Econometrica, Vol. 70, No. 6 (November, 2002), 2421–2453

INFORMATIONAL SIZE AND INCENTIVE COMPATIBILITY

BY RICHARD MCLEAN AND ANDREW POSTLEWAITE[1]

We examine a general equilibrium model with asymmetrically informed agents. The presence of asymmetric information generally presents a conflict between incentive compatibility and Pareto efficiency. We present a notion of informational size and show that the conflict between incentive compatibility and efficiency can be made arbitrarily small if agents are of sufficiently small informational size.

KEYWORDS: Incentive compatibility, mechanism design, incomplete information.

1. INTRODUCTION

THE INCOMPATIBILITY OF PARETO EFFICIENCY and incentive compatibility is a central theme in economics and game theory. The issues associated with this incompatibility are particularly important in the design of resource allocation mechanisms in the presence of asymmetrically informed agents where the need to acquire information from agents in order to compute efficient outcomes and the incentives agents have to misrepresent that information for personal gain come into conflict. Despite a large literature that focuses on these issues, there has been little work aimed at understanding those situations in which informational asymmetries are quantitatively important.

Virtually every transaction is characterized by some asymmetry of information: any investor who buys or sells a share of stock generally knows *something* relevant to the value of the share that is not known to the person on the other side of the transaction. In order to focus on more salient aspects of the problem, many models (rightly) ignore the incentive problems associated with informational asymmetries in the belief that, for the problem at hand, agents are "informationally small." However, few researchers have investigated the circumstances under which an analysis that ignores these incentive problems will yield results similar to those obtained when these problems are fully accounted for.

In this paper, we study a class of mechanism design problems. Our goal is to formalize informational size in a way that, when agents are informationally small,

[1] Early versions of this paper circulated under the titles "Informational Smallness and Pareto Efficiency" and "Informational Smallness and Incentive Compatibility." We thank Jacques Cremer, Dean Foster, Andreu Mas-Colell, Stephen Morris, David Schmeidler, and particularly, Faruk Gul for helpful conversations and Rajiv Vohra for a very careful reading. We also thank the numerous participants in seminars at which this paper was presented for helpful comments. We also thank an editor and the referees for very helpful comments, and thank Ichiro Obara for pointing out an error in an earlier version. Postlewaite gratefully acknowledges support from the National Science Foundation.

one can ignore the incentive issues associated with the presence of asymmetric information without substantially affecting the analysis of these problems.

We analyze a pure exchange economy with incomplete information in which there is uncertainty regarding the characteristics of the goods that are traded and hence, uncertainty regarding the utility agents will derive from the goods. In our model, the set of states of nature is finite, with each state corresponding to a given profile of characteristics for the goods. Hence, each state of nature corresponds to a complete information Arrow-Debreu pure exchange economy. Agents do not know the state of nature, but each agent privately observes a signal that is correlated with the state of nature.

Our objective is to determine when an arbitrary allocation, conditional on the unobservable state, can be approximated in utility by an incentive compatible allocation. We focus on the case of *negligible aggregate uncertainty:* the state of nature can be inferred with high precision from all agents' signals. We show that approximations are possible when (i) each agent is informationally small in the sense that the conditional distribution on the state of nature does not vary much in that agent's signal if other agents' signals are known, and (ii) for each agent, the distributions on the state space, conditional on different signals the agent might receive, are not "too close." More specifically, we show that any given precision of approximation is possible if each agent is sufficiently informationally small relative to the variability of the conditional distributions on the state space conditioned on his possible signals.

In mechanism design problems, truthful reporting can be ensured with a scheme of the kind suggested by Cremer-McLean (1985): each agent is rewarded when he announces a signal that is likely given other agents' signals, and punished otherwise. Very large rewards and punishments may be necessary to ensure truthful reporting and may limit the applicability of such mechanisms for two reasons. First, large payments may be inefficient when agents are risk averse, and second, limited liability may preclude large punishments. In this paper, we identify conditions under which truthful reporting can be assured with small payments and small informational size plays an important role.

Agents will be informationally small in our sense in two natural economic settings. When all agents receive noisy signals of the state that are independent conditioned on the state and if each agent's signal is very accurate, then agents will be informationally small regardless of the number of agents. Alternatively, agents will become informationally small as the number of agents increases, regardless of the (fixed) accuracy of the agents' signals. This is a consequence of the law of large numbers and plays a crucial role in the replica theorem of Section 5.

We present our basic model in the next section, and in Section 3 we present an example illustrating the model and our results. Section 4 contains our result for economies of fixed size, and Section 5 contains our theorem for replica economies. We discuss possible extensions of our work in Section 6, related literature in Section 7, and close with a discussion section. All proofs are contained in the Appendix.

2. PRIVATE INFORMATION ECONOMIES

Let $N = \{1, 2, \dots, n\}$ denote the set of *economic agents*. Let $\Theta = \{\theta_1, \dots, \theta_m\}$ denote the (finite) *state space*, and let $T_1, T_2, \dots, T_n$ be finite sets where T_i represents the set of possible *signals* that agent i might receive. Let $J_m = \{1, \dots, m\}$. Let $T \equiv T_1 \times \cdots \times T_n$ and $T_{-i} \equiv \times_{j \neq i} T_i$. If $t \in T$, then we will often write $t = (t_{-i}, t_i)$. If X is a finite set, we will denote by Δ_X the set of probability distributions on X. If $x \in \Re^k$ for some positive integer k, then $\|x\|$ will denote the ℓ_1-norm of x and $\|x\|_2$ will denote the ℓ_2-norm of x.

In our model, nature chooses a state $\theta \in \Theta$. All uncertainty is embedded in θ: if θ were known, then information would be complete and symmetric. Examples of uncertainty of this kind include problems in which different θ's correspond to different quantities (or qualities) of oil in a field, different outcomes of a research and development program, or different underlying qualities of objects that have been manufactured in a particular way. The state of nature is unobservable but each agent i receives a "signal" t_i that is correlated with nature's choice of θ. More formally, let $(\tilde{\theta}, \tilde{t}_1, \tilde{t}_2, \dots, \tilde{t}_n)$ be an $(n+1)$-dimensional random vector taking values in $\Theta \times T$ with associated distribution $P \in \Delta_{\Theta \times T}$ where

$$P(\theta, t_1, \dots, t_n) = \text{Prob}\{\tilde{\theta} = \theta, \tilde{t}_1 = t_1, \dots, \tilde{t}_n = t_n\}.$$

We assume that for each θ, $\text{Prob}\{\tilde{\theta} = \theta\} > 0$ and for each $t \in T$, $\text{Prob}\{\tilde{t} = t\} > 0$. For $t \in T$, let $P_\Theta(\cdot|t) \in \Delta_\Theta$ denote the induced conditional probability measure on Θ, and let $I_\theta \in \Delta_\Theta$ denote the degenerate measure that puts probability one on state θ.

The *consumption set* of each agent is $\Re^\ell_+$ and $w_i \in \Re^\ell_+$, $w_i \neq 0$, denotes the *initial endowment* of agent i (an agent's initial endowment is independent of the state θ). For each $\theta \in \Theta$, let $u_i(\cdot, \theta)$: $\Re^\ell_+ \to \Re$ be the utility function of agent i in state θ. In this specification, the utility that an agent derives from a given bundle of goods is determined by the state. The utility from owning an oil field (or a share of the field) will be determined by the quantity and quality of the oil in the field, the utility of a share of a company engaging in a research and development project will be determined by the product that emerges from the project, and the utility from a used car depends on the quality of the engineering design of the car. The assumption that u depends only on the bundle of goods and on the state θ, but not on the profile of agents' types t, is clearly without loss of generality, since one can always include t as part of θ. We will, however, make assumptions below regarding how much different agents know about the state θ. Given these assumptions, our model captures better the case in which θ embodies uncertainty about the characteristics of goods that might be of direct interest to many agents rather than the case in which θ embodies uncertainty about a single agent's utility function.

We will assume that each $u_i(\cdot, \theta)$ is continuous, $u_i(0, \theta) = 0$ and satisfies the following monotonicity assumption: if $x, y \in \Re^\ell_+$, $x \geq y$ and $x \neq y$, then $u_i(x, \theta) > u_i(y, \theta)$.

Each $\theta \in \Theta$ gives rise to a pure exchange economy and these economies will play an important role in the analysis that follows. Formally, let $e(\theta) = \{w_i, u_i(\cdot, \theta)\}_{i\in N}$ denote the *complete information economy* (CIE) corresponding to state θ. For each $\theta \in \Theta$, a *complete information economy (CIE) allocation* for $e(\theta)$ is a collection $\{x_i(\theta)\}_{i\in N}$ satisfying $x_i(\theta) \in \Re^\ell_+$ for each i and $\sum_{i\in N}(x_i(\theta) - w_i) \le 0$. For each $\theta \in \Theta$, a CIE allocation $\{x_i(\theta)\}_{i\in N}$ for the complete information economy $e(\theta)$ is *efficient* if there is no other CIE allocation $\{y_i(\theta)\}_{i\in N}$ for $e(\theta)$ such that

$$u_i(y_i(\theta), \theta) > u_i(x_i(\theta), \theta)$$

for each $i \in N$.

The collection $(\{e(\theta)\}_{\theta\in\Theta}, \tilde{\theta}, \tilde{t}, P)$ will be called a *private information economy* (PIE for short). An *allocation* $x = (x_1, x_2, \dots, x_n)$ for the PIE $(\{e(\theta)\}_{\theta\in\Theta}, \tilde{\theta}, \tilde{t}, P)$ is a collection of functions x_i: $T \to \Re^\ell_+$ satisfying $\sum_{i\in N}(x_i(t) - w_i) \le 0$ for all $t \in T$. We will not distinguish between w_i and the constant allocation that assigns the bundle w_i to agent i for all $t \in T$.

We next introduce standard notation in order to define the properties of allocations. For a given PIE allocation $x = (x_1, x_2, \dots, x_n)$ define

$$\begin{aligned} U_i(x_i, t_i'|t_i) &= \sum_{\theta\in\Theta}\sum_{t_{-i}\in T_{-i}} u_i(x_i(t_{-i}, t_i'), \theta)P(\theta, t_{-i} \mid t_i) \\ &= E[u_i(x_i(\tilde{t}_{-i}, t_i'), \tilde{\theta})|\tilde{t}_i = t_i] \end{aligned}$$

for each $t_i', t_i \in T_i$ and

$$\begin{aligned} U_i(x_i \mid t) &= \sum_{\theta\in\Theta} u_i(x_i(t), \theta)P_\Theta(\theta \mid t) \\ &= E[u_i(x_i(\tilde{t}), \tilde{\theta}) \mid \tilde{t} = t] \end{aligned}$$

for each $t \in T$.

A PIE allocation $x = (x_1, x_2, \dots, x_n)$ is said to be:

- *incentive compatible* (IC) if

$$U_i(x_i, t_i|t_i) \ge U_i(x_i, t_i'|t_i)$$

for all $i \in N$, and all $t_i, t_i' \in T_i$;

- *ex post individually rational* (XIR) if

$$U_i(x_i \mid t) \ge U_i(w_i \mid t)$$

for all $i \in N$ and for all $t \in T$;

- *ex post ε-efficient* ($\mathrm{X}_\varepsilon\mathrm{E}$) if there exists $E \subseteq T$ such that $\mathrm{Prob}\{\tilde{t} \in E\} \ge 1 - \varepsilon$ and for no other PIE allocation $y(\cdot)$ is it true that, for some $t \in E$,

$$U_i(y_i(t) \mid t) > U_i(z_i(t) \mid t) + \varepsilon$$

for all $i \in N$.

Note that allocations can depend on agents' types (their information) but not on θ, which is assumed to be unobservable. Hence, our use of the term "ex post" refers to events that occur *after* the realization of the signal profile t but *before* the realization of the state θ.

3. EXAMPLE

There are six agents, three of whom are potential buyers of cars (B) and three of whom are potential sellers (S). The engineering design of the car is either flawed or not flawed with equal probability. Let F denote the state in which the design is flawed and let N denote the state in which the design is not flawed. Agents cannot observe whether the design is flawed or not, but sellers have private information that we represent as signals (G or B) correlated with the state of nature. Buyers receive no signal and, therefore, have no private information. The sellers' signals are independent conditional on the state and the matrix of conditional probabilities of the signals given the state is:

	State	
	N	F
Signal		
G	ρ	$1-\rho$
B	$1-\rho$	ρ

All agents have linear, separable utility functions. Buyers and sellers of the cars have respective utilities of $u_B(m, x; \theta)$ and $u_S(m, x; \theta)$ for m units of money and x cars in the two states, $\theta = F$ and $\theta = N$. These utilities are given in the following table:

	State	
	N	F
Agent		
Buyer	$m+24x$	$m+8x$
Seller	$m+20x$	$m+4x$

Each buyer has an initial endowment of money and no car. Each seller initially has no money and one car. Given the utility functions, ex post efficiency dictates that all cars be transferred from sellers to buyers. If the goal is to effect a transaction that is ex post individually rational and ex post Pareto efficient, we must induce the sellers to truthfully reveal their signals in order to determine whether the payment should be relatively high (when the design is not flawed) or relatively low (when it is flawed). An obvious incentive compatibility problem arises since the sellers have a clear interest in making it appear that the design is not flawed.

Consider the following revelation mechanism. Sellers announce their signals and the state of nature is "estimated" to be N if a majority of the sellers announce G, and F if a majority of the sellers announce B. Each seller will then transfer his car to a buyer in return for a payment that depends on both the estimated state and his announcement according to the following table:

Seller's Own Announcement (t_i)	Estimated State (θ)	Payment
G	N	22
B	N	21
G	F	5
B	F	6

For example, if the sellers other than i both announce G, then N is the estimated state independent of i's announcement. In this case, i receives a payment of 22 if he announces G and 21 if he announces B. One can interpret the mechanism as specifying a payment that depends on the majority announcement and "punishes" a seller (by lowering the transfer price by 1) whose announcement differs from the majority.

