Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 15-011

"Informational size and two-stage mechanisms"

by

Richard P. McLean and Andrew Postlewaite

http://ssrn.com/abstract=2573892

Informational size and two-stage mechanisms*

Richard P. McLean
Rutgers University

Andrew Postlewaite
University of Pennsylvania

March, 2015

Abstract

We showed in McLean and Postlewaite (2014) that when agents are informationally small, there exist small modifications to VCG mechanisms in interdependent value problems that restore incentive compatibility. This paper presents a two-stage mechanism that similarly restores incentive compatibility. The first stage essentially elicits that part of the agents' private information that induces interdependence and reveals it to all agents, transforming the interdependent value problem into a private value problem. The second stage is a VCG mechanism for the now private value problem. Agents typically need to transmit substantially less information in the two stage mechanism than would be necessary for a single stage mechanism. Lastly, the first stage that elicits the part of the agents' private information that induces interdependence can be used to transform certain other interdependent value problems into private value problems.

Keywords: Auctions, Incentive Compatibility, Mechanism Design, Interdependent Values, Ex Post Incentive Compatibility, Informational Size

JEL Classifications: C70, D44, D60, D82

1. Introduction

The Vickrey-Clarke-Groves mechanism (hereafter VCG) for private values environments is a classic example of a mechanism for which truthful revelation is ex post incentive compatible. It is well-known, however, that truthful revelation is generally no longer incentive compatible when we move from a private values environment to an interdependent values environment. In McLean and Postlewaite (2014) (henceforth MP (2014)) we showed that, when agents are informationally small in the sense of McLean and Postlewaite (2002), there exists a modification of a generalized VCG mechanism using small additional transfers that restores incentive compatibility. This paper presents an alternative, two-stage, mechanism that accomplishes the same goal – restoring incentive compatibility for interdependent value problems. The advantage of the two stage mechanism relative to a single stage mechanism is that, for typical problems, agents need to transmit substantially less information.

We will explain intuitively the nature of the savings in transmitted information. Consider a problem in which there is uncertainty about the state of nature. An agent's private information

*We thank the National Science Foundation for financial support. We thank Rakesh Vohra for helpful conversations and Zehao Hu for excellent proofreading.

consists of a state dependent payoff function and a signal correlated with the state. A single stage mechanism that delivers an efficient outcome for any realization of agents' types must do two things. First, it must elicit the information agents have about the state of nature to determine the posterior probability distribution given that information. Second, it must elicit agents' privately known state dependent payoffs. A two stage mechanism can separate the two tasks. First, elicit the information about the state of nature, but relay to agents the posterior distribution on the state of nature before collecting any additional information. When agents are induced to reveal their information about the state of nature truthfully, relaying the posterior distribution on the state of nature converts the interdependent value problem into a private value problem. When agents know the probability distribution on the set of states of nature, they need only report their expected utility for each possible social outcome rather than their utility for every social outcome in each of the states. Essentially, by moving from a one stage mechanism to a two stage mechanism, we can shift the job of computing expected utilities given the posterior from the mechanism to the agents. Doing this reduces the information that agents must report to the mechanism; we discuss this in the last section.

In this paper, we construct a two stage game in which the second stage is modeled as a standard private values VCG mechanism. The basic mechanics of first eliciting the information correlated with the state of nature in order to convert an interdependent values problem into private values problem can be applied to certain other environments as well. In particular, our first stage can be combined with certain other mechanisms with desirable properties in private value problems to address implementation problems in the presence of interdependent values.

We provide an example of how our mechanism works in the next section, and present the general mechanism after that.

2. Example

The following single object auction example, a modification of the example in McLean and Postlewaite (2004), illustrates our basic idea. An object is to be sold to one of three agents. There are two equally likely states of the world, θ_1 and θ_2, and an agent's value for the object depends on the state of the world. Agent i's state dependent utility function can written as $v_i = (v_i^1, v_i^2) = (v_i(\theta_1), v_i(\theta_2))$ where v_i^j is his utility of the object in state θ_j. An agent's utility function is private information. In addition, each agent i receives a private signal $s_i \in \{a_1, a_2\}$ correlated with the state. These signals are independent conditional on the state and the conditional probabilities are as shown in the following table.

	signal	a_1	a_2
state			
θ_1		ρ	$1-\rho$
θ_2		$1-\rho$	ρ

where $\rho > \frac{1}{2}$. Consequently, an agent's private information, his type, is a pair (s_i, v_i) and we make two assumptions. First, for any type profile $(s_i, v_i)_{i=1}^3$, the conditional distribution on the state space given $(s_i, v_i)_{i=1}^3$ depends only on the signals (s_1, s_2, s_3). Therefore, the agents' utility functions provide no information relevant for predicting the state that is not already contained in the signal profile alone. Second, we assume that for any type (s_i, v_i) of agent i, the conditional distribution on the signals s_{-i} of the other two agents given (s_i, v_i) depends only on i's signal s_i. Note that the conditional distribution on the state space given (s_1, s_2, s_3) and the conditional distribution on the signals s_{-i} given s_i can be computed using the table above.

Suppose the objective is to allocate the object to the agent for whom the expected value, conditional on the agents' true signal profile, is highest. This is a problem with interdependent values. Agent i's conditional expected value for the object depends on the probability distribution on the states, conditional on the signals of all three agents. In McLean and Postlewaite (2004), it is shown how one can design a mechanism to allocate the object to the highest value agent for problems of this sort. Consider a direct mechanism in which each agent reports his type (s_i, v_i). The mechanism uses the reported signals about the state, $s = (s_1, s_2, s_3)$ to compute the posterior distribution $(\pi(\theta_1|s), \pi(\theta_2|s))$ on Θ. This posterior is used along with agent i's announced state dependent utilities to compute i's expected utility: $\hat{v}_i(s) = v_i^1\pi(\theta_1|s) + v_i^2\pi(\theta_2|s)$. The mechanism then awards the object to the agent with the highest expected value $\bar{v}_i(s)$ and that agent pays the second highest expected value.[1]

The mechanism can be thought of as eliciting all the information available about the unknown state θ and then using a Vickrey auction based on the expected values that are computed using this information about θ. If agents always reported their true signals and if the true signal profile is $s = (s_1, s_2, s_3)$, then by the well known property of Vickrey auctions, it is a dominant strategy for each agent i to truthfully report his expected payoff $\bar{v}_i(s)$. However, agents may indeed have an incentive to misreport their signals in order to manipulate the conditional expected valuations that are used to determine the winner and the price. For example, if all agents' have state dependent values that are lower in state θ_1 than in state θ_2 ($x_i < y_i,\ i = 1, 2, 3$), then an agent who has received signal a_2 may have an incentive to report a_1. Such a misreport will increase the probability weight that the posterior assigns to θ_1. Consequently, this will lower all agents' expected values which, in turn, will affect the price paid by the winner of the object.

To induce agents to truthfully announce their signals, McLean and Postlewaite add a reward z to the payment of agent i if his report about the state is in the majority. Since agents receive conditionally independent signals about the state, an agent maximizes the probability that he gets the reward z by announcing truthfully if other agents are doing so. Since the maximal possible gain from misreporting is bounded, then a sufficiently large z will make truthful announcement an equilibrium.

The reward z need not be very large if ρ is close to 1. When ρ is close to 1, it is very likely that the all agents received "correct" signals about the state. Therefore, conditional on his own signal, agent i believes that a lie will, with high probability, have only a small effect on the posterior distribution on Θ. But the expected gain from misreporting will be small if the expected change in the posterior is small. Thus, when agents are receiving very accurate signals about θ, small rewards will support truthful announcement as an equilibrium.

Our aim in this paper is to demonstrate a two stage modification of this kind of mechanism that accomplishes the same goal but requires less information to be transmitted. The two stages correspond to the two parts of the mechanism discussed above. The first stage elicits the information about the state and publicly posts the posterior distribution on Θ given the reports. The agents use that posterior to compute their expected values of the object, and then report those expected values. As before, the object is sold to the agent with the highest reported

[1] If there are more than one agent with the highest expected value the object is awarded to each with equal probability.

expected value at the second highest reported expected value.

The important difference between the one stage mechanism and the two stage mechanism is that, in the one stage mechanism, agent i reports his type – his signal s_i and his vector of state dependent payoffs (x_i, y_i) – while in the two stage mechanism he does *not* report his type. Instead he reports his signal s_i and his *expected* utility given the posted posterior π, a single number rather than a pair of numbers. In more general problems with m states of the world and k outcomes, an agent's type would include a $m \times k$ matrix consisting of m state contingent payoffs for each of the k outcomes.

The difference in the dimensionality of these announcements is tied to where the expectation of utility given the posterior on Θ is computed. In a one stage mechanism it must be done within the mechanism. To compute the expectation, the entire matrix of state-outcome payoffs of each agent must be transmitted to the mechanism. In the two stage mechanism, only the information necessary to determine the posterior on Θ given all available private information is transmitted in the first stage. The resulting posterior is then transmitted to the agents, who compute the expectations and return those expected values to the mechanism. In terms of the amount of information that must be transmitted, it is more efficient for the agents to compute the expectations.

3. Preliminaries

In this section, we review the structure and salient results from MP (2014). If K is a finite set, let $|K|$ denote the cardinality of K and let $\Delta(K)$ denote the set of probability measures on K. Throughout the paper, $||\cdot||_2$ will denote the 2-norm and, for notational simplicity, $||\cdot||$ will denote the 1-norm. The real vector spaces on which these norms are defined will be clear from the context. Let $\Theta = \{\theta_1, .., \theta_m\}$ represent the finite set of states of nature and let T_i denote the finite set of types of player i. Let $\Delta^*(\Theta \times T)$ denote the set of $P \in \Delta(\Theta \times T)$ whose marginals on Θ and T satisfy the following full support assumptions: $P(\theta) > 0$ for each $\theta \in \Theta$ and $P(t) > 0$ for each $t \in T$. The conditional distribution induced by P on Θ given $t \in T$ (resp., the conditional distribution induced by (the marginal of) P on T_{-i} given $t_i \in T_i$) is denoted $P_\Theta(\cdot|t)$ (resp., $P(\cdot|t_i)$). Let C denote the finite set of social alternatives. Agent $i's$ payoff is represented by a nonnegative valued function $v_i : C \times \Theta \times T_i \to \mathbb{R}_+$ and we assume that for all i, $v_i(\cdot, \cdot, \cdot) \leq M$ for some $M \geq 0$.

