CARESS Working Paper # 99-14
Informational Size and Incentive Compatibility*

Richard McLean
Rutgers University

Andrew Postlewaite
University of Pennsylvania

October 12, 1999

Abstract

We examine a general equilibrium model with asymetrically 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 sufficiently small informationally .

1

*Early versions of this paper circulated under the titles "Informational Smallness and Pareto Efficiency" and "Informational Smallness and Incentive Compatibility." We thank Dean Foster, Andreu MasColell, Stephen Morris, David Schmeidler and particularly Faruk Gul for helpful conversations, and numerous participants of seminars at which this paper was presented for helpful comments. We also thank Ichiro Obara for pointing out an error in an earlier version. Postlewaite gratefully acknowledges support from the National Science Foundation.

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 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 when 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, our goal is to formalize informational size in a way that, when agents are informationally small, one can ignore the incentive problems associated with the presence of asymmetric information without substantially affecting the resulting analysis. We analyze an Arrow-Debreu pure exchange economy in which the agents are asymmetrically informed. Specifically, the agents' utility functions will depend on an underlying but unobserved state of nature and each agent will receive a private signal that is correlated with the state of nature. Roughly speaking, this corresponds to a "common value" model in which signals do not directly affect the underlying payoff functions but do affect expected utilities. This model encompasses the typical "finance" problem in which agents' expected utilities depend on their information only to the extent that it is correlated with the value of some asset.

For this common value model, we will provide a plausible definition of informational size and we will show that certain Pareto efficient allocations can be approximated in utility by an incentive compatible allocation if agents have sufficiently small informational size. We further show that there are at least two situations in which agents will be informationally small in our sense, and to which our theorems apply. The first is the case in which a given economy with private information is replicated, and the second is the case in which agents' have largely redundant information.

In the next section, we provide a brief review of related literature. Section 3 provides a formal description of our model. In section 4, we present an example that illustrates the ideas and logic of our results and we present the formal definition of informational size. In section 5, we present our main result: if agents are sufficiently informationally small and if the agents' information in the aggregate resolves nearly all the uncertainty with respect to Θ, then there exists an incentive compatible allocation that is ex post individually rational and approximately ex post efficient. Our use of the term "ex post" refers to events that occur after the realization of the agents' signals but before the realization of the state θ. We then show how this theorem applies to replica economies. In section 6, we extend these results to the case in which nonnegligible residual uncertainty is present, even when all agents' information is known. Section 7 concludes with a discussion of our results. All proofs are contained in the appendix.

2 Related Literature

Gul and Postlewaite (1992) is the closest work to ours. They consider an economy with asymmetric information that is replicated. Each replica is an independent draw from the probability distribution over agents' types. A given agent's utility function depends on his type and the types of the agents in his replica, but not on the types of agents in other replicas. While that paper did not provide a formal definition, it used the following informal notion of informational size: an agent is informationally small if the incremental impact of each agent's information (given the information of others) on the demand for every good is small. In a replica economy, the fraction of agents whose utility functions depend on a given agent's type goes to zero and, as a result, agents are informationally small in the sense described above. Gul and Postlewaite 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 ways. In that paper, the (informal) notion of informational size was motivated by the particular replication process considered there. In this paper, we provide a formal definition of informational size that is applicable to general asymmetric information economies with a common value structure. Replication will be but one natural setting in which agents will be informationally small. Furthermore, this definition depends only on the information structure of the economy (i.e., the joint probability of states and signals) and is independent of endowments and utility functions.

There are other important differences in addition to the different notions of informational size. In the economies analyzed in Gul - Postlewaite, the agents' utilities may have a common value component but an individual agent's utility cannot be independent of his own type (i.e., his signal). This excludes from consideration the case of pure common values in which there is an underlying value for any good with agents' information being of use only in predicting that value. In this paper, we treat this common value problem but it should be emphasized that we exclude any private value component. The last relevant difference is that 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 can be approximated by incentive compatible allocations when agents are sufficiently informationally small.

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 roughly corresponds to the case in which agents' informational size is 0.

3 Private Information Economies:

Throughout the paper, $J_q = \{1, .., q\}$ for each positive integer q and $||\cdot||$ will denote the 1-norm unless specified otherwise. Let $N = \{1, 2, ..., n\}$ denote the set of **economic agents**. Let $\Theta = \{\theta_1, .., \theta_m\}$ denote the (finite) **state space** and let $T_1, T_2, ..., T_n$ be finite sets where T_i represents the set of possible **signals** that agent i might receive. 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, define

$$\Delta_X := \{\rho \in \Re^{|X|} | \rho(x) \geq 0, \sum_{x \in X} \rho(x) = 1\}.$$

In our model, nature chooses an element $\theta \in \Theta$. 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 $(\widetilde{\theta}, \widetilde{t}_1, \widetilde{t}_2, ..., \widetilde{t}_n)$ be an (n+1)-dimensional random vector taking values in $\Theta \times T$ with associated distribution $P \in \Delta_{\Theta \times T}$ where

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

Without loss of generality, we will make the following assumption regarding the marginal distributions:

full support: $\text{supp}(\widetilde{\theta}) = \Theta$ i.e. for each $\theta \in \Theta$,

$$P(\theta) = \text{Prob}\{\widetilde{\theta} = \theta\} > 0$$

and for each $i \in N$, $\text{supp}(\widetilde{t}_i) = T_i$ i.e. for each $t_i \in T_i$,

$$P(t_i) = \text{Prob}\{\widetilde{t}_i = t_i\} > 0.$$

Let $T^* = \text{supp}(\tilde{t}_1, \tilde{t}_2, ..., \tilde{t}_n) = \{t \in T | P(t) > 0\}$. We note that T^* need *not* be equal to T. If $t \in T^*$, let $P_\Theta(\cdot|t) \in \Delta_\Theta$ denote the induced conditional probability measure on Θ. Let $\chi_\theta \in \Delta_\Theta$ denote the degenerate measure that puts probability one on state θ.

The **consumption set** of each agent is $\Re^\ell_+$ and $w_i \in \Re^\ell_+ \backslash \{0\}$ denotes the **initial endowment** of agent i (an agent's initial endowment is independent of the state θ). For each $\theta \in \Theta$, let $u_i(\cdot, \theta) : \Re^\ell_+ \rightarrow \Re$ be the utility function of agent i in state θ. 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 $\sum_{i \in N}(x_i(\theta) - w_i) \leq 0$. 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** $z = (z_1, z_2, ..., z_n)$ for the PIE is a collection of functions z_i: $T \rightarrow \Re^\ell_+$ satisfying $\sum_{i \in N}(z_i(t) - w_i) \leq 0$ for all $t \in T$. We will not distinguish between $w_i \in \Re^\ell_+$ and the constant allocation that assigns the bundle w_i to agent i for all $t \in T$.

If $z = (z_1, z_2, ..., z_n)$ is a PIE allocation, then define

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

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

$$U_i(z_i \mid t) = \sum_{\theta \in \Theta} u_i(z_i(t), \theta) P(\theta \mid t)$$

$$= E[u_i(z_i(\tilde{t}), \tilde{\theta}) \mid \tilde{t} = t]$$

for each $t \in T$.

A PIE allocation z is said to be:

(*incentive compatible)* (IC) if

$$U_i(z_i, t_i | t_i) \geq U_i(z_i, t_i' | t_i)$$

for all $i \in N$, and all $t_i, t'_i \in T_i$.

(*ex post individually rational)* (XIR) if

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

for all $i \in N$ and for all $t \in T^*$.

(*ex post* $\varepsilon-$*efficient*)(X_εE) if there exists $E \subseteq T^*$ such that $P(E) \geq 1 - \varepsilon$ 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 vector t but *before* the realization of the state θ.

4 Informational Size

As we stated in the introduction, our goal is to provide a notion of "informational size" having the property that, when agents are informationally small, one can ignore the informational asymmetries and the analysis will be nearly the same as that which would result if these asymmetries were fully accounted for. We will illustrate our model and the notion of informational smallness with the following example.

4.1 An Example

The economy consists of three pairs of agents, each consisting of an agent A who initially has money and an agent B who has one unit of a second good. All agents have linear, separable utility functions. The utilities $u_A(m, x; \theta)$ and $u_B(m, x; \theta)$ for each type of agent for m units of money and x units of the second good in each of the two states, θ_1 and θ_2 are given in the following table.

agent \ *state*	θ_1	θ_2
A	$m + 23x$	$m + 7x$
B	$m + 20x$	$m + 4x$

Each type A agent has an initial endowment of $\overline{m}$ units of money and zero units of the second good. Each type B agent has an initial endowment of 0 units of money and one unit of the second good. Efficiency dictates that all of the second good be transferred from the B agents to the A agents; to be individually rational, each of the B agents must receive an amount of money to compensate him for giving up the second good and these payments will depend on the state. If trade takes place at payments m_1 and m_2, the resulting utilities are

	state	θ_1	θ_2
agent			
A		$\overline{m} - m_1 + 23$	$\overline{m} - m_2 + 7$
B		m_1	m_2

If $20 \leq m_1 \leq 23$ and $4 \leq m_2 \leq 7$, then the trades are efficient and individually rational.

We will next describe the information structure of the economy. Type B agents receive a signal (α or β) correlated with the state of nature while type A agents receive no signal. Each state is equally likely, the agents' signals are independent conditional on the state and the matrix of conditional probabilities $P(\theta|t_i)$ for $\theta \in \{\theta_1, \theta_2\}$ and $t_i \in T_i = \{\alpha, \beta\}$ is

	signal	α	β
state			
θ_1		$\frac{1+r}{2}$	$\frac{1-r}{2}$
θ_2		$\frac{1-r}{2}$	$\frac{1+r}{2}$

where $0 \leq r \leq 1$

If the goal is to effect a transaction that is individually rational and Pareto efficient, we must induce the B agents to truthfully reveal their signals in order to determine whether the payment should be relatively high (when θ_1 is the likely state) or relatively low (when θ_2 is the likely state). An obvious incentive compatibility problem arises since the informed type B agents have a clear interest in making θ_1 seem the likely state.

Consider the following revelation mechanism. The type B agents announce their signals and the true state of nature is "estimated" to be θ_1 if a majority of the B's announce α, and θ_2 if a majority of the B's announce β. Each B agent will then transfer his one unit of the second good to an A agent in return for a payment that depends on both the estimated state and his announcement as in the following table.

own announcement(t_i)	*estimated state*(θ)	*payment*
α	θ_1	22
β	θ_1	21
α	θ_2	5
β	θ_2	6

For example, if $t_{-i} = (\alpha, \alpha)$, then the estimated state is θ_1, independent of i's announcement. In this case, i receives a payment of 22 if $t_i = \alpha$ and 21 if $t_i = \beta$.

Hence, one can think of the mechanism as specifying a payment that depends on the estimated state, and then "punishes" an agent if his announcement differs from the majority. We note several things about the mechanism. First, if $r \approx 1$, then the information of the three B agents is enough to predict the state nearly perfectly. In particular, if $r \approx 1$, then $P(\theta_1|t) \approx 1$ if a majority of B's announce α while $P(\theta_2|t) \approx 1$ if a majority of B's announce β. Hence, the mechanism is XIR and (because of the linear utilities) satisfies X_0E.

The mechanism is also incentive compatible when $r \approx 1$. Suppose that a B agent sees signal α. Then a misreport of β will change the estimated state or leave it unchanged. The estimated state will change only when the other two B agents have received different signals. The probability that the other two B agents receive different signals, conditional on a B agent receiving signal α, is close to zero when $r \approx 1$. (In fact, $P(\alpha, \alpha|\alpha) \approx 1$ when $r \approx 1$.) In our example, the maximum possible gain from lying is bounded. Therefore, a misreport that changes the estimated state may be profitable but the *expected* gain from misreporting approaches zero as r approaches 1. While the expected gain from a misreport that changes the estimated state may be negligible (for $r \approx 1$), a type B agent must still be induced to report truthfully and we accomplish this by "punishing" him when his misreport does not change the state, i.e., when his announcement is not in the majority. In the example, it is very likely that, conditional on having observed α, the signals of the other two type B agents are identical when $r \approx 1$. A truthful report when he sees α yields an expected payment close to 22 while a misreport of β yields an expected payment close to 21. Hence, a misreport that does not change the estimated state results in an approximate loss of 1. When the two effects are combined, we conclude that, for $r \approx 1$, a misreport results in an approximate decrease in expected utility equal to 1. The same argument works for a B agent who observes β but reports α and the mechanism is incentive compatible when $r \approx 1$.

