pattern. Given a direct return $\underline{V}$, we define a modified direct return $\underline{V}'$ such that V_i' gives agent i the same direct return for actions in the same β-alikeness class:

$$(5.2) \qquad V_i'(z, \underline{a}, \underline{r}) = \sum_{a_i' \beta a_i} P(a_i' | \beta[a_i]) V_i(z, a_{-i}, a_i', \underline{r}).$$

The sufficient conditions for budget-balancing $\underline{A}$-inducement given in Theorem 5.1 fail to be necessary. This is not surprising, since these conditions do not involve any utility comparisons between agents, while budget balancing is a matter of making transfers between agents. Our next theorem derives necessary conditions for budget balancing in the case of consistent beliefs and responsiveness.

THEOREM 5.5: *Under (CON) and (RESP), if $\underline{V}$ is such that*

$$(5.3) \qquad E\left(\sum_i V_i(\tilde{z}, \underline{A}(\tilde{\underline{r}}), \tilde{\underline{r}})\right) < E\left(\sum_i V_i'(\tilde{z}, \underline{A}(\tilde{\underline{r}}), \tilde{\underline{r}})\right),$$

then there is no budget-balancing transfer payment scheme which is $\underline{A}$-inducing.

PROOF: Suppose $\underline{t}$ is such a payment scheme. Let $a_i = A_i(r_i)$ and $a_i' = A_i(r_i')$ to simplify notation. As t_i is A_i-inducing, we must have

$$EV_i(a_i | r_i) + Et_i(a_i | r_i) \geq EV_i(a_i' | r_i) + Et_i(a_i' | r_i)$$

and

$$EV_i(a_i' | r_i') + Et_i(a_i' | r_i') \geq EV_i(a_i | r_i') + Et_i(a_i | r_i').$$

Adding these together and multiplying by $P(r_i)P(r_i')$ gives:

$$\sum_{z, a_{-i}} [P(z, \underline{a}) P(r_i') - P(z, a_{-i}, a_i') P(r_i)](t_i(z, \underline{a}) - t_i(z, a_{-i}, a_i'))$$

$$\geq \sum_{z, r_{-i}} [P(z, \underline{r}) P(r_i')(V_i(z, A_i(r_{-i}), a_i', \underline{r}) - V_i(z, A_{-i}(\underline{r}_i), a_i, \underline{r}))$$

$$- P(z, r_{-i}, r_i') P(r_i)(V_i(z, A_{-i}(r_{-i}), a_i', r_{-i}, r_i')$$

$$- V_i(z, A_{-i}(r_{-i}), a_i, r_{-i}, r_i'))].$$

Summing this over all pairs r_i and r_i' such that $r_i \beta r_i'$, and dividing by $P(\beta[r_i])$, gives

$$\sum_{z,\underline{a}} \sigma_i(z,\underline{a}) t_i(z,\underline{a}) \geqslant EV_i'(\tilde{z},\underline{A}(\tilde{\underline{r}}),\tilde{\underline{r}}) - EV_i(\tilde{z},\underline{A}(\tilde{\underline{r}}),\tilde{\underline{r}})$$

where

$$\sigma_i(z,\underline{a}) = P(z,\underline{a}) - P(z,a_{-i},\beta[a_i])P(a_i|\beta[a_i]).$$

The proof will be complete once we have shown that $\sigma_i(z,\underline{a}) = \sigma(z,\underline{a})$ does not depend on i, for then we will have

$$0 = \sum_{z,\underline{a}} \sigma(z,\underline{a}) \sum_i t_i(z,\underline{a})$$
$$\geqslant E\left(\sum_i V_i'(\tilde{z},\underline{A}(\tilde{\underline{r}}),\tilde{\underline{r}})\right) - E\left(\sum_i V_i(\tilde{z},\underline{A}(\tilde{\underline{r}}),\tilde{\underline{r}})\right).$$

