Individually Rational, Balanced-Budget Bayesian Mechanisms and the Informed Principal Problem.

Grigory Kosenok and Sergei Severinov *

First Version: June 3, 2002
This Version: October 7, 2004

Abstract

We investigate the issue of implementation via individually rational ex-post budget-balanced Bayesian mechanisms. We show that all decision rules generating a nonnegative expected social surplus are implementable if and only if the probability distribution of the agents' types satisfies two conditions: the well-known condition of Crémer and McLean (1985) and the *Identifiability* condition introduced in this paper. These conditions are also necessary for ex-post efficiency to be attainable. The expected social surplus in these mechanisms can be distributed in any desirable way. The Identifiability condition, as well as Crémer-McLean condition, are generic when there are at least three agents, and none of them has more types than the number of type profiles of the other agents. Also generically, any ex-post efficient decision rule can be implemented in an informed principal framework, i.e. when the mechanism is offered by an informed participant. Only ex-post efficient mechanisms allocating all surplus to the party designing the mechanism are both sequential equilibrium outcomes and *neutral optima*, i.e. can never be blocked. Thus, an informed principal can also extract all surplus from the other agents in a Bayesian mechanism.

JEL Nos: C72, D82. Keywords: mechanism design, Bayesian implementation, individual rationality, ex-post budget balancing, surplus allocation, informed principal.

1   Introduction.

This paper focuses on three issues. First, we explore the existence of individually rational ex-post budget-balanced Bayesian mechanisms for the implementation of a broad class of decision rules[1], including ex-post efficient ones. Then we study the related issue of surplus allocation in such mechanisms. We also explore the issue of implementation in the informed principal framework where the mechanism is designed not by an outsider, but by one of the agents after she has learned her private information.

The theory of Bayesian mechanism design provides a universally accepted implementation tool used to study a large variety of environments, including contracting, auctions, bargaining, etc. For this reason, it is important to understand the scope and limits of Bayesian implementation. One of the well-known issues is the tension between budget balancing, individual rationality and efficiency which one can reasonably view as desirable properties of a mechanism. Myerson and Satterthwaite (1983) have shown that Bayesian mechanisms possessing these properties could fail to exist when private information is independently distributed across agents.[2] More precisely, Makowski and Mezzetti (1994) demonstrate that such mechanisms exist if and only if the ex-ante expected deficit in the corresponding Groves mechanism does not exceed the sum of the fees that the agents can be charged ex-ante for participating in the mechanism. Krishna and Perry (1998) and Williams (1999) extend this result to the multidimensional setting.

This condition fails in many economically important situations. For example, Rob (1989) demonstrates that public good would not be provided efficiently. Mailath and Postlewaite (1990) show that in any feasible mechanism the probability of the public good provision goes to zero as the number of agents increases. With multidimensional independently distributed types, Jehiel and Moldovanu (2001) establish that efficient mechanisms exist only in nongeneric situations.

Relaxing one of these three requirements makes it possible to obtain positive results. Various sufficient conditions for efficient Bayesian implementation with ex-post budget balancing but without individual rationality requirement have been derived by d'Aspremont and Gérard-Varet (1979), d'Aspremont, Crémer and Gérard-Varet (1990) and (2003) (Compatibility), Matsushima (1991), Aoyagi (1998), and Chung (1999) (Weak and Strict Regularity), Fudenberg, Levine and Maskin (1996) (Pairwise Identifiability).

Finally, d'Aspremont, Crémer and Gérard-Varet (2004) present a necessary and sufficient condition for balanced-budget Bayesian implementation without individual rationality. We discuss the connection between their and our results in Subsection 3.3.

Crémer and McLean (1985) and (1988) and McAfee and Reny (1992) demonstrate that an uninformed mechanism designer can implement an ex-post efficient decision rule, preserve individual rationality, and extract all expected surplus from the agents if, relative to the prior, an agent's type contains additional information about the other agents' types. The mechanism designer exploits this statistical interdependence to cross-check agents' reports, thereby inducing each agent to reveal her type truthfully without leaving any informational rent to her. Naturally, such mechanism is not ex-post budget-balanced. The uninformed mechanism designer plays an important role of a budget breaker. She collects transfers from

---

[1] A decision rule is a mapping from the set of agents' type profiles into the set of public decisions.

[2] Myerson and Satterthwaite (1983) focus on the case of continuous probability distribution of private information. But their result also extends to the case of discrete probability distributions.

the agents, and may also have to pay them in some states of the world.[3] Without imposing ex-post budget balance, McLean and Postlewaite (2002) and (2003) show that only small transfers from/to each agent are needed to construct individually rational, ex-post efficient mechanisms when each agent is 'informationally small.' That is, even if an agent misrepresents her private information, the state of the world can still be inferred with a high degree of accuracy provided that all other agents report truthfully.

We study implementation under all the three properties in question - interim individual rationality, ex-post budget balancing and ex-post efficiency. Our main result establishes necessary and sufficient conditions for the existence of mechanisms possessing these three properties in the transferable utility framework. The first is already mentioned condition of Crémer and McLean (1985). The second is *Identifiability* condition introduced in this paper.

In fact, these conditions guarantee that not only ex-post efficient but all *ex-ante socially rational*[4] decision rules- the ones that generate a nonnegative expected social surplus- are implementable via an interim individually rational and ex-post budget-balanced mechanism. It is easy to see that ex-ante social rationality of the decision rule is necessary for interim individually rationality and ex-post budget-balancing, so our sufficiency result cannot be extended further.

Intuitively, the Identifiability condition says the following. Suppose that the agents' reporting strategies in a direct mechanism induce a probability distribution of reported type profiles that is different from the true prior from which the types are drawn. Then there exists an agent and her type such that, when this type is announced, the mechanism designer knows that this agent has not committed a unilateral deviation from truthtelling. Thus, for any probability distribution over reported type profiles, identifiability allows to determine a non-empty set of *agent types* who have not deviated unilaterally. The Identifiability condition is generic when there are at least three agents and none of them has more types than the number of type profiles of all other agents.

To understand the role of the Identifiability condition, consider Crémer-McLean mechanism as a benchmark. In this mechanism, the principal extracts information about an agent's type by offering her a lottery which has zero expected value if this agent reports her type truthfully and a negative expected value if the agent misrepresents her type.

If we attempt to use this approach in an environment where an outside budget-breaker is not available and so ex-post budget balance has to hold, then we have to allocate the transfers from such lotteries to agents in an incentive compatible way. This creates an additional incentive problem. For example, designating agent j to receive transfers from the lottery given to agent i may induce j to misrepresent her type in a way that makes a truthful report by i to appear untruthful which would cause i to make transfers to j.

The Identifiability Condition allows to resolve this issue and preserve individual rationality. A detailed explanation of our mechanism is provided in Section 3. Here we just point out that the identifiability condition allows to construct an individually rational, budget-balanced system of transfers that not only gives a negative expected payoff to an agent misrepresenting her type when other agents report truthfully, but also does not allow any agent to benefit by imitating a deviation by another agent.

Importantly, the Identifiability and Crémer-McLean conditions are necessary. If either

---

[3]Ex-ante budget balance can be attained in Crémer-McLean mechanism if the mechanism designer pays each agent an amount equal to her ex-ante expected transfer in the ensuing mechanism.

[4]To the best our knowledge, this term was coined by d'Aspremont and Gérard-Varet (1979). Obviously, the class of ex-ante socially rational decision rules includes ex-post efficient ones as a special case.

one fails, then efficient, budget-balanced implementation preserving individual rationality is impossible under some profiles of the utility functions.

Under ex-post budget balance all surplus generated by the mechanism is distributed among the agents and is not extracted by the outside mechanism designer as in Crémer and McLean (1985) and (1988) and McAfee and Reny (1992). So, it is natural to consider how the surplus can be allocated among the agents. We show that generically this can be done in an arbitrary way. That is, when the Identifiability and Crémer-McLean conditions hold, there is a mechanism distributing (ex-ante) expected social surplus across *agent types* in any desired way.

The allocation of surplus result has important implications for the analysis of the so-called informed principal problem. This problem arises when an uninformed principal is not available to design a mechanism, and it has to be designed by one of the agents who has already learned her private information. In this case, the choice of a mechanism by an agent serves as a signal to the other agents about the designer's type, and so the equilibrium mechanism has to balance the interests of different types of the designer. These aspects make the analysis of the informed principal problem more complex. Mechanism design by informed principal has been studied by Myerson (1983), and Maskin and Tirole (1990) and (1992) who propose several solution concepts to this problem. However, except for some cases, the issue of characterizing the outcomes of this game has not been completely resolved.

We contribute to the study of the informed principal problem by demonstrating that an ex-post efficient decision rule can generically be implemented in this environment. Specifically, an ex-post efficient allocation profile that gives all social surplus to the agent designing the mechanism so that each type of her gets all expected social surplus conditional on her type can generically be supported as a sequential equilibrium outcome and a core solution. Moreover, only such allocation profiles constitute neutral optima, i.e. 'the smallest possible set of un-blocked mechanisms' (Myerson 1983). Any other allocation profile (i.e. an ex-post inefficient one or one in which some type of the informed principal gets less then expected social surplus conditional on her type) is not a neutral optimum: it can be blocked by some types of the mechanism designer.

Thus, the informed principal problem need not cause any loss of efficiency. Also, an informed principal, like an uninformed one, can extract all surplus from the mechanism. These conclusions are sufficiently robust, since they hold in a sequential equilibrium, core solution and a unique neutral optimum outcome.

We briefly address the issue of implementation via ex-post budget-balanced mechanisms without individual rationality requirement. We show that a weaker version of the Identifiability condition - *Weak Identifiability*- is necessary and sufficient for any implementable decision rule to be implementable with ex-post budget balance. This condition also has an intuitive economic interpretation relying on the concept of probability distributions of the reported type profiles induced by the players' strategies. A different necessary and sufficient condition for such implementation - Condition C- has been previously derived by d'Aspremont, Crémer and Gérard-Varet (2004). Naturally, our Weak Identifiability is equivalent to Condition C.

Lastly, the modelling approach in this paper and in the related literature is based on the assumption that there is a one-to-one relationship between an agent's payoff-relevant type and her beliefs about other agents' types. Although there may be situations where this is not so, this approach is plausible in many economically significant environments, such as competition for mineral rights where a firm's private signal about the amount of mineral resources in the ground determines both its expected profits and also its beliefs about the competitors.

3

Similarly, when there is uncertainty about market conditions, a firm's information about the future demand for its product is both her payoff-relevant type, as well as the determinant of its beliefs about the competitors' demands.

The rest of the paper is organized as follows. In Section 2 we develop the model. Subsection 3.1 introduces the concept of Identifiability. In Subsection 3.2 we establish out cental implementation results. Subsection 3.3 studies implementation without individual rationality requirement. Section 4 deals with the informed principal problem. All proofs are in the Appendix.

2   The Model

There are n agents in the economy. Agent i has privately known type θ_i which belongs to the type space $\Theta_i \equiv \{\theta_i^1, ..., \theta_i^{m_i}\}$ of cardinality $m_i < \infty$. A state of the world is characterized by a profile of types $\theta = (\theta_1, ..., \theta_n)$. The set of type profiles is given by $\Theta \equiv \prod_{i=1,..n} \Theta_i$ which has cardinality $L \equiv \prod_{i=1,...,n} m_i$. When focussing on agent i, we will use the notation (θ_{-i}, θ_i) for the profile of agent types, where θ_{-i} stands for the profile of types of agents other than i. Let $\Theta_{-i} = \prod_{j \neq i} \Theta_j$ and $L_{-i} = \prod_{j \neq i} m_j$. θ_{-i-j}, Θ_{-i-j} and L_{-i-j} are defined similarly. In Section 3 we assume the presence of a mechanism designer who does not possess any private information. In section 4 we analyze the informed principal problem where such mechanism designer is not available and the mechanism has to be designed by one of the informed agents.

Let X denote the set of public decisions controlled by the mechanism designer, and let x denote a typical element of X. Agent i's utility function is quasilinear in the decision x and transfer t_i that she receives from the mechanism and is given by $u_i(x, \theta) + t_i$. Without loss of generality, an agent's reservation utility is normalized to zero.[5] A (social) decision rule $x(.)$ is a function mapping the type space Θ into the set of public decisions X.[6] $t_i(\theta)$ is a transfer function to agent i, and $t(\theta) = (t_1(\theta), ..., t_n(\theta))$ is a collection of transfer functions to all agents. An allocation profile is a combination of a decision rule $x(\theta)$ with a collection of transfer functions $t(\theta)$.

Finally, let $p(\theta)$ denote the probability distribution of the type profiles. $p(.)$ is common knowledge. Also, $p_i(\theta_i)$, $p_{i,j}(\theta_i, \theta_j)$ and $p(\theta_{-i}|\theta_i)$ denote the marginal probability distribution of agent i's type, the marginal probability distribution of types of agents i and j and the probability distribution of type profiles of agents other than i conditional on the type of agent i, respectively. We use a similar system of notation for any other probability distribution function introduced in the text. We will assume that $p_{i,j}(\theta_i, \theta_j) > 0$ for any $\theta_i \in \Theta_i, \theta_j \in \Theta_j$ of any two agents i and j. This condition is not needed to prove our main result -Theorem 1- if we restrict attention to ex-post efficient decision rules. Furthermore, it is clearly generic. A straightforward but lengthy argument which we omit to conserve space shows the following. If this condition fails, then all our results hold, but the analysis becomes more cumbersome because we need to operate on the common knowledge partition dividing the type space into disjoint subsets which each agent knows have occurred on the basis of her type.[7]

---

[5] Suppose that agent i's utility from her outside option is equal to $w_i(\theta_i, \theta_{-i})$. Such environment is equivalent to the environment where i's utility function is given by $u_i(x, \theta) - w_i(\theta) + t_i$ and her outside option is 0. Note that the sets of ex-post efficient decision rules and the notions of social surplus are the same in both environments.

[6] Note that randomization in public decisions is implicitly allowed, since X can be regarded as a set of probability distributions over some set of "pure" outcomes.

[7] A description of this argument is provided in Appendix B available at http://faculty.fuqua.duke.edu/%7Esseverin/mechsupp.pdf.

In the case of an uninformed mechanism designer, we rely on the Revelation Principle to restrict the analysis to the class of direct mechanisms in which the principal offers an allocation profile to the agents. The agents then decide whether to participate in the mechanism. If they decide to participate, the agents report their types and the allocation corresponding to the reported type profile is implemented.

Our main goal is to provide necessary and sufficient conditions for the existence of interim individually rational and ex-post budget-balanced Bayesian mechanisms implementing desirable decision rules. Let us describe these properties formally.

We will say that the allocation profile $(x(\theta), t(\theta))$ is incentive compatible (IC) (and so, can be implemented via a Bayesian mechanism) if for all i and $k, k' \in \{1, ..., m_i\}$ the following *Interim Incentive Constraint* $IC_i(\theta_i^k, \theta_i^{k'})$ (or $IC_i(k, k')$) holds:

$$\sum_{\theta_{-i} \in \Theta_{-i}} \left(u_i(x(\theta_{-i}, \theta_i^k), (\theta_{-i}, \theta_i^k)) + t_i(\theta_{-i}, \theta_i^k) - u_i(x(\theta_{-i}, \theta_i^{k'}), (\theta_{-i}, \theta_i^k)) - t_i(\theta_{-i}, \theta_i^{k'}) \right) p(\theta_{-i}|\theta_i^k) \geq 0 \tag{1}$$

A decision rule $x(.)$ is *implementable* if there exists a profile of transfer functions $t(.)$ such that $(x(.), t(.))$ is incentive compatible.

Interim Individual Rationality (IR) requires the following $IR_i(\theta_i^k)$ constraint to hold for all i and $\theta_i^k \in \Theta_i$:

$$\sum_{\theta_{-i} \in \Theta_{-i}} \left(u_i(x(\theta_{-i}, \theta_i^k), (\theta_{-i}, \theta_i^k)) + t_i(\theta_{-i}, \theta_i^k) \right) p(\theta_{-i}|\theta_i^k) \geq 0 \tag{2}$$

Ex-post Budget Balancing (BB) constraint can be written as follows:

$$\sum_{i=1}^{n} t_i(\theta) = 0 \quad \forall \theta \in \Theta \tag{3}$$

A decision rule $x(.)$ is *ex-post efficient* if $x(\theta) \in \arg\max_{x \in X} \sum_{i=1}^{n} u_i(x, \theta)$ for all $\theta \in \Theta$, i.e. $x(\theta)$ maximizes ex-post *social surplus*. Since the principal always has an option to disband the mechanism and cause the agents to take their outside options, we assume without loss of generality that $\max_{x \in X} \sum_{i=1}^{n} u_i(x, \theta) \geq 0$ for all $\theta \in \Theta$. Finally, IR and BB together imply the following *Ex-Ante Social Rationality (EASR)* condition:

$$\sum_{\theta \in \Theta} \sum_{i=1}^{n} u_i(x(\theta), \theta) p(\theta) \geq 0 \tag{4}$$

EASR simply says that a decision rule must generate a nonnegative (ex ante) expected surplus. Clearly, this is a very weak requirement. It is satisfied by a large variety of decision rules including the ex-post efficient ones. Having established *EASR* as a necessary condition, in the next section we characterize necessary and sufficient conditions for IR and BB implementation of *EASR* and ex-post efficient decision rules.