This example exhibits three essential features that will play an important role in the general results that we prove later. First, the conditional distributions $P_\Theta(\cdot|t)$ on Θ are nearly degenerate for each t if $r \approx 1$. This property is a special case of "*negligible aggregate uncertainty*" that we will introduce in the general model below. The example exhibits a second crucial feature: *informational smallness.* If $r \approx 1$, then the probability that, conditional on his true signal, a misreport by a B agent will change the estimated state is small. A third property of the distribution P is also important but less obvious: the posterior distributions $P_\Theta(\cdot|\alpha)$ and $P_\Theta(\cdot|\beta)$ on Θ are different. When $r \approx 1$, $P(\alpha, \alpha|\alpha) \approx P_\Theta(\theta_1|\alpha) \approx 1$ and $P(\beta, \beta|\beta) \approx P_\Theta(\theta_2|\beta) \approx 1$. Each agent knows that the probability that he can change the estimated state is close to zero and, therefore, focuses on avoiding the penalty for not announcing with the majority. An agent who has received signal α believes that state θ_1 is more likely than θ_2, and is strictly better off by announcing truthfully than by misreporting. If $P_\Theta(\cdot|\alpha) = P_\Theta(\cdot|\beta)$, the agent would be equally likely to be punished by a truthful announcement as by misreporting.

The linear utilities of the example make it possible to construct a mechanism that is incentive compatible, individually rational and Pareto efficient. In the case of general (nonlinear) utilities, 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 for that state, modified so as to induce truthful revelation.

4.2 The Definition of Informational Size

In the mechanism of the example, agents reveal their types, and the announced types are used to estimate the state of nature. The mechanism is incentive compatible because agent i is informationally small in the following sense: i does not have a "large" influence on the estimate of θ conditioned on agents' announcements when other agents announce truthfully.

To investigate these issues in a more general framework, we need to formalize the idea of *informational size.* If $t \in T^*$, recall that $P_\Theta(\cdot|t) \in \Delta_\Theta$ denotes the induced conditional probability measure on Θ and $\chi_\theta \in \Delta_\Theta$ denotes the measure that puts probability one on θ. Our example indicates that a natural notion of an agent's informational size is the degree to which he can alter the best estimate of the state θ when other agents are announcing truthfully. In our setup, that estimate is the conditional probability distribution on Θ given a vector of types t. 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. We consider agent i to be informationally small if, for each t_i, there is a "small" probability that he can induce a "large" change in the induced conditional distribution on Θ by changing his announced type from t_i to some other t'_i. We formalize this in the following definition.

Definition: Let

$$I^i_\varepsilon(t'_i, t_i) = \{t_{-i} \in T_{-i} | (t_{-i}, t_i) \in T^*, (t_{-i}, t'_i) \in T^* \text{ and } ||P_\Theta(\cdot|t_{-i}, t_i) - P_\Theta(\cdot|t_{-i}, t'_i)|| > \varepsilon\}$$

The *informational size* of agent i is defined as

$$\nu_i^P = \max_{t_i \in T_i} \max_{t'_i \in T_i} \inf\{\varepsilon > 0 | \text{ Prob}\{\tilde{t}_{-i} \in I^i_\varepsilon(t'_i, t_i) | \tilde{t}_i = t_i\} \leq \varepsilon\}$$

Loosely speaking, we will say that agent i is *informationally small* with respect to P if his informational size ν_i^P is "small." If agent i receives signal t_i but reports $t'_i \neq t_i$, then the effect of this misreport is a change in the conditional distribution on Θ from $P_\Theta(\cdot|t_{-i},t_i)$ to $P_\Theta(\cdot|t_{-i},t'_i)$. If $t_{-i}\in I_\varepsilon(t'_i,t_i)$, then this change is "large" in the sense that $||P_\Theta(\cdot|\hat{t}_{-i},t_i)-P_\Theta(\cdot|\hat{t}_{-i},t'_i)||>\varepsilon$. Therefore, $\text{Prob}\{\tilde{t}_{-i}\in I_\varepsilon(t'_i,t_i)|\tilde{t}_i=t_i\}$ is the probability that i can have a "large" influence on the conditional distribution on Θ by reporting t'_i instead of t_i when his observed signal is t_i. 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.

Informational smallness is not related to the "quality" of an agent's information. In the example of section 4.1, $P_\Theta(\cdot|t_i)$ is nearly degenerate for each t_i when r is close to 1. Hence, agents have good estimates of the true state conditional on their signals, yet each agent is informationally small.

5 The Case of Negligible Aggregate Uncertainty

In this section, we will study a general problem (motivated by the example of section 4.1) in which the agents' aggregate information "almost" resolves the uncertainty regarding the state θ and the agents have "small" but nonzero informational size.

5.1 An outline of the approach

In order to convey the underlying ideas, we will first isolate the role of informational size by considering an example in which $P_\Theta(\cdot|t)$ is a vertex of Δ_Θ for every $t\in T^*$. Let $(\{e(\theta)\}_{\theta\in\Theta},\widetilde{\theta},\tilde{t},P)$ be a PIE. Suppose that $T=T^*$ and that T is partitioned into nonempty sets $\{T(\theta)\}_{\theta\in\Theta}$ with the property that $P_\Theta(\cdot|t)=\chi_\theta$ if and only if $t\in T(\theta)$. (Recall that $\chi_\theta\in\Delta_\Theta$ is the measure that puts probability one on state θ.) We will refer to this example as the case of "zero aggregate uncertainty": if agents truthfully reveal their signals, then the state of nature is known with certainty. What is an agent's informational size in this example? If $\varepsilon>0$, then $||P_\Theta(\cdot|t_{-i},t_i)-P_\Theta(\cdot|t_{-i},t'_i)||>\varepsilon$ if and only if there exist θ and θ' with $\theta\neq\theta'$ such that $(t_{-i},t_i)\in T(\theta)$ and $(t_{-i},t'_i)\in T(\theta')$. Therefore,

$$\begin{aligned}\inf\{\varepsilon &> 0|\ \text{Prob}\{\tilde{t}_{-i}\in I_\varepsilon(t'_i,t_i)|\tilde{t}_i=t_i\}\le\varepsilon\}\\ &= \sum_{\theta\in\Theta}\text{Prob}\{(\tilde{t}_{-i},t_i)\in T(\theta)\text{ and }(\tilde{t}_{-i},t'_i)\notin T(\theta)|\tilde{t}_i=t_i\}\end{aligned}$$

from which it follows that

$$\nu_i^P=\max_{t_i\in T_i}\max_{t'_i\in T_i}\sum_{\theta\in\Theta}\text{Prob}\{(\tilde{t}_{-i},t_i)\in T(\theta)\text{ and }(\tilde{t}_{-i},t'_i)\notin T(\theta)|\tilde{t}_i=t_i\}.$$

For each $\theta\in\Theta$, let $\zeta(\theta)$ be an efficient, strictly individually rational complete information allocation for the CIE $e(\theta)$. First, consider a very simple mechanism

$z^*(\cdot)$ where $z^*(t) = \zeta(\theta)$ whenever $t \in T(\theta)$. The efficiency and strict individual rationality of the CIE allocation $\zeta(\theta)$ in $e(\theta)$ for each θ will imply that $z^*(\cdot)$ will be ex-post efficient and ex-post individually rational for the PIE. Incentive compatibility is less obvious. From the definitions, it follows that

$$\begin{aligned}
& \sum_{\theta\in\Theta}\sum_{t_{-i}\in T_{-i}} \left[u_i(z_i^*(t_{-i},t_i),\theta) - u_i(z_i^*(t_{-i},t_i'),\theta)\right] P(\theta,t_{-i}|t_i) \\
= & \sum_{\theta\in\Theta}\sum_{\substack{t_{-i}\in T_{-i} \\ :(t_{-i},t_i)\in T(\theta)}}\sum_{\hat\theta\in\Theta} \left[u_i(z_i^*(t_{-i},t_i),\hat\theta) - u_i(z_i^*(t_{-i},t_i'),\hat\theta)\right] P(\hat\theta|t)P(t_{-i}\mid t_i) \\
= & \sum_{\theta\in\Theta}\sum_{\substack{t_{-i}\in T_{-i} \\ :(t_{-i},t_i)\in T(\theta)}} \left[u_i(z_i^*(t_{-i},t_i),\theta) - u_i(z_i^*(t_{-i},t_i'),\theta)\right] P(t_{-i}\mid t_i) \\
= & \sum_{\theta\in\Theta}\sum_{\substack{t_{-i}\in T_{-i} \\ :(t_{-i},t_i)\in T(\theta) \\ :(t_{-i},t_i')\notin T(\theta)}} \left[u_i(z_i^*(t_{-i},t_i),\theta) - u_i(z_i^*(t_{-i},t_i'),\theta)\right] P(t_{-i}\mid t_i) \qquad (*)
\end{aligned}$$

If $\nu_i^P = 0$, then (*) is equal to zero and the simple mechanism $z^*(\cdot)$ will be incentive compatible. However, the mechanism $z^*(\cdot)$ may not be incentive compatible if $\nu_i^P > 0$. In this case, there exist θ and $\hat\theta$ with $\theta \neq \hat\theta$ and a set $A \subseteq T_{-i}$ such that (i) $\text{Prob}(\tilde t_{-i} \in A|\tilde t_i = t_i) > 0$ and (ii) $(t_{-i}, t_i) \in T(\theta)$ and $(t_{-i}, t_i') \in T(\hat\theta)$ whenever $t_{-i} \in A$. If $u_i(\zeta_i(\theta);\theta) < u_i(\zeta_i(\hat\theta);\theta)$, then i gains by lying whenever $t_{-i} \in A$. In this case,

$$\left[u_i(\zeta_i(\theta);\theta) - u_i(\zeta_i(\hat\theta);\theta)\right] P(\tilde t_{-i} \in A|\tilde t_i = t_i)$$

is negative, (*) may be negative and incentive compatibility may be violated.

The construction of $z^*(\cdot)$ above is intended to give agent i a bundle that depends only on the "estimate" of the most likely state of nature. Now suppose that we construct a new set of allocations by perturbing the members of the collection $\{\zeta(\theta)\}_{\theta\in\Theta}$ (the size of the perturbation depending on i's announced type, t_i and the estimated state, θ). Let $z_i(\theta, t_i)$ denote the perturbed bundle and suppose that the $z_i(\theta,t_i)'s$ satisfy two conditions: (i) $(z_1(\theta,t_1),..,z_n(\theta,t_n))$ is a feasible allocation for the CIE $e(\theta)$ and (ii) $z_i(\theta,t_i) \approx \zeta_i(\theta)$.[1] Since the allocation $\zeta(\theta)$ is strictly individually rational and efficient in $e(\theta)$, it follows that $(z_1(\theta,t_1),..,z_n(\theta,t_n))$ will be strictly individually rational and approximately efficient in $e(\theta)$. We can define a new mechanism where i's bundle depends on the estimated state and i's announcement as follows: if $t \in T$ is the announced vector of signals, then i receives $\hat z_i(t) = z_i(\theta, t_i)$ if $t \in T(\theta)$. The mechanism $\hat z(\cdot)$ will be ex-post IR and approximately ex-post efficient.

[1] In a more general approach, one would choose an allocation $z_i(\theta, t)$ rather than $z_i(\theta, t_i)$. That is, the perturbed allocation would depend on the estimated state θ and the entire vector of announced types rather than only agent i's announced type. The advantages of this are outlined in the second point of the discussion section at the end of the paper.