A *social choice problem* is a collection $(v_1, .., v_n, P)$ where $P \in \Delta^*(\Theta \times T)$. An *outcome function* is a mapping $q : T \to C$ that specifies an outcome in C for each profile of announced types. A *mechanism* is a collection $(q, x_1, .., x_n)$ (written simply as $(q, (x_i))$ where $q : T \to C$ is an outcome function and the functions $x_i : T \to \mathbb{R}$ are transfer functions. For any profile of types $t \in T$, let

$$\hat{v}_i(c; t) = \hat{v}_i(c; t_{-i}, t_i) = \sum_{\theta \in \Theta} v_i(c, \theta, t_i) P_\Theta(\theta | t_{-i}, t_i).$$

Although $\hat{v}$ depends on P, we suppress this dependence for notational simplicity as well. Finally, we make the simple but useful observation that the pure private value model is mathematically identical to a model in which $|\Theta| = 1$.

Definition 1: Let $(v_1, .., v_n, P)$ be a social choice problem. A mechanism $(q, (x_i))$ is:

ex post incentive compatible if truthful revelation is an *ex post Nash equilibrium*: for all $i \in N$, all $t_i, t_i' \in T_i$ and all $t_{-i} \in T_{-i}$,

$$\hat{v}_i(q(t_{-i}, t_i); t_{-i}, t_i) + x_i(t_{-i}, t_i) \geq \hat{v}_i(q(t_{-i}, t_i'); t_{-i}, t_i) + x_i(t_{-i}, t_i').$$

interim incentive compatible if truthful revelation is a *Bayes-Nash equilibrium*: for each $i \in N$ and all $t_i, t_i' \in T_i$

$$\sum_{t_{-i} \in T_{-i}} \left[\hat{v}_i(q(t_{-i}, t_i); t_{-i}, t_i) + x_i(t_{-i}, t_i)\right] P(t_{-i}|t_i) \geq \sum_{t_{-i} \in T_{-i}} \left[\hat{v}_i(q(t_{-i}, t_i'); t_{-i}, t_i) + x_i(t_{-i}, t_i')\right] P(t_{-i}|t_i)$$

ex post individually rational if

$$\hat{v}_i(q(t); t) + x_i(t) \geq 0 \text{ for all } i \text{ and all } t \in T.$$

feasible if for each $t \in T$,

$$\sum_{j \in N} x_j(t) \leq 0.$$

outcome efficient if for each $t \in T$,

$$q(t) \in \arg\max_{c \in C} \sum_{j \in N} \hat{v}_j(c; t).$$

In our framework, ex post means ex post the realization of the agents' information profile. All activity takes place after players learn their private information but before the realization of θ is known. If, for all i, $\hat{v}_i(c; t)$ does not depend on t_{-i}, then the notions of ex post Nash equilibrium and dominant strategy equilibrium coincide.

4. The Model

4.1. Information Decompositions

In this section, we show how the information structure for general incomplete information problems, even those without a product structure, can be represented in a way that separates out an agent's information about the state θ. This is important because it is this part of his type that affects other agents' valuations of the social alternatives. The example of Section 2 illustrates how we can elicit truthful reporting of agents' signals about the state when they are correlated.

In that example, an agent has beliefs about other agents signals that depend on his own signal, and it is important that the beliefs are different for different signals the agent may receive. In the example, agent i's type consists of a signal a_i and a state dependent utility function that is independent of his signal. Consequently agent i has multiple types consisting of the same signal but different utility functions, and all of these types will necessarily have the same beliefs about other agents' signals.

It isn't necessary to elicit that part of an agent's type that doesn't affect other agents' valuations (e.g., his utility function in the example) to cope with the interdependence, only the part related to the state θ. To formalize this idea, we recall the notion of information decomposition from McLean and Postlewaite (2004).[2]

Definition 2: Suppose that $P \in \Delta^*(\Theta \times T)$. An information decomposition for P is a collection $\mathbb{D} = ((A_i, f_i)_{i \in N}, Q)$ satisfying the following conditions:

(i) For each i, A_i is a finite set and $f_i : T_i \to A_i$ is a function and $Q \in \Delta(\Theta \times A_1 \times \cdots \times A_n)$.[3] For each $t \in T$, define $f(t) := (f_1(t_1), .., f_n(t_n))$ and $f_{-i}(t_{-i}) := (f_1(t_1), .., f_{i-1}(t_{i-1}), f_{i+1}(t_{i+1}), .., f_n(t_n))$.

(ii) For each $t \in T$,

$$P_\Theta(\theta|t) = Q_\Theta(\theta|f(t))$$

(iii) For each i, $t_i \in T_i$ and $a \in A$,

$$\sum_{\substack{t_{-i} \in T_{-i} \\ : f_{-i}(t_{-i}) = a_{-i}}} P(t_{-i}|t_i) = Q(a_{-i}|f_i(t_i))$$

If $t_i \in T_i$, we will interpret $f_i(t_i) \in A_i$ as the "informationally relevant component" of t_i and we will refer to A_i as the set of agent i's "signals." Condition (ii) states that a type profile $t \in T$, contains no information beyond that contained in the signal profile $f(t)$ that is useful in predicting the state of nature. Condition (iii) states that a specific type $t_i \in T_i$ contains no information beyond that contained in the signal $f_i(t_i)$ that is useful in predicting the signals of other agents.

Every $P \in \Delta^*(\Theta \times T)$ has at least one information decomposition in which $A_i = T_i, f_i = id$, and $Q = P$ which we will refer to as the *trivial decomposition*. However, the trivial decomposition may not be the only one (or the most useful one as we will show below). For example, suppose that each agent's type set has a product structure $T_i = X_i \times Y_i$ and that $P \in \Delta^*(\Theta \times T)$ satisfies

$$P(\theta, x_1, y_1, .., x_n, y_n) = P_1(\theta, x_1, .., x_n)P_2(y_1, .., y_n)$$

for each $(x_1, y_1, .., x_n, y_n)$ where $P_1 \in \Delta(\Theta \times X)$ and $P \in \Delta(Y)$. Then defining the projection map $p_{X_i}(x_i, y_i) = x_i$, it follows that $\mathbb{D} = ((X_i, p_{X_i})_{i \in N}, P_1)$ is an information decomposition for P.

Remark: If $(v_1, .., v_n, P)$ is a social choice problem, then it follows from the definition that any two information decompositions $\mathbb{D} = ((A_i, f_i)_{i \in N}, Q)$ and $\mathbb{D}' = ((A'_i, f'_i)_{i \in N}, Q')$ for P give rise to the same $\hat{v}_i$, i.e., for all $t \in T$, we have

$$\sum_{\theta \in \Theta} v_i(c, \theta, t_i) Q_\Theta(\theta|f(t)) = \sum_{\theta \in \Theta} v_i(c, \theta, t_i) P_\Theta(\theta|t_{-i}, t_i) = \sum_{\theta \in \Theta} v_i(c, \theta, t_i) Q_\Theta(\theta|f'(t)).$$

[2] This definition is equivalent to the partition formulation in McLean and Postlewaite (2004).

[3] The conditional distribution induced by Q on Θ given $a \in A$ (resp., the conditional distribution induced by (the marginal of) Q on A_{-i} given $a_i \in A_i$) is denoted $Q_\Theta(\cdot|a)$ (resp., $Q(\cdot|a_i)$).

4.2. Informational Size

In this paper, a fundamental role is played by the notion of informational size. Suppose that $\mathbb{D} = ((A_i, f_i)_{i\in N}, Q)$ is an information decomposition for $P \in \Delta^*(\Theta\times T)$. In a direct mechanism, agent i reports an element of T_i to the mechanism. Consider an alternative scenario in which each agent i reports a signal $a_i \in A_i$ to the mechanism. If i reports a_i and the remaining agents report a_{-i}, it follows that the profile $a = (a_{-i}, a_i) \in A$ will induce a conditional distribution on Θ (computed from Q) and, if agent i's report changes from a_i to a'_i, then this conditional distribution will (in general) change. We consider agent i to be *informationally small* if, for each a_i, agent i ascribes "small" probability to the event that he can effect a "large" change in the induced conditional distribution on Θ by changing his announced type from a_i to some other a'_i. This is formalized in the following definition.

Definition 3: Suppose that $\mathbb{D} = ((A_i, f_i)_{i\in N}, Q)$ is an information decomposition for $P \in \Delta^*(\Theta \times T)$. Let

$$I^i_\varepsilon(a'_i, a_i) = \{a_{-i} \in A_{-i} |\ ||Q_\Theta(\cdot|a_{-i}, a_i) - Q_\Theta(\cdot|a_{-i}, a'_i)|| > \varepsilon\}$$

and

$$\nu^Q_i(a'_i, a_i) = \min\{\varepsilon \geq 0 | \sum_{a_{-i}\in I^i_\varepsilon(a'_i,a_i)} Q_\Theta(a_{-i}|a_i) \leq \varepsilon\}$$

The *informational size* of agent i is defined as

$$\nu^Q_i = \max_{a_i\in A_i} \max_{a'_i\in A_i} \nu^Q_i(a'_i, a_i).$$

Loosely speaking, we will say that agent i is *informationally small* with respect to Q if $\nu^Q_i(a'_i, a_i)$ is small for all $a'_i, a_i \in A_i$. If agent i receives signal a_i but reports $a'_i \neq a_i$, the effect of this misreport is a change in the conditional distribution on Θ from $Q_\Theta(\cdot|a_{-i}, a_i)$ to $Q_\Theta(\cdot|a_{-i}, a'_i)$. If $a_{-i} \in I_\varepsilon(a'_i, a_i)$, then this change is "large" in the sense that $||Q_\Theta(\cdot|a_{-i}, a_i) - Q_\Theta(\cdot|a_{-i}, a'_i)|| > \varepsilon$. Therefore, $\sum_{a_{-i}\in I^i_\varepsilon(a'_i,a_i)} Q(a_{-i}|a_i)$ is the probability that i can have a "large" influence on the conditional distribution on Θ by reporting a'_i instead of a_i conditional on his observed signal a_i. An agent is informationally small if for each of his possible types a_i, he assigns small probability to the event that he can have a "large" influence on the distribution $Q_\Theta(\cdot|a_{-i}, a_i)$, given his observed type.[4]

4.3. Variability of Beliefs

The example of Section 2 illustrates how one might induce truthful announcement of agents' signals about the state. An agent who receives the signal a_1 believes that the state is more likely to be θ_1 than θ_2. Given that agents' signals are conditionally independent, he believes that each of the other agents is more likely to have received signal a_1 than a_2. Hence, if those agents are announcing truthfully, he maximizes his chance of receiving the reward z by announcing truthfully as well. More generally, the key to constructing rewards for an agent who might

[4]Informational size is closely related to the notion of nonexclusive information as well the Ky Fan distance between random variables. See MP (2014) for an elaboration.

receive signal a or a' is that the agent's beliefs about other agents' signals when he receives signal a differ from his beliefs when he receives signal a'. Moreover, the magnitude of the difference matters in inducing truthful announcement. We turn next to defining a measure of the variation of an agent's beliefs.

