The Full Surplus Extraction Theorem with Hidden Actions

Ichiro Obara*†
Department of Economics
UCLA

First version: January 2003
This version: May, 2003
PRELIMINARY AND INCOMPLETE
COMMENTS WELCOME

Abstract

I consider a situation in which a principal commits to a mechanism first and then agents choose unobservable actions before they draw their types. The agents' actions affect not only their payoffs directly but also a distribution of private types as well. Thus, the distribution of types is determined endogenously rather than exogeneously unlike standard mechanism design literature.

Then I extend Cremer and McLean's full surplus extraction theorem [10] to such a setting, that is, identify a necessary and sufficient condition on the information structure for the full surplus extration. More importantly, it implies that the full surplus extraction may not be obtained for generic set of information structure in this more general setting. This contrasts with the standard full surplus extraction theorem which holds generically.

I show, however, that the principal can extract all the surplus for any completely mixed action profile for almost all information structure by using more general mechanisms in which agents announce both their types and the realizations of their mixed actions. Since any pure action profile can be arbitrarily approximated by a completely mixed action profile, the principal can virtually extracts all the maximized social surplus.

JEL classification: C72, D44, D82

Keywords: auction, communication, costly information acquision, full surplus extraction, mechanism design, mixed strategy, moral hazard, private strategy, virtual implementation.

*E-mail: iobara@econ.ucla.edu

†I am grateful for helpful comments to Sandeep Baliga, Tomas Sjöström, and seminar paricipants at NBER/NSF Decentralization Conference 2003 at Purdue, MEDS and Penn State.

1 Introduction

Since the seminal paper by Hurwicz[15], mechanism design literature has been dealing with a variety of situation of which asymmetric information among agents is an important aspect. Standard assumption in this literature is that there is some exogenously given distribution of private types, which is common knowledge among all agents and a principal. In this paper, I consider a situation in which a principal commits to a mechanism first and then agents choose unobservable actions before they draw their types. The agents' actions may affect not only their value of allocations but also a distribution of their private types as well. Thus the distribution of types is determined endogenously rather than exogenously unlike standard mechanism design literature. One special example of such situation would be an auction in which the bidders can gather private information about the object by spending some costs.[1] There have been several works on costly information acquisition such as Matthews[20], Stegeman[27] and more recently, Persico[25], Bergemann and Välimäki[3].

In this more general setting in which players' type distribution is determined endogenously, I consider a possibility of full surplus extraction, in particular, a possibility of full surplus extraction when the surplus-maximizing action profile is implemented. Cremer and McLean [10] (CM) identified a necessary and sufficient condition on the information structure for a principal to be able to extract all the rents from the agents in auction. I extend their full surplus extraction theorem to the current setting.

More importantly, my characterization result shows that full surplus extraction does not hold for a generic set of information structure in this more general setting. CM showed that their necessary and sufficient condition on information structure is satisfied generically and this result was extended by McAfee and Reny[21] to cases with continuous type under weak conditions. These papers cast a doubt on the robustness of the results obtained in standard mechanism design framework where agents are risk neutral and their type distributions are independent. On the other hand, casual observation shows that a principal never be able to extract all the rents from the agents, or there is just no such sophisticated mechanism as CM's mechanism to try to extract all the surplus. Indeed, mechanisms we see are often very simple as assumed in standard mechanism literature. This calls for an explanation, or more substantially, a foundation for standard mechanism design framework. To address this question, some papers take CM's model as a benchmark, then introduced a new feature so that full surplus extraction may fail. One paper introduced risk averseness/limited liability (Robert[26]) and other papers questioned the routine use of common prior assumption (Neeman[23]). This paper contributes to such strand of literature by adding one more reason for the failure of full surplus extraction; if agents can affect the underlying type distribution through their actions, the principal may be kept from extracting the full surplus and even forced to use a simple mechanism.

[1] It is special in the sense that the value of allocation is not directly affected by actions.

I illustrate this point using a simple example of auction with costly information acquisition. It is an auction in which two bidders can pay some cost to obtain information about the object for themselves or pay the same cost to steal the other bidder's information. The auction can be relatively more common value or private value depending on one parameter. When there is enough common component, no information acquisition is optimal as the allocation of good is not a big problem. On the other hand, when the auction is more private, costly information acquisition is optimal as it facilitates more efficient allocation. Either way, the mechanism to extract all the surplus necessarily takes a form of sophisticated lottery as in CM given the optimal action profile. When a bidder announces her type, she receives a lottery, the outcome of which depend on the other bidder's announcements. As is well known, the seller cannot extract all the surplus when the bidder's types are independent. But, when there is no action, the seller can extract all the surplus by using such sophisticated lottery when the bidders' types are enough correlated. However, this is no longer true when there are available actions. As the stake of the auction becomes big, each bidder's payment becomes necessarily more and more sensitive to the other bidder's announcement and monetary transfer tends to be a huge amount. This breaks down the bidders' incentive for costly information acquisition or no costly information acquisition because stealing information becomes mote attractive given such extreme payment scheme.

I show, however, that the principal can extract all the surplus for almost all information structure if there is more flexibility in terms of available mechanisms. It is shown that the principal can extract all the surplus for any completely mixed action profile for almost all information structure by using more general mechanisms in which agents announce both their types and *the realizations of their mixed actions*. Since the efficient action profile can be approximated arbitrarily closely by a completely mixed action profile, the principal can virtually extract all the maximized social surplus. This result reinforces the results obtained by Cremer and McLean and McAfee and Reny.

The basic model and definitions are introduced in the next section. CM's full surplus extraction theorem is extended to the current setting in Section 3. It is also shown that full surplus extraction is not a generic result anymore unlike standard cases with a fixed type distribution. An explicit example of an auction with costly information acquisition is given in Section 4 to illustrate this failure of the full extraction. Section 5 proves a strengthened full surplus extraction theorem which works for almost all information structure, and Section 6 concludes.

2 The Model and Definitions

2.1 The Model

There is one principal and $n\,(\geq 2)$ agents; $N = \{1, 2, ..., n\}$. Agent i takes an action from a finite set A_i, then draws her type according to some joint distribution over $S = \prod_{i=1}^{n} S_i$ conditional on all agents' actions, where S_i is a finite set of possible player i's types. *Information structure* (π, S, A) is defined as the set of conditional distributions on S given action profile; $\{\pi(s|a)\}_{s\in S, a\in A}$. Thus the space of information structure Ψ can be identified as $|A|$-time products of $|S|-1$ dimensional simplex in $\Re_+^{|A|\times(|S|-1)}$.[2] In some contexts, a more specific distribution form $\pi(s|a) = \sum_{v\in V} p(v) \prod_{i=1}^{n} f(s_i|v_i, a_i)$ is useful. This is called *Information Acquisition Model.* Interpretation is that $v = (v_1, ..., v_n) \in V$ is a value vector and each agent receives a conditionally independent signal about her own value and informativeness of i's signal depends on i's action a_i. For this specific class of information structure, the space of information structure is identified as the product space of $p(v)$ and $f(s_i|v_i, a_i)$, $i = 1, ..., n$. Let $p(s_j|a) = \sum_{s_{-j}\in S_{-j}} \pi(s_j, s_{-j}|a)$ be the marginal distribution on S_j given a. It is assumed that $p(s_i|a) > 0$ for all $a \in A = \prod_{i=1}^{n} A_i$ and $s_i \in S_i$. Let $\pi(s_{-i}|s_i, a) = \frac{\pi(s|a)}{p(s_i|a)}$ be agent i's subjective distribution over the other agents' types given private signal s_i and action profile a. Let $p(s_j|s_i, a) = \frac{\pi(s_i, s_j, s_{-i,j}|a)}{p(s_i|a)}, j \neq i$ be agent i's subjective distribution over agent j's types given (s_i, a).

The agents choose their actions simultaneously, and their actions and types are observable neither to the principal nor the other agents. In the beginning of the game, the principal chooses a (direct) mechanism (x, t) which maps a profile of announcements $\widetilde{s} \in S$ into an allocation $x(\widetilde{s}) \in X$ and transfers $t(\widetilde{s}) = (t_1(\widetilde{s}_{-1}|\widetilde{s}_1), ..., t_n(\widetilde{s}_{-n}|\widetilde{s}_n)) \in \Re^n$, where $t_i(\widetilde{s}_{-i}|\widetilde{s}_i)$ is a transfer for agent i.

The principal's revenue is simply given by $\sum_{i=1}^{n} t_i$. Agent i's payoff consists of three components. Agent i's *utility* from allocations is given by $V_i : X \times S \times A \to \Re^+$. In general, it depends on the other agents' types s_{-i} (mutually payoff-relevant) and action profile a. However, since agents' utility may take more specific forms for specific applications, I often focus on a particular class of utility functions. The class of utility functions which only depends on x and s_i is called TYPE I and denoted as $V_i(x, s_i) \in V^I$. The class of utility functions which depends on x and s is called TYPE II and denoted as $V_i(x, s) \in V^{II}$. Finally, the most general utility functions $V_i(x, s, a)$ are called TYPE III.[3] Clearly, $V^I \subset V^{II} \subset V^{III}$. Utility V_i and transfer t_i are two components of agent i's payoff which matter if and only if agent i plays the mechanism offered by the principal. The third component of agent i's payoff is $g_i(a_i)$ which is independent of i's decision to enter or stay out of the mechanism. This term reflects a direct cost or profit associated with one's own action [4]. Let $g_i^* = \max_{a_i} g_i(a_i)$.

[2] $\left\{\pi(s|a) \in \Re_+ | \ \sum_{s\in S} \pi(s|a) = 1, s \in S, a \in A\right\}$

[3] As an example of such utility function which depends on a and s, consider a procurement auction in which a principal is selling a project which requires a new, sophisticated technology. Bidders invest their resources to this new technology before they play the auction. Suppose that their investment has two effects; information gathering effect and value-enhancing effect. Given that every bidder's information is relevant to the other bidders, agent i's utility is given by $V_i(x, s, a_i)$ (or $V_i(x, s, a)$ if bidders actions have direct externality on the other bidders) for this example.

[4] It is easy to have g_i depend on the other agents' actions as well. In such a case, however, i's outside option ($= \max_{a_i} g_i(a_i, a_{-i})$ +constant) depends on which action profile to implement.

Given (x, s, a) and t_i, $V_i(x, s, a) - t_i + \{g_i(a_i) - g_i^*\}$ is called agent i's *total rent.* I call $V_i(x, s, a) - t_i(s)$ simply agent i's *rent* and $V_i(x, s, a) + \{g_i(a_i) - g_i^*\}$ is called agent i's *surplus.* Ex-ante refers to a value before type is drawn and interim means a conditional value given agent's type. For example, $\sum_{s \in S} \pi(s|a) V_i(x(s), s, a)$ is an ex-ante expected utility and $\sum_{s \in S} \pi(s_{-i}|s_i, a)$ $(V_i(x(s), s, a) - t_i(s_{-i}|s_i))$ is an interim expected rent.

The time line of the game is shown in Figure 1.[5] Since the principal can commit to a mechanism in the beginning of the game, it is without loss of generality to focus exclusively on direct mechanisms in terms of equilibrium payoff characterization. There are two different scenarios depending on the timing of agents' participation decision. Interim participation case is the case in which the agents' participation decision comes after their private types are realized. Ex-Ante participation case is the case in which the agents have to decide to participate before any action is taken. In either case, each agent's reservation value is normalized to 0 without loss of generality. As explained later, both cases fit in my model by taking into account appropriate individually rational constraints.