We note several things about the mechanism. First, if ρ is close to 1, then the information of the three sellers is sufficient to predict the state nearly perfectly. In particular, when ρ is close to 1, $P_\Theta(N|t_1, t_2, t_3) \approx 1$ if a majority of sellers receive the signal G, while $P_\Theta(F|t_1, t_2, t_3) \approx 1$ if a majority of sellers receive B. Hence, the mechanism yields an ex post efficient allocation for every vector of agents' types when $\rho \approx 1$. Furthermore, the allocation is ex post individually rational if ρ is close to 1. If the profile of announced signals is $t = (t_1, t_2, t_3)$, then ex post individual rationality requires that the payment lie between $24P_\Theta(N|t) + 8P_\Theta(F|t)$ and $20P_\Theta(N|t) + 4P_\Theta(F|t)$. When $\rho \approx 1$ and a majority of sellers announce G, the estimated state is N, in which case, $24P_\Theta(N|t) + 8P_\Theta(F|t) \approx 24$ and $20P_\Theta(N|t) + 4P_\Theta(F|t) \approx 20$. A seller will receive 21 or 22, depending on his own announcement but the transfer will be ex post individually rational in either case. A similar argument applies when a majority of sellers announce B.

The mechanism is incentive compatible for ρ sufficiently close to 1. To see this, suppose that a seller receives signal B. A false report of G may change the estimated state or may leave it unchanged. The estimated state will change only when the other two sellers receive different signals. The conditional probability that the other two sellers receive different signals approaches zero as ρ approaches 1. In our example, the gain in revenue from lying when the other two sellers receive different signals is 16 and, therefore, a misreport that changes the estimated state is profitable. However, these gains contribute very little to the total expected gain in revenue from misreporting since they will be weighted by probabilities that are close to zero when ρ is close to one.

What happens when a misreport does not change the estimated state? There are two possibilities. If the other two agents receive signal B, then a false report of G results in a loss of 1. If the other two agents receive signal G, then a false report of G results in a gain of 1. When ρ is close to one, a seller who observes B will believe it very likely that the other sellers' signals are both B and he will believe it very unlikely that the other sellers' signals are both G. Hence, the contribution to the total expected gain in revenue from a misreport that does not change the estimated state is close to -1. Therefore, the total expected gain from false reporting is close to -1 and we see that for ρ sufficiently close to one, a misreport leads to an expected decrease in utility. The same argument holds for the case in which a seller observes G but falsely reports B; hence the mechanism is incentive compatible.

We are able to induce truthful revelation of information (and, consequently, ensure a Pareto efficient and individually rational outcome) in the case when ρ is close to 1 as a consequence of three features of this example. First, agents are *informationally small:* with high probability, sellers are not able to change the estimated state by misreporting their signals. Second, sellers' types are correlated. Despite the fact that a seller who receives the signal B can increase the expected price he will receive by falsely announcing the signal G, an offsetting benefit for truthful announcement is possible because the most likely signal received by either of the other two sellers is also B. If agents' types had been independent, it would be impossible to construct such an offsetting benefit. Finally, the combined information of all but one agent will, with very high probability, resolve nearly all the uncertainty about the state of nature.

The linear utilities of the example make it possible to construct a mechanism that is incentive compatible, ex post individually rational, and ex post Pareto efficient for ρ close to 1. In the case of general (nonlinear) utilities, exact Pareto efficiency will not be obtained. However, we will demonstrate that, when appropriate versions of the three conditions above hold, there will exist incentive compatible, individually rational allocations that are nearly Pareto efficient. The proof of this result will roughly parallel the construction of the mechanism of the example. The agents' announcements will be used to estimate the state of nature and, for each estimated state of nature, the outcome will be an allocation that is efficient and individually rational for that state, modified slightly so as to induce truthful revelation.

It is important to mention several features of the example that do *not* play any role in our results. To illustrate the basic idea in a straightforward way, we constructed an example in which (i) the agents' information had the form of a noisy signal of the state of nature, (ii) agents' information was independent, conditional on the state of nature, and (iii) each agent's information alone provided a very accurate estimate of the state of nature. Our analysis includes information structures with features such as these, but is not restricted to such structures.

4. ECONOMIES OF FIXED SIZE

Before stating the main result, we will discuss the three features mentioned above that are key to ensuring that an incentive compatible, individually rational and approximately Pareto efficient PIE allocation exists.

4.1. *Informational Size*

In the mechanism of the example, sellers reveal their signals and the announced signals are used to estimate the state of nature. The mechanism is incentive compatible because each seller is informationally small in the following sense: with high probability, he does not have a "large" influence on the conditional probability distribution over states when other sellers announce truthfully.

We will formalize this notion of *informational size* for general problems. If $t \in T$, recall that $P_\Theta(\cdot|t) \in \Delta_\Theta$ denotes the induced conditional probability measure on Θ. Our example suggests that a natural notion of an agent's informational size is the degree to which he can alter the posterior distribution on Θ when other agents are announcing truthfully. Any vector of agents' types $t = (t_{-i}, t_i) \in T$ induces a conditional distribution on Θ and, if agent i unilaterally changes his announced type from t_i to t_i', this conditional distribution will (in general) change. If i's type is t_i but he announces $t_i' \neq t_i$, the set

$$\{t_{-i} \in T_{-i} | \|P_\Theta(\cdot|t_{-i}, t_i) - P_\Theta(\cdot|t_{-i}, t_i')\| > \varepsilon\}$$

consists of those t_{-i} for which agent i's misrepresentation will have (at least) an "ε-effect" on the conditional distribution. Let $\nu_i^P(t_i, t_i')$ be defined as the smallest nonnegative ε (formally, the infimum over all $\varepsilon \geq 0$) such that

$$\text{Prob}\{\|P_\Theta(\cdot|\tilde{t}_{-i}, t_i) - P_\Theta(\cdot|\tilde{t}_{-i}, t_i')\| > \varepsilon | \tilde{t}_i = t_i\} \leq \varepsilon$$

and define the *informational size* of agent i as

$$\nu_i^P = \max_{t_i, t_i'} \nu_i^P(t_i, t_i').^2$$

Loosely speaking, we will say that agent i is *informationally small* with respect to P if his informational size ν_i^P is "small." An agent is informationally small if for each of his possible types t_i, he assigns small probability to the event that he can have a "large" influence on the distribution $P_\Theta(\cdot|t_{-i}, t_i)$, given his observed type.

There are several important aspects of this definition of informational size. First, note that $\nu_i^P = 0$ for every i if and only if for every $t \in T$, the probability distribution on Θ given t is the same as the probability distribution on Θ given t_{-i}. More formally, $\nu_i^P = 0$ for every i if and only if $P_\Theta(\cdot|t) = P_\Theta(\cdot|t_{-i})$ for each $t \in T$ and each i. Second, informational smallness is not determined by the "quality" of an agent's information in isolation. In the example of Section 3, $P_\Theta(\cdot|t_i)$ is nearly degenerate for each t_i when ρ is close to 1. In this case, agents have very good estimates of the true state conditional only on their own signals, yet each agent is informationally small. However, it is true that, holding other agents' information fixed, an increase in the accuracy of a given agent's signal will increase that agent's informational size.

[2] That ε occurs twice in the definition is not important. There are many functions f such that an alternative definition of informational size using $\text{Prob}\{\|P_\Theta(\cdot|\tilde{t}_{-i}, t_i) - P_\Theta(\cdot|\tilde{t}_{-i}, t_i')\| > \varepsilon | \tilde{t}_i = t_i\} \leq f(\varepsilon)$ would not qualitatively change our results.

4.2. *Negligible Aggregate Uncertainty*

In the example, the information of any pair of sellers will "almost" resolve the uncertainty regarding the state θ. We will introduce a measure that quantifies this aggregate uncertainty that we will use in our theorem. Recall that I_θ is the degenerate probability distribution on Θ that puts probability 1 on the state θ. For any $t \in T$, $\|P_\Theta(\cdot|t) - I_\theta\|$ is a measure of the degree to which the posterior on Θ resolves completely the uncertainty regarding the state. A measure of an agent's estimate of the aggregate uncertainty when agent i is of type t_i is then the probability that, conditional on t_i, the posterior on Θ is not close to I_θ for any θ. Formally, we have the following definition.

DEFINITION: Let

$$\mu_i^P = \max_{t_i \in T_i} \inf\{\varepsilon \geq 0 | \operatorname{Prob}\{\|P_\Theta(\cdot|\tilde{t}) - I_\theta\| > \varepsilon \text{ for all } \theta \in \Theta | \tilde{t}_i = t_i\} \leq \varepsilon\}.$$

We define the *aggregate uncertainty* as $\mu^P \equiv \max_i \mu_i^P$ and we will say that P exhibits *negligible aggregate uncertainty* if μ^P is small. In this case, each agent knows that, conditional on his own signal, the aggregate information of all agents will, with high probability, provide a good prediction of the true state.[3]

4.3. *Distributional Variability*

In the example in Section 3, a car seller is induced to truthfully reveal his signal by conditioning the price at which his car will be sold on whether or not his reported signal is equal to the estimated state. Denote the set of states in that example by $\Theta = \{N, F\}$. When the accuracy of the signals that sellers receive is close to 1, the probability distributions on Θ given the private signals G and B will be approximately the degenerate distributions that put probability close to 1 on the states N and F respectively. That is, $P_\Theta(N|G) \approx 1$ and $P_\Theta(F|B) \approx 1$. Hence, the difference between $P_\Theta(\cdot|G)$ and $P_\Theta(\cdot|B)$ increases as the accuracies of the sellers' signals converge to 1. This is a feature of this specific example in which agents receive noisy signals that are independent conditional on the state.

In more general problems, our ability to construct a mechanism that will give agent i an incentive to reveal his information will depend on the magnitude of the difference between $P_\Theta(\cdot|t_i)$ and $P_\Theta(\cdot|t_i')$, the conditional distributions on the states of nature given different types t_i and t_i' for agent i. We will refer to this magnitude informally as the *variability of agents' beliefs*.

To formally define the measure of variability that is convenient for our purposes, we first define a metric d on Δ_Θ as follows: for each $\alpha, \beta \in \Delta_\Theta$, let

$$d(\alpha, \beta) = \left\| \frac{\alpha}{\|\alpha\|_2} - \frac{\beta}{\|\beta\|_2} \right\|_2$$

[3] There are alternative notions of aggregate uncertainty that one might consider, some of which are somewhat weaker than the concept presented here.

where $\|\cdot\|_2$ denotes the 2-norm. Hence, $d(\alpha, \beta)$ measures the Euclidean distance between the Euclidean normalizations of α and β.

If $P \in \Delta_{\Theta\times T}$, recall that $P_\Theta(\cdot|t_i) \in \Delta_\Theta$ is the conditional distribution on Θ given that i receives signal t_i and define

$$\Lambda_i^P = \min_{t_i\in T_i} \min_{t_i'\in T_i\setminus t_i} d(P_\Theta(\cdot|t_i), P_\Theta(\cdot|t_i'))^2.$$

This is the measure of the "variability" of the conditional distribution $P_\Theta(\cdot|t_i)$ as a function of t_i. Let

$$\Delta^*_{\Theta\times T} = \{P \in \Delta_{\Theta\times T} | \text{ for each } i, P_\Theta(\cdot|t_i) \neq P_\Theta(\cdot|t_i') \text{ whenever } t_i \neq t_i'\}.$$

The set $\Delta^*_{\Theta\times T}$ is the collection of distributions on $\Theta \times T$ for which the induced conditionals are different for different types. Hence, $\Lambda_i^P > 0$ for all i whenever $P \in \Delta^*_{\Theta\times T}$.

4.4. *Results*

In this section, we present our main result on the existence of incentive compatible, individually rational, and nearly Pareto efficient allocations when aggregate uncertainty and the agents' informational sizes are both small relative to the variability of agents' beliefs. This will follow from the stronger result (Theorem 1 below) that any collection of complete information economy allocations, $\mathcal{A} = \{x(\theta)\}_{\theta\in\Theta}$, can be approximated in utility under these conditions.

THEOREM 1: *Let $\Theta = \{\theta_1, \dots, \theta_m\}$. Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and suppose that $\mathcal{A} = \{x(\theta)\}_{\theta\in\Theta}$ is a collection of associated CIE allocations with $x_i(\theta) \neq 0$ for each i and θ. For every $\varepsilon > 0$, there exists a $\delta > 0$ such that, whenever $P \in \Delta_{\Theta\times T}$ and satisfies*

$$\max_i \mu_i^P \leq \delta \min_i \Lambda_i^P,$$

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

there exists an incentive compatible PIE allocation $z(\cdot)$ for the PIE $(\{e(\theta)\}_{\theta\in\Theta}, \tilde\theta, \tilde t, P)$ and a collection $A_1, \dots, A_m$ of disjoint subsets of T satisfying:

(i) $\text{Prob}\{\tilde t \in \cup_{k=1}^m A_k\} \geq 1 - \varepsilon$;

(ii) $\text{Prob}\{\tilde\theta = \theta_k | \tilde t = t\} \geq 1 - \varepsilon$ *for each $k \in J_m$ and $t \in A_k$*;

(iii) *for all $i \in N$*,

$$u_i(x_i(\theta_k); \theta_k) \geq u_i(z_i(t); \theta_k) \geq u_i(x_i(\theta_k); \theta_k) - \varepsilon$$

for each $k \in J_m$ and $t \in A_k$.

To understand Theorem 1, first note that δ depends on ε, the collection $\{e(\theta)\}_{\theta\in\Theta}$, and the collection $\mathcal{A}$, but is independent of the distribution P. The theorem requires that the measures of aggregate uncertainty (μ_i^P) and informational size (ν_i^P) be sufficiently small relative to the variability of beliefs (Λ_i^P). For any distribution P for which these conditions hold, we can find an incentive compatible PIE allocation $z(\cdot)$ and sets $A_1, \ldots, A_m$ such that $\mathrm{Prob}\{\tilde{t} \in \bigcup_{k=1}^m A_k\} \approx 1$. Furthermore, $P_\Theta(\theta_k|t) \approx 1$ and $u_i(z_i(t); \theta_k) \approx u_i(x_i(\theta_k); \theta_k)$ for each i whenever $t \in A_k$.