Hence let us calculate (making use of (CON))

$$\begin{aligned}
&\sigma_i(z,\underline{a}) - \sigma_j(z,\underline{a})\\
&\quad = P(a_j|\beta[a_j])P(z,a_{-j},\beta[a_j])\\
&\qquad\qquad - P(a_i|\beta[a_i])P(z,a_{-i},\beta[a_i])\\
&\quad = P(z,a_{-i-j},\beta[a_j],\beta[a_i])\\
&\qquad \times\Big(P(a_j|\beta[a_j])P(a_i|z,a_{-i-j},\beta[a_i],\beta[a_j])\\
&\qquad\qquad - P(a_i|\beta[a_i])P(a_j|z,a_{-i-j},\beta[a_i],\beta[a_j])\Big).
\end{aligned}$$

Now from (5.1), for all i and $j \neq i$, $(\tilde{z}, A_{-i-j}(\tilde{r}_{-i-j}), \beta[A_j(\tilde{r}_j)])$ is independent of $A_i(\tilde{r}_i)$ given $A_i(\tilde{r}_i)\beta a_i$. Thus

$$P(a_j|\beta[a_j]) = P(a_j|z,a_{-i-j},\beta[a_i],\beta[a_j]) \quad \text{and}$$
$$P(a_i|\beta[a_i]) = P(a_i|z,a_{-i-j},\beta[a_i],\beta[a_j]),$$

so that $\sigma_i(z,\underline{a}) = \sigma_j(z,\underline{a})$. *Q.E.D.*

5.2. *Conditions on Beliefs for Budget Balancing*

We get as an easy corollary of our results in the responsive case:

THEOREM 5.6: *Under (RESP) and (CON), $\beta = \alpha \Leftrightarrow$ for all $\underline{V}$ there is a budget-balancing, $\underline{A}$-inducing payment scheme.*

PROOF: $\Rightarrow$: By Theorem 5.1 it suffices to have permutation dominance on all β-alikeness classes, but each of these corresponds to a single action if $\beta = \alpha$, so permutation dominance is trivial.

$\Leftarrow$: Let $\underline{V}$ be as in the proof of Proposition 3.7 so that, in particular, $V_i(z,\underline{A}(\underline{r}),\underline{r}) = 0$ for all z and $\underline{r}$, while $V_i'(z,\underline{A}(\underline{r}),\underline{r}) = \sum_{a_i' \beta a_i:\, a_i' \neq a_i} P(a_i'|\beta[a_i])$

$\geqslant 0$ where $\underline{V}'$ is defined by (5.2) and $a_i = A_i(r_i)$. If $\beta \neq \alpha$, then there are $a_i' \beta a_i$ with $a_i' \neq a_i$, so that $V_i'(z, \underline{A}(\underline{r}), \underline{r}) > 0$ for some i, z, and $\underline{r}$ with $P(z, \underline{r}) > 0$. Thus inequality (5.3) of Theorem 5.5 holds. *Q.E.D.*

In fact, in the case of truth revelation ($\alpha = \Delta$), consistent beliefs, and no $\tilde{z}$, the condition $\beta = \alpha$ becomes $\beta = \Delta$ which is then equivalent to Condition B of d'Aspremont and Gérard-Varet (1982). This follows from Proposition 3.7, Theorem 5.6, and the fact that in this case, condition (c) of Proposition 3.7 is the Condition B just mentioned. Because β is computable from its inductive definition, this shows exactly how much stochastic relevance between the agents' types is needed for Condition B to hold, in the finite case of consistent beliefs.

Our last theorem characterizes the condition (LINK) as the condition that budget balancing always comes for free.