3 Analysis.

3.1 Identifiability.

We start by introducing a condition which plays a major role in our analysis:

Definition 1 _Identifiability._ *The probability distribution $p(.) \in \Delta^{(\Pi_l m_l)-1}$ of type profiles is identifiable if for any $\hat{\pi}(.) \in \Delta^{(\Pi_l m_l)-1}$, $\hat{\pi}(.) \neq p(.)$ there is an agent i and her type θ_i^k such that $\hat{\pi}_i(\theta_i^k) > 0$ and for any collection of nonnegative coefficients $c_i^{k'k}$, $k' \in \{1, ... m_i\}$,*

$$\hat{\pi}(.|\theta_i^k) \neq \sum_{k' \in \{1, ... m_i\}} c_i^{k'k} p(.|\theta_i^{k'}) \tag{5}$$

Consider also the familiar condition of Crémer and McLean (1985) and (1988) which is necessary and sufficient for full surplus extraction by the mechanism designer:

Definition 2 *Say that Crémer-McLean condition holds for agent i if for any type $\theta_i^k \in \Theta_i$, $p(.|\theta_i^k)$ cannot be expressed as a positive linear combination of $p(.|\theta_i^{k'})$, $k' \neq k$, i.e. for any collection of nonnegative coefficients $c_i^{k'k}$, $k' \in \{1, ... m_i\}$, $k' \neq k$,*

$$p(.|\theta_i^k) \neq \sum_{k' \in \{1, ..., m_i\}, k' \neq k} c_i^{k'k} p(.|\theta_i^{k'})$$

In the next subsection we show that Identifiability together with Crémer-McLean condition are necessary and sufficient for BB, IR, efficient implementation. Although Crémer-McLean condition is well-understood[8], Identifiability is a new condition introduced in this paper. So, before we exhibit and explain our results, let us explore it in greater detail.

First, let us examine the relationship between the Identifiability condition and the notion of strategies chosen by the agents in a direct mechanism. For this, we need some additional notation. Agent i's strategy s_i in a direct mechanism is a vector of size m_i^2 such that its $(k-1)m_i + k'$-th entry $s_i^{kk'}$ denotes the probability with which agent i of type θ_i^k reports type $\theta_i^{k'}$ when following s_i. Note that $s_i^{kk'} \in [0, 1]$ and $\sum_{k'=1}^{m_i} s_i^{kk'} = 1$ for all $k, k' \in \{1, ..., m_i\}$. Let S_i be the set of all such strategies s_i. Formally, $S_i = \left(\Delta^{m_i-1}\right)^{m_i}$ where Δ^{m_i-1} is an $m_i - 1$ dimensional unit simplex in $\mathbf{R}_+^{m_i}$. A truthful strategy s_i^* of agent i is such that $s_i^{kk} = 1$, $s_i^{kk'} = 0$ for all $k, k' \in \{1, ..., m_i\}$, $k' \neq k$.

A strategy profile $\mathbf{s} \equiv (s_1, ..., s_n)$ is a collection of strategies followed by the agents. A strategy profile such that agent i follows strategy s_i and all other agents follow truthful strategies is denoted by (s_i, s_{-i}^*).

Definition 3 *Say that the strategy profile $\mathbf{s} \equiv (s_1, ..., s_n)$ induces the probability distribution of the reported type profiles $\pi(.|\mathbf{s})$ if the type profile θ is reported with probability $\pi(\theta|\mathbf{s})$ when the agents follow strategies $\mathbf{s} = (s_1, ..., s_n)$ and the types are drawn from the prior $p(.)$.*

Thus, $\pi(.|\mathbf{s})$ is a probability distribution vector of size $\prod_l m_l$ with each entry corresponding to some type profile in the natural order induced by the ordering of agents and their types. To compute $\pi(.|\mathbf{s})$, note that for any $\hat{\theta} \equiv (\theta_1^{h_1}, ..., \theta_n^{h_n})$, $\pi(\hat{\theta}|\mathbf{s}) = \sum_{(\theta_1^{k_1}, ..., \theta_n^{k_n}) \in \Theta} \left(p(\theta_1^{k_1}, ..., \theta_n^{k_n}) \prod_{i=1}^n s_i^{k_i h_i}\right)$.

Further, a strategy component vector $\pi(.|k, k', \mathbf{s}_{-i}^*)$ of agent i is a vector of size $\prod_l m_l$ with entries $p(\theta_{-i}|\theta_i^k)$ in positions corresponding to type profiles $(\theta_{-i}, \theta_i^{k'})$ for all $\theta_{-i} \in \Theta_{-i}$ and zeroes in all other positions. It corresponds to the probability distribution of reported type profiles conditional on agent i having type θ_i^k and reporting her type as $\theta_i^{k'}$, and all other agents types' being drawn from $p(.|\theta_i^k)$ and reported truthfully. To differentiate between truthful and

---

[8] See also McAfee and Reny (1992) for an intuitive discussion of this condition.

non-truthful reporting by agent i, we will say that a strategy component $\pi(.|k, k', \mathbf{s}^*_{-i})$ of agent i is a deviation component if $k' \neq k$ and is a truthful component if $k' = k$. Note that for any $s_i \in S_i$, $\pi(.|s_i, \mathbf{s}^*_{-i}) = \sum_{k,k' \in \{1,...,m_i\}} s_i^{kk'} \pi(.|k, k', \mathbf{s}^*_{-i}) p_i(\theta_i^k)$.

It is natural to interpret the coefficients $c_i^{k'k}$ in the definition of identifiability as stemming from i's reporting strategy i.e. $c_i^{k'k} = \frac{s_i^{k'k} p_i(\theta_i^{k'})}{\sum_{k''} s_i^{k''k} p_i(\theta_i^{k''})}$. Then the identifiability requires that for each $\hat{\pi}(.) \neq p(.)$ there exists an agent i who does not have a strategy such that when i reports some θ_i^k according to this strategy and the other agents report truthfully, the induced probability distribution of the other agents' type profiles coincides with $\hat{\pi}(.|\theta_i^k)$. That is, $\frac{\sum_{k' \in \{1,...,m_i\}} s_i^{k'k} p(\theta_i^{k'}) p(.|\theta_i^{k'})}{\sum_{k' \in \{1,...,m_i\}} s_i^{k'k} p(\theta_i^{k'})} \neq \hat{\pi}(.|\theta_i^k)$ for all s_i.

Thus, under identifiability, when the agents' strategies induce probability distribution $\hat{\pi}(.) \neq p(.)$ of the reported types, the principal can identify a non-empty set of *agent types* such that the agents reporting these types have not unilaterally deviated from truthtelling. This interpretation provides a rationale for the use of the term 'identifiability' to name this condition.

In comparison, Crémer-McLean condition requires that no agent possess a non-truthful strategy which, in combination with truthtelling by other agents, induces a probability distribution of the reported type profiles such that its conditional on *any* type θ_i of agent i coincides with the conditional probability distribution $p(.|\theta_i)$ derived from the prior.

Next, let us introduce the following weakening of the Identifiability condition:

Definition 4 *Weak Identifiability.* *The prior $p(.)$ is weakly identifiable if for any agent i and any $s_i \in S_i$ s.t. $\pi(.|s_i, \mathbf{s}^*_{-i}) \neq p(.)$ there exists an agent j s.t. for all $s_j \in S_j$*

$$\pi(.|s_i, \mathbf{s}^*_{-i}) \neq \pi(.|s_j, \mathbf{s}^*_{-j}) \tag{6}$$

In Subsection 3.3 we show that Weak Identifiability is necessary and sufficient for budget-balanced implementation when individual rationality is not required. To understand why Identifiability is stronger than Weak Identifiability, note that Weak Identifiability requires that for every $\hat{\pi}(.) \neq p(.)$ there exists some agent i who cannot induce $\hat{\pi}(.)$ by a unilateral deviation from truthtelling, i.e. $\hat{\pi}(.) \neq \pi(.|s_i, \mathbf{s}^*_{-i})$ for any $s_i \in S_i$. But even if $p(.)$ is weakly identifiable, it may fail to be identifiable if for every agent i there exists a collection of strategies $\{s_i(r)\}_{r=1,...,m_i}$ s.t. the conditional of $\pi(.|s_i(r), \mathbf{s}^*_{-i})$ on θ_i^r is equal to $\hat{\pi}(.|\theta_i^r)$ for some $\hat{\pi}(.)$ (although for some i there is no strategy $s_i \in S_i$ such that $\pi(.|s_i, \mathbf{s}^*_{-i}) = \hat{\pi}(.)$).

The following Lemma further highlights the distinction between Identifiability and Weak Identifiability. More importantly, the Lemma will be useful for explaining the necessity part of Theorem 1 in the next subsection.

Lemma 1 *The probability distribution $p(.)$ is identifiable if and only if for any agent i, any $s_i \in S_i$ s.t. $\pi(.|s_i, \mathbf{s}^*_{-i}) \neq p(.)$ and any collection of coefficients $b_i^k \geq 0$, $k \in \{1,...,m_i\}$, there exists an agent j s.t. for any $s_j \in S_j$ and any collection of coefficients $b_j^h \geq 0$, $h \in \{1,...,m_j\}$,*

$$\pi(.|s_i, \mathbf{s}^*_{-i}) + \sum_{k \in \{1,...,m_i\}} b_i^k \pi(.|k, k, \mathbf{s}^*_{-i}) \neq \pi(.|s_j, \mathbf{s}^*_{-j}) + \sum_{h \in \{1,...,m_j\}} b_j^h \pi(.|h, h, \mathbf{s}^*_{-j}) \tag{7}$$

We conclude this subsection with the following lemma which connects the Identifiability condition with the familiar notion of linear independence. It implies that the Identifiability condition is easy to satisfy.

Lemma 2 *The probability distribution $p(.)$ is identifiable if for some agent i and each $k \in \{1, ..., m_i\}$ there exists an agent j_k and $h_k \in \{1, ..., m_{j_k}\}$ such that $m_i + m_{j_k} - 1$ conditional probability vectors $p(.|\theta_i^{k'}, \theta_{j_k}^{h_k})$, $p(.|\theta_i^{k}, \theta_{j_k}^{h'})$, $k' \in \{1, ..., m_i\}$, $h' \in \{1, ..., m_{j_k}\}$, $h' \neq h_k$, are linearly independent.*

Proof: See the appendix.

The condition of Lemma 2 is significantly stronger than Identifiability, as one can infer from the respective dimensionality requirements. Indeed, Lemma 2 implies that $p(.)$ is identifiable if for some i and j the collection of $m_i + m_j - 1$ vectors $p(.|\theta_i^k, \theta_j^h)$, $k \in \{1, ..., m_i\}$, $h \in \{1, ..., m_j\}$ is linearly independent. For this, it is necessary that $\prod_{l \neq i,j} m_l \geq m_i + m_j - 1$. In contrast, Lemma 3 below demonstrates that agent i is generically identifiable when $\prod_{l \neq i} m_l \geq m_i$.

3.2 Main Result.

The main result of this paper is stated in the following theorem:

Theorem 1 ***Sufficiency.*** *Any ex-ante socially rational decision rule is implementable via an interim individually rational and ex-post budget balanced Bayesian mechanism if the prior $p(.)$ is identifiable and Crémer-McLean condition holds for all agents.*

   Necessity. *An ex-post efficient decision rule is implementable via an interim individually rational, ex-post budget-balanced Bayesian mechanism under any profile of the utility functions (quasilinear in transfers) only if the prior $p(.)$ is identifiable and Crémer-McLean condition holds for all agents.*

Remark. The necessity part of the Theorem is established for ex-post efficient, rather than ex-ante socially rational decision rules because failure of ex-post efficiency is arguably a more significant problem. A careful reading of the Step 6 in the proof should convince a reader that decision rules that are not ex-post efficient but are ex-ante socially rational may also fail to be implementable if the identifiability condition does not hold.

To explain the intuition behind Theorem 1, it is useful to start the discussion from a natural benchmark. So, let us highlight the relationship between our results and the work of Crémer and McLean (1985) and (1988) on the surplus extraction by the mechanism designer. Below, we will also discuss the link to the literature on implementation with ex-post budget balancing, but without interim individual rationality.

The mechanism of Crémer and McLean (1988) elicits the agents' private information without leaving any informational rents to them by means of lotteries - systems of transfers to/from each agent which depend on the whole profile of the reported types. The lotteries are constructed in such a way that the loss from a misrepresentation in a lottery for any agent always exceeds any potential gain from a better allocation $x(.)$. Crémer and McLean condition described in Definition 2 is necessary and sufficient for such lotteries to exist.

This method works successfully in the presence of an outside principal who acts as a budget-breaker, or a residual claimant, for the lotteries offered to the agents. In contrast, in our budget-balanced framework all payments to/by any particular agent have to be made by/distributed to other agents. Consequently, we must designate residual claimant(s) for the

lottery given to each agent, which raises the issue of the incentives of the residual claimants. For example, designating agent j as a residual claimant for the lottery given to agent i may generate incentives for j to 'rig the lottery:' misrepresent her type in a way that makes a truthful report by i to appear untruthful and, thus, to cause i to make transfers to j.

It is natural to consider the following candidate for an optimal mechanism which allows to resolve the issue of the incentives of residual claimants. Suppose that we divide all agents into groups and run a separate Crémer-McLean mechanism in each group using an agent outside the group as a residual claimant for the balance of transfers from/to this group. The incentive compatibility of this mechanism can be ensured by making the lotteries in each group independent of the residual claimant's reported type. Clearly, such mechanism will also be budget balanced. However, it could fail to be individually rational for the residual claimants, because the aggregate transfer between each group and its residual claimant will be determined independently of the residual claimant's reported type. But, given statistical interdependence between the agents' types, the probability distribution of these aggregate transfers and hence the expected utility of the residual claimant will be *dependent on her type.* In particular, for some values of her private information the residual claimant could expect a negative payoff, in which case she would refuse to participate in this mechanism. Therefore, such mechanism would not work in general. Its failure highlights the role of individual rationality, and illustrates the difference between our environment and the ones with ex-post budget-balancing but without individual rationality requirement as in Aoyagi (1998) and d'Aspremont, Crémer and Gérard-Varet (2004). Indeed, a mechanism of this kind is used by Aoyagi (1998). In contrast, in our mechanism every agent's allocation typically has to depend on the reported types of all agents, which could generate an opportunity for residual claimants to increase their payoffs by mimicking other agents' deviations.

To summarize, we need to deal with three issues within a balanced budget framework: (i) an agent misrepresenting her type in a way which is indistinguishable from truthtelling; (ii) "Mimicry:" an agent misrepresenting her type in a way which is statistically indistinguishable from some misrepresentation by another agent; (iii) the individual rationality. The tension between the second and the third issues was described above.

Consider how these issues can be resolved when Identifiability and Crémer-McLean conditions hold. Under Identifiability, if the probability distribution of reported type profiles is $\hat{\pi}(.) \neq p(.)$, then the mechanism designer can identify *a non-empty set of agent types* such that the agents reporting these types are surely truthful.

Moreover, since the Identifiability condition requires this to be true for all probability distributions of reported type profiles, there are sufficiently many such agent types to cover the whole type space under any $\hat{\pi}(.)$. Precisely, suppose that $\hat{\pi}(\hat{\theta}) \neq p(\hat{\theta})$ for some $\hat{\theta} \equiv (\hat{\theta}_1, ..., \hat{\theta}_n)$. Then there exists an agent i s.t. $\hat{\pi}(.|\hat{\theta}_i) \neq \pi[(.|s_i, s^*_{-i})|\hat{\theta}_i]$ for all $s_i \in S_i$ (where $\pi[(.|s_i, s^*_{-i})|\hat{\theta}_i]$ denotes the conditional of $\pi(.|s_i, s^*_{-i})$ on $\hat{\theta}_i$).

In turn, Crémer-McLean condition implies that no agent can perfectly 'hide' her deviation from truth-telling, i.e. no agent i has a non-truthful reporting strategy s_i s.t. $s_i^{k'k} > 0$ for some $k \neq k'$ but which in combination with truthtelling by other agents generates probability distribution $\pi[(.|s_i, s^*_{-i})$ satisfying $\pi[(.|s_i, s^*_{-i})|\theta_i^k] = p(.|\theta_i^k)$.

Therefore, we can construct a budget-balanced system of transfers $t(\theta) = (t_1(\theta), ..., t_n(\theta))$ s.t. (i) $\sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i) p(\theta_{-i}|\theta_i) = 0$ for all i, $\theta_i \in \Theta_i$, (ii) $\sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i) \hat{\pi}(\theta_{-i}|\theta_i) > 0$ if $\hat{\pi}(.|\theta_i) \neq p(.|\theta_i)$ and $\hat{\pi}(.|\theta_i) \neq \pi[(.|s_i, s^*_{-i})|\theta_i]$ for all $s_i \in s_i^*$, (iii) $\sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i) \hat{\pi}(\theta_{-i}|\theta_i) < 0$ if $\hat{\pi}(.|\theta_i) \neq p(.|\theta_i)$ and $\hat{\pi}(.|\theta_i) = \pi[(.|s_i, s^*_{-i})|\theta_i]$ for some s_i.

Thus, the system of transfers $t(.)$ always punishes an agent who *could have* misrepresented her type unilaterally by giving her a negative expected transfer, and gives a strictly positive payoff to an agent who *could not have* done so[9].

This implies that $t(.)$ is both individually rational and strictly incentive compatible, i.e.

$$\sum_{\theta_{-i}\in\Theta_{-i}} t_i(\theta_{-i},\theta_i)p(\theta_{-i}|\theta_i) = 0 \text{ and } \sum_{\theta_{-i}\in\Theta_{-i}} t_i(\theta_{-i},\theta_i')p(\theta_{-i}|\theta_i) < 0 \text{ for all } \theta_i \neq \theta_i' \quad (8)$$

To reiterate, is it possible to ensure that $t(.)$ is budget balanced because for any probability distribution of reported type profiles there is a sufficiently large set of agent types who could not have unilaterally deviated, and who are therefore given a positive expected transfer.

By scaling the transfers up appropriately, the mechanism designer can elicit type information from the agents at no cost, irrespective of the decision rule $x(.)$. To these transfers, she can also add budget-balanced transfers that compensate agents for the utility consequences of the public decision $x(.)$ and allocate the social surplus without affecting the incentive compatibility of the mechanism. This explains the sufficiency of the Identifiability and Crémer-McLean conditions.