Let Γ be the set of all direct mechanisms, that is, all pair of mappings $x : S \to X$ and $t : S \to \Re^n$. The principal's strategy is simply to commit to one mechanism in the beginning of the game. Agent i's strategy $\sigma_i = (\sigma_i^a, \sigma_i^s)$ simply consists of two mappings; $\sigma_i^a : \Gamma \to A_i$ which maps a mechanism into an action and an announcement strategy $\sigma_i^s : \Gamma \times S_i \to S_i$.[6] Note that agents are only allowed to reveal their types even though their actions are also their private information. Of course, this is without loss of generality as long as the implementation of pure action profiles is concerned, which is a natural starting point given that I am concerned with the implementation of efficient outcomes. I will come back to this issue when I analyze a more general class of mechanism in which the agents play mixed action profiles and announce both their private types and their realized actions.

Finally, I use perfect Bayesian Nash equilibrium as an equilibrium notion for this model.[7]

Since this may make an analysis look unnecessarily more complicated and blur the notion of full surplus extraction, I decided not to include the other agents' actions in g_i.

[5] It is a strong assumption that the principal can choose a mechanism before any action is taken, thus before agents draw their types. This assumption may be satisfied, for example, for repeated auctions for which a principal need to choose a format in the beginning for every auction.

I also try to justify this assumption later by showing that the principal indeed has a very strong incentive to committ to a mechanism before any action is taken.

[6] Since it is not efficeint to keep some agents out of the mechanism, every agent enters the mechanism in any efficient equilibrium. So I ignore their entry strategy and just check their interim IR conditions.

[7] Since I focus on the failure of full surplus extraction (except for Section 5), I prefer to use less restrictive notion of equilibrium here. Hence I do not use a stronger notion of equilibrium such as dominant strategy equilibrium or ex-post equilibrium.

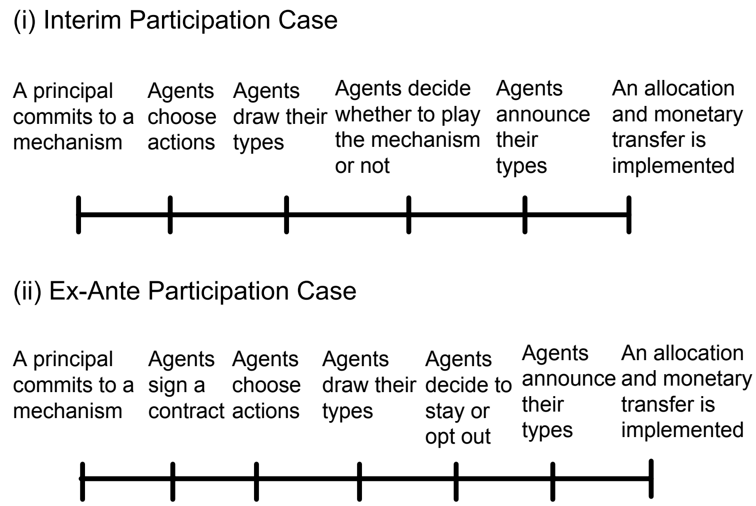

Figure 1:

2.2 Relevant Constraints

To check whether (x, t) and a can be implemented as an outcome of Bayesian Nash equilibrium, I only need to verify the following constraints. First, all agents play the mechanism on the equilibrium path without loss of generality;

$$\sum_{s_{-i} \in S_{-i}} \pi\left(s_{-i} | a, s_i\right)\left\{V_i\left(x\left(s\right), s, a\right) - t_i\left(s_{-i} | s_i\right)\right\} \geq -\eta \text{ for all } s_i \in S_i, i \in N \tag{$IR_{interim}$}$$

If agent i's participation decision comes after s_i is drawn, then η is simply the reservation value 0. On the other hand, if agent i needs to sign a contract before taking action, but can opt out after she learns her type, then η may be strictly positive. It reflects an upper bound on the amount of fine which can imposed for a breach of contracts. All the theorems are valid for these two different scenarios; Interim participation and Ex-Ante participation.

Second, I need the following truth telling constraint on the equilibrium path;

$$\begin{aligned} & \sum_{s_{-i} \in S_{-i}} \pi\left(s_{-i} | a, s_i\right)\left\{V_i\left(x\left(s\right), s, a\right) - t_i\left(s_{-i} | s_i\right)\right\} \qquad (IC_{s_i}) \\ \geq & \sum_{s_{-i} \in S_{-i}} \pi\left(s_{-i} | a, s_i\right)\left\{V_i\left(x\left(\left(s_i', s_{-i}\right)\right), s, a\right) - t_i\left(s_{-i} | s_i'\right)\right\} \text{ for all } s_i, s_i' \left(\neq s_i\right) \in S_i \end{aligned}$$

Third, agents have to have incentive to play a;

$$
\begin{aligned}
& \sum_{s\in S} \pi\left(s|a\right)\left(V_i\left(x\left(s\right), s, a\right) - t_i\left(s_{-i}|s_i\right)\right) + g_i\left(a_i\right) \qquad (IC_{a_i}) \\
\geq \quad & \sum_{s_i\in S_i} p\left(s_i|a_i', a_{-i}\right) \\
& \max\left\{ \sum_{s_{-i}\in S_{-i}} p\left(s_{-i}|s_i, \left(a_i', a_{-i}\right)\right) \left\{ \begin{array}{c} V_i\left(x\left(\left(s_i', s_{-i}\right)\right), s, \left(a_i', a_{-i}\right)\right) \\ -t_i\left(s_{-i}|s_i'\right) \end{array} \right\}, -\eta \right\} \\
& + g_i\left(a_i'\right) \text{ for all } a_i' \left(\neq a_i\right), i \in N.
\end{aligned}
$$

This means that an agent would not gain by deviating to play some action a_i' and use any announcement/participation strategy.

Finally, the following Ex-Ante participation constraint needs to be satisfied additionally for Ex-Ante participation scenario;

$$
\sum_{s\in S} \pi\left(s|a\right)\left(V_i\left(x\left(s\right), s, a\right) - t_i\left(s_{-i}|s_i\right)\right) + g_i\left(a_i\right) \geq g_i^* \qquad (IR_{ExAnte})
$$

This constraint is redundant for Interim Participation scenario or when $\eta = 0$.

2.3 Examples

Some simple examples are provided below to motivate the above general model.

Example 1: Auction with Interdependent Values

One object is to be allocated among n bidders. True value of the object consists of n components $v = \left(v_1, ..., v_n\right)$, which is stochastic and whose distribution is known to everyone. The value of the object for bidder i is given by $\sum_{k=1}^{n} a_{i,k} v_k$, where coefficients $a_{i,k}$ are exogenous and may be different across the bidders. Bidder i cannot observe her true value, but can obtain some information s_i. The quality of i's information depends on i's effort level. Each agent may be able to make s_i to be a more accurate signal about v_i or a more accurate signal about the other variables such as $v_j, s_j, j \neq i$ by choosing an appropriate type of effort. The distribution of signal profile conditional on any action profile and $v = \left(v_1, ..., v_n\right)$ is commonly known to all bidders

For this example, $X = \{1, ..., n\}$. Bidder i's utility is $V_i\left(x, s\right) = E\left[\sum_{k=1}^{n} a_{i,k} v_k | s\right]$ if $x = i$ and 0 otherwise. Since it is natural to assume that the auctioneer cannot identify the bidders in advance, it is appropriate to assume that only interim IR constraint binds $\left(\eta = 0\right)$.

Example 2: Efficient Task Allocation Problem

A principal is trying to assign n different tasks to n agents with different characteristics. Only one agent can be assigned to each task. Thus X is a space of all permutations mapping $\{1, ..., n\}$ to $\{1, ..., n\}$. The resulting profit from this team production is independent of task allocation, but task allocation affects the cost each agent has to incur. Each agent's cost $C_{i,k}(\theta_i, \tau)$ of conducting task k is a function of her own characteristic θ_i and all task's characteristics $\tau = (\tau_1, ..., \tau_n)$. Agent i can obtain some information $s_{i,k}$ about a characteristic of task k, whose quality can be improved if agents pay some costs. Let $s_i = (s_{i,1}, ..., s_{i,n})$ be agent i's information vector. These information may be potentially useful to implement the efficient cost-saving task allocations. For this example, it is natural to assume that the principal makes a take-it-or-leave-it offer before any information gathering activity takes place. Each agent can opt out by paying a penalty for breach of the contract after she learns her type.

For this example, agent i's utility function is given by $V_i(x, s) = -E[C_{i,k}(\theta_i, \tau)|s]$ if task k is assigned to i. The principal's objective function is $Profit + \sum_{i=1}^n t_i$.

2.4 Some Definitions

Optimal Allocation

It is assumed that $V_i(x, s, a)$ is continuous with respect to x for each $(s, a) \in S \times A$ for all i, and that X is a compact metric space. Then, the optimal allocation function $x^a(\cdot)$ exists for each action profile $a \in A$.[8] When every agent plays the mechanism and (x, s, a) is realized, the social surplus is $\sum_{i=1}^n V_i(x, s, a) + \sum_{i=1}^n \{g_i(a_i) - g_i^*\}$. Thus, $x^a(\cdot)$ maximizes

$$\sum_{s \in S} \pi(s|a) \sum_{i=1}^n V_i(x(s), s, a) + \sum_{i=1}^n g_i(a_i)$$

Let $W(a) = \sum_{i=1}^n \sum_{s \in S} \pi(s|a) V_i(x^a(s), s, a) + \sum_{i=1}^n \{g_i(a_i) - g_i^*\}$ be the optimal social surplus given $a \in A$. Let $A^*(V, g)$ be the set of action profiles which maximizes $W(a)$ given $V = (V_1, ..., V_n)$ and $g = (g_1, ..., g_n)$.

FSE-implementability

Definition 1 *An action profile $a \in A$ is FSE-implementable if there exists a perfect Bayesian Nash equilibrium $((x^a, t), \sigma)$ in which a is played and all the social surplus is extracted by the principal, that is, $\sum_{s \in S} \pi(s|a) \sum_{i=1}^n t_i(s_{-i}|s_i) = W(a)$.*

Definition 2 *An information structure $(\pi, (S, A))$ has FSE property in V^X if, for any $V \in V^X$ and g, there exists $a \in A^*(V, g)$ which is FSE implementable.*

[8] The optimal allocation function is independent of a for TYPE I and II utility funtions.

3 The Full Surplus Extraction Theorem

3.1 Cremer and McLean (1988)

If the action set is singleton (or equivalently, V_i, π, and g_i do not depend on actions), this problem reduces to a standard full surplus extraction problem. Thus the original surplus extraction theorem can be stated in the current framework. Suppose that the action set is singleton and denote $\pi(\cdot|s_i, a) = \pi(\cdot|s_i)$. CM showed that the full surplus can be extracted in Bayesian Nash equilibrium if and only if $\pi(\cdot|s_i)$ is not a positive linear combination of $\left\{\pi(\cdot|s_i') \in \Re_+^{|S_{-i}|} : s_i' \in S_i / \{s_i\}\right\}$ for any $s_i \in S_i$.[9] See Figure 2 (a) to get a geometric image of this condition. Denote the convex full of a collection of m vectors x_i in $\Re^n$ by $co\{x_i \in \Re^n : i = 1, 2, ..., m\}$. Then, their theorem can be stated as follows.

Theorem 3 *(Cremer and McLean [10]) An information structure* (π, S) *has FSE property in* V^I *and* V^{II} $\left(= V^{III}\right)$ *if and only if* $\pi(\cdot|s_i) \notin co\{\pi(\cdot|s_i') : s_i' \in S_i / \{s_i\}\}$ *for all* $s_i \in S_i$ *and all* $i \in N$.

The original theorem was proved only for V^I. But as the authors claimed in [9], it can be easily extended to V^{II}. Another difference is that η can be positive here. This is also easily accommodated to the original full surplus extraction theorem.

