Individually rational, budget-balanced mechanisms and allocation of surplus

Grigory Kosenok[a], Sergei Severinov[b],[*],[1]

[a]*New Economic School, Moscow, Russia*
[b]*Department of Economics, University of Essex, UK*

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 via such mechanisms if and only if the probability distribution of the agents' type profiles satisfies two conditions: the well-known condition of Crémer and McLean [1988. Full extraction of the surplus in Bayesian and dominant strategy auctions, Econometrica 56, 1247–1257] and the *Identifiability* condition introduced in this paper. We also show that these conditions are necessary for ex-post efficiency to be attainable with budget balance and individual rationality, and that the expected social surplus in these mechanisms can be distributed in any desirable way. Lastly, we demonstrate that, like Crémer–McLean condition, the Identifiability condition is generic if there are at least three agents.
© 2007 Elsevier Inc. All rights reserved.

JEL classification: C72; D82

Keywords: Mechanism design; Bayesian implementation; Individual rationality; Ex-post budget balancing; Surplus allocation

1. Introduction

The theory of Bayesian mechanism design provides a universally accepted implementation tool for a large variety of environments, such as contracting, auctions, and bargaining. For this reason, it is important to understand the scope and limits of Bayesian implementation. In this regard, it is reasonable to consider budget balance, individual rationality and efficiency as desirable properties

---

of a mechanism. Examples of environments, where one would like these properties to hold jointly, include standard and double auctions, public good provision, various trading situations. On a more general level, imposing ex-post budget balance allows to separate the design of a mechanism from its operation, since the mechanism designer is no longer needed to play the role of a budget-breaker. This is important when the mechanism takes the form of a constitution, a law, regulations or a code of behavior designed by the previous generations, or in any other situation where the mechanism designer imposes the mechanism but is not present when the mechanism operates.

Unfortunately, the tension between efficiency, individual rationality and budget-balance is a well-known issue. Myerson and Satterthwaite [25] have shown that Bayesian mechanisms possessing these three properties can fail to exist in private values environments with independent types.[2] More precisely, Makowski and Mezzetti [15] demonstrate that such mechanisms exist in these environments, with types of arbitrary finite dimension, if and only if the ex-ante expected deficit in the corresponding Groves–Clarke mechanism does not exceed the sum of the fees that agents can be charged for participating in the mechanism. This condition does not hold in many economically important situations. For example, Rob [27] demonstrates that public good would not be provided efficiently, while Mailath and Postlewaite [14] show that in any feasible mechanism the probability of public good provision goes to zero as the number of agents increases. Continuing this line of research, Williams [28] and Krishna and Perry [13] establish the equivalence of efficient Bayesian and Clarke–Groves mechanisms.

With multidimensional independently distributed types and interdependent values, Jehiel and Moldovanu [12] show that efficient mechanisms exist only in nongeneric situations. However, Mezzetti [23,24] demonstrates that in this case the mechanism designer can achieve efficiency and extract all surplus, if she can run a two-stage mechanism in which the agents first report their types, and then report their realized payoffs.

Relaxing either the individual rationality or budget balance requirement makes it possible to obtain positive results. Various sufficient conditions for efficient Bayesian implementation with ex-post budget balance but without individual rationality have been derived by d'Aspremont and Gérard-Varet [5], d'Aspremont, Crémer and Gérard-Varet [2,3] (Compatibility), Matsushima [17], Aoyagi [1], and Chung [6] (Weak and Strict Regularity), Fudenberg, Levine and Maskin [9,10] (Pairwise Identifiability). Some of these conditions hold with independent types, whereas others require some correlation between the types. Finally, d'Aspremont, Crémer and Gérard-Varet [4] present Condition C which they show to be necessary and sufficient for balanced-budget Bayesian implementation without individual rationality. Condition C holds with both independent and correlated types.

Crémer and McLean [8] demonstrate that an uninformed mechanism designer can implement an ex-post efficient decision rule, preserve individual rationality, and extract all surplus from the agents if the prior probability distribution, from which the agents' type profiles are drawn, satisfies what we henceforth refer to as Crémer–McLean condition. Following McAfee and Reny [19], this condition can be intuitively described as follows: 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 a mechanism is not ex-post budget-balanced. The uninformed mechanism designer plays the important role of a budget-breaker. She collects transfers from the agents, and may also have to pay them in some

---

[2] Myerson and Satterthwaite [25] focus on the case of a continuum of agents' types. But their result also extends to the case with a finite number of types.

states of the world.[3] McAfee and Reny [19] extend the surplus extraction result to the case of a continuous distribution of types.

Without imposing ex-post budget balance, McLean and Postlewaite [20,21] show that the mechanism designer needs to make only small transfers in an individually rational, ex-post efficient mechanism if each agent is "informationally small". That is, if only one agent misrepresents her private information, the state of the world can still be inferred with a high degree of accuracy.

The main result of this paper establishes necessary and sufficient conditions for the existence of an interim individually rational, ex-post budget-balanced, and ex-post efficient mechanisms in the transferable utility framework. The first condition is the previously mentioned condition of Crémer and McLean [8]. The second is the *Identifiability* condition introduced in this paper. These two conditions guarantee that not only ex-post efficient but all *ex ante socially rational*[4] decision rules—those that generate a nonnegative expected social surplus—are implementable via an interim individually rational and ex-post budget-balanced mechanism. We show that ex-ante social rationality of the decision rule is necessary for interim individual rationality and ex-post budget-balancing to hold jointly, so our sufficiency result cannot be extended to a larger set of decision rules.

It is worth noting that our existence proof is constructive, as we develop a method for computing the desired mechanism. In Section 3, we provide a detailed description of our mechanism and of the role which the Identifiability and Crémer–McLean conditions play in it.

The necessity part of our results shows that an efficient, individually rational, budget-balanced mechanism fails to exist under some profiles of the utility functions, if either Identifiability or Crémer–McLean condition does not hold. This result is important, since it implies that these conditions cannot be relaxed further.

Intuitively, the Identifiability condition says the following. Observe that any profile of the agents' reporting strategies in a direct mechanism, in combination with the prior from which the types are drawn, induces a probability distribution over the reported type profiles. The prior is said to be identifiable if, for any probability distribution over the agents' type profiles, q, different from the prior, there exists some agent and her type such that the conditional probability distribution over the other agents' reported type profiles corresponding to q could not have been induced by this agent unilaterally deviating from truthtelling and reporting this type untruthfully. Thus, under any probability distribution q over the agents' type profiles, the mechanism designer can identify at least one agent-type the report of which is surely truthful. Such agent-type will be referred to as a nondeviator under q.

We show that Identifiability condition is generic if there are at least three agents and at least two agents have (weakly) less types than the type profiles of all other agents. Crémer–McLean condition is generic when no agent has more types than the type profiles of all other agents, so our mechanism exists generically.

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 [7,8] and McAfee and Reny [19]. So, it is natural to consider how this surplus can be allocated. We show that under our conditions there are no restrictions on this: the expected social surplus can be allocated across *agent-types* in any desired way.

---

[3] Ex-ante budget balance can be attained in the Crémer–McLean mechanism if the mechanism designer starts by paying each agent an amount equal to the latter's ex-ante expected transfer in the ensuing mechanism.

[4] To the best of our knowledge, this term was coined by d'Aspremont and Gérard-Varet [5]. The class of ex-ante socially rational decision rules includes ex-post efficient ones as a special case.

In a recent paper, Matsushima [18] has presented alternative sufficient conditions for budget-balanced, individually rational implementation of ex-ante socially rational decision rules. His conditions are strictly stronger than ours and, hence, are not necessary. Also, more restrictive dimensionality requirements have to be imposed for his conditions to hold and to be generic (for details, see Lemma 1 below and the discussion that follows it).

We also address the issue of ex-post budget-balanced implementation without the individual rationality requirement. We show that a weakening of the Identifiability condition, *Weak Identifiability*, is necessary and sufficient for any implementable decision rule to be implementable with budget balance, but without individual rationality. As the Identifiability Condition, Weak Identifiability 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 [4]. Naturally, our Weak Identifiability is equivalent to Condition C.

Lastly, the modeling 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 the other agents' types. Although there may be situations where this is not so, our approach is applicable in many economically important 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. Similarly, when there is an uncertainty about market conditions, a firm's information about future demand for its product is both its payoff-relevant type and the determinant of its beliefs about the demand for the competitors' products.

The rest of the paper is organized as follows. In Section 2 we develop the model. Section 3.1 introduces the concept of Identifiability. In Section 3.2 we establish our main results. Section 3.3 presents an example of our mechanism. Section 3.4 studies implementation without individual rationality requirement. All proofs are relegated to Appendix A.

2. The model

There are n agents in the economy. Agent i has privately known type which belongs to the type space $\Theta_i \equiv \{\theta_i^1, \dots, \theta_i^{m_i}\}$ of cardinality m_i, $2 \leqslant m_i < \infty$. A generic element of Θ_i will be denoted by θ_i or θ_i'. A state of the world is characterized by a type profile $\theta = (\theta_1, \dots, \theta_n)$. The set of type profiles is given by $\Theta \equiv \prod_{i=1,\dots,n} \Theta_i$, with cardinality $L \equiv \prod_{i=1,\dots,n} m_i$. 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 $\Theta_{-i} = \prod_{l \neq i} \Theta_l$, $L_{-i} = \prod_{l \neq i} m_l$, $\Theta_{-i-j} = \prod_{l \notin \{i,j\}} \Theta_l$, and $L_{-i-j} = \prod_{l \notin \{i,j\}} m_l$. A generic element of Θ_{-i-j} is denoted by θ_{-i-j}.

The (true) probability distribution of the agents' type profile θ is denoted by $p(\theta)$, with $p_i(\theta_i)$ and $p_{i,j}(\theta_i, \theta_j)$ denoting the corresponding marginal probability distribution of agent i's type and the marginal probability distribution of types of agents i and j, respectively. We assume that $p(\theta)$ is common knowledge. We also 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.[5] Further, let $p_{-i}(\theta_{-i}|\theta_i)$ $(p_j(\theta_j|\theta_i))$ denote the probability distribution

---

[5] This condition is clearly generic, and is employed only in the proof of Theorem 1 to establish that any allocation of surplus between agent-types is feasible. If this condition fails, then a straightforward but lengthy argument (which we omit for brevity) can be used to establish the existence of a partition of the type space Θ such that the surplus can be freely reallocated between the agent-types in each element of this partition. Detailed description of this argument is provided in Appendix B available at http://www.severinov.com/AppendixB_mechanisms.pdf.

of type profiles of agents other than *i* (agent *j*'s type) conditional on the type of agent *i*. We use a similar system of notation for other probability distributions over Θ that will be introduced below. The set of all probability distributions over Θ is denoted by $\mathcal{P}(\Theta)$.