The necessity part of Theorem 1 can be explained using a similar logic. If the Identifiability condition fails, then by Lemma 1 there is a collection of strategies $(s_1, ..., s_n)$ with $s_i \neq s_i^*$ for some i and nonnegative coefficients $\{b_1^{k_1}, ..., b_n^{k_n}\}$ s.t.

$$\sum_{k_1\in\{1,...,m_1\}} b_1^{k_1}\pi(.|k_1,k_1,s_{-1}^*) + \pi(.|s_1,s_{-1}^*) = ... = \sum_{k_n\in\{1,...,m_n\}} b_n^{k_n}\pi(.|k_n,k_n,s_{-n}^*) + \pi(.|s_n,s_{-n}^*) = \hat{\pi}$$

$$(9)$$

Let us show that in this case there is no budget-balanced system of transfers $(t_1(\theta), ..., t_n(\theta))$ which is strictly incentive compatible and individually rational, i.e. satisfies (8). Indeed, since $s_i \neq s_i^*$, (8) implies that $t_i(.) \times \pi(.|s_i, s_{-i}^*) < 0$. Also, (8) implies that $t_l(.) \times \pi(.|s_l, s_{-l}^*) \leq 0$ for all $l \neq i$ and $s_l \in S_l$, and $t_l(.) \times \pi(.|k_l, k_l, s_{-l}^*) = 0$ for all l. So, if both (8) and (9) hold, then $t_i(.)\hat{\pi} < 0$ and $t_l(.)\hat{\pi} \leq 0$ for all $l \neq i$. Summing up, we get $\left(\sum_l t_l(.)\right)\hat{\pi} < 0$ which contradicts budget-balancing.

Similarly, if Crémer-McLean condition fails for agent i, then there is a strategy $s_i \neq s_i^*$ and $b_i^{k_i} \geq 0$, $k_i = 1, ..., m_i$ such that $\sum_{k_i\in\{1,...,m_i\}} b_i^{k_i}\pi(.|k_i,k_i,s_{-i}^*) + \pi(.|s_i,s_{-i}^*) = \alpha p(.)$ for some $\alpha > 0$. But this contradicts (8) which implies that $t_i(.) \times \left(\sum_{k_i\in\{1,...,m_i\}} b_i^{k_i}\pi(.|k_i,k_i,s_{-i}^*) + \pi(.|s_i,s_{-i}^*)\right) < 0$ and, on the other hand, $t_i(.)p(.) = 0$.

The non-existence of a system of transfers satisfying (8) means that the mechanism has to provide informational rents to prevent agents from engaging in some deviations. We then show that the sum of expected informational rents that need to be given to the agents exceeds the expected social surplus for a class of utility function profiles, even though the decision rule is ex-post efficient. Then efficient implementation is impossible. Note that the same logic is behind the impossibility result of Myerson and Satterthwaite (1983). The difference is that in our case the mechanism may have to pay informational rents only to some subset of agent types (this could even be a single type). Yet, as our proof demonstrates, this could be enough to make implementation impossible.

---

[9]Informally, we can say that Crémer-McLean condition makes it possible to detect the deviators and punish them, while the Identifiability condition makes it possible to identify those who could not have deviated and, therefore, could be given the proceeds from punishments.

Note that Crémer and McLean (1988) show that their condition is necessary for surplus extraction in their environment, but not for implementation per se. In contrast, we establish that this condition is, in fact, necessary for *implementation*.

With ex-post budget balance, all social surplus generated by the mechanism is allocated to the agents. Thus, it is natural to investigate how this surplus can be distributed across the agents and agent types. The next result shows that this can be done in an arbitrary way.

Theorem 2 *Consider any EASR decision rule $x(\theta)$, and suppose that the prior $p(.)$ is identifiable and Crémer-McLean condition holds for all agents. Then for any collection of $\sum_i m_i$ nonnegative constants $v_i(\theta_i)$ satisfying:*

$$\sum_{i\in\{1,\dots,n\},\theta_i\in\Theta_i} v_i(\theta_i)p_i(\theta_i) = \sum_{i\in\{1,\dots,n\},\theta\in\Theta} u_i(x(\theta),\theta)p(\theta) \tag{10}$$

there exists an IC, BB, and IR Bayesian mechanism $(x(\theta),t(\theta))$ s.t. the expected surplus of type θ_i of agent i is equal to $v_i(\theta_i)$, i.e.

$$\sum_{\theta_{-i}\in\Theta_{-i}} (u_i(x(\theta_{-i},\theta_i),(\theta_{-i},\theta_i)) + t_i(\theta_{-i},\theta_i))p(\theta_{-i}|\theta_i) = v_i(\theta_i)$$

Proof: *See the Appendix.*

An interesting aspect of Theorem 2 is that, compared to Theorem 1, no additional conditions are required to achieve the desired allocation of surplus. To understand why this is so, refer to the method of lotteries described above. These lotteries guarantee strict incentive compatibility of truthtelling and binding individual rationality for all types. Given that the expected losses from non-truthful reporting can be made arbitrarily large, the mechanism will remain incentive compatible even if we add to it transfers distributing the expected social surplus in any desirable way.

This result is of independent interest because the existence of an efficient, individually rational, balanced-budget mechanism does not by itself guarantee that the social surplus can be allocated arbitrarily. For example, one can show that such a mechanism exists if there is an agent $\hat{i}$ whose type is distributed independently of all other agents' types, the prior over other agents' types $p_{-\hat{i}}$ is identifiable and Crémer-McLean condition holds for all agents other than $\hat{i}$. However, agent $\hat{i}$ must receive at least some, and sometimes all social surplus. So, identifiability is an essential requirement guaranteeing the freedom in surplus allocation.

Next, we study whether a sufficiently large set of probability distributions is identifiable and satisfies Crémer-McLean conditions for all agents.

Using Lemma 2, it is relatively straightforward to show that identifiability is generic if $L_{-i-j} \geq m_i + m_j - 1$ for some i and j. However, the following Lemma demonstrates that this dimensionality requirement can be significantly relaxed.

Lemma 3 *Genericity of Identifiability.* *Suppose that there are at least three agents ($n \geq 3$) and $\Pi_{j\neq i}m_j \geq m_i$ for all i. Also, if $n = 3$ then at least one of the agents has at least three types. Then almost all probability distributions $p(.)$ are identifiable.*

In the proof of Lemma 3 we use a measure-theoretic definition of genericity and show that the set of probability distributions which are not identifiable has (Lebesgue) measure zero.

Precisely, the proof establishes that $p(.)$ is not identifiable only if a set of non-degenerate polynomials in the entries of $p(.)$ are non-zero. This argument also implies that identifiability is generic in the topological sense, i.e. holds on an open dense set in the topology generated by the Euclidian metric.

It is easy to see that Crémer-McLean condition for agent i holds generically when $m_i \leq \prod_{l \neq i} m_l$. This also follows from *Fact 2* in Step 3 of the proof of Lemma 3. Since a finite union of sets of measure zero has measure zero, while an intersection of a finite number of open and dense sets is open and dense, we conclude that both conditions of Theorems 1 and 2 are generic when the dimensionality requirements of Lemma 3 hold.

The proof of Lemma 3 is quite involved because there is no simple method to construct an efficient, IR, BB mechanism. In contrast, if individual rationality requirement is left out, this can be done by relying either on Cauchy-Schwartz inequality or the scoring rule (see Aoyagi (1998) or d'Aspremont, Crémer and Gérard-Varet (2004)). But if we take the mechanism constructed in either of these papers and attempt to adjust it to make the individual rationality hold, then either incentive compatibility or budget-balancing would be violated.

When $m_i > \Pi_{j \neq i} m_j$ for some i, then Crémer McLean condition is not generic for i, which also implies that i is not generically identifiable. It is possible to show directly that the identifiability condition fails when $n = 2$ and $m_1 = m_2$. We conjecture that the same result holds for $m_1 \neq m_2$. Finally, for $n = 3$ and $m_1 = m_2 = m_3 = 2$ we can show by direct computation that $p(.)$ is identifiable if at least one of the following inequalities is satisfied (where $p_{k_1 k_2 k_3}$ denotes the probability that the agents 1,2 and 3 have types k_1, k_2, and k_3 respectively): $(p_{211}p_{222} - p_{212}p_{221})(p_{121}p_{222} - p_{122}p_{221}) < 0$;
$(p_{121}p_{112} - p_{122}p_{111})(p_{121}p_{222} - p_{122}p_{221}) < 0$; $(p_{212}p_{111} - p_{112}p_{211})(p_{121}p_{222} - p_{122}p_{221}) < 0$.
In this case, and provided that the elements of $p(.)$ are drawn uniformly, 88% of the probability distributions are identifiable.

We complete this section by illustrating our results with two examples.

Example 1. Bargaining and Trade. There are 2 buyers and 2 sellers each of whom can produce one unit of the good. A buyer's valuation is equal to $\underline{v}$ with probability p and $\overline{v}$ with probability $(1 - p)$. A seller's cost is equal to $\underline{c}$ with probability q and $\overline{c}$ with probability $(1 - q)$. Further, $\overline{v} > \overline{c} > \underline{v} > \underline{c}$. Let $\Delta_v = \overline{v} - \underline{v}$, and $\Delta_c = \overline{c} - \underline{c}$.

If the costs and valuations are distributed independently, then following the method of Myerson and Satterthwaite (1983), we can show that an efficient, IR, BB Bayesian mechanism does not exist if the sum of (ex-ante) expected informational rents of all traders $E\pi$ in an efficient, IR mechanism exceeds total (ex-ante) expected surplus ES. Using standard techniques, we find that $E\pi = 2(1 - p)q(1 + p(1 - q))(\Delta_v + \Delta_c)$ and
$ES = 2(\overline{v} - \overline{c})(1 - q)^2(1 - p) + 2(\underline{v} - \underline{c})p^2 q + 2(\overline{v} - \underline{c})(1 - p)q(1 - q + p)$. Fixing $p, q \in (0, 1)$, it is easy to see that $E\pi > ES$ if $\overline{c} - \underline{v}$ is sufficiently large compared to both $\overline{v} - \overline{c}$ and $\underline{v} - \underline{c}$.

However, by Lemma 3 all traders are generically identifiable. In particular, let us modify the joint probability distribution of their types slightly and suppose that the most efficient type profile $(\overline{v}, \overline{v}, \underline{c}, \underline{c})$ occurs with probability $(1 - p)^2 q^2 + \delta$, the most inefficient type profile $(\underline{v}, \underline{v}, \overline{c}, \overline{c})$ occurs with probability $p^2(1 - q)^2 - \delta$, for $\delta \neq 0$, $\delta \in (-(1 - p)^2 q^2, p^2(1 - q)^2)$, and all other type profiles occur with the same probabilities as in the independence case. Then it is easy to check that the probability distribution over the buyers' values and the sellers costs is identifiable (one can use Condition G in the proof of Lemma 3 to do so) and Crémer-McLean condition holds for all traders, and so an ex-post efficient, IR, BB mechanism exists.

Example 2. Public Good. A number of authors starting from Rob (1989) and Mailath and Postlewaite (1990) have established the impossibility of efficient public good provision via an IR and BB mechanism when the agents' valuations for the public good are privately known and independently distributed. The inefficiency becomes more severe as the number of agents increases. We use the following simple example to illustrate this.

There are 4 consumers. A consumer's valuation of the public good is $\underline{v}$ with probability p and $\overline{v}$ with probability $(1-p)$, where $\overline{v} > \underline{v}$. The cost of the public good is C satisfying $2(\overline{v} + \underline{v}) < C = 3\overline{v} + \underline{v} - \epsilon$ for $\epsilon > 0$. Thus, the public good should be provided if at least three consumers have valuations $\overline{v}$ for it.

Let us consider whether the public good can be provided efficiently when the valuations are distributed independently. By IR, a consumer with valuation $\underline{v}$ cannot pay more than $\underline{v}$ if the public good is provided. If every consumer reports truthfully and the mechanism is efficient, then a consumer with valuation $\overline{v}$ must get an expected payoff of at least $(\overline{v} - \underline{v})(1-p)^3$. Thus, the mechanism must generate at least $SP = 4(1-p)^4(\overline{v} - \underline{v})$ of (ex-ante) expected surplus to cover the informational rents. However, the total amount of expected surplus available in an efficient mechanism is no more than $SA = 3\epsilon p(1-p)^3 + (1-p)^4(\overline{v} - \underline{v} + \epsilon)$.

Clearly, $SP > SA$ when ϵ is small enough. So, an efficient IR, BB mechanism does not exist when valuation are independently distributed. On the other hand, by Lemma 3 almost all probability distributions of the type profiles are identifiable. For example, suppose that the type profiles $(\underline{v}, \underline{v}, \underline{v}, \underline{v})$ and $(\overline{v}, \overline{v}, \overline{v}, \overline{v})$ occur with probabilities $p^4 + \delta$ and $(1-p)^4 + \delta$ respectively, where $|\delta|$ is positive and not too large, while the probability of any other type profile differs from its probability in the independence case by $-\delta/7$. The parameter δ measures the strength of correlation between types. It is easy to check that in this case $p(.)$ is identifiable (use Condition G in the proof of Lemma 3) and Crémer-McLean condition holds for all agents, and so an ex-post efficient, BB, IR mechanism exists.

3.3   Mechanisms without Individual Rationality.

Several authors have explored the issue of implementation via ex-post budget-balanced mechanisms, without imposing individual rationality. This literature includes d'Aspremont and Gérard-Varet (1979), d'Aspremont, Crémer and Gérard-Varet (1990), (2003) and (2004), Matsushima (1991), Fudenberg, Levine and Maskin (1996), Aoyagi (1998), and Chung (1999).

d'Aspremont, Crémer and Gérard-Varet (2004) provide necessary and sufficient conditions for balanced-budget Bayesian implementation and show that these conditions are strictly weaker than the ones studied by the other authors. Their Lemma 1 says that any Bayesian implementable allocation profile can also be implemented with ex-post budget balance if and only if the following **Condition C** is satisfied:

For every function $R(.) : \Theta \mapsto \mathbf{R}$, there exists a transfer rule $t(\theta) \equiv (t_1(\theta), ..., t_n(\theta))$ such that:

$(i) \quad \sum_{i \in \{1,...,n\}} t_i(\theta) = R(\theta) \ for \ all \ \theta \in \Theta$

$(ii) \quad \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i)p(\theta_{-i}|\theta_i) \geq \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i')p(\theta_{-i}|\theta_i) \ for \ all \ i \in \{1,...,n\}, \ \theta_i, \theta_i' \in \Theta_i, \theta_i \neq \theta_i'$

Using the approach developed in the previous sections, we can show that Weak Identifiability is an alternative condition for such implementation. Recall that Weak Identifiability says that any probability distribution of reported type profiles different from the prior, which

can be induced by a unilateral deviation from truthtelling by some agent i, cannot be induced by such deviation of some other agent j (see Definition 4). Then we have:

Lemma 4 *Condition C holds if and only if Weak Identifiability Condition holds.*

Proof: See the Appendix.

d'Aspremont, Crémer and Gérard-Varet (2004) have also shown that any decision rule is implementable via an ex-post budget-balanced Bayesian mechanism if and only if both condition C and the following 'no free beliefs' condition hold: $p(.|\theta_i) \neq p(.|\theta_i')$ for all i, $\theta_i, \theta_i' \in \Theta_i$. Intuitively, the 'no free beliefs' condition guarantees that the incentive constraints in part (ii) of Condition C can be made to hold strictly.

Thus, by Lemma 4, Weak Identifiability together with the 'no free beliefs' condition are also necessary and sufficient for ex-post budget-balanced Bayesian implementation of all decision rules. The intuition for this result is similar to that for Theorem 1, but there are also important differences between the two.

In particular, since the IR constraints do not have to hold, it is sufficient to determine the set of *agents* -not *agent types*- who could not have deviated, reward them and punish the agents who could have deviated. Weak Identifiability condition allows to do so, as it guarantees that the set of agents who are surely not unilateral deviators from truthtelling is non-empty under any probability distribution of the reported type profiles. Formally, for *any* probability distribution $\hat{\pi} \neq p(.)$ of reported type profiles, we can divide all agents into two sets: $D(\hat{\pi}) \equiv \{i | \exists s_i \in S_i \text{ s.t. } \hat{\pi} = \pi(.|s_i, s^*_{-i})\}$ and $ND(\hat{\pi}) \equiv \{i | \hat{\pi} \neq \pi(.|s_i, s^*_{-i}) \ \forall \ s_i \in S_i\}$. The set $D(\hat{\pi})$ includes potential deviators who could have induced $\hat{\pi}$ by a unilateral deviation from truthtelling, while the set $ND(\hat{\pi})$ includes agents who are surely not unilateral deviators from truthtelling when the probability distribution of the reported type profiles is $\hat{\pi}$. Weak Identifiability guarantees that $ND(\hat{\pi})$ is non-empty for all $\hat{\pi} \neq p(.)$.

Therefore, we can construct a system of transfers such that under $\hat{\pi} \neq p(.)$ an agent from $ND(\hat{\pi})$ $(D(\hat{\pi}))$ gets a positive (negative) expected transfer, and rewards and punishments offset each other to balance the budget. Also, under $p(.)$ every agent's (ex-ante) expected transfer is set to be zero. Then some agent types may end up with negative expected transfers even if everyone is truthful. But, nevertheless, each type gets a higher payoff by reporting truthfully than by misrepresenting.

Finally, the 'no free beliefs' condition guarantees that the agents are *strictly* (and not just weakly) better off when reporting their types truthfully.