3.2 Main Theorem

Now I extend the above theorem to the case with hidden actions. First, I introduce a property of information structure which corresponds to the above necessary and sufficient condition when hidden actions are present.

$(*)$ $\pi(\cdot|s_i, a) \notin co\{\pi(\cdot|s_i', (a_i', a_{-i})) : (s_i', a_i') \in (S_i \times A_i) / \{(s_i, a_i)\}\}$ for all $s_i \in S_i$, $a \in A$, and all $i \in N$.

The condition means that, for any $s_i \in S_i$ and $a \in A$, $\pi(\cdot|s_i, a)$ cannot be a positive linear combination of a set of vectors, which include not only $\pi(\cdot|s_i', a)$ as in CM, but also $\pi(\cdot|s_i', (a_i', a_{-i}))$; conditional signal distributions after a deviation. If the set of conditional distributions after a deviation looks like Figure 2 (b), then this condition is satisfied at a. If it is as shown in Figure 2(c), then the condition is violated. Intuitively, the condition $(*)$ is violated if an information obtained by a deviation is in some sense essential about the other agents' types. This intuition motivates a simple model in the next section in which an agent can steal the other agents' information.

I need the following two regularity conditions to prove the theorem.[10]

[9] They also provided a necessary and sufficient condition for full surplus extraction with dominant strategy equilibrium.

[10] These conditions are much stronger than necessary, but they help me to simplify the proof in the paper considerably. They will be replaced by weaker and simpler (if any) conditions in later versions. Note also that these conditions are trivially satisfied in Cremer and McLean's setting with no action.

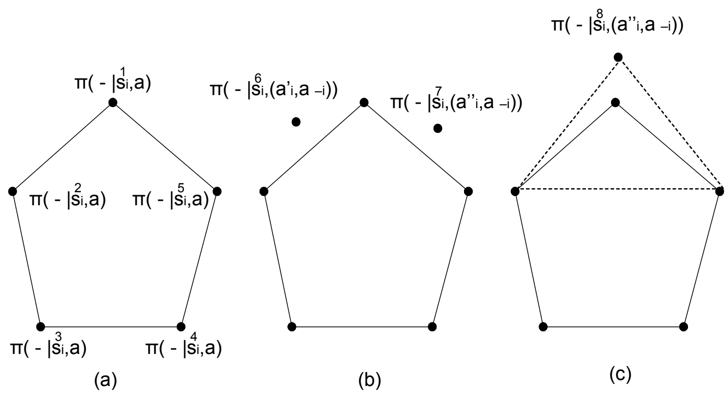

Figure 2:

Condition 4 *For any a and $i \in N$, there does not exist $\lambda : A_i/\{a_i\} \to \Re_+$ such that $p(s_j|a) = \sum_{a_i' \in A_i/\{a_i\}} \lambda(a_i')\, p(s_j|a_i', a_{-i})$ for all $j \neq i$.*

Condition 5 *$\pi(\cdot|a) \notin co\{\pi(\cdot|a_i', a_{-i}) : a_i' \in A_i/\{a_i\}\}$ for all $a \in A$, and all $i \in N$.*

These are relatively weak regularity conditions. For example, both conditions are generic in Ψ if $|A_i| \leq \max_{j \neq i} |S_j|$.

The first main theorem of this paper is the following extension of the full surplus extraction theorem in V^{II} and V^{III}.

Theorem 6 *An information structure (π, S, A) has FSE property in V^{III} if and only if it satisfies $(*)$. An information structure (π, S, A) has FSE property in V^{II} if it satisfies $(*)$, and $(*)$ is implied by FSE property if Condition 5 holds.*

Proof. See Appendix. ■

The basic idea of the proof is as follows. Since $V^{II} \subset V^{III}$, I only need to prove sufficiency for V^{III}. The sufficiency part is a relatively straightforward extension of CM's idea. Suppose that the condition $(*)$ holds. Then, a function $w_i(\cdot|s_i)$ for each s_i can be constructed so that $\sum_{s_{-i} \in S_{-i}} \pi(s_{-i}|s_i, a)\, w_i(s_{-i}|s_i) = 0$ and $\sum_{s_{-i} \in S_{-i}} \pi(s_{-i}|s_i, (a_i', a_{-i}))\, w_i(s_{-i}|s_i') > 0$ for all $(s_i', a_i') \neq (s_i, a_i)$ by Farkas' lemma. Then, transfer $t_i(s_{-i}|s_i)$ is defined as the sum of $w_i(s_{-i}|s_i)$ and the constant term which extracts agent i's expected surplus ($= \sum_{s \in S} \pi(s|a)$

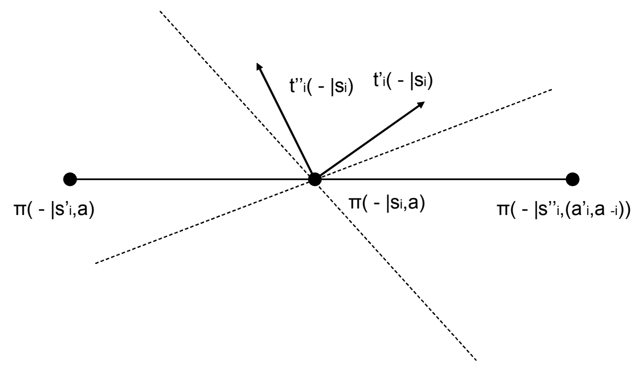

Figure 3:

$V_i(x(s), s, a) + \{g_i(a_i) - g_i^*\})$. These $w_i(s_{-i}|s_i)$ can be chosen so that agent i loses a huge amount of money from deviating from a_i and/or lying $(s_i' \neq s_i)$. Given such transfer functions, she does not have incentive to lie on the equilibrium path. She does not have incentive to deviate from a_i either because the expected loss from playing the mechanism is huge enough to overwhelm any gain from distorting the distribution of private signals and final allocations. Hence the only thing agent i can do is to stay out of the mechanism after a deviation. This means that she can obtain at most g_i^* from such deviation.

On the other hand, the necessity part is more involved. Below I illustrate the logic of the proof for TYPE II utility functions using a simple example. Suppose that the condition $(*)$ is violated for some $(s_{i,k}, a_{i,k}) \neq (s_i, a_i), k = 1, ..., K$. Condition 5 guarantees that there exists some $s_{i,k} \neq s_i$. For simplicity, I assume that $\pi(\cdot|s_i', a)$ $(s_i' \neq s_i)$, and $\pi(\cdot|s_i'', (a_i', a_{-i}))$ are two such vectors as in Figure 3. Choose continuous functions V_i and g_i , $i = 1, ..., n$ so that a is the unique surplus maximizing actions. Modify the utility of agent i by $\widehat{V}_i(x, (s_i', s_{-i})) = V_i(x, (s_i', s_{-i})) + \gamma\,(> 0)$ and the utility of some agent $j \neq i$ by $\widehat{V}_j(x, (s_i', s_{-i})) = V_j(x, (s_i', s_{-i})) - \gamma$ only for s_i'. Note that the optimal action profile and the optimal allocation function remains the same for such modifications of utility functions. By assumption, there should still exist a mechanism to implement a and extract the full surplus.

Now pick γ large enough so that $\widehat{V}_i(x, (s_i', s_{-i})) > V_i(x, (s_i, s_{-i}))$ for all s_{-i} and x. First consider the case where the transfer function is such as $t_i'(s_{-i}|s_i)$ in Figure 3 when s_i is announced. If this is the case, then type s_i' can pretend to be s_i to obtain more interim expected utility with less interim expected payments than type s_i. This implies that the interim expected rent of s_i' should be higher than the interim expected rent of type s_i in equilibrium. Since the interim expected rent of s_i is bounded below by 0, agent i's ex-ante expected rent from the mechanism exceeds $g_i^* - g_i(a_i)$ and the total rent becomes strictly positive if γ is large enough. Since full surplus extraction is equivalent to 0 total rent for all agents, the principal cannot extract the full surplus with a transfer function such as $t_i'(s_{-i}|s_i)$ when γ is large. Therefore, the equilibrium transfer function should look like $t_i''(s_{-i}|s_i)$ to make the expected payment of type s_i' from announcing s_i larger than the expected payment of type s_i in equilibrium. As γ becomes larger, the size of $t_i''(s_{-i}|s_i)$ must become larger so that s_i' needs to make a large expected payment from pretending to be s_i. Note, however, that a large expected payment for s_i' implies a large expected profit for s_i'' who pretends to be s_i after a deviation to a_i'. Since every other type $s_i \neq s_i''$ can guarantee at least $-\eta$ after the deviation to a_i', the ex-ante expected rent from such deviation exceeds $g_i^* - g_i(a_i')$ at some point as γ becomes large, hence agent i's total rent becomes strictly positive in equilibrium. Since either one of these two cases applies for any transfer function, the principal fails to extract the full surplus from agent i if γ is large enough.

TYPE I Utility Functions

What happens if agent i's utility only depends on her own type ($V_i(x, s_i)$)? Since $V^I \subset V^{II}$, $(*)$ is still sufficient. However, necessary condition can be in principle weaker as less flexible utility functions lead to less implication on information structure which has FSE property. Indeed a problem arises in the above "proof" when γ is subtracted from agent j's utility because j's utility is not a function of i's type. The problem is, if γ is just added to agent i's utility function with type s_i', a may cease to be optimal when γ is large. In particular, any action profile for which the probability of s_i' is high becomes more attractive.

Condition 4 leads to the following useful lemma.

Lemma 7 *There exists $v_j : S_j \to \Re, j \neq i$ for any $a \in A$ and $i \in N$ such that $\sum_{s\in S} \pi(s|a) \sum_{j\neq i} v_j(s_j) > \sum_{s\in S} \pi(s|a_i', a_{-i}) \sum_{j\neq i} v_j(s_j)$ for all $a_i' \neq a_i$ if and only if Condition 4 is satisfied.*

Proof. The condition

$$\sum_{s\in S} \pi(s|a) \sum_{j\neq i} v_j(s_j) > \sum_{s\in S} \pi(s|a_i', a_{-i}) \sum_{j\neq i} v_j(s_j) \text{ for } a_i' \neq a_i$$

is identical to

$$\begin{aligned} \sum_{j\neq i} \sum_{s_j\in S_j} p(s_j|a) v_j(s_j) &> \sum_{j\neq i} \sum_{s_j\in S_j} p(s_j|a_i', a_{-i}) v_j(s_j) \\ \sum_{j\neq i} \sum_{s_j\in S_j} (p(s_j|a) - p(s_j|a_i', a_{-i})) v_j(s_j) &> 0 \text{ for } a_i' \neq a_i \end{aligned}$$

Let $p_{i,j}(a, a_i') = \left(p(s_{j,1}|a) - p(s_{j,1}|a_i', a_{-i}), ..., p\left(s_{j,|S_j|}|a\right) - p\left(s_{j,|S_j|}|a_i', a_{-i}\right)\right)$ be a $|S_j|$ dimensional vector for $j \neq i$ and let A be $(|A_i| - 1) \times \sum_{j \neq i} |S_j|$ matrix whose row vectors are given by $(p_{i,1}(a, a_i'), ..., p_{i,n}(a, a_i'))$ for $a_i' \neq a_i$. Let $v = (v_1(\cdot), ..., v_n(\cdot))^\top \in \Re^{\sum_{j \neq i}|S_j|}$ be a column vector. Then the above condition can be stated compactly as

$$Av > \mathbf{0}$$

By Theorem 2.9. in Gale [13], such v exists if and only if there does not exist nonzero $\lambda \in \Re_+^{|A_i|-1}$ such that $\lambda^\top A = 0$. Thus the lemma is proved. ■

This lemma can be used to fix this problem. By Lemma 7, there exists $v_j : S_j \to \Re, j \neq i$ to satisfy $\sum_s \pi(s|a) \sum_{j \neq i} v_j(s_j) > \sum_s \pi(s|a_i', a_{-i}) \sum_{j \neq i} v_j(s_j)$. Modify the utility of agent $j \neq i$ by $\widehat{V}_j(x, s_j) = V_j(x, s_j) + \lambda v_j(s_j)$. For each γ, λ can be chosen large enough so that a is still optimal among all $(a_i', a_{-i}), a_i' \in A_i$. As for the other players' action profile $a_{-i}' \neq a_{-i}$, $g_j\left(a_j'\right)$ can be set to be very small for all such actions to make sure that nothing but a_{-i} can be optimal. In this way, a set of (V, g) can be constructed for any level of γ while maintaining that a is the unique optimal action. The rest of the proof is exactly the same as before.