Suppose that $Q \in \Delta(\Theta \times A)$.We can find rewards that will induce agent i to reveal his information if $Q(\cdot|a_i)$ and $Q(\cdot|a_i')$, the distributions on A_{-i} given signals a_i and a_i' for agent i, are different when $a_i \neq a_i'$. The size of the rewards that will induce truthful reporting will depend on the magnitude of the difference between $Q(\cdot|a_i)$ and $Q(\cdot|a_i')$ for different types a_i and a_i' of agent i.

To define formally the measure of variability, we treat each conditional $Q(\cdot|a_i) \in \Delta(A_{-i})$ as a point in a Euclidean space of dimension equal to the cardinality of A_{-i}. Our measure of variability is defined as

$$\Lambda_i^Q = \min_{a_i \in A_i} \min_{a_i' \in A_i \setminus a_i} \left\| \frac{Q(\cdot|a_i)}{||Q(\cdot|a_i)||_2} - \frac{Q(\cdot|a_i')}{||Q(\cdot|a_i')||_2} \right\|^2 .$$

If $\Lambda_i^Q > 0$, then the agents' signals cannot be stochastically independent with respect to Q. We will exploit this correlation in constructing Bayesian incentive compatible mechanisms. For a discussion of the relationship between this notion of correlation and that found in the full extraction literature, see MP (2014).

It is important to point out that Λ_i^Q and $\Lambda_i^{Q'}$ are generally different for two decompositions $D = ((A_i, f_i)_{i \in N}, Q)$ and $D' = ((A_i', f_i')_{i \in N}, Q')$ for P. When an agent's type set has a product structure $T_i = X_i \times Y_i$ as in the example of Section 4.1 and $D = ((T_i, id)_{i \in N}, P)$ is the trivial decomposition, then $\Lambda_i^Q = \Lambda_i^P = 0$ for all i. However, for the decomposition $D = ((X_i, p_{X_i})_{i \in N}, P_1)$ of that example, it may in fact be the case that $\Lambda_i^{P_1} > 0$. The utility of decompositions will become apparent when we state Theorem B below.

5. The One Stage Implementation Game

5.1. The Generalized VCG Mechanism

We now adapt some of our previous results on implementation with interdependent values to the model of this paper. In the special case of pure private values, i.e., when $|\Theta| = 1$, it is well known that the classical VCG transfers will implement an outcome efficient social choice function: in the induced direct revelation game, it is a dominant strategy to honestly report one's type. In the general case of interdependent values, the situation is more delicate.

Let $q : T \to C$ be an outcome efficient social choice function for the problem $(v_1, .., v_n, P)$. For each t, define transfers as follows:

$$\alpha_i^q(t) = \sum_{j \in N \setminus i} \hat{v}_j(q(t); t) - \max_{c \in C} \left[\sum_{j \in N \setminus i} \hat{v}_j(c; t) \right]$$

Note that $\alpha_i^q(t) \leq 0$ for each i and t. The resulting mechanism $(q, (\alpha_i^q))$ is the *generalized VCG mechanism with interdependent valuations* (GVCG for short) studied in MP(2014). It is straightforward to show that the GVCG mechanism is ex post individually rational and feasible.

In the pure private value case where $|\Theta| = 1$, it follows that for an outcome efficient social choice function $q : T \to C$, we have

$$q(t) \in \arg\max_{c \in C} \sum_{i \in N} v_i(c, t_i)$$

and the GVCG transfers reduce to

$$\alpha_i^q(t) = \sum_{j \in N \setminus i} v_j(c, t_j) - \max_{c \in C} \left[\sum_{j \in N \setminus i} v_j(c, t_j) \right].$$

which are precisely the classical VCG transfers. Honest reporting of one's type is, of course, a dominant strategy in this private values setup. Unfortunately, the GVCG mechanism does not inherit the very attractive dominant strategy property of the pure private values special case. It is tempting to conjecture that the GVCG mechanism satisfies ex post incentive compatibility or perhaps the weaker notion of Bayesian incentive compatibility but even the latter need not hold. There are, however, certain positive results. In MP (2014), it is shown that the GVCG mechanism is ex post incentive compatible when the problem satisfies a condition called non-exclusive information. This property is satisfied by all pure private values models, in which case, ex post incentive compatibility reduces to dominant strategy incentive compatibility.

This observation follows as an immediate consequence of the definition of GVCG mechanisms but it can also be deduced from a result that will play a crucial role in our analysis. This is the "gain-bounded" property of the GVCG mechanism proved in McLean and Postlewaite (2014).

Lemma A: Suppose that $q : T \to C$ is outcome efficient for the problem $(v_1, .., v_n, P)$ and suppose that $\mathbb{D} = ((A_i, f_i)_{i \in N}, Q)$ is an information decomposition for P. Then for all $t_{-i} \in T_{-i}$ and $t_i, t_i' \in T_i$,

$$\begin{aligned}
&[\hat{v}_i(q(t_{-i}, t_i'); t_{-i}, t_i) + \alpha_i^q(t_{-i}, t_i')] - [\ \hat{v}_i(q(t_{-i}, t_i); t_{-i}, t_i) + \alpha_i^q(t_{-i}, t_i)] \\
&\quad \leq 2M(n-1)||Q_\Theta(\cdot|f_{-i}(t_{-i}), f_i(t_i) - Q_\Theta(\cdot|f_{-i}(t_{-i}), f_i(t_i')||.
\end{aligned}$$

In the case of the GVCG mechanism, Lemma A provides an upper bound on the "ex post gain" to agent i when i's true type is t_i but i announces t_i' and others announce truthfully. An important implication of Lemma A is that an agent's gain by misreporting his type is essentially bounded by the degree to which his type affects the posterior probability distribution on Θ; we return to this below. In the pure private values model where $|\Theta| = 1$, we conclude that $||Q_\Theta(\cdot|f_{-i}(t_{-i}), f_i(t_i) - Q_\Theta(\cdot|f_{-i}(t_{-i}), f_i(t_i')|| = 0$ and we recover the classic dominant strategy result. There is a second ramification of Lemma A: when agents are informationally small, honest reporting is an approximate ex post Nash equilibrium in the GVCG mechanism. See MP (2014) for a discussion of this result.

Lemma A has a third important consequence: if agent i is informationally small, then truth is an approximate Bayes-Nash equilibrium in the GVCG mechanism so the mechanism is approximately interim incentive compatible. More precisely, we can deduce from Lemma A that the interim expected gain from misreporting one's type is essentially bounded from above by one's informational size. If we want the mechanism to be exactly interim incentive compatible, then we must alter the mechanism (specifically, construct an augmented GVCG mechanism) in order to provide the correct incentives for truthful behavior. We turn to this next.

5.2. A Direct Mechanism for the One Stage Implementation Game

If we use the GVCG mechanism to define a direct revelation game, then we show in MP (2014) that honest reporting is an approximate ex post Nash equilibrium and an approximate Bayes-Nash equilibrium when agents are informationally small. In MP (2014), we also consider a modification of the GVCG mechanism that is both approximately ex post incentive compatible and *exactly*, rather than *approximately*, interim incentive compatible when agents are informationally small. To state the main result of MP (2014), we need the notion of an augmented mechanism.

Definition 4: Let $(z_i)_{i \in N}$ be an n-tuple of functions $z_i : T \to \mathbb{R}_+$ each of which assigns to each $t \in T$ a nonnegative number, interpreted as a "reward" to agent i. If $(q, x_1, .., x_n)$ is a mechanism, then the associated *augmented* mechanism is defined as $(q, x_1 + z_1, .., x_n + z_n)$ and will be written simply as $(q, (x_i + z_i))$.

Using precisely the same techniques found in MP (2014), we can prove the following result for direct mechanisms.

Theorem A : Let $(v_1, .., v_n)$ be a collection of payoff functions.

(i) Suppose that $q : T \to C$ is outcome efficient for the problem $(v_1, .., v_n, P)$. Suppose that $\mathbb{D} = ((A_i, f_i)_{i \in N}, Q)$ is an information decomposition for $(v_1, .., v_n, P)$ satisfying $\Lambda_i^Q > 0$ for each i. Then there exists an augmented GVCG mechanism $(q, (\alpha_i^q + z_i))$ for the social choice problem $(v_1, .., v_n, P)$ satisfying ex post IR and interim IC.

(ii) For every $\varepsilon > 0$, there exists a $\delta > 0$ such that the following holds: whenever $q : T \to C$ is outcome efficient for the problem $(v_1, .., v_n, P)$ and whenever $\mathbb{D} = ((A_i, f_i)_{i \in N}, Q)$ is an information decomposition for $(v_1, .., v_n, P)$ satisfying

$$\max_i \nu_i^Q \leq \delta \min_i \Lambda_i^Q,$$

there exists an augmented GVCG mechanism $(q, (\alpha_i^q + z_i))$ with $0 \leq z_i(t) \leq \varepsilon$ for every i and t satisfying ex post IR and interim IC.[5]

Part (i) of Theorem 2 states that, as long as $Q(\cdot|a_i) \neq Q(\cdot|a_i')$ whenever $a_i \neq a_i'$, then irrespective of the agents' informational sizes, the augmenting transfers can be chosen so that the augmented mechanism satisfies Bayesian incentive compatibility. However, the required augmenting transfers will be large if the agents have large informational size. Part (ii) states that the augmenting transfers will be small if the agents have informational size that is small enough relative to our measure of variation in the agents' beliefs.