Although this modified mechanism sacrifices some efficiency, it may allow us to attain incentive compatibility. To see this, note that zero aggregate uncertainty implies that for each θ,

$$P_\Theta(\theta|t_i) = \sum_{\hat{\theta}\in\Theta} \sum_{\substack{t_{-i}\in T_{-i} \\ :(t_{-i},t_i)\in T(\hat{\theta})}} P(\theta|t)P(t_{-i} \mid t_i) = \sum_{\substack{t_{-i}\in T_{-i} \\ :(t_{-i},t_i)\in T(\theta)}} P(t_{-i} \mid t_i)$$

Therefore,

$$\sum_{\theta\in\Theta} \sum_{t_{-i}\in T_{-i}} \left[u_i(\hat{z}_i(t_{-i},t_i),\theta) - u_i(\hat{z}_i(t_{-i},t_i'),\theta)\right] P(\theta, t_{-i} \mid t_i) = L_1 + L_2$$

where

$$\begin{aligned} L_1 &= \sum_{\theta\in\Theta} \sum_{\substack{t_{-i}\in T_{-i} \\ :(t_{-i},t_i)\in T(\theta)}} \left[u_i(\hat{z}_i(t_{-i},t_i),\theta) - u_i(z_i(\theta,t_i');\theta)\right] P(t_{-i} \mid t_i) \\ &= \sum_{\theta} [u_i(z_i(\theta,t_i);\theta) - u_i(z_i(\theta,t_i');\theta)] P_\Theta(\theta|t_i) \end{aligned}$$

and

$$L_2 = \sum_{\theta\in\Theta} \sum_{\substack{t_{-i}\in T_{-i} \\ :(t_{-i},t_i)\in T(\theta) \\ (t_{-i},t_i')\notin T(\theta)}} \left[u_i(z_i(\theta,t_i');\theta) - u_i(\hat{z}_i(t_{-i},t_i'),\theta)\right] P(t_{-i} \mid t_i)$$

In the simple mechanism z^*, the L_1 term is equal to zero and does not appear in equation (*). If agent i is informationally small, then $\nu_i^P \approx 0$ and $L_2 \approx 0$. However, as in the case of the simple mechanism z^*, we are not guaranteed that $L_2 \geq 0$. If the $z_i(\theta, t_i)$ could be chosen so that $L_1 > 0$, then we might be able to force the *sum* $L_1 + L_2$ to be nonnegative, thereby obtaining incentive compatibility. This was exactly the situation in our example of section 4.1 when $r \approx 1$. When agent i receives signal t_i but contemplates a report of t_i', he must weigh two possible consequences of this lie: the case in which (t_{-i}, t_i) and (t_{-i}, t_i') belong to the *same* $T(\theta)$ and the case in which (t_{-i}, t_i) and (t_{-i}, t_i') belong to *different* $T(\theta)'s$. The number L_1 corresponds to the first case while L_2 corresponds to the second case. While we cannot rule out the existence of t_{-i} for which (t_{-i}, t_i) and (t_{-i}, t_i') belong to different $T(\theta)'s$, we do know that these values of t_{-i} are unlikely given t_i when agent i is informationally small. Hence, a deviation may be profitable when (t_{-i}, t_i) and (t_{-i}, t_i') belong to different $T(\theta)'s$ and L_2 may be negative but L_2 will be a small negative number as a consequence of informational smallness. This means that we must make it sufficiently costly for i to misreport in the event that (t_{-i}, t_i) and (t_{-i}, t_i') belong to the same $T(\theta)$ so that the total effect of a lie is a loss in utility, i.e., we must make L_1 sufficiently large so that $L_1 + L_2$ is nonnegative. Hence, we must identify conditions under which there exist perturbed allocations $(z_1(\theta, t_1), .., z_n(\theta, t_n))$ satisfying

$$\sum_{\theta} [u_i(z_i(\theta,t_i);\theta) - u_i(z_i(\theta,t_i');\theta)] P_\Theta(\theta|t_i) > 0$$

The key to the existence of such perturbed allocations is a condition concerning the variability in the conditional distributions $P_\Theta(\cdot|t_i)$ as t_i ranges over T_i and, in the next section, we provide the appropriate quantification of this variability.

The discussion up to this point only covers the special case of zero aggregate uncertainty but the extension to the case of "negligible aggregate uncertainty" uses the same ideas. Suppose that aggregate uncertainty is small but not zero in the sense that, conditional on $\tilde{t}_i = t_i$,the random variable $P_\Theta(\cdot|\tilde{t})$ is "close" to a vertex of Δ_Θ with high probability. Then

$$P_\Theta(\theta|t_i) \approx \sum_{\substack{t_{-i}\in T_{-i} \\ :(t_{-i},t_i)\in T(\theta)}} P(t_{-i} \mid t_i)$$

for each t_i and a third "error term" L_3 is added to the RHS of the incentive compatibility inequality. This error term will become smaller as aggregate uncertainty becomes smaller. Hence, our strategy in the presence of *negligible aggregate uncertainty* may be summarized as follows: identify conditions under which there exist perturbed allocations $(z_1(\theta,t_1),..,z_n(\theta,t_n))$ that will ensure that $L_1 > 0$ so that $L_1 + L_2 + L_3 \geq 0$ when agents are informationally small (so that $L_2 \approx 0$) and aggregate uncertainty is small (so that $L_3 \approx 0$).

In summary, the construction of a mechanism satisfying XIR, $\mathrm{X}_\varepsilon E$ and IC will require a delicate balance between three quantities: *informational size*, *aggregate uncertainty* and a measure of the *variability* in the conditional distributions $\{P_\Theta(\cdot|t_i)\}_{t_i\in T_i}$. We have defined informational smallness in section 4.2 and, in the next section, we quantify the aggregate uncertainty and variability in the conditional distributions.

5.2 Negligible Aggregate Uncertainty

We will next quantify aggregate uncertainty.

Definition: Let

$$\mu_i^P = \max_{t_i\in T_i} \inf\{\varepsilon > 0|\mathrm{Prob}\{\tilde{t}\in T^* \text{ and } ||P_\Theta(\cdot|\tilde{t}) - \chi_\theta|| > \varepsilon \text{ for all } \theta\in\Theta|\tilde{t}_i = t_i\} \leq \varepsilon\}$$

If μ_i^P is small for each i, then we will say that P exhibits *negligible aggregate uncertainty*. 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.

5.3 Distributional Variability

In the presence of positive but small aggregate uncertainty, we will construct a mechanism $z(\cdot)$ for the PIE satisfying XIR, IC and $\mathrm{X}_\varepsilon E$. We have noted that the construction of such a mechanism will require a delicate balance between informational size, aggregate uncertainty and a measure of the *variability* in the conditional distributions $\{P_\Theta(\cdot|t_i)\}_{t_i \in T_i}$. To define this measure of variability, let $P \in \Delta_{\Theta\times T}$ and let $P_\Theta(\cdot|t_i) \in \Delta_\Theta$ be the conditional distribution on Θ given that i receives signal t_i. Next, define

$$\Lambda_i^P = \min_{t_i \in T_i} \min_{t_i' \in T_i \setminus t_i} \frac{||P_\Theta(\cdot|t_i)||_2 ||P_\Theta(\cdot|t_i')||_2 - \sum_\theta [P_\Theta(\theta|t_i) P_\Theta(\theta|t_i')]}{||P_\Theta(\cdot|t_i')||_2}$$

where $||\cdot||_2$ denotes the 2-norm. This is precisely a measure of the difference in the distributions on $P_\Theta(\cdot|t_i)$ and $P_\Theta(\cdot|t_i')$. 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 $P \in \Delta^*_{\Theta\times T}$.

5.4 The Main Result for the case of Negligible Aggregate Uncertainty

The next theorem is the main result for problems exhibiting negligible aggregate uncertainty. Recall that $J_m = \{1, .., m\}$.

Theorem 1: Let $\Theta = \{\theta_1, .., \theta_m\}$. Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and suppose that $\mathcal{A} = \{\zeta(\theta)\}_{\theta\in\Theta}$ is a collection of associated CIE allocations with $\zeta_i(\theta) \neq 0$ 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 $(\{e(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}, P)$ and a collection $A_1, .., A_m$ of disjoint subsets of T^* such that $\text{Prob}\{\tilde{t} \in \cup_{k=1}^m A_k\} \geq 1-\varepsilon$ and, for all $k \in J_m$ and all $t \in A_k$,

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

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

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

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. Furthermore, the theorem requires that the measures of aggregate uncertainty (μ_i^P) and informational size (ν_i^P) be sufficiently small relative to the measure of variability (Λ_i^P). This is the balance between informational size, aggregate uncertainty and the variability in the conditional distributions $\{P_\Theta(\cdot|t_i)\}$ to which we alluded above. Under these conditions, we can find an incentive compatible PIE allocation $z(\cdot)$ and sets $A_1, .., A_m$ such that, whenever $t \in A_k$, $P_\Theta(\theta_k|t) \approx 1$ and $u_i(z_i(t); \theta_k) \geq u_i(\zeta_i(\theta_k); \theta_k) - \varepsilon$ for each i. If the collection $\mathcal{A} = \{\zeta(\theta)\}_{\theta\in\Theta}$ is chosen so that each $\zeta(\theta)$ is a strictly individually rational, Pareto efficient allocation for $e(\theta)$, then $z(\cdot)$ will satisfy XIR and $\mathrm{X}_\varepsilon E$. Thus we have the following result.

Corollary 1: Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and suppose that for each θ, there exists a Pareto efficient, strictly individually rational CIE allocation for the CIE $e(\theta)$. 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 a PIE allocation $z(\cdot)$ for the PIE $(\{e(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}, P)$ satisfying XIR, XIC and $\mathrm{X}_\varepsilon E$.

While the details of the proof of Theorem 1 are somewhat complicated, we will provide an informal presentation of the basic ideas. Suppose that $\mathcal{A} = \{\zeta(\theta)\}_{\theta\in\Theta}$ is a collection of CIE allocations and suppose that $\varepsilon > 0$. The first step is to show that there exists a positive number $k(\varepsilon, \mathcal{A})$ and 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) $u_i(z_i(\theta, t_i); \theta) \approx u_i(\zeta_i(\theta); \theta)$ for all $t_i \in T_i$ and all $\theta \in \Theta$
(iii) for each $t_i, t_i' \in T_i$ with $t_i \neq t_i'$,

$$\sum_\theta \left[u_i(z_i(\theta, t_i); \theta) - u_i(z_i(\theta, t_i'); \theta)\right] P(\theta|t_i) \geq k(\varepsilon, \mathcal{A}) \min_i \Lambda_i^P$$

We call a collection with these properties a *quasi-mechanism* and quasi-mechanisms play a fundamental role in our approach. (For a formal definition, see the appendix.) In our model, the final allocation is a function of t (the vector of announced types) but cannot depend on θ; if the allocation could depend on θ, the mechanism design problem would be trivial. Suppose for a moment, though, that we could let allocations depend on θ, and that $\{\{z_i(\theta, t_i)\}_{(t_i,\theta)\in T_i\times\Theta}\}_{i\in N}$ satisfies conditions (i)-(iii) of the definition. If agent i announces $t_i' \in T_i$, then he receives the bundle $z_i(\theta, t_i')$. If i's true type is t_i, he (typically) does not know θ and, from his point of view, he faces a lottery (i.e., a random variable) $z_i(\tilde{\theta}, t_i')$ with expected payoff

$$\sum_\theta u_i(z_i(\theta, t_i'); \theta) P(\theta|t_i).$$

Condition (i) guarantees that, for each θ, the collection $\{z_i(\theta, t_i)\}_{i\in N}$ is a complete information allocation for $e(\theta)$ irrespective of the announced vector of types. The inequality in Condition (ii) says that, for each θ and t_i, $u_i(z_i(\theta, t_i); \theta) \approx u_i(\zeta_i(\theta); \theta)$ so the complete information allocation $\{z_i(\theta, t_i)\}_{i \in N}$ cannot be "much worse" than $\zeta_i(\theta)$. Finally, Condition (iii) is an incentive compatibility condition: when given a choice of the lotteries $\{z_i(\tilde{\theta}, t_i')\}_{t_i' \in T_i}$, agent i will choose $z_i(\tilde{\theta}, t_i)$ if he observes signal t_i. Furthermore, the expected loss associated with any announcement different from t_i is at least $k(\varepsilon, \mathcal{A}) \min_i \Lambda_i^P$.