Theorem 8 *An information structure (π, S, A) has FSE property in V^I if it satisfies $(*)$, and $(*)$ is implied by FSE property if Condition 4 and 5 holds.*

Proof. See Appendix. ■

3.3 Genericity of FSE Property

This theorem which extends Cremer and McLean has a corollary which is more important than the theorem itself. How strong is this necessary and sufficient condition $(*)$? Suppose that the action set is singleton. Since $\pi(\cdot|s_i), s_i \in S_i$ are vectors on $|S_{-i}| - 1$ dimensional simplex, they are generically linearly independent (hence satisfy $(*)$) if and only if $|S_i| \leq |S_{-i}|$. In particular, this condition is trivially satisfied if the number of types is the same for all agents. The condition $(*)$ suggest that this is no longer the case when there are many actions. When there are $|A_i|$ actions, the set of vectors $\pi(\cdot|s_i, a), s_i \in S_i, a \in A$ is generically linearly independent if and only if $|S_i| \times |A_i| \leq |S_{-i}|$, which may be violated with large action spaces even when the number of the type is the same for all agents. For example, when $|S_i| = m$ and $|A_i| = K$ for all i, the condition reduces to $mK \leq m(n-1)$, thus $K \leq n - 1$. If $n - 1 < K$, then there is an open set of information structure for which the set of conditional distributions is not linearly independent. Indeed, it is not difficult to show that there is an open set of information structure for which $(*)$, an even weaker condition, is violated. Thus, full surplus extraction result may not be obtained for a generic subset of Ψ if $n - 1 < K$. I summarize these arguments below.

Theorem 9 *Suppose that $|S_i| = m$ and $|A_i| = K$ for all i. FSE is a generic property if and only if $K \leq n-1$.*

Note that one can derive a corresponding result for information acquision models. For this class of information structure, FSE is a generic property if and only if $K \leq n-1$ and $K \times m \leq |V|$ (to be proved).

3.4 Discussions

Timing of Commitment by Principal

The crucial assumption of this model is that the principal can commit to a mechanism before any action is taken by the agents. It is also possible to think about a different model in which a principal proposes a mechanism after actions are chosen, or a model in which the principal proposes a mechanism and the agents take actions simultaneously. Which model is more appropriate may depend on each specific context. For example, the latter may be a more reasonable assumption if action is very difficult to observe by nature. Both assumptions have been used in the literature. Lewis and Sappington[18] used the former assumption in a multi-agent procurement model. Crémer and Khalil[7] also uses the former assumption. Crémer, Khalil, and Rochet[8] used the latter assumption in an interesting principal-agent model in which a mixed strategy equilibrium is implemented.

I would like to point out that, in multi-agents model, the principal has a very strong incentive to commit to a mechanism before any action is taken. Suppose that the principal cannot commit to a mechanism. Since the principal knows which action is taken by the agents in equilibrium, what matters for full surplus extraction is CM's necessary and sufficient condition given the equilibrium action, which holds generically. Thus the principal can subtract *all the ex-ante expected utility* $(+\eta)$ *(*not the ex-ante surplus*)* generically because any cost associated with actions is sunk. This implies that the only implementable action is the cheapest action profile. In the context of costly information acquisition, this means that no agent spends money to obtain useful information. So the principal needs to commit in advance to extract the most surplus with costly information acquisition.[11]

Interim IR Constraint

Suppose that agents sign a contract before taking any action and that the contract cannot be unilaterally disposed without the principal's consent. This leads to a model with only ex-ante IR constraint and no interim IR constraint. Although its assumption is strong, such model may allow much larger room for the full surplus extraction.

[11] This would not be a problem if an optimal action profile is the cheapest action. Indeed, the principal may have incentive not to commit in the beginning for such a case as it may leave some rents to the agents. See a common value example in Section 4.

For this case, necessary and sufficient condition for the full surplus extraction reduces to a much weaker condition, namely, necessary and sufficient condition for the existence of efficient Bayesian incentive compatible mechanism. This is because the principal can extract the full surplus as "entrance fee" when there is no interim IR constraint.

Since VCG mechanism (Vickrey[28], Clarke[6], Groves[14]) can always implement an efficient outcome for TYPE I utility functions in dominant strategy equilibrium, the full surplus extraction result is obtained without any condition for TYPE I utility functions. For TYPE II (mutually payoff relevant) or TYPE III utility functions, it can be shown that an assumption similar to Assumption I(ii) in Aoyagi[1] is a sufficient condition for the existence of efficient Bayesian incentive compatible mechanism, which is satisfied generically.[12][13]

Common Prior Assumption and Related Literature

This result is related to a violation of common prior assumption. As CM has already pointed out in [10], one of the assumptions driving their full surplus extraction result is common prior assumption. In this paper, all the agents and the principal share the common prior on the equilibrium path, but they do not off the equilibrium path. A situation without common prior is not just exogenously given, rather endogenously generated off the equilibrium path.

There are other papers which examined the relationship of common prior assumption and full surplus extraction. A recent paper by Neeman[23] questioned the notion of common prior as is used in standard mechanism design literature. Neeman showed that the full surplus extraction fails when there are two different types (with different utilities) who share the same belief about the other types. Although such information structure is nongeneric in standard settings, he argued that such structure indeed might be generic in universal type space à la Mertens and Zamir [22].[14] On the other hand, the full surplus extraction fails to hold generically *even in standard settings with common prior* in this paper. Also note that two types ((s'_i, a) and (s''_i, a'_i)) in the example who are responsible for the failure of the full surplus extraction have different beliefs about the other types.[15]

[12] Assumption 1 (ii) corresponds to $\pi(s_{-i}|s_i, a) \neq \pi\left(s_{-i}|s'_i, \left(a'_i, a_{-i}\right)\right)$ for any $\left(s'_i, a'_i\right) \neq (s_i, a_i)$ and $i \in N$ here.

[13] There are many papers which study the existence of efficient Bayesian incentive compatible (BIC-) mechanism *with budget balance* $\left(\sum_{i=1}^{n} t_i(s_{-i}|s_i) = 0 \text{ for all } s \in S\right)$. For TYPE I, many sufficient conditions are available. Most general sufficient conditions are Condition C* (d'Aspremont and Gérard-Varet[12]), LINK (Johnson, Pratt and Zeckhauser[16]), and Condition C (d'Aspremont,Crémer and Gérard-Varet[11]) , whcih turn out to be equivalent (see [11] for detail). See also Matsushima[19] and Chung[5]. For mutually payoff-relevant cases (TYPE II), see [16] and Aoyagi[1].

[14] Bergemann and Morris[2] analyzes general implications of relaxing an assumption of common prior on payoff relevant "naive" type spaces. They treat not only universal type space, but also a variety of type spaces between universal type space and naive type spaces and show that how larger type spaces lead to a stronger notion of equilibrium in usual naive type spaces.

[15] The similar point was first made by Parreiras[24].

The paper which is more closely related to this paper is Parreiras[24]. In his model, each agent draws two private information $(t_i, \theta_i) \in T_i \times \Theta_i$, where t_i is a payoff relevant type (which corresponds to s_i in this paper) and θ_i is a parameter which reflects informativeness of t_i about t_{-i}. Since θ_i is private information, agent i's true conditional distribution on T_{-i} given t_i is not common knowledge. He showed that FSE property does not hold if there are θ_i' and θ_i'' such that i's conditional distribution on T_{-i} given (t_i, θ_i') is more informative than i's conditional distribution on T_{-i} given (t_i, θ_i'') in the sense of Blackwell. In this paper, action plays a similar role to θ_i. Indeed, Blackwell condition C2 in [24] is roughly a special case of (negation of) $(*)$. So, this paper can be regarded as endogenizing such additional type as θ and generalizing the result in [24]. There is also important difference between this paper and [24]. In this paper, FSE property fails to hold for a generic subset of the space of information structure, while FSE is still a generic property for such augmented type space as $T_i \times \Theta_i$ because CM's necessary and sufficient condition still holds generically in $\Pi_{i=1}^{n} (T_i \times \Theta_i)$. The failure of full surplus extraction theorem in this paper is due to the fact that all actions but one action is not taken on the equilibrium path unlike such additional type as θ_i.[16] Some crucial type with extreme conditional belief is not present on the equilibrium path, but generated endogenously by deviations to off the equilibrium path.

4 Information Acquisition or Information Stealing?

In this section, I illustrate how a principal may fail to extract full surplus by a simple example of auction with costly information acquisition. As shown in the last section, full surplus extraction may fail when an agent can deviate to gain "essential" information about the other agents, which is not available on the equilibrium path. In the example in this section, bidders can spend their resources to either (1) gather useful information about their own value or (2) steal the other bidder's information. When the private component of the bidders' value is dominant, it is optimal for the bidders to spend their resources to acquire useful information. When the common component of the bidders' value is dominant, it is optimal not to waste any resource for information gathering activity as allocation is not an issue. For either case, it is shown that the bidders have strong incentive to steal the other bidders' information if the principal tries to implement such efficient action profile and extract the full surplus from the bidders. This example also complement the general analysis in the last section because it has much more specific structure. Although the necessity part of the last section relies on a flexibility of utility functions, this example illustrates

[16] This motivates an introduction of mixed action profile in Section 5 because a realization of action indeed serves as an additional type. Then $(*)$ is essentially CM's necessary and sufficient condition for full surplus extraction in the extended type space $\Pi_{i=1}^{n} (S_i \times A_i)$. Moreover, it is shown that $(*)$ holds generically in Section 5.

that the same results can be obtained even for utility functions which has more explicit structure.

Consider the following auction with two bidders $i = 1, 2$. Each bidder's true value is either $v > 0$ or 0. The joint distribution for $v = (v_1, v_2)$ is given by $\Pr(v_1 = v, v_2 = v) = \Pr(v_1 = 0, v_2 = 0) = \frac{1}{4} + \lambda$ and $\Pr(v_1 = v, v_2 = 0) = \Pr(v_1 = 0, v_2 = v) = \frac{1}{4} - \lambda$. Bidder i does not observe her true value before the auction, but receives a private signal $s_i \in S_i = \{h, l\}$. The accuracy of bidders' private signals depends on agents' actions. Bidder i's action set is $A_i = \{NA, IA, IS\}$, where NA stands for *No Action, IA* for *Information Acquisition,* and *IS* for *Information Stealing.* It is assumed that the cost of NA is 0, and the cost of IA and IS is $c > 0$. When IA or $NA$is chosen by agent i, s_i only depends on v_i and independent of s_j, v_j. Note that this auction is an independent private value auction when $\lambda = 0$ and a common value auction when $\lambda = \frac{1}{4}$. I assume that $\Pr(s_i = h|v_i = v)$ and $\Pr(s_i = l|v_i = 0) = \beta > \frac{1}{2}$ when NA is chosen by bidder i. When IA is chosen by one or both agents, private signals become more informative about their true values. I assume that $\Pr(s_i = h|v_i = v)$ and $\Pr(s_i = l|v_i = 0)$ is $\beta' > \beta$ if IA is chosen by both agents, and that $\Pr(s_i = h|v_i = v)$ and $\Pr(s_i = l|v_i = 0)$ is $\frac{1}{2}(\beta' + \beta)$ if only bidder i chooses IA.[17] When IS is chosen, each bidder does not create any new information about her true value, but steals the other agent's private signal. For example, if (IS, IA) is chosen, then $\Pr(s_1 = s_2) = 1$, while s_2 follows the distribution described above. I leave the type distribution conditional given (IS, IS) and v unspecified at this point as it is not relevant. Let $\Pi_i(a)$ be 2×2 conditional distribution matrix $\left(\pi\left(s_{-i}^n | s_i^m, a\right)\right)_{m,n=1}^2$ for agent i and $\Pi(a)$ be 2×2 joint distribution matrix $\left(\pi\left(s_1^m, s_2^n | a\right)\right)_{m,n=1}^2$.