To understand this condition, we first note that, if agent i is informationally small, then truth is an approximate Bayes-Nash equilibrium in the GVCG mechanism so the mechanism is approximately interim incentive compatible. More precisely, we can deduce from Lemma A that the interim expected gain from misreporting one's type is essentially bounded from above by

[5] In MP (2014), it is also shown that the augmented mechanism is approximately ex post incentive compatible in the sense defined that that paper.

one's informational size. If we want the mechanism to be exactly interim incentive compatible, then we must alter the mechanism (specifically, construct an augmented GVCG mechanism) in order to provide the correct incentives for truthful behavior. In MP (2014), the augmenting z_i rewards are defined using a spherical scoring rule and the difference in an agent's expected reward between honest reporting and lying is (essentially) bounded from below by Λ_i^Q. Thus, the number Λ_i^Q quantifies the expected gain from honest reporting of one's type and the condition of part (ii) of Theorem A simply requires that the gain to lying quantified in terms of informational size ν_i^Q be sufficiently outweighed by the gain to honest reporting as quantified by Λ_i^Q.[6]

The two conclusions of Theorem A illustrate the value of information decompositions for implementation with interdependent valuations. The number Λ_i^Q depends only on the informationally relevant component of an agent's type. As we have already indicated, it is possible that $\Lambda_i^P = 0$ for the trivial decomposition $D = ((T_i, id)_{i\in N}, P)$ while $\Lambda_i^P > 0$ for some other decomposition $D = ((A_i, f_i)_{i\in N}, Q)$.

As we have mentioned previously, the agents' beliefs cannot be independent if $\Lambda_i^Q > 0$ for each i. Correlated information also plays a significant role in the full surplus extraction problem in the mechanism design literature (see Cremer and McLean (1985, 1989).) Those papers (and subsequent work by McAfee and Reny (1992)) demonstrated how one can use correlation to fully extract the surplus in certain mechanism design problems. The problems, however, are quite different. Surplus extraction is a mechanism design problem while our problem is an implementation problem. We do not look for transfers and an allocation scheme that solves a mechanism design problem of the type presented in, for example, Myerson (1981) or Cremer and McLean (1985, 1988). Instead, we study the problem of implementation of a given efficient social choice function and it is important to explicate the differences. See MP (2014) for a discussion of this issue.

5.3. An Indirect Mechanism for the One Stage Implementation Problem

The direct revelation game is the most common formulation of the implementation problem and is, in a well known sense, without loss of generality as a result of the revelation principle. However, there are interesting "indirect" mechanisms in which the message space M_i of agent i is different from T_i. Suppose that $(v_1, .., v_n, P)$ is a problem and suppose that $\mathbb{D} = ((A_i, f_i)_{i\in N}, Q)$ is an information decomposition for P. Let G_i denote the collection of all functions $g_i : C \times \Theta \to \mathbb{R}_+$. Consequently we identify G_i with $\mathbb{R}_+^{C\times\Theta}$. Suppose that i's message space is $M_i = A_i \times G_i$.

Let $Z = (\zeta_i)_{i\in N}$ be an n-tuple of functions $\zeta_i : A \to \mathbb{R}_+$ each of which assigns to each $a \in A$ a nonnegative number $\zeta_i(a)$ interpreted as a "reward" to agent i. If the agents report the profile $(a, g) \in A \times G$,then the mediator chooses a social outcome

$$\psi(a, g) \in \arg\max_{c\in C} \sum_{i\in N} \sum_{\theta} g_i(c, \theta) Q_\Theta(\theta|a)$$

and augmented transfers

$$\eta_i(a, g) = \sum_{j\in N\setminus i} \sum_{\theta} g_j(\psi(a, g), \theta) Q_\Theta(\theta|a) - \max_{c\in C} \left[\sum_{j\in N\setminus i} \sum_{\theta} g_j(c, \theta) Q_\Theta(\theta|a) \right] + \zeta_i(a).$$

[6] The proof of the main result in MP(2014) or Part 3 of the proof of Theorem B in the Appendix demonstrates the spherical scoring rule construction as well as the precise roles of ν_i^Q and Λ_i^Q in this tradeoff.

This indirect mechanism induces a game of incomplete information. A strategy for agent i in this game is a pair (α_i, γ_i) where $\alpha_i : T_i \to A_i$ specifies a type dependent report $\alpha_i(t_i) \in A_i$ and $\gamma_i : T_i \to G_i$ specifies a type dependent valuation function $\gamma_i(t_i) \in G_i$.

If $a_i = f_i(t_i)$ and $g_i = v_i(\cdot, \cdot, t_i)$ for each i, then

$$\psi(a, g) \in \arg\max_{c \in C} \sum_{i \in N} \sum_{\theta} v_i(c, \theta, t_i) Q_\Theta(\theta | f(t))$$

and

$$\eta_i(a, g) = \sum_{j \in N \setminus i} v_j(c, \theta, t_j) Q_\Theta(\theta | f(t)) - \max_{c \in C} \left[\sum_{j \in N \setminus i} v_j(c, \theta, t_j) Q_\Theta(\theta | f(t)) \right] + \zeta_i(f(t))$$

A strategy (α_i, γ_i) for player i is *truthful* for i if $\alpha_i(t_i) = f_i(t_i)$ and $\gamma_i(t_i) = v_i(\cdot, \cdot, t_i)$ for all $t_i \in T_i$. A strategy profile $(\alpha_i, \gamma_i)_{i \in N}$ is *truthful* if (α_i, γ_i) is truthful for each player i. If $(\alpha_i, \gamma_i)_{i \in N}$ is a truthful Bayes-Nash equilibrium in this one stage game, then the revelation principle ensures that the outcome of the game coincides with an outcome of the revelation game of Section 5.2. Under the hypotheses of (i) or (ii) of Theorem A, one can construct a system of augmenting transfers ζ_i ensuring that a truthful Bayes-Nash equilibrium exists in this game of incomplete information.

6. The Two Stage Implementation Game

6.1. Preliminaries

Suppose that $(v_1, .., v_n, P)$ is a social choice problem and suppose that $\mathbb{D} = ((A_i, f_i)_{i \in N}, Q)$ is an information decomposition for P.

Throughout this section, we will use the following notational convention:

$$\rho(a_{-i}, a_i) = Q_\Theta(\cdot | a_{-i}, a_i) \text{ and } \rho_\theta(a_{-i}, a_i) = Q_\Theta(\theta | a_{-i}, a_i).$$

Let H_i denote the collection of all functions $u_i : C \to \mathbb{R}_+$. Consequently we identify H_i with $\mathbb{R}_+^C$. For each profile $u = (u_1, .., u_n) \in H := H_1 \times \cdots \times H_n$, let

$$\varphi(u) \in \arg\max_{c \in C} \sum_{i \in N} u_i(c)$$

and define

$$\hat{y}_i(u) = \sum_{j \in N \setminus i} u_j(\varphi(u)) - \max_{c \in C} \left[\sum_{j \in N \setminus i} u_j(c) \right] .$$

Therefore, $(\varphi, \hat{y}_1, .., \hat{y}_n)$ defines the classic private values VCG mechanism and it follows that

$$u_i \in \arg\max_{u_i' \in H_i} u_i(\varphi(u_{-i}, u_i')) + \hat{y}_i(u_{-i}, u_i')$$

for all $u_i \in H_i$ and all $u_{-i} \in H_{-i}$.

We wish to formulate our implementation problem with interdependent valuations as a two stage problem in which honest reporting of the agents' signals in stage one resolves the "interdependency" problem so that the stage two problem is a simple implementation problem with private values to which the classic VCG mechanism can be immediately applied. We now define an extensive form game that formalizes the two stage process that lies behind this idea. As in the indirect mechanism for the one stage problem, let $Z = (\zeta_i)_{i\in N}$ be an n-tuple of functions $\zeta_i : A \to \mathbb{R}_+$ each of which assigns to each $a \in A$ a nonnegative number $\zeta_i(a)$ again interpreted as a "reward" to agent i. These rewards are designed to induce agents to honestly report their signals in stage 1.

Given an information decomposition $\mathbb{D}$ and a reward system Z, we define an extensive form game $\Gamma(\mathbb{D}, Z)$ that unfolds in the following way.

Stage 1: Each agent i learns his type $t_i \in T_i$ and makes a (not necessarily honest) report $r_i \in A_i$ of his signal to the mechanism designer. If $(r_1, .., r_n)$ is the profile of stage 1 reports, then agent i receives the nonnegative payment $\zeta_i(r_1, .., r_n)$ and the game moves to stage 2.

Stage 2: If $(r_1, .., r_n) = r \in A$ is the reported type profile in stage 1, the mechanism designer publicly posts the conditional distribution $\rho(r) = Q_\Theta(\cdot|r)$. Agents observe this posted distribution (but not the profile r) and make a second (not necessarily honest) report from H_i to the mechanism designer. If $(u_1, .., u_n) = u \in H$ is the second stage profile of reports, then the mechanism designer chooses the social alternative $\varphi(u) \in C$, each agent i receives the transfer $\hat{y}_i(u)$, and the game ends.

We wish to design the rewards ζ_i so as to accomplish two goals. In stage 1, we want to induce agents to report honestly so that the reported stage 1 profile is exactly $f(t) = (f_1(t_1), .., f_n(t_n))$ when the true type profile is t. In stage 2, upon observing the posted posterior distribution $Q_\Theta(\cdot|f(t))$, we want each agent i to report the payoff function

$$u_i^*(\cdot) = \sum_{\theta\in\Theta} v_i(\cdot, \theta, t_i) Q_\Theta(\theta|f(t)).$$

If these twin goals are accomplished in a Bayes-Nash equilibrium, then the social outcome is

$$\varphi(u^*) \in \arg\max_c \sum_{i\in N} \sum_{\theta\in\Theta} v_i(c, \theta, t_i) Q_\Theta(\theta|f(t)),$$

the transfers are

$$\hat{y}_i(u^*) = \sum_{j\in N\backslash i} v_j(\varphi(u^*), \theta, t_j) Q_\Theta(\theta|f(t)) - \max_{c\in C} \left[\sum_{j\in N\backslash i} v_j(c, \theta, t_j) Q_\Theta(\theta|f(t)) \right].$$

and the ex post payoff to agent i of type t_i is

$$\sum_{i\in N} \sum_{\theta\in\Theta} v_i(\varphi(u^*), \theta, t_j) Q_\Theta(\theta|f(t)) + \hat{y}_i(u^*) + \zeta_i(f(t)).$$

Note that these transfers and payoffs are precisely the GVCG transfers and payoffs defined in Section 5 for the one stage implementation problem.