4   The Informed Principal Problem.

In this section we study an environment where an uninformed mechanism designer is not available, and so the mechanism has to be designed by one of the agents (referred to as the primary agent in the sequel) after all the agents have already learned their private information.[10] In the literature this is known as an 'informed principal problem.' Since different types of the primary agent may decide to offer different mechanisms, the choice of a mechanism provides a signal on the basis of which the other agents update their beliefs about the primary agent's type. Naturally, the outcome of this inference process could affect the agents' incentive. As a

---

[10]The absence of a mechanism designer is natural in many contexts. For example, in collusion context the parties have to agree upon a mechanism for collusion.

result, the informed principal problem is more complex to analyze and solve than a standard mechanism design problem.[11]

In this section we advance the investigation of the informed principal problem by showing that generically it possesses an ex-post efficient solution allocating all social surplus to the primary agent. The explanation for this result is fairly simple. By Theorem 2, generically there exists a class of ex-post-efficient, IR and BB mechanisms M^{c1} in which one agent (say, agent 1) gets all expected social surplus. It is intuitive to conjecture that agent 1 would choose a mechanism from class M^{c1} when she acts as an informed principal. Clearly, agent 1 cannot obtain more surplus, as this would violate the individual rationality, and getting less would be dominated. As far as the distribution of surplus, a natural outcome is for each type of agent 1 to obtain all expected social surplus conditional on her type. In this section we confirm this conjecture formally, and show that the described outcome is fairly robust to the choice of a solution concept.

We start by describing the informed principal game Γ and introducing the necessary notation. Without loss of generality assume that agent 1 is the primary agent who has the authority to propose and implement the mechanism. The timeline of the game Γ is as follows:

- Stage 1. All agents learn their types.

- Stage 2. Agent 1 proposes a mechanism M which can be represented as an outcome function $M : \prod_{i=1,...,n} S_i \mapsto Y$ from the set of agents' strategy profiles into the set of outcomes Y. Y is the set of all budget-balanced allocations $(x, t_1, ..., t_n)$, with $x \in X$.

- Stage 3. Agents 2 to n simultaneously decide whether to participate in the mechanism.

- Stage 4. If all agents have agreed to participate, the mechanism M is implemented.[12] The outcome is determined by the agents' strategy choices and the outcome function of M.

Let Z denote the set of admissible mechanisms. We will require the continuation game following the offer of a mechanism from the set Z to possess a sequential equilibrium for arbitrary agents' beliefs at stage 3. This can be ensured by simply assuming that all mechanisms in Z are finite, i.e. have a finite set of terminal nodes, or a finite set of outcomes assigned to the terminal nodes.

Without loss of generality, we can assume that all mechanisms in Z are direct, so that any mechanism M from Z can be represented as an outcome function $(x(.), t(.))$ mapping the agents' reported type profiles into the allocation profiles. The individual rationality of the mechanism M has to be defined with respect to the agents' beliefs at stage 3 which we denote by $b_i^3(\theta_{-i}|\theta_i, M)$, while the incentive compatibility of the mechanism has to be defined with

---

[11]If a mechanism designer is not available, but the mechanism can be designed at an ex-ante stage when no agent has learned her private information, then it is reasonable to expect that an ex-post efficient decision rule would be implemented and the allocation of surplus would be determined by the ex-ante distribution of the bargaining power.

[12]For simplicity, we assume that the mechanism is not implemented and all agents get their outside options if at least one of them drops out at stage 3. However, the outcomes of our mechanism can be obtained under the same solution concepts if we make alternative assumptions regarding the play of the mechanism following the refusal of some subset of agents to participate in it. To do so, we can simply define the outcome function of a mechanism as a mapping from $\prod_{i=1,...,n} (\emptyset \cup S_i)$ into Y where $\emptyset$ denotes an agent's refusal to participate in the mechanism at stage 3, and the outcome function is restricted to have zero transfers for the agents who have refused to participate in the mechanism.

respect to their posterior beliefs at stage 4 denoted by $b_i^4(\theta_{-i}|\theta_i, M)$. That is, the individual rationality (incentive) constraints are given by the same inequalities as in (2) ((1)), with the only difference that for $i \in \{1, ..., n\}$ we use her beliefs $b_i^3(\theta_{-i}|\theta_i, M)$ at stage 3 ($b_i^4(\theta_{-i}|\theta_i, M)$ at stage 4[13]), instead of her prior beliefs $p_i(\theta_{-i}|\theta_i)$.

The Inscrutability Principle of Myerson (1983) says that without loss of generality on the equilibrium path all types of the primary agent offer the same mechanism, so that the other agents do not change their prior beliefs at stage 3. A mechanism with this property is called inscrutable. The Inscrutability Principle holds because for any mechanism in which the primary agent reveals some information about her type at the mechanism-proposal stage 2, there is an outcome-equivalent inscrutable mechanism in which the primary agent reveals her private information only through her type announcement in stage 4.

The Inscrutability Principle is useful for characterizing solutions to the informed principal problem and describing the corresponding mechanisms. However, we still need to consider all possible deviations from the candidate mechanism, including deviations to non-inscrutable mechanisms. In particular, we have to consider deviations where *some, but not all* types of the primary agent choose to offer a mechanism different from the candidate solution. Such a deviation would cause the other agents to update their priors beliefs in a non-trivial manner at stage 3. Their updated beliefs will then be concentrated on the set of deviating types. The following concept will be useful in the analysis of such situations. Say that a mechanism $(x(\theta), t(\theta))$ is *incentive compatible given a subset R of* Θ_1 if $(x(\theta), t(\theta))$ satisfies standard interim incentive constraints of agent 1 given by (1), while for any agent $i \in \{2, ..., n\}$ the following incentive constraints hold $\forall \theta_i \in \Theta_i$:

$$\sum_{\theta_{-1-i} \in \Theta_{-1-i}, \theta_1 \in R} \big(u_i(x(\theta_{-i}, \theta_i), (\theta_{-i}, \theta_i)) + t_i(\theta_{-i}, \theta_i) - u_i(x(\theta_{-i}, \theta_i'), (\theta_{-i}, \theta_i)) - t_i(\theta_{-i}, \theta_i')\big) p(\theta_{-i}|\theta_i) \geq 0$$

(11)

We will now introduce our solution concepts. A standard solution concept for extensive form games with incomplete information is sequential equilibrium. It is well-known that sequential equilibrium allows too much freedom in the specification of posterior beliefs off the equilibrium path. Therefore, we strengthen our analysis by relying on two additional solution concepts: core mechanism and neutral optimum proposed by Myerson (1983).

The neutral optimum concept is based on the notion of blocking. To define it, let $\{U_i(\theta_i|M)\}_{\theta_i \in \Theta_i}$ be the vector of expected payoffs of agent i in an incentive-compatible mechanism M. Next, let $B(\Gamma)$ denote a set of blocked expected payoff vectors of agent 1 in game Γ. Following Myerson (1983), we require $B(\Gamma)$ to satisfy the following axioms:

Axiom 1 *(Domination) For any vectors $w(.)$ and $z(.)$ in $\mathbb{R}^{m_1}$, if $w(.) \in B(\Gamma)$, and $z(\theta_1) \leq w(\theta_1)$ for every $\theta_1 \in \Theta_1$, then $z(.) \in B(\Gamma)$.*

Axiom 2 *(Openness) $B(\Gamma)$ is open in the set of feasible expected payoff vectors.*[14]

---

[13]Since stage 4 is reached only if at stage 3 all agents have agreed to participate in the mechanism, we can omit the dependence of stage 4 beliefs on stage 3 participation decisions.

[14]Here we make a slight departure from Myerson's definition which requires $B(\Gamma)$ to be open in $\mathbb{R}^{m_1}$. Note that Myerson's proofs of existence of a neutral optimum and characterization results apply verbatim with our notion of openness. It appears quite natural to require $B(\Gamma)$ to be open relative to the set of feasible payoff vectors. Particularly, since otherwise it will never be possible to block any point on the 'upper' boundary of the set of feasible expected payoff vectors of agent 1. This boundary is given by $\{z(.) \in \mathbb{R}_+^{m_1} : \sum_{\theta_1 \in \Theta_1} z(\theta_1)p_1(\theta_1) = \sum_{i, \theta \in \Theta} u_i(x^*(\theta), \theta)p(\theta)\}$, where $x^*(.)$ is an ex-post efficient decision rule.

Axiom 3 *(Extension) Let $\bar{\Gamma}$ be an informed principal problem that differs from Γ only because its feasible action set $\bar{X}$ includes the feasible action set X of Γ, i.e. $X \subset \bar{X}$. Then $B(\Gamma) \subset B(\bar{\Gamma})$.*

Axiom 4 *(Strong Solutions) If mechanism M is incentive compatible given any $\theta_1 \in \Theta_1$ and there does not exist another incentive compatible mechanism M' satisfying $U_1(M|\theta_1) \leq U_1(M'|\theta_1)$ for all $\theta_1 \in \Theta_1$ with strict inequality for at least one θ_1, then $U_1(M|\theta_1) \notin B(\Gamma)$. Such mechanism is called a strong solution.*

These axioms do not define the set of blocked payoff vectors $B(\Gamma)$ uniquely. Rather, there may be several sets of blocked allocations $B(\Gamma)$. To avoid ambiguity, let $\mathcal{I}$ be the index set of all sets of payoff vectors blocked by some concept of blocking satisfying Axioms 1-4, and let $B^*(\Gamma)$ denote the union of all sets of blocked payoff vectors, i.e. $B^*(\Gamma) = \cup_{k \in \mathcal{I}} B_k(\Gamma)$. Thus, a payoff vector is not in $B^*(\Gamma)$ if it cannot be blocked by any concept of blocking satisfying Axioms 1-4.

Definition 5 *(Myerson 1983) An IR, BB mechanism $\tilde{M}$ is a neutral optimum if it is incentive compatible with respect to prior beliefs and the vector $\{U(\tilde{M}|\theta_1)\}$ of the expected payoffs of agent 1 does not belong to $B^*(\Gamma)$.*

The core mechanism is defined as follows:

Definition 6 *(Myerson 1983) An IR, BB mechanism $\tilde{M}$ is a core mechanism if it is incentive compatible (with respect to prior beliefs) and there does not exist any other IR, BB mechanism $\hat{M}$ such that*

$$\left\{ \theta_1 \in \Theta_1 | U_1(\hat{M}|\theta_1) > U_1(\tilde{M}|\theta_1) \right\} \neq \emptyset$$

and $\hat{M}$ is incentive compatible given any set S satisfying

$$\left\{ \theta_1 \in \Theta_1 | U_1(\hat{M}|\theta_1) > U_1(\tilde{M}|\theta_1) \right\} \subset S \subset \Theta_1.$$

The attractiveness of the neutral optimum as a solution concept stems from the fact that a number of other solution concepts, including sequential equilibrium and core mechanisms, give rise to sets of outcomes that can be described via some concept of blocking satisfying the Axioms 1-4. Since neutral optima correspond to the smallest set of mechanisms unblocked according to these axioms, it follows that a neutral optimum also constitutes a solution according to those other solution concepts. In particular, Theorem 5 in (Myerson 1983) shows that *any neutral optimum is also a sequential equilibrium and a core mechanism*.

The main result of this section provides a generic characterization of the set of neutral optima.

Theorem 3 *Suppose that $p(.)$ is identifiable and Crémer-McLean condition holds. Then the set of neutral optima consists of budget-balanced mechanisms $(x(\theta), t(\theta))$ which are incentive compatible and individually rational with respect to prior beliefs and possess the following properties:*

$(i) \quad x(\theta) \in \arg \max_{x \in X} \sum_i u_i(x, \theta) \quad \forall \theta \in \Theta$

$(ii) \; \forall \theta_1 \in \Theta_1 \quad \sum_{\theta_{-1}} \left(u_1(x(\theta_{-1}, \theta_1), (\theta_{-1}, \theta_1)) + t_1(\theta_{-1}, \theta_1) \right) p(\theta_{-1}|\theta_1) = \sum_{i, \theta_{-1}} u_i(x(\theta_{-1}, \theta_1), (\theta_{-1}, \theta_1)) p(\theta_{-1}|\theta_1)$

Theorem 3 says that a neutral optimum outcome is generically unique. It must be ex-post efficient and allocate all surplus to the primary agent in such a way that each type of the primary agent gets all the expected social surplus conditional on her type. By Theorems 1 and 2 and Lemma 3, the existence of such mechanism is generic.

In the appendix we also demonstrate how a neutral optimum mechanism can be supported as a sequential equilibrium.

5 Conclusions.

In this paper we have characterized necessary and sufficient conditions for ex-post budget-balanced, interim individually rational efficient Bayesian implementation. These conditions are a well-known Crémer-McLean condition and Identifiability condition introduced in this paper. They hold generically when there are at least three agents and none of them has more types than the number of different type profiles of the other agents. We have provided an intuitive explanation of the mechanisms that have the above properties and shown that these mechanisms allow to distribute the social surplus between agent types in any desirable way.

An extension of our results demonstrates that, also generically, efficiency is attainable in an informed principal framework, and the informed principal can extract all surplus from the other agents. We believe that our results can be extended to the case of continuous type distribution, and we plan to address this issue in our future work.

6 Appendix

Proof of Lemma 1: Suppose that $p(.)$ is not identifiable, i.e. for some $\hat{\pi}(.) \neq p(.)$ there exist collections of nonnegative coefficients $c_i^{kk'}$ s.t. $\hat{\pi}(.|\theta_i^{k'}) = \sum_{k\in\{1,...,m_i\}} c_i^{kk'} p(.|\theta_i^k)$ for all i and $k, k' \in \{1, ..., m_i\}$.

Let $L = \max_i \left\{ \max_k \frac{\sum_{k'} c_i^{kk'}}{p(\theta_i^k)} \right\}$, and for every i, $k, k' \in \{1, ..., m_i\}$, $k \neq k'$ set:

$$s_i^{kk'} = \frac{c_i^{kk'} \hat{\pi}_i(\theta_i^{k'})}{p(\theta_i^k)L}; \quad s_i^{kk} = 1 - \sum_{k'\neq k} s_i^{kk'} \geq 0; \quad b_i^k = p_i(\theta_i^k)(1 - s_i^{kk}) + \frac{c_i^{kk} \hat{\pi}_i(\theta_i^k)}{L} \geq 0$$

Then for all i we have $\pi(.|s_i, s_{-i}^*) + \sum_k b_i^k \pi(.|k, k, s_{-i}^*) = \frac{\hat{\pi}(.)}{L} + p$.

Conversely, suppose that for some $\hat{\pi}(.) \neq \alpha p(.)$, $\alpha \geq 0$, and every agent i there exists $s_i \in S_i$ and $b_i^k \geq 0$, $k \in \{1, ..., m_i\}$ s.t. $\pi(.|s_i, s_{-i}^*) + \sum_k b_i^k \pi(.|k, k, s_{-i}^*) = \hat{\pi}(.)$. Then $\sum_k b_i^k = \sum_h b_j^h$ for all i and j, and $\tilde{\pi}(.) \equiv \frac{\hat{\pi}(.)}{1+\sum_k b_i^k} \in \Delta^{\prod_l m_l - 1}$. Then we have:

$$\sum_{k,k'} \frac{s_i^{kk'} p_i(\theta_i^k)}{1 + \sum_k b_i^k} \pi(.|k, k', s_{-i}^*) + \sum_k \frac{b_i^k}{1 + \sum_k b_i^k} \pi(.|k, k, s_{-i}^*) = \tilde{\pi}(.)$$

So, for all i and all $k' \in \{1, ..., m_i\}$ we have:

$$\tilde{\pi}(.|\theta_i^{k'}) = \sum_{k\neq k'} \frac{s_i^{kk'} p_i(\theta_i^k)}{(1 + \sum_k b_i^k)p_i(\theta_i^{k'})} p(.|\theta_i^k) + \left(\frac{s_i^{k'k'} p(\theta_i^{k'}) + b_i^{k'}}{(1 + \sum_k b_i^k)p(\theta_i^{k'})} \right) p(.|\theta_i^{k'})$$

Thus, $p(.)$ is not identifiable. Q.E.D.

18

Proof of Lemma 2: Suppose that $p(.)$ is not identifiable. Let us show that for any agent i there exists $k \in \{1, ..., m_i\}$ such that for all $j \neq i$ and $h \in \{1, ..., m_j\}$ the conditional probability vectors $p_{-i-j}(.|\theta_i^{k'}, \theta_j^h)$, $p_{-i-j}(.|\theta_i^k, \theta_j^{h'})$, $k' \in \{1, ..., m_i\}$, $h' \in \{1, ..., m_{j_k}\}$, $h' \neq h$, are not linearly independent.

Since $p(.)$ is not identifiable, there exists $\pi(.) \in \Delta^{\Pi_l m_l - 1}$, $\pi(.) \neq p(.)$, such that for any i and $k \in \{1, ..., m_i\}$ there is a collection of nonnegative coefficients $\{c_i^{k'k}\}$, $k' \in \{1, ..., m_i\}$, satisfying $\pi(.|\theta_i^k) = \sum_{k' \in \{1, ... m_i\}} c_i^{k'k} p(.|\theta_i^{k'})$. In particular, choose k such that $\pi(.|\theta_i^k) \neq p(.|\theta_i^k)$ (such k exists because $\pi(.) \neq p(.)$). Then $c_i^{k'k} > 0$ for some $k' \neq k$. Thus, for all $\theta_{-i} \in \Theta_{-i}$ we have:

$$\pi(\theta_{-i}, \theta_i^k) = \pi_i(\theta_i^k) \sum_{k' \in \{1, ... m_i\}} c_i^{k'k} p(\theta_{-i}|\theta_i^{k'}).$$

Similarly, for any other agent j, any $h \in \{1, ..., m_j\}$ and any $\theta_{-j} \in \Theta_{-j}$ we have:

$$\pi(\theta_{-j}, \theta_j^h) = \pi_j(\theta_j^h) \sum_{h' \in \{1, ... m_j\}} c_j^{h'h} p(\theta_{-j}|\theta_j^{h'}).$$

Combining the last two expressions and using the fact that $p(\theta_{-i-j}, \theta_j^h|\theta_i^k) = p(\theta_{-i-j}|\theta_i^k, \theta_j^h) p(\theta_j^h|\theta_i^k)$ and similarly $p(\theta_{-i-j}, \theta_i^k|\theta_j^h) = p(\theta_{-i-j}|\theta_i^k, \theta_j^h) p(\theta_i^k|\theta_j^h)$, we obtain that for all $\theta_{-i-j} \in \Theta_{-i-j}$:

$$\pi(\theta_{-i-j}, \theta_i^k, \theta_j^h) = \pi_i(\theta_i^k) \sum_{k' \in \{1, ... m_i\}} c_i^{k'k} p(\theta_j^h|\theta_i^k) p(\theta_{-i-j}|\theta_i^{k'}, \theta_j^h) = \pi_j(\theta_j^h) \sum_{h' \in \{1, ... m_j\}} c_j^{h'h} p(\theta_i^k|\theta_j^h) p(\theta_{-i-j}|\theta_i^k, \theta_j^{h'})$$

$$(12)$$

Since $c_i^{k'k} > 0$ for some $k' \neq k$, (12) implies that $p_{-i-j}(.|\theta_i^{k'}, \theta_j^h)$, $p_{-i-j}(.|\theta_i^k, \theta_j^{h'})$, $k' \in \{1, ..., m_i\}$, $h' \in \{1, ..., m_j\}$, $h' \neq h$, are not linearly independent. $\quad Q.E.D.$

Proof of Theorem 1: The theorem will be proved in a number of steps. Throughout, $x(.)$ is an arbitrary EASR decision rule.