The optimal action-allocation pair depends on the level of λ. For example, if $\lambda = \frac{1}{4}$, then $a^* = (NA, NA)$ and any allocation is optimal. The optimal social surplus is $\frac{v}{2}$. Needless to say, IS is not socially desirable for any λ as it takes the same cost as IA but does not create any useful information. If λ is small enough (values are enough independent), then (IA, IA) could be optimal. Suppose that (IA, IA) is chosen. When $(v_1, v_2) = (1, 1)$ or $(0, 0)$, the allocation of the good does not matter. When $(v_1, v_2) = (v, 0)$ or $(0, v)$, the good is allocated to the high type with probability $\beta'^2 + 2\beta'(1 - \beta') \times \frac{1}{2} = \beta'$, assuming that the auctioneer flips a coin to decide who win the good when $(s_1, s_2) = (l, l)$ or (h, h).[18] The optimal surplus is thus $\left\{\left(\frac{1}{4} + \lambda\right) + \left(\frac{1}{4} - \lambda\right) 2\beta'\right\} v$. Similarly, the optimal surplus would be $\left\{\left(\frac{1}{4} + \lambda\right) + \left(\frac{1}{4} - \lambda\right)\left(\beta' + \beta\right)\right\} v$ if only one bidder chooses IA and $\left\{\left(\frac{1}{4} + \lambda\right) + \left(\frac{1}{4} - \lambda\right) 2\beta\right\} v$ if (NA, NA) is chosen. It is assumed that v is so large that it is socially optimal for both bidders to choose IA when

[17] This informational externality with respect to IA is not crucial for the result. Its role is just to make (IA, IA) socially optimal action profile. This informational externality is not needed if there are more than two bidders. I restrict the number of bidders to two and employ this assumption just to make this example as simple as possible.

[18] Any other way to allocate the good in the case of tie leads to the same optimal surplus.

λ is small enough, that is

$$\begin{aligned} & \left(\frac{1}{4}+\frac{\beta'}{2}\right)v-2c \\ > & \left(\frac{1}{4}+\frac{\beta'+\beta}{4}\right)v-c \\ > & \left(\frac{1}{4}+\frac{\beta}{2}\right)v \end{aligned}$$

Using these values, the optimal action profile can be characterized for each λ.

Lemma 10 *There exists λ^* such that (IA, IA) is optimal for $\lambda \in [0, \lambda^*]$ and (NA, NA) is optimal for $\lambda \in \left[\lambda^*, \frac{1}{4}\right]$*

Proof. By assumption, $\left\{\left(\frac{1}{4}+\lambda\right)+\left(\frac{1}{4}-\lambda\right)2\beta'\right\}v-2c$ and $\left\{\left(\frac{1}{4}+\lambda\right)+\left(\frac{1}{4}-\lambda\right)2\beta\right\}v$ crosses at some $\lambda^* \in \left(0, \frac{1}{4}\right)$, which is obtained by solving

$$\begin{aligned} & \left\{\left(\frac{1}{4}+\lambda\right)+\left(\frac{1}{4}-\lambda\right)2\beta'\right\}v-2c \\ & = \left\{\left(\frac{1}{4}+\lambda\right)+\left(\frac{1}{4}-\lambda\right)2\beta\right\}v \end{aligned}$$

Thus $\lambda^* = \frac{1}{4} - \frac{c}{(\beta'-\beta)v}$. It is also easy to show that $\left\{\left(\frac{1}{4}+\lambda\right)+\left(\frac{1}{4}-\lambda\right)\left(\beta'+\beta\right)\right\}v-c$ cross with these two lines at the same λ^*. ■

Suppose that IA is the only available action, then the auctioneer can extract all the surplus as long as λ is strictly positive. Note that private signals are affiliated given (IA, IA). In particular,

$$\begin{aligned} & \Pi_i\left((IA, IA)\right) \\ = & \begin{pmatrix} \Pr\left(s_j = h | s_i = h\right), & \Pr\left(s_j = l | s_i = h\right) \\ \Pr\left(s_j = h | s_i = l\right), & \Pr\left(s_j = l | s_i = l\right) \end{pmatrix} \\ = & \begin{pmatrix} 2A & 2B \\ 2B & 2A \end{pmatrix} \end{aligned}$$

where

$$\begin{aligned} A &= \left(\frac{1}{4}+\lambda\right)\left\{\beta'^2+\left(1-\beta'\right)^2\right\}+\left(\frac{1}{4}-\lambda\right)2\beta'\left(1-\beta'\right) \\ B &= \left(\frac{1}{4}+\lambda\right)2\beta'\left(1-\beta'\right)+\left(\frac{1}{4}-\lambda\right)\left\{\beta'^2+\left(1-\beta'\right)^2\right\} \end{aligned}$$

Since $A \neq B$ if $\lambda > 0$, this matrix is full rank and so clearly satisfies CM's condition for full surplus extraction.

Now I argue that the full surplus extraction is impossible if v is large enough. Suppose that the auctioneer is able to extract the full surplus from the bidders. If a bidder is a high type, she can pretend to be a low type and get more expected utility than a low type would get. If the high type pay less than a low type would pay in equilibrium by pretending to be a low type, the high type's interim rent should be larger by certain amount than the low type's one. Since the low type's interim rent is bounded below by 0 by IR constraint, this gives a lower bound for the high type's interim rent. Since this lower bound increases as v increases, the bidder's ex-ante rent exceeds c at some point if v is large enough, contradicting the assumption of full surplus extraction. This implies that the high type needs to pay a certain amount more than a low type would pay in equilibrium by mimicking a low type.

However, it becomes more and more difficult to satisfy this incentive constraint as v becomes larger. Let $x = (\pi(l|l), \pi(h|l))$ be a vector corresponding to the conditional distribution given $s_i = l$, and $y = (\pi(l|h), \pi(h|h))$. As discussed above, the transfer function $t(\cdot|l)$ for low type has to be as in Figure 4 to make the high type pay more than a low type would pay when announcing l. It turns out that, as $v \to \infty$, the auctioneer need to choose even an extreme transfer such as $t'(\cdot|l)$ whose component parallel to $y - x$ is very large. This implies that the low type receives a large amount of money when the other bidder is also a low type and pay a large amount of money when a high type. Intuitively, the scheme extracting the full surplus needs to reward a positive correlation of announcements and punish negative correlation of announcements to cope with the high type's incentive to announce l.

However, if this is the case, bidder i can deviate to play the following strategy profitably; play IS, announce l if $s_i (= s_j)$ is l and stay out of the mechanism if s_i is h. Since a transfer from the auctioneer to agent i after $(s_i, s_j) = (l, l)$ becomes larger and larger as $v \to \infty$, the expected transfer from this strategy exceeds the cost from stealing information at some point.

Since the same argument applies when (NA, NA) is the optimal action profile (when $\lambda > \lambda^*$). The following proposition is obtained;

Proposition 11 *There exists a* $\underline{v}$ *such that no optimal action profile is FSE-implementable for any* $\lambda \in \left[0, \frac{1}{4}\right]$ *if* $v > \underline{v}$

Proof. See Appendix. ■

5 Private Strategy and Curse of Full Extraction Theorem

In the last sections, the main interest is in implementing an efficient action profile while extracting the full surplus. Thus the focus is naturally exclusively on pure strategies. This section considers an implementation of mixed action profiles. Notice that if players play a mixed action profile, they have additional private information or endogenously generated "types" before the mechanism is

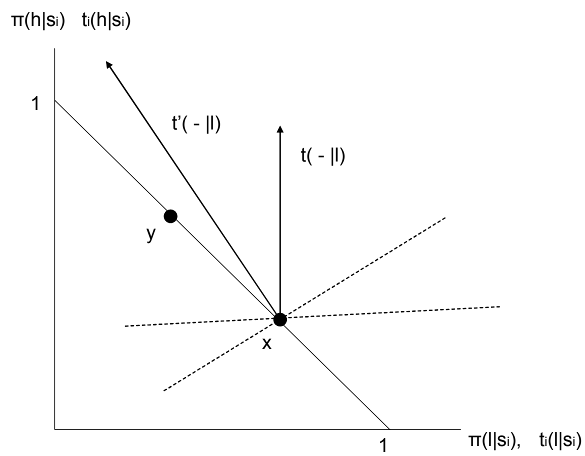

Figure 4:

played; a realization of their own behavior strategies. Since there is no reason for the principal not to use such additional private information, I allow more general mechanism in which the agents announce both their private signals and *realizations of their behavior strategies* and the agent's continuation strategies depend on both private information. This kind of strategies is sometimes called *private strategy*.[19] This generalization of the mechanism implies an expansion of players' type spaces. Then it turns out that a set of conditional distributions in such larger type space satisfies CM's necessary and sufficient condition for generic behavior strategy profiles for generic information structure. One implication of this theorem is that for any efficient action profile, there is a nearby mixed action profile which is FSE-implementable for generic information structure, thus an information structure has *virtual FSE property* generically. This result may reinforce a view expressed by Cremer and McLean and McAfee and Reny that the standard model needs some additional feature which prevents the principal from using such sophisticated scheme.

I use the last example of auction with costly information acquisition to motivate such general mechanism. Suppose that the two bidders randomize between

[19] It has been shown that private strategies are useful in many different contexts for differet reason. What is relevant here is the insight that even if private signals are independent given any action profile, a combination of private signals and a mixed action profile can be correlated as shown by Bhaskar and van Damme[4]. The private strategy in this paper, which uses the same idea but based on announcements, is similar to private strategies by Kandori[17] in the context of repeated games with imperfect public monitoring.

IA and IS. Player i plays IA with probability $\alpha_i \left(\in (0,1)\right)$ and IS with probability $1-\alpha_i$ for $i=1,2$. I assume that NA is not available for the sake of simplicity. The definition of $\Pi_i(a)$ and $\Pi(a)$ is generalized in a natural way in the extended type space $A_i \times S_i, i=1,2$. Then, bidder i's conditional distributions $\Pi_i(a)$ in bidder $j\,(\neq i)$'s extended type space $A_j \times S_j$ is obtained as follows.

	(IA,h)	(IA,l)	(IS,h)	(IS,l)
(IA,h)	$2\alpha_j A$	$2\alpha_j B$	$(1-\alpha_j)$	0
(IA,l)	$2\alpha_j B$	$2\alpha_j A$	0	$(1-\alpha_j)$
(IS,h)	$\frac{0.5\alpha_j}{0.5\alpha_j+(1-\alpha_j)(x+y)}$	0	$\frac{x(1-\alpha_j)}{0.5\alpha_j+(1-\alpha_j)(x+y)}$	$\frac{y(1-\alpha_j)}{0.5\alpha_j+(1-\alpha_j)(x+y)}$
(IS,l)	0	$\frac{0.5\alpha_j}{0.5\alpha_j+(1-\alpha_j)(w+z)}$	$\frac{w(1-\alpha_j)}{0.5\alpha_j+(1-\alpha_j)(w+z)}$	$\frac{z(1-\alpha_j)}{0.5\alpha_j+(1-\alpha_j)(w+z)}$

where $\pi\left((h,h)\,|\,(IS,IS)\right)=x, \pi\left((h,l)\,|\,(IS,IS)\right)=y, \pi\left((l,h)\,|\,(IS,IS)\right)=w$, and $\pi\left((l,l)\,|\,(IS,IS)\right)=z$.