A mechanism designer, who does not possess any private information, controls the set of public decisions *X*. Let *x* denote a generic 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.[6] A (social) decision rule $x(.)$ is a function mapping the type space Θ into the set of public decisions *X*.[7] Also, $t(.) = (t_1(.), \ldots, t_n(.))$ is a collection of transfer functions to all agents, where $t_i(.) \colon \Theta \mapsto \mathbf{R}$ is a transfer function to agent *i*. An allocation profile is a combination of a decision rule $x(.)$ with a collection of transfer functions $t(.)$.

By the Revelation Principle we can restrict the analysis to direct mechanisms in which the mechanism designer offers an allocation profile to the agents. If the agents, informed of their types, decide to participate in this mechanism, they report their types to the mechanism designer and the allocation corresponding to the reported type profile is implemented.[8]

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 $(x(.), t(.))$ is incentive compatible if the following *Interim Incentive Constraint* $IC_i(\theta_i, \theta_i')$ holds for all $i \in \{1, \ldots, n\}$ and $\theta_i, \theta_i' \in \Theta_i$:

$$
\sum_{\theta_{-i} \in \Theta_{-i}} \big(u_i(x(\theta_{-i}, \theta_i), (\theta_{-i}, \theta_i)) + t_i(\theta_{-i}, \theta_i) - u_i(x(\theta_{-i}, \theta_i'), (\theta_{-i}, \theta_i))
$$
$$
- t_i(\theta_{-i}, \theta_i') \big) \, p_{-i}(\theta_{-i}|\theta_i) \geqslant 0. \tag{1}
$$

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

Interim Individual Rationality (*IR*) requires the following $IR_i(\theta_i)$ constraint to hold for all $i \in \{1, \ldots, n\}$ and $\theta_i \in \Theta_i$:

$$
\sum_{\theta_{-i} \in \Theta_{-i}} (u_i(x(\theta_{-i}, \theta_i), (\theta_{-i}, \theta_i)) + t_i(\theta_{-i}, \theta_i)) \, p_{-i}(\theta_{-i}|\theta_i) \geqslant 0. \tag{2}
$$

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

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

---

[6] 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.

[7] 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.

[8] As mentioned above, Mezzetti [23,24] shows that, with interdependent values and independently distributed types, and without budget balance, a mechanism designer can attain efficiency by running a mechanism in which the agents, first, report their types, and then report their realized utilities after the allocation is made. We focus on mechanisms without ex-post utility reporting, and attain efficiency with budget balance and nonindependently distributed types. Also, we allow for private values environments where ex-post utility reporting has no advantages.

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* $\sum_{i=1}^{n} u_i(x, \theta)$. 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 $\max_{x \in X} \sum_{i=1}^{n} u_i(x, \theta) \geqslant 0$ for all $\theta \in \Theta$. Finally, *IR* and *BB* together imply the following *Ex-Ante Social Rationality* (*EASR*) condition:

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

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 $p(.)$ of the agents' type profiles is identifiable if, for any probability distribution $q(.) \in \mathcal{P}(\Theta)$, $q(.) \neq p(.)$, there is an agent i and her type θ_i', with $q_i(\theta_i') > 0$, such that for any collection of nonnegative coefficients $c_{\theta_i \theta_i'}, \theta_i, \theta_i' \in \Theta_i$,

$$q_{-i}(.|\theta_i') \neq \sum_{\theta_i \in \Theta_i} c_{\theta_i \theta_i'} p_{-i}(.|\theta_i). \tag{5}$$

Consider also the familiar condition of Crémer and McLean [8] 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 $\theta_i' \in \Theta_i$, $p_{-i}(.|\theta_i')$ cannot be expressed as a positive linear combination of $p_{-i}(.|\theta_i)$, $\theta_i \neq \theta_i'$, i.e. for any collection of nonnegative coefficients $c_{\theta_i \theta_i'}$, where $\theta_i, \theta_i' \in \Theta_i$,

$$p_{-i}(.|\theta_i') \neq \sum_{\theta_i \in \Theta_i \setminus \theta_i'} c_{\theta_i \theta_i'} p_{-i}(.|\theta_i).$$

In the next subsection we show that the Identifiability of the prior $p(\theta)$ together with Crémer–McLean condition are necessary and sufficient for *BB*, *IR*, efficient implementation. Although Crémer–McLean condition is well-understood, [9] Identifiability is a new condition introduced in this paper. So, before we exhibit and explain our results, let us explore this condition 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.

---

[9] See also McAfee and Reny [19] for an intuitive discussion of this condition.

Agent i's strategy s_i in a direct mechanism is a vector of size m_i^2 such that its entry $s_{\theta_i \theta_i'}$ denotes the probability with which agent i of type θ_i reports type θ_i'. Note that $s_{\theta_i \theta_i'} \in [0, 1]$ and $\sum_{\theta_i' \in \Theta_i} s_{\theta_i \theta_i'} = 1$ for all $\theta_i \in \Theta_i$. Let S_i be the set of all such strategies s_i. A truthful strategy s_i^* of agent i is such that $s_{\theta_i \theta_i} = 1$ and $s_{\theta_i \theta_i'} = 0$ for all $\theta_i, \theta_i' \in \Theta_i$ s.t. $\theta_i \neq \theta_i'$. A strategy profile $\mathbf{s} \equiv (s_1, \ldots, s_n)$ 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 $\mathbf{s} \equiv (s_1, \ldots, s_n)$ induces the probability distribution over the reported type profiles $\pi(.|\mathbf{s})$ if type profile $\theta' \in \Theta$ is reported with probability $\pi(\theta'|\mathbf{s})$ when the agents follow strategies $\mathbf{s} = (s_1, \ldots, s_n)$ and the types are drawn from the prior $p(.)$.

To compute $\pi(.|\mathbf{s})$, note that for any $\theta' = (\theta_1', \ldots, \theta_n')$,

$$\pi(\theta'|\mathbf{s}) = \sum_{(\theta_1, \ldots, \theta_n) \in \Theta} p(\theta_1, \ldots, \theta_n) \prod_{i=1}^{n} s_{\theta_i \theta_i'}.$$

It is natural to interpret the coefficients $c_{\theta_i \theta_i'}$ in the definition of Identifiability as stemming from i's reporting strategy i.e.

$$c_{\theta_i \theta_i'} = \frac{s_{\theta_i \theta_i'} p_i(\theta_i)}{\sum_{\theta_i'' \in \Theta_i} s_{\theta_i'' \theta_i'} p_i(\theta_i'')}.$$

Then the Identifiability requires that for each $q(.) \in \mathcal{P}(\Theta)$, $q(.) \neq p(.)$, there exists an agent i who does not have a strategy such that when i reports some θ_i' according to this strategy and the other agents report truthfully, the induced probability distribution of the other agents' type profiles coincides with $q_{-i}(.|\theta_i')$. That is, for all $s_i \in S_i$,

$$q_{-i}(.|\theta_i') \neq \frac{\sum_{\theta_i \in \Theta_i} s_{\theta_i \theta_i'} p_i(\theta_i) p_{-i}(.|\theta_i)}{\sum_{\theta_i \in \Theta_i} s_{\theta_i \theta_i'} p_i(\theta_i)}.$$

Thus, if $p(.)$ is identifiable and the agents' strategies induce probability distribution $q(.)$ over the reported type profiles, with $q(.) \neq p(.)$, then the mechanism designer can identify a nonempty set of *agent-types* such that the agents who report these types have not unilaterally deviated from truthtelling. This interpretation provides a rationale for the use of the term 'identifiability' in reference to this condition.

In comparison, Crémer–McLean condition for agent i says that i does not possess a nontruthful strategy which, in combination with truthtelling by other agents, induces a probability distribution over the reported type profiles, $q(.)$, the conditional of which on *any* type θ_i, $q_{-i}(.|\theta_i)$, coincides with $p_{-i}(.|\theta_i)$, the conditional probability distribution derived from the prior.

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 utility functions* (*quasilinear in transfers*) *only if the prior* $p(.)$ *is identifiable and Crémer–McLean condition holds for all agents.*

Remark. To make our result stronger, we prove that the conditions of the Theorem are necessary to attain ex-post efficiency, and not just ex-ante social rationality. It is worth noting that socially rational decision rules that are not efficient may also fail to be implementable if either Identifiability or Crémer–McLean condition fails to hold.[10]

To explain the intuition behind Theorem 1, it is useful to highlight the relationship between our results and those of Crémer and McLean [7,8]. Crémer and McLean show that the mechanism designer can implement an efficient decision rule and extract all surplus from the agents. Their mechanism relies on lotteries—systems of transfers between each agent and the mechanism designer that depend on the whole profile of the reported types. The lotteries are constructed so that the loss from a misrepresentation in a lottery for any agent-type always exceeds any potential gain from a better allocation $x(.)$. Crémer and McLean condition described in Definition 2 is necessary and sufficient for the existence of such lotteries.[11]

Importantly, in their mechanism, all transfers are made between each agent and the mechanism designer who acts as a budget-breaker, or residual claimant, for the lotteries offered to the agents. In contrast, in our ex-post budget-balanced framework all payments are made between the agents. This difference is substantial, as simply redistributing the transfers, which the mechanism designer exchanges with a particular agent in the Crèmer–McLean mechanism, to the other agents, while ensuring ex-post budget balance, would undermine the incentive compatibility of this mechanism.

For example, designating agent j, instead of the mechanism designer, to balance the budget of agent i's lottery 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, thereby causing i to make transfers to j. Thus, our mechanism has to overcome an additional incentive problem that arises from the budget-balance requirement.

Consider how the Identifiability condition allows us to solve this problem. Take any probability distribution over type profiles, $q(.)$, such that $q(.) \neq p(.)$. By Identifiability, there exists a nonempty set of agent-types $\mathcal{I}_q$ such that for any θ_i' in this set, agent i cannot induce the reported type profile of the other agents to be distributed according to $q_{-i}(.|\theta_i')$ by unilaterally deviating from truthtelling and reporting θ_i' untruthfully. (i.e. $q_{-i}(.|\theta_i') \neq \sum_{\theta_i \in \Theta_i} c_{\theta_i \theta_i'} p_{-i}(.|\theta_i)$ for any collection of coefficients $c_{\theta_i \theta_i'} \geqslant 0$). Thus, agent i's report of type θ_i' is surely truthful under $q(.)$, and so we refer to θ_i' as a nondeviator under $q(.)$.