Step 1. Restatement of the Problem. To begin with, let us rewrite IC, IR and BB constraints given by (1), (2) and (3) respectively in a matrix form. Let B_i be a matrix of size $m_i(m_i - 1) \times L$ each row of which corresponds to a different IC constraint of agent i. All rows from $(k-1)(m_i - 1) + 1$ to $k(m_i - 1)$ of B_i correspond to incentive constraints of type $k \in \{1, ..., m_i\}$ of agent i, and each column corresponds to one of $L \equiv \prod_l m_l$ possible type profiles in the natural order induced by the ordering of agents and their types. Then, the row corresponding to $IC_i(k, k')$ is equal to $\pi(.|k, k, \mathbf{s}_{-i}^*) - \pi(.|k, k', \mathbf{s}_{-i}^*)$, the difference of the corresponding truthtelling and deviation components of agent i. Its entry in the column corresponding to the type profile $(\theta_{-i}, \theta_i^k)$ for some $\theta_{-i} \in \Theta_{-i}$ is equal to $p(\theta_{-i}|\theta_i^k)$, its entry in the column corresponding to the type profile $(\theta_{-i}, \theta_i^{k'})$ is equal to $-p(\theta_{-i}|\theta_i^k)$, while the entries in all other columns are zero.

To rewrite the IR constraints, we use $m_i \times L$ matrix P_i formed by stacking the truthtelling components of agent i one on top of another. The k-th row of this matrix corresponds to $IR_i(\theta_i^k)$ and is equal to $\pi(.|k, k, \mathbf{s}_{-i}^*)$. Its entry in the column corresponding to the type profile $(\theta_{-i}, \theta_i^k)$ for some $\theta_{-i} \in \Theta_{-i}$ is equal to $p(\theta_{-i}|\theta_i^k)$, and entries in all other columns are zero.

Also, for all $i \in \{1, ..., n\}$ and $k, k' \in \{1, ..., m_i\}$ s.t. $k \neq k'$, let

$$\bar{u}_{ik} = \sum_{\theta_{-i} \in \Theta_{-i}} u_i(x(\theta_{-i}, \theta_i^k), (\theta_{-i}, \theta_i^k)) p(\theta_{-i}|\theta_i^k) \tag{13}$$

$$\hat{u}_{ikk'} = \sum_{\theta_{-i} \in \Theta_{-i}} \left(u_i(x(\theta_{-i}, \theta_i^k), (\theta_{-i}, \theta_i^k)) - u_i(x(\theta_{-i}, \theta_i^{k'}), (\theta_{-i}, \theta_i^k)) \right) p(\theta_{-i}|\theta_i^k) \tag{14}$$

Construct vectors $\bar{u}_i$ and $\hat{u}_i$ by concatenating $\bar{u}_{ik}$ and $\hat{u}_{ikk'}$, i.e. $\bar{u}_i = (\bar{u}_{i1}, ..., \bar{u}_{im_i})'$ and $\hat{u}_i = (\hat{u}_{i12}, ..., \hat{u}_{i1m_i}, ..., \hat{u}_{im_i1}, ..., \hat{u}_{im_i(m_i-1)})'$ where prime denotes a transpose. Let I be an identity matrix of size $L \times L$. Then (1), (2) and (3) can respectively be rewritten as follows:

$$(IC) \quad \begin{Vmatrix} B_1 & 0 & 0 & 0 \\ 0 & B_2 & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & 0 & B_n \end{Vmatrix} \times \begin{bmatrix} t_1 \\ t_2 \\ ... \\ t_n \end{bmatrix} \geq \begin{bmatrix} -\hat{u}_1 \\ -\hat{u}_2 \\ ... \\ -\hat{u}_n \end{bmatrix}$$

$$(IR) \quad \begin{Vmatrix} P_1 & 0 & 0 & 0 \\ 0 & P_2 & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & 0 & P_n \end{Vmatrix} \times \begin{bmatrix} t_1 \\ t_2 \\ ... \\ t_n \end{bmatrix} \geq \begin{bmatrix} -\bar{u}_1 \\ -\bar{u}_2 \\ ... \\ -\bar{u}_n \end{bmatrix} \tag{15}$$

$$(BB) \quad \begin{Vmatrix} I & I & ... & I \end{Vmatrix} \times \begin{bmatrix} t_1 \\ t_2 \\ ... \\ t_n \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$$

Thus, the decision rule $x(\theta)$ is implementable via a BB and IR Bayesian mechanism if and only if there exists a solution $(t_1, ..., t_n)$ to the system (15).

Step 2. *Necessary and Sufficient Conditions for the Existence of a solution to (15).*

The Theorem of The Alternative: Consider a system of linear inequalities:

$$\begin{aligned} Sx &\geq a \\ Tx &= b \end{aligned} \tag{16}$$

where a and b are fixed vectors of size l_1 and l_2 respectively, while S and T are fixed matrices of size $l_1 \times l_3$ and $l_2 \times l_3$ respectively.

System (16) has a solution x^* if and only if for any row vector λ_S of size l_1 with nonnegative entries and any row vector λ_T of size l_2 such that $\lambda_S S + \lambda_T T = \mathbf{0}$, we have:

$$\lambda_T b + \lambda_S a \leq 0.$$

For the proof of the Theorem of the Alternative see Mangasarian (1969) (page 34). Using the Theorem and exploiting the block structure of matrices in (15), we conclude that system (15) has a solution if and only if the following property holds:

Property D. *Consider any collection of row vectors $\{\gamma_1, ..., \gamma_n\}$, $\{\lambda_1, ..., \lambda_n\}$, μ where γ_i is of size $m_i(m_i - 1)$, λ_i is of size m_i, $\gamma_i \geq 0$, $\lambda_i \geq 0$ for all $i = 1, ..., n$, and μ is of size L*

$$\text{such that} \quad \gamma_i B_i + \lambda_i P_i + \mu = \mathbf{0} \quad \text{for all} \quad i = 1, ..., n, \tag{17}$$

$$\text{Then,} \qquad \sum_{i=1,...,n} \gamma_i \hat{u}_i + \sum_{i=1,...,n} \lambda_i \bar{u}_i \geq 0$$

In the rest of the proof we will show that **Property D** holds under the conditions of the Theorem. So, *let γ_i, λ_i, μ be an arbitrary collection of vectors satisfying the conditions of Property D.* Let $\gamma_i^{kk'}$ denote the entry of vector γ_i corresponding to the incentive constraint $IC_i(k, k')$. Note that this entry is $(k-1)(m_i-1)+k'$-th if $k' < k$ and $(k-1)(m_i-1)+k'-1$-th if $k' > k$. Also, let λ_i^k denote the k-th entry of vector λ_i corresponding to the individual rationality constraint $IR_i(k)$.

Step 3. $\gamma_i = \mathbf{0}$ for all i.
Proof: The proof is by contradiction. So, suppose that $\gamma_j^{h_1 h_2} > 0$ for some j, and $h_1, h_2 \in \{1, ..., m_j\}$, $h_1 \neq h_2$. Let us show that either $p(.)$ is not identifiable or Crémer-McLean condition fails for j. By (17), we have:

$$\gamma_i B_i + \lambda_i P_i = \gamma_l B_l + \lambda_l P_l \quad for \ all \ i, l \in \{1, ..., n\} \tag{18}$$

By definition, for all i, $\gamma_i B_i = \sum_{k,k' \in \{1,...,m_i\}, k' \neq k} \gamma_i^{kk'} (\pi(.|k, k, \mathbf{s}^*_{-i}) - \pi(.|k, k', \mathbf{s}^*_{-i}))$,
and $\lambda_i P_i = \sum_{k \in \{1,...,m_i\}} \lambda_i^k \pi(.|k, k, \mathbf{s}^*_{-i})$.

Let $F = \max_i \left\{ \max_k \frac{\sum_{k':k' \neq k} \gamma_i^{kk'} + \lambda_i^k}{p_i(\theta_i^k)} \right\} > 0$, and set γ_i^{kk} to satisfy $\gamma_i^{kk} = Fp(\theta_i^k) - \sum_{k':k' \neq k} \gamma_i^{kk'} - \lambda_i^k \geq 0$. Then $\gamma_i B_i + \lambda_i P_i = Fp(.) - \sum_{k,k' \in \{1,...,m_i\}} \gamma_i^{kk'} \pi(.|k, k', \mathbf{s}^*_{-i})$. Thus, (18) implies that $\sum_{k,k' \in \{1,...,m_i\}} \gamma_i^{kk'} \pi(.|k, k', \mathbf{s}^*_{-i}) = \sum_{h,h' \in \{1,...,m_l\}} \gamma_l^{hh'} \pi(.|h, h', \mathbf{s}^*_{-l})$ for all i, l. Summing up over all $\theta \in \Theta$, we conclude that for all i $\sum_{k,k' \in \{1,...,m_i\}} \gamma_i^{kk'} = K$ for some $K > 0$. $K > 0$ since $\gamma_j^{h_1 h_2} > 0$ for some $h_1 \neq h_2$. Hence, $\frac{\sum_{k,k' \in \{1,...,m_i\}} \gamma_i^{kk'} \pi(.|k, k', \mathbf{s}^*_{-i})}{K} = \frac{\sum_{h,h' \in \{1,...,m_l\}} \gamma_l^{hh'} \pi(.|h, h', \mathbf{s}^*_{-l})}{K}$ is an element of $\Delta^{\prod_l m_l - 1}$ which we denote by $\hat{\pi}(.)$.

If $\hat{\pi} \neq p(.)$, then $p(.)$ is not identifiable, because $\hat{\pi}_{-i}(.|\theta_i^k) = \sum_{k' \in \{1,...,m_i\}} \frac{\gamma_i^{kk'}}{K} p_{-i}(.|\theta_i^{k'})$ for all i and k. If $\hat{\pi} = p(.)$, then Crémer-McLean condition fails for j because $\gamma_j^{h_1 h_2} > 0$ j for some $h_1 \neq h_2$.

Step 4. *There exists $\bar{\Lambda}$, s.t. $\frac{\lambda_i^k}{p_i(\theta_i^k)} = \bar{\Lambda} \geq 0$ for all i and $k \in \{1, ..., m_i\}$.*
Step 3 and equation (18) imply that

$$\lambda_i P_i = \lambda_j P_j \quad for \ all \ i, j \in \{1, ..., n\}$$

which is equivalent to: $\lambda_i^k p(\theta_{-i-j}, \theta_j^h | \theta_i^k) = \lambda_j^h p(\theta_{-i-j}, \theta_i^k | \theta_j^h)$ for all $k \in \{1, ..., m_i\}$, $h \in \{1, ..., m_j\}$ and $\theta_{-i-j} \in \Theta_{-i-j}$. Then the statement of Step 4 follows because $p_{i,j}(\theta_i, \theta_j) > 0$ for all $\theta_i \in \Theta_i$, $\theta_j \in \Theta_j$ and $p(\theta_{-i-j}, \theta_j^h | \theta_i^k) p_i(\theta_i^k) = p(\theta_{-i-j}, \theta_i^k | \theta_j^h) p_j(\theta_j^h)$.

Step 5. To establish **Property D**, we need to show that $\sum_{i=1,...,n} \gamma_i \hat{u}_i + \sum_{i=1,...,n} \lambda_i \bar{u}_i \geq 0$.
By Step 3, $\gamma_i = \mathbf{0}$ for all i. Step 4 implies that $\sum_{i=1,...,n} \lambda_i \bar{u}_i = \bar{\Lambda} \sum_{i=1,...,n} \sum_{k=1,...,m_i} \bar{u}_i^k p_i(\theta_i^k)$. Recall that $\bar{\Lambda} \geq 0$ by assumption of Property D, while $\sum_{i=1,...,n} \sum_{k=1,...,m_i} \bar{u}_i^k p_i(\theta_i^k) = \sum_{\theta \in \Theta} \sum_{i=1}^n u_i(x(\theta), \theta) p(\theta) \geq 0$ by EASR. $\qquad$ *Q.E.D.*

Step 6. Necessity:

If either $p(.)$ is not identifiable or Crémer-McLean condition fails for some agent i, then there exists a profile of utility functions s.t. an ex-post efficient, IR, BB mechanism does not exist.

If $p(.)$ is not identifiable, there exists $\pi(.) \in \Delta^{\Pi_l m_l - 1}$, $\pi(.) \neq p(.)$, s.t. for all l and h' $\pi(.|\theta_l^{h'}) = \sum_{h \in \{1,...,m_l\}} c_l^{hh'} p(.|\theta_l^h)$ for some collection of nonnegative coefficients $c_l^{hh'}$. Since $\pi(.) \neq p$, for all l there exists $h_1, h_2 \in \{1, ..., m_l\}$ s.t. $c_l^{h_1 h_2} > 0$. Set $\hat{\gamma}_l^{hh'} = c_l^{hh'} p_l(\theta_l^{h'})$ for all $h \neq h'$, $\hat{F} = \max_l \{\max_h \sum_{h'} \frac{c_l^{hh'} p_l(\theta_l^{h'})}{p_l(\theta_l^h)}\}$, and $\hat{\lambda}_l^h = \hat{F} p_l(\theta_l^h) - \sum_{h'} c_l^{hh'} p_l(\theta_l^{h'}) \geq 0$ for all h. Then for all l:

$$\hat{\gamma}_l B_l + \hat{\lambda}_l P_l = \sum_{h,h':h \neq h'} c_l^{hh'} p_l(\theta_l^{h'})(\pi(.|h, h, s_{-l}^*) - \pi(.|h, h', s_{-l}^*)) +$$

$$\sum_h (\hat{F} p_l(\theta_l^h) - \sum_{h'} c_l^{hh'} p_l(\theta_l^{h'}))\pi(.|h, h, s_{-l}^*) = \hat{F} p(.) - \pi(.)$$

Similarly, if Crémer-McLean condition fails for some agent i then there exist $r' \in \{1, ..., m_i\}$ s.t. $p(.|\theta_i^{r'}) = \sum_{k \in \{1,...,m_i\}} c_i^{kr'} p(.|\theta_i^k)$ with $c_i^{kr'} \geq 0$ for all k and $c_i^{rr'} > 0$ for some $r \neq r'$.

Then set $\hat{\gamma}_i^{kr'} = c_i^{kr'} p_i(\theta_i^{r'})$ for all $k \neq r'$, $\hat{\gamma}_i^{kk'} = 0$ for $k' \neq r'$ and $\hat{\gamma}_j^{hh'} = 0$ for all $j \neq i$, and $h, h' \in \{1, ..., m_j\}$. Also, let $\tilde{F} = p(\theta_i^{r'}) \max_k \frac{c_i^{kr'}}{p(\theta_i^k)}$, and set $\hat{\lambda}_i^k = \tilde{F} p(\theta_i^k) - c_i^{kr'} p(\theta_i^{r'})$ for $k \neq r'$ and $\hat{\lambda}_i^{r'} = (\tilde{F} + 1 - c_i^{r'r'}) p(\theta_i^{r'})$. Finally, set $\hat{\lambda}_j^h = \tilde{F} p(\theta_j^h)$ for $j \neq i$ and all $h \in \{1, ..., m_j\}$. Then, $\hat{\gamma}_l B_l + \hat{\lambda}_l P_l = \tilde{F} p(.)$ for all l. Thus, in both cases we have $\hat{\gamma}_i B_i + \hat{\lambda}_i P_i = \hat{\gamma}_l B_l + \hat{\lambda}_l P_l$ for all l with $\hat{\gamma}_i^{rr'} > 0$ for some r, r'.

Next, let $X \equiv \{x_1, x_2\}$ and consider the following profile of the utility functions:
(i) $u_l(x_1, \theta) = a > 0$ for all $l \in \{1, ..., n\}$, $\theta \in \Theta$.
(ii) for all $l \neq i$ and $\theta_{-i} \in \Theta_{-i}$: $u_l(x_2, (\theta_{-i}, \theta_i)) = a$ if $\theta_i \neq \theta_i^r$; $u_l(x_2, (\theta_{-i}, \theta_i^r)) = a - zA > 0$.
(iii) for all $\theta_{-i} \in \Theta_{-i}$ $u_i(x_2, (\theta_{-i}, \theta_i^{r'})) = a + \epsilon > 0$, $u_i(x_2, (\theta_{-i}, \theta_i^r)) = a + A$, $u_i(x_2, (\theta_{-i}, \theta_i)) = a - \delta$ if $\theta_i \notin \{\theta_i^r, \theta_i^{r'}\}$.
where $A > 0$, $z > 1/(n-1)$, $\epsilon > 0$, $\delta > 0$.