THEOREM 5.7: *Under (RESP) and (CON) the following are equivalent*: (a) $\sim\, = \beta$; (b) (*LINK*); (c) *for all* $\underline{V}$, *if there is an* $\underline{A}$-*inducing payment scheme then there is a budget-balancing*, $\underline{A}$-*inducing payment scheme*; (d) *for all standard* $\underline{V}$, *if there is an* $\underline{A}$-*inducing payment scheme then there is a budget-balancing*, $\underline{A}$-*inducing payment scheme.*

PROOF: (a) $\Rightarrow$ (b) and (c) $\Rightarrow$ (d) are trivial, while (b) $\Rightarrow$ (c) follows at once from Corollary 5.2. To show (d) $\Rightarrow$ (a), let $\underline{V}$ be given by

$$V_i(z, \underline{a}, \underline{r}) = \begin{cases} -2 & \text{if } a_j \neq A_j(r_j) \text{ some } j \neq i, \\ 1 & \text{if } a_i \nsim A_i(r_i) \text{ and } a_{-i} = A_{-i}(r_{-i}) \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $\underline{V}$ is standard. Also, $V_i(z, \underline{A}(\underline{r}), \underline{r}) = 0$ for all z and $\underline{r}$, while

$$V_i'(z, \underline{A}(\underline{r}), \underline{r}) = \sum_{\substack{a_i' \beta a_i \\ a_i' \nsim a_i}} P(a_i' | \beta[a_i]) \geqslant 0 \quad \text{where} \quad a_i = A_i(r_i).$$

If $\beta \neq \sim$ then there are $a_i' \beta a_i$ with $a_i' \nsim a_i$ so that $V_i'(a, \underline{A}(\underline{r}), \underline{r}) > 0$ for some i, z, and $\underline{r}$ with $P(z, \underline{r}) > 0$. Thus $E(\Sigma_i V_i'(\tilde{z}, \underline{A}(\tilde{\underline{r}}), \tilde{\underline{r}})) > E(\Sigma_i V_i(\tilde{z}, \underline{A}(\tilde{\underline{r}}), \tilde{\underline{r}}))$ and by Theorem 5.5 there fails to be a budget-balancing, $\underline{A}$-inducing payment scheme. On the other hand, $EV_i(a_i | r_i) = 0$ whenever $a_i \sim A_i(r_i)$, so permutation dominance holds trivially on all $\sim$-alikeness classes and hence there is an $\underline{A}$-inducing payment scheme which cannot be budget balancing. *Q.E.D.*

6. CONCLUSION

This analysis demonstrates the possibility of inducing risk neutral agents to take actions and reveal private information in a manner that achieves a specified outcome even though the acts and information are relevant to the payoffs of other agents and interests diverge. A central authority oversees the process and makes relevant transfer payments. A necessary condition (in the finite case) is that, in the absence of any transfer payments, no agent should prefer to permute the acts he is called upon to make given an equal-chance lottery over an alikeness class (an information set that the center cannot monitor, even probabilistically, on the basis of its information and other agents' actions). Under this condition, strategy-inducing transfers exist if in addition the conditional distributions of the center's monitoring information given the agents' alikeness classes are linearly independent. Any agent for which this fails can merely be called upon to reveal his alikeness class. (Analogous conditions apply when not all other agents' acts are used in monitoring.) The budget can be balanced as well, if the above conditions hold with a weaker notion of alikeness.

Here we have dealt only with the finite case. Unfortunately, results for the continuous case do not follow by taking limits (or by analogy) because the transfer functions may have infinite limits or no limits as the finite approximation approaches a continuous problem. We have begun work on the continuous case and also on special conditions that permit efficient coordination despite risk aversion or possibly collusive behavior on the part of the agents.

A central challenge in the design of an economic system is to develop procedures that are effective despite privately-held information that is important to the well-being of others. The information may relate to such matters as future market conditions, pollution effects, or anti-competitive behavior. Here we show under what particular information conditions the use of financial incentives can accomplish this task, and when they cannot.

Department of Mathematics, Australian National University, Canberra, Australia,
Harvard Business School, Soldiers Field, Boston, MA 02163, U.S.A.
and
John F. Kennedy School of Government, Harvard University, Cambridge, MA 02138, U.S.A.

Manuscript received November, 1986; final revision received May, 1989.

APPENDIX

The version of Lemma 3.2 which we prove here is stated in more general terms, primarily to avoid a clutter of symbols. If B and A_0 are finite sets, if λ is a function from B to A_0, and if V is any real-valued function on $B \times A_0$, then we can clearly generalize the notions of transfer maximum and permutation dominance.