Exploiting the nonemptiness of the set of nondeviators $\mathcal{I}_q$ under any $q(.)$, the mechanism designer can construct a budget-balanced system of transfers, $t(.)$, with negative expected value for every agent who unilaterally deviates from truthtelling (i.e. $\sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i') p_{-i}(\theta_{-i}|\theta_i) < 0$ for all $i = 1, \ldots, n$, and $\theta_i, \theta_i' \in \Theta_i$, $\theta_i' \neq \theta_i$), and zero expected value for every agent when

---

[10] This can be shown by slightly modifying the decision rule, $x^*(.)$, and the profile of the utility functions in the necessity part of the proof, and checking that its conclusion still holds.

[11] It is worth pointing out that Crémer and McLean [8] 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*.

everyone reports truthfully (i.e. $\sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i) p_{-i}(\theta_{-i}|\theta_i) = 0$ for all i and θ_i). This is the first step in the proof of sufficiency part of Theorem 1 (see Lemma A2 in Appendix A).

Particularly, the system of transfers $t(.)$ is constructed so that the expected transfer to agent i announcing type θ_i' is negative under any probability distribution of the other agents' reported type profile $q_{-i}(.|\theta_i')$ which i can induce by unilaterally deviating from truthtelling and reporting θ_i' untruthfully. To balance the budget, positive expected transfers under $q(.)$ are given to some nondeviator types under $q(.)$. Such agent-types exist by Identifiability. Since any agent-type who gets a positive expected transfer under $q(.)$ is a nondeviator, she does not have an incentive to deviate from truthtelling, because she cannot induce $q(.)$ unilaterally.

Crémer–McLean condition also plays a role in ensuring that the transfers $t(.)$ have the described properties. Precisely, it guarantees that, if agent i unilaterally deviates from truthtelling and reports some type θ_i' untruthfully, the probability distribution of the reported type profiles of the other agents will differ from $p_{-i}(.|\theta_i')$. Therefore, every unilateral deviation is detectable and distinguishable from truthtelling.

To the system of transfers $t(.)$, we add budget-balanced transfers $\tau(.)$ that compensate agents for the utility consequences of the public decision $x(.)$ and allocate the social surplus as desired by the mechanism designer. The system of transfers $\tau(.)$ is constructed in Lemma A3 in Appendix A. Finally, the incentive compatibility of the mechanism is ensured by scaling up the first system of transfers $t(.)$. This scaling up is done so that any gain in utility from a social decision or from transfers $\tau(.)$, which an agent can obtain by deviating from truthtelling, is dominated by a negative expected value of the scaled transfer $t(.)$ that she gets in this case.

The necessity part of Theorem 1 can be explained as follows. If Crémer–McLean condition fails, then the mechanism designer does not have the ability to detect some deviations from truthtelling, even if they take place. On the other hand, if Identifiability fails, then the designer lacks the ability to punish all potential deviators by large negative expected transfers without violating budget balance. Thus, in either case, there is no budget-balanced system of transfers with zero expected value for every agent-type when everyone reports truthfully, and with a negative expected value for any agent-type unilaterally deviating from truthtelling (see Step 3 in Lemma A2 in Appendix A). Consequently, the mechanism designer has to provide sufficient surplus to some agent-types to induce them to report truthfully, but this could be infeasible. Specifically, as we demonstrate in the proof, under certain profiles of the utility functions some agent-types can obtain very large utility gains by deviating, and to prevent them from doing so, the mechanism would have to provide them with net payoffs exceeding all the expected social surplus. In such cases budget-balanced implementation is impossible.

Let us now exhibit an important corollary of Theorem 1. With ex-post budget balance, all social surplus generated by the mechanism is allocated to the agents. Therefore, it is natural to inquire how this surplus can be distributed among the agents and their types. The following corollary shows that this can be done in any way desired by the mechanism designer.

Corollary 1. *Consider any ex-ante socially rational 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 $\sum_{i \in \{1,...,n\}} m_i$ nonnegative constants $v_i(\theta_i)$ satisfying*:

$$\sum_{i \in \{1,...,n\}} \sum_{\theta_i \in \Theta_i} v_i(\theta_i) p_i(\theta_i) = \sum_{i \in \{1,...,n\}} \sum_{\theta \in \Theta} u_i(x(\theta), \theta) p(\theta), \tag{6}$$

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 in this mechanism is equal to $v_i(\theta_i)$, i.e.

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

Corollary 1 directly follows from the intermediate steps in the proof of the sufficiency part of Theorem 1. The Corollary 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 the 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, our conditions are essential to guarantee the freedom in surplus allocation.

Next, we examine whether sufficiently many probability distributions are identifiable and satisfy Crémer–McLean condition for all agents. It is immediate to check that any independent distribution, $p(\theta) = \prod_{i=1,\dots,n} p_i(\theta_i)$, is identifiable for any $n \geqslant 2$. However, Crémer–McLean condition fails in this case. Further, only independent probability distributions are identifiable when $n = 2$. The proof of this assertion is straightforward, but computationally tedious.

Importantly, the following Theorem provides the dimensionality requirements under which the Identifiability condition holds generically.

Theorem 2 (*Genericity of Identifiability*). *Suppose that there are at least three agents ($n \geqslant 3$). Also, if $n = 3$, then at least one of the agents has at least three types. Then almost all probability distributions $p(.)$ are identifiable.*

The proof of Theorem 2 shows that, under these conditions, the set of probability distributions which are not identifiable has (Lebesgue) measure zero. The argument in the proof 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. [12]

It is well-known that Crémer–McLean condition for agent i holds generically when $m_i \leqslant \prod_{j \neq i} m_j$. Since a finite union of sets of measure zero has measure zero, and an intersection of a finite number of open and dense sets is open and dense, we obtain that the two conditions of Theorem 1 hold generically when $m_i \leqslant \prod_{j \neq i} m_j$ for all $i \in \{1, \dots, n\}$.

We conclude this subsection with the following Lemma which connects the Identifiability condition with the familiar notion of linear independence, and allows us to highlight the distinction between this paper and Matsushima [18].

Lemma 1. *The probability distribution $p(.)$ is identifiable if there is an agent i such that for each $\theta_i \in \Theta_i$ there exists an agent j and $\theta_j \in \Theta_j$ such that $m_i + m_j - 1$ vectors of conditional probability distributions $p_{-i-j}(.|\theta_i', \theta_j)$, $p_{-i-j}(.|\theta_i, \theta_j')$, $\theta_i' \in \Theta_i$, $\theta_j' \in \Theta_j$, $\theta_j' \neq \theta_j$, are linearly independent.*

---

[12] In our model, as in most related literature, an agent's type is equivalent to her preference parameter and uniquely determines her beliefs about the types of others. Neeman [26] and Heifetz and Neeman [11] show that standard genericity results no longer hold, if an agent with the same preference parameter can have different beliefs.

Matsushima [18] establishes that a stronger version of the condition in Lemma 1 (his Condition 1) together with a strengthening of Crémer–McLean condition (his Condition 2) are sufficient for budget-balanced, individually rational, efficient implementation. So, both our conditions are strictly weaker than their counterparts in Matsushima [18]. [13]

3.3. Example

To illustrate our results, let us explicitly construct the mechanism in the special case with three agents, each of whom has two types i.e., $\Theta_i = \{\theta_i^1, \theta_i^2\}$ for $i \in \{1, 2, 3\}$. Let $p_{k_1 k_2 k_3}$ be the probability of the type profile $(\theta_1^{k_1}, \theta_2^{k_2}, \theta_3^{k_3})$, where $k_1, k_2, k_3 \in \{1, 2\}$. The following corollary relies directly on the intermediate results established in the proof of Theorem 1:

Corollary 2. *Suppose that there are three agents, each with two types. If $p(.)$ is identifiable and Crémer–McLean condition holds for all three agents, then the agents and their types can be relabeled so that the following inequalities hold:*

$$p_{111}p_{122} - p_{112}p_{121} < 0, \tag{7}$$
$$p_{111}p_{212} - p_{112}p_{211} > 0. \tag{8}$$

A careful reading of the proof of this corollary in Appendix A reveals that Crémer–McLean condition implies that the expressions on the left-hand sides of (7) and (8) are nonzero, while Identifiability of $p(.)$ implies the sign restriction specified by (7) and (8).

Importantly, inequalities (7) and (8) imply that agent-types θ_1^2 and θ_2^2 cannot induce the same probability distribution of reported type profiles by unilateral deviations from truthtelling. To see this, consider a geometric illustration in Fig. 1. Inequalities (7) and (8) imply that vectors (p_{211}, p_{212}) and (p_{121}, p_{122}) lie on the different sides of the vector (p_{111}, p_{112}) in the north-east quadrant. Therefore, if agent-type θ_1^2 is the only type who deviates from truthtelling and reports θ_1^1 with probability $\alpha \in (0, 1]$, then the vector of probabilities of agent 3's reported types, $r^3(\alpha)$, conditional on agents 1 and 2 reporting a pair of types (θ_1^1, θ_2^1), will be a convex combination of vectors (p_{211}, p_{212}) and (p_{111}, p_{112}), and so will lie strictly between them.

On the other hand, if agent-type θ_2^2 is the only one who deviates from truthtelling and reports θ_2^1 with probability $\beta \in (0, 1]$, then the vector of probabilities of agent 3's reported types, $r^3(\beta)$, conditional on agents 1 and 2 reporting a pair of types (θ_1^1, θ_2^1), will be a convex combination of vectors (p_{121}, p_{122}) and (p_{111}, p_{112}) and will lie strictly between them. Hence, $r^3(\alpha)$ is not equal to $r^3(\beta)$ for any α and $\beta \in (0, 1]$. So, under any probability distribution of reported type profiles, either agent 1's report of type θ_1^1, or agent 2's report of type θ_2^1 must be truthful, i.e. either agent-type θ_1^1 or θ_2^1 is a nondeviator.

We will use this property to construct a budget-balanced system of transfers $t_i^{k_1 k_2 k_3}$, $i \in \{1, 2, 3\}$ and $k_i \in \{1, 2\}$, such that under any probability distribution of reported type profiles different

---

[13] Specifically, his Condition 1 requires that for some pair of agents i and j and *all pairs of their types θ_i and θ_j*, $m_i + m_j - 1$ vectors of conditional probability distributions $p_{-i-j}(.|\theta_i', \theta_j)$, $p_{-i-j}(.|\theta_i, \theta_j')$, where $\theta_i' \in \Theta_i$ and $\theta_j' \in \Theta_j$, $\theta_j' \neq \theta_j$, are linearly independent. This condition can hold only if $\prod_{l \neq i, j} m_l \geqslant m_i + m_j - 1$, and fails if there are three agents with equal numbers of types. In contrast, our Identifiability condition is always generic when $n \geqslant 3$ except if there are three agents with two types each. In the latter case, Identifiability holds for most probability distributions, in fact, 87.5% of them under Lebesgue measure, as follows from Corollary 2 in Section 3.3. Also, our regularity assumption requiring that $p_{i,j}(\theta_i, \theta_j) > 0$ for all $i, j \in \{1, \ldots, n\}$ and $\theta_i \in \Theta_i$, $\theta_j \in \Theta_j$ is strictly weaker than Matsushima's full support assumption that $p(\theta) > 0$ for all $\theta \in \Theta$.