Then the unique ex-post efficient decision rule $x^*(\theta)$ is such that:

$$x^*(\theta_{-i}, \theta_i) = \begin{cases} x_1, & \text{if} \quad \theta_i \neq \theta_i^{r'} \\ x_2, & \text{if} \quad \theta_i = \theta_i^{r'} \end{cases} \tag{19}$$

By **Property D** in Step 2, $x^*(\theta)$ is not implementable if:

$$\sum_{l \in \{1,...,n\}} \sum_{k,k' \in \{1,...,m_l\}} \hat{\gamma}_l^{kk'} \hat{u}_{lkk'} + \sum_{l \in \{1,...,n\}} \sum_{k \in \{1,...,m_l\}} \hat{\lambda}_l^k \bar{u}_{lk} < 0 \tag{20}$$

By the definitions of $\hat{u}_{lkk'}$, $\bar{u}_{lk}$ (see (13) and (14) respectively) and $x^*(.)$, we obtain:

$\bar{u}_{lk} = a$ for all $l \neq i$ and $k \in \{1, ..., m_l\}$, $\qquad \bar{u}_{ik} = \begin{cases} a, & \text{if} \quad k \neq r' \\ (a + \epsilon), & \text{if} \quad k = r' \end{cases}$

$\hat{u}_{lkk'} = 0$ for all $l \neq i$ and $k, k' \in \{1, ..., m_l\}$, $\qquad \hat{u}_{ikk'} = \begin{cases} -A, & \text{if} \quad k = r, k' = r' \\ \epsilon, & \text{if} \quad k = r' \text{ and any } k' \neq r' \\ \delta, & \text{if} \quad k \notin \{r, r'\}, k' = r' \\ 0, & \text{otherwise} \end{cases}$

Substituting these values into (20), we obtain that $x^*(\theta)$ is not implementable if

$$-\hat{\gamma}_i^{rr'} A + \epsilon \left(\sum_{k=\{1,...,m_i\}} \hat{\gamma}_i^{r'k} + \hat{\lambda}_i^{r'} \right) + a \sum_{l\in\{1,...,n\},k\in\{1,...,m_l\}} \hat{\lambda}_l^k + \delta \sum_{k\notin\{r,r'\}} \hat{\gamma}_i^{kr'} < 0 \qquad (21)$$

Since $\hat{\gamma}_i^{rr'} > 0$, (21) holds when A is large enough and a, ϵ and δ are sufficiently small. *Q.E.D.*

Proof of Theorem 2:

Step 1. *Fix some agent $j \in \{1, ..., n\}$ and her type $\theta_j^h \in \Theta_j$. Let us show that there exists an IC, IR and BB mechanism $(x(\theta), t(\theta))$ with $v_j(\theta_j^h) p_j(\theta_j^h) = \sum_{\theta\in\Theta, i\in\{1,...,n\}} u_i(x(\theta), \theta) p(\theta)$, $v_j(\theta_j) = 0$ for $\theta_j \neq \theta_j^h$, and $v_i(\theta_i) = 0$ for all $\theta_i \in \Theta_i$, $i \neq j$, i.e. type θ_j^h obtains all expected surplus and all other agent types are held at their reservation utility levels.*

Such mechanism exists if there exists a solution to the system of inequalities and equalities obtained from the system (15) by changing all inequalities in its subsystem IR, except for the one corresponding to $IR_j(\theta_j^h)$, to strict equalities, and leaving unchanged the subsystems of inequalities BB and IC, as well as the inequality corresponding to $IR_j(\theta_j^h)$.

By the Theorem of the Alternative (see Step 2 of the proof of Theorem 1), the modified system has a solution if and only if the following **Property D'** holds:

For all families of row vectors $\{\gamma_i\}_{i=1,...,n} \geq 0$, $\{\lambda_i\}_{i=1,...,n}$ s.t. $\lambda_j^h \geq 0$ and vectors μ :

$$\gamma_i B_i + \lambda_i P_i + \mu = \mathbf{0} \quad \forall i = 1, ..., n \implies \sum_{i=1,...,n} \gamma_i \hat{u}_i + \sum_{i=1,...,n} \lambda_i \bar{u}_i \geq 0 \qquad (22)$$

Note that Property D' differs from Property D (see Step 2 of Theorem 1) only insofar that the sign of all λ's except λ_j^h is now unrestricted.

The rest of the proof can be completed by repeating Steps 3-5 in the proof of Theorem 1. Steps 3-4, which establish that all γ's are zero and λ's are equal to each other, apply verbatim. The arguments in these Steps rely on the nonnegativity of the entries of vectors γ_i, but not λ_i. Since $\lambda_j^h \geq 0$, we conclude that all λ's, as well as Λ specified in Step 4 are nonnegative. So, the argument in Step 5 of Theorem 1 also applies.

Step 2. Fix a collection of $\sum_i m_i$ nonnegative constants $v_i(\theta_i^k)$ satisfying (10). Define $\alpha(i,k) = \frac{v_i(\theta_i^k) p_i(\theta_i^k)}{\sum_{i\in\{1,...,n\},\theta\in\Theta} u_i(x(\theta),\theta) p(\theta)}$. Let $(x(\theta), t^{(i,k)}(\theta))$ be an IC, IR and BB direct mechanism which implements decision rule $x(\theta)$ and allocates all surplus to type θ_i^k of agent i. By Step 1 such mechanism exists for all i and $k \in \{1, ..., m_i\}$.

Now consider direct mechanism $(x(\theta), \bar{t}(\theta))$ where

$$\bar{t}_j(\theta) = \sum_{i=1,...,n} \sum_{k=1,...,m_i} \alpha(i,k) t_j^{(i,k)}(\theta)$$

$(x(\theta), \bar{t}(\theta))$ is IC, IR and BB mechanism because IC, IR and BB constraints are linear in transfers, $\sum_{i\in\{1,...,n\},k\in\{1,...,m_i\}} \alpha(i,k) = 1$ and all $\alpha(i,k)$ are nonnegative. By linearity, $\sum_{\theta_{-j}\in\Theta_{-j}} (u_j(x(\theta_{-j}, \theta_j^h), (\theta_{-j}, \theta_j^h)) + \bar{t}_j(\theta_{-j}, \theta_j^h)) p(\theta_{-j}|\theta_j^h) = v_j(\theta_j^h)$ for all j and $h \in \{1, ..., m_j\}$. *Q.E.D.*

Proof of Lemma 3:
Step 1. Definition of genericity. The set of L-point probability distributions $\{p(.) \in \mathbf{R}_+^L | \sum_{\theta \in \Theta} p(\theta) = 1, \ p(\theta) \geq 0\}$ is equivalent to an $L-1$ dimensional simplex $\Delta^{L-1} \subset \mathbf{R}_+^L$. We say that a property holds generically if it fails on a subset $S \subset \Delta^{L-1}$ of Lebesgue measure 0.

Thus, we need to show that the set of probability distributions in Δ^{L-1} that are not identifiable has Lebesgue measure zero. To avoid operating on a simplex, let us introduce the following convenient transformation. Consider an L-vector $q(.) \in [0,1]^L$. To transform it into a probability distribution vector $p(.) \in \Delta^{L-1}$, let $p(q(.))(\theta) \equiv \frac{q(\theta)}{\sum_{\theta \in \Theta} q(\theta)}$. This transformation is a continuous open map from $[0,1]^L \setminus \mathbf{0}$ onto Δ^{L-1}. Although the definition of identifiability in the text is given for elements of Δ^{L-1}, it applies *vertbatim* to any $q(.) \in [0,1]^L \setminus \mathbf{0}$. Then it is easy to see that, if $q(.) \in [0,1]^L$ is identifiable, then so is $p(q(.))$. So we can assume without loss of generality that $p(.) \in [0,1]^L$ and establish that almost all such $p(.)$ are identifiable. With a slight abuse of terminology, we will still refer to $p(.) \in [0,1]^L$ as a probability distribution vector.

Step 2. Let W_i (P_i) be an $m_i(m_i-1) \times L$ ($m_i \times L$) matrix formed by stacking all deviation (truthtelling) component vectors of agent i on top of each other in the natural order of i's types. So, for $k \in \{1, ..., m_i\}$, the $((k-1)m_i + k')$-th row of W_i is equal to $\pi(.|k, k', s_{-i}^*)$ if $k' < k$ and is equal to $\pi(.|k, k'+1, s_{-i}^*)$ if $k' \in \{k, ..., m_i\}$, while the k-th row of P_i is equal to $\pi(.|k, k, s_{-i}^*)$. Let us define **Condition G** as follows:

$$If \quad \psi_j W_j + \zeta_j P_j + \psi_i W_i + \zeta_i P_i = \mathbf{0} \tag{23}$$
$$for \ some \ row \ vectors \ \psi_i, \psi_j, \zeta_i \ and \ \zeta_j, \quad then \quad \psi_j \equiv \mathbf{0}.$$

Comparing Condition G to expression (7) in Lemma 1, it is easy to see that Condition G implies that $p(.)$ is identifiable. In fact, Condition G is significantly stronger. So, to complete the proof, we will show that Condition G holds under almost all $p(.)$ for $j \in \arg\min_{l \in \{1,...,n\}} m_l$ and $i \in \arg\min_{l \neq j} m_l$.

Step 3. Preliminary Facts.
Fact 1. A set $\{(x_1, ..., x_L) \in [0,1]^L | (x_1, ..., x_L)$ satisfies a finite number of polynomial equations$\}$ has measure zero.
Fact 2. For any $i, j \in \{1, ..., n\}$, let M_i^{jh} (M_{ik}^j) be an $m_i \times L_{-i-j}$ ($m_j \times L_{-i-j}$) matrix s.t.

$$M_i^{jh} = \left\| \begin{array}{c} \mathbf{p}(.; \theta_i^1, \theta_j^h) \\ \mathbf{p}(.; \theta_i^2, \theta_j^h) \\ ... \\ \mathbf{p}(.; \theta_i^{m_i}, \theta_j^h) \end{array} \right\| ; \quad M_{ik}^j = \left\| \begin{array}{c} \mathbf{p}(.; \theta_i^k, \theta_j^1) \\ \mathbf{p}(.; \theta_i^k, \theta_j^2) \\ ... \\ \mathbf{p}(.; \theta_i^k, \theta_j^{m_j}) \end{array} \right\|$$

where $\mathbf{p}(.; \theta_i^k, \theta_j^h)$ is a vector of size L_{-i-j} each entry of which is equal to $p(\theta_{-i-j}, \theta_i^k, \theta_j^h)$ for some $\theta_{-i-j} \in \Theta_{-i-j}$ arranged in the natural order of agents other than i and j and their types.

Also, let $M_{i(-k)}^{jh}$ be an $(m_j-1) \times L_{-i-j}$ matrix obtained from M_i^{jh} by removing its k-th row. Finally, let $Y_{i(k)}^j$ be an $(m_j^2 + m_i - 1) \times L_{-i}$ matrix such that:

$$Y_{i(k)}^j = \left\| \begin{array}{cccc} M_{ik}^j & 0 & 0 & 0 \\ 0 & M_{ik}^j & 0 & 0 \\ ... & ... & ... & ... \\ 0 & ... & 0 & M_{ik}^j \\ M_{i(-k)}^{j1} & ... & ... & M_{i(-k)}^{jm_j} \end{array} \right\|$$

24

Then for almost all $p(.) \in [0,1]^L$:

(i) If $m_i \leq L_{-i-j}$ ($m_j \leq L_{-i-j}$), then M_i^{jh} (M_{ik}^j) has full row rank, i.e. all its rows are linearly independent, and any of its principal minors (square submatrices in the top-left corner) are non-singular.

(ii) If $m_j^2 + m_i - 1 \leq L_{-i}$, then the matrix $Y_{i(k)}^j$ has full row rank.

Proof of Fact 2. We will prove (i) for M_{ik}^j. The proof for M_i^{jh} is identical. First, we provide a heuristic explanation of the argument. Let $\{U_1, U_2, ..., U_{m_i}\}$ be a collection of principal minors of M_{ik}^j. All such minors exist because $m_j \leq L_{-i-j}$, i.e. M_{ik}^j has more columns than rows. The determinant $(\det(.))$ of U_l is a non-degenerate polynomial. The non-degeneracy follows from the fact that it does not contain any entry of $p(.)$ in more than one term. So, by *Fact 1* the equation $\det(U_l) = 0$ holds on a set of $p(.) \in [0,1]^L$ of measure zero.

Now, let us exhibit the proof in full detail. Note that U_1 is equivalent to $p(\theta_{-i-j}^1, \theta_i^k, \theta_j^1)$ where θ_{-i-j}^1 is a profile of types of players other than i and j s.t. $\theta_l = \theta_l^1$ for all $l \notin \{i,j\}$. Clearly, the set of $p(.) \in [0,1]^L$ s.t. $p(\theta_{-i-j}^1, \theta_i^k, \theta_j^1) = 0$ has measure zero. To proceed by induction, suppose that the determinant of U_l for some $l \in \{1, ..., m_j - 1\}$ is equal to zero on a subset of $[0,1]^L$ of measure zero and consider the determinant of U_{l+1}. We have:

$$\det(U_{l+1}) = \sum_{t=1}^{l+1} (-1)^{l+1+t} p(\theta_{-i-j}^{(t)}, \theta_i^k, \theta_j^{l+1}) \det(U_{(\theta_{-i-j}^{(t)}, \theta_i^k, \theta_j^{l+1})}. \tag{24}$$

where $p(\theta_{-i-j}^{(t)}, \theta_i^k, \theta_j^{l+1})$ is the t-th entry of the vector $\mathbf{p}(.; \theta_i^k, \theta_j^{l+1})$ and $U_{(\theta_{-i-j}^{(t)}, \theta_i^k, \theta_j^{l+1})}$ is the minor of U_{l+1} complementary to $p(\theta_{-i-j}^{(t)}, \theta_i^k, \theta_j^{l+1})$.

Note that for all t, $U_{(\theta_{-i-j}^{(t)}, \theta_i^k, \theta_j^{l+1})}$ does not contain any entry of the vector $\mathbf{p}(.; \theta_i^k, \theta_j^{l+1})$ and $U_{(\theta_{-i-j}^{(l+1)}, \theta_i^k, \theta_j^{l+1})} = U_l$. So, $\det(U_{l+1})$ is linear in the entries of $\mathbf{p}(.; \theta_i^k, \theta_j^{l+1})$, with coefficients -the determinants of the complementary minors- which are not all equal to zero for almost all $p(.) \in [0,1]^L$. Since a finite intersection of sets of full measure has full measure, we conclude that $\det(U_{l+1}) \neq 0$ for almost all $p \in [0,1]^L$ also. Proceeding by induction, we conclude that $\det(U_l) \neq 0$ for all $l \in \{1, ..., m_j\}$ for almost all $p(.)$.

(ii) Applying the same inductive method as in (i), let $\left(Y_1, ..., Y_{m_j^2 + m_i - 1} \right)$ be a collection of all principal minors of the matrix $Y_{i(k)}^j$. Part (i) implies that the first m_j^2 minors $(Y_1, ..., Y_{m_j^2})$ are non-singular for almost all $p(.) \in [0,1]^L$.

Now suppose that $Y_{m_j^2 + s}$, $s \in \{0, m_i - 2\}$, is non-singular for almost all $p(.) \in [0,1]^L$. Let us show that $\det(Y_{m_j^2 + s + 1}) \neq 0$ for almost all $p(.) \in [0,1]^L$. The expansion on the last row of $Y_{m_j^2 + s + 1}$ yields:

$$\det(Y_{m_j^2 + s + 1}) = \sum_{t=1}^{m_j^2 + s + 1} (-1)^{m_j^2 + s + 1 + t} g^{(t)} \det(Y_{m_j^2 + s + 1}^{(t)}). \tag{25}$$

where $g^{(t)}$ is the t-th element of the last row of $Y_{m_j^2 + s + 1}$ $\{\mathbf{p}(\theta_{-i-j}, \theta_i^r, \theta_j^1), ..., \mathbf{p}(\theta_{-i-j}, \theta_i^r, \theta_j^{m_j})\}$, where $r = s + 1$ if $s < k - 1$ and $r = s + 2$ if $s \geq k - 1$, while $Y_{m_j^2 + s + 1}^{(t)}$ is the minor of $U_{m_j^2 + s + 1}$ complementary to $g^{(t)}$. Importantly, for any $t \in \{1, ..., m_j^2 + s + 1\}$, $Y_{m_j^2 + s + 1}^{(t)}$ does not contain any elements of the row $\{\mathbf{p}(\theta_{-i-j}, \theta_i^r, \theta_j^1), ..., \mathbf{p}(\theta_{-i-j}, \theta_i^r, \theta_j^{m_j})\}$. So, $\det(Y_{m_j^2 + s + 1})$ is linear in

the elements of this row, with coefficients -the determinants of the complementary minors- which are not all equal to zero for almost all $p(.) \in [0,1]^L$. Particularly, $\det(Y_{m_j^2+s+1}^{(m_j^2+s+1)}) = \det(Y_{m_j^2+s}) \neq 0$ for almost all $p(.)$. Since a finite intersection of sets of full measure has full measure, $\det(Y_{m_j^2+s+1}) \neq 0$ for almost all $p \in [0,1]^L$. Proceeding by induction, we conclude that $\det(Y_{m_j^2+s'}) \neq 0$ for all $s' \leq m_i - 1$ for almost all $p(.)$. Q.E.D.

Step 3. Consider the matrix $X_i^j \equiv \left\| \begin{matrix} W_j \\ P_j \\ W_i \\ P_i \end{matrix} \right\|$ formed by stacking the matrices W_i, P_i, W_j, P_j on top of each other. Let us modify X_i^j using a series of the following rank-preserving elementary transformations: (i) interchanging some pair of its rows or columns; (ii) multiplying all entries in some row (or column) by a non-zero constant.