If the collection $\mathscr{A} = \{x(\theta)\}_{\theta\in\Theta}$ in the statement of Theorem 1 has the property that each $x(\theta)$ is a strictly individually rational, Pareto efficient allocation for $e(\theta)$, then $z(\cdot)$ will satisfy XIR and X$_\varepsilon$E. More formally, we have the following result.

COROLLARY 1: *Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and suppose that there exists a strictly individually rational, efficient allocation for each PIE $e(\theta)$. Then for every $\varepsilon > 0$, there exists a $\delta > 0$ such that, whenever $P \in \Delta_{\Theta\times T}$ and satisfies*

$$\max_i \mu_i^P \leq \delta \min_i \Lambda_i^P,$$
$$\max_i \nu_i^P \leq \delta \min_i \Lambda_i^P,$$

there exists an allocation $z(\cdot)$ for the PIE$(\{e(\theta)\}_{\theta\in\Theta}, \tilde{\theta}, \tilde{t}, P)$ satisfying XIR, XIC, and X$_\varepsilon$E.

A second consequence that immediately follows from Theorem 1 is the following corollary.

COROLLARY 2: *Let $\Theta = \{\theta_1, \ldots, \theta_m\}$. Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and suppose that $\mathscr{A} = \{x(\theta)\}_{\theta\in\Theta}$ is a collection of associated CIE allocations such that $x_i(\theta) \neq 0$ for each i and θ. For every $\varepsilon > 0$, there exists a $\delta > 0$ such that, whenever $P \in \Delta_{\Theta\times T}$ and satisfies*

$$\max_i \mu_i^P \leq \delta \min_i \Lambda_i^P,$$
$$\max_i \nu_i^P \leq \delta \min_i \Lambda_i^P,$$

there exists an incentive compatible allocation $z(\cdot)$ for the PIE $(\{e(\theta)\}_{\theta\in\Theta}, \tilde{\theta}, \tilde{t}, P)$ such that for all $i \in N$ and all $\theta \in \Theta$,

$$\sum_{t\in T} u_i(z_i(t); \theta) P(t|\theta) \geq u_i(x_i(\theta); \theta) - \varepsilon.$$

The left-hand side of the inequality, $\sum_{t\in T} u_i(z_i(t); \theta) P(t|\theta)$, is agent i's conditional expected utility from the allocation z when the state of nature is θ. Thus, Corollary 2 states that, if aggregate uncertainty and agents' informational size are sufficiently small relative to the variability of beliefs, we can find an incentive compatible allocation that assures every agent in every state θ an expected utility that is nearly as large as his utility from the CIE allocation $x(\theta)$.

The details of the proof of Theorem 1 and Corollary 1 are left to the Appendix, but we will sketch the proof of Theorem 1 informally. Suppose that $\mathcal{A} = \{x(\theta)\}_{\theta\in\Theta}$ is a collection of CIE allocations and suppose that $\varepsilon > 0$. We partition T into $m+1$ disjoint sets with $A_k = \{t \in T \,|\, \|P(\cdot|t) - I_{\theta_k}\| \le \max_i \mu_i^P\}$ for $k = 1, \dots, m$, and $A_0 = T\backslash[\bigcup_{k\ge1} A_k]$. In words, A_k is the set of $t \in T$ for which the posterior distribution on Θ is close to the degenerate distribution that puts probability 1 on θ_k. Therefore, A_0 is the set of $t \in T$ for which the posterior is not close to I_θ for any θ.

We begin with a PIE allocation y with $y(t) = x(\theta_k)$ for $t \in A_k$, $k = 1, \dots, m$, and $y(t) = 0$ for $t \in A_0$. When aggregate uncertainty is small, the profile of agents' information $t \in T$ will, with high probability, resolve most of the uncertainty regarding the state of nature θ. There are two consequences of small aggregate uncertainty: the probability that $\tilde{t} \in A_0$ is small and, for each $t \in A_k$, $P_\Theta(\theta_k|t)$ is close to 1. Since $P_\Theta(\theta_k|t)$ is close to 1 whenever $t \in A_k$, it follows that $\sum_k u_i(y_i(t); \theta_k)P_\Theta(\theta_k|t)$ is close to $u_i(x_i(\theta_k); \theta_k)$ whenever $t \in A_k$. However, $y(\cdot)$ is not incentive compatible in general.

Suppose that i receives signal t_i and the other agents truthfully report t_{-i}. It may be the case that $(t_{-i}, t_i) \in A_j$ while $(t_{-i}, t_i') \in A_k$, $j \ne k$. Hence, i receives $x_i(\theta_j)$ if he reports t_i, while he receives $x_i(\theta_k)$ if he reports t_i'. If $x_i(\theta_k)$ results in higher utility than $x_i(\theta_j)$, agent i may have an incentive to misreport. To say that agent i is informationally small means that there is a low probability that the posteriors on Θ given (t_{-i}, t_i) and (t_{-i}, t_i') put probability close to 1 on θ_j and θ_k respectively. If an agent's informational size is small, then the expected gain to that agent from a misreported type is also small. In order to offset this (small) potential gain that i might enjoy by misreporting, we modify the bundle $x_i(\theta_k)$ that i receives when $t \in A_k$. If agent i's posteriors on Θ for different types t_i and t_i' are different for any $t_i \ne t_i'$, we can construct bundles $z_i(t)$ with the properties that (i) $z_i(t)$ is close to $x_i(\theta_k)$ for every i and for every $t \in A_k$, and (ii) the mechanism $z(\cdot)$ thus defined is incentive compatible. The fact that $z_i(t)$ is close to $x_i(\theta_k)$ for $t \in A_k$ implies that i's utility from $z_i(t)$ is close to his utility from $x_i(\theta_k)$ and the conclusion of the theorem follows.

5. LARGE ECONOMIES

In the presence of a large number of agents, we might expect agents to be informationally small. However, the presence of many agents by itself is clearly not enough for agents to be informationally small. An economy with a large number of agents who have no information, and one agent who is perfectly informed would provide a trivial counterexample to any such conjecture.

The fact that one of the agents in the example above is informationally large even though the economy is also large is not at all surprising given the asymmetry of the agents. Even in the presence of a large number of symmetric agents, all agents may be informationally large as the next example illustrates.

EXAMPLE: Let the number of agents, n, be odd. Let $\Theta = \{\alpha, \beta\}$ and let $T_i = \{a, b\}$. For each $t \in T$, let $\hat{a}(t) = \#\{i | t_i = a\}$. We now define $P \in \Delta_{\Theta \times T}$ as follows:

$$\begin{aligned} P(\alpha, t) &= \left(\frac{1}{2}\right)^n && \text{if } \hat{a}(t) \text{ is odd,} \\ P(\alpha, t) &= 0 && \text{if } \hat{a}(t) \text{ is even,} \\ P(\beta, t) &= \left(\frac{1}{2}\right)^n && \text{if } \hat{a}(t) \text{ is even,} \\ P(\beta, t) &= 0 && \text{if } \hat{a}(t) \text{ is odd.} \end{aligned}$$

It is straightforward to verify that $P(\alpha) = P(\beta) = 1/2$ and $P(t) = (1/2)^n$ for each $t \in T$. Hence, the random variable t has full support. Since $P(\alpha|t) = 1$ if $\hat{a}(t)$ is odd and $P(\beta|t) = 1$ if $\hat{a}(t)$ is even, the measure P exhibits zero aggregate uncertainty, so that the agents' signals completely determine the state of nature.

For our purposes, this example exhibits another interesting feature: while the profile of signals of the n agents completely resolves all uncertainty, the signals of any $n-1$ agents resolve nothing. Indeed, the random variable $\tilde{\theta}$ and the random vector $\tilde{t}_{-i}$ are stochastically independent for each i since

$$P(\alpha|t_{-i}) = \frac{P(\alpha|t_{-i}, a)P(t_{-i}, a) + P(\alpha|t_{-i}, b)P(t_{-i}, b)}{P(t_{-i})} = \frac{1}{2} = P(\alpha),$$

and

$$P(\beta|t_{-i}) = \frac{P(\beta|t_{-i}, a)P(t_{-i}, a) + P(\beta|t_{-i}, b)P(t_{-i}, b)}{P(t_{-i})} = \frac{1}{2} = P(\beta).$$

Thus, with probability 1, every agent will be able to "maximally" affect the posterior on Θ: the posterior probability distribution will put probability 1 on one state when he announces truthfully (the correct state if others also announce truthfully), and will put probability 1 on the other state if he misreports his signal. Hence, even in arbitrarily large, symmetric economies, informational smallness is not assured.

There are circumstance, however, under which a large number of agents will ensure small informational size. Roughly speaking, if any single agent's information adds little to the aggregate information, agents will become informationally small when the number of agents increases. We investigate next a replica framework in which this sort of substitutability of agents' information naturally occurs.

5.1. *Replica Economies*

Let $\{e(\theta)\}_{\theta \in \Theta}$ be a collection of complete information economies and recall that $J_r = \{1, 2, \ldots, r\}$. For each positive integer r and each θ, let $e^r(\theta) = \{w_{is}, u_{is}(\cdot, \theta)\}_{(i,s) \in N \times Jr}$ denote the r replicated complete information economy (r-CIE) corresponding to state θ satisfying:

(i) $w_{is} = w_i$ for all $s \in J_r$;

(ii) $u_{is}(\cdot, \theta) = u_i(\cdot, \theta)$ for all $s \in J_r$.

For any positive integer r, let $T^r = T \times \cdots \times T$ denote the r-fold Cartesian product and let $t^r = (t^r_{\cdot 1}, \ldots, t^r_{\cdot r})$ denote a generic element of T^r where $t^r_{\cdot s} = (t^r_{1s}, \ldots, t^r_{ns})$. If $P^r \in \Delta_{\Theta \times T^r}$, then $e^r = (\{e^r(\theta)\}_{\theta \in \Theta}, \tilde{\theta}, \tilde{t}^r, P^r)$ is a PIE with nr agents. If $\mathscr{A} = \{x(\theta)\}_{\theta \in \Theta}$ is a collection of CIE allocations for $\{e(\theta)\}_{\theta \in \Theta}$, let $\mathscr{A}^r = \{x^r(\theta)\}_{\theta \in \Theta}$ be the associated "replicated" collection where $x^r(\theta)$ is the CIE allocation for $e^r(\theta)$ satisfying

$$x^r_{is}(\theta) = x_i(\theta) \quad \text{for each} \quad (i, s) \in N \times J_r.$$

Definition: A sequence of replica economies $\{(\{e^r(\theta)\}_{\theta \in \Theta}, \tilde{\theta}, \tilde{t}^r, P^r)\}_{r=1}^{\infty}$ is a *conditionally independent sequence* if there exists a $P \in \Delta^*_{\Theta \times T}$ such that:

(a) for each r, each $s \in J_r$, and each $(\theta, t_1, \ldots, t_n) \in \Theta \times T$,

$$\text{Prob}\{\tilde{\theta} = \theta, \tilde{t}^r_{1s} = t_1, \tilde{t}^r_{2s} = t_2, \ldots, \tilde{t}^r_{ns} = t_n\} = P(\theta, t_1, t_2, \ldots, t_n);$$

(b) for each r and each θ, the random vectors

$$(\tilde{t}^r_{11}, \tilde{t}^r_{21}, \ldots, \tilde{t}^r_{n1}), \ldots, (\tilde{t}^r_{1r}, \tilde{t}^r_{2r}, \ldots, \tilde{t}^r_{nr})$$

are independent conditional on the event $\tilde{\theta} = \theta$;

(c) for every $\theta, \hat{\theta}$ with $\theta \neq \hat{\theta}$, there exists a $t \in T$ such that $P(t|\theta) \neq P(t|\hat{\theta})$.

Thus a conditionally independent sequence is a sequence of PIE's with nr agents containing r "copies" of each agent $i \in N$. Each copy of an agent i is identical, i.e., has the same endowment and the same utility function. Furthermore, the realizations of type profiles across cohorts are independent given the true value of $\tilde{\theta}$. As r increases each agent is becoming "small" in the economy in terms of endowment, and we show that each agent is also becoming informationally small. Note that, for large r, an agent may have a small amount of private information regarding the preferences of everyone through his information about $\tilde{\theta}$. We now state our main result for replica economies.

Theorem 2: *Let $\{e(\theta)\}_{\theta \in \Theta}$ be a collection of CIE's such that each $u_i(\cdot; \theta)$ is concave, and suppose that $\mathscr{A} = \{x(\theta)\}_{\theta \in \Theta}$ is a collection of associated CIE allocations with $x(\theta)$ strictly individually rational and efficient for $e(\theta)$. Let $\{(\{e^r(\theta)\}_{\theta \in \Theta}, \tilde{\theta}, \tilde{t}^r, p^r)\}_{r=1}^{\infty}$ be a conditionally independent sequence. Then for every $\varepsilon > 0$, there exists an integer $\hat{r} > 0$ such that for all $r > \hat{r}$, there exists an incentive compatible allocation z^r for the PIE $(\{e^r(\theta)\}_{\theta \in \Theta}, \tilde{\theta}, \tilde{t}^r, P^r)$ satisfying XIR, XIC, and $X_\varepsilon E$.*

It is important to point out that Theorem 2 is not an immediate application of Theorem 1. As the economy is replicated, agents become informationally small and aggregate uncertainty converges to zero. In addition, the measure of variability is independent of r, the size of the replication. When Corollary 2 is applied to the r-replicated PIE $(\{e^r(\theta)\}_{\theta \in \Theta}, \tilde{\theta}, \tilde{t}^r, P^r)$, the number δ can depend on r. To prove Theorem 2, we must show that δ can be chosen to depend only on the collection $\mathscr{A}$ and not on r.