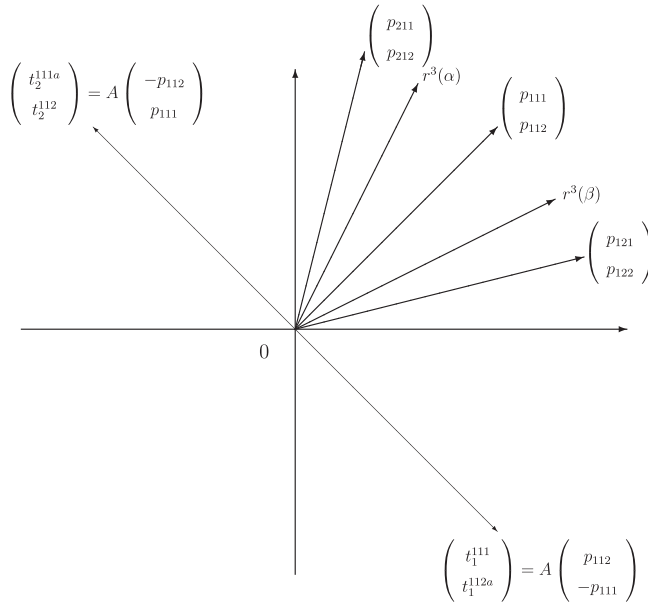Fig. 1. Example: budget-balanced transfers for agent-types θ_1^1 and θ_2^1 satisfying $IR^0(\theta_1^1)$, $IR^0(\theta_2^1)$, $IC^0(\theta_1^2, \theta_1^1)$ and $IC^0(\theta_2^2, \theta_2^1)$.

from the prior $p(.)$, any agent-type who could induce this probability distribution by a unilateral deviation from truthtelling gets a negative expected payoff, while the nondeviator under this probability distribution, which is either θ_1^1 or θ_2^1, gets a positive expected payoff. Also, the expected transfer to every agent-type under $p(.)$ should be equal to zero.

Referring to the proof of Theorem 1, the system of transfers $t_i^{k_1 k_2 k_3}$ will be constructed to satisfy conditions (12)–(14), providing incentives for truth-telling. That is, the system $t_i^{k_1 k_2 k_3}$ will satisfy constraints $IC^0(\theta_i^k, \theta_i^{k'})$ and $IR^0(\theta_i^k)$ for $i \in \{1, 2, 3\}$ and $k, k' \in \{1, 2\}, k \neq k'$, where $IC^0(\theta_1^k, \theta_1^{k'})$ is given by $\sum_{k_2, k_3 \in \{1,2\}} t_1^{k_1' k_2 k_3} p_{k_1 k_2 k_3} < 0$, and $IR^0(\theta_1^{k_1})$ is given by $\sum_{k_2, k_3 \in \{1,2\}} t_1^{k_1 k_2 k_3} p_{k_1 k_2 k_3} = 0$. $IC^0(\theta_i^k, \theta_i^{k'})$ and $IR^0(\theta_i^k)$ for $i \in \{2, 3\}$ are defined similarly.

Recall that constructing transfers satisfying these conditions is the first step in the proof of sufficiency in Theorem 1 (see also the discussion in Section 3.2). [14]

We design this system in two steps. The transfers $t_1^{112}, t_1^{122}, t_2^{111}$ and t_2^{211} will consist of two parts—part (a) and part (b)—defined separately. Specifically, we will have: $t_1^{112} = t_1^{112a} + t_1^{112b}$, $t_1^{122} = t_1^{122a} + t_1^{122b}$, $t_2^{111} = t_2^{111a} + t_2^{111b}$, and $t_2^{211} = t_2^{211a} + t_2^{211b}$. The other transfers will be defined in a single step.

Step 1: In this key step, we specify a set of budget-balanced transfers between agents 1 and 2, who report types θ_1^1 and θ_2^1, respectively. These transfers guarantee that agent-type θ_1^2 (θ_2^2) gets a

---

[14] Lemma A2 in Appendix A shows that Identifiability and Crémer–McLean conditions are necessary and sufficient for the existence of such transfers. Thus, by constructing such transfers, we also establish that Corollary 2 holds in the opposite direction, i.e. inequalities (7) and (8) imply Identifiability and Crémer–McLean conditions.

negative expected payoff when she misreports her type as θ_1^1 (θ_2^1). Particularly, let A be a positive constant which will be specified later and set:

$$t_1^{111} = -t_2^{111a} = Ap_{112}, \quad t_1^{112a} = -t_2^{112} = -Ap_{111}.$$

Then, by (7) and (8), we have

$$IC^0(\theta_1^2, \theta_1^1): \quad p_{211}t_1^{111} + p_{212}t_1^{112a} = A(p_{211}p_{112} - p_{212}p_{111}) < 0,$$
$$IC^0(\theta_2^2, \theta_2^1): \quad p_{121}t_2^{111a} + p_{122}t_2^{112} = A(-p_{121}p_{112} + p_{122}p_{111}) < 0.$$

The individual rationality constraints $IR^0(\theta_1^1)$ and $IR^0(\theta_2^1)$ hold, since we have

$$p_{111}t_1^{111} + p_{112}t_1^{112a} = -\left(p_{111}t_2^{111a} + p_{112}t_2^{112}\right) = Ap_{111}p_{112} - Ap_{112}p_{111} = 0.$$

In defining these transfers, we have exploited the fact that agent-types θ_1^2 and θ_2^2 cannot induce the same probability distribution of reported type profiles by unilaterally deviating from truthtelling and reporting types θ_1^1 and θ_2^1, respectively. As noted above, this property is implied by inequalities (7) and (8), which geometrically say that vectors (p_{211}, p_{212}) and (p_{121}, p_{122}) lie on the different sides of the vector (p_{111}, p_{112}) (see Fig. 1).

By choosing (t_1^{111}, t_1^{112a}) to be orthogonal to (p_{111}, p_{112}) and to form an obtuse angle with (p_{211}, p_{212}), we ensure that (t_1^{111}, t_1^{112a}) has zero inner product with the former and a negative inner product with the latter. So both $IR^0(\theta_1^1)$ and $IC^0(\theta_1^2, \theta_1^1)$ hold.

Further, (t_2^{111a}, t_2^{112}), being the negative of (t_1^{111}, t_1^{112a}), is also orthogonal to (p_{111}, p_{112}) (so $IR^0(\theta_2^1)$ holds), and forms an obtuse angle with (p_{121}, p_{122}) (so $IC^0(\theta_2^2, \theta_2^1)$ holds).

Thus, the transfers constructed in Step 1 punish both types θ_1^2 or θ_2^2 for a deviation, and allocate proceeds from these punishments to types θ_2^1 and θ_1^1, respectively.

Step 2: In this step, we construct pairs of transfers between four different pairs of agent-types. Each pair of agent-types includes either type θ_1^1 or type θ_2^1, who play the role of residual claimants balancing the budget. Specifically, agent-type θ_1^1 (θ_2^1) will play the role of a residual claimant vis-a-vis agent-types θ_2^2 and θ_3^2 (θ_1^2 and θ_3^2). Exploiting inequalities (7) or (8), we construct these transfers so that they satisfy IC^0 constraints of *the types who could imitate the latter types* i.e., of types $\theta_2^1, \theta_3^1, \theta_1^1$ and θ_3^2. We also make sure that IR^0 constraints of all involved types hold. Although the transfers received by the "residual claimants" θ_1^1 and θ_2^1 may violate incentive constraints $IC^0(\theta_1^2, \theta_1^1)$ and $IC^0(\theta_2^2, \theta_2^1)$, we will take care of these constraints later by choosing a sufficiently large constant A introduced in Step 1. All transfers constructed in this way are described in Table 1.

Table 1
Transfers of Step 2

Transfers between:	Transfers	IC^0 constr.	IR^0 constraints
$\theta_2^2 \leftrightarrow \theta_1^1$	$t_1^{121} = -t_2^{121} = -p_{122}, t_1^{122a} = -t_2^{122} = p_{121}$	$IC^0(\theta_2^1, \theta_2^2)$	$IR^0(\theta_2^2), IR^0(\theta_1^1)$
$\theta_3^2 \leftrightarrow \theta_1^1$	$t_1^{112b} = -t_3^{112} = -p_{122}, t_1^{122b} = -t_3^{122} = p_{112}$	$IC^0(\theta_1^1, \theta_3^2)$	$IR^0(\theta_3^2), IR^0(\theta_1^1)$
$\theta_1^2 \leftrightarrow \theta_2^1$	$t_1^{211} = -t_2^{211a} = -p_{212}, t_1^{212} = -t_2^{212} = p_{211}$	$IC^0(\theta_1^1, \theta_1^2)$	$IR^0(\theta_1^2), IR^0(\theta_2^1)$
$\theta_3^1 \leftrightarrow \theta_2^1$	$t_3^{111} = -t_2^{111b} = p_{211}, t_3^{211} = -t_2^{211b} = -p_{111}$	$IC^0(\theta_3^2, \theta_3^1)$	$IR^0(\theta_3^1), IR^0(\theta_2^1)$

Table 2
Aggregate transfers

	t_i^{111}	t_i^{112}	t_i^{121}	t_i^{122}	t_i^{211}	t_i^{212}	t_i^{221}	t_i^{222}
$i=1$:	$Ap_{112},$	$-Ap_{111}-p_{122},$	$-p_{122},$	$p_{121}+p_{112},$	$-p_{212},$	$p_{211},$	$0,$	0
$i=2$:	$-Ap_{112}-p_{211},$	$Ap_{111},$	$p_{122},$	$-p_{121},$	$p_{212}+p_{111},$	$-p_{211},$	$0,$	0
$i=3$:	$p_{211},$	$p_{122},$	$0,$	$-p_{112},$	$-p_{111},$	$0,$	$0,$	0

Consider the first line of Table 1. It describes the transfers between agents 1 and 2 when they announce types θ_1^1 and θ_2^2, respectively. Incentive constraint $IC^0(\theta_2^1, \theta_2^2)$ holds because by (7) we have

$$p_{111}t_2^{121} + p_{112}t_2^{122} = p_{111}p_{122} - p_{112}p_{121} < 0.$$

$IR^0(\theta_1^1)$ and $IR^0(\theta_2^2)$ hold, since we have

$$p_{121}t_1^{121} + p_{122}t_1^{122a} = -\left(p_{121}t_2^{121} + p_{122}t_2^{122}\right) = -p_{121}p_{122} + p_{122}p_{121} = 0.$$

The other lines of Table 1 describe the transfers between the other pairs of types listed in the first column.