First, a row of X_i^j is equal to either a strategy component $\pi(.|h, h', s_{-j}^*)$ of player j, for $h, h' \in \{1, ..., m_j\}$ (the top m_j^2 rows), or a strategy component $\pi(.|k, k', s_{-i}^*)$ of player i, with $k, k' \in \{1, ..., m_i\}$ (the bottom m_i^2 rows). So, multiply the row equal to the strategy component $\pi(.|h, h', s_{-j}^*)$ by $p_j(\theta_j^h)$ and the row equal to the strategy component $\pi(.|k, k', s_{-i}^*)$ by $p_i(\theta_i^k)$.

Second, reorder the top m_j^2 rows of X_i^j corresponding to $\pi(.|h, h', s_{-j}^*)$, first by h, and then by h'. Also, reorder the bottom m_i^2 rows corresponding to $\pi(.|k, k', s_{-j}^*)$, first by k, and then by k'. Third, reorder the columns of X_i^j by agent types in the following sequence: $i, j, 1,...,n$, so that the first L_{-i-j} columns correspond to types $(\theta_{-i-j}, \theta_j^1, \theta_i^1)$ for all $\theta_{-i-j} \in \Theta_{-i-j}$, and so on. It is easy to see that the second and third steps can be done via a sequence of pairwise interchanges of rows and columns respectively. As a result, we get the following matrix $\bar{Y}_i^j$:

$$\bar{Y}_i^j = \left\| \begin{matrix} M_{i1}^j & 0 & 0 & M_{i2}^j & 0 & 0 & .... & M_{im_i}^j & 0 & 0 \\ ... & ... & ... & ... & ... & ... & .... & ... & ... & ... \\ 0 & 0 & M_{i1}^j & 0 & 0 & M_{i2}^j & .... & 0 & 0 & M_{im_i}^j \\ M_i^{j1} & ... & M_i^{jm_j} & 0 & 0 & 0 & .... & 0 & 0 & 0 \\ 0 & 0 & 0 & M_i^{j1} & ... & M_i^{jm_j} & .... & 0 & 0 & 0 \\ ... & ... & ... & ... & ... & ... & .... & ... & ... & ... \\ 0 & 0 & 0 & 0 & 0 & 0 & .... & M_i^{j1} & ... & M_i^{jm_j} \end{matrix} \right\|$$

Note that the matrices X_i^j and $\bar{Y}_i^j$ have the same ranks because the employed elementary transformations are rank-preserving. Also, there is a one-to-one relationship between the set of solutions $(\psi_j, \zeta_j, \psi_i, \zeta_i)$ to the system (23) and the set of solutions to the following system:

$$\left(\delta_1, ..., \delta_{m_j}, -\eta_1, ..., -\eta_{m_i} \right) \bar{Y}_i^j = \mathbf{0} \tag{26}$$

where δ_h and η_k are row vectors of size m_j and m_i respectively.

In particular, since the top m_j^2 rows of the matrix X_i^j correspond to the top m_j^2 rows of the matrix $\bar{Y}_i^j$, the subvector (ψ_j, ζ_j) $((\psi_i, \zeta_i))$ in a solution to (23) corresponds to the subvector $(\delta_1, ..., \delta_{m_j})$ $(-\eta_1, ..., -\eta_{m_i})$ in a solution to (26). So, if there is a solution to (23) s.t. $(\psi_j, \zeta_j) \neq \mathbf{0}$, then $(\delta_1, ..., \delta_{m_j}) \neq \mathbf{0}$ in the corresponding solution to (23), and vice versa.

Recall that (23) has the following solution: $\psi_i = \psi_j \equiv \mathbf{0}$, $\zeta_i^k = p_i(\theta_i^k)$ and $\zeta_j = -p_j(\theta_j^h)$, which corresponds to the solution of (26) in which only the k-th (h-th) entry of δ_k (η_h) is

non-zero and is equal to 1, and which we refer to as 'basic solution.' Therefore, Condition G holds if the only solution to (26) such that $(\delta_1, ..., \delta_{m_j}) \neq \mathbf{0}$ is the 'basic solution.'

So, to complete the proof let us show that for almost all $p(.) \in [0, 1]^L$ the only solution to (26) such that $(\delta_1, ..., \delta_{m_j}) \neq \mathbf{0}$ is the 'basic solution.' The proof is given separately for two cases. In Step 5, we deal with the case of $n = 3$ and $m_1 = m_2 = m_3 = m \geq 3$, while in Step 4 we deal with all other cases.

Step 4. *Suppose that either $n \geq 4$, or $n = 3$ and it is not true that $m_1 = m_2 = m_3$.*

Fix some type $k \in \{1, ..., m_i\}$, and consider $L_{-i} \times (m_j^2 + m_i)$ submatrix $\hat{Y}_{ik}^j$ of $\bar{Y}_{ij}$ which consists of the columns from $(k-1)L_{-i}$-th to kL_{-i} of $\bar{Y}_{ij}$, i.e. the set of columns which have submatrices M_{ik}^j in their 'upper' part, and all rows which have non-zero entries in these columns (which are the rows from 1-st to m_j^2-th and from $(k-1)m_i$-th to km_i-th). Obviously, if $(\delta_1, ..., \delta_{m_j}, -\eta_1, ..., -\eta_{m_i})$ solves (26), then $(\delta_1, ..., \delta_{m_j}, -\eta_k)\hat{Y}_{ik}^j = 0$.

So, it is sufficient to show that the row rank of $\hat{Y}_{ik}^j$ is equal to $m_j^2 + m_i - 1$ for almost all $p(.) \in [0, 1]^L$. Eliminating the $m_j^2 + k$-th row of $\hat{Y}_{ik}^j$ we get the matrix $Y_{i(-k)}^j$. Recall that $j \in \arg\min_{h \in \{1,...,n\}} m_h$ and $i \in \arg\min_{h \in \{1,...,n\}, h \neq j} m_j$. Therefore, $m_j < L_{-i-j}$ and $m_j^2 + m_i - 1 \leq L_{-i}$. So, $Y_{i(-k)}^j$ has full row rank equal to $m_j^2 + m_i - 1$ for almost all $p(.) \in [0, 1]^L$.

Step 5. $n = 3$ and $m_1 = m_2 = m_3 = m \geq 3$.

To show that the system (26) has a unique non-zero solution, first, notice that (26) is equivalent to: $\delta_h M_{ik}^j = \eta_k M_i^{jh}$ for all $k \in \{1, ..., m\}$ and $h \in \{1, ..., m\}$. By Fact 2 and $m_1 = m_2 = m_3 = m$, M_{ik}^j and M_i^{jh} are $m \times m$ non-singular matrices for almost all $p(.)$. Hence, we can assume their non-singularity in the rest of the proof. Then, (26) is equivalent to the following system: $\eta_k = \delta_h M_{ik}^j (M_i^{jh})^{-1}$ for all k and h, which implies that:

$$(\delta_1, -\delta_h) \begin{pmatrix} M_{i1}^j (M_i^{j1})^{-1} & M_{i2}^j (M_i^{j1})^{-1} & ... & M_{im}^j (M_i^{j1})^{-1} \\ M_{i1}^j (M_i^{jh})^{-1} & M_{i2}^j (M_i^{jh})^{-1} & ... & M_{im}^j (M_i^{jh})^{-1} \end{pmatrix} = 0 \quad \text{for all } h \in \{2, ..., m\} \tag{27}$$

So, (26) has a unique solution iff (27) has a unique solution. In turn, (27) has a unique solution if the following square submatrix of it has rank $2m - 1$ for almost all $p(.)$:

$$\begin{pmatrix} M_{i1}^j (M_i^{j1})^{-1} & M_{im}^j (M_i^{j1})^{-1} \\ M_{i1}^j (M_i^{jm})^{-1} & M_{im}^j (M_i^{jm})^{-1} \end{pmatrix} \tag{28}$$

To complete the proof, let us show that the principal minors $\{Z_m, Z_{m+1}, ..., Z_{2m-1}\}$ of (28) (Z_l is an $l \times l$ matrix consisting of the elements of the first l rows and l columns of (28)) have non-zero determinants for almost all $p(.)$. We will proceed by induction. First, $Z_m = M_{i1}^j (M_i^{j1})^{-1}$ is non-singular by Fact 2. Next, suppose that $\det(Z_{m+s-1}) \neq 0$ for some $s \in \{1, ..., m-1\}$. Let us show that $\det(Z_{m+s}) \neq 0$ for almost all $p(.)$. We have:

$$\det(Z_{m+s}) = \sum_{t=1}^{m+s} (-1)^{m+s+t} b_{m+s}^t \det(Z_{m+s}^{-t}). \tag{29}$$

where b_{m+s}^t is the t-th entry in the $m + s$-th row of Z_{m+s} and Z_{m+s}^{-t} is a minor of Z_{m+s} complementary to b_{m+s}^t. Note that $Z_{m+s}^{-(m+s)} = Z_{m+s-1}$.

Let d_{ht} be the entry at the intersection of h-th row and t-th column of $(M_i^{jm})^{-1}$. Then $b_{m+s}^t = \sum_{h=1}^m p(\theta_a^h, \theta_i^1, \theta_j^s) d_{ht}$ (where $a \in \{1,2,3\} \setminus \{i,j\}$) for all $t \in \{1,...,m\}$, and $b_{m+s}^{m+t'} = \sum_{h=1}^m p(\theta_a^h, \theta_i^m, \theta_j^s) d_{ht'}$ for $t' \in \{1,...,s\}$. These are non-zero for almost all $p(.)$. Since $\det\left(Z_{m+s}^{-(m+s)}\right) = \det(Z_{m+s-1}) \neq 0$, the last entry $b_{m+s}^{m+s} \det(Z_{m+s}^{-(m+s)})$ of (29) is non-zero. If all other entries are zero, then $\det(Z_{m+1}) \neq 0$ for almost all $p()$. If there are other non-zero entries in the summation in (29), then with the new notation we can rewrite it as follows:

$$\det(Z_{m+s}) = \sum_{h=1}^m p(\theta_a^h, \theta_i^1, \theta_j^s) \left(\sum_{t=1}^m d_{ht} \det(Z_{m+s}^{-t})(-1)^{m+s+t} \right) + \sum_{h=1}^m p(\theta_a^h, \theta_i^m, \theta_j^s) \left(\sum_{t=1}^s d_{ht} \det(Z_{m+s}^{-(m+t)})(-1)^{s+t} \right)$$

(30)

Note that for almost all $p(.)$, $\frac{p(\theta_a^h, \theta_i^1, \theta_j^s)}{p(\theta_a^{h'}, \theta_i^m, \theta_j^s)} \neq \frac{p(\theta_a^h, \theta_i^1, \theta_j^s)}{p(\theta_a^{h'}, \theta_i^m, \theta_j^s)}$ for all $h, h' \in \{1, ..., m\}$, $s \neq m$, and importantly, the entries of the matrix $(M_i^{jm})^{-1}$ do not depend on any entries of the vector $p(.)$ contained in the first $m-1$ rows of the matrix M_{im}^j. Therefore, for $s \neq m$, (30) is a nondegenerate polynomial in the entries (d_{ht}) of the matrix $(M_i^{jm})^{-1}$, with coefficients which are not all equal to zero. So, by *Fact 1* $\det(Z_{m+s}) \neq 0$ for almost all $p(.) \in [0,1]^L$. Note that this argument does not apply for $s = m$, since the m-th rows of M_{im}^j and M_i^{jm} coincide, and so $\det(Z_{2m})$ is a degenerate polynomial. Thus, the rank of the matrix (28) is $2m - 1$ for almost all $p(.) \in [0,1]^L$.

Proof of Lemma 4: First, let us consider the problem of existence of a system of transfers satisfying Condition C. Using the notation introduced in Step 1 of Theorem 1, this problem can be rewritten as follows. For all $R(\theta) \in \mathbf{R}^L$, we need to establish the existence of a system of transfers $(t_1, ..., t_n)$, $t_i \in \mathbf{R}^L$, such that:

$$(IC) \quad \left\| \begin{matrix} B_1 & 0 & 0 & 0 \\ 0 & B_2 & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & 0 & B_n \end{matrix} \right\| \times \begin{bmatrix} t_1 \\ t_2 \\ ... \\ t_n \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ ... \\ 0 \end{bmatrix}$$

(31)

$$(BB) \quad \left\| \begin{matrix} I & I & ... & I \end{matrix} \right\| \times \begin{bmatrix} t_1 \\ t_2 \\ ... \\ t_n \end{bmatrix} = [\, R(\theta) \,]$$

By the Theorem of Alternative (see Step 2 in the proof of Theorem 1), this system has a solution if and only if for any collection of vectors μ and $\{\gamma_1, ..., \gamma_n\}$, with $\gamma_i \geq 0$ for all i, the following is true:

$$\textbf{If} \quad \gamma_i B_i + \mu = \mathbf{0} \quad \textbf{for all} \quad i, \quad \textbf{Then} \quad \sum_{\theta \in \Theta} \mu(\theta) R(\theta) \geq 0$$

Since this condition has to hold for any $R(\theta)$, we must have $\mu(\theta) \equiv 0$. Hence, the system (31) has a solution for any $R(\theta)$ if and only if

$$\gamma_1 B_1 = \gamma_2 B_2 = ... = \gamma_n B_n \Longrightarrow \gamma_i B_i = \mathbf{0} \quad \text{for all } i$$

(32)

Next, let $V_i^s = \{\beta\pi(.|s_i, \mathbf{s}_{-i}^*)|\beta \geq 0, s_i \in S_i\}$. It is easy to see that the Weak Identifiability Condition is equivalent to the following: $\cap_i V_i^s = \{\alpha p(.)|\alpha \geq 0\}$.

So, to complete the proof we will establish the following claim: (32) holds if and only if $\cap_i V_i^s = \{\alpha p(.)|\alpha \geq 0\}$. Recall that $\gamma_i B_i = \sum_{k,k':k'\neq k} \gamma_i^{kk'}(\pi(.|k,k,s_{-i}^*) - \pi(.|k,k',s_{-i}^*))$. Also, observe that

$$V_i^s = \left\{ \sum_{k,k'\in\{1,...,m_i\}} c_i^{kk'}\pi(.|k,k',s_{-i}^*) \,\middle|\, c_i^{kk'} \geq 0; \text{ there exist } c_i \geq 0 \text{ s.t. for all } k, \sum_{k'} c_i^{kk'} = c_i p_i(\theta_i^k) \right\}$$
(33)

This follows from the fact that $v_i \in V_i^s$ iff $v_i = \alpha\pi(.|s_i, s_{-i}^*) \equiv \alpha \sum_{k,k'} s_i^{kk'}\pi(.|k,k',s_{-i}^*)p_i(\theta_i^k)$ where $\sum_{k'\in\{1,...,m_i\}} s_i^{kk'} = 1$ for all k and $\alpha \geq 0$. Now, to the proof of the claim.

"Only If:" Suppose that there exists a vector $\tilde{v} \in \cap_i V_i^s$ s.t. $\tilde{v} \neq \alpha p(.)$. Then, by (33), for all i there exists a collection of coefficients $\{c_i^{kk'} \geq 0\}_{k,k'\in\{1,...,m_i\}}$ satisfying $\sum_{k'\in\{1,...,m_i\}} c_i^{kk'} = c_i p_i(\theta_i^k) > 0$ for some c_i s.t. $\tilde{v} = \sum_{k,k'\in\{1,...,m_i\}} c_i^{kk'}\pi(.|k,k',s_{-i}^*)$. Observe that $c_i = \tilde{c}$ for all i. This is so because $\sum_{\theta\in\Theta}\sum_{k,k'\in\{1,...,m_i\}} c_i^{kk'} p(\theta|k,k',s_{-i}^*) = c_i$. So, if $c_i \neq c_j$ for some i,j, then $\sum_{k,k'\in\{1,...,m_i\}} c_i^{kk'}\pi(.|k,k',s_{-i}^*) \neq \sum_{h,h'\in\{1,...,m_j\}} c_j^{hh'}\pi(.|h,h',s_{-j}^*)$.

Now, consider the vector $\hat{v} = p(.)\tilde{c} - \tilde{v} \neq 0$. Since $p(.) = \sum_{k\in\{1,...,m_i\}} \pi(.|k,k,s_{-i}^*)p_i(\theta_i^k)$, we conclude that $\hat{v} = \sum_{k,k'\in\{1,...,m_i\},k\neq k'} c_i^{kk'}(\pi(.|k,k,s_{-i}^*) - \pi(.|k,k',s_{-i}^*))$. Hence, setting $\gamma_i^{kk'} = c_i^{kk'}$ for $k,k' \in \{1,...,m_i\}$, $k' \neq k'$, we obtain that $\gamma_i B_i = \hat{v} \neq 0$ for all i, which contradicts (32).

"If:" Suppose that (32) fails, i.e. there exists $w \in \mathbf{R}^L$ s.t. $\gamma_1 B_1 = ... = \gamma_n B_n = w$, $w \neq \mathbf{0}$. Let e_L be an L vector of units. Then $B_i e_L = \mathbf{0}$, so $w \times e_L = \mathbf{0}$. Hence, w has positive and negative entries and so $w \neq \beta p(.)$ for any β. Let $\eta = \max_i \max_k \frac{\sum_{k':k'\neq k}\gamma_i^{kk'}}{p_i(\theta_i^k)} > 0$. Then for all i,

$$-w+\eta p(.) = -\gamma_i B_i + \eta p(.) = \sum_{k,k':k'\neq k} \gamma_i^{kk'}\pi(.|k,k',s_{-i}^*) + \sum_k \left(\eta p_i(\theta_i^k) - \sum_{k':k'\neq k}\gamma_i^{kk'}\right)\pi(.|k,k,s_{-i}^*)$$

So, by (33), $-w+\eta p(\theta) \in V_i^s$ for all i. Since $-w+\eta p(.) \neq \alpha p(.)$, we have $\cap_i V_i^s \neq \{\alpha p(.)|\alpha \geq 0\}$. Q.E.D.