6.2. Strategies and Equilibria in the Two Stage Game

Define

$$\Pi := \{Q_\Theta(\cdot|a) : a \in A\}.$$

Given the specification of the extensive form, it follows that the second stage information sets of agent i are indexed by the elements of $A_i \times \Pi \times T_i$. A strategy for agent i in this game is a pair (α_i, β_i) where $\alpha_i : T_i \to A_i$ specifies a type dependent report $\alpha_i(t_i) \in A_i$ in stage 1 and $\beta_i : A_i \times \Pi \times T_i \to H_i$ specifies a second stage report $\beta_i(r_i, \pi, t_i) \in H_i$ as a function of i's first stage report $r_i \in A_i$, the posted distribution π, and $i's$ type $t_i \in T_i$.

We are interested in a Perfect Bayesian Equilibrium (PBE) assessment for the two stage implementation game $\Gamma(\mathbb{D}, Z)$ consisting of a strategy profile $(\alpha_i, \beta_i)_{i \in N}$ and a system of second stage beliefs in which players truthfully report their private information at each stage.

Definition 5: A strategy (α_i, β_i) for player i is *truthful* for i if $\alpha_i(t_i) = f_i(t_i)$ for all $t_i \in T_i$ and

$$\beta_i(f_i(t_i), \pi, t_i) = \sum_{\theta \in \Theta} v_i(\cdot, \theta, t_i)\pi(\theta)$$

for all $\pi \in \Pi$ and $t_i \in T_i$. A strategy profile $(\alpha_i, \beta_i)_{i \in N}$ is *truthful* if (α_i, β_i) is truthful for each player i.

Formally, a **system of beliefs** for player i is a collection of probability measures on $\Theta \times T_{-i}$ indexed by $A_i \times \Pi \times T_i$, i.e., a collection of the form

$$\{\mu_i(\cdot|r_i, \pi, t_i) \in \Delta(\Theta \times T_{-i}) : (r_i, \pi, t_i) \in A_i \times \Pi \times T_i\}.$$

with the following interpretation: when player i of type t_i reports r_i in Stage 1 and observes the posted distribution π, then player i assigns probability mass $\mu_i(\theta, t_{-i}|r_i, \pi, t_i)$ to the event that other players have true types t_{-i} and that the state of nature is θ. As usual, an *assessment* is a pair $\{(\alpha_i, \beta_i)_{i \in N}, (\mu_i)_{i \in N}\}$ consisting of a strategy profile and a system of beliefs for each player.

Definition 6: An assessment $\{(\alpha_i, \beta_i)_{i \in N}, (\mu_i)_{i \in N}\}$ is an *incentive compatible Perfect Bayesian equilibrium (ICPBE) assessment* if $(\alpha, \beta, \mu) = \{(\alpha_i, \beta_i)_{i \in N}, (\mu_i)_{i \in N}\}$ is a Perfect Bayesian Equilibrium assessment and the profile $(\alpha_i, \beta_i)_{i \in N}$ is truthful.

6.3. The Main Result

Theorem B: Let $(v_1, .., v_n)$ be a collection of payoff functions.

(a) Suppose that $q : T \to C$ is outcome efficient for the problem $(v_1, .., v_n, P)$. Suppose that $\mathbb{D} = (A_i, f_i)_{i \in N}, Q)$ is an information decomposition for P satisfying $\Lambda_i^Q > 0$ for each i. Then there exists a reward system $Z = (\zeta_i)_{i \in N}$ such that the two stage game $\Gamma(\mathbb{D}, Z)$ has an ICPBE (α^*, β^*, μ).

(b) For every $\varepsilon > 0$, there exists a $\delta > 0$ such that the following holds: whenever $q : T \to C$ is outcome efficient for the problem $(v_1, .., v_n, P)$ and $\mathbb{D} = (A_i, f_i)_{i \in N}, Q)$ is an information decomposition for P satisfying

$$\max_i \nu_i^Q \leq \delta \min_i \Lambda_i^Q,$$

there exists a reward system $Z = (\zeta_i)_{i \in N}$ such that the two stage game $\Gamma(\mathbb{D}, Z)$ has an ICPBE (α^*, β^*, μ). Furthermore, $0 \leq \zeta_i(a) \leq \varepsilon$ for every i and a.

To prove Theorem B, we proceed in several steps which we outline here. Suppose that $\mathbb{D} = (A_i, f_i)_{i \in N}, Q)$ is an information decomposition for P.

Step 1: Suppose that (α, β) is a strategy profile with $\alpha_i(t_i) = f_i(t_i)$ for all i and t_i. Suppose that player i is of true type t_i, the other players have true type profile t_{-i}, player i reports r_i in stage 1. Given the definition of $(\alpha_{-i}, \beta_{-i})$, it follows that $\alpha_j(t_j) = f_j(t_j)$ for each $j \neq i$. Therefore, player i of type t_i who has submitted report r_i in stage 1 and who observes $\pi \in \Pi$ at stage 2 will assign positive probability

$$\sum_{\hat{t}_{-i} : \rho(f_{-i}(\hat{t}_{-i}), r_i) = \pi} P(\hat{t}_{-i} | t_i) > 0$$

to the event

$$\{\hat{t}_{-i} \in T_{-i} : \rho(f_{-i}(\hat{t}_{-i}), r_i) = \pi\}.$$

Therefore, i's updated beliefs regarding (θ, t_{-i}) consistent with (α, β) are given by

$$\begin{aligned} \mu_i(\theta, t_{-i} | r_i, \pi, t_i) &= \frac{\rho_\theta(f_{-i}(t_{-i}), f_i(t_i)) P(t_{-i} | t_i)}{\sum_{\hat{t}_{-i} : \rho(f_{-i}(\hat{t}_{-i}), r_i) = \pi} P(\hat{t}_{-i} | t_i)} \text{ if } \rho(f_{-i}(t_{-i}), r_i) = \pi \\ &= 0 \text{ otherwise.} \end{aligned}$$

Step 2: Let

$$\overline{w}_i(\cdot, \pi, t_i) := \sum_{\theta \in \Theta} v_i(\cdot, \theta, t_i) \pi(\theta)$$

and for each π and t_{-i} let $w^*_{-i}(\pi, t_{-i}) \in H_{-i}$ be defined as

$$w^*_{-i}(\pi, t_{-i}) := (\overline{w}_j(\cdot, \pi, t_j))_{j \in N \setminus i}.$$

Next, we define the following particular second stage component β_i^* of agent i's strategy as follows: for each $(r_i, \pi, t_i) \in A_i \times \Pi \times T_i$, let

$$\beta_i^*(r_i, \pi, t_i) \in \arg \max_{u_i \in H_i} \sum_{t_{-i} \in T_{-i}} \sum_{\theta \in \Theta} \left[v_i(\varphi(u_i, w^*_{-i}(\pi, t_{-i})), \theta, t_i) + \hat{y}_i(u_i, w^*_{-i}(\pi, t_{-i})) \right] \mu_i(\theta, t_{-i} | r_i, \pi, t_i).$$

where $\mu_i(\cdot | r_i, \pi, t_i)$ is defined in Step 1[7]. We then show that

$$\beta_i^*(f_i(t_i), \pi, t_i) = \sum_{\theta \in \Theta} v_i(\cdot, \theta, t_i) \pi(\theta) = \overline{w}_i(\cdot, \pi, t_i).$$

[7] Note that $\beta_i^*(r_i, \pi, t_i)$ exists since each v_i takes on only finitey many values for each i and the set Π is also finite.

Step 3: If $\alpha_i^*(t_i) = f_i(t_i)$ for all i and t_i and β_i^* is defined as in Step 2, then (α^*, β^*) is a truthful strategy profile. The proof is completed by constructing a system of rewards $Z = (\zeta_i)_{i \in N}$ such that (α^*, β^*, μ) is an ICPBE of the two stage game $\Gamma(\mathbb{D}, Z)$. To accomplish this, we define

$$\zeta_i(a_{-i}, a_i) = \varepsilon \frac{Q(a_{-i}|a_i)}{||Q(\cdot|a_i)||_2}.$$

To prove part (a) of Theorem B, we show that one can find $\varepsilon > 0$ so that (i) deviations at second stage information sets are unprofitable given the beliefs μ defined in step 1 and (ii) coordinated deviations across the two stages are unprofitable. This latter argument depends crucially on the special transfers ζ_i. It is these transfers that induce truthful reporting in stage 1, thus reducing the second stage to a simple implementation problem with private values. To prove part (b), we show that ε can be chosen to be small when each agent's informational size is small enough relative to the variation in his beliefs.

7. Discussion

7.1. Information Transmitted

We have presented three approaches to implementation with interdependent values: a one stage direct mechanism, a one stage indirect mechanism and a two stage game. In each of these approaches, the "quantity" and "complexity" of the information that is transmitted by the agents to the mechanism is different. Because of the revelation principle, it is commonplace to conduct analyses using direct mechanisms. However, a direct mechanism may not be the agents' preferred mechanism when, e.g., privacy is a consideration. In our one stage direct mechanism, an agent reports his type. In the one stage indirect mechanism, he reports signal and a payoff as a function of the social outcome and the state. At the end of the two stage game, he has reported a signal and a payoff as a function only of the social outcome.

For many problems agents would like to send as little information as possible. They may be concerned that the information may be misused by the mechanism designer, or they might be concerned that others may learn something about the agent because of the information transmitted. The mechanism itself may reveal, at least partially, the information an agent sends; alternatively, the agent may be concerned that reported information may be inadvertently leaked.

Our three mechanisms provide a hierarchy of implementation games in terms of privacy protection and we can use the product structure example of Section 4.3 to illustrate this hierarchy. In an equilibrium of the one stage direct mechanism, agent i reveals his true type $t_i = (x_i, y_i)$ and the mechanism learns both coordinates. In an equilibrium of the one stage indirect mechanism, agent i reveals his true signal x_i and his true payoff function $g_i = v_i(\cdot, \cdot, t_i)$ but not his true y_i. Although the mechanism learns i's true value of x_i, the mechanism can only make an inference regarding the second coordinate y_i of agent i's type using the information contained in the pair (x_i, g_i). In an equilibrium of the two stage game, agent i reveals his true x_i and his true expected payoff function

$$u_i = \sum_{\theta \in \Theta} v_i(\cdot, \theta, t_i) P_1(\theta | x_1, .., x_n).$$

Again, the mechanism learns i's true value of x_i and again, the mechanism can only make an inference regarding the second coordinate y_i of agent i's type. However, the information with respect to which this inference is made, i.e., (x_i, u_i) is coarser that the information available in the one stage indirect mechanism.