Finally, Table 2 combines Steps 1 and 2 and describes the aggregate transfers of all types.

By construction, the system of transfers in Table 2 is budget-balanced and satisfies $IR^0(\theta_i^k)$ and $IC^0(\theta_i^k, \theta_i^{k'})$ for all $i \in \{1, 2, 3\}$, $k, k' \in \{1, 2\}$, except for possibly $IC^0(\theta_1^2, \theta_1^1)$ and $IC^0(\theta_2^2, \theta_2^1)$ which could fail because of additional transfers to agent-types θ_1^1 and θ_2^1 in Step 2 (see Table 1). To ensure that the latter constraints hold, we need to set constant A to be sufficiently large. Specifically, using (7) and (8) it is easy to check, that it is sufficient to set:

$$A > \max\left\{ \frac{|p_{112}p_{222} + p_{121}p_{222} - p_{122}p_{212} - p_{122}p_{221}|}{p_{111}p_{212} - p_{112}p_{211}}, \frac{|p_{111}p_{221} + p_{212}p_{221} - p_{211}p_{122} - p_{211}p_{222}|}{p_{112}p_{121} - p_{111}p_{122}} \right\}.$$

Finally, let us bring the utility terms $u_i(x, \theta)$ into the picture and show how one can implement an arbitrary *EASR* decision rule $x(.)$ and allocate the social surplus. By Corollary 1 (see also Lemma A3 in the proof of Theorem 1), the mechanism designer can allocate the social surplus generated by $x(.)$ in any desired way. Below, we will demonstrate how she can allocate all expected social surplus to agent-type θ_1^1 and zero surplus to any other agent-type. Using a similar mechanism, the designer can allocate all surplus to another type, and then take a convex combination of such mechanisms to achieve any desired allocation of surplus. To complete these steps, we need to introduce another system of transfers $\tau_i^{k_1 k_2 k_3}$, $i \in \{1, 2, 3\}$, $k_i \in \{1, 2\}$. This system is a particular case of transfers $\tau(\theta)$ characterized in Lemma A3, and will be used in our mechanism in combination with transfers $t_i^{k_1 k_2 k_3}$.

Abbreviating the notation, let $\bar{u}_i(k_1 k_2 k_3) = u_i(x(\theta_1^{k_1}, \theta_2^{k_2}, \theta_3^{k_3}), \theta_1^{k_1}, \theta_2^{k_2}, \theta_3^{k_3})$. Then, set:

$$\tau_1^{1k_2k_3} = -\tau_2^{1k_2k_3} - \tau_3^{1k_2k_3}, \quad \tau_1^{2k_2k_3} = -\bar{u}_1(2k_2k_3);$$

$$\tau_2^{1k_2k_3} = -\bar{u}_2(1k_2k_3) - \frac{p_1(\theta_1^2|\theta_2^{k_2})}{p_1(\theta_1^1|\theta_2^{k_2})}\bar{u}_2(2k_2k_3), \quad \tau_2^{2k_2k_3} = 0;$$

$$\tau_3^{1k_2k_3} = -\bar{u}_3(1k_2k_3) - \frac{p_1(\theta_1^2|\theta_3^{k_3})}{p_1(\theta_1^1|\theta_3^{k_3})}(\bar{u}_1(2k_2k_3) + \bar{u}_3(2k_2k_3)), \quad \tau_3^{2k_2k_3} = \bar{u}_1(2k_2k_3).$$

The system $\tau_i^{k_1k_2k_3}$ requires all types of agents 2 and 3 to transfer their expected utility from the decision rule $x(.)$ to agent-type θ_1^1. Agent 1 of type θ_1^2 transfers her expected utility to type θ_1^1 of herself via agent 3. So all agent-types, except θ_1^1, get zero expected surplus. Further, let

$$
D_\tau^1 = \max_{k_1,k_1' \in \{1,2\}, k_1' \neq k_1} \sum_{k_2,k_3 \in \{1,2\}} \left(\tau_1^{k_1'k_2k_3} + u_1(x(\theta_1^{k_1'}, \theta_2^{k_2}, \theta_3^{k_3}), \theta_1^{k_1}, \theta_2^{k_2}, \theta_3^{k_3}) \right.
$$
$$
\left. - \tau_1^{k_1k_2k_3} - \bar{u}_i(k_1k_2k_3) \right) p_{k_1k_2k_3}.
$$

Intuitively, D_τ^1 is the maximal expected utility gain which some type of agent 1 could obtain by unilaterally deviating from truthtelling in a direct mechanism with decision rule $x(.)$ and transfers $\tau_i^{k_1k_2k_3}$, normalized by the probability of this type. [15] Define D_τ^2 and D_τ^3 similarly for agents 2 and 3, and let $\bar{D}_\tau = \max\left\{ D_\tau^1, D_\tau^2, D_\tau^3 \right\}$. Next, let

$$
\bar{Q}_1 = \max_{k_1,k_1' \in \{1,2\}, k_1 \neq k_1'} \sum_{k_2,k_3 \in \{1,2\}} t_1^{k_1'k_2k_3} p_{k_1k_2k_3}.
$$

Define $\bar{Q}_2$ and $\bar{Q}_3$ similarly for agents 2 and 3. Note that $\bar{Q}_i < 0$ for $i \in \{1, 2, 3\}$ by construction. So, $\bar{Q} \equiv \max\{\bar{Q}_1, \bar{Q}_2, \bar{Q}_3\} < 0$. Finally, choose a constant b to satisfy $b \times \bar{Q} + \bar{D}_\tau < 0$, and define a new system of transfers $\check{t}_i(.)$ so that $\check{t}_i(\theta_1^{k_1}\theta_2^{k_2}\theta_3^{k_3}) = bt_i^{k_1k_2k_3} + \tau_i^{k_1k_2k_3}$. Then the direct mechanism $(x(.), \check{t}_i(.))$ is budget-balanced, individually rational, and incentive compatible. The latter property holds because $b \times \bar{Q} + \bar{D}_\tau < 0$. This mechanism allocates all expected social surplus to agent-type θ_1^1.

3.4. Mechanisms without individual rationality

As pointed out in the Introduction, several authors have explored the issue of implementation via ex-post budget-balanced mechanisms without imposing individual rationality. d'Aspremont, Crémer and Gérard-Varet [4] provide a necessary and sufficient condition for such implementation and show that this condition is strictly weaker than the ones studied by the other authors. Their Lemma 1 says that any Bayesian implementable allocation profile can 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 $t(\theta) \equiv (t_1(\theta), \dots, t_n(\theta))$ such that:

(i) $\sum_{i \in \{1,\dots,n\}} t_i(\theta) = R(\theta)$ *for all* $\theta \in \Theta$;

(ii) $\sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i) p_{-i}(\theta_{-i}|\theta_i) \geqslant \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i') p_{-i}(\theta_{-i}|\theta_i)$ *for all* $i \in \{1, \dots, n\}$, *and* $\theta_i, \theta_i' \in \Theta_i$.

Using the approach developed in the previous sections, we can provide an alternative necessary and sufficient condition for budget-balanced implementation without individual rationality. Our condition, Weak Identifiability, depends only on the properties of the probability distribution $p(.)$ and, as we show below, is obtained by weakening the Identifiability condition. Recall that, according to Definition 3, $\pi(.|s_1, \dots, s_n)$ stands for the probability distribution of type profiles

---

[15] Precisely, D_τ^1 is equal to the expected utility gain of type θ_i^k multiplied by $p_1(\theta_1^{k_1})$.

induced by the strategy profile $(s_1, \ldots, s_n)$. Then we have:

Definition 4 (*Weak Identifiability*). The prior $p(.)$ is weakly identifiable if there is no profile of agents' strategies $(s_1, \ldots, s_n) \in \prod_{i \in \{1, \ldots, n\}} S_i$ such that

$$\pi(.|s_1, s^*_{-n}) = \cdots = \pi(.|s_n, s^*_{-n}) \neq p(.) \qquad (9)$$

Importantly, we have:

Lemma 2. *Condition C holds if and only if Weak Identifiability condition holds.*

The intuition for this result is similar to that for Theorem 1. When Weak Identifiability holds, then for any probability distribution of reported type profiles different from the prior, there is a nonempty set of agents who could not have induced this probability distribution by a unilateral deviation from truthtelling. So, the mechanism designer can ensure the incentive compatibility of the mechanism by constructing a budget-balanced system of transfers such that, under any probability distribution of reported type profiles $q(.)$, $q(.) \neq p(.)$, the expected transfer to any agent-type who could (could not) have induced $q(.)$ by a unilateral deviation from truthtelling is less (greater) than this type's expected transfer under $p(.)$. So, each type gets a higher payoff by reporting truthfully than by misrepresenting itself. In contrast with the transfers constructed in Section 3.2, here we no longer require that each agent-type get a nonnegative payoff under $p(.)$, since individual rationality does not have to hold.

To confirm that Weak Identifiability is, indeed, a weakening of the Identifiability condition, we provide an alternative formulation of Identifiability in the following Lemma.

Lemma 3. *The probability distribution $p(.)$ is identifiable if and only if there do not exist a profile of agents' strategies $(s_1, \ldots, s_n) \in \prod_{i \in \{1, \ldots, n\}} S_i$ and a collection of functions $b_i(.) : \Theta_i \mapsto \mathbf{R}_+$ for $i = 1, \ldots, n$, such that for all $\theta' \equiv (\theta'_1, \ldots, \theta'_n) \in \Theta$ we have:*

$$\pi(\theta'|s_1, s^*_{-1}) + b_1(\theta'_1)p(\theta') = \cdots = \pi(\theta'|s_n, s^*_{-n}) + b_n(\theta'_n)p(\theta') = q(\theta'), \qquad (10)$$

where $q(.) : \Theta \mapsto \mathbf{R}_+$ is such that $q(.) \neq \alpha p(.)$ for all $\alpha \geqslant 0$.

Observe that condition (9) in the definition of Weak Identifiability is obtained from (10) in Lemma 3 by setting $b_i(\theta'_i) = 0$ for all $i \in \{1, \ldots, n\}$ and $\theta'_i \in \Theta_i$. So Weak Identifiability is implied by Identifiability.

4. Conclusions

In this paper we have characterized necessary and sufficient conditions for ex-post budget-balanced, interim individually rational Bayesian implementation of ex-ante socially rational and efficient decision rules. These conditions are the well-known Crémer–McLean condition and Identifiability condition introduced here. We have provided an intuitive explanation of our mechanism and have shown that the social surplus can be distributed across agent-types in any desirable way.