6. EXTENSIONS

1. In the presence of negligible aggregate uncertainty and informationally small agents, we prove the existence of nearly efficient, incentive compatible allocations. The assumption of negligible aggregate uncertainty is not a necessary condition for the existence of approximately efficient incentive compatible allocations, however. Consider, for example, the case in which the profile of agents' types $\tilde{t}$ is independent of the state $\tilde{\theta}$. In this case, the agents' information can be ignored with no loss in efficiency and a simple incentive compatible mechanism satisfying individual rationality and (exact) ex post efficiency is available. Simply choose an allocation $\bar{x}$ that is individually rational and efficient for the economy in which agent i has initial endowment w_i and utility $\bar{u}_i(\cdot) = \sum_\theta u_i(\cdot;\theta)P_\Theta(\theta)$ where P_Θ is the marginal measure on Θ. Now define the mechanism $z(t) = \bar{x}$ for all $t \in T$.

In this example, aggregate uncertainty is nonnegligible if P_Θ is not degenerate. Furthermore, the variability $\Lambda_i^P = 0$ for each i. Therefore, examples of this kind are not covered by Corollary 1 even though an incentive compatible mechanism yielding exact ex post efficiency exists. However, the construction of the mechanism in the presence of negligible aggregate uncertainty provides a clue to an approach that can be used in certain problems exhibiting non-negligible aggregate uncertainty. This approach appeared in an earlier version of this paper.

Suppose for a given information structure, we can find a set of distributions $\mathcal{P} = \{\pi_1, \pi_2, \ldots, \pi_m\}$ on Θ, with the property that (i) with high probability $P_\Theta(\cdot|t)$ is close to some $\pi \in \mathcal{P}$ and (ii) with high probability, the conditional distribution on Θ does not change much when an individual agent's type changes. We are treating the probability distributions in $\mathcal{P}$ as something like quasi-states and this property may be interpreted as negligible aggregate uncertainty with respect to quasi-states. This paper treats the problem of negligible aggregate uncertainty with respect to *actual* states and corresponds to the particular set of probability distributions $\{I_\theta\}_{\theta\in\Theta}$ on Θ. Learning which of the quasi-states is the "true" conditional distribution over Θ is all that one can hope for given the information structure. For each k, let $e(\pi_k)$ denote the PIE in which agent i has endowment w_i and utility function $v_i(\cdot|\pi_k)$ defined by

$$v_i(x_i|\pi_k) = \sum_{\theta\in\Theta} u_i(x_i, \theta)\pi_k(\theta).$$

Let $\zeta(\pi_k)$ be strictly individually rational and efficient in $e(\pi_k)$. Now define $z_i(t) = \zeta_i(\pi_k)$ if $P_\Theta(\cdot|t) = \pi_k$ for some $\pi_k \in \mathcal{P}$ and $z_i(t) = w_i$ otherwise. If we define variability in terms of quasi-states, then we can modify the $z(\cdot)$ so that the resulting mechanism is incentive compatible, ex post individually rational and approximately efficient if (i) informational size is small relative to variability and (ii) aggregate uncertainty with respect to quasi-states is small relative to variability. The proof is essentially the same as that of Corollary 1 but at a substantial cost in notational complexity.

In McLean and Postlewaite (2001), we present a different approach to the case of non-negligible aggregate uncertainty. In that paper, we begin with a definition of informational size that allows for $P(t)=0$ for some $t \in T$. Using this extended definition, we can prove that all agents have zero informational size if and only if the information structure exhibits a property that Postlewaite and Schmeidler (1986) called *nonexclusive information* (NEI). We then define variability in terms of the distributions $P_{T_{-i}}(\cdot|t_i)$ rather than the distributions $P_\Theta(\cdot|t_i)$. We can then prove versions of Theorem 1 and Corollary 1 without any reference to aggregate uncertainty under one more crucial assumption. We require that the correspondence mapping $\pi \in \Delta_\Theta$ to efficient allocations of the economy $e(\pi)$ admit certain Lipschitz continuous selections, where $e(\pi)$ denotes the PIE in which agent i has endowment w_i and utility function $v_i(\cdot|\pi)$ defined by

$$v_i(x_i|\pi)=\sum_{\theta\in\Theta} u_i(x_i,\theta)\pi(\theta).$$

Lipschitz selections play no role in the current work and, while the results are related, they are not nested.

2. When $P_\Theta(\cdot|t_i)\neq P_\Theta(\cdot|t_i')$, we can find punishments depending on i's announcement and the estimated state that gave i a strict incentive to truthfully announce his type. When $P_\Theta(\cdot|t_i)=P_\Theta(\cdot|t_i')$, we may still be able to construct more elaborate punishments that might provide agents with a strict incentive to truthfully reveal their types. These punishments are based on other conditional distributions associated with a measure $P\in\Delta_{\Theta\times T}$. For example, we could use $P_{T_{-i}}(\cdot|t_i)$, the conditional distribution on T_{-i} given t_i. If $P_{T_{-i}}(\cdot|t_i)\neq P_{T_{-i}}(\cdot|t_i')$ for each i and $t_i, t_i'\in T_i$, then we can find punishments $z_i(t_{-1},t_i)$ with the property that

$$\sum_{t_{-1}} z_i(t_{-i},t_i)P_{T_{-i}}(t_{-i}|t_i) > \sum_{t_{-i}} z_i(t_{-i},t_i')P_{T_{-i}}(t_{-i}|t_i) \quad \text{if} \quad t_i\neq t_i'.$$

These punishments peg the payoff of agent i on the complete vector of announced types, rather than simply on his announced t_i and some "estimate" of the state θ. We should note, however, that when the number of agents is large relative to the number of states, the vectors of punishments that depend on t_{-i} are commensurately larger than the vectors of punishments depending on θ. In other words, the mechanisms constructed in this way are somewhat more complicated than those constructed in this paper.

We could even use the conditionals $P_{\Theta\times T_{-i}}(\cdot|t_i)$ on $\Theta\times T_{-i}$ to construct punishments. Indeed, these are the best in the sense that it is possible for $P_{\Theta\times T_{-i}}(\cdot|t_i)\neq P_{\Theta\times T_{-i}}(\cdot|t_i')$ each i and $t_i,t_i'\in T_i$ even if $P_\Theta(\cdot|t_i)=P_\Theta(\cdot|t_i')$ for each i and $t_i,t_i'\in T_i$ and $P_{T_{-i}}(\cdot|t_i)=P_{T_{-i}}(\cdot|t_i')$ for each i and $t_i,t_i'\in T_i$. These issues are discussed more thoroughly in McLean and Postlewaite (2001).

7. RELATED LITERATURE

1. As mentioned in the introduction, our work is closely related to that of Cremer and McLean (1985, 1988). Those papers, and subsequent work by McAfee and Reny (1992), demonstrated how one can use correlation to obtain

full extraction of surplus in certain mechanism design problems. The key ingredient there is the assumption that the collection of conditional distributions $\{P_{T_{-i}}(\cdot|t_i)\}_{t_i\in T_i}$ is a linearly independent set for each i (where $P_{T_{-i}}(\cdot|t_i)$ is the conditional distribution on T_{-i} given t_i). Linear independence implies that the elements of the collection $\{P_{T_{-i}}(\cdot|t_i)\}_{t_i\in T_i}$ must be different, but they can be arbitrarily "close" and full extraction will be possible. In their quasilinear framework, Cremer and McLean use the full rank condition (or a weaker cone condition) to construct rewards and punishments $z_i(t_{-i}, t_i)$ with the following features:

$$\sum_{t_{-i}} z_i(t_{-i}, t_i) P_{T_{-i}}(t_{-i}|t_i) = 0$$

and

$$\sum_{t_{-i}} z_i(t_{-i}, t_i') P_{T_{-i}}(t_{-i}|t_i) < 0 \quad \text{if} \quad t_i \neq t_i'.$$

These rewards/punishments can then be used to ensure incentive compatibility.

In the present work, the collection $\{P_\Theta(\cdot|t_i)\}_{t_i\in T_i}$ need not be linearly independent and we can always find rewards and punishments $z_i(t_{-i}, t_i)$ satisfying the weaker property that

$$\sum_{\theta} z_i(\theta, t_i) P_\Theta(\theta|t_i) \geq \sum_{\theta} z_i(\theta, t_i') P_\Theta(\theta|t_i) \quad \text{if} \quad t_i \neq t_i'.$$

However, the "closeness" of the members of $\{P_\Theta(\cdot|t_i)\}_{t_i\in T_i}$ is an important issue. If the posteriors $\{P_\Theta(\cdot|t_i)\}_{t_i\in T_i}$ are all distinct, then the incentive compatibility inequalities are strict but the inequalities become weaker as the posteriors get closer. The *difference* in the expected reward from a truthful report and false report will be very small if the conditional posteriors are very close to each other. Our results require that informational size and aggregate uncertainty be small relative to the variation in these posteriors.

Weak incentives for truthful reporting are not a serious problem in the surplus extraction problem studied by Cremer and McLean since the rewards and punishments can be rescaled so that a false report results in a large negative expected payment. Of course, the punishments themselves may then become very large.

However, such rescaling is not possible in our framework for two reasons. First, we deal with pure exchange economies where the feasibility requirement limits the size of punishments. Second, we do not restrict attention to quasilinear preferences. Since agents may be risk averse, punishments and rewards that have small (or zero) expected value can have large negative welfare effects.

2. Gul and Postlewaite (1992) considered a model similar to that in this paper in which an economy with asymmetric information is replicated. They show that, when an economy is replicated sufficiently often in their framework, an allocation that is approximately Walrasian for the replica economy will be incentive compatible.

Our work differs from Gul-Postlewaite in several important ways. First, unlike this paper, Gul-Postlewaite dealt with replica economies. They did not formalize the notion of informational size (although they did discuss the idea informally in the context of the replication process considered there). An important part of our paper is the formalization of informational size, independent of agents' endowments and utility functions. This notion of informational size allows us to determine the circumstances in which asymmetry of information is important in general frameworks. While the informational size of agents decreases when an economy is replicated, the applicability of the concept is not limited to that case. Our theorem can be interpreted as providing limits on the informational rents due to private information when agents are informationally small. This implies that in situations in which there is a small number of agents with similar, but not identical information, informational size captures the degree to which the asymmetry of information leads to inefficiency.

In addition to the formalization of informational size, the model in the present paper treats an important class of economies excluded by Gul and Postlewaite. In the economies analyzed there, agents' utilities may depend on the state θ, but an individual agent's utility *cannot* be independent of his own type (i.e., his signal). This eliminates from consideration problems in which the uncertainty stems from characteristics of the object(s) being traded. If, for example, the only uncertainty pertains to the quantity of oil in a given tract to be traded, agents' utility naturally depends *only* on the state θ.[4]

Finally, Gul and Postlewaite demonstrate the existence of an incentive compatible, nearly Walrasian allocation for sufficiently large replica economies. In this paper, we show that a large class of allocations (including Walrasian allocations) can be approximated by incentive compatible allocations when agents are of sufficiently small informational size.

3. Our measure of informational size is motivated in part by the concept of nonexclusive information introduced in Postlewaite and Schmeidler (1986), which was shown to be a sufficient condition for the implementation of social choice correspondences satisfying Bayesian monotonicity. An economy with asymmetric information exhibits nonexclusive information if we can exclude any single agent's information and use only the information of the remaining agents to predict the economically relevant state of nature. Loosely speaking, our measure of informational size will be the "degree" to which an agent's information affects the prediction of the economically relevant state of nature, given other agents' information. The case of nonexclusive information is precisely the case in which each agent has zero informational size.

4. In a mechanism design framework, Al-Najjar and Smorodinsky (2000) study the circumstances under which an agent is pivotal in a mechanism in the sense that an agent can nontrivially affect the outcome of the mechanism through his reports. They provide conditions under which the proportion of agents who are pivotal must go to zero as the number of agents goes to infinity.[5] Our measure

[4] It should be noted that our model does not just allow for the case that agents' utility functions depend only on the state θ, but requires that they depend only on θ, and not also on their own signal.

[5] See Fudenberg, Levine, and Pesendorfer (1998) for an analysis of a similar problem in a game-theoretic rather than a mechanism design framework.

of an agent's informational size is related somewhat to the notion of pivoting presented in Al-Najjar and Smorodinski, but differs in several ways. Our notion of informational size measures (loosely speaking) the degree to which an agent is "pivotal" with respect to the conditional distribution on states. In the setup of Al-Najjar and Smorodinski, an agent is pivotal with respect to a particular mechanism. For example, in a voting model, we can compute the probability that an agent will affect the outcome under some voting rule, say majority rule. However, the probability of affecting the outcome might be quite different if the voting rule were unanimity rather than majority rule. Our definition of informational size depends only on the information structure, and is independent of any particular mechanism.

8. DISCUSSION

1. We were motivated in this paper by the question of how an agent's informational size would affect the degree to which efficient reallocation was possible. Our analysis depends on the construction of incentive compatible mechanisms that generate nearly ex post efficient allocations. We should emphasize that, while this provides a relatively clear understanding of the degree to which inefficiency will stem from informational asymmetries alone, it does not shed much light on how much inefficiency will result from asymmetric information *within a specific institutional setting*. The fact that an optimally designed mechanism will result in a nearly efficient outcome for a particular informational structure tells us little about how a specific institution, for example an anonymous market, will perform. We believe that it is important to identify those institutions that will do well, relative to the theoretical bounds we establish, in the face of uncertainty.[6]

2. Suppose that $(\{e(\theta)\}_{\theta\in\Theta}, \tilde{\theta}, \tilde{t}, P)$ is a PIE. If some agent is "informationally large," then our Corollary 1 will generally not be useful in determining whether or not an allocation satisfying the desired efficiency, individual rationality, and incentive properties will exist for this PIE. However, the following example suggests a way to improve the theorem to encompass certain problems with informationally large agents. Consider the following simple replica example. There are two equally likely states of nature, θ_1 and θ_2. In the nth economy, there are n agents, each of whom receives a noisy signal of the state. That is, each agent will receive a signal s_1 or s_2, with $P(s_i|\theta_i) = q$ where $.5 \leq q \leq 1$. Agents' signals are i.i.d. conditional on the state. When n is large, the economy will exhibit negligible aggregate uncertainty and agents will be informationally small, both consequences of the law of large numbers. We could then use the vector of announced types t to estimate the probability distribution over Θ, and choose an allocation that is approximately optimal for the most likely state; this is exactly what we did in Theorem 2.