Notice that this conditional distribution matrix $\Pi_i(\alpha)$ is full rank if and only if the following joint distribution $\Pi(a)$ on $\Pi_{i=1,2}(A_i \times S_i)$ is full rank.

$$\begin{pmatrix} \alpha_i\alpha_j\Pi\left((IA,IA)\right) & \alpha_i(1-\alpha_j)\Pi\left((IA,IS)\right) \\ (1-\alpha_i)\alpha_j\Pi\left((IS,IA)\right) & (1-\alpha_i)(1-\alpha_j)\Pi\left((IS,IS)\right) \end{pmatrix}$$

The rank of this matrix is the same as the rank of

$$\begin{pmatrix} \Pi\left((IA,IA)\right) & \Pi\left((IA,IS)\right) \\ \Pi\left((IS,IA)\right) & \Pi\left((IS,IS)\right) \end{pmatrix}$$

which is

$$\begin{pmatrix} A & B & 0.5 & 0 \\ B & A & 0 & 0.5 \\ 0.5 & 0 & x & y \\ 0 & 0.5 & w & z \end{pmatrix}$$

After a couple of rank preserving manipulation, I get

$$\begin{pmatrix} B & A & 0 & 0.5 \\ A & B & 0.5 & 0 \\ 0.5-2Ax-2By & -2Ay-2Bx & 0 & 0 \\ -2Aw-2Bz & 0.5-2Az-2Bw & 0 & 0 \end{pmatrix}$$

This matrix turns out to be full rank for almost all $x,y,w,z \geq 0$ such that $x+y+w+z=1$ if $A>B$. Since CM' s condition is satisfied in this extended type space, payment transfers can be constructed in such a way that no surplus is left to agents after any realization of action. This in turn guarantees that the agents indeed have an incentive to use a completely mixed action profile in the first place. This result is easily generalized to general games. Since any pure action profile can be arbitrary approximated by a completely mixed action profile and the full surplus can be extracted, the principal can capture virtually all the social surplus associated with any pure action profile, including the surplus-maximizing action profile.

Let me introduce some notations to introduce my final result.

Definition 12 *An action profile* $\alpha \in \Pi \triangle A_i$ *is FSE-implementable if there exists a perfect Bayesian Nash equilibrium* $((x^{\alpha}, t), \sigma)$ *in which* α *is played and all the surplus is extracted by the principal, that is,* $\sum_{s \in S} \pi(s|\alpha) \sum_{i=1}^{n} t_i(s) = W(a)$

Definition 13 *An information structure* (π, S) *has virtual FSE property in* V^X *if, for any* $V \in V^X$ *and* g*, there exists* $\alpha \in \Pi \triangle A_i$ *which is FSE implementable and satisfies* $|W(a^*) - W(\alpha)| < \varepsilon$ *for* $a^* \in A^*(V, g)$.

Then, the following general result is obtained.

Theorem 14 *Suppose that* $|S_i| \cdot |A_i| \leq \sum_j |S_j| \cdot |A_j|$ *for all* i*. Then* (π, S, A) *has virtual FSE property for any action profile for almost all information structure.*

Proof. To be completed. ■

Note that the above condition is satisfied if the number of types and actions is the same for all players.

Comments

- Suppose that the joint type distribution is independent given the optimal action profile. Then full surplus extraction fails even without hidden actions. However, this theorem implies that even in such a case, the principal may be able to extract almost all the surplus by letting the agents to play some inefficient actions with some small probability and introducing enough correlation on the extended type space. Note that *an existence of hidden actions is beneficial to the principal* for such a case contrary to the conclusion in the last sections.

- It may seem awkward to implement a mixed action profile because the principal needs to force the agents to randomize in particular way. However, the following purification argument can justify this implementation of mixed profile. Suppose that there is incomplete information in each agent's payoff and assume that the distribution of private payoff is common knowledge. Given such incomplete information, if any mechanism keeps every agent indifferent among all actions without payoff perturbation, each agent is going to randomize over actions in a particular way which is common knowledge to every agent and the principal. Thus everything would be consistent if the principal constructs a mechanism which indeed implements this particular mixed action profile. In this story, implementing a mixed action profile is totally natural. It may well be the case that the resulting mixed action profile is far from the efficient action profile. If this is the case, the principal can adjust the rent each agent would get from the mechanism in such a way that the agent will play a completely mixed action profile close to the action profile. Of course, some agents' surplus may need to be strictly positive, hence full surplus extraction may fail. However, this loss of surplus should be negligible if payoff incomplete information is very small in its size as usually assumed.

- The timing of commitment is again very important. If the principal chooses a mechanism after an action profile is chosen, then the principal can subtract all the agent's utility (not surplus) generically because she knows the joint distribution of the extended type space (although the joint distribution on S is not common knowledge).

- The above genericity result is with respect to the space Ψ. Let's examine the case of *information acquisition model.* Suppose that the model is symmetric; $|A_i| = K$ and $|S_i| = m$ for all agents. I conjecture that, if $|V|$ is larger than or equal to Km, then virtual FSE property is generic even for information acquisition models.

  There are some nongeneric cases which are worth mentioning. First, when the values are independent, virtual FSE property cannot be obtained because $\pi_i(s_{-i}|s_i, a)$ are independent of (s_i, a). For this case, the rank of the joint distribution matrix is 1. Second, if there are some two actions a_i, a_i' such that s_i given a_i' is more informative about s_{-i} in Blackwell sense than s_i given a_i, then virtual FSE property may fail (this example can be found in Parreiras[24]). This happens when s_i given a_i' is more informative about v_{-i} than s_i given a_i. For this case, the rank of the joint distribution matrix is at most $(K-1)\,m$.

  To see this, first note that $\Pr(s_{-i}|s_i, a_i) = \sum_{v_{-i}} \Pr(v_{-i}|s_i, a) \prod_{j \neq i} f(s_j|v_j, a_j)$. Since s_i given a_i' is more informative about v_{-i} in Blackwell's sense than s_i given a_i, there exists $\lambda_{s_i} \in \triangle S_i$ such that $\Pr(v_{-i}|s_i, a) = \sum_{s_i'} \lambda_{s_i}(s_i')$

$\Pr\left(v_{-i}|s_i', (a_i', a_{-i})\right)$.[20] Thus,

$$\begin{aligned}
\Pr\left(s_{-i}|s_i, a\right) &= \sum_{v_{-i}}\sum_{s_i'}\lambda_{s_i}\left(s_i'\right)\Pr\left(v_{-i}|s_i', (a_i', a_{-i})\right)\prod_{j\neq i} f\left(s_j|v_j, a_j\right) \\
&= \sum_{s_i'}\lambda_{s_i}\left(s_i'\right)\sum_{v_{-i}}\Pr\left(v_{-i}|s_i', (a_i', a_{-i})\right)\prod_{j\neq i} f\left(s_j|v_j, a_j\right) \\
&= \sum_{s_i'}\lambda_{s_i}\left(s_i'\right)\Pr\left(s_{-i}|s_i', (a_i', a_{-i})\right)
\end{aligned}$$

Finally, let $\alpha_{-i} \in \prod_{j\neq i} \triangle A_j$ be any mixed action profile by $j \neq i$. Since agent i's signal does not contain any information about the other players realized action profile, $\Pr\left(s_{-i}, a_{-i}|s_i, a_i\right) = \alpha_{-i}\left(a_{-i}\right)\Pr\left(s_{-i}|s_i, a\right)$. Therefore,

$$\begin{aligned}
\Pr\left(s_{-i}, a_{-i}|s_i, a_i\right) &= \alpha_{-i}\left(a_{-i}\right)\Pr\left(s_{-i}|s_i, a\right) \\
&= \alpha_{-i}\left(a_{-i}\right)\sum_{s_i'}\lambda_{s_i}\left(s_i'\right)\Pr\left(s_{-i}|s_i', (a_i', a_{-i})\right) \\
&= \sum_{s_i'}\lambda_{s_i}\left(s_i'\right)\alpha_{-i}\left(a_{-i}\right)\Pr\left(s_{-i}|s_i', (a_i', a_{-i})\right) \\
&= \sum_{s_i'}\lambda_{s_i}\left(s_i'\right)\Pr\left(s_{-i}, a_{-i}|s_i', a_i'\right)
\end{aligned}$$

This implies that $\Pr\left(s_{-i}, a_{-i}|s_i, a_i\right)$ is a linear combination of $\left\{\Pr\left(s_{-i}, a_{-i}|s_i', a_i'\right)\right\}_{s_i'\in S_i}$ for any s_i and any mixed action profile α_{-i}. Thus the rank of the joint distribution matrix needs to be reduced by at least m.

6 Conclusion

This paper extended Cremer and McLean's full surplus extraction theorem to the situation in which a principal commits to a mechanism first, then agents

[20] To see this, let $\widetilde{Y}$ is a garbling of Y with respect to X. The density functions for X, Y given X, and $\widetilde{Y}$ given X are given by $f\left(x\right), g\left(y|x\right), \sum_y h\left(\widetilde{y}|y\right) g\left(y|x\right) f\left(x\right)$, where $\sum_{\widetilde{y}} h\left(\widetilde{y}|y\right) = 0$ for each y. Then,

$$\begin{aligned}
\Pr\left(X = x'|\widetilde{Y} = y'\right) &= \frac{\sum_y h\left(y'|y\right) g\left(y|x'\right) f\left(x'\right)}{\sum_x\sum_y h\left(y'|y\right) g\left(y|x\right) f\left(x\right)} \\
&= \frac{\sum_y h\left(y'|y\right)\sum_x g\left(y|x\right) f\left(x\right)\Pr\left(X = x'|Y = y\right)}{\sum_x\sum_y h\left(y'|y\right) g\left(y|x\right) f\left(x\right)} \\
&= \sum_y \lambda_{y'}\left(y\right)\cdot\Pr\left(X = x'|Y = y\right)
\end{aligned}$$

where $\lambda_{y'}\left(y\right) = \frac{h\left(y'|y\right)\sum_x g\left(y|x\right)f\left(x\right)}{\sum_x\sum_y h\left(y'|y\right)g\left(y|x\right)f\left(x\right)}$ $\left(= \Pr\left(Y = y|\widetilde{Y} = y'\right)\right)$.

take actions and draw their types. The agents' actions not only affects the value of the allocation but also the distribution of private types as well. Hence, the distribution of types is determined endogenously rather than exogenously. One special example of such situation is auction with costly information acquisition.

By doing so, it is shown that full surplus extraction does not hold generically as it does in the standard mechanism design framework where the distribution of types is given exogenously. This paper suggests that a possibility of manipulating an underlying type distribution keeps principal from extracting the full surplus and even force her to use a simple mechanism.

However, it is also shown that full surplus extraction is possible for generic information structure if the principal can use a more general mechanism in which the agents play a mixed action profile and announce both realization of their actions and private types.