Call $\langle V, \lambda \rangle$ a *transfer maximum* if there exists a real-valued function c on A_0 such that $V(r, a) + c(a)$ is maximized at $a = \lambda(r)$ for all r. Call λ *permutation dominant* for V if, for all permutations π of $\lambda(B)$ $(= \{\lambda(r): r \in B\})$ and for all representative functions $r, \Sigma_{a \in \lambda(B)} V(r_a, a) \geqslant \Sigma_{a \in \lambda(B)} V(r_a, \pi(a))$ (where r is a *representative function* if $\lambda(r_a) = a$ for all $a \in \lambda(b)$).

Lemma 3.2: *$\langle V, \lambda \rangle$ is a transfer maximum iff λ is permutation dominant for V. In this case we may take*

$$c(a) = \min \left\{ \sum_{k=1}^{m} [V(r_k, \lambda(r_k)) - V(r_k, \lambda(r_{k+1}))] : \lambda(r_{m+1}) = a,\ m \geqslant 1, \right.$$
$$\left. \text{and } \lambda(r_1), \ldots, \lambda(r_m) \text{ are distinct} \right\}$$

for $a \in \lambda(B)$ (provided we take $c(a)$ sufficiently small for $a \notin \lambda(B)$).

PROOF ⇒: If $\langle V, \lambda \rangle$ is a transfer maximum, then for any $\{r_a: a \in A_0\}$ such that $\lambda(r_a) = a$ for all a, and any permutation π of A_0:

$$\begin{aligned}\sum_{a \in A_0} V(r_a, a) &= -\sum_{a \in A_0} c(a) + \sum_{a \in A_0} \left(V(r_a, a) + c(a)\right) \\ &\geqslant -\sum_{a \in A_0} c(a) + \sum_{a \in A_0} \left(V(r_a, \pi(a)) + c(\pi(a))\right) \\ &= \sum_{a \in A_0} V(r_a, \pi(a)).\end{aligned}$$

⇐: Suppose $V(r, \lambda(r)) + c(\lambda(r)) < V(r, a^0) + c(a^0)$ for some $r \in B$, $a^0 \in A_0$, where c is defined as in the statement of the lemma. Then for some $r_1, \dots, r_{m+1}$ with $\lambda(r_1), \dots, \lambda(r_m)$ distinct and $\lambda(r_{m+1}) = \lambda(r)$:

$$V(r, \lambda(r)) + \sum_{k=1}^{m} \left[V(r_k, \lambda(r_k)) - V(r_k, \lambda(r_{k+1}))\right] < V(r, a^0) + c(a^0).$$

In fact we may take $r_{m+1} = r$. So $c(a^0) > \sum_{k=1}^{m+1} V(r_k, \lambda(r_k)) - V(r_k, \lambda(r_{k+1}))$ where $\lambda(r_{m+2}) = a^0$. This contradicts the definition of $c(a^0)$ unless $\lambda(r_k) = \lambda(r)$ for some (unique) $k_0 \leqslant m$.

By permutation dominance and the fact that $\lambda(r_{k_0}) = \lambda(r_{m+1})$,

$$\sum_{k=k_0}^{m} V(r_k, \lambda(r_k)) \geqslant \sum_{k=k_0}^{m} V(r_k, \lambda(r_{k+1})), \quad \text{so}$$

$$\begin{aligned}c(a^0) > \sum_{k=1}^{k_0-1} &\left(V(r_k, \lambda(r_k)) - V(r_k, \lambda(r_{k+1}))\right) \\ &+ V(r_{m+1}, \lambda(r_{m+1})) - V(r_{m+1}, \lambda(r_{m+2})).\end{aligned}$$

This contradicts the definition of $c(a^0)$ (even if $k_0 = 1$). Hence $V(r, \lambda(r)) + c(\lambda(r)) \geqslant V(r, a) + c(a)$ for all $r \in B$ and $a \in \lambda(B)$.