It is worth noting that, along certain important lines, our results cannot be extended further. In particular, only ex-ante socially rational decision rules can be implemented with budget balance and individual rationality. It is also easy to show that interim individual rationality cannot be

strengthened by replacing it with ex-post individual rationality without sacrificing the scope of implementability, even if we only insist on ex-ante budget balance. [16]

Our results can be applied to study the problem of mechanism design by an informed principal. In particular, they suggest that, also generically, efficiency is attainable in an informed principal framework, and that the informed principal will be able to extract all surplus from the other agents. We are pursuing this line of inquiry in related work.

Acknowledgments

Appendix A

In the proofs of Theorem 1 and Lemma 1 we will use the following result:

Lemma A1. *Suppose that $p(.)$ is not identifiable, i.e. for some $q(.) \in \mathcal{P}(\Theta)$, $q(.) \neq p(.)$, there exists a collection of nonnegative coefficients $\{c_{\theta_i \theta_i'} | i = 1, \ldots, n, \theta_i, \theta_i' \in \Theta_i\}$ s.t. $q_{-i}(.|\theta_i') = \sum_{\theta_i \in \Theta_i} c_{\theta_i \theta_i'} p_{-i}(.|\theta_i)$ for all $i \in \{1, \ldots, n\}$ and $\theta_i' \in \Theta_i$ with $q_i(\theta_i') > 0$. Also, suppose that for some $j \in \{1, \ldots, n\}$, we have $c_{\theta_j \theta_j'} = 0$ for all $\theta_j, \theta_j' \in \Theta_j$ s.t. $q_j(\theta_j') > 0$ and $\theta_j \neq \theta_j'$. Then for all $\theta_i' \in \Theta_i$, with $i \neq j$ and $q_i(\theta_i') > 0$, there exists $\theta_i \in \Theta_i$ such that $\theta_i \neq \theta_i'$ and $c_{\theta_i \theta_i'} > 0$.*

Proof of Lemma A1. The proof is by contradiction. So suppose that: (i) there exists $j \in \{1, \ldots, n\}$ s.t. $c_{\theta_j \theta_j'} = 0$ for all $\theta_j, \theta_j' \in \Theta_j$, $\theta_j \neq \theta_j'$; (ii) there exist $h \in \{1, \ldots, n\}$, $h \neq j$, and $\theta_h' \in \Theta_h$ s.t. $q_h(\theta_h') > 0$ and $c_{\theta_h \theta_h'} = 0$ for all $\theta_h \in \Theta_h \backslash \theta_h'$. Let us show that (i) and (ii) together imply that $q(.) = p(.)$.

First, supposition (ii) implies that $q_{-h}(.|\theta_h') = p_{-h}(.|\theta_h')$. This, in combination with our regularity assumption that $p_{jh}(\theta_j, \theta_h') > 0$ for all $\theta_j \in \Theta_j$, implies that $q_j(\theta_j) > 0$ for all $\theta_j \in \Theta_j$.

Next, fix some $\theta_j' \in \Theta_j$. Since $q_j(\theta_j') > 0$, supposition (i) yields $q_{-j}(.|\theta_j') = p_{-j}(.|\theta_j')$. Further, by our regularity assumption $p_{hj}(\theta_h', \theta_j') > 0$, and so there exists $\theta_{-j-h} \in \Theta_{-j-h}$ s.t. $p(\theta_{-j-h}, \theta_j', \theta_h') > 0$. Hence,

$$q(\theta_{-j-h}, \theta_j', \theta_h') = \frac{q_j(\theta_j')}{p_j(\theta_j')} p(\theta_{-j-h}, \theta_j', \theta_h') = \frac{q_h(\theta_h')}{p_h(\theta_h')} p(\theta_{-j-h}, \theta_j', \theta_h') > 0. \qquad (11)$$

Eq. (11) implies that $\frac{q_j(\theta'_j)}{p_j(\theta'_j)} = \frac{q_h(\theta'_h)}{p_h(\theta'_h)}$. Note that the right-hand side of this equality is independent of θ'_j and recall that θ'_j was chosen arbitrarily. So, $\frac{q_j(\theta'_j)}{p_j(\theta'_j)}$ is constant in $\theta'_j \in \Theta_j$. Combining this with the fact that $q_j(\theta'_j) > 0$ for all $\theta'_j \in \Theta_j$ and $\sum_{\theta'_j \in \Theta_j} p_j(\theta'_j) = \sum_{\theta'_j \in \Theta_j} q_j(\theta'_j) = 1$, we obtain that $q_j(\theta'_j) = p_j(\theta'_j)$ for all $\theta'_j \in \Theta_j$. Since we also have $q_{-j}(.|\theta'_j) = p_{-j}(.|\theta'_j)$ for all $\theta'_j \in \Theta_j$, it follows that $q(.) = p(.)$. Contradiction. $\square$

Proof of Theorem 1. *Sufficiency*: The proof proceeds as follows. First, in Lemma A2 we show that Identifiability and Crémer-McLean conditions hold if and only if there exists a budget-balanced system of transfers, $t(\theta)$, with a negative (zero) expected value for any agent-type who misrepresents her type (reports her type truthfully) in a direct mechanism when all other agents report truthfully.

Next, we show that it is possible to construct another system of transfers, $\tau(\theta)$, to allocate the surplus in the mechanism in any desirable way (Lemma A3). Finally, we will use the aggregate of these two systems of transfers to construct a mechanism implementing an arbitrary *EASR* decision rule $x(.)$.

Let us first introduce some vector notation. First, for a given system of transfers $t(\theta) = (t_1(\theta), \ldots, t_n(\theta))$, let $\mathbf{t}_i$ be a vector of size L, the entries of which are equal to the elements of $t_i(\theta)$ ordered in the natural order of type profiles. [17] Also, let $\mathbf{t}$ be a vector of size nL formed by stacking vectors $\mathbf{t}_1, \ldots, \mathbf{t}_n$ together.

Further, let $\mathbf{e}_\theta$ be a vector in R^{nL}, each entry of which corresponds to some agent $i \in \{1, \ldots, n\}$ and some type profile $\theta \in \Theta$ (ordered in the natural order of agents and type profiles) with the n entries corresponding to state θ being equal to 1 and all other entries being equal to zero. Also, for $i \in \{1, \ldots, n\}$ and $\theta_i, \theta'_i \in \Theta_i$, let $\mathbf{p}_{\theta_i \theta'_i}$ be a vector in R^{nL} s.t. its entry corresponding to agent i and type profile (θ'_i, θ_{-i}) is equal to $p_{-i}(\theta_{-i}|\theta_i)$ for all $\theta_{-i} \in \Theta_{-i}$, and all other entries are zero. Thus, a vector $\mathbf{p}_{\theta_i \theta'_i}$ has at most L_{-i} nonzero entries. [18]

Let Ω be a linear subspace of $\mathbf{R}^{nL}$ spanned by the collection of vectors $\{\mathbf{e}_\theta, \mathbf{p}_{\theta_i \theta_i}|\theta \in \Theta, i = 1, \ldots, n, \theta_i \in \Theta_i\}$, i.e. Ω is a set of all linear combinations of the vectors in this collection. Also, let S be a set of all convex combinations (convex hull) of vectors $\mathbf{p}_{\theta_i \theta'_i}$ for all $i = 1, \ldots, n$ and all $\theta_i, \theta'_i \in \Theta_i$ s.t. $\theta'_i \neq \theta_i$. The following Lemma provides a key step for the proof of sufficiency.

---

[17] Recall that $\Theta_i \equiv \{\theta_i^1, \ldots, \theta_i^{m_i}\}$, and so the natural order of agent i's types is given by the types' superscripts.

[18] As an illustration, consider the case of $n = 3$ and $\Theta_i = \{\theta_i^1, \theta_i^2\}$ for $i = 1, 2, 3$. Then, letting $\mathbf{0}_k$ be a vector consisting of k zeros, we have

$$\mathbf{t}_i = (t_i(\theta_1^1, \theta_2^1, \theta_3^1), t_i(\theta_1^1, \theta_2^1, \theta_3^2), t_i(\theta_1^1, \theta_2^2, \theta_3^1), t_i(\theta_1^1, \theta_2^2, \theta_3^2), t_i(\theta_1^2, \theta_2^1, \theta_3^1), t_i(\theta_1^2, \theta_2^1, \theta_3^2),$$
$$t_i(\theta_1^2, \theta_2^2, \theta_3^1), t_i(\theta_1^2, \theta_2^2, \theta_3^2)),$$

$$\mathbf{e}_{(\theta_1^1, \theta_2^2, \theta_3^2)} = (0, 0, 0, 1, \mathbf{0}_7, 1, \mathbf{0}_7, 1, 0, 0, 0, 0),$$

$$\mathbf{p}_{\theta_1^1 \theta_1^2} = (0, 0, 0, 0, p_{-1}(\theta_2^1, \theta_3^1|\theta_1^1), p_{-1}(\theta_2^1, \theta_3^2|\theta_1^1), p_{-1}(\theta_2^2, \theta_3^1|\theta_1^1), p_{-1}(\theta_2^2, \theta_3^2|\theta_1^1), \mathbf{0}_{16}),$$

$$\mathbf{p}_{\theta_1^1 \theta_1^1} = (p_{-1}(\theta_2^1, \theta_3^1|\theta_1^1), p_{-1}(\theta_2^1, \theta_3^2|\theta_1^1), p_{-1}(\theta_2^2, \theta_3^1|\theta_1^1), p_{-1}(\theta_2^2, \theta_3^2|\theta_1^1), \mathbf{0}_{20}).$$

Lemma A2. *The following three statements are equivalent*:

(A) $p(.)$ *is identifiable and Crémer–McLean condition holds for all $i \in \{1, \ldots, n\}$*;
(B) $S \cap \Omega = \emptyset$;
(C) *there exists a system of transfers $t(.)$ which is*:
  (i) *ex-post budget-balanced i.e., satisfies*

$$\sum_i t_i(\theta) = \mathbf{e}_\theta \cdot \mathbf{t} = 0 \quad \text{for all } \theta \in \Theta; \tag{12}$$

  (ii) *has zero expected value for any agent-type, when all agents report their types truthfully i.e.,*

$$\sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i) p_{-i}(\theta_{-i}|\theta_i) = \mathbf{p}_{\theta_i \theta_i} \cdot \mathbf{t} = 0$$

$$\text{for all } i = 1, \ldots, n, \text{ and } \theta_i \in \Theta_i; \tag{13}$$

  (iii) *strictly incentive compatible i.e., satisfies*

$$\sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_{-i}, \theta_i') p_{-i}(\theta_{-i}|\theta_i) = \mathbf{p}_{\theta_i \theta_i'} \cdot \mathbf{t} < 0 \quad \text{for all } i = 1, \ldots, n,$$

$$\text{and all } \theta_i, \theta_i' \in \Theta_i, \theta_i' \neq \theta_i. \tag{14}$$

Proof of Lemma A2. In step 1 we will show that statement (B) implies (C). Step 2 shows that (A) implies (B). Finally, Step 3 establishes that (C) implies (A).