7 Appendix: Proof of the Theorem

Proof of Theorem 6

Proof. (The sufficiency part): I only need to prove sufficiency for TYPE III utility functions. Suppose that the assumption $(*)$ is satisfied. Then, for each s_i, there exists $w_i(\cdot|s_i) \in \Re^{|S_{-i}|}$ such that $\sum_{s_{-i}\in S_{-i}} \pi(s_{-i}|s_i,a)\, w_i(s_{-i}|s_i) = 0$ and $\sum_{s_{-i}\in S_{-i}} \pi(s_{-i}|s_i',(a_i',a_{-i}))\, w_i(s_{-i}|s_i) > 0$ for all $(s_i',a_i') \neq (s_i,a_i)$.[21] Let $q_i(s_i,a) = \sum_{s\in S} \pi(s_{-i}|a,s_i)\, V_i(x^a(s),s,a) + \{g_i(a_i) - g_i^*\}$ be type s_i's interim expected surplus given $a \in A$. Define the transfer for agent i by $t_i(s_{-i}|s_i) = q_i(s_i,a) + \lambda w_i(s_{-i}|s_i)$ where λ is some positive number. By definition, this mechanism extracts the full surplus from agent i if a and x are successfully implemented.

Three conditions should be checked; (1) agent i has to play $a_i(IC_{a_i})$ (2) enter the mechanism ($IR_{interim}$), and (3) reveal her true type (IC_{s_i}) on the equilibrium path (after a_i is chosen). For IC_{s_i}, the following inequality should be satisfied for all $s_i \in S_i$;

$$\begin{aligned} & \sum_{s_{-i}\in S_{-i}} \pi(s_{-i}|a,s_i)\{V_i(x^a(s),s,a) - t_i(s_{-i}|s_i)\} \\ \geq & \sum_{s_{-i}\in S_{-i}} \pi(s_{-i}|a,s_i)\{V_i(x^a(s_i',s_{-i}),s,a) - t_i(s_{-i}|s_i')\} \end{aligned}$$

These constraints are clearly satisfied if λ is chosen large enough.

It is also easy to see that $IR_{interim}$ constraints are satisfied on the equilibrium path. For off the equilibrium path, note that the following inequalities are satisfied for any $a_i' \neq a_i$ and all s_i' and s_i'' if λ is large enough;

$$-\eta > \sum_{s_{-i}\in S_{-i}} \pi(s_{-i}|(a_i',a_{-i}),s_i)\{V_i(x^a(s_i'',s_{-i}),(s_i',s_{-i}),(a_i',a_i)) - t_i(s_{-i}|s_i'')\}$$

This means that agent i's $IR_{interim}$ constraints are violated for every type if a_i is not chosen. Thus agent i's most attractive deviation with respect to action is to choose the action which maximizes g_i and stay out of the mechanism independent of her realized types. Since this leads to the same ex-ante expected utility ($= g_i^*$) as agent i would get by playing a_i on the equilibrium path, IC_{a_i} is also satisfied.

This proves that any $a \in A$ can be implemented and all the surplus can be extracted by the principal if λ is chosen large enough.

(The necessity part) Suppose that there exists $\{\pi(\cdot|s_{i,k},(a_{i,k},a_{-i}))\}_{k=1}^K$ for which $(*)$ is violated for some s_i^* and a^*. It can be taken to be a minimal set of such conditional distributions without loss of generality. I derive a contradiction by assuming that the principal can extract the full surplus for any (V,g). I first prove necessity for TYPE III utility functions.

[21] This is easily derived, for example, from Theorem 2.6(P.44) in Gale [13].

(TYPE III utility functions): Choose a set of continuous functions V_i and $g_i, i = 1, ..., n$ so that a^* is the unique surplus maximizing action. Such V_i and g_i clearly exists. Now modify agent i's utility by $\widehat{V}_i(x, s, a) = V_i(x, s, a) + \gamma\delta(s_i, a_i)$ and some agent j's utility ($j \neq i$) by $\widehat{V}_j(x, s, a) = V_j(x, s, a) - \gamma\delta(s_i, a_i)$, where $\delta : S_i \times A_i \to \{0, 1\}$ and $\delta(s_i, a_i) = 1$ if and only if $(s_i, a_i) = (s_{i,1}, a_{i,1})$. Note that such change of utility functions affects neither the unique optimal action profile nor the optimal allocation $x^a(s)$.

One possible deviation for agent i would be to play $a_{i,1}$ and announce s_i^* when $s_{i,1}$ is observed and stay out of the mechanism otherwise. So the following inequality should be satisfied;

$$\begin{aligned}
& \sum_{s_{-i}\in S_{-i}} \pi(s_{-i}|s_i^*, a^*) \left\{ V_i\left(x^{a^*}(s_i^*, s_{-i}), s, a^*\right) - t_i(s_{-i}|s_i^*) \right\} + \eta \\
\geq\ & 0 \\
\geq\ & p\left(s_{i,1}|\left(a_{i,1}, a_{-i}^*\right)\right) \sum_{s_{-i}\in S_{-i}} \left\{ \begin{array}{c} \pi\left(s_{-i}|s_{i,1}, \left(a_{i,1}, a_{-i}^*\right)\right) \cdot \\ \left(\begin{array}{c} V_i\left(x^{a^*}(s_i^*, s_{-i}), (s_{i,1}, s_{-i}), \left(a_{i,1}, a_{-i}^*\right)\right) \\ +\gamma - t_i(s_{-i}|s_i^*) \end{array} \right) \end{array} \right\} \\
& + \left(g_i(a_{i,1}) - g_i^*\right) \\
\geq\ & \sum_{s_{-i}\in S_{-i}} \left\{ \pi\left(s_{-i}|s_{i,1}, \left(a_{i,1}, a_{-i}^*\right)\right) \left(\begin{array}{c} V_i\left(x^{a^*}(s_i^*, s_{-i}), (s_{i,1}, s_{-i}), \left(a_{i,1}, a_{-i}^*\right)\right) \\ +\gamma - t_i(s_{-i}|s_i^*) \end{array} \right) \right\} \\
& + \frac{g_i(a_{i,1}) - g_i^*}{p\left(s_{i,1}|\left(a_{i,1}, a_{-i}^*\right)\right)}
\end{aligned}$$

where the first inequality is s_i's $IR_{interim}$, the second inequality comes from the assumption of full surplus extraction and IC_{a_i}.

This implies that

$$\sum_{s_{-i}\in S_{-i}} \left\{ \pi\left(s_{-i}|s_{i,1}, \left(a_{i,1}, a_{-i}^*\right)\right) - \pi(s_{-i}|s_i^*, a^*) \right\} t_i(s_{-i}|s_i^*) \geq \gamma + R_1 \qquad (1)$$

where R_1 is some number independent of γ and $t(\cdot|\cdot)$.

For each $k > 1$, consider a similar deviation; play $a_{i,k}$ and announce s_i when $s_{i,k}$ is observed and stay out of the mechanism otherwise. Then, again by $IR_{interim}$ & full surplus extraction & IC_{a_i}, at least the following inequalities must be satisfied for $k = 2, ..., K$.[22]

$$\begin{aligned}
& \sum_{s_{-i}\in S_{-i}} \pi(s_{-i}|s_i^*, a^*) \left\{ V_i\left(x^{a^*}((s_i^*, s_{-i})), s, a^*\right) - t_i(s_{-i}|s_i^*) \right\} + \eta \\
\geq\ & 0 \\
\geq\ & p\left(s_{i,k}|\left(a_{i,k}, a_{-i}^*\right)\right) \sum_{s_{-i}\in S_{-i}} \left\{ \begin{array}{c} \pi\left(s_{-i}|s_{i,k}, \left(a_{i,k}, a_{-i}^*\right)\right) \cdot \\ \left(\begin{array}{c} V_i\left(x^{a^*}(s_i^*, s_{-i}), (s_{i,k}, s_{-i}), \left(a_{i,k}, a_{-i}^*\right)\right) \\ -t_i(s_{-i}|s_i^*) \end{array} \right) \end{array} \right\}
\end{aligned}$$

[22] Note that it could be the case that $s_{i,k} = s_{i,1}$. The incentive constraint is even stronger in such a case.

$$
\begin{aligned}
&+ \left(g_i\left(a_{i,k}\right) - g_i^*\right) \\
\geq & \sum_{s_{-i}\in S_{-i}} \pi\left(s_{-i}|s_{i,k},\left(a_{i,k},a_{-i}^*\right)\right)\left\{\begin{array}{c} V_i\left(x^{a^*}\left(s_i^*,s_{-i}\right),\left(s_{i,k},s_{-i}\right),\left(a_{i,k},a_{-i}^*\right)\right) \\ -t_i\left(s_{-i}|s_i^*\right)\end{array}\right\} \\
&+ \frac{g_i\left(a_{i,k}\right) - g_i^*}{p\left(s_{i,k}|\left(a_{i,k},a_{-i}^*\right)\right)}
\end{aligned}
$$

which leads to similar inequalities for $k = 2, ..., K$.

$$
\sum_{s_{-i}\in S_{-i}} \left\{\pi\left(s_{-i}|s_{i,k},\left(a_{i,k},a_{-i}^*\right)\right) - \pi\left(s_{-i}|a^*,s_i^*\right)\right\} t_i\left(s_{-i}|s_i^*\right) \geq R_k \qquad (2)
$$

for some R_k independent of γ and $t(\cdot|\cdot)$.

By assumption, there exists $\alpha_k > 0$ such that $\pi(\cdot|a^*, s_i) = \sum_{k=1}^{K} \alpha_k \pi\left(\cdot|s_{i,k},\left(a_{i,k},a_{-i}^*\right)\right)$. Summing up (1) and (2), the following inequality is obtained for any $t(\cdot|\cdot)$,

$$
0 \geq \alpha_1 \gamma + \sum_{k=1}^{K} \alpha_k R_k
$$

This cannot be satisfied for any transfer function if γ is large enough. This is a contradiction.[23]

(TYPE II utility functions): By Condition R^{II}, it should be the case that $s_{i,k} \neq s_i^*$ for some k. It is assumed that $k = 1$ without loss of generality. Choose a set of continuous functions V_i and $g_i, i = 1, ..., n$ so that a^* is the unique surplus maximizing action. Modify agent i's utility by $\widehat{V}_i(x, s_i) = V_i(x, s_i) + \gamma\delta(s_i)$ and some agent j's utility $(j \neq i)$ by $\widehat{V}_j(x, s) = V_j(x, s) - \gamma\delta(s_i)$, where $\delta : S_i \to \{0, 1\}$ and $\delta(s_i) = 1$ if and only if $s_i = s_{i,1}$. Note that such change of utility functions affects neither the unique optimal action profile nor the optimal allocation $x^a(s)$. The rest of the proof is exactly the same as the proof for TYPE III. ■

Proof of Theorem 8

Proof. Since sufficiency is already proved for TYPE III utility functions, I only need to prove necessity. Since condition R^I implies R^{II}, it should be the case that $s_{i,k} \neq s_i^*$ for some k. It is again assumed that $k = 1$ without loss of generality. Choose a set of continuous functions V_i and $g_i, i = 1, ..., n$ so that a^* is the unique surplus maximizing action. Modify agent i's utility by $\widehat{V}_i(x, s_i) = V_i(x, s_i) + \gamma\delta(s_i)$, where $\delta : S_i \to \{0, 1\}$ and $\delta(s_i) = 1$ if and only if $s_i = s_{i,1}$. Also modify agent j's utility $(j \neq i)$ by $V_j(x, s_j) = V_j(x, s_j) + \lambda v_j(s_j)$, where $v_j : S_j \to \Re^+, j \neq i$ satisfies $\sum_s \pi(s|a^*) \sum_{j\neq i} v_j(s_j) > \sum_s \pi\left(s|\left(a_i, a_{-i}^*\right)\right) \sum_{j\neq i} v_j(s_j)$

[23] When γ is large, V_j may become negative, thus violate the original assumption $V_j(x, s, a) \geq 0$ for all (x, s, a). In such a case, a positive constant term can be added to every agent's utility function so that nonnegative utility is obtained without affecting the proof.