Proof of Theorem 3: By the Inscrutability Principle we can restrict our analysis to direct, IR and BB mechanisms incentive compatible with respect to prior beliefs. By Theorems 1 and 2, this class contains mechanism $M^* = (x^*(\theta), t^*(\theta))$ such that $x^*(\theta)$ is ex-post efficient and each type of agent 1 gets all social surplus conditional on her type, i.e. $U(\theta_1|M^*) = \sum_{i,\theta_{-1}\in\Theta_{-1}} u_i(x^*(\theta_{-1},\theta_1),(\theta_{-1},\theta_1))p(\theta_{-1}|\theta_1)$, while all types of all other agents earn zero.

Let us show that M^* is a neutral optimum, and no mechanism where agent 1's expected payoff vector is different from $\{U(\theta_1|M^*)\}_{\theta_1\in\Theta_1}$ is a neutral optimum. The proof is by exclusion[15]. First, by Theorem 6 in Myerson (1983) a neutral optimum exists.

Also, by Theorem 8 of Myerson (1983), if $(x(.),t(.))$ is a neutral optimum, then there exists

---

[15]It is possible to prove directly that M^* is a direct optimum. We provide such proof in Appendix B available at http://faculty.fuqua.duke.edu/%7Esseverin/mechsupp.pdf

a collection of strictly positive multipliers $\lambda(.)$ s.t. $(x(.), t(.))$ maximizes

$$\sum_{\theta_1 \in \Theta_1} \left(\lambda(\theta_1) \sum_{\theta_{-1} \in \Theta_{-1}} \left(u_1(x(\theta), \theta) + t_1(\theta) \right) p(\theta_{-1}|\theta_1) \right) \tag{34}$$

subject to IC, IR and BB constraints (1)-(3).

By Theorem 2 any allocation of ex-ante expected social surplus can be supported by an IC, IR and BB mechanism. Therefore, in a neutral optimum player 1 must obtain all expected social surplus, i.e. $\sum_{\theta \in \Theta} \left(u_1(x(\theta), \theta) + t_1(\theta) \right) p(\theta) = \sum_{i, \theta \in \Theta} u_1(x(\theta), \theta)p(\theta)$. For if $\sum_{\theta \in \Theta} \left(u_1(x(\theta), \theta) + t_1(\theta) \right) p(\theta) < \sum_{i, \theta \in \Theta} u_1(x(\theta), \theta)p(\theta)$, then $(x(.), t(.))$ cannot be maximizing (34), because there exists another IR, BB mechanism $(x(.), t'(.))$ s.t. $\sum_{\theta_{-1} \in \Theta_{-1}} t'_1(\theta)p(\theta_{-1}|\theta_1) > \sum_{\theta_{-1} \in \Theta_{-1}} t_1(\theta)p(\theta_{-1}|\theta_1)$ for all $\theta_1 \in \Theta_1$. Note that, because of IR and BB, it is infeasible that $\sum_{\theta \in \Theta} \left(u_1(x(\theta), \theta) + t_1(\theta) \right) p(\theta) > \sum_{i, \theta \in \Theta} u_1(x(\theta), \theta)p(\theta)$.

Also, $x(\theta)$ must be ex-post efficient. Otherwise, the value of (34) can be increased by using a mechanism implementing an ex-post efficient decision rule.

It remains to show that in a neutral optimum each type of agent 1 gets all expected social surplus conditional on her type, i.e. $U(\theta_1|(x(.), t(.)) = \sum_{i, \theta_{-1} \in \Theta_{-1}} u_i(x(\theta_{-1}, \theta_1), (\theta_{-1}, \theta_1))p(\theta_{-1}|\theta_1)$ for all $\theta_1 \in \Theta_1$. For this, a different line of argument is required. Define a set of blocked payoff vectors of agent 1 $\hat{B}_k$ as follows:

$$\hat{B}_k \equiv \left\{ z(.) \in \mathbb{R}_+^{m_1} \mid \sum_{\theta_1 \in \Theta_1} z(\theta_1)p_1(\theta_1) \leq \sum_{i; \theta \in \Theta} u_i(x^*(\theta), \theta)p(\theta); \quad z(\theta_1^k) < \sum_{i; \theta_{-1} \in \Theta_{-1}} u_i(x^*(\theta), \theta)p(\theta_{-1}|\theta_1^k) \right\}$$

Clearly, any IC, IR, BB mechanism which is not blocked by $\cup_k \hat{B}_k$ satisfies the conditions of Theorem 3 because there is no IC, IR, BB mechanism $\tilde{M}$ such that $\sum_{\theta_1 \in \Theta_1} U_1(\theta_1|\tilde{M})p_1(\theta_1) > \sum_{i; \theta \in \Theta} u_i(x^*(\theta), \theta)p(\theta)$. So, to complete the proof, it is sufficient to show that $\hat{B}_k$ satisfies Axioms 1-4 for all $k \in \{1, ..., m_1\}$.

It is immediate that the Domination Axiom is satisfied. The Openness Axiom is also satisfied because the 'upper' boundary of the set of feasible payoff vectors of agent 1 is given by $\{y(.) \in \mathbf{R}_+^{m_1} \mid \sum_{\theta_1 \in \Theta_1} y(\theta_1)p_1(\theta_1) = \sum_{i; \theta \in \Theta} u_i(x^*(\theta), \theta)p(\theta)\}$.

The Extension Axiom holds because adding more actions to X may change the set of ex-post efficient decision rules and cause an increase of $\sum_{i=1,...,n; \theta \in \Theta} u_i(x^*(\theta), \theta)p(\theta)$ and, thus, of $\sum_{i=1,...,n; \theta_{-1} \in \Theta_{-1}} u_i(x^*(\theta), \theta)p(\theta_{-1}|\theta_1)$ for some $\theta_1 \in \Theta_1$.

Finally, let us show that $\hat{B}_k$ satisfies Strong Solutions Axiom, i.e. a strong solution, if exists, does not belong to $\hat{B}_k$ for all k. Suppose that mechanism M^s is a strong solution. Then for all $\theta_1 \in \Theta_1$, $U_1(\theta_1|M^s) \leq \sum_{i=1,...,n; \theta_{-1} \in \Theta_{-1}} u_i(x^*(\theta), \theta)p(\theta_{-1}|\theta_1)$. For suppose there is a type $\theta'_1 \in \Theta_1$ such that $U_1(\theta'_1|M^s) > \sum_{i=1,...,n; \theta_{-1} \in \Theta_{-1}} u_i(x^*(\theta), \theta)p(\theta_{-1}|\theta'_1)$. Then in M^s some type $\hat{\theta}_i$ of some agent $i \in \{2, ..., n\}$ earns a negative expected payoff conditional on $(\hat{\theta}_i, \theta'_1)$. But this contradicts the fact that M^s is a strong solution, i.e. is incentive compatible and individually rational for all agent types given any type of agent 1.

But then it could not be that $U_1(\theta_1|M^s) < \sum_{i=1,...,n; \theta_{-1} \in \Theta_{-1}} u_i(x^*(\theta), \theta)p(\theta_{-1}|\theta_1)$ for some $\theta_1 \in \Theta_1$, because in this case M^s would be dominated by the mechanism $M^* = (x^*(\theta), t^*(\theta))$.

Construction of a sequential equilibrium in the informed principal problem.

Consider a neutral optimum mechanism $M^* \equiv (x^*(\theta), t^*(\theta))$ specified in Theorem 3. Recall that $x^*(\theta)$ is ex-post efficient and the vector of the expected payoffs of different types of agent i

in this mechanism $\{U_i(\theta_i|M^*)\}_{\theta_i\in\Theta_i}$ is such that $U_i^*(\theta_i|M^*) = 0$ for all $\theta_i \in \Theta_i$ and $i \in \{2,...,n\}$, and $U_1^*(\theta_1|M^*) = \sum_{i,\theta_{-1}\in\Theta_{-1}} u_i(x^*(\theta_{-1},\theta_1),(\theta_{-1},\theta_1))p(\theta_{-1}|\theta_1)$.

Let us demonstrate the existence of a sequential equilibrium (i.e. a collection of sequentially rational strategies and a consistent belief system) on the equilibrium path of which all types of agent 1 offer M^*, in stage 2 the agents 2 to n do not change their prior beliefs about the type of agent 1 and agree to participate in the mechanism, while in stage 3 all agents, including agent 1, report their types truthfully.

Since M^* is incentive compatible given prior beliefs, we only need to rule out deviations by any type of agent 1 to an alternative mechanism $\nu \in Z$. For this, let us start by considering a modified game Γ^ν which differs from Γ in one aspect: agent 1's only choice at stage 2 is either to exit and earn the *modified reservation payoff* $U_1^*(\theta_1|M^*)$ or to offer the mechanism ν.

Consider any sequential equilibrium of Γ^ν (which exists because $\nu \in Z$). Let $q(\theta_1)$ denote the probability that type θ_1 of agent 1 chooses mechanism ν in this equilibrium, σ^ν denote the agents' strategy profile in the continuation game after ν is offered, and $\{P_i^\nu(\theta_{-i}|\theta_i), O_i^\nu(\theta_{-i}|r,\theta_i)\}$ denote the agents' belief system in this continuation game. $P_i^\nu(\theta_{-i}|\theta_i)$ stands for the posterior beliefs of agent i in stage 3 after the mechanism ν is offered, while $O_i^\nu(\theta_{-i}|r,\theta_i)$ stands for the posterior beliefs of agent i in stage 4 after the agents have taken participation decisions described by the vector r. Note that for agent 1 $P_1^\nu(\theta_{-1}|\theta_1) = p(\theta_{-1}|\theta_1)$. Also, let $W_i(\theta_i|\nu,\sigma^\nu,P_i^\nu(.|\theta_i))$ denote the expected payoff of type θ_i of agent i in this equilibrium.

Sequential rationality requires that:

$$q(\theta_1) = \begin{cases} 1 & \text{if } W_1(\theta_1|\nu,\sigma^\nu,p(.|\theta_1)) > U_1^*(\theta_1|M^*) \\ any\ x \in [0,1] & \text{if } W_1(\theta_1|\nu,\sigma^\nu,p(.|\theta_1)) = U_1^*(\theta_1|M^*) \\ 0 & \text{if } W_1(\theta_1|\nu,\sigma^\nu,p(.|\theta_1)) < U_1^*(\theta_1|M^*) \end{cases} \tag{35}$$

Consistency of beliefs requires that for $i = 2,...,n$:

$$P_i^\nu(\theta_1,\theta_{-1-i}|\theta_i) = \frac{p(\theta_1,\theta_{-1-i}|\theta_i)Q(\theta_1)}{\sum_{\theta'_{-1-i}\in\Theta_{-1-i},\theta'_1\in\Theta_1} p(\theta'_1,\theta'_{-i}|\theta_i)Q(\theta'_1)} \tag{36}$$

for some $Q(.)$ satisfying: $Q(\theta_1)\left(\sum_{\theta'_1\in\Theta_1} q(\theta'_1)\right) = q(\theta_1)$ for all $\theta_1 \in \Theta_1$.

If $W_1(\theta_1|\nu,\sigma^\nu,p(.|\theta_1)) \le U_1^*(\theta_1|M^*)$ for all $\theta_1 \in \Theta_1$, then the strategies σ^ν and the belief system $\{P_i^\nu(\theta_{-i}|\theta_i), O_i^\nu(\theta_{-i}|r,\theta_i)\}$ support a sequential equilibrium of Γ^ν in which all types of agent 1 choose the reservation payoff $U_1^*(.|M^*)$ with probability 1, i.e. $q(\theta_1) = 0$ for all $\theta_1 \in \Theta_1$.

So, to complete the construction we need to rule out that $W_1(\hat\theta_1|\nu,\sigma^\nu,p(.|\hat\theta_1)) > U_1^*(\hat\theta_1|M^*)$ for some $\hat\theta_1 \in \Theta_1$. The proof is by contradiction, so suppose that such $\hat\theta_1$ exists. Then $q(\hat\theta_1) = 1$.

The optimality of agent 1's strategy implies that $W_1(\theta_1|\nu,\sigma^\nu,p(.|\theta_1)) \ge U_1^*(\theta_1|M^*)$ for all $\theta_1 \in \Theta_1$ such that $q(\theta_1) > 0$. Also, by (36) and the definition of $Q(.)$, $P_i^\nu(\theta_1,\theta_{-1-i}|\theta_i) > 0$ only if $q(\theta_1) > 0$. Therefore, since $U_1^*(\theta_1|M^*) = \sum_{i,\theta_{-1}\in\Theta_{-1}} u_i(x^*(\theta_{-1},\theta_1),(\theta_{-1},\theta_1))p(\theta_{-1}|\theta_1)$, there must exist an agent $i \ne 1$ and her type $\theta_i \in \Theta_i$ s.t. $W_i(\theta_i|\nu,\sigma^\nu,P_i^\nu(.|\theta_i)) < 0$. However, this is impossible because this agent type can always ensure herself a reservation payoff of zero by dropping out in stage 3. Hence, $W_1(\theta_1|\nu,\sigma^\nu,p(.|\theta_1)) \le U_1^*(\theta_1|M^*)$ for all $\theta_1 \in \Theta_1$.

Therefore, the strategies $q(\theta_1) = 0$ for all $\theta_1 \in \Theta_1$ and σ^ν, together with the belief system $\{P_i^\nu(\theta_{-i}|\theta_i), O_i^\nu(\theta_{-i}|r,\theta_i)\}$ constitute a sequential equilibrium of the game Γ^ν in which all types of agent 1 choose the reservation payoff $U_1^*(\theta_1|M^*)$ with probability 1 in stage 2. Recall than ν is an arbitrary mechanism in Z.

It follows that the 'informed principal' game Γ^* possesses the following sequential equilibrium. All types of agent 1 offer $M^* = (x^*(\theta), t^*(\theta))$ with probability 1. All agents accept it

31

and report truthfully in stage 3. Off the equilibrium path, if agent 1 offers a mechanism $\nu \in Z$, then the strategies and the beliefs in the continuation game are the same as in the sequential equilibrium of the game Γ^ν, i.e. σ^ν and $\{P_i^\nu(.), O_i^\nu(.)\}$ for $i = 1, ..., n$. $Q.E.D.$

References

Aoyagi, M. (1998), "Correlated Types and Bayesian Incentive Compatible Mechanisms with Budget Balance, " *Journal of Economic Theory*, 79, pp 142-151.

d'Aspremont, C. and L.-A. Gérard-Varet (1979), "Incentive and Incomplete Information," *Journal of Public Economics*, 11, pp. 25-45.

d'Aspremont, C., J. Crémer and L.-A. Gérard-Varet (1990) "Incentive and the Existence of Pareto-Optimal Revelation Mechanisms," *Journal of Economic Theory*, 51, pp. 233-254.

d'Aspremont, C., J. Crémer and L.-A. Gérard-Varet, mimeo., (2003) "Correlation, Independence and Bayesian Incentives," Social Choice and Welfare, 21, pp. 1-30.

d'Aspremont, C., J. Crémer and L.-A. Gérard-Varet (2004) "Balanced Bayesian Mechanisms," forthcoming, Journal of Economic Theory.

Chung, K-S. (1999) "A Note on Matsushima's Regularity Condition," *Journal of Economic Theory*, 87, pp. 429-433.

Cremer, J., and R. McLean (1985), "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent," *Econometrica*, 53, pp. 345-361.

Cremer, J., and R. McLean (1988), "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, pp. 1247-1257.

Fudenberg, D., D. Levine and E. Maskin (1994), "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, pp. 997-1039.

Fudenberg, D., D. Levine and E. Maskin (1996), "Balanced-Budget Mechanisms with Incomplete Information," mimeo. Department of Economics, UCLA.

Jehiel, P., and B. Moldovanu (2001), "Efficient Design with Interdependent Valuations," *Econometrica*, 69, pp. 1237-1261.

Krishna, V., and M. Perry (1998), "Efficient Mechanism Design," *mimeo*, Dept. of Economics, Pennsylvannia State University.

Mailath, G., and A. Postlewaite (1990), "Asymmetric Information Bargaining Problems with Many Agents," *Review of Economic Studies*, 57, pp. 351-367.

Makowski, L., and C. Mezzetti (1994), "Bayesian and Weakly Robust First Best Mechanisms: Characterizaions," *Journal of Economic Theory*, 64, pp. 500-520.

Mangasarian, O. (1969) "Nonlinear Programming," McGraw-Hill, New York.

Maskin, E., and J. Tirole (1990), "The Principal-Agent Relationship with an Informed Principal: The Case of Private Values," *Econometrica*, 58, pp. 379-409.

Maskin, E., and J. Tirole (1992), "The Principal-Agent Relationship with an Informed Principal, II: Common Values," *Econometrica*, 60, pp. 1-42.

Matsushima, H. (1991), "Incentive Compatible Mechanisms with Full Transferability," *Journal of Economic Theory*, 79, pp. 198-203.

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

McLean, R., and A. Postlewaite (2002), "Informational Size and Incentive Compatibility," *Econometrica*, 70, pp. 2421-2455.

McLean R., and A. Postlewaite (2003), "Informational Size and Efficient Auctions," mimeo, University of Pennsylvannia.

McLean, R., and A. Postlewaite (2003), "Informational Size and and Incentive Compatibility with Aggregate Uncertainty," forthcoming in *Games and Economic Behavior.*

Myerson, R., and M. Satterthwaite (1983), "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 28, pp. 265-281.

Myerson, R. (1983), "Mechanism Design by an Informed Principal," *Econometrica*, 51, pp. 1767-1797.

Rob, R. (1989), "Pollution Claim Settlements Under Private Information," *Journal of Economic Theory*, 47, pp. 307-333.

Williams, S. (1999), "A Characterization of Efficient, Bayesian Incentive Compatible Mechanisms," *Economic Theory*, 14, pp. 155-180.