Step 1: If $S \cap \Omega = \emptyset$, then there exists a system of transfers $t(.)$ satisfying conditions (12)–(14). Let $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ constitute a solution to the following problem:

$$(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \in \arg \min_{\mathbf{x} \in \Omega, \mathbf{y} \in S} \|\mathbf{x} - \mathbf{y}\|, \tag{15}$$

where $\|.\|$ is the standard Euclidian norm. Since the Euclidean norm is a continuous function, S is compact and nonempty, and Ω is a linear subspace, a solution to problem (15) exists by Weierstrass's Theorem. Let $\hat{\mathbf{t}} = \hat{\mathbf{x}} - \hat{\mathbf{y}}$. From $S \cap \Omega = \emptyset$ it follows that $\|\hat{\mathbf{t}}\| > 0$.

Let us show that $\hat{\mathbf{t}} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in \Omega$ and $\hat{\mathbf{t}} \cdot \mathbf{y} < 0$ for all $\mathbf{y} \in S$. By definition of Ω and S, this would establish that $\hat{\mathbf{t}}$ satisfies (12)–(14).

To see that $\hat{\mathbf{t}} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in \Omega$, choose an arbitrary $\mathbf{x} \in \Omega$, and for any $\alpha \in \mathbb{R}$ let $\varphi(\alpha) = \|\hat{\mathbf{t}} - \alpha \mathbf{x}\|^2 = \|\hat{\mathbf{t}}\|^2 - 2\alpha \hat{\mathbf{t}} \cdot \mathbf{x} + \alpha^2 \|\mathbf{x}\|^2$. Note that $\hat{\mathbf{t}} - \alpha \mathbf{x} = (\hat{\mathbf{x}} - \alpha \mathbf{x}) - \hat{\mathbf{y}}$, with $(\hat{\mathbf{x}} - \alpha \mathbf{x}) \in \Omega$ and $\hat{\mathbf{y}} \in S$. Then problem (15) implies that $\varphi(\alpha)$ reaches a minimum at $\alpha = 0$. Hence, $\varphi'(0) \equiv -2\hat{\mathbf{t}} \cdot \mathbf{x} = 0$.

Next, let us show that $\hat{\mathbf{t}} \cdot \mathbf{y} < 0$ for all $\mathbf{y} \in S$. First, since $\hat{\mathbf{t}} \cdot \hat{\mathbf{x}} = 0$ and $\hat{\mathbf{t}} = \hat{\mathbf{x}} - \hat{\mathbf{y}}$ we have $\hat{\mathbf{t}} \cdot \hat{\mathbf{y}} = -\|\hat{\mathbf{y}}\|^2 < 0$. (Recall that $\hat{\mathbf{y}} \neq \mathbf{0}$ since $\mathbf{0} \notin S$.)

Now take some $\mathbf{y} \in S \backslash \hat{\mathbf{y}}$ and $\alpha \in [0, 1]$. Then $(1 - \alpha)\hat{\mathbf{y}} + \alpha \mathbf{y} \in S$. Note that $\hat{\mathbf{x}} - [(1 - \alpha)\hat{\mathbf{y}} + \alpha \mathbf{y}] = \hat{\mathbf{t}} + \alpha(\hat{\mathbf{y}} - \mathbf{y})$. Since $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ solves (15), for all $\alpha \in [0, 1]$ we have

$$\|\hat{\mathbf{t}}\|^2 \leqslant \|\hat{\mathbf{t}} + \alpha(\hat{\mathbf{y}} - \mathbf{y})\|^2 = \|\hat{\mathbf{t}}\|^2 + 2\alpha \hat{\mathbf{t}} \cdot (\hat{\mathbf{y}} - \mathbf{y}) + \alpha^2 \|\hat{\mathbf{y}} - \mathbf{y}\|^2. \tag{16}$$

From (16) it follows that $\hat{\mathbf{t}} \cdot (\hat{\mathbf{y}} - \mathbf{y}) \geqslant 0$. For, if $\hat{\mathbf{t}} \cdot (\hat{\mathbf{y}} - \mathbf{y}) < 0$, then (16) fails for $\alpha = \min \left\{ 1, -[\hat{\mathbf{t}} \cdot (\hat{\mathbf{y}} - \mathbf{y})]/\|\hat{\mathbf{y}} - \mathbf{y}\|^2 \right\}$. Finally, $\hat{\mathbf{t}} \cdot (\hat{\mathbf{y}} - \mathbf{y}) \geqslant 0$ implies that: [19]

$$\hat{\mathbf{t}} \cdot \mathbf{y} \leqslant \hat{\mathbf{t}} \cdot \hat{\mathbf{y}} < 0. \tag{17}$$

Step 2: If $p(.)$ is identifiable and Crémer–McLean condition holds for all i, then $S \cap \Omega = \emptyset$.

The proof is by contrapositive. So, suppose that $S \cap \Omega \neq \emptyset$ and take some $\bar{\mathbf{x}} \in S \cap \Omega$. Since any element in S is a convex combination of nonzero vectors with nonnegative entries, we have $\mathbf{0} \notin S$, and hence $\bar{\mathbf{x}} \neq 0$. Then by definition of S and Ω:

$$\bar{\mathbf{x}} = \sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \sum_{\theta_i' \in \Theta_i : \theta_i' \neq \theta_i} \tilde{\gamma}_{\theta_i \theta_i'} \mathbf{p}_{\theta_i \theta_i'} = \sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \tilde{\gamma}_{\theta_i \theta_i} \mathbf{p}_{\theta_i \theta_i} + \sum_{\theta \in \Theta} \tilde{\mu}_\theta \mathbf{e}_\theta \tag{18}$$

for some collection of coefficients $\{\tilde{\mu}_\theta | \theta \in \Theta\}$ and $\{\tilde{\gamma}_{\theta_i \theta_i'} | i = 1, \ldots, n, \theta_i, \theta_i' \in \Theta_i\}$, with $\tilde{\gamma}_{\theta_i \theta_i'} \geqslant 0$ when $\theta_i \neq \theta_i'$. The sign restriction on $\tilde{\gamma}_{\theta_i \theta_i'}$, for $\theta_i \neq \theta_i'$, follows by definition of S.

Let us modify (18) by replacing the coefficients $\tilde{\mu}_\theta$ and $\tilde{\gamma}_{\theta_i \theta_i}$ with the ones that satisfy an additional nonnegativity restriction. By the definition of $\mathbf{e}_\theta$ and $\mathbf{p}_{\theta_i \theta_i}$, we have

$$\sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} p_i(\theta_i) \mathbf{p}_{\theta_i \theta_i} = \sum_{\theta \in \Theta} p(\theta) \mathbf{e}_\theta. \tag{19}$$

Let $C = \max \left\{ \max_{i \in \{1,\ldots,n\}, \theta_i \in \Theta_i} \{\tilde{\gamma}_{\theta_i \theta_i}/p_i(\theta_i)\}, \max_{\theta \in \Theta}\{-\tilde{\mu}_\theta/p(\theta)\}, 0 \right\}$. Next, add the left-hand side of (19) multiplied by C to the middle expression in (18), and add the right-hand side of (19) multiplied by C to the expression on the right-hand side of (18). Then rearrange to obtain

$$\sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \sum_{\theta_i' \in \Theta_i : \theta_i' \neq \theta_i} \tilde{\gamma}_{\theta_i \theta_i'} \mathbf{p}_{\theta_i \theta_i'} + \sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \left(C p_i(\theta_i) - \tilde{\gamma}_{\theta_i \theta_i} \right) \mathbf{p}_{\theta_i \theta_i}$$
$$= \sum_{\theta \in \Theta} \left(C p(\theta) + \tilde{\mu}_\theta \right) \mathbf{e}_\theta. \tag{20}$$

---

[19] Thus, we have established that a system of transfers satisfying the inequalities (12)–(14) can be derived by solving the convex minimization problem in (15). This problem has $nL + \sum_{i \in \{1,\ldots,n\}} m_i^2$ variables—the coefficients on the vectors $\mathbf{e}_\theta$, $\mathbf{p}_{\theta_i \theta_i}$ and $\mathbf{p}_{\theta_i \theta_i'}$. So deriving a solution to (15) is computationally straightforward. To provide a simple illustration of this method, consider a special case with $n = 3$, $\Theta_i = \{\theta_i^1, \theta_i^2\}$, $p(\theta_1^1, \theta_2^1, \theta_3^1) = p(\theta_1^2, \theta_2^2, \theta_3^2) = \frac{1}{8} + 3a$, and $p(\theta_1^{k_1}, \theta_2^{k_2}, \theta_3^{k_3}) = \frac{1}{8} - a$ for $(k_1, k_2, k_3) \notin \{(1,1,1),(2,2,2)\}$, $a \in [-1/24, 1/8]$. By symmetry of this case and convexity of the problem, there must exist a symmetric solution. That is, there exist $(b, c, d) \in \mathbb{R}^3$ such that $\hat{x} = b(\mathbf{e}_{\theta_1^1 \theta_2^1 \theta_3^1} + \mathbf{e}_{\theta_1^2 \theta_2^2 \theta_3^2}) + c \sum_{k_1, k_2, k_3 \in \{1,2\}} \mathbf{e}_{\theta_1^{k_1} \theta_2^{k_2} \theta_3^{k_3}} + d \sum_{i \in \{1,2,3\}, k_i \in \{1,2\}} \mathbf{p}_{\theta_i^{k_i} \theta_i^{k_i}}$, and $\hat{y} = \frac{1}{6} \sum_{i \in \{1,2,3\}} (\mathbf{p}_{\theta_i^1 \theta_i^2} + \mathbf{p}_{\theta_i^2 \theta_i^1})$.
Substituting these values into (15) we obtain the following reduced optimization problem

$$\min_{(b,c,d) \in \mathbb{R}^3} \frac{1}{384} + \frac{1}{2}a^2 + \left(\frac{3}{8} + a\right)c + \left(\frac{1}{8} - a\right)b + \left(\frac{3}{4} + 18a\right)b \times d + \left(\frac{9}{4} - 18a\right)c \times d$$
$$+ \left(\frac{1}{32} - 2a^2\right)d + \left(\frac{3}{32} + 18a^2\right)d^2 + 6b^2 + 18c^2.$$

The solution to this minimization problem is given by $b = \frac{1}{9}a$, $c = -\frac{1}{96} - \frac{1}{36}a$, $d = 0$. Hence the vector of transfers satisfying (12)–(14) is given by $\mathbf{t} = \frac{a}{9}(0, 1, 1, -2, -2, 1, 1, 0, 0, 1, -2, 1, 1, -2, 1, 0, 0, -2, 1, 1, 1, 1, -2, 0)$.