for any $a_i \neq a_i^*$. Such $v_j, j \neq i$ exists by Lemma 7. Agent j's g_j are also modified by $\widehat{g}_j(a_j) = g_i(a_j) - \xi\delta_j(a_j)$ for $j \neq i$, where $\delta_j : A_j \rightarrow \{0,1\}$ and $\delta_j(a_j) = 1$ if and only if $a_j \neq a_j^*$. Three parameters (γ, λ, ξ) are chosen as follows. For each $\gamma > 0$, λ is chosen large enough to make a^* optimal among all $\left(a_i, a_{-i}^*\right)$. Next, choose ξ large enough to offset any gain from $v_j(s)$ for $a_{-i} \neq a_{-i}^*$ so that any such action profile by $j \neq i$ cannot be optimal. This guarantees that a^* continues to be the unique surplus-maximizing action and the optimal allocation remains the same for any level of $\gamma > 0$. The rest of the proof is again exactly the same as before. ■

Proof of Proposition 11

Proof. I first prove that (IA, IA) cannot be implementable when v is large. Suppose that (IA, IA) is FSE-implementable. Agent i's incentive constraint given $s_i = h$ is

$$\left\{\Pr\left(\left(s_j, v_i\right) = (l, v) \,|h\right) + \frac{1}{2}\Pr\left(\left(s_j, v_i\right) = (h, v) \,|h\right)\right\} - \sum \pi\left(s_j|h\right) t_i\left(s_j|h\right)$$
$$\geq \frac{1}{2}\Pr\left(\left(s_j, v_i\right) = (l, v) \,|h\right) v - \sum \pi\left(s_j|h\right) t_i\left(s_j|l\right)$$

So,

$$\sum \pi\left(s_j|h\right)\left(t_i\left(s_j|l\right) - t_i\left(s_j|h\right)\right) \tag{3}$$
$$\geq -\frac{1}{2}\left\{\Pr\left(\left(s_j, v_i\right) = (l, v) \,|h\right) + \Pr\left(\left(s_j, v_i\right) = (h, v) \,|h\right)\right\} v$$
$$= -\frac{\beta' v}{2}$$

Let R_h be agent i's interim expected rent conditional on $s_i = h$. Then,

$$\sum \pi\left(s_j|h\right) t_i\left(s_j|h\right) \tag{4}$$
$$= \left\{\Pr\left(\left(s_j, v_i\right) = (l, v) \,|h\right) + \frac{1}{2}\Pr\left(\left(s_j, v_i\right) = (h, v) \,|h\right)\right\} v - R_h$$
$$= \left\{\left(\frac{1}{4} + \lambda\right)\left(2\beta' - \beta'^2\right) + \left(\frac{1}{4} - \lambda\right)\left(\beta' + \beta'^2\right)\right\} v - R_h$$

Let R_l be agent i's interim expected rent conditional on $s_i = l$. Then,

$$\sum \pi\left(s_j|l\right) t_i\left(s_j|l\right) \tag{5}$$
$$= \frac{1}{2}\Pr\left(\left(s_j, v_i\right) = (l, v) \,|l\right) v - R_l$$
$$= \left\{\left(\frac{1}{4} + \lambda\right)\left(1 - \beta'\right)^2 + \left(\frac{1}{4} - \lambda\right)\beta'\left(1 - \beta'\right)\right\} v - R_l$$

Since the auctioneer extracts the full surplus of agent i, her ex-ante expected rent from the mechanism should be exactly c. Since s_i is 1 or 0 with equal probability, $\frac{1}{2}R_h + \frac{1}{2}R_l = c$. Note also that R_h and R_l are nonnegative because of the individual rationality constraint. This implies that $R_h, R_l \in [0, 2c]$

From (3), (4) and (5), I can obtain

$$\sum \left(\pi\left(s_j|h\right) - \pi\left(s_j|l\right)\right) t_i\left(s_j|l\right) \tag{6}$$
$$= \sum \pi\left(s_j|h\right)\left(t_i\left(s_j|l\right) - t_i\left(s_j|h\right)\right) + \sum \pi\left(s_j|h\right) t_i\left(s_j|h\right) - \sum \pi\left(s_j|l\right) t_i\left(s_j|l\right)$$
$$\geq \left\{\frac{2\beta' - 1}{4} - \lambda\left(2\beta' - 1\right)^2\right\} v - 2c$$

Since $\pi(h|h) - \pi(h|l)$ is $4\lambda\left(2\beta' - 1\right)^2$ and similarly $\pi(l|h) - \pi(l|l)$ is $-4\lambda\left(2\beta' - 1\right)^2$, $\sum\left(\pi(s_j|h) - \pi(s_j|l)\right)t_i(s_j|l)$ is $4\lambda\left(2\beta' - 1\right)^2\left(t_i(h|l) - t_i(l|l)\right)$.[24] Thus the above inequality can be simplified to be

$$4\lambda\left(2\beta' - 1\right)\left(t_i(h|l) - t_i(l|l)\right) \geq \left\{\frac{1}{4} - \lambda\left(2\beta' - 1\right)\right\}v - 2c \tag{7}$$

Since $\left\{\frac{1}{4} - \lambda\left(2\beta' - 1\right)\right\} > 0$, $t_i(h|l) - t_i(l|l)$ goes to infinity as v goes to infinity. Now I show that in particular $t_i(l|l)$ needs to go to $-\infty$. By (5),

$$\begin{aligned} & \pi(h|l)\left(t_i(h|l) - t_i(l|l)\right) + t_i(l|l) \\ = \; & \frac{1}{2}\Pr\left((s_j, v_i) = (l, v)\,|l\right)v - R_l \end{aligned}$$

Thus,

$$t_i(h|l) - t_i(l|l) = \frac{\frac{1}{2}\Pr\left((s_j, v_i) = (l, v)\,|l\right)v - R_l - t_i(l|l)}{\pi(h|l)}$$

Substituting this into (7) and simplifying it, I can obtain

$$\begin{aligned} & -4\lambda\left(2\beta' - 1\right)\frac{t_i(l|l) + R_l}{\pi(h|l)} \\ \geq \; & \frac{\left\{\frac{1}{4} - \lambda\left(2\beta' - 1\right)\right\}^2}{\frac{1}{4} - \lambda\left(2\beta' - 1\right)^2}v - 2c \end{aligned} \tag{8}$$

This inequality is clearly violated for $\lambda = 0$, but satisfied for small enough $t_i(l|l)$ for each v, which goes to $-\infty$ as $v \to \infty$. Moreover, for any $K > 0$, I can choose v large enough so that $-t_i(l|l)$ needs to be larger than K for full surplus extraction for any $\lambda > 0$. Now pick $\underline{v}$ so that $-\frac{1}{2}t_i(l|l) - c > 0$ for any $v > \underline{v}$ and any $\lambda > 0$. Given such $t_i(l|l)$, agent i has a profitable deviation. She can deviate to play IS to obtain a positive surplus. After $s_i = h$, agent i is guaranteed to receive at least 0. Agent i can receive $-t_i(l|l)$ after $s_i = l$, which occurs with probability $\frac{1}{2}$. Hence agent i's expected total rent is at least $-\frac{1}{2}t_i(l|l) - c$, which is strictly positive by assumption. Thus (8) is not satisfied for any λ, which is a contradiction.

The same proof can show that (NA, NA) is not implementable either (although the range for which (NA, NA) is optimal $\left(\left[\lambda^*, \frac{1}{4}\right]\right)$ vanishes as $v \to \infty$) by setting $A = B = 0$ and using β instead of β' in the above proof. ■

---

[24]

$$\begin{aligned} & \pi(h|h) - \pi(h|l) \\ = 2 & \left\{ \begin{array}{c} \left(\frac{1}{4} + \lambda\right)\left\{(1 - \beta')^2 + \beta'^2 - 2\beta'(1 - \beta')\right\} \\ + \left(\frac{1}{4} - \lambda\right)\left\{2\beta'(1 - \beta') - (1 - \beta')^2 - \beta'^2\right\} \end{array} \right\} \\ = \; & 4\lambda\left\{\left(1 - \beta'\right)^2 + \beta'^2 - 2\beta'\left(1 - \beta'\right)\right\} \\ = \; & 4\lambda\left(2\beta' - 1\right)^2 \end{aligned}$$

References

[1] M. Aoyagi. Correlated types and bayesian incentive compatible mechanisms with budget balance. *Journal of Economic Theory*, 79:142–151, 1998.

[2] D. Bergemann and S. Morris. Robust mechanism design. 2002.

[3] D. Bergemann and J. Valimaki. Information acquisition and efficient mechanism design. *Econometrica*, 70:1007–1033, 2002.

[4] V. Bhaskar and E. van Damme. Moral hazard and private monitoring. *Journal of Economic Theory*, 102:16–39, 2002.

[5] K.-S. Chung. A note on matsushima's regularity condition. *Journal of Economic Theory*, 87:429–433, 1999.

[6] E. H. Clarke. Multipart pricing of public goods. *Public Choice*, 2:19–33, 1971.

[7] J. Cremer and F. Khalil. Gathering information before signing a contract. *American Economic Review*, 82:566–578, 1992.

[8] J. Cremer, F. Khalil, and J.-C. Rochet. Strategic information gathering before a contract is offered. *Journal of Economic Theory*, 81:163–200, 1998.

[9] J. Cremer and R. McLean. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53:345–361, 1985.

[10] J. Cremer and R. McLean. Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56:1247–1257, 1988.

[11] C. d'Aspremont, J. Cremer, and L.-A. Gerard-Varet. Balanced bayesian mechanisms. 2002. Mimeo.

[12] C. d'Aspremont and L.-A. Gerard-Varet. Bayesian incentive compatible beliefs. *Journal of Mathematical Economics*, 10:83–103, 1982.

[13] D. Gale. *The Theory of Linear Economic Models*. The University of Chicago Press, 1960.

[14] T. Groves. Incentives in teams. *Econometrica*, 41:617–631, 1973.

[15] L. Hurwicz. On information decentralized systems. In R. Radner and C. McGuire, editors, *Decision and Organization*, pages 297–336. Amsterdam: North Holland, 1972.

[16] S. Johnson, J. W. Pratt, and R. Zeckhauser. Efficiency despite mutually payoff-relevant private information: The finite case. *Econometrica*, 58:873–900, 1990.

[17] M. Kandori. Randomization, communication, and efficiency in repeated games with imperfect public monitoring. forthcoming in Econometrica, 2003.

[18] T. R. Lewis and D. E. M. Sappington. Information management in incentive problems. *Journal of Political Economy*, 105:796–821, 1997.

[19] H. Matsushima. Incentive compatible mechanisms with full transferability. *Journal of Economic Theory*, 54:198–203, 1991.

[20] S. Matthews. Information acquisition in competitive bidding process. In M. Boyer and R. Kihlstrom, editors, *Bayesian Models in Economic Theory*, pages 181–207. Amsterdam: North-Holland, 1984.

[21] R. P. McAfee and P. J. Reny. Correlated information and mechanism design. *Econometrica*, 60:395–421, 1992.

[22] J. F. Mertens and S. Zamir. Formulation of bayesian analysis for games with incomplete information. *International Journal of Game Theory*, 10:619–632, 1985.

[23] Z. Neeman. The relevance of private information in mechanism design. 2003. Mimeo.

[24] S. Parreiras. Correlated information, mechanism design and informational rents. 2002. Mimeo.

[25] N. Persico. Information acquisition in auctions. *Econometrica*, 68:135–148, 2000.

[26] J. Robert. Continuity in auction design. *Journal of Economic Theory*, 55:169–179, 1991.

[27] M. Stegeman. Participation costs and efficient auctions. *Journal of Economic Theory*, 71:228–259, 1996.

[28] W. Vickrey. Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance*, 16:8–37, 1961.