Since a mechanism for a PIE cannot, of course, prescribe allocations that depend on θ, the second step uses the quasi-mechanism to construct a mechanism for the PIE with the desired properties. To define this mechanism, let

$$A_k = \{t \in T^* | ||P(\cdot|t) - \chi_{\theta_k}|| \leq \max_i \mu_i^P\}$$

and let $A_0 = T \backslash [\cup_{k \geq 1} A_k]$. If $\max_i \mu_i^P$ is sufficiently small, the collection $\{A_0, A_1, .., A_m\}$ will be a partition of T. Define a PIE allocation $z(\cdot)$ as follows:

$$\begin{aligned} z_i(t) &= z_i(\theta_k; t_i) \text{ if } t \in A_k \text{ and } k \geq 1 \\ &= w_i \text{ if } t \in A_0 \end{aligned}$$

and consider incentive compatibility. For each $i \in N$ and each $t_i \in T_i$, it can be shown that

$$\sum_{k=1}^{m} \sum_{t_{-i}} [u_i(z_i(t_{-i}, t_i); \theta_k) - u_i(z_i(t_{-i}, t_i'); \theta_k)] P(\theta, t_{-i}|t_i) \geq k(\varepsilon, \mathcal{A}) \min_i \Lambda_i^P - K_1 \mu_i^P - K_2 \nu_i^P]$$

where K_1 and K_2 are positive constants depending on $\{e(\theta)\}_{\theta \in \Theta}, \mathcal{A}$ and ε but not on P. It follows that, given $\mathcal{A}$ and ε, we can choose δ small enough so that we obtain IC whenever P satisfies the conditions of the theorem.

Two remarks are in order. First, the distribution P in the statement of Theorem 1 is not required to be in $\Delta^*_{\Theta \times T}$ so that $k(\varepsilon, \mathcal{A}) \min_i \Lambda_i^P$ need not be positive. In the replica theorem below, however, we will require that $P \in \Delta^*_{\Theta \times T}$. Second, this theorem does not cover the case in which the random variable $\tilde{\theta}$ and the random vector $\tilde{t}$ are stochastically independent except in uninteresting cases. If $\tilde{\theta}$ and $\tilde{t}$ are independent, then each $\Lambda_i^P = 0$. However, the distributions $P_\Theta(\cdot|t)$ are all equal and unless this common conditional distribution is degenerate, the theorem does not say anything about the existence of a mechanism with the desired properties. In the independent case, however, there *is* a simple mechanism satisfying XIR, IC and $\mathrm{X}_0 E$. Let $\overline{e}$ denote the CIE with endowments w_i with utilities $\overline{u}_i$ where

$$\overline{u}_i(x_i) = \sum_{\theta \in \Theta} u_i(x_i; \theta) P(\theta).$$

Choose an individually rational, Pareto efficient allocation $\overline{x}$ for $\overline{e}$ and define a mechanism $z(\cdot)$ as follows:

$$z_i(t) = \overline{x}_i \text{ for all } t \in T \ .$$

Then $z(\cdot)$ satisfies XIR, IC and X_0E. Aesthetically, it is desirable to have a result that would include the case of negligible aggregate uncertainty (Corollary 1) and the independent case as special cases. This will be accomplished by the generalization to nonnegligible aggregate uncertainty given in section 6 below.

5.5 Nearly Redundant Information

There are two natural economic problems to which the theorem above applies. The first is the case in which agents have nearly redundant information as in the example of section 4.1. Suppose that there is a finite number of agents, each of whom receives a noisy signal of the true state of nature. When an agent's information is sufficiently correlated with that of some subset of the other agents, he will be informationally small. If, in addition, there is little aggregate uncertainty, the theorem applies.

5.6 The Replica Problem

In the presence of a large number of agents, we might expect any single agent to be informationally small, and replica economies are a natural framework in which to investigate this conjecture.

5.6.1 Notation and Definitions:

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

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

(2) $u_{is}(z, \theta) = u_i(z, \theta)$ for all $z \in \Re^\ell_+, i \in N$ and $s \in J_r$.

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

$$\zeta^r_{is}(\theta) = \zeta_i(\theta) \text{ for each } (i, s) \in N \times J_r$$

5.7 Replica Economies and the Replica Theorem

Definition: A sequence of replica economies $\{(\{e^r(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}^r, P^r)\}_{r=1}^{\infty}$ 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 $(\theta, t_1, .., t_n) \in \Theta \times T$,

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

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

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

are independent given $\widetilde{\theta} = \theta$.

(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 $\widetilde{\theta}$. As r increases each agent is becoming "small" in the economy in terms of endowment, and we will 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 $\widetilde{\theta}$.

Theorem 2: Let $\Theta = \{\theta_1, .., \theta_m\}$. Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and suppose that $\mathcal{A} = \{\zeta(\theta)\}_{\theta\in\Theta}$ is a collection of associated CIE allocations with $\zeta_i(\theta) \neq 0$ for each i and θ. For every $\varepsilon > 0$, there exists a $\delta > 0$ such for each $r \geq 1$ and each $P^r \in \Delta_{\Theta\times T^r}$ satisfying

$$\max_i \mu_i^{P^r} \leq \delta \min_i \Lambda_i^{P^r}$$

$$\max_i \nu_i^{P^r} \leq \delta \min_i \Lambda_i^{P^r},$$

there exists an incentive compatible PIE allocation $z^r(\cdot)$ for the PIE $e^r = (\{e^r(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}^r, P^r)$ and a collection $A^r_1, .., A^r_m$ of disjoint subsets of $(T^r)^*$ such that $\text{Prob}\{\tilde{t}^r \in \cup_{k=1}^m A^r_k\} \geq 1 - \varepsilon$ and, for all $k \in J_m$ and all $t \in A^r_k$,

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

(ii) For all $(i, s) \in N \times J_r$,

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

In the statement of Theorem 2 above, we have made no special assumptions regarding P^r. We now state an analogue of Corollary 1 for replica economies.

Corollary 2: Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and suppose that for each θ, there exists a Pareto efficient, strictly individually rational CIE allocation for the CIE $e(\theta)$. Let $\{(\{e^r(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}^r, P^r)\}_{r=1}^{\infty}$ be a conditionally independent sequence and suppose that each $u_i(\cdot;\theta)$ is concave. Then for every $\varepsilon > 0$, there exists an integer $\hat{r} > 0$ such that for all $r > \hat{r}$, there exists an allocation z^r for the PIE $(\{e^r(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}^r, P^r)$ which satisfies IC, XIR and $X_\varepsilon E$.

We leave the proofs of Theorem 2 and Corollary 2 to the appendix but we can present the ideas. The proof of Theorem 2 is identical to that of Theorem 1 after showing that δ can be chosen to be independent of r. In fact, Theorem 1 is the special case of Theorem 2 in which $r = 1$. To deduce Corollary 2, we first show that $\nu_i^{P^r}$ and $\nu_i^{P^r}$ converge to zero as $r \to \infty$ (this is a consequence of the law of large numbers.) Hence, aggregate uncertainty and informational influence are small for large r. Next, we observe that $\Lambda_i^{P^r}$ is independent of r in a conditionally independent sequence and $\Lambda_i^{P^r} > 0$ since $P^r \in \Delta^*_{\Theta\times T^r}$. Let $\mathcal{A} = \{\zeta(\theta)\}_{\theta\in\Theta}$ be a collection and suppose that for each θ, $\zeta(\theta)$ is a Pareto efficient, strictly individually rational CIE allocation for the CIE $e(\theta)$. Applying Theorem 2, we can find, for all sufficiently large r, an incentive compatible mechanism satisfying conditions (i) and (ii) of Theorem 2 for the collection $\mathcal{A} = \{\zeta(\theta)\}_{\theta\in\Theta}$. Ex post individual rationality follows from the strict individual rationality of each $\zeta(\theta)$ in the CIE $e(\theta)$. Ex post $\varepsilon-$ Efficiency follows from the concavity assumption and the Pareto efficiency of $\zeta(\theta)$ in the CIE $e(\theta)$.

6 The General Theorem: Nonnegligible Aggregate Uncertainty

In this section, we will extend our results for the case of negligible aggregate uncertainty above to the case in which there may be nonnegligible aggregate uncertainty. We will first review the logic of the result for negligible aggregate uncertainty so as to identify the problems that may arise in the more general case. In the presence of negligible aggregate uncertainty, we identified a particular set of probability distributions $\{\chi_\theta\}_{\theta\in\Theta}$ on Θ with the property that, for most $t \in T$, there exists a θ such that $P(\cdot|t) \approx \chi_\theta$. When agent i is informationally small, there is high probability that for any two types t_i and $t_i^{'}$ of agent i, the conditional distributions on Θ given these types, $P(\cdot|t_{-i}, t_i)$ and $P(\cdot|t_{-i}, t_i^{'})$, are close. Given a distribution P with these properties and given a set of individually rational, Pareto efficient CIE allocations $\{\zeta(\theta)\}_{\theta\in\Theta}$, we constructed an allocation z satisfying $z_i(t) \approx \zeta_i(\theta)$ if $P(\cdot|t)$ is close to χ_θ, and $z_i(t) = w_i$ otherwise. The allocation z constructed in this way is approximately ex post efficient because of the assumed negligible aggregate uncertainty and the efficiency of $\{\zeta(\theta)\}_{\theta\in\Theta}$, while incentive compatibility for agent i will follow from the individual rationality of $\{\zeta(\theta)\}_{\theta\in\Theta}$, negligible aggregate uncertainty and informational smallness.

When we drop the assumption of negligible aggregate uncertainty, the set of probability distributions $\{\chi_\theta\}$ on Θ will not have the properties that allowed us to construct the desired allocation. However, other sets of distributions on Θ may permit an analogous construction. Consider the following example. There are two equally likely states of the world, θ_1 and θ_2 and there are 3 agents, each of whom receives one of two signals, a or b. Suppose that the signals are perfectly correlated. That is, all agents receive the same signal, either a or b, and suppose that $P(\theta_1|a) = .3$ and $P(\theta_1|b) = .7$. Negligible aggregate uncertainty clearly fails, since given the two possible signal vectors, (a, a, a) and (b, b, b), the conditional distributions on Θ are $(.3, .7)$ and $(.7, .3)$ respectively.

Nevertheless, we can construct an allocation that is nearly efficient for each $t \in T^*$ and incentive compatible. Consider a set of distributions on Θ defined by $\mathcal{P} \equiv \{\pi_1, \pi_2\}$ where $\pi_1 = (.3, .7)$ and $\pi_2 = (.7, .3)$. 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 $\zeta(\pi)$ and $\zeta(\pi)$ be allocations with $\zeta(\pi_k)$ 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. The allocation z will be individually rational and efficient for each t by construction, and z will be incentive compatible since any misrepresentation by agent i will change the bundle he gets from one which is strictly better than his initial endowment to his initial endowment.

In this construction, we are treating the probability distributions in $\mathcal{P}$ as something like quasi-states. Learning which of these distributions is the "true" conditional distribution over Θ is all that one can hope for given the information structure. This example illustrates a condition that can serve as an analogue to negligible aggregate uncertainty. Suppose for a given information structure, we can find a set of distributions $\mathcal{P} = \{\pi_1, \pi_2, ..., \pi_m\}$ 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 could then mimic the construction in the example above and we will formalize this next.

Let m be a positive integer, let $Q \in \Delta_{J_m \times T}$ (recall that $J_m = \{1, ..m\}$ and let $\chi_k \in \Delta_{J_m}$ denote the measure that puts probability one on k. Next, let $\mathcal{P} = \{\pi_1, .., \pi_m\}$ be a collection of measures in Δ_Θ and define a measure $\mathcal{P} * Q \in \Delta_{\Theta \times T}$ as

$$P(\theta, t) = \sum_{k=1}^{m} \pi_k(\theta)Q(k, t).$$

Let

$$\Lambda_i^Q = \min_{t_i \in T_i} \min_{t'_i \in T_i \setminus t_i} \frac{||Q_{J_m}(\cdot|t_i)||_2 ||Q_{J_m}(\cdot|t'_i)||_2 - \sum_{k \in J_m} [Q_{J_m}(k|t_i) Q_{J_m}(k|t'_i)]}{||Q_{J_m}(\cdot|t'_i)||_2}$$

where $||\cdot||$ denotes the 2-norm. Abusing notation slightly, define

$$\mu_i^Q = \max_{t_i \in T_i} \inf\{\varepsilon > 0 | \text{Prob}\{\tilde{t} \in T^* \text{ and } ||Q_{J_m}(\cdot|\tilde{t}) - \chi_k|| > \varepsilon \text{ for all } k \in J_m | \tilde{t}_i = t_i\} \leq \varepsilon\}$$

Note that, since the marginals of $P = \mathcal{P} * Q$ and Q on T are the same, the symbol Prob$\{\cdot\}$ in the definition of μ_i^Q above refers to this common marginal.