Furthermore, while mechanism design typically focuses on the incentive problems of eliciting information, there is ultimately a cost of running these mechanisms. One of the costs of operating a mechanism is related to the complexity of the information that agents must transmit to the mechanism.[8] Agents transmit less information in our two stage mechanism than they transmit in the one stage indirect mechanism described above. In the one stage indirect mechanism agents need to transmit their entire state dependent utility function for the mechanism to compute agents' expected utility functions on C**.** In the two stage mechanism, they need only transmit the *expected* utility function, a vector, a substantial savings in message size, particularly when there are many states.

7.2. Informational Requirements

In mechanism design theory, it is commonly assumed that the data defining the problem is common knowledge, that is, that the utility functions and probability distributions are commonly known by the mechanism designer and the agents who are to participate in the mechanism. This is a heroic assumption in nearly all problems and there is always a desire to decrease the reliance of mechanisms on the common knowledge assumption. Dominant strategy mechanisms, when possible, are attractive for this reason. Similarly mechanisms for which truthful announcement is an ex post Nash equilibrium are valued. We note that in our two stage mechanism, the first stage announcements are approximately ex post Nash equilibria in the sense defined in MP (2014) when agents are informationally small.

There are interesting problems that are covered by our model for which the usual common knowledge assumption can be relaxed. Specifically, for the case of noisy signals about two states, it is enough that the agents and the mechanism designer know the maximum and minimum values of the accuracy parameter ρ, the agents know their payoffs and and the mechanism designer knows the maximum and minimum values of the possible payoffs. With this information, the mechanism can compute augmenting transfers that reward an agent whose report is a majority report as in the example of Section 2. Furthermore, these transfers can be chosen to induce honest reporting for any distribution generated by accuracy levels in the given range. Consequently, the agents will report honestly in stage 1 irrespective of their beliefs regarding the true value of the accuracy knowing only the range of possible values. Once the informationally relevant information is revealed in stage 1, it is a dominant strategy to truthfully report expected payoff functions in stage 2.

[8] The seminal paper on message size requirements of mechanisms is Mount and Reiter (1974); see Ledyard and Palfrey (1994, 2002) for more recent investigations of the message size requirements of mechanisms.

8. Appendix

8.1. Preparatory Lemmas

Lemma B: Let M be a nonnegative number and let $\{g_i : C \times \Theta \to \mathbb{R}_+ : i \in N\}$ be a collection of functions satisfying $g_i(\cdot,\cdot) \leq M$ for all i. For each $S \subseteq \{1, .., n\}$ and for each $\pi \in \Delta(\Theta)$, let

$$F_S(\pi) = \max_{c \in C} \sum_{i \in S} \sum_{\theta \in \Theta} g_i(c, \theta)\pi(\theta).$$

Then for each $\pi, \pi' \in \Delta(\Theta)$,

$$|F_S(\pi) - F_S(\pi')| \leq |S|M||\pi - \pi'||.$$

Proof: See MP(2014)

Lemma C: Let M be a nonnegative number and let $\{g_i : C \times \Theta \to \mathbb{R}_+ : i \in N\}$ be a collection of functions satisfying $g_i(\cdot, \cdot) \leq M$ for all i. For each $\pi \in \Delta(\Theta)$, let

$$\xi(\pi) \in \arg\max_{c \in C} \sum_{i \in N} \sum_{\theta \in \Theta} g_i(c.\theta)\pi(\theta),$$

and

$$\eta_i(\pi) = \sum_{j \in N \setminus i} \sum_{\theta \in \Theta} g_j(\xi(\pi), \theta)\pi(\theta) - \max_{c \in C} \left[\sum_{j \in N \setminus i} \sum_{\theta \in \Theta} g_j(c, \theta)\pi(\theta). \right].$$

Then for each $t \in T$ and all $\pi, \pi' \in \Delta(\Theta)$,

$$\left[\sum_{\theta \in \Theta} g_i(\xi(\pi'), \theta)\pi'(\theta) + \eta_i(\pi') \right] - \left[\sum_{\theta \in \Theta} g_i(\xi(\pi), \theta)\pi(\theta) + \eta_i(\pi) \right] \leq (2n-1)M||\pi - \pi'||$$

Proof:

$$\begin{aligned}
&\left[\sum_{\theta \in \Theta} g_j(\xi(\pi'), \theta)\pi'(\theta) + \eta_i(\pi') \right] - \left[\sum_{\theta \in \Theta} g_j(\xi(\pi), \theta)\pi(\theta) + \eta_i(\pi) \right] \\
&= \sum_{k \in N} \sum_{\theta \in \Theta} g_k(\xi(\pi'), \theta)\pi'(\theta) - \sum_{k \in N} \sum_{\theta \in \Theta} g_k(\xi(\pi), \theta)\pi(\theta) \\
&\qquad + \max_{c \in C} \left[\sum_{j \in N \setminus i} \sum_{\theta \in \Theta} g_j(c, \theta)\pi(\theta). \right] - \max_{c \in C} \left[\sum_{j \in N \setminus i} \sum_{\theta \in \Theta} g_j(c, \theta)\pi'(\theta) \right] \\
&= \max_{c \in C} \left[\sum_{k \in N} \sum_{\theta \in \Theta} g_k(c, \theta)\pi'(\theta) \right] - \max_{c \in C} \left[\sum_{k \in N} \sum_{\theta \in \Theta} g_k(c, \theta)\pi(\theta) \right] \\
&\qquad + \max_{c \in C} \left[\sum_{j \in N \setminus i} \sum_{\theta \in \Theta} g_j(c, \theta)\pi(\theta) \right] - \max_{c \in C} \left[\sum_{j \in N \setminus i} \sum_{\theta \in \Theta} g_j(c, \theta)\pi'(\theta) \right] \\
&\leq nM||\pi - \pi'|| + (n-1)M||\pi - \pi'||
\end{aligned}$$

where the last inequality follows from Lemma B.

Lemma D: For each $a_i', a_i \in A_i$,

$$\sum_{a_{-i}} ||\rho(a_{-i}, a_i) - \rho(a_{-i}, r_i))||Q(a_{-i}|a_i) \leq 3\nu_i^Q$$

Proof: Recall that

$$\nu_i^Q = \max_{a_i \in A_i} \max_{a_i' \in A_i} \nu_i^Q(a_i', a_i).$$

where

$$I_\varepsilon^i(a_i', a_i) = \{a_{-i} \in A_{-i} | \ ||Q_\Theta(\cdot|a_{-i}, a_i) - Q_\Theta(\cdot|a_{-i}, a_i')|| > \varepsilon\}$$

and

$$\nu_i^Q(a_i', a_i) = \min \left\{ \varepsilon \geq 0 \ \middle| \ \sum_{s_{-i} \in I_\varepsilon^i(a_i', a_i)|} Q(a_{-i}|a_i\} \leq \varepsilon \right\}.$$

Therefore,

$$\begin{aligned}
& \sum_{a_{-i}} ||\rho(a_{-i}, a_i) - \rho(a_{-i}, r_i))||Q(a_{-i}|a_i) \\
= & \sum_{\substack{a_{-i} \\ :||\rho(a_{-i},a_i)-\rho(a_{-i},r_i))||>\nu_i^Q}} ||\rho(a_{-i}, a_i) - \rho(a_{-i}, r_i))||Q(a_{-i}|a_i) \\
& \quad + \sum_{\substack{a_{-i} \\ :||\rho(a_{-i},a_i)-\rho(a_{-i},r_i))||\leq\nu_i^Q}} ||\rho(a_{-i}, a_i) - \rho(a_{-i}, r_i))||Q(a_{-i}|a_i) \\
\leq & \ 2\nu_i^Q + \nu_i^Q \\
= & \ 3\nu_i^Q
\end{aligned}$$

Lemma E: Suppose that $Q \in \Delta(\Theta \times A)$ and define a system of rewards $Z = (\zeta_i)_{i \in N}$ where

$$\zeta_i(a_{-i}, a_i) = \varepsilon \frac{Q(a_{-i}|a_i)}{||Q(\cdot|a_i)||_2}$$

for each $(a_{-i}, a_i) \in A$. Then for each $a_i, a_i' \in A_i$,

$$\sum_{a_{-i}} [\zeta_i(a_{-i}, a_i) - \zeta_i(a_{-i}, a_i')] \, Q(a_{-i}|a_i) \geq \frac{\varepsilon}{2\sqrt{|A|}} \Lambda_i^Q.$$

Proof: Since

$$\left\| \frac{Q(\cdot|a_i)}{||Q(\cdot|a_i)||_2} - \frac{Q(\cdot|a_i')}{||Q(\cdot|a_i')||_2} \right\|^2 = 2 \left[1 - \frac{Q(\cdot|a_i') \cdot Q(|a_i)}{||Q(\cdot|a_i)||_2 ||Q(\cdot|a_i')||_2} \right]$$

and $||Q(\cdot|a_i)||_2 \geq \frac{1}{\sqrt{|A-i|}} \geq \frac{1}{\sqrt{|A|}}$, we conclude that

$$\begin{aligned}
\sum_{a_{-i}} [\zeta_i(a_{-i},a_i) - \zeta_i(a_{-i},a_i')]\, Q(a_{-i}|a_i) &= \sum_{a_{-i}} \left[\varepsilon \frac{Q(a_{-i}|a_i)}{||Q(\cdot|a_i)||_2} - \varepsilon \frac{Q(a_{-i}|a_i')}{||Q(\cdot|a_i')||_2}\right] Q(a_{-i}|a_i) \\
&= \varepsilon \left[\frac{Q(\cdot|a_i)\cdot Q(|a_i)}{||Q(\cdot|a_i)||_2} - \frac{Q(\cdot|a_i')\cdot Q(|a_i)}{||Q(\cdot|a_i')||_2}\right] \\
&= \varepsilon ||Q(\cdot|a_i)||_2 \left[1 - \frac{Q(\cdot|a_i')\cdot Q(|a_i)}{||Q(\cdot|a_i)||_2 ||Q(\cdot|a_i')||_2}\right] \\
&= \frac{\varepsilon}{2} ||Q(\cdot|a_i)||_2 \left\| \frac{Q(\cdot|a_i)}{||Q(\cdot|a_i)||_2} - \frac{Q(\cdot|a_i')}{||Q(\cdot|a_i')||_2} \right\|^2 \\
&\geq \frac{\varepsilon}{2\sqrt{|A|}} \Lambda_i^Q.
\end{aligned}$$