[6] Along these lines, Krasa and Shafer (1998) analyze a related notion of informational smallness in a Walrasian market.

Suppose now that we alter this example by letting agent 1 receive a *perfect* signal of the correct state, while all other agents continue to receive the noisy signal. In this case, $P_\Theta(\cdot|t)$ will be either $(1,0)$ or $(0,1)$, depending only on agents 1's signal, since his is the only non-noisy signal. It is clear that with this modification, our Theorem 2 no longer applies. Aggregate uncertainty will still be negligible but the assumption that agents are informationally small no longer holds since agent 1's announcement alone determines whether the conditional distribution on Θ is $(1,0)$ or $(0,1)$. However, it is important to note that this *does not* preclude our finding an incentive compatible allocation that is individually rational and ex post nearly efficient. A mediator could simply ignore agent 1's announcement and estimate the distribution on Θ using only the other agent's announcements. When this distribution puts probability close to 1 on some state θ, the allocation for that state would be assigned. In this way, we can construct an incentive compatible allocation that is individually rational and nearly ex post efficient despite the fact that agent one is not informationally small.

This example suggests a way to extend our results. Our proof uses the Bayesian posterior given the agents' announcements as an estimate of the state of nature. The above example illustrates how one could find a mechanism with the desired properties using a subset of the agents' announcements. More generally, one could estimate the state of nature using a general function of the agents' announcements. This is a topic for further research.

3. We assumed that both Θ and T were finite. In general, it should be possible to extend the results to the case in which Θ is a compact subset of R^l. If the utility functions are uniformly continuous on $\Re_+^\ell \times \Theta$, one could take a finite partition of Θ and use agents' announcements to estimate the most likely cell in the partition. For each estimated cell, one could prescribe a given allocation for that cell, with appropriate punishments to induce truthful announcements. There would be an additional efficiency loss in that the allocation so constructed would be constant across any cell in the partition, but this utility loss can be made arbitrarily small by constructing increasingly finer partitions.

The situation with respect to T is much more delicate, however. In our construction, the ability to give any agent an incentive to announce his type truthfully depends on the variation in the distributions $P_\Theta(\cdot|t_i)$ and $P_\Theta(\cdot|t_i')$ on Θ, conditional on different types t_i and t_i'. If the T_i are intervals and the conditionals $P_\Theta(\cdot|t_i)$ are continuous in t_i, then $P_\Theta(\cdot|t_i)$ and $(P_\Theta(\cdot|t_i')$ will be close when t_i and t_i' are close. Hence, the required "balance" between informational smallness, aggregate uncertainty and variability in the conditional distributions is more complicated. This is also a problem for further research.

4. There is a possible generalization of our results related to the previous point. Consider a PIE allocation that satisfies the assumptions of Theorem 1. Now alter the PIE in the following way. Choose an agent i and some type $\hat{t}_i$ for that agent and suppose that his signal $\hat{t}_i$ is replaced by two signals, t_i', or t_i''.

Furthermore, suppose that the new information structure $\widehat{P}$ is defined as

$$\widehat{P}(\theta, t_{-i}, t_i') = \widehat{P}(\theta, t_{-i}, t_i'') = \frac{P(\theta, t_{-i}, \hat{t}_i)}{2} \qquad \text{for all } \theta \text{ and } t_{-i},$$

$$\widehat{P}(\theta, t_{-i}, t_i) = P(\theta, t_{-i}, t_i) \qquad \text{for all } \theta \text{ and } t_{-i} \text{ and all } t_i \neq \hat{t}_i.$$

That is, we have taken the original information structure and altered it by separating one signal for agent i into two different signals in a way that has no effect on the information conveyed by those signals. In particular, $\widehat{P}_\Theta(\cdot|t_{-i}, t_i') = \widehat{P}_\Theta(\cdot|t_{-i}, t_i'') = P_\Theta(\cdot|t_{-i}, \hat{t}_i)$ for all t_{-i} and $\widehat{P}_\Theta(\cdot|t_i') = \widehat{P}_\Theta(\cdot|t_i'') = P_\Theta(\cdot|\hat{t}_i)$.

One can think of this as agent i flipping a coin after he receives signal $\hat{t}_i$ and labeling the outcomes $t_i' = (t_i$ and heads) and $t_i'' = (t_i$ and tails). For this altered PIE, the assumptions of Theorem 1 will generally not hold since $\min_i \Lambda_i^{\widehat{P}} = 0$. Clearly, however, this alteration should not affect what outcomes can be approximated. We can, in fact, still approximate an allocation by treating the two signals t_i' and t_i'' as a single signal, $\hat{t}_i$. The crucial feature of this simple splitting example is the fact that $\widehat{P}_\Theta(\cdot|t_{-i}, t_i') = \widehat{P}_\Theta(\cdot|t_{-i}, t_i'')$ for all t_{-i}. Whenever this is true, we can collapse types into equivalence classes and treat each class as a single type. With appropriate modifications of the definitions of informational size and aggregate uncertainty, we would expect to be able to prove a result analogous to our Theorem 1 when each agent's type set can be partitioned so that, within each element of the partition, the types are sufficiently similar.[7]

5. We used the revelation principle to analyze the constraints imposed by incentive compatibility on the set of incentive compatible utility vectors. As is often the case with revelation games, there are additional equilibria in our mechanism different from the truthful reporting equilibrium. We do not view this as problematic since we do not propose the mechanism as one to be used in practice; we use the revelation mechanism simply to determine the degree to which incentive constraints limit the utilities that can be obtained. The issue of multiplicity of equilibria in settings such as ours has been addressed, however. Postlewaite and Schmeidler (1986) and Jackson (1991) demonstrate how revelation mechanisms can be augmented so as to eliminate nontruthful equilibria in a large set of problems. While we do not do so here, there is reason to expect that those techniques could be similarly applied to our setting.

Dept. of Economics, New Jersey Hall, Rutgers University, New Brunswick, NJ 08901-1248, U.S.A; rpmclean@rci.rutgers.edu

and

Dept. of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, U.S.A.; apostlew@econ.sas.upenn.edu.

Manuscript received October, 1999; final revision received April, 2001.

[7] We thank Ichiro Obara for making this point.

APPENDIX

A.1. Preliminary Definitions and Lemmas

Throughout this appendix, we will assume that $|\Theta| = m$. Suppose that $\mathscr{A} = \{x(\theta)\}_{\theta\in\Theta}$ is a collection of CIE allocations such that, for each $\theta \in \Theta$, the allocation $(x_i(\theta))_{i\in N}$ is a CIE allocation for $e(\theta)$ with $x_i(\theta) \neq 0$ for all $\theta \in \Theta$ and for all i. For each $\eta \geq 0$, let

$$c(\eta, \mathscr{A}) = \min_i \min_\theta \{u_i(x_i(\theta); \theta) - u_i(\beta_i(\theta, \eta, \mathscr{A})x_i(\theta); \theta)\}$$

where

$$\beta_i(\theta, \eta, \mathscr{A}) = \min\{\beta | 1/2 \leq \beta \leq 1, u_i(x_i(\theta); \theta) - u_i(\beta x_i(\theta); \theta) \leq \eta\}.$$

Since $x_i(\theta) \neq 0$ for each i and each θ, it follows from the monotonicity assumption that $c(0, \mathscr{A}) = 0$ and that $c(\eta, \mathscr{A}) > 0$ whenever $\eta > 0$.

Finally, recall that

$$\Lambda_i^P = \min_{t_i\in T_i} \min_{t_i'\in T_i\backslash t_i} d(P_\Theta(\cdot|t_i), P_\Theta(\cdot|t_i'))^2$$

where

$$d(\alpha, \beta) = \left\| \frac{\alpha}{\|\alpha\|_2} - \frac{\beta}{\|\beta\|_2} \right\|_2$$

for each $\alpha, \beta \in \Delta_\Theta$ and $\|\cdot\|_2$ denotes the 2-norm.

LEMMA A.1: *Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and suppose that $P \in \Delta_{\Theta\times T}$ with conditionals $P_\Theta(\cdot|t_i) \in \Delta_\Theta$ for all i and $t_i \in T_i$. Furthermore, suppose that $\mathscr{A} = \{x(\theta)\}_{\theta\in\Theta}$ is a collection of CIE allocations such that, for each $\theta \in \Theta$, the allocation $\{x_i(\theta)\}_{i\in N}$ is a CIE allocation for $e(\theta)$ with $x_i(\theta) \neq 0$ for all θ and for all i. For each $\eta \geq 0$, there exists a collection $\{\{z_i(\theta, t_i)\}_{(t_i,\theta)\in T_i\times\Theta}\}_{i\in N}$ satisfying*:

(i) *$z_i(\theta, t_i) \in \Re_+^\ell$ and $\sum_{i\in N}(z_i(\theta, t_i) - w_i) \leq 0$ for all $t_i \in T_i$ and all $\theta \in \Theta$;*

(ii) *$u_i(x_i(\theta); \theta) \geq u_i(z_i(\theta, t_i); \theta) \geq u_i(x_i(\theta); \theta) - \eta$ for all $t_i \in T_i$ and all $\theta \in \Theta$;*

(iii) *for each $t_i, t_i' \in T_i$ with $t_i \neq t_i'$,*

$$\sum_\theta [u_i(z_i(\theta, t_i); \theta) - u_i(z_i(\theta, t_i'); \theta)] P_\Theta(\theta|t_i) \geq \frac{c(\eta, \mathscr{A})}{2\sqrt{m}} \min_i \Lambda_i^P.$$

PROOF: Suppose that $P \in \Delta_{\Theta\times T}$ with conditionals $P_\Theta(\cdot|t_i) \in \Delta_\Theta$ for all i and $t_i \in T_i$. Next, define

$$\alpha_i(\theta, t_i) = \frac{P_\Theta(\theta|t_i)}{\|P_\Theta(\cdot|t_i)\|_2}$$

for each $\theta \in \Theta$. Hence,

$$\Lambda_i^P = \min_{t_i\in T_i} \min_{t_i'\in T_i\backslash t_i} \|\alpha_i(\cdot, t_i) - \alpha_i(\cdot, t_i')\|_2^2.$$

Let $\mathscr{A} = \{x(\theta)\}_{\theta\in\Theta}$ be a collection of CIE allocations with $x_i(\theta) \neq 0$ for all θ and for all i. If $\eta = 0$, then $c(\eta, \mathscr{A}) = 0$ and the result is trivial (let $z_i(\theta, t_i) = x_i(\theta)$). So suppose that $\eta > 0$. For each i, t_i, and θ, there exists a number $\tau_i(\theta, t_i) \geq 0$ such that

$$u_i((1 + \tau_i(\theta, t_i))\beta_i(\theta, \eta, \mathscr{A})x_i(\theta); \theta) - u_i(\beta_i(\theta, \eta, \mathscr{A})x_i(\theta); \theta) = c(\eta, \mathscr{A})\alpha_i(\theta, t_i).$$

(This is possible because $0 \le c(\eta, \mathscr{A})\alpha_i(\theta, t_i) \le c(\eta, \mathscr{A})$ and $\beta_i(\theta, \eta, \mathscr{A})x_i(\theta) \neq 0$.) Furthermore, $(1+\tau_i(\theta, t_i))\beta_i(\theta, \eta, \mathscr{A}) \le 1$. (If $(1+\tau_i(\theta, t_i))\beta_i(\theta, \eta, \mathscr{A}) > 1$, then monotonicity implies that

$$\begin{aligned}
&u_i((1+\tau_i(\theta, t_i))\beta_i(\theta, \eta, \mathscr{A})x_i(\theta); \theta) - u_i(\beta_i(\theta, \eta, \mathscr{A})x_i(\theta); \theta) \\
&\quad > u_i(x_i(\theta); \theta) - u_i(\beta_i(\theta, \eta, \mathscr{A})x_i(\theta); \theta) \\
&\quad \ge c(\eta, \mathscr{A}) \\
&\quad \ge c(\eta, \mathscr{A})\alpha_i(\theta, t_i),
\end{aligned}$$

a contradiction.) Defining

$$z_i(\theta, t_i) = (1+\tau_i(\theta, t_i))\beta_i(\theta, \eta, \mathscr{A})x_i(\theta),$$

it follows that the collections $\{z_i(\theta, t_i)\}_{t_i, \theta}$ satisfy

$$z_i(\theta, t_i) \in \Re^\ell_+ \qquad \text{and} \qquad \sum_{i\in N}(z_i(\theta, t_i) - w_i) \le 0$$

and part (i) is satisfied. From the definitions of $\beta_i(\theta, \eta, \mathscr{A})$ and $z_i(\theta, t_i)$ and the fact that

$$1 \ge (1+\tau_i(\theta, t_i))\beta_i(\theta, \eta, \mathscr{A}) \ge \beta_i(\theta, \eta, \mathscr{A}),$$

we conclude that

$$u_i(x_i(\theta); \theta) \ge u_i(z_i(\theta, t_i); \theta) \ge u_i(\beta_i(\theta, \eta, \mathscr{A})x_i(\theta); \theta) \ge u_i(x_i(\theta); \theta) - \eta$$

and part (ii) is satisfied. Finally, part (iii) follows from the observation that

$$\begin{aligned}
&\sum_\theta [u_i(z_i(\theta, t_i); \theta) - u_i(z_i(\theta, t_i'); \theta)]P_\Theta(\theta|t_i) \\
&\quad = \sum_\theta [c(\eta, \mathscr{A})\alpha_i(\theta, t_i) - c(\eta, \mathscr{A})\alpha_i(\theta, t_i')]P_\Theta(\theta|t_i) \\
&\quad = c(\eta, \mathscr{A}) \sum_\theta [\alpha_i(\theta, t_i) - \alpha_i(\theta, t_i')]P_\Theta(\theta|t_i) \\
&\quad = c(\eta, \mathscr{A})\|P_\Theta(\cdot|t_i)\|_2 \sum_\theta [\alpha_i(\theta, t_i) - \alpha_i(\theta, t_i')]\alpha_i(\theta, t_i) \\
&\quad = \frac{c(\eta, \mathscr{A})\|P_\Theta(\cdot|t_i)\|_2}{2}\|\alpha_i(\cdot, t_i) - \alpha_i(\cdot, t_i')\|^2 \\
&\quad \ge \frac{c(\eta, \mathscr{A})}{2\sqrt{m}}\Lambda_i^P.
\end{aligned}$$