For those a in A_0 which are not in $\lambda(B)$, we can clearly make $c(a)$ negative enough so that $\langle V, \lambda \rangle$ is a transfer maximum. *Q.E.D.*

We next prove Proposition 5.4 in a more general setting. We assume responsiveness, so that probabilities conditional on acts make sense. Let Λ_i denote the set of all nonnegative real valued functions $\lambda_i(a_i, a_i')$ where a_i and a_i' are distinct acts for agent i. Let Λ_i^s denote the subset of Λ_i consisting of those λ_i which are symmetric, i.e. $\lambda_i(a_i, a_i') = \lambda_i(a_i', a_i)$. We consider the following conditions on beliefs (always assuming (RESP)).

(C) (compatibility):

For all $\underline{\lambda} \in \Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n$, if for all $i, z, \underline{a}$,

$$\sum_{a_i': a_i' \neq a_i} \left(P(z, a_{-i} | a_i) \lambda_i(a_i', a_i) - P(z, a_{-i} | a_i') \lambda_i(a_i, a_i')\right) = \kappa(z, \underline{a})$$

where κ does not depend on i, then $\kappa \equiv 0$.

(STRC) (strong compatibility):

Replace $\lambda_i(a_i', a_i)$ by $\lambda_i(a_i, a_i')$ in (C).

(SYMC) (symmetric compatibility):

Replace $\Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n$ by $\Lambda_1^s \times \Lambda_2^s \times \cdots \times \Lambda_n^s$ in (C).

Step 3 of the proof of Theorem 7 of d'Aspremont and Gérard-Varet (1979) (which is clearly valid for our slightly more general setting) shows:

if for all $i, z, \underline{a}$

$$(*) \qquad \sum_{a_i': a_i' \neq a_i} \left(P(z, a_{-i} | a_i) \lambda_i(a_i, a_i) - P(z, a_{-i} | a_i') \lambda_i(a_i', a_i')\right) = \kappa(z, \underline{a})$$

where κ does not depend on i, then for all i, a_i,

$$\sum_{a_i': a_i' \neq a_i} \lambda_i(a_i', a_i) = \sum_{a_i': a_i' \neq a_i} \lambda_i(a_i, a_i').$$

From $(*)$ it is clear that (STRC) $\Rightarrow$ (C) while it is trivial that (C) $\Rightarrow$ (SYMC). In the case of (TR) and no $\tilde{z}$, (C) is the discrete case of the compatibility condition of d'Aspremont and Gérard-Varet (1982), and (STRC) is the compatibility condition of d'Aspremont and Gérard-Varet (1979). We will show below that (STRC) is *strictly* stronger than (C) even under (CON). First we have the following proposition.

PROPOSITION 5.4: *Under (RESP), each of the following is equivalent to (C).* (a) *For all $\underline{V}$, if there is an $\underline{A}$-inducing payment scheme then there is a budget-balancing, $\underline{A}$-inducing payment scheme*; (b) *there is a budget-balancing payment scheme $\underline{t}$ such that for all $i, a_i, a_i', Et_i(a_i|a_i) - Et_i(a_i'|a_i) \geqslant 0$ with equality only if $a_i \sim a_i'$. Under (CON), these conditions are also equivalent to (LINK) and to (SYMC).*

PROOF: (C) $\Rightarrow$ (a): This is essentially proved in d'Aspremont and Gérard-Varet (1982) since everything proved there about public $\underline{V}$ applies equally well to those $\underline{V}$ for which there exists an $\underline{A}$-inducing payment scheme. In particular, the compatibility condition implies the existence of a *budget-balancing*, $\underline{A}$-inducing payment scheme for such $\underline{V}$.

(a) $\Rightarrow$ (b): Consider the utility profile

$$V_i(z, \underline{a}, \underline{r}) = \begin{cases} 1 & \text{if } a_i \nsim A_i(r_i), \\ 0 & \text{if } a_i \sim A_i(r_i). \end{cases}$$

There is an $\underline{A}$-inducing payment scheme by Theorem 3.5, so under (a) there is a budget-balancing, $\underline{A}$-inducing payment scheme $\underline{t}$. Clearly $\underline{t}$ must satisfy the inequalities of (b).