For a collection $\mathcal{P} = \{\pi_1, .., \pi_m\}$ and a collection of CIE's $\{e(\theta)\}_{\theta \in \Theta}$, let $\{e(\pi_k)\}_{k \in J_m}$ denote a new collection where $e(\pi_k)$ is the PIE in which agent i has endowment w_i and utility function $v_i(\cdot|\pi_k)$ defined by

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

We can now state generalizations of Theorem 1 and Corollary 1.

Theorem 3 : Let $\{e(\theta)\}_{\theta \in \Theta}$ be a collection of CIE's and let $\mathcal{P} = \{\pi_1, .., \pi_m\}$ be a collection of measures in Δ_Θ. Furthermore, suppose that $\mathcal{A} = \{\zeta(k)\}_{k \in J_m}$ is a collection where each $\zeta(k)$ is a CIE allocation for the CIE $e(\pi_k)$ satsfying $\zeta_i(\pi_k) \neq 0$ for each i and k. For every $\varepsilon > 0$, there exists a $\delta > 0$ such that, whenever $Q \in \Delta_{J_m \times T}$, $P = \mathcal{P} * Q$ and satisfies

$$\max_i \hat{\mu}_i^Q \leq \delta \min_i \Lambda_i^Q$$

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

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

(i) $Q_{J_m}(k|t) \geq 1 - \varepsilon$

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

$$v_i(z_i(t); \pi_k) \geq v_i(\zeta_i(k); \pi_k) - \varepsilon.$$

When $P = \mathcal{P} * Q$, it is not difficult to show that $\hat{\mu}_i^P$ will be small when $\hat{\mu}_i^Q$ is small. In fact, we could define a notion of informational size $\hat{\nu}_i^Q$ with respect to Q in an obvious way and, according to this definition, $\hat{\nu}_i^P$ will be small when $\hat{\nu}_i^Q$ is small if $P = \mathcal{P} * Q$. We have chosen to work directly with $\hat{\nu}_i^P$ (in condition (ii) above) since we believe that informational size is most naturally defined in term of the distribution P rather than the "auxiliary" distribution Q.

Corollary 3: Let $\{e(\theta)\}_{\theta \in \Theta}$ be a collection of CIE's and let $\mathcal{P} = \{\pi_1, .., \pi_m\}$ be a collection of measures in Δ_Θ. Furthermore, suppose that for each $k \in J_m$, there exists a Pareto efficient, strictly individually rational CIE allocation for the CIE $e(\pi_k)$. Then for every $\varepsilon > 0$, there exists a $\delta > 0$ such that, whenever $Q \in \Delta_{J_m \times T}$, $P = \mathcal{P} * Q$ and satisfies

$$\max_i \hat{\mu}_i^Q \leq \delta \min_i \Lambda_i^Q$$

$$\max_i \hat{\nu}_i^P \le \delta \min_i \Lambda_i^Q$$

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

If $\Theta = \{\theta_1, .., \theta_m\}$, then we obtain Theorem 1 from Theorem 3 and Corollary 1 from Corollary 3 by setting $\pi_k = \chi_{\theta_k}$ and $Q(k,t) = P(\theta_k, t)$ for each k and t. We also obtain the independent case as the special case in which $m = 1$. Note that $m = 1$ implies that $\hat{\nu}_i^P = \hat{\mu}_i^Q = \Lambda_i^Q = 0$ for all i.

7 Discussion

1. The role played by the variability of the conditional distributions in the construction of quasi-mechanisms is reminiscent of the work on full surplus extraction (see, e.g., Cremer and McLean (1985,1988) and McAfee and Reny (1992)). In a finite type, private values framework, Cremer and McLean (1985, 1988) demonstrate 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 the present work, the collection $\{P_\Theta(\cdot|t_i)\}_{t_i\in T_i}$ need not be linearly independent but the "closeness" of the members of $\{P_\Theta(\cdot|t_i)\}_{t_i\in T_i}$ is an important issue. Since Pareto efficiency and full extraction are in some sense related, it is reasonable to conjecture that full extraction might be possible in a common value mechanism design framework satisfying assumptions like (e.g.) those of Theorem 1. This is indeed the case and in McLean and Postlewaite (1999), we address this question and provide a more precise comparison with the results of Cremer and McLean.

There are other technical differences between the setup of Cremer and McLean and that of this paper. First, Cremer and McLean assume, as does much of the mechanism design literature, that agents have quasi-linear utilities. In addition, Cremer and McLean ignore any "endowment" constraints and allow agents to make possibly large payments. In this paper, we do not restrict attention to quasi-linear utilities and we have assumed that all outcomes, both in and out of equilibrium, are feasible for the given initial endowments.

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 provide agents with a strict incentive to truthfully reveal their

types and we will provide an example that demonstrates how this might be accomplished. In the example, every signal that an agent receives will generate a posterior distribution on Θ that is the same as the prior, yet there is no aggregate uncertainty. Although by itself, each agent's information provides no information about the state of nature, we will illustrate how punishments that depend on other agents' announcements, in addition to his own announcement, can be constructed so as to induce a strict incentive for truthful announcement.

There are four agents and two states of nature, θ_1 and θ_2. The probability that state θ_k is the true state is $1/2$ for $k = 1, 2$. Each agent will receive a signal in the set $\{a, b\}$; that is $T_i = \{a, b\}, i = 1, 2, 3, 4$. The information structure is as follows. Suppose that the true state of nature is θ_1. With probability $1/3$, all four agents receive signal a and with probability $2/3$, exactly one agent will receive signal a, with each of the agents being equally likely to receive the signal a. Symmetrically, if the state of nature is θ_2, then with probability $1/3$ all agents will receive signal b and with probability $2/3$ exactly one agent will receive signal b, with each of the agents being equally likely to receive the signal b.

For any agent, $P(a|\theta_1) = 1/2$, and similarly $P(a|\theta_2) = 1/2$. Hence, $P(\theta_i|a) = P(\theta_i|b) = 1/2$, $i = 1, 2$. For this information structure, an individual agent's signal provides no information. It is clear, however, that there is no aggregate uncertainty. If all agents receive signal a or exactly one agent receives signal a, then $\theta = \theta_1$. If all agents receive signal b or exactly one agent receives signal b, then $\theta = \theta_2$.

We will next construct punishments for agent i, depending on all agents' announcements, that will provide a strict incentive for i to truthfully announce his information. Consider a collection $\{\tau_i(t_{-i}, t_i)\}$ defined as follows. Give agent i no punishment if all agents announce the same signal and a small punishment otherwise. An agent who truthfully reports his type will avoid punishment with probability $1/3$, while misreporting results in punishment with probability 1.

In this example, agents' posteriors on the state of nature are independent of their signal, but their posteriors on the signals that other agents receive are not independent of their own signal. We could generalize slightly our results if, instead of assuming $P_\Theta(\cdot|t_i) \neq P_\Theta(\cdot|t_i^{'})$ for each i and $t_i, t_i^{'}$, we assumed $P_{T_{-i}}(\cdot|t_i) \neq P_{T_{-i}}(\cdot|t_i^{'})$ for each i and $t_i, t_i^{'} \in T_i$ (where $P_{T_{-i}}(\cdot|t_i)$ is the conditional distribution on T_{-i} given t_i). 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 punishments depending on θ. In other words, the mechanisms constructed in this way are somewhat more complicated than those constructed in this paper.

Agents in the example above are not informationally small. In fact, with probability 1 any agent who misreports his signal will change the estimated state. However, it is easy to see that the method of constructing punishments to induce truthful revelation is unrelated to this and could be applied regardless of agents' informational size.

3. 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. In an interesting work along these lines, Krasa and Shafer (1998) show that a variant of our notion of informational smallness is both necessary and sufficient for a particular robustness test of Walrasian equilibria.

4. Suppose that $(\{e(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}, P)$ is a PIE. If some agent is "informationally large," then our Theorem 1 will 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 get a signal s_1 or s_2, with $P(s_i|\theta_i) = q$ where $.5 < q < 1$. 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 that the most likely state; this is exactly what we did in Theorem 2.

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 agent 1's signal, since his is the only non-noisy signal. It is clear that with this modification, our Theorem 1 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 agents' 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.

5. 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 in θ, 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 a problem for further research.

6. 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 t_i for that agent, suppose that instead of receiving t_i, agent i receives one of two signals, t_i' or t_i''. Furthermore, suppose that $P_\Theta(\cdot|t_i') = P_\Theta(\cdot|t_i'') = P_\Theta(\cdot|t_i)$. That is, we have taken the original PIE 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. One can think of this as agent i flipping a coin after he receives signal 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 $\min_i \Lambda_i^P = 0$. 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, t_i. More generally, we could partition agent's type set into subsets, treating each subset as an "approximate 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.[2]

[2] **We thank Ichiro Obara for making this point.**

8 Bibliography

Cremer, J. and R. McLean (1985), "Optimal Selling Strategies under Uncertainty for a Monopolist When Demands are Interdependent," *Econometrica* **53**, 345-61.

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

Gul, F. and A. Postlewaite (1992), "Asymptotic Efficiency in Large Economies with Asymmetric Information," *Econometrica* **60**, 1273-1292.

Krasa, S. and W. Shafer (1998), "Informational Robustness of Competitive Equilibria," mimeo, 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 (1999), Informational Size and Mechanism Design with Common Values, mimeo.

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

9 Appendix:

9.1 Preliminary Definitions and Lemmas:

Let Δ_{J_m} denote the set of probability measures on the finite set $J_m = \{1, .., m\}$.

Definition: Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's. Suppose that $\mathcal{P} = \{\pi_1, .., \pi_m\}$ is collection of measures in Δ_Θ and suppose that $Q \in \Delta_{J_m\times T}$ with conditionals $Q_{J_m}(\cdot|t_i) \in \Delta_{J_m}$ for all i and $t_i \in T_i$. Furthermore, suppose that $\mathcal{A} = \{\zeta(k)\}_{k\in J_m}$ is a collection of CIE allocations such that, for each $k \in J_m$, the allocation $\{\zeta_i(k)\}_{i\in N}$ is a CIE allocation for $e(\pi_k)$ with $\zeta_i(k) \neq 0$ for all $k \in J_m$ and for all i. If $\eta \geq 0$ and $\lambda \geq 0$, then a collection $\{\{z_i(k, t_i)\}_{(t_i,\theta)\in T_i\times\Theta}\}_{i\in N}$ is a $(Q, \mathcal{P}, \eta, \lambda, \mathcal{A})$ quasi-mechanism for $\{e(\theta)\}_{\theta\in\Theta}$ if

(i) $z_i(k, t_i) \in \Re^\ell_+$ and $\sum_{i\in N}(z_i(k, t_i) - w_i) \leq 0$ for all $t_i \in T_i$ and all $k \in J_m$.