A.2. Proof of Theorem 1

Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and suppose that $P \in \Delta_{\Theta\times T}$ with conditionals $P_\Theta(\cdot|t_i) \in \Delta_\Theta$ for all i and $t_i \in T_i$. Furthermore, suppose that $\mathscr{A} = \{x(\theta)\}_{\theta\in\Theta}$ is a collection of CIE allocations such that, for each $\theta \in \Theta$, the allocation $\{x_i(\theta)\}_{i\in N}$ is a CIE allocation for $e(\theta)$ with $x_i(\theta) \neq 0$ for all θ and for all i. Choose $\varepsilon > 0$. Let

$$K_1 = \max_\theta \max_i \left\{u_i\left(\sum_{j\in N} w_j; \theta\right)\right\}$$

and choose δ so that

$$0 < \delta < \min\left\{\frac{c(\varepsilon, \mathscr{A})}{20\sqrt{m}K_1}, \frac{\varepsilon}{2}, \frac{1}{3}\right\}.$$

(The monotonicity assumption implies that $K_i > 0$ since $\sum_{j\in N} w_j \neq 0$.) Finally, define $\hat{\mu}^P = \max_i \mu_i^P$, $\hat{\nu}^P = \max_i \nu_i^P$, and $\Lambda^P = \min_i \Lambda_i^P$, and suppose that

$$\hat{\nu}^P \leq \delta \Lambda^P,$$
$$\hat{\mu}^P \leq \delta \Lambda^P.$$

If $\Lambda^P = 0$, then $\hat{\nu}^P = 0$ and $\hat{\mu}^P = 0$. Since $\hat{\nu}^P = 0$ and $\text{Prob}\{\tilde{t} = t\} > 0$ for all $t \in T$, it follows that $\tilde{\theta}$ and $\tilde{t}$ are independent. Hence, $\hat{\mu}^P = 0$ implies that there exists $\hat{\theta} \in \Theta$ such that $P_\Theta(\cdot|t) = I_{\hat{\theta}}$ for all $t \in T$. Now choose an efficient, individually rational CIE allocation $(x_i(\theta))_{i\in N}$ for $e(\hat{\theta})$ and define $z_i(t) = x_i(\hat{\theta})$ for all i and all $t \in T$. The PIE allocation $z(\cdot)$ is ex post efficient, ex post individually rational, and incentive compatible.

Now suppose that $\Lambda^P > 0$. For each k, let

$$A_k = \{t \in T \mid \|P_\Theta(\cdot|t) - I_{\theta_k}\| \leq \hat{\mu}^P\}$$

and let

$$A_0 = T \Big\backslash \Big[\bigcup_k A_k\Big].$$

Since $\Lambda^P \leq 2$, it follows that

$$\hat{\mu}^P \leq \delta \Lambda^P < \frac{1}{3}\Lambda^P \leq \frac{2}{3}$$

and the collection $\Pi = \{A_0, A_1, \ldots, A_m\}$ is a partition of T.

Applying Lemma A.1, there exists a collection $\{\{z_i(\theta, t_i)\}_{(\theta,t_i)\in\Theta\times T_i}\}_{i\in N}$ satisfying:

(i) $z_i(\theta, t_i) \in \Re_+^\ell$ and $\sum_{i\in N}(z_i(\theta, t_i) - w_i) \leq 0$ for all $t_i \in T_i$ and all $\theta \in \Theta$;

(ii) $u_i(x_i(\theta); \theta) \geq u_i(z_i(\theta, t_i); \theta) \geq u_i(x_i(\theta); \theta) - \varepsilon$ for all $t_i \in T_i$ and all $\theta \in \Theta$;

(iii) for each $t_i, t_i' \in T_i$,

$$\sum_\theta [u_i(z_i(\theta, t_i); \theta) - u_i(z_i(\theta, t_i'); \theta)] P_\Theta(\theta|t_i) \geq \frac{c(\varepsilon, \mathscr{A})}{2\sqrt{m}} \Lambda^P.$$

Next, let $z(\cdot)$ be the PIE allocation for $(\{e(\theta)\}_{\theta\in\Theta}, \tilde{\theta}, \tilde{t}, P)$ defined as[8]

$$\begin{aligned} z_i(t) &= z_i(\theta_k, t_i) \quad \text{if } t \in A_k, \\ &= 0 \qquad\qquad\quad \text{if } t \in A_0. \end{aligned}$$

Before proving that the PIE allocation $z(\cdot)$ is incentive compatible, we first prove two claims.

CLAIM 1: *For each i and each $t_i \in T_i$,*

$$\sum_k |P_\Theta(\theta_k|t_i) - \text{Prob}\{\tilde{t} \in A_k | \tilde{t}_i = t_i\}| \leq 2\hat{\mu}^P.$$

PROOF OF CLAIM 1: First, note that

$$P_\Theta(\theta_k|t_i) = \sum_{\ell=1}^m \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_\ell}} P_\Theta(\theta_k|t_{-i}, t_i) P(t_{-i}|t_i) + \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_0}} P(\theta_k, t_{-i}|t_i)$$

[8] If for each θ the CIE allocation $\{x_i(\theta)\}_{i\in N}$ is strictly individually rational for $e(\theta)$, then a simple modification of the proof would allow us to define $z_i(t) = w_i$ when $t \in A_0$.

and

$$\text{Prob}\{\tilde{t} \in A_k | \tilde{t}_i = t_i\} = \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_k}} P(t_{-i}|t_i).$$

Therefore,

$$\begin{aligned}
&P_\Theta(\theta_k|t_i) - \text{Prob}\{\tilde{t} \in A_k | \tilde{t}_i = t_i\} \\
&\quad = \left[\sum_{\ell=1}^m \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_\ell}} P_\Theta(\theta_k|t_{-i}, t_i) P(t_{-i}|t_i)\right] - \left[\sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_k}} P(t_{-i}|t_i)\right] + \left[\sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_0}} P(\theta_k, t_{-i}|t_i)\right] \\
&\quad = \left[\sum_{\ell=1}^m \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_\ell}} P_\Theta(\theta_k|t_{-i}, t_i) P(t_{-i}|t_i)\right] - \left[\sum_{\ell=1}^m I_{\theta_\ell}(\theta_k) \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_\ell}} P(t_{-i}|t_i)\right] \\
&\qquad + \left[\sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_0}} P(\theta_k, t_{-i}|t_i)\right] \\
&\quad = \left[\sum_{\ell=1}^m \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_\ell}} [P_\Theta(\theta_k|t_{-i}, t_i) - I_{\theta_\ell}(\theta_k)] P(t_{-i}|t_i)\right] + \left[\sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_0}} P(\theta_k, t_{-i}|t_i)\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
&\sum_{k=1}^m |P_\Theta(\theta_k|t_i) - \text{Prob}\{\tilde{t} \in A_k | \tilde{t}_i = t_i\}| \\
&\qquad \le \sum_{\ell=1}^m \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_\ell}} \sum_{k=1}^m |P(\theta_k|t_{-i}, t_i) - I_{\theta_\ell}(\theta_k)| P(t_{-i}|t_i) + \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_0}} P(t_{-i}|t_i) \\
&\qquad \le \hat{\mu}^P \sum_{\ell=1}^m \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_\ell}} P(t_{-i}|t_i) + \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_0}} P(t_{-i}|t_i) \\
&\qquad \le \hat{\mu}^P + \mu_i^P \\
&\qquad \le 2\hat{\mu}^P
\end{aligned}$$

and the proof of Claim 1 is complete.

CLAIM 2: *For each i, t_i, and t_i',*

$$\sum_{k=1}^m \sum_{\substack{t_{-i} \\ (t_{-i}, t_i) \in A_k \\ (t_{-i}, t_i') \notin A_k \cup A_0}} P(t_{-i}|t_i) \le \hat{\nu}^P.$$

PROOF OF CLAIM 2: Choose $t_i, t_i' \in T_i$ and define

$$\Psi = \bigcup_{\ell \in J_m} \{t_{-i} \in T_{-i} | (t_{-i}, t_i) \in A_\ell \text{ and } (t_{-i}, t_i') \notin A_\ell \cup A_0\}$$

and

$$\Phi = \{t_{-i} \in T_{-i} \mid \|P_\Theta(\cdot|t_{-i}, t_i) - P_\Theta(\cdot|t_{-i}, t_i')\| > \hat{\nu}^P\}.$$

Since

$$\text{Prob}\{\tilde{t} \in \Psi | \tilde{t}_i = t_i\} = \sum_{k=1}^{m} \sum_{\substack{t_{-i} \\ (t_{-i}, t_i) \in A_k \\ (t_{-i}, t_i') \notin A_k \cup A_0}} P(t_{-i}|t_i)$$

and

$$\text{Prob}\{\tilde{t} \in \Phi | \tilde{t}_i = t_i\} \leq \hat{\nu}^P,$$

it suffices to prove that $\Psi \subseteq \Phi$. Suppose that $t_{-i} \in \Psi$ but $t_{-i} \notin \Phi$. Then there exist $\ell, k \in J_m$ with $k \neq \ell$ such that $(t_{-i}, t_i) \in A_\ell$ and $(t_{-i}, t_i') \in A_k$ and $\|P_\Theta(\cdot|t_{-i}, t_i) - P_\Theta(\cdot|t_{-i} t_i')\| \leq \hat{\nu}^P$. Since $\Lambda^P \leq 2$, it follows that

$$\hat{\mu}^P \leq \delta \Lambda^P < \frac{1}{3} \Lambda^P \leq \frac{2}{3}$$

and that

$$\hat{\nu}^P \leq \delta \Lambda^P < \frac{1}{3} \Lambda^P \leq \frac{2}{3}.$$

Therefore,

$$\begin{aligned} \|I_{\theta_\ell} - I_{\theta_k}\| &\leq \|P_\Theta(\cdot|t_{-i}, t_i) - I_{\theta_\ell}\| + \|P_\Theta(\cdot|t_{-i}, t_i) - P_\Theta(\cdot|t_{-i}, t_i')\| \\ &\quad + \|P_\Theta(\cdot|t_{-i}, t_i') - I_{\theta_k}\| \\ &\leq \hat{\mu}^P + \hat{\nu}^P + \hat{\mu}^P \\ &< 3\frac{2}{3} \\ &= 2, \end{aligned}$$

an impossibility. This completes the proof of Claim 2.

Next, we observe that

$$u_i(z_i(t); \theta) \leq K_1$$

for all $t \in T$ and all $\theta \in \Theta$ since

$$u_i(z_i(t); \theta_k) = u_i(z_i(\theta_k, t_i); \theta_k) \leq u_i(x_i(\theta); \theta) \leq K_1$$

if $t \in A_k$ and

$$u_i(z_i(t); \theta_k) = u_i(0; \theta_k) \leq K_1,$$

$t \in A_0$.

To prove incentive compatibility, note that

$$\begin{aligned} &\sum_\theta \sum_{t_{-i}} [u_i(z_i(t_{-i}, t_i); \theta) - u_i(z_i(t_{-i}, t_i'); \theta)] P(\theta, t_{-i}|t_i) \\ &\quad = \sum_{\substack{t_{-i} \\ :(t_{-i}, t_i) \in A_0}} \sum_\theta [u_i(z_i(t_{-i}, t_i); \theta) - u_i(z_i(t_{-i}, t_i'); \theta)] P(\theta, t_{-i}|t_i) \end{aligned}$$

$$
\begin{aligned}
&+\sum_{k=1}^{m}\sum_{\substack{t_{-i}\\:(t_{-i},t_i)\in A_k}}\sum_{\theta}[u_i(z_i(\theta_k,t_i);\theta)-u_i(z_i(t_{-i},t_i');\theta)]P(\theta|t_{-i},t_i)P(t_{-i}|t_i)\\
&\geq -2K_1\mu_i^P+\sum_{k=1}^{m}\sum_{\substack{t_{-i}\\:(t_{-i},t_i)\in A_k}}\sum_{\theta}[u_i(z_i(\theta_k;t_i);\theta)-u_i(z_i(t_{-i},t_i');\theta)]P(\theta|t_{-i},t_i)P(t_{-i}|t_i)\\
&\geq -2K_1[\mu_i^P+\hat{\mu}^P]+\sum_{k=1}^{m}\sum_{\substack{t_{-i}\\:(t_{-i},t_i)\in A_k}}[u_i(z_i(\theta_k;t_i);\theta_k)-u_i(z_i(t_{-i},t_i');\theta_k)]P(t_{-i}|t_i)\\
&= -2K_1[\mu_i^P+\hat{\mu}^P]+\sum_{k=1}^{m}[u_i(z_i(\theta_k;t_i);\theta_k)-u_i(z_i(\theta_k;t_i');\theta_k)]\left[\sum_{\substack{t_{-i}\\:(t_{-i},t_i)\in A_k}}P(t_{-i}|t_i)\right]\\
&\quad+\sum_{k=1}^{m}\sum_{\substack{t_{-i}\\:(t_{-i},t_i)\in A_k}}[u_i(z_i(\theta_k;t_i');\theta_k)-u_i(z_i(t_{-i},t_i');\theta_k)]P(t_{-i}|t_i)\\
&\geq -2K_1[\mu_i^P+\hat{\mu}^P+2\hat{\mu}^P]+\sum_{k=1}^{m}[u_i(z_i(\theta_k;t_i);\theta_k)-u_i(z_i(\theta_k;t_i');\theta_k)]P_\Theta(\theta_k|t_i)\\
&\quad+\sum_{k=1}^{m}\sum_{\substack{t_{-i}\\:(t_{-i},t_i)\in A_k\\(t_{-i},t_i')\notin A_k\cup A_0}}[u_i(z_i(\theta_k;t_i');\theta_k)-u_i(z_i(t_{-i},t_i');\theta_k)]P(t_{-i}|t_i)\\
&\qquad\qquad\text{(applying Claim 1)}\\
&\geq \frac{c(\varepsilon,\mathscr{A})}{2\sqrt{m}}\Lambda^P-2K_1[4\hat{\mu}^P+\hat{\nu}^P]\quad\text{(applying Claim 2)}\\
&\geq \frac{c(\varepsilon,\mathscr{A})}{2\sqrt{m}}\Lambda^P-2K_1\left[5\frac{c(\varepsilon,\mathscr{A})}{20\sqrt{m}K_1}\Lambda^P\right]\\
&=0.
\end{aligned}
$$