(b) $\Rightarrow$ (C): By finiteness, (b) is equivalent to the existence of a $\underline{t}$ such that $\Sigma_{i=1}^{n} t_i(z, \underline{a}) = 0$ for all z and $\underline{a}$, and

$$\sum_{z, a_{-i}} P(z, a_{-i}|a_i)(t_i(z, \underline{a}) - t_i(z, a_{-i}, a_i')) \geqslant \begin{cases} 1 & \text{if } a_i' \nsim a_i, \\ 0 & \text{if } a_i' \sim a_i. \end{cases}$$

By Theorem 1 of Ky Fan (1956), a necessary and sufficient condition for the above system of inequalities to be consistent is that the "if" clause of (C) imply $\lambda_i(a_i', a_i) = 0$ whenever $a_i' \nsim a_i$. By $(*)$ and the definition of $\sim$ this implies (C).

Now assume (CON). We already know (by Theorem 5.7) that (LINK) $\Leftrightarrow$ (a). Also (C) $\Rightarrow$ (SYMC) trivially. Thus it suffices to show that (SYMC) $\Rightarrow$ (LINK).

Let

$$\lambda_i(a_i, a_i') = \begin{cases} P(a_i)P(a_i')/P(\beta[a_i]) & \text{if } a_i \beta a_i', \\ 0 & \text{otherwise.} \end{cases}$$

Then for all i, a_i, a_i',

$$\sum_{a_i' :\, a_i' \neq a_i} \lambda_i(a_i, a_i')(P(z, a_{-i}|a_i) - P(z, a_{-i}|a_i'))$$
$$= \sum_{a_i' :\, a_i' \beta a_i} P(a_i')P(a_i|\beta[a_i])(P(z, a_{-i}|a_i) - P(z, a_{-i}|a_i'))$$
$$= P(z, \underline{a}) - P(z, a_{-i}, \beta[a_i])P(a_i|\beta[a_i]) = \sigma(z, \underline{a}),$$

which from the proof of Theorem 5.5 does not depend on i. From (SYMC) we get $\sigma(z, \underline{a}) = 0$ for all $z, \underline{a}$. But then

$$\sigma(z, \underline{a}) = P(a_i)(P(z, a_{-i}|a_i) - P(z, a_{-i}|\beta[a_i])) = 0$$

for all $z, \underline{a}$, which can only happen if $\beta = \sim$. *Q.E.D.*

The following example shows that (C) $\nRightarrow$ (STRC) even under (CON). Assume (TR) and no $\tilde{z}$ and suppose there are three agents with two types each (types 0 and 1 say). Suppose beliefs are derived from the following joint prior distribution:

$\tilde{r}_3 = 0$:

	$\tilde{r}_2 = 0$	$\tilde{r}_2 = 1$
$\tilde{r}_1 = 0$	1/16	3/16
$\tilde{r}_1 = 1$	3/16	1/16

$\tilde{r}_3 = 1$:

	$\tilde{r}_2 = 0$	$\tilde{r}_2 = 1$
$\tilde{r}_1 = 0$	1/8	1/8
$\tilde{r}_1 = 1$	1/8	1/8

Then, calculating β from its definition in Section 5.1, we get $\rho^1 = \nabla$; $\rho_1^2 = \Delta_1$, $\rho_2^2 = \Delta_2$ and $\rho_3^2 = \nabla_3$; and $\rho_3 = \Delta = \beta = \sim$. Hence (LINK) and therefore (C) hold. In fact, following the proof of Theorem 5.1, we construct the payment scheme:

$$t_1(\underline{a}) = L \ln P(a_2|a_1) - \ln P(a_1a_2|a_3), \quad t_2(\underline{a}) = L \ln P(a_1|a_2),$$
$$t_3(\underline{a}) = \ln P(a_1a_2|a_3) - L(\ln P(a_1|a_2) + \ln P(a_2|a_1)).$$