(ii) $v_i(\zeta_i(k); \pi_k) \geq v_i(z_i(k, t_i); \pi_k) \geq v_i(\zeta_i(k); \pi_k) - \eta$ for all $t_i \in T_i$ and all $k \in J_m$

(iii) for each $t_i, t'_i \in T_i$

$$\sum_{k\in J_m} [v_i(z_i(k, t_i); \pi_k) - v_i(z_i(k, t'_i); \pi_k)]\, Q_{J_m}(k|t_i) \geq \lambda$$

Definition: If $Q \in \Delta_{J_m\times T}$ is a measure with conditionals $Q_{J_m}(\cdot|t_i) \in \Delta_{J_m}$ for all i and $t_i \in T_i$, then define

$$\Lambda^Q_i(t'_i|t_i) = \frac{||Q_{J_m}(\cdot|t_i)||_2||Q_{J_m}(\cdot|t'_i)||_2 - \sum_{k\in J_m}[Q_{J_m}(k|t_i)Q_{J_m}(k|t'_i)]}{||Q_{J_m}(\cdot|t'_i)||_2}$$

and

$$\Lambda^Q_i = \min_{t_i} \min_{t'_i\neq t_i} \Lambda^Q_i(t'_i|t_i)$$

Definition: Suppose that $\mathcal{P} = \{\pi_1, .., \pi_m\}$ is a collection of measures in Δ_Θ and suppose that $\mathcal{A} = \{\zeta(k)\}_{k\in J_m}$ is a collection of CIE allocations such that, for each $k \in J_m$, the allocation $\{\zeta_i(k)\}_{i\in N}$ is a CIE allocation for $e(\pi_k)$ with $\zeta_i(k) \neq 0$ for all $k \in J_m$ and for all i. For each $\eta \geq 0$, let

$$c(\eta, \mathcal{P}, \mathcal{A}) = \min_i \min_k \{v_i(\zeta_i(k); \pi_k) - v_i(\beta_i(k)\zeta_i(k); \pi_k)\}$$

where

$$\beta_i(k) = \min\{\beta | 1/2 \leq \beta \leq 1, v_i(\zeta_i(k); \pi_k) - v_i(\beta\zeta_i(k); \pi_k) \leq \eta\}.$$

Lemma A: Let $(\{e(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}, P)$ be a PIE and suppose that $Q \in \Delta_{J_m\times T}$ with conditionals $Q_{J_m}(\cdot|t_i) \in \Delta_{J_m}$ for all i and $t_i \in T_i$. Furthermore, suppose that $\mathcal{P} = \{\pi_1, .., \pi_m\}$ is a collection of measures in Δ_Θ and suppose that $\mathcal{A} = \{\zeta(k)\}_{k\in J_m}$ is a collection of CIE allocations such that, for each $k \in J_m$, the allocation $\{\zeta_i(k)\}_{i\in N}$ is a CIE allocation for $e(\pi_k)$ with $\zeta_i(k) \neq 0$ for all $k \in J_m$ and for all i. Then for all $\eta \geq 0$, there exists a $(Q, \mathcal{P}, \eta, \lambda, \mathcal{A})$ quasi-mechanism for $\{e(\theta)\}_{\theta\in\Theta}$ with

$$\lambda = c(\eta, \mathcal{P}, \mathcal{A}) \min_i \Lambda_i^Q$$

Proof: Suppose that $Q \in \Delta_{J_m\times T}$ with conditionals $Q_{J_m}(\cdot|t_i) \in \Delta_{J_m}$ for all i and $t_i \in T_i$. Next, define

$$\alpha_i(k, t_i) = \frac{Q_{J_m}(k|t_i)}{||Q_{J_m}(\cdot|t_i)||_2}$$

for each $k \in J_m$. Hence,

$$\Lambda_i^Q(t_i'|t_i) = \sum_{k\in J_m} [Q_{J_m}(k|t_i)\alpha_i(k, t_i)] - \sum_{k\in J_m} [Q_{J_m}(k|t_i)\alpha_i(k, t_i')] \geq 0.$$

Let $\mathcal{A} = \{\zeta(k)\}_{k\in J_m}$ be a collection of CIE allocations with $\zeta_i(k) \neq 0$ for all $k \in J_m$ and for all i. If $\eta = 0$, then $c(\eta, \mathcal{A}, \mathcal{P}) = 0$ and the result is trivial (let $z_i(k, t_i) = \zeta_i(k)$ and $\lambda = 0$). So suppose that $\eta > 0$. From the continuity and monotonicity assumptions, it follows that $c(\eta, \mathcal{A}, \mathcal{P}) > 0$. For each i, t_i and k, there exists a number $\tau_i(k, t_i) \geq 0$ such that

$$v_i((1 + \tau_i(k, t_i))\beta_i(k)\zeta_i(k); \pi_k) - v_i(\beta_i(k)\zeta_i(k); \pi_k) = c(\eta, \mathcal{A}, \mathcal{P})\alpha_i(k, t_i).$$

[This is possible because $0 \leq c(\eta, \mathcal{A}, \mathcal{P})\alpha_i(k, t_i) \leq c(\eta, \mathcal{A}, \mathcal{P})$]. Furthermore, $(1 + \tau_i(k, t_i))\beta_i(k) \leq 1$. [If $(1 + \tau_i(k, t_i))\beta_i(k) > 1$, then monotonicity implies that

$$\begin{aligned} v_i((1 + \tau_i(k, t_i))\beta_i(k)\zeta_i(k); \pi_k) - v_i(\beta_i(k)\zeta_i(k); \pi_k) &> v_i(\zeta_i(k); \pi_k) - v_i(\beta_i(k)\zeta_i(k); \pi_k) \\ &\geq c(\eta, \mathcal{A}, \mathcal{P}) \\ &\geq c(\eta, \mathcal{A}, \mathcal{P})\alpha_i(k, t_i) \end{aligned}$$

a contradiction.] Defining

$$z_i(k, t_i) = (1 + \tau_i(k, t_i))\beta_i(k)\zeta_i(k)$$

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

$$z_i(k, t_i) \in \Re_+^\ell \text{ and } \sum_{i\in N}(z_i(k, t_i) - w_i) \leq 0$$

and part (i) of the definition is satisfied. Furthermore,

$$v_i(z_i(k,t_i);\pi_k) - v_i(\beta_i(k)\zeta_i(k);\pi_k) = c(\eta,\mathcal{A},\mathcal{P})\alpha_i(k,t_i)$$

for all $t_i \in T_i$ and all $k \in J_m$. Therefore,

$$v_i(\zeta_i(k);\pi_k) \geq v_i(z_i(k,t_i);\pi_k) = v_i(\beta_i(k)\zeta_i(k);\pi_k) + c(\eta,\mathcal{A},\mathcal{P})\alpha_i(k,t_i) \geq v_i(\zeta_i(k);\pi_k) - \eta$$

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

$$\begin{aligned}
& \sum_{k\in J_m} [v_i(z_i(k,t_i);\pi_k) - v_i(z_i(k,t_i');\pi_k)]Q_{J_m}(k|t_i) \\
= & \sum_{k\in J_m} [c(\eta,\mathcal{A},\mathcal{P})\alpha_i(k,t_i) - c(\eta,\mathcal{A},\mathcal{P})\alpha_i(k,t_i')]Q_{J_m}(k|t_i) \\
= & \ c(\eta,\mathcal{A},\mathcal{P}) \sum_{k\in J_m} [\alpha_i(k,t_i) - \alpha_i(k,t_i')]Q_{J_m}(k|t_i) \\
= & \ c(\eta,\mathcal{A},\mathcal{P})\Lambda_i^Q(t_i'|t_i) \\
\geq & \ c(\eta,\mathcal{A},\mathcal{P}) \min_i \Lambda_i^Q
\end{aligned}$$

9.2 Conditioning Systems:

Definition: Let m be a positive integer. A *conditioning system* for T is a triple $\mathcal{S} = (\Pi,\mathcal{P},Q)$ where

(i) $\Pi = \{A_0, A_1, .., A_m\}$ is a partition of T with $\emptyset \neq A_k \subseteq T^*$ for each $k \geq 1$. (We allow for the possibility that $A_0 = \emptyset$.)

(ii) $\mathcal{P} = \{\pi_1, .., \pi_m\}$ a collection of measures in Δ_Θ.

(iii) $Q \in \Delta_{J_m \times T}$ with conditionals $Q_{J_m}(\cdot|t_i) \in \Delta_{J_m}$ for all i and $t_i \in T_i$.

Definition: Let $P \in \Delta_{\Theta\times T}$ and let $\mathcal{S} = (\Pi,\mathcal{P},Q)$ be a conditioning system. Let Prob$\{\cdot\}$refer to the measure P.

The $P-$*mesh* of $\mathcal{S}$ is defined as

$$\gamma^P(\mathcal{S}) = \max_{1\leq k\leq m} \max_{t\in A_k} ||\pi_k - P_\Theta(\cdot|t)||$$

The $P-$*coverage* of agent i at t_i in $\mathcal{S}$ is defined as

$$\mu_i^P(\mathcal{S}|t_i) = Prob\{\tilde{t} \in A_0|\tilde{t}_i = t_i\}$$

The $P-$*influence* of agent i at t_i in $\mathcal{S}$ is defined as

$$\nu_i^P(\mathcal{S}|t_i) = \max_{t_i'\in T_i} \sum_{k=1}^m Prob\{(\tilde{t}_{-i},t_i) \in A_k \text{ and } (\tilde{t}_{-i},t_i') \notin A_k \cup A_0|\tilde{t}_i = t_i\}$$

The $P-accuracy$ of $\mathcal{S}$ is defined as

$$\sigma^P(\mathcal{S}) = \max_{i\in N}\max_{t_i\in T_i}\sum_{k=1}^{m}|\Pr ob\{\tilde{t}\in A_k|\tilde{t}_i = t_i\} - Q_{J_m}(k|t_i)|$$

Proposition A: Suppose that

(i) $(\{e(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}, P)$ is a PIE

(ii) $(\Pi, \mathcal{P}, Q)$ is a conditioning system for T

(iii) $\mathcal{A} = \{\zeta(k)\}_{k\in J_m}$ is a collection where each $\zeta(k)$ is a CIE allocation for the CIE $e(\pi_k)$ satisfying $\zeta_i(\pi_k) \neq 0$ for each i and k.

(iv) $\{\{z_i(k,t_i)\}_{t_i\in T_i, k\in J_m}\}_{i=1}^{n}$ is a $(Q, \mathcal{P}, \eta, \lambda, \mathcal{A})$ quasi-mechanism for the PIE $\{e(\theta)\}_{\theta\in\Theta}$.

Let $z(\cdot)$ be an allocation for the PIE defined by

$$\begin{aligned} z_i(t) &= z_i(k;t_i) \text{ if } t\in A_k \text{ and } k\geq 1 \\ &= w_i \text{ if } t\in A_0 \end{aligned}$$

and define

$$M = \max_{\theta}\max_{i} u_i(\sum_{j=1}^{n} w_j;\theta).$$

Then for all $i\in N$ and all $t_i.t_i'\in T_i$, $z(\cdot)$ satisfies the following *approximate incentive compatibility* condition:

$$\sum_{k=1}^{m}\sum_{t_{-i}}[u_i(z_i(t_{-i},t_i);\theta_k) - u_i(z_i(t_{-i},t_i');\theta_k)]P(\theta,t_{-i}|t_i) \geq \lambda - 2M[\gamma_i(\mathcal{S}|t_i) + \mu(\mathcal{S}) + \sigma(\mathcal{S}) + \nu_i(\mathcal{S}|t_i)].$$

Proof: First, we show that for all $t\in T$ and all $\theta\in\Theta$,

$$u_i(z_i(t);\theta)\leq M.$$

Suppose that $t\in A_k$. From the definition of quasimechanism, it follows that

$$\zeta_i(k)\leq\sum_{j=1}^{n} w_j$$

for each i and k. This observation, together with the monotonicity assumption implies that

$$u_i(z_i(t);\theta) = u_i(z_i(k,t_i);\theta)\leq u_i(\zeta_i(k);\theta)\leq M$$

for each i and k. If $t\in A_0$, then monotonicity implies that

$$u_i(z_i(t);\theta) = u_i(w_i;\theta)\leq M.$$

To prove approximate incentive compatibility, note that

$$
\begin{aligned}
& \sum_{\theta}\sum_{t_{-i}}[u_i(z_i(t_{-i},t_i);\theta) - u_i(z_i(t_{-i},t_i');\theta)]P(\theta,t_{-i}|t_i) \\
= & \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_0}} \sum_{\theta}[u_i(z_i(t_{-i},t_i);\theta) - u_i(z_i(t_{-i},t_i');\theta)]P(\theta,t_{-i}|t_i) \\
& + \sum_{k=1}^{m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_k}} \sum_{\theta}[u_i(z_i(k,t_i);\theta) - u_i(z_i(t_{-i},t_i');\theta)]P(\theta|t_{-i},t_i)P(t_{-i}|t_i) \\
\geq & -2M\mu_i(\mathcal{S}|t_i) + \sum_{k=1}^{m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_k}} \sum_{\theta}[u_i(z_i(k;t_i);\theta) - u_i(z_i(t_{-i},t_i');\theta)]P(\theta|t_{-i},t_i)P(t_{-i}|t_i)
\end{aligned}
$$

$$
\begin{aligned}
\geq & -2M[\mu_i(\mathcal{S}|t_i)+\gamma(\mathcal{S})] + \sum_{k=1}^{m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_k}} \sum_{\theta}[u_i(z_i(k;t_i);\theta) - u_i(z_i(t_{-i},t_i');\theta)]\pi_k(\theta)P(t_{-i}|t_i) \\
= & -2M[\mu_i(\mathcal{S}|t_i)+\gamma(\mathcal{S})] + \sum_{k=1}^{m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_k}} \sum_{\theta}[u_i(z_i(k;t_i);\theta) - u_i(z_i(k;t_i');\theta)]\pi_k(\theta)P(t_{-i}|t_i) \\
& + \sum_{k=1}^{m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_k}} \sum_{\theta}[u_i(z_i(k;t_i');\theta) - u_i(z_i(t_{-i},t_i');\theta)]\pi_k(\theta)P(t_{-i}|t_i) \\
\geq & -2M[\mu_i(\mathcal{S}|t_i)+\gamma(\mathcal{S})+\sigma(\mathcal{S})] + \sum_{k=1}^{m}[v_i(z_i(k;t_i);\pi_k) - v_i(z_i(k;t_i');\pi_k)]Q_{J_m}(k|t_i) \\
& + \sum_{k=1}^{m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_k \\ (t_{-i},t_i')\notin A_k\cup A_0}} \sum_{\theta}[u_i(z_i(k;t_i');\theta) - u_i(z_i(t_{-i},t_i');\theta)]\pi_k(\theta)P(t_{-i}|t_i) \\
\geq & \ \lambda - 2M[\mu_i(\mathcal{S}|t_i)+\gamma(\mathcal{S})+\sigma(\mathcal{S})+\nu_i(\mathcal{S}|t_i)]
\end{aligned}
$$