8.2. Proof of Theorem B

We will prove part (b) of Theorem B first. To begin, define beliefs $\mu_i(\cdot|r_i,\pi,t_i) \in \Delta(\Theta \times T_{-i})$ for agent i at each information set $(r_i,\pi,t_i) \in A_i \times \Pi \times T_i$ as in Section 5. In addition, define $\overline{w}_i(\cdot,\pi,t_i)$ and $w^*_{-i}(\pi,t_{-i})$ as in Section 5. Let $\alpha^*_i(t_i) = f_i(t_i)$ and recall that β^*_i is defined for agent i as follows: for each $(r_i,\pi,t_i) \in A_i \times \Pi \times T_i$, let

$$\beta^*_i(r_i,\pi,t_i) \in \arg\max_{u_i \in H_i} \sum_{t_{-i}\in T_{-i}} \sum_{\theta\in\Theta} \Big[v_i(\varphi(u_i, w^*_{-i}(\pi,t_{-i})),\theta,t_i) + \hat{y}_i(u_i, w^*_{-i}(\pi,t_{-i}))\Big]\mu_i(\theta,t_{-i}|r_i,\pi,t_i).$$

Choose $\varepsilon > 0$ and define a system of rewards $Z = (\zeta_i)_{i\in N}$ where

$$\zeta_i(a_{-i},a_i) = \varepsilon \frac{Q(a_{-i}|a_i)}{||Q(\cdot|a_i)||_2}.$$

Since $0 \leq \frac{Q(a_{-i}|a_i)}{||Q(\cdot|a_i)||_2} \leq 1$, for all i, a_{-i} and a_i.it follows that

$$0 \leq \zeta_i(a_{-i},a_i) \leq \varepsilon.$$

Next suppose that

$$0 < \delta < \frac{\varepsilon}{12nM\sqrt{|A|}}.$$

We will show that (α^*,β^*,μ) is an ICPBE in the game $\Gamma(\mathbb{D},Z)$ whenever $\max_i \nu_i^Q \leq \delta \min_i \Lambda_i^Q$. To accomplish this, we must show that (α^*,β^*) is truthful, that first stage deviations are unprofitable and that coordinated deviations across stages are unprofitable.

Part 1: To show that (α^*,β^*) is truthful, we must show that

$$\beta^*_i(f_i(t_i),\pi,t_i) = \sum_{\theta\in\Theta} v_i(\cdot,\theta,t_i)\pi(\theta) = \overline{w}_i(\cdot,\pi,t_i)$$

for all $\pi \in \Pi$ and $t_i \in T_i$. i.e., that

$$\overline{w}_i(\cdot, \pi, t_i) \in \arg\max_{u_i \in H_i} \sum_{t_{-i} \in T_{-i}} \sum_{\theta \in \Theta} \left[v_i(\varphi(u_i, w^*_{-i}(\pi, t_{-i}), \theta, t_i) + \hat{y}_i(u_i, w^*_{-i}(\pi, t_{-i})) \right] \mu_i(\theta, t_{-i} | f_i(t_i), \pi, t_i)$$

for each t_i and each $\pi \in \Pi$. To see this, note that for each $u_i \in H_i$,

$$\begin{aligned}
&\sum_{t_{-i} \in T_{-i}} \sum_{\theta \in \Theta} \Big[v_i(\varphi(u_i, w^*_{-i}(\pi, t_{-i})), \theta, t_i) + \hat{y}_i(u_i, w^*_{-i}(\pi, t_{-i}) \Big] \mu_i(\theta, t_{-i} | f_i(t_i), \pi, t_i) \\
&= \sum_{\substack{t_{-i} \in T_{-i} \\ : \rho(f_{-i}(t_{-i}), f_i(t_i)) = \pi}} \sum_{\theta \in \Theta} \Big[v_i(\varphi(u_i, w^*_{-i}(\pi, t_{-i})), \theta, t_i) + \hat{y}_i(u_i, w^*_{-i}(\pi, t_{-i}) \Big] \left[\frac{\rho_\theta(f_{-i}(t_{-i}), f_i(t_i)) P(t_{-i} | t_i)}{\sum_{\hat{t}_{-i} : \rho(f_{-i}(\hat{t}_{-i}), f_i(t_i)) = \pi} P(\hat{t}_{-i} | t_i)} \right] \\
&= \sum_{\substack{t_{-i} \in T_{-i} \\ : \rho(f(t)) = \pi}} \left[\sum_{\theta \in \Theta} v_i(\varphi(u_i, w^*_{-i}(\pi, t_{-i})), \theta, t_i) \rho_\theta(f(t)) + \hat{y}_i(u_i, w^*_{-i}(\pi, t_{-i})) \right] \left[\frac{P(t_{-i} | t_i)}{\sum_{\hat{t}_{-i} : \rho(f_{-i}(\hat{t}_{-i}), f_i(t_i)) = \pi} P(\hat{t}_{-i} | t_i)} \right] \\
&= \sum_{\substack{t_{-i} \in T_{-i} \\ : \rho(f(t)) = \pi}} \left[\sum_{\theta \in \Theta} v_i(\varphi(u_i, w^*_{-i}(\pi, t_{-i})), \theta, t_i) \pi(\theta) + \hat{y}_i(u_i, w^*_{-i}(\pi, t_{-i})) \right] \left[\frac{P(t_{-i} | t_i)}{\sum_{\hat{t}_{-i} : \rho(f_{-i}(\hat{t}_{-i}), f_i(t_i)) = \pi} P(\hat{t}_{-i} | t_i)} \right] \\
&= \sum_{\substack{t_{-i} \in T_{-i} \\ : \rho(f(t)) = \pi}} \Big[\overline{w}_i(\varphi(u_i, w^*_{-i}(\pi, t_{-i})), \pi, t_i) + \hat{y}_i(u_i, w^*_{-i}(\pi, t_{-i}) \Big] \left[\frac{P(t_{-i} | t_i)}{\sum_{\hat{t}_{-i} : \rho(f_{-i}(\hat{t}_{-i}), f_i(t_i)) = \pi} P(\hat{t}_{-i} | t_i)} \right] \\
&\leq \sum_{\substack{t_{-i} \in T_{-i} \\ : \rho(f(t)) = \pi}} \Big[\overline{w}_i(\varphi(\overline{w}_i(\cdot, \pi, t_i), w^*_{-i}(\pi, t_{-i})), \pi, t_i) + \hat{y}_i(\overline{w}_i(\cdot, \pi, t_i, w^*_{-i}(\pi, t_{-i})) \Big] \left[\frac{P(t_{-i} | t_i)}{\sum_{\hat{t}_{-i} : \rho(f_{-i}(\hat{t}_{-i}), f_i(t_i)) = \pi} P(\hat{t}_{-i} | t_i)} \right] \\
&= \sum_{t_{-i} \in T_{-i}} \sum_{\theta \in \Theta} \Big[v_i(\varphi(\overline{w}_i(\cdot, \pi, t_i), w^*_{-i}(\pi, t_{-i})), \theta, t_i) + \hat{y}_i(\overline{w}_i(\cdot, \pi, t_i), w^*_{-i}(\pi, t_{-i}) \Big] \mu_i(\theta, t_{-i} | r_i, \pi, t_i).
\end{aligned}$$

Therefore, (α^*, β^*) is truthful.

Part 2: To show that deviations at second stage information sets are unprofitable, suppose that all players use α^*_i in stage 1 and players $j \neq i$ use β^*_j in stage 2. Then, upon observing π and having reported truthfully in stage 1, it follows from the definition of β^*_j that each player $j \neq i$ reports $\beta^*_j(f_j(t_j), \pi, t_j) = \overline{w}_j(\cdot, \pi, t_j)$ in stage 2. Therefore, the second stage expected payoff to player i who reports $u_i \in H_i$ given the beliefs μ_i defined above is

$$\sum_{t_{-i} \in T_{-i}} \sum_{\theta \in \Theta} \left[v_i(\varphi(u_i, w^*_{-i}(\pi, t_{-i})), \theta, t_i) + \hat{y}_i(u_i, w^*_{-i}(\pi, t_{-i}) \right] \mu_i(\theta, t_{-i} | f_i(t_i), \pi, t_i)$$

so the definition of β^*_i implies that

$$\overline{w}_i(\cdot, \pi, t_i) \in \arg\max_{u_i \in H_i} \sum_{t_{-i} \in T_{-i}} \sum_{\theta \in \Theta} \left[v_i(\varphi(u_i, w^*_{-i}(\pi, t_{-i})), \theta, t_i) + \hat{y}_i(u_i, w^*_{-i}(\pi, t_{-i})) \right] \mu_i(\theta, t_{-i} | f_i(t_i), \pi, t_i).$$

Part 3: To show that coordinated deviations across stages are unprofitable for player i, we assume that other players use $(\alpha^*_{-i}, \beta^*_{-i})$ and we must show that, for all $t_i \in T_i$ and all $r_i \in A_i$, we have

$$\sum_{t_{-i}\in T_{-i}} \left[\sum_{\theta\in\Theta} v_i(\varphi(\overline{w}_i(\cdot, \rho(f(t)), t_i)), w^*_{-i}(\rho(f(t), t_{-i})), \theta, t_i)\rho_\theta(f(t)) + \hat{y}_i(\overline{w}_i(\cdot, \rho(f(t)), t_i)), w^*_{-i}(\rho(f(t)), t_{-i}))\right] P(t_{-i}|t_i)$$
$$+ \sum_{t_{-i}\in T_{-i}} \zeta_i(f(t))]P(t_{-i}|t_i)$$
$$\geq \sum_{t_{-i}\in T_{-i}} \max_{u_i\in H_i} \left[\sum_{\theta\in\Theta} v_i(\varphi(u_i, w^*_{-i}(f_{-i}(t_{-i}), r_i), t_{-i})), \theta, t_i)\rho_\theta(f(t)) + \hat{y}_i(u_i, w^*_{-i}(f_{-i}(t_{-i}), r_i), t_{-i}))\right] P(t_{-i}|t_i)$$
$$+ \sum_{t_{-i}\in T_{-i}} \zeta_i(f_{-i}(t_{-i}), r_i)P(t_{-i}|t_i).$$