To complete the proof of Theorem 1, we must show that $z(\cdot)$ satisfies conditions (i), (ii), and (iii) in the statement of the theorem. To prove (i), note that $\text{Prob}\{\tilde{t}\in A_0|\tilde{t}_i=t_i\}\leq\hat{\mu}^P$ for each i and t_i. Hence,

$$\text{Prob}\{\tilde{t}\in A_0\}=\sum_{t_i\in T_i}\text{Prob}\{\tilde{t}\in A_0|\tilde{t}_i=t_i\}P(t_i)\leq\hat{\mu}^P\leq\delta\Lambda^P\leq\frac{\varepsilon}{2}\Lambda^P\leq\varepsilon$$

from which we conclude that

$$\text{Prob}\left\{\tilde{t}\in\bigcup_{k=1}^{m}A_k\right\}=1-\text{Prob}\{\tilde{t}\in A_0\}\geq 1-\varepsilon.$$

To prove (ii), suppose that $t\in A_k$. Since

$$[1-P_\Theta(\theta_k|t)]+\sum_{\ell\neq k}P_\Theta(\theta_\ell|t)=\|P_\Theta(\cdot|t)-I_{\theta_k}\|\leq\hat{\mu}^P\leq\delta\Lambda^P\leq\frac{\varepsilon}{2}\Lambda^P\leq\varepsilon,$$

it follows that

$$1-\varepsilon\leq P_\Theta(\theta_k|t).$$

Finally, (iii) is satisfied since the construction of $z(\cdot)$ implies that for all $i \in N$,

$$u_i(x_i(\theta_k); \theta_k) \geq u_i(z_i(t); \theta_k) \geq u_i(x_i(\theta_k); \theta_k) - \varepsilon$$

whenever $t \in A_k$.

A.3. Proof of Corollary 1

Let $\{e(\theta)\}_{\theta \in \Theta}$ be a collection of CIE's and suppose that $\mathscr{A} = \{x(\theta)\}_{\theta \in \Theta}$ is a collection where, for each θ, $x(\theta)$ is a Pareto efficient, strictly individually rational CIE allocation for the CIE $e(\theta)$. Let

$$K_1 = \max_\theta \max_i \left\{ u_i\left(\sum_{j \in N} w_j; \theta \right) \right\}$$

and let

$$K_2(\mathscr{A}) = \min_i \min_\theta [u_i(x_i(\theta); \theta) - u_i(w_i; \theta)].$$

Since each $x(\theta)$ is strictly individually rational for the CIE $e(\theta)$, it follows that $K_2(\mathscr{A}) > 0$. Choose $\varepsilon > 0$ and choose $\hat{\varepsilon}$ so that

$$0 < \hat{\varepsilon} < \min\left\{ \frac{K_2(\mathscr{A})}{4K_1 + 1}, \frac{\varepsilon}{4K_1 + 1} \right\}.$$

Applying Theorem 1, there exists a $\hat{\delta} > 0$ such that, whenever $P \in \Delta_{\Theta \times T}$ and

$$\max_i \hat{\mu}_i^P \leq \hat{\delta} \min_i \Lambda_i^P,$$
$$\max_i \hat{\nu}_i^P \leq \hat{\delta} \min_i \Lambda_i^P,$$

there exists an incentive compatible PIE allocation $z(\cdot)$ for the PIE $(\{e(\theta)\}_{\theta \in \Theta}, \tilde{\theta}, \tilde{t}, P)$ and a collection $A_1, \dots, A_m$ of disjoint subsets of T such that $\text{Prob}\{\tilde{t} \in \bigcup_{k=1}^m A_k\} \geq 1 - \hat{\varepsilon}$ and for all $k = 1, \dots, m$ and all $t \in A_k$:

(i) $\text{Prob}\{\tilde{\theta} = \theta_k | \tilde{t} = t\} \geq 1 - \hat{\varepsilon}$;
(ii) for all $i \in N$,

$$u_i(x_i(\theta_k); \theta_k) \geq u_i(z_i(t); \theta_k) \geq u_i(x_i(\theta_k); \theta_k) - \hat{\varepsilon}.$$

If $t \in A_k$ for some $k \geq 1$, then $P_\Theta(\theta_k | t) \geq 1 - \hat{\varepsilon}$ implies that

$$\|P_\Theta(\cdot | t) - I_{\theta_k}\| = [1 - P_\Theta(\theta_k | t)] + \sum_{\ell \neq k} P_\Theta(\theta_\ell | t) \leq 2\hat{\varepsilon}.$$

To prove XIR, suppose that $t \in A_k$ and note that

$$\begin{aligned}
\sum_\theta [u_i(z_i(t); \theta) - u_i(w_i; \theta)] P_\Theta(\theta | t) &= \sum_\ell [u_i(z_i(t); \theta_\ell) - u_i(w_i; \theta_\ell)] P_\Theta(\theta_\ell | t) \\
&\geq u_i(z_i(t); \theta_k) - u_i(w_i; \theta_k) - (2K_1)(2\hat{\varepsilon}) \\
&\geq u_i(x_i(\theta_k); \theta_k) - u_i(w_i; \theta_k) - \hat{\varepsilon} - 4K_1\hat{\varepsilon} \\
&\geq K_2(\mathscr{A}) - (4K_1 + 1)\hat{\varepsilon} \\
&> 0.
\end{aligned}$$

Hence, $z(\cdot)$ satisfies XIR.

To show that $z(\cdot)$ satisfies $X_\varepsilon E$, let $E = \bigcup_{k=1}^m A_k$ and note that

$$\text{Prob}\{\tilde{t} \in E\} = \text{Prob}\left\{\tilde{t} \in \bigcup_{k=1}^m A_k\right\} \geq 1 - \hat{\varepsilon} \geq 1 - \frac{\varepsilon}{4K_1 + 1} \geq 1 - \varepsilon.$$

Now suppose that $y(\cdot)$ is a feasible PIE allocation and that

$$\sum_\theta [u_i(y_i(t); \theta) - u_i(z_i(t); \theta)] P(\theta|t) > \varepsilon$$

for each $i \in N$. Since $y(\cdot)$ is a feasible PIE allocation, it follows that

$$u_i(y_i(t); \theta) \leq K_1$$

for all i, t, and θ. If $t \in A_k$ for some $k \in J_m$, then for each $i \in N$, it follows that

$$\begin{aligned}
\varepsilon &< \sum_\theta [u_i(y_i(t); \theta) - u_i(z_i(t); \theta)] P(\theta|t) \\
&\leq (2K_1)(2\hat{\varepsilon}) + u_i(y_i(t); \theta_k) - u_i(z_i(t); \theta_k) \\
&= 4K_1\hat{\varepsilon} + [u_i(y_i(t); \theta_k) - u_i(x_i(\theta_k); \theta_k)] \\
&\quad + [u_i(x_i(\theta_k); \theta_k) - u_i(z_i(t); \theta_k)] \\
&\leq 4K_1\hat{\varepsilon} + [u_i(y_i(t); \theta_k) - u_i(x_i(\theta_k); \theta_k)] + \hat{\varepsilon}.
\end{aligned}$$

Therefore,

$$0 < \varepsilon - 4K_1\hat{\varepsilon} - \hat{\varepsilon} < [u_i(y_i(t); \theta_k) - u_i(x_i(\theta_k); \theta_k)]$$

for each i, contradicting the assumption that $\{x_i(\theta_k)\}_{i \in N}$ is Pareto optimal in $e(\theta_k)$. Therefore, $t \notin E = \bigcup_{k=1}^m A_k$ and $z(\cdot)$ satisfies $X_\varepsilon E$.

A.4. Proof of Theorem 2

Let $\{(\{e^r(\theta)\}_{\theta \in \Theta}, \tilde{\theta}, \tilde{t}^r, P^r)\}_{r=1}^\infty$ be a conditionally independent sequence and suppose that each $u_i(\cdot; \theta)$ is concave.

Step 1: For each $t^r \in T^r$, let $\varphi(t^r)$ denote the "empirical frequency distribution" that t^r induces on T. More formally, $\varphi(t^r)$ is a probability measure on T defined for each $\tau \in T$ as follows:

$$\varphi(t^r)(\tau) = \frac{|\{s \in J_r | (t_{1,s}^r, \ldots, t_{n,s}^r) = \tau\}|}{r}.$$

(We suppress the dependence of φ on r for notational convenience.)

CLAIM: *For every $\rho > 0$, there exists an integer $\hat{r}$ such that for all $r > \hat{r}$,*

$$\nu_{i,s}^{P^r} \leq \rho \qquad \textit{and} \qquad \mu_{i,s}^{P^r} \leq \rho.$$

PROOF OF CLAIM: Choose $\rho > 0$. Applying the argument in the Appendix to Gul-Postlewaite (1992) (see the analysis of their equation (9)), together with the definition of φ and the law of large numbers, it follows that there exists $\lambda > 0$ and an integer $\hat{r}$ such that for all $r > \hat{r}$,

$$\|\varphi(t^r) - P_T(\cdot|\theta_k)\| < \lambda \Rightarrow \|P_\Theta^r(\cdot|t^r) - I_{\theta_k}\| < \rho/2 \quad \text{for all } t^r \text{ and } k \geq 1,$$

$$\|\varphi(t_{-is}^r, t_i) - \varphi(t_{-is}^r, t_i')\| < \lambda/2 \quad \text{for all } t_i, t_i' \in T_i \text{ and all } t^r \text{ and all } i,$$

and

$$\text{Prob}\{\|\varphi(\tilde{t}^r) - P_T(\cdot|\theta_k)\| < \lambda/2|\tilde{t}^r_{is} = t_i, \tilde{\theta} = \theta_k\} > 1-\rho \quad \text{for all } t_i, t'_i \in T_i \text{ and } k \geq 1.$$

Choose $t_i, t'_i \in T_i$, $k \geq 1$, and $r > \hat{r}$. Then

$$\begin{aligned}
&\text{Prob}\{\|P^r_\Theta(\cdot|\tilde{t}^r_{-is}, t_i) - P^r_\Theta(\cdot|\tilde{t}^r_{-is}, t'_i)\| < \rho|\tilde{t}^r_{is} = t_i, \tilde{\theta} = \theta_k\}\\
&\quad \geq \text{Prob}\{\|\varphi(\tilde{t}^r_{-is}, t_i) - P_T(\cdot|\theta_k)\| < \lambda/2 \text{ and } \|\varphi(\tilde{t}^r_{-is}, t'_i) - P_T(\cdot|\theta_k)\| < \lambda|\tilde{t}^r_{is} = t_i, \tilde{\theta} = \theta_k\}\\
&\quad \geq \text{Prob}\{\|\varphi(\tilde{t}^r_{-is}, t_i) - P_T(\cdot|\theta_k)\| < \lambda/2 \text{ and } \|\varphi(\tilde{t}^r_{-is}, t_i) - \varphi(\tilde{t}^r_{-is}, t'_i)\| < \lambda/2|\tilde{t}^r_{is} = t_i, \tilde{\theta} = \theta_k\}\\
&\quad = \text{Prob}\{\|\varphi(\tilde{t}^r_{-is}, t_i) - P_T(\cdot|\theta_k)\| < \lambda/2|\tilde{t}^r_{is} = t_i, \tilde{\theta} = \theta_k\}\\
&\quad \geq 1-\rho.
\end{aligned}$$

Hence,

$$\text{Prob}\{\|P^r_\Theta(\cdot|\tilde{t}^r_{-is}, t_i) - P^r_\Theta(\cdot|\tilde{t}^r_{-is}, t'_i)\| < \rho|\tilde{t}^r_{is} = t_i\} \geq 1-\rho$$

and we conclude that $\nu^{P^r}_{i,s} \leq \rho$. Since

$$\begin{aligned}
\|\varphi(t^r) - P_T(\cdot|\theta_k)\| < \lambda/2 &\Rightarrow \|\varphi(t^r) - P_T(\cdot|\theta_k)\| < \lambda\\
&\Rightarrow \|P^r_\Theta(\cdot|t^r) - I_{\theta_k}\| < \rho/2 < \rho \quad \text{for all } t^r,
\end{aligned}$$

whenever $r > \hat{r}$ and $k \geq 1$, it follows that

$$\begin{aligned}
&\text{Prob}\{\|P^r_\Theta(\cdot|\tilde{t}^r) - I_{\theta_k}\| < \rho|\tilde{t}^r_{is} = t_i, \tilde{\theta} = \theta_k\}\\
&\quad \geq \text{Prob}\{\|\varphi(\tilde{t}^r) - P_T(\cdot|\theta_k)\| < \lambda/2|\tilde{t}^r_{is} = t_i, \tilde{\theta} = \theta_k\}\\
&\quad > 1-\rho.
\end{aligned}$$

Hence,

$$\sum_{k=1}^{m} \text{Prob}\{\|P^r_\Theta(\cdot|\tilde{t}^r) - I_{\theta_k}\| < \rho|\tilde{t}^r_{is} = t_i\} \geq 1-\rho$$

and we conclude that $\mu^{P^r}_{i,s} \leq \rho$.