9.3 Proof of Theorem 3:

Let $\mathcal{P} = \{\pi_1, .., \pi_m\}$ be a collection of distinct measures in Δ_Θ and let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's. Furthermore, suppose that $\mathcal{A} = \{\zeta(k)\}_{k\in J_m}$ is a collection of CIE allocations such that, for each $k \in J_m$, $\zeta_i(k)$ is an allocation for $e(\pi_k)$ with $\zeta_i(k) \neq 0$ for all i. Choose $\varepsilon > 0$. Let

$$D(\mathcal{P}) = \min_{k} \min_{\substack{\ell \\ :\ell\neq k}} ||\pi_\ell - \pi_k||.$$

and

$$M = \max_\theta \max_i u_i(\sum_{j=1}^n w_j;\theta).$$

and choose δ so that

$$0 < \delta < \min\{\frac{c(\varepsilon,\mathcal{A},\mathcal{P})}{10M},\varepsilon,\frac{D(\mathcal{P})}{3}\}$$

Suppose that $Q \in \Delta_{J_m\times T}$ and $P = \mathcal{P} * Q$, let ν_i^P be defined as in section 5.1 and let μ_i^Q and Λ_i^Q be defined as in section 6. Finally, define $\hat{\mu}^Q = \max_i \mu_i^Q$, $\hat{\nu}^P = \max_i \nu_i^P$and $\Lambda^Q = \min_i \Lambda_i^Q$ and suppose that,

$$\begin{aligned} \hat{\nu}^P &\leq \delta\Lambda^Q \\ \hat{\mu}^Q &\leq \delta\Lambda^Q \ . \end{aligned}$$

For each $k \in J_m$, let

$$A_k = \{t \in T^*|||Q_{J_m}(\cdot|t) - \chi_k|| \leq \hat{\mu}^Q\}$$

and let

$$A_0 = T\backslash[\cup_{k\in J_m} A_k].$$

Since $\Lambda^Q \leq 1$, it follows that

$$\hat{\mu}^Q \leq \delta\Lambda^Q < \frac{D(\mathcal{P})}{3}\Lambda^Q \leq \frac{D(\mathcal{P})}{3}$$

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

The partition Π, together with $\mathcal{P} = \{\pi_1, .., \pi_m\}$ and Q , define a conditioning system $\mathcal{S} = (\Pi,\mathcal{P},Q)$. Applying Lemma 1, there exists a $(Q,\mathcal{P},\varepsilon,\lambda,\mathcal{A})$ quasimechanism $\{\{z_i(k,t_i)\}_{(t_i,\theta)\in T_i\times J_m}\}_{i\in N}$ with

$$\lambda = c(\varepsilon,\mathcal{A},\mathcal{P})\Lambda^Q$$

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

$$\begin{aligned} z_i(t) &= z_i(k,t_i) \text{ if } t \in A_k \\ &= w_i \text{ if } t \in A_0 \end{aligned}$$

Since the marginals of $P = \mathcal{P} * Q$ and Q on T are equal, it follows that

$$\mu_i^P(\mathcal{S}|t_i) \leq \hat{\mu}^Q \leq \delta\Lambda^Q$$

for all i and all $t_i \in T_i$. Next, we prove three claims.

Claim 1: For each i,

$$\sigma^P(\mathcal{S}) \leq 2\delta\Lambda^Q.$$

Proof of Claim 1: First, note that

$$Q_{J_m}(k|t_i) = \sum_{\ell \in J_m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_\ell}} Q(k|t_{-i},t_i)Q(t_{-i}|t_i) + \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_0}} Q(k,t_{-i}|t_i)$$

and

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

Therefore,

$$\begin{aligned}
|Q_{J_m}(k|t_i) - \text{Prob}\{\tilde{t} &\in A_k|\tilde{t}_i = t_i\}| \\
&= |[\sum_{\ell \in J_m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_\ell}} Q(k|t_{-i},t_i)Q(t_{-i}|t_i)] - [\sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_k}} Q(t_{-i}|t_i)] \\
&\quad +[\sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_0}} Q(k|t_{-i},t_i)]| \\
&= |[\sum_{\ell \in J_m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_\ell}} Q(k|t_{-i},t_i)Q(t_{-i}|t_i)] - [\sum_{\ell \in J_m} \chi_k(\theta_\ell) \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_\ell}} Q(t_{-i}|t_i)] \\
&\quad +[\sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_0}} Q(k|t_{-i},t_i)]| \\
&= |[\sum_{\ell \in J_m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_\ell}} [Q(k|t_{-i},t_i) - \chi_\ell(\theta_k)]Q(t_{-i}|t_i)] + [\sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_0}} Q(k|t_{-i},t_i)]|.
\end{aligned}$$

Hence,

$$\begin{aligned}
\sum_{k \in J_m} |Q_{J_m}(k|t_i) - \text{Prob}\{\tilde{t} &\in A_k|\tilde{t}_i = t_i\}| \\
&\leq \sum_{\ell \in J_m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_\ell}} \sum_{k \in J_m} |Q(k|t_{-i},t_i) - \chi_\ell(\theta_k)|Q(t_{-i}|t_i) + \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_0}} Q(t_{-i}|t_i) \\
&\leq \hat{\mu}^Q \sum_{\ell \in J_m} \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_\ell}} Q(t_{-i}|t_i) + \sum_{\substack{t_{-i} \\ :(t_{-i},t_i)\in A_0}} Q(t_{-i}|t_i) \\
&\leq \hat{\mu}^Q + \mu_i^Q \\
&\leq 2\hat{\mu}^Q \\
&\leq 2\delta\Lambda^Q.
\end{aligned}$$

Claim 2: For each i and each t_i,

$$\nu_i^P(\mathcal{S}|t_i) \leq \delta\Lambda^Q.$$

Proof of Claim 2: Define

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

and note that

$$\nu_i^P(\mathcal{S}|t_i) = \max_{t_i' \in T_i} \Pr\text{ob}\{\tilde{t} \in \Psi_i(t_i', t_i) | \tilde{t}_i = t_i\}.$$

Also, recall that

$$I^i_{\hat{\nu}^P}(t_i', t_i) = \{t_{-i} \in T_{-i} | (t_{-i}, t_i) \in T^*, (t_{-i}, t_i') \in T^* \text{ and } ||P_\Theta(\cdot|t_{-i}, t_i) - P_\Theta(\cdot|t_{-i}, t_i')|| > \hat{\nu}^P\}$$

We claim that $\Psi_i(t_i', t_i) \subseteq I^i_{\hat{\nu}^P}(t_i', t_i)$. To see this, suppose that $t_{-i} \in \Psi_i(t_i', t_i)$ but $t_{-i} \notin I^i_{\hat{\nu}^P}(t_i', t_i)$. 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 $\Lambda^Q \leq 1$, it follows that

$$\hat{\mu}^Q \leq \delta \Lambda^Q < \frac{D(\mathcal{P})}{3} \Lambda^Q \leq \frac{D(\mathcal{P})}{3}$$

and that

$$\hat{\nu}^P \leq \delta \Lambda^Q < \frac{D(\mathcal{P})}{3} \Lambda^Q \leq \frac{D(\mathcal{P})}{3}.$$

Therefore,

$$\begin{aligned} ||\pi_\ell - \pi_k|| &\leq ||P_\Theta(\cdot|t_{-i}, t_i) - \pi_\ell|| + ||P_\Theta(\cdot|t_{-i}, t_i) - P_\Theta(\cdot|t_{-i}, t_i')|| + ||P_\Theta(\cdot|t_{-i}, t_i') - \pi_k|| \\ &\leq \hat{\mu}^Q + \hat{\nu}^P + \hat{\mu}^Q \\ &< 3\frac{D(\mathcal{P})}{3} \\ &= D(\mathcal{P}) \end{aligned}$$

an impossibility. Hence, we conclude that

$$\begin{aligned} \nu_i^P(\mathcal{S}|t_i) &= \max_{t_i' \in T_i} Prob\{\tilde{t}_{-i} \in \Psi_i(t_i', t_i) | \tilde{t}_i = t_i\} \\ &\leq \max_{t_i' \in T_i} Prob\{\tilde{t}_{-i} \in I^i_{\hat{\nu}^P}(t_i', t_i) | \tilde{t}_i = t_i\} \\ &\leq \hat{\nu}^P \\ &\leq \delta \Lambda^Q. \end{aligned}$$

Claim 3: For the conditioning system $\mathcal{S}$,

$$\gamma^P(\mathcal{S}) \leq \delta \Lambda^Q$$

Proof of Claim 3: From the definition of $\mathcal{P} * Q$, it follows that

$$\begin{aligned} ||\pi_k - P_\Theta(\cdot|t)|| &= \sum_\theta |\sum_\ell \pi_\ell(\theta)\chi_k(\ell) - \sum_\ell \pi_\ell(\theta)Q(\ell|t)| \\ &\leq \sum_\theta \sum_\ell \pi_\ell(\theta)|\chi_k(\ell) - Q(\ell|t)| \\ &= \sum_\ell |\chi_k(\ell) - Q(\ell|t)|. \end{aligned}$$

Therefore,

$$\gamma^P(\mathcal{S}) = \max_{1\leq k\leq m} \max_{t\in A_k} ||\pi_k - P_\Theta(\cdot|t)|| \leq \max_{1\leq k\leq m} \max_{t\in A_k} ||\chi_k - Q(\cdot|t)|| \leq \hat{\mu}^Q \leq \delta\Lambda^Q.$$

Incentive compatibility now follows from Proposition A since

$$\begin{aligned} \lambda - 2M[\mu_i(\mathcal{S}|t_i) + \gamma(\mathcal{S}) + \sigma(\mathcal{S}) + \nu_i(\mathcal{S}|t_i)] &\geq c(\varepsilon, \mathcal{A}, \mathcal{P})\Lambda^Q - 10M\delta\Lambda^Q \\ &= \Lambda^Q[c(\varepsilon, \mathcal{A}, \mathcal{P}) - 10M\delta] \\ &\geq 0 \end{aligned}$$

Now suppose that $\text{Prob}\{\tilde{t} \in A_0|t_i\} \leq \hat{\mu}^Q$ for each i and t_i. Hence,

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

from which it follows that

$$\text{Prob}\{\tilde{t} \in \cup_{k=1}^m A_k\} = 1 - \sum_{t_i\in T_i} \text{Prob}\{\tilde{t} \in A_0|t_i\}P(t_i) \geq 1 - \varepsilon.$$

Finally, suppose that $t \in A_k$. Then

$$[1 - Q_{J_m}(k|t)] + \sum_{\ell\neq k} Q_{J_m}(\ell|t) = ||Q_{J_m}(\cdot|t) - \chi_k|| \leq \hat{\mu}^Q \leq \delta\Lambda^Q \leq \varepsilon\Lambda^Q \leq \varepsilon$$

and we conclude that

$$1 - \varepsilon \leq Q_{J_m}(k|t).$$

Furthermore, the definition of $(Q, \mathcal{P}, \varepsilon, \lambda, \mathcal{A})$ quasi-mechanism and the construction of $z(\cdot)$ imply that for all $i \in N$,

$$v_i(z_i(t); \pi_k) \geq v_i(\zeta_i(k); \pi_k) - \varepsilon.$$

This completes the proof of the the Theorem.