Since

$$\sum_{\theta\in\Theta} v_i(\varphi(\overline{w}_i(\cdot, \rho(f_{-i}(t_{-i}), r_i), t_i), w^*_{-i}(\rho(f_{-i}(t_{-i}), r_i), t_{-i})), \theta, t_i)\rho_\theta(f_{-i}(t_{-i}), r_i)$$
$$+ \hat{y}_i(\overline{w}_i(\cdot, \rho(f_{-i}(t_{-i}), r_i), t_i), w^*_{-i}(\rho(f_{-i}(t_{-i}), r_i), t_{-i}))$$
$$\geq \sum_{\theta\in\Theta} v_i(\varphi(u_i, w^*_{-i}(\rho(f_{-i}(t_{-i}), r_i), t_{-i})), \theta, t_i)\rho_\theta(f_{-i}(t_{-i}), r_i)$$
$$+ \hat{y}_i(u_i, w^*_{-i}(\rho(f_{-i}(t_{-i}), r_i), t_{-i}))$$

for all $u_i \in H_i$, it follows from Lemma C that for each t_{-i} and each u_i, r_i and t_i,

$$\sum_{\theta\in\Theta} v_i(\varphi(\overline{w}_i(\cdot,\rho(f(t)),t_i)), w^*_{-i}(\rho(f(t)),t_{-i})),\theta,t_i)\rho_\theta(f(t)) + \hat{y}_i(\overline{w}_i(\cdot,\rho(f(t)),t_i)), w^*_{-i}(\rho(f(t)),t_{-i}))$$

$$- \left[\sum_{\theta\in\Theta} v_i(\varphi(u_i, w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i})),\theta,t_i)\rho_\theta(f(t)) + \hat{y}_i(u_i, w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i}))\right]$$

$$= \sum_{\theta\in\Theta} v_i(\varphi(\overline{w}_i(\cdot,\rho(f(t)),t_i)), w^*_{-i}(\rho(f(t)),t_{-i})),\theta,t_i)\rho_\theta(f(t)) + \hat{y}_i(\overline{w}_i(\cdot,\rho(f(t)),t_i)), w^*_{-i}(\rho(f(t),t_{-i})))$$

$$- \Bigg[\sum_{\theta\in\Theta} v_i(\varphi(\overline{w}_i(\cdot,\rho(f_{-i}(t_{-i}),r_i),t_i), w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i})),\theta,t_i)\rho_\theta(f_{-i}(t_{-i}),r_i)$$

$$+ \hat{y}_i(\overline{w}_i(\cdot,\rho(f_{-i}(t_{-i}),r_i),t_i), w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i}))\Bigg]$$

$$+ \Bigg[\sum_{\theta\in\Theta} v_i(\varphi(\overline{w}_i(\cdot,\rho(f_{-i}(t_{-i}),r_i),t_i), w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i})),\theta,t_i)\rho_\theta(f_{-i}(t_{-i}),r_i)$$

$$+ \hat{y}_i(\overline{w}_i(\cdot,\rho(f_{-i}(t_{-i}),r_i),t_i), w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i}))\Bigg]$$

$$- \left[\sum_{\theta\in\Theta} v_i(\varphi(u_i, w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i})),\theta,t_i)\rho_\theta(f_{-i}(t_{-i}),r_i) + \hat{y}_i(u_i, w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i}))\right]$$

$$+ \sum_{\theta\in\Theta} v_i(\varphi(u_i, w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i})),\theta,t_i)[\rho_\theta(f_{-i}(t_{-i}),r_i) - \rho_\theta(f(t))]$$

$$\geq -(2n-1)M||\rho(f(t)) - \rho(f_{-i}(t_{-i}),r_i))|| - M||\rho(f(t)) - \rho(f_{-i}(t_{-i}),r_i))||$$

Therefore,

$$\sum_{t_{-i}\in T_{-i}}\sum_{\theta\in\Theta}\left[v_i(\varphi(\overline{w}_i(\cdot,\rho(f(t),t_i)), w^*_{-i}(\rho(f(t),t_{-i})),\theta,t_i)\rho_\theta(f(t)) + \hat{y}_i(\overline{w}_i(\cdot,\rho(f(t),t_i)), w^*_{-i}(\rho(f(t),t_{-i})))\right]P(t_{-i}|t_i)$$

$$+ \sum_{t_{-i}\in T_{-i}} \zeta_i(f(t))P(t_{-i}|t_i)$$

$$\geq \sum_{t_{-i}\in T_{-i}} \max_{u_i\in H_i}\left[\sum_{\theta\in\Theta} v_i(\varphi(u_i, w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i})),\theta,t_i)\rho_\theta(f(t)) + \hat{y}_i(u_i, w^*_{-i}(\rho(f_{-i}(t_{-i}),r_i),t_{-i}))\right]P(t_{-i}|t_i)$$

$$+ \sum_{t_{-i}\in T_{-i}} \zeta_i(f_{-i}(t_{-i}),r_i)P(t_{-i}|t_i)$$

$$+ \sum_{t_{-i}\in T_{-i}} \left[\zeta_i(f(t)) - \zeta_i(f_{-i}(t_{-i}),r_i)\right]P(t_{-i}|t_i) - 2nM\sum_{t_{-i}\in T_{-i}} ||\rho(f(t)) - \rho(f_{-i}(t_{-i}),r_i))||P(t_{-i}|t_i)$$

To complete the proof, we must show that

$$\sum_{a_{-i}} \left[\zeta_i(a_{-i},f_i(t_i)) - \zeta_i(a_{-i},r_i)\right]Q(a_{-i}|f_i(t_i)) \geq 2nM\sum_{a_{-i}} ||\rho(a_{-i},f_i(t_i)) - \rho(a_{-i},r_i))||Q(a_{-i}|f_i(t_i))$$

Applying Lemmas D and E and the definition of information decomposition, we conclude that

$$
\begin{aligned}
&\sum_{t_{-i}\in T_{-i}} \left[\zeta_i(f(t)) - \zeta_i(f_{-i}(t_{-i}), r_i)\right] P(t_{-i}|t_i) - 2nM \sum_{t_{-i}\in T_{-i}} ||\rho(f(t)) - \rho(f_{-i}(t_{-i}), r_i))||P(t_{-i}|t_i) \\
=\ & \sum_{a_{-i}} \sum_{\substack{t_{-i}\in T_{-i} \\ :f_{-i}(t_{-i})=a_{-i}}} \left[\zeta_i(f(t)) - \zeta_i(f_{-i}(t_{-i}), r_i)\right] P(t_{-i}|t_i) \\
& \qquad -2nM \sum_{a_{-i}} \sum_{\substack{t_{-i}\in T_{-i} \\ :f_{-i}(t_{-i})=a_{-i}}} ||\rho(f(t)) - \rho(f_{-i}(t_{-i}), r_i))||P(t_{-i}|t_i) \\
=\ & \sum_{a_{-i}} \left[\zeta_i(a_{-i}, f_i(t_i)) - \zeta_i(a_{-i}, r_i)\right] \left[\sum_{\substack{t_{-i}\in T_{-i} \\ :f_{-i}(t_{-i})=a_{-i}}} P(t_{-i}|t_i) \right] \\
& \qquad -2nM \sum_{a_{-i}} ||\rho(a_{-i}, f_i(t_i)) - \rho(a_{-i}, r_i))|| \left[\sum_{\substack{t_{-i}\in T_{-i} \\ :f_{-i}(t_{-i})=a_{-i}}} P(t_{-i}|t_i) \right] \\
=\ & \sum_{a_{-i}} \left[\zeta_i(a_{-i}, f_i(t_i)) - \zeta_i(a_{-i}, r_i)\right] Q(a_{-i}|f_i(t_i)) - 2nM \sum_{a_{-i}} ||\rho(a_{-i}, f_i(t_i)) - \rho(a_{-i}, r_i))||Q(a_{-i}|f_i(t_i)) \\
\geq\ & \frac{\varepsilon}{2\sqrt{|A|}} \Lambda_i^Q - (2nM)(3\nu_i^Q) \\
\geq\ & 0
\end{aligned}
$$

To prove part (a) of the theorem, note that part 3 above shows that

$$
\begin{aligned}
&\sum_{t_{-i}\in T_{-i}} \left[\zeta_i(f(t)) - \zeta_i(f_{-i}(t_{-i}), r_i)\right] P(t_{-i}|t_i) - 2nM \sum_{t_{-i}\in T_{-i}} ||\rho(f(t)) - \rho(f_{-i}(t_{-i}), r_i))||P(t_{-i}|t_i) \\
\geq\ & \frac{\varepsilon}{2\sqrt{|A|}} \Lambda_i^Q - 2nM \sum_{t_{-i}\in T_{-i}} ||\rho(f(t)) - \rho(f_{-i}(t_{-i}), r_i))||P(t_{-i}|t_i).
\end{aligned}
$$

If $\Lambda_i^Q > 0$, then choosing $\varepsilon > 0$ sufficiently large proves the result.

9. bibliography

References

[1] Clarke, E. (1971), "Multipart Pricing of Public goods," *Public Choice* 8, 19-33.

[2] Cremer, J. and R. P. McLean, (1985), "Optimal Selling Strategies under Uncertainty for a Discriminatory Monopolist when Demands Are Interdependent," *Econometrica,* 53, 345-61.

[3] Cremer, J. and R. P. McLean, (1988), "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica,* 56, 1247-57.

[4] Green, J. and J.-J. Laffont (1979), *Incentives in Public Decision Making.* Amsterdam: North-Holland.

[5] Groves, T. (1973), "Incentives in Teams," *Econometrica* 41, 617-631.

[6] Ledyard, J.O. and T.R. Palfrey (1994), "Voting and Lottery Drafts as Efficient Public Goods Mechanisms," *Review of Economic Studies* 61, 327–355.

[7] Ledyard, J.O. and T.R. Palfrey (2002), "The Approximation of Efficient Public Good Mechanisms by Simple Voting Schemes," *Econometrica* 67, 435–448.

[8] McAfee, P. and P. Reny (1992), "Correlated Information and Mechanism Design," *Econometrica* 60, 395-421.

[9] McLean, R. and A. Postlewaite (2002), "Informational Size and Incentive Compatibility," *Econometrica 70, 2421-2454.*

[10] McLean, R. and A. Postlewaite (2004), "Informational Size and Efficient Auctions," *Review of Economic Studies* 71, 809-827.

[11] McLean, R. and A. Postlewaite (2014), "Implementation with Interdependent Valuations," *forthcoming, Theoretical Economics.*

[12] Mount, K. and S. Reiter (1974), "The Informational Size of Message Spaces," *Journal of Economic Theory* 8, 161–192.

[13] Postlewaite, A. and D. Schmeidler, (1986), "Implementation in Differential Information Economies," *Journal of Economic Theory*, June 1986, 39, 14-33.

[14] Vickrey, W. (1961), "Counterspeculation, Auctions and Competitive Sealed tenders," *Journal of Finance* 16, 8-37.