Step 2: For a conditionally independent sequence,

$$P^r_\Theta(\cdot|t_{i,s}) = P_\Theta(\cdot|t_{i,s})$$

for all r and all $t_{i,s} \in T_i$. In particular, $P^r_\Theta(\cdot|t_{i,s})$ is independent of r and it follows that

$$\Lambda^{P^r}_{i,s} = \Lambda^P_i$$

for all r and $s \in J_r$. Furthermore, $\Lambda^P_i > 0$ since $P \in \Delta^*_{\Theta \times T}$.

Step 3:

CLAIM: *Suppose that* $\mathscr{A} = \{x(\theta)\}_{\theta\in\Theta}$ *is a collection of CIE allocations such that, for each* $\theta \in \Theta$*, the allocation* $\{x_i(\theta)\}_{i\in N}$ *is a CIE allocation for* $e(\theta)$ *with* $x_i(\theta) \neq 0$ *for all* θ *and for all i. For every* $\varepsilon > 0$*, there exists an* $\hat{r} > 0$ *such that, for all* $r > \hat{r}$*, there exists an incentive compatible PIE allocation* $z^r(\cdot)$ *for the PIE* $(\{e^r(\theta)\}_{\theta\in\Theta}, \tilde{\theta}, \tilde{t}^r, P^r)$ *and a collection* $B^r_1, \ldots, B^r_m$ *of disjoint subsets of* T^r *such that* $\text{Prob}\{\tilde{t}^r \in \bigcup_{k=1}^m B^r_k\} \geq 1-\varepsilon$ *and, for all* $k \in J_m$ *and all* $t^r \in B^r_k$*:*

(i) $\text{Prob}\{\tilde{\theta} = \theta_k|\tilde{t}^r = t^r\} \geq 1-\varepsilon$;

(ii) *for all* $i \in N$,

$$u_i(x_i(\theta_k); \theta_k) \geq u_i(z^r_{is}(t^r); \theta_k) \geq u_i(x_i(\theta_k); \theta_k) - \varepsilon.$$

PROOF OF CLAIM: We will sketch the proof since the details, while notationally cumbersome, are identical to those in the proof of Theorem 1. Suppose that $\mathscr{A} = \{x(\theta)\}_{\theta \in \Theta}$ is a collection of CIE allocations such that, for each $\theta \in \Theta$, the allocation $\{x_i(\theta)\}_{i \in N}$ is a CIE allocation for $e(\theta)$ with $x_i(\theta) \neq 0$ for all θ and for all i. Choose $\varepsilon > 0$. As in the proof of Theorem 1, let

$$K_1 = \max_{\theta} \max_{i} \left\{ u_i \left(\sum_{j \in N} w_j ; \theta \right) \right\}$$

and choose δ so that

$$0 < \delta < \min \left\{ \frac{c(\varepsilon, \mathscr{A})}{20\sqrt{m} K_1}, \frac{\varepsilon}{2}, \frac{1}{3} \right\}.$$

Finally, define $\hat{\mu}^{P^r} = \max_{i,s} \mu_{i,s}^{P^r}$, $\hat{\nu}^{P^r} = \max_{i,s} \nu_{i,s}^{P^r}$, $\Lambda^{P^r} = \min_{i,s} \Lambda_{i,s}^{P^r}$, and $\Lambda^P = \min_i \Lambda_i^P$. Applying Steps 1 and 2, there exists an r such that for all $r > \hat{r}$,

$$\hat{\nu}^{P^r} \leq \delta \Lambda^P = \delta \Lambda^{P^r},$$
$$\hat{\mu}^{P^r} \leq \delta \Lambda^P = \delta \Lambda^{P^r}.$$

For each k, let

$$B_k^r = \{ t^r \in T^r \mid \| P_\Theta^r(\cdot | t^r) - I_{\theta_k} \| \leq \hat{\mu}^{P^r} \}$$

and let

$$B_0^r = T^r \Big\backslash \Big[\bigcup_k B_k^r \Big].$$

Since $\Lambda^P \leq 2$, it follows that

$$\hat{\mu}^{P^r} \leq \delta \Lambda^P < \frac{1}{3} \Lambda^P \leq \frac{2}{3}$$

and the collection $\Pi = \{B_0^r, B_1^r, \ldots, B_m^r\}$ is a partition of T^r.

Since $\Lambda^P = \Lambda^{P^r}$, we can apply Lemma A.1 and conclude that there exists a collection $\{\{z_i(\theta, t_i)\}_{(\theta, t_i) \in \Theta \times T_i}\}_{i \in N}$ satisfying:

(i) $z_i(\theta, t_i) \in \Re_+^l$ and $\sum_{i \in N} (z_i(\theta, t_i) - w_i) \leq 0$ for all $t_i \in T_i$ and all $\theta \in \Theta$;

(ii) $u_i(x_i(\theta); \theta) \geq u_i(z_i(\theta, t_i); \theta) \geq u_i(x_i(\theta); \theta) - \varepsilon$ for all $t_i \in T_i$ and all $\theta \in \Theta$;

(iii) for each $t_i, t_i' \in T_i$,

$$\sum_\theta [u_i(z_i(\theta, t_i); \theta) - u_i(z_i(\theta, t_i'); \theta)] P_\Theta(\theta | t_i) \geq \frac{c(\varepsilon, \mathscr{A})}{2\sqrt{m}} \Lambda^{P^r}.$$

Next, let $z^r(\cdot)$ be the PIE allocation for $(\{e^r(\theta)\}_{\theta \in \Theta}, \tilde{\theta}, \tilde{t}^r, P^r)$ defined as

$$\begin{aligned} z_{is}^r(t^r) &= z_i(\theta_k, t_i) && \text{if } t^r \in B_k^r \text{ and } t_{is}^r = t_i, \\ &= 0 && \text{if } t^r \in B_0^r. \end{aligned}$$

Note that

$$u_i(z_{is}^r(t^r); \theta) \leq K_1.$$

The proof of the claim is now completed using exactly the same arguments as those used in the proof of Theorem 1.

Step 4: We now complete the proof of Theorem 2. Let $\varepsilon > 0$ be given. Suppose that $\mathscr{A} = \{x(\theta)\}_{\theta\in\Theta}$ is a collection where, for each θ, $x(\theta)$ is a Pareto efficient, strictly individually rational CIE allocation for the CIE $e(\theta)$. Let

$$K_1 = \max_\theta \max_i \left\{ u_i\left(\sum_{j\in N} w_j; \theta\right)\right\}$$

and let

$$K_2(\mathscr{A}) = \max_i \max_\theta [u_i(x_i(\theta); \theta) - u_i(w_i; \theta)].$$

Since each $x(\theta)$ is strictly individually rational for the CIE $e(\theta)$, it follows that $K_2(\mathscr{A}) > 0$. Choose $\hat{\varepsilon}$ so that

$$0 < \hat{\varepsilon} < \min\left\{\frac{K_2(\mathscr{A})}{4K_1+1}, \frac{\varepsilon}{4K_1+1}\right\}.$$

Applying Step 3, there exists an $\hat{r} > 0$ such that, for all $r > \hat{r}$, there exists an incentive compatible PIE allocation $z^r(\cdot)$ for the PIE $(\{e^r(\theta)\}_{\theta\in\Theta}, \tilde{\theta}, \tilde{t}^r, P^r)$ and a collection $B_1^r, \ldots, B_m^r$ of disjoint subsets of T^r such that $\text{Prob}\{\tilde{t}^r \in \bigcup_{k=1}^m B_k^r\} \geq 1 - \hat{\varepsilon}$ and, for all $k \in J_m$ and all $t^r \in B_k^r$:

(i) $\text{Prob}\{\tilde{\theta} = \theta_k | \tilde{t}^r = t^r\} \geq 1 - \hat{\varepsilon}$;

(ii) for all $i \in N$,

$$u_i(x_i(\theta_k); \theta_k) \geq u_{is}(z_{is}^r(t^r); \theta_k) \geq u_i(x_i(\theta_k); \theta_k) - \hat{\varepsilon}.$$

Suppose that $r > \hat{r}$. If $t^r \in B_k^r$ for some $k \geq 1$, then $P_\Theta^r(\theta_k|t) \geq 1 - \hat{\varepsilon}$ implies that

$$\|P_\Theta^r(\cdot|t^t) - I_{\theta_k}\| = [1 - P_\Theta^r(\theta_k|t^r)] + \sum_{\ell\neq k} P_\Theta(\theta_\ell|t^r) \leq 2\hat{\varepsilon}.$$

To prove that $z^r(\cdot)$ satisfies XIR, suppose that $t \in B_k^r$ and note that

$$\begin{aligned}\sum_\theta [u_{is}(z_{is}(t^r); \theta) - u_{is}(w_i; \theta)]P(\theta|t^r) &= \sum_\ell [u_i(z_{is}(t^r); \theta_\ell) - u_i(w_i; \theta_\ell)]P_\Theta(\theta_\ell|t^r) \\ &\geq u_i(z_{is}(t); \theta_k) - u_i(w_i; \theta_k) - (2K_1)(2\hat{\varepsilon}) \\ &\geq u_i(x_i(\theta_k); \theta_k) - u_i(w_i; \theta_k) - \hat{\varepsilon} - 4K_1\hat{\varepsilon} \\ &\geq K_2(\mathscr{A}) - (4K_1+1)\hat{\varepsilon} \\ &> 0.\end{aligned}$$

Hence, $z^r(\cdot)$ satisfies XIR.

To show that $z^r(\cdot)$ satisfies $\text{X}_\varepsilon\text{E}$, let $E^r = \bigcup_{k=1}^m B_k^r$ and note that

$$\text{Prob}\{\tilde{t}^r \in E^r\} = \text{Prob}\left\{\tilde{t}^r \in \bigcup_{k=1}^m B_k^r\right\} \geq 1 - \hat{\varepsilon} \geq 1 - \frac{\varepsilon}{4K_1+1} \geq 1 - \varepsilon.$$

Now suppose that $y^r(\cdot)$ is a feasible PIE allocation for e^r satisfying

$$\sum_\theta [u_{is}(y_{is}^r(t^r); \theta) - u_{is}(z_{is}^r(t^r); \theta)]P(\theta|t^r) > \varepsilon$$

for each (i, s). For each i, let

$$\bar{y}_i = \frac{1}{r}\sum_{s=1}^r y_{is}^r(t^r)$$

and therefore,

$$\sum_{i=1}^{n} \bar{y}_i = \frac{1}{r} \sum_{i=1}^{n} \sum_{s=1}^{r} y_{is}^r(t^r) \le \sum_{i=1}^{n} w_i.$$

Note that $u_i(\bar{y}_i; \theta) \le K_1$ since $\sum_{i=1}^n \bar{y}_i \le \sum_{i=1}^n w_i$.

Suppose that $t^r \in B_k^r$ for some $k \ge 1$. Then for each $i \in N$,

$$\sum_{\theta} u_i(z_{is}^r(t^r); \theta) P(\theta|t^r) \ge u_i(z_{is}^r(t^r); \theta_k) - 2K_1\hat{\varepsilon} \ge u_i(x_i(\theta_k); \theta_k) - 2K_1\hat{\varepsilon} - \hat{\varepsilon}$$

and

$$\sum_{\theta} u_i(\bar{y}_i; \theta) P(\theta|t^r) \le u_i(\bar{y}_i; \theta_k) + 2K_1\hat{\varepsilon}.$$

Combining these inequalities and using the concavity of each $u_i(\cdot; \theta)$, we conclude that

$$\begin{aligned} u_i(\bar{y}_i; \theta_k) + 2K_1\hat{\varepsilon} &\ge \sum_{\theta} u_i(\bar{y}_i; \theta) P(\theta|t^r) \\ &\ge \sum_{\theta} \left(\frac{1}{r} \sum_{s=1}^{n} u_i(y_{is}^r(t^r); \theta) \right) P(\theta|t^r) \\ &> \frac{1}{r} \sum_{s=1}^{r} \left[\sum_{\theta} u_i(z_{is}^r(t^r); \theta) P(\theta|t^r) \right] + \varepsilon \\ &\ge u_i(x_i(\theta_k); \theta_k) - 2K_1\hat{\varepsilon} - \hat{\varepsilon} + \varepsilon. \end{aligned}$$

Therefore,

$$0 < \varepsilon - (4K_1 + 1)\hat{\varepsilon} < [u_i(\bar{y}_i; \theta_k) - u_i(x_i(\theta_k); \theta_k)]$$

for each i. Since $(\bar{y}_i)_{i \in N}$ is feasible for the CIE $e(\theta_k)$, we conclude that $(x_i(\theta_k))_{i \in N}$ is not Pareto optimal in $e(\theta_k)$, a contradiction. Hence, $t^r \notin \bigcup_{k=1}^m B_k^r = E^r$ and the proof is complete.

REFERENCES

AL-NAJJAR, N., AND R. SMORODINSKY (2000): "Pivotal Players and the Characterization of Influence," *Journal of Economic Theory*, 92, 318–342.

CREMER, J., AND R. MCLEAN (1985): "Optimal Selling Strategies under Uncertainty for a Monopolist when Demands are Interdependent," *Econometrica*, 53, 345–361.

——— (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, 1247–1257.

FUDENBERG, D., D. LEVINE, AND W. PESENDORFER (1998): "When Are Nonanonymous Players Negligible?" *Journal of Economic Theory*, 79, 46–71.

GUL, F., AND A. POSTLEWAITE (1992): "Asymptotic Efficiency in Large Economies with Asymmetric Information," *Econometrica*, 60, 1273–1292.

JACKSON, M. (1991): "Bayesian Implementation," *Econometrica*, 59, 461–477.

KRASA, S., AND W. SHAFER (1998): "Informational Robustness of Competitive Equilibria," Unpublished Manuscript, University of Illinois.

MCAFEE, R. P., AND P. RENY (1992): "Correlated Information and Mechanism Design," *Econometrica*, 60, 395–421.

MCLEAN, R. P., AND A. POSTLEWAITE (2001): "Informational Size and Incentive Compatibility with Aggregate Uncertainty," Unpublished Manuscript, University of Pennsylvania.

POSTLEWAITE, A., AND D. SCHMEIDLER (1986): "Implementation in Differential Information Economies," *Journal of Economic Theory*, 39, 14–33.