9.4 Proof of Corollary 2:

Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and let $\mathcal{P} = \{\pi_1, .., \pi_m\}$ be a collection of measures in Δ_Θ. Furthermore, suppose that $\mathcal{A} = \{\zeta(k)\}_{k\in J_m}$ is a collection where, for each $k \in J_m$, $\zeta(k)$ is a Pareto efficient, strictly individually rational CIE allocation for the CIE $e(\pi_k)$. Let

$$K(\mathcal{A}) = \min_i \min_k [v_i(\zeta_i(k); \pi_k) - v_i(w_i; \pi_k)]$$

and

$$M = \max_\theta \max_i u_i(\sum_{j=1}^n w_j; \theta).$$

Since each $\zeta(k)$ is strictly individually rational for the CIE $e(\pi_k)$, it follows that $K(\mathcal{A}) > 0$. Choose $\varepsilon > 0$ and choose $\hat{\varepsilon}$ so that

$$0 < \hat{\varepsilon} < \min\{\frac{K(\mathcal{A})}{4M+1}, \frac{\varepsilon}{4M+1}\}$$

Applying Theorem 3, there exists a $\hat{\delta} > 0$ such that, whenever $Q \in \Delta_{J_m\times T}$, $P = \mathcal{P} * Q$ and satisfies

$$\max_i \hat{\mu}_i^Q \leq \hat{\delta} \min_i \Lambda_i^Q$$

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

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

(i) $Q(k|t) \geq 1 - \hat{\varepsilon}$

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

$$v_i(z_i(t); \pi_k) \geq v_i(\zeta_i(k); \pi_k) - \hat{\varepsilon}.$$

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

$$||Q_{J_m}(\cdot|t) - \chi_k|| = [1 - Q_{J_m}(k|t)] + \sum_{\ell\neq k} Q_{J_m}(\ell|t) \leq 2\hat{\varepsilon}.$$

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

$$\begin{aligned}
\sum_\theta [u_i(z_i(t); \theta) - u_i(w_i; \theta)]\, P(\theta|t) &= \sum_\ell [v_i(z_i(t); \pi_\ell) - v_i(w_i; \pi_\ell)]\, Q(\ell|t) \\
&\geq v_i(z_i(t); \pi_k) - v_i(w_i; \pi_k) - (2M)(2\hat{\varepsilon}) \\
&\geq v_i(\zeta_i(k); \pi_\ell) - v_i(w_i; \pi_\ell) - \hat{\varepsilon} - 4M\hat{\varepsilon} \\
&\geq K(\mathcal{A}) - (4M+1)\hat{\varepsilon} \\
&> 0.
\end{aligned}$$

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

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

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

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 $i \in N$, it follows that

$$\begin{aligned}
\varepsilon &< \sum_\theta [u_i(y_i(t);\theta) - u_i(z_i(t);\theta)]P(\theta|t) \\
&= \sum_\ell [v_i(y_i(t);\pi_\ell) - v_i(z_i(t);\pi_\ell)]Q(\ell|t) \\
&\leq (2M)(2\hat{\varepsilon}) + v_i(y_i(t);\pi_k) - v_i(z_i(t);\pi_k) \\
&= 4M\hat{\varepsilon} + [v_i(y_i(t);\pi_k) - v_i(\zeta_i(k);\pi_k)] + [v_i(\zeta_i(k);\pi_k) - v_i(z_i(t);\pi_k)] \\
&\leq 4M\hat{\varepsilon} + [v_i(y_i(t);\pi_k) - v_i(\zeta_i(k);\pi_k)] + \hat{\varepsilon}.
\end{aligned}$$

Therefore,

$$0 < \varepsilon - 4M\hat{\varepsilon} - \hat{\varepsilon} < [v_i(y_i(t);\pi_k) - v_i(\zeta_i(k);\pi_k)]$$

for each i, contradicting the assumption that $\{\zeta_i(k)\}_{i \in N}$ is Pareto optimal in $e(\pi_k)$. Therefore, $t \notin E$ and $z(\cdot)$ satisfies X$_\varepsilon$E.

9.5 Proofs of Theorem 1 and Corollary 1:

These are immediate consequences of Theorem 3 and Corollary 2. If $\Theta = \{\theta_1, .., \theta_m\}$, then let $\pi_k = \chi_{\theta_k}$ and $Q(k,t) = P(\theta_k, t)$ for each k and t.

9.6 Proof of Theorem 2:

The proof is also essentially identical to that of Theorem 3. Suppose $\Theta = \{\theta_1, .., \theta_m\}$ and let $\pi_k = \chi_{\theta_k}$ and $Q(k,t) = P^r(\theta_k, t)$ for each k and t so that $\hat{\mu}^Q = \hat{\mu}^{P^r}$, $\hat{\nu}^Q = \hat{\nu}^{P^r}$and $\Lambda^Q = \Lambda^{P^r}$. Suppose that $\mathcal{A} = \{\zeta(\theta)\}_{\theta \in \Theta}$ is a collection of CIE allocations for the collection $\{e(\theta)\}_{\theta \in \Theta}$. Choose $r \geq 1$ and let $\{e^r(\theta)\}_{\theta \in \Theta}$ denote the associated replica economy and let $\mathcal{A}^r = \{\zeta^r(\theta)\}_{\theta \in \Theta}$ denote the associated replica allocation. It follows from the definitions that

$$c(\varepsilon, \mathcal{A}, \mathcal{P}) = c(\varepsilon, \mathcal{A}^r, \mathcal{P})$$

where $c(\varepsilon, \mathcal{A}, \mathcal{P})$ is defined for $\{e(\theta)\}_{\theta\in\Theta}, \mathcal{A}$ and $\mathcal{P}$ and $c(\varepsilon, \mathcal{A}^r, \mathcal{P})$ is defined for $\{e^r(\theta)\}_{\theta\in\Theta}, \mathcal{A}^r$ and $\mathcal{P}$. If

$$M = \max_{\theta} \max_{i} u_i(\sum_{j=1}^{n} w_j; \theta),$$

then any $(Q, \mathcal{P}, \varepsilon, \lambda, \mathcal{A}^r)$ quasimechanism $\{\{z^r_{is}(\theta_k, t_i)\}_{(t_i,k)\in T_i\times J_m}\}_{(i,s)\in N\times J_r}$ for the PIE $(\{e^r(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}^r, P^r)$ has the property that

$$u_{is}(z^r_{is}(\theta_\ell, t_i), \theta_k) \leq u_{is}(\zeta^r_{is}(\theta_\ell, t_i), \theta_k) = u_i(\zeta_i(\theta_\ell, t_i), \theta_k) \leq M.$$

Hence, for each $\varepsilon > 0$, we may choose δ to be independent of r. More precisely, choose δ so that

$$0 < \delta < \min\{\frac{c(\varepsilon, \mathcal{A}, \mathcal{P})}{10M}, \varepsilon, \frac{D(\mathcal{P})}{3}\}.$$

Now the remainder of the proof is a verbatim copy of the proof of Theorem 3.

9.7 Proof of Corollary 2:

Let $\{(\{e^r(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}^r, P^r)\}_{r=1}^{\infty}$ 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:

$$\varphi(t^r)(\tau) = \frac{|\{s \in J_r | t^r_{\cdot s} = \tau\}|}{r}$$

(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 $r > \hat{r}$,

$$\nu_i^{P^r} \leq \rho \text{ and } \mu_i^{P^r} < \rho.$$

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 $r > \hat{r}$,

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

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

and

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

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

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

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

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 $\nu_i^{P^r} \leq \rho$. Since

$$||\varphi(\tilde{t}^r) - P_T(\cdot|\theta_k)|| < \lambda/2 \Rightarrow ||\varphi(t^r) - P_T(\cdot|\theta_k)|| < \lambda \Rightarrow ||P^r_\Theta(\cdot|t^r) - \chi_{\theta_k}|| < \rho/2 < \rho \text{ for all } t^r,$$

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

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

Hence,

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

and we conclude that $\mu_i^{P^r} \leq \rho$.

Step 2:

Let $\{e(\theta)\}_{\theta\in\Theta}$ be a collection of CIE's and let $\mathcal{P} = \{\chi_{\theta_1}, .., \chi_{\theta_m}\}$. Furthermore, suppose that $\mathcal{A} = \{\zeta(\theta_k)\}_{k\in J_m}$ is a collection where, for each $k \in J_m$, $\zeta(\theta_k)$ is a Pareto efficient, strictly individually rational CIE allocation for the CIE $e(\theta_k)$. Let

$$K(\mathcal{A}) = \min_i \min_k [u_i(\zeta_i(\theta_k); \theta_k) - u_i(w_i; \theta_k)]$$

and

$$M = \max_\theta \max_i u_i(\sum_{j=1}^{n} w_j; \theta).$$

Since each $\zeta(\theta_k)$ is strictly individually rational for the CIE $e(\theta_k)$, it follows that $K(\mathcal{A}) > 0$. Choose $\varepsilon > 0$ and choose $\hat{\varepsilon}$ so that

$$0 < \hat{\varepsilon} < \min\{\frac{K(\mathcal{A})}{4M+1}, \frac{\varepsilon}{4M+1}\}.$$

For a conditionally independent sequence,

$$\Lambda_{i,s}^{P^r} = \Lambda_i^P$$

for all r and s. In particular, $\Lambda_{i,s}^{P^r}$ is independent of r. Furthermore, $\Lambda_i^P > 0$ since $P \in \Delta^*_{\Theta \times T}$. Applying Theorem 2 and the claim in step 1 above, we conclude that 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 $(\{e^r(\theta)\}_{\theta\in\Theta}, \widetilde{\theta}, \tilde{t}^r, P)$ and a collection $A_1^r, .., A_m^r$ of disjoint subsets of $(T^r)^*$ such that $\text{Prob}\{\tilde{t}^r \in \cup_{k=1}^m A_k^r\} \geq 1 - \hat{\varepsilon}$ and, for all $k \in J_m$ and all $t \in A_k^r$,

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

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

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

To show that $z^r(\cdot)$ satisfies XIR, suppose that $t^t \in A_k^r$ for some $k \geq 1$. Then

$$\begin{aligned}
\sum_\ell [u_{is}(z_{is}^r(t^r); \theta_\ell) - u_{is}(w_{is}; \theta_\ell)] P(\theta_\ell | t^r) &\geq u_i(z_{is}^r(t^r); \theta_k) - u_i(w_i; \theta_k) - (2M)(2\hat{\varepsilon}) \\
&\geq u_i(\zeta_i(\theta_k); \theta_k) - u_i(w_i; \theta_k) - \hat{\varepsilon} - 4M\hat{\varepsilon} \\
&\geq K(\mathcal{A}) - (4M+1)\hat{\varepsilon} \\
&> 0.
\end{aligned}$$

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

$$\text{Prob}\{\tilde{t}^r \in E^r\} = \text{Prob}\{\tilde{t}^r \in \cup_{k=1}^m A_k^r\} \geq 1 - \hat{\varepsilon} \geq 1 - \frac{\varepsilon}{4M+1} \geq 1 - \varepsilon.$$

To complete the proof, suppose $t^r \in (T^r)^*$ and that $y^r(\cdot)$ is a 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

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

and therefore,

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

Since each $u_i(\cdot; \theta)$ is concave and $z_{is}^r(t^r) = z_{is'}^r(t^r)$, it follows that

$$\sum_\theta [u_i(\overline{y}_i; \theta) - u_i(z_{is}^r(t^r); \theta)] P(\theta | t^r) > \varepsilon.$$

If $t^r \in A_k^r$ for some k, then for each $i \in N$,

$$\begin{aligned}
\varepsilon &< \sum_{\theta}[u_i(\overline{y}_i;\theta) - u_i(z_{is}^r(t^r);\theta)]P(\theta|t^r) \\
&\leq u_i(\overline{y}_i;\theta_k) - u_i(z_{is}^r(t^r);\theta_k) + (2M)(2\hat{\varepsilon}) \\
&\leq 4M\hat{\varepsilon} + u_i(\overline{y}_i;\theta_k) + \hat{\varepsilon} - u_i(\zeta_i(\theta_k);\theta_k).
\end{aligned}$$

Therefore,

$$0 < \varepsilon - (4M+1)\hat{\varepsilon} < [u_i(\overline{y}_i;\theta_k) - u_i(\zeta_i(\theta_k);\theta_k)]$$

for each i and we conclude that $\{\zeta_i(\theta_k)\}_{i \in N}$ is not Pareto optimal in $e(\theta_k)$, a contradiction. Hence, $t^r \notin E^r$ and the proof is complete.