This is clearly budget-balancing and it satisfies condition B of d'Aspremont and Gérard-Varet (1982), $Et_i(r_i|r_i) > Et_i(a_i|r_i)$ for all i, r_i, and $a_i \neq r_i$, if L is sufficiently large ($L = 3$ will do for this case). However, if

$$\lambda_1(0,1) = \lambda_1(1,0) = \lambda_2(0,1) = \lambda_2(1,0) = 1, \quad \lambda_3(0,1) = 2, \quad \text{and} \quad \lambda_3(1,0) = 0,$$

then the "if" clause of (STRC) holds with

$$\kappa(0,0,0) = \kappa(1,1,0) = -1/4, \quad \kappa(0,1,0) = \kappa(1,0,0) = -1/4, \quad \kappa(\underline{a}) = 0 \quad \text{otherwise},$$

and (STRC) fails.

REFERENCES

ARROW, K. J. (1979): "The Property Rights Doctrine and Demand Revelation Under Incomplete Information," in *Economics and Human Welfare: Essays in Honor of Tibor Scitovsky*, ed. by M. J. Boskin. New York: Academic Press.

CRÉMER, J., AND R. P. MCLEAN (1985): "Optimal Selling Strategies Under Uncertainty for a Discriminating Monopolist When Demands are Interdependent," *Econometrica*, 53, 345–361.

D'ASPREMONT, C., AND L. GÉRARD-VARET (1979): "Incentives and Incomplete Information," *Journal of Public Economics*, 11, 25–45.

——— (1982): "Bayesian Incentive Compatible Beliefs," *Journal of Mathematical Economics*, 10, 83–103.

GOOD, I. J. (1952): "Rational Decisions," *Journal of the Royal Statistical Society*, Series B, 107–114.

GREEN, J. A., A. HYLLAND, J. W. PRATT, AND R. J. ZECKHAUSER (1984): "Incentives to Tell the Truth," Harvard University mimeo.

HARSANYI, J. C. (1967/8): "Games With Incomplete Information Played by 'Bayesian' Players," *Management Science*, 14, 159–182, 320–334, 486–502.

JOHNSON, S., J. W. PRATT, AND R. J. ZECKHAUSER (1989): "Efficiency Despite Mutually Payoff-Relevant Private Information: The Finite Case," Harvard Business School Working Paper, #89-067.

KY FAN (1958): "On Systems of Inequalities," in *Linear Inequalities and Related Systems*, ed. by H. W. Kuhn and A. W. Tucker. Princeton, NJ: Princeton University Press.

LAWLER, E. L., J. K. LENSTRA, A. H. G. RINNOOY KAN, AND D. B. SHMOYS (1985): *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*. New York: John Wiley & Sons.

MOSTELLER, F., AND D. L. WALLACE (1964): *Inference and Disputed Authorship: The Federalist*. Reading, MA: Addison-Wesley. Second edition (1984), *Applied Bayesian and Classical Inference: The Case of the Federalist Papers*. New York: Springer-Verlag.

MYERSON, R. B. (1982): "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, 10, 67–81.

——— (1985): "Bayesian Equilibrium and Incentive-Compatibility: An Introduction," in *Social Goals and Social Organization: Essays in Memory of Elisha Pazner*, ed. by L. Hurwicz, D. Schmeidler, and H. Sonnenschein. New York: Cambridge University Press.

PRATT, J. W., AND R. J. ZECKHAUSER (1987): "Incentive-Based Decentralization: Expected Externality Payments Induce Efficient Behavior in Groups," in *Arrow and the Ascent of Economic Theory*, ed. by G. Feiwel. London: MacMillan.

VAN NAERSSEN, R. F. (1962): "A Scale for the Measurement of Subjective Probability," *Acta Psychologia*, 20, 159–166.

WILSON, R. (1985): "Efficient Trading," in *Issues in Contemporary Microeconomics and Welfare*, ed. by G. R. Feiwel. Albany: State University of New York Press.

WINKLER, R. L. (1967): "The Quantification of Judgment: Some Methodological Suggestions," *Journal of the American Statistical Association*, 62, 1105–1120.