By definition of C, $Cp_i(\theta_i) - \tilde{\gamma}_{\theta_i\theta_i} \geqslant 0$ for all i and $\theta_i \in \Theta_i$ and $Cp(\theta) + \tilde{\mu}_\theta \geqslant 0$ for all $\theta \in \Theta$. Furthermore, since $\bar{\mathbf{x}} \neq \mathbf{0}$ in (18), there exist $j \in \{1, \dots, n\}$ and $\hat{\theta}_j, \hat{\theta}'_j \in \Theta_j$, with $\hat{\theta}_j \neq \hat{\theta}'_j$, s.t. $\tilde{\gamma}_{\hat{\theta}_j\hat{\theta}'_j} > 0$. Therefore, $Cp(\theta) + \tilde{\mu}_\theta > 0$ for some $\theta \in \Theta$. Next, let $G = \sum_{\theta \in \Theta}(Cp(\theta) + \tilde{\mu}_\theta) > 0$ and let $\gamma_{\theta_i\theta'_i} = \frac{\tilde{\gamma}_{\theta_i\theta'_i}}{G}$, $\gamma_{\theta_i\theta_i} = \frac{Cp_i(\theta_i) - \tilde{\gamma}_{\theta_i\theta_i}}{G}$ for all i and all $\theta_i, \theta'_i \in \Theta_i$, $\theta_i \neq \theta'_i$, and $\mu(\theta) = \frac{Cp(\theta) + \tilde{\mu}_\theta}{G}$ for all $\theta \in \Theta$. Then (20) implies that

$$\sum_{i=1,\dots,n} \sum_{\theta_i \in \Theta_i} \sum_{\theta'_i \in \Theta_i} \gamma_{\theta_i\theta'_i} \mathbf{p}_{\theta_i\theta'_i} = \sum_{\theta \in \Theta} \mu_\theta \mathbf{e}_\theta, \quad \text{where } \sum_{\theta \in \Theta} \mu_\theta = 1. \tag{21}$$

We need to consider two possible cases.

Case 1: $\mu_\theta = p(\theta)$ for all $\theta \in \Theta$.

Recall that $\gamma_{\hat{\theta}_j\hat{\theta}'_j} > 0$ for some j, $\hat{\theta}_j, \hat{\theta}'_j \in \Theta_j$, $\hat{\theta}_j \neq \hat{\theta}'_j$, and that the left-hand and the right-hand sides of (21) are vectors of size nL, with each entry corresponding to one of the agents and one of type profiles $\theta \in \Theta$. Consider L_{-j} entries of these vectors corresponding to agent j and type profiles $(\theta_{-j}, \hat{\theta}'_j)$, $\theta_{-j} \in \Theta_{-j}$. By (21), in those entries we have

$$\sum_{\theta_j \in \Theta_j} \gamma_{\theta_j\hat{\theta}'_j} p_{-j}(\theta_{-j}|\theta_j) = p(\theta_{-j}, \hat{\theta}'_j) \quad \text{for all } \theta_{-j} \in \Theta_{-j}. \tag{22}$$

Using $p(\theta_{-j}, \hat{\theta}'_j) = p_{-j}(\theta_{-j}|\hat{\theta}'_j) p_j(\hat{\theta}'_j)$ and rearranging (22), we get:

$$\sum_{\theta_j \in \Theta_j : \theta_j \neq \hat{\theta}'_j} \frac{\gamma_{\theta_j\hat{\theta}'_j}}{p_j(\hat{\theta}'_j) - \gamma_{\hat{\theta}'_j\hat{\theta}'_j}} p_{-j}(\theta_{-j}|\theta_j) = p_{-j}(\theta_{-j}|\hat{\theta}'_j) \text{for all} \theta_{-j} \in \Theta_{-j}.$$

Note that $p_j(\hat{\theta}'_j) - \gamma_{\hat{\theta}'_j\hat{\theta}'_j} > 0$ because $\gamma_{\hat{\theta}_j\hat{\theta}'_j} > 0$ and Eq. (22) holds. Hence, $p_{-j}(.|\hat{\theta}'_j)$ is a convex combination of $\{p_{-j}(.|\theta_j): \theta_j \in \Theta_j, \theta_j \neq \hat{\theta}'_j\}$, i.e. Crémer–McLean condition fails for j.

Case 2: $\mu_\theta \neq p(\theta)$ for some $\theta \in \Theta$.

Consider the probability distribution $q(.)$ over Θ s.t. $q(\theta) = \mu_\theta$ for all $\theta \in \Theta$. Since $\sum_{\theta \in \Theta} \mu_\theta = 1$, we indeed have $q(.) \in \mathcal{P}(\Theta)$ and $q(.) \neq p(.)$. By (21), for any agent i and any $\theta'_i \in \Theta_i$ we have

$$\sum_{\theta_i \in \Theta_i} \gamma_{\theta_i\theta'_i} p_{-i}(\theta_{-i}|\theta_i) = q(\theta_{-i}, \theta'_i) \quad \text{for all } \theta_{-i} \in \Theta_{-i}.$$

Thus, for all i and θ'_i s.t. $q_i(\theta'_i) > 0$ we have: $\sum_{\theta_i \in \Theta_i} \frac{\gamma_{\theta_i\theta'_i}}{q_i(\theta'_i)} p_{-i}(\theta_{-i}|\theta_i) = q_{-i}(\theta_{-i}|\theta'_i)$ for all $\theta_{-i} \in \Theta_{-i}$. So, $p(.)$ is not identifiable.

Step 3: If there exists a system of transfers $t(.) = (t_1(.), \dots, t_1(.))$ satisfying inequalities (12)–(14), then $p(.)$ is identifiable and Crémer–McLean condition holds for all agents.

The proof is by contradiction. So suppose, first, that a system of transfers $t(.)$ satisfying (12)–(14) exists, but $p(.)$ is not identifiable i.e., there exists a probability distribution $q(.) \in \mathcal{P}(\theta)$, $q(.) \neq p(.)$, and a collection of nonnegative coefficients $\{c_{\theta_i\theta'_i}|i \in \{1, \dots, n\}, \theta_i, \theta'_i \in \Theta_i\}$ s.t.

for all i and $\theta_i' \in \Theta_i$ with $q_i(\theta_i') > 0$, we have: $q_{-i}(.|\theta_i') = \sum_{\theta_i \in \Theta_i} c_{\theta_i \theta_i'} p_{-i}(.|\theta_i)$. Then,

$$\sum_{\theta \in \Theta} \sum_{i=1,\ldots,n} t_i(\theta) q(\theta)$$

$$= \sum_{i=1,\ldots,n} \sum_{\theta_i' \in \Theta_i : q_i(\theta_i')>0} q_i(\theta_i') \sum_{\theta_{-i} \in \Theta_{-i}} q_{-i}(\theta_{-i}|\theta_i') t_i(\theta_{-i}, \theta_i')$$

$$= \sum_{i=1,\ldots,n} \sum_{\theta_i' \in \Theta_i : q_i(\theta_i')>0} q_i(\theta_i') \sum_{\theta_i \in \Theta_i} \sum_{\theta_{-i} \in \Theta_{-i}} c_{\theta_i \theta_i'} p_{-i}(\theta_{-i}|\theta_i) t_i(\theta_{-i}, \theta_i') < 0. \qquad (23)$$

The first equality holds by definition. The second equality holds because $p(.)$ is not identifiable. The inequality holds because the system of transfers $t(.)$ satisfies inequalities (13) and (14) and, by Lemma A1, $c_{\theta_i \theta_i'} > 0$ for some i and $\theta_i, \theta_i' \in \Theta_i$, $\theta_i \neq \theta_i'$. Hence, (23) contradicts the fact that $t(.)$ is budget-balanced, i.e. $\sum_{i \in \{1,\ldots,n\}} t_i(\theta) = 0$ for all $\theta \in \Theta$.

If Crémer–McLean condition fails for some agent i, then it is easy to show that for this agent (13) and (14) cannot hold together. We omit this step for brevity. □

Let us briefly comment on the proof of Lemma A2. Steps 1 and 2 have established that a system of transfers satisfying (12)–(14) exists and can be derived by solving the minimization problem (15), if Identifiability and Crémer–McLean conditions hold. Eq. (21) highlights the role of these two conditions. In particular, when Identifiability holds, then for every probability distribution of reported type profiles (represented by vector μ_θ in Eq. (21)) different from the prior, there is an agent-type θ_i for whom (21) cannot hold. This agent-type is a nondeviator under this distribution, i.e. the report of this type is surely truthful.

To understand why a system of transfers satisfying conditions (12)–(14) requires that, under any probability distribution of type profiles different from $p(.)$, there exists a nondeviator type who gets a positive expected transfer, let us argue by contradiction. So, suppose that under some probability distribution $q(.)$, $q(.) \neq p(.)$, either there are no nondeviator types or each of them gets a nonpositive expected transfer. All other agent-types are potential deviators, i.e. could induce the conditional probability distributions corresponding to $q(.)$ by unilateral deviations. So, by (13) and (14), they should get negative expected transfers under $q(.)$. But then there is no type left to receive a positive expected transfer and balance the budget. Inequality (23) establishes this failure of budget-balance when there are no nondeviator types i.e., when the Identifiability condition fails.

To complete the construction of our mechanism we need another system of transfers $\tau(\theta) = (\tau_1(\theta), \ldots, \tau_n(\theta))$ to redistribute the generated surplus among agent-types in the way specified by the mechanism designer. The existence of such system of transfers is established in Lemma A3. We will then combine it with transfers satisfying (12)–(14) to obtain the desired mechanism.

Lemma A3. *Consider a collection of coefficients $\{w_i(\theta_i)|i = 1, \ldots, n, \theta_i \in \Theta_i\}$ such that*

$$\sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} p_i(\theta_i) w_i(\theta_i) = 0. \qquad (24)$$

Then there exists a system of transfers $\tau(\theta) = (\tau_1(\theta), \ldots, \tau_n(\theta))$ satisfying $\sum_i \tau_i(\theta) = 0$ for all $\theta \in \Theta$ and

$$\sum_{\theta_{-i} \in \Theta_{-i}} \tau_i(\theta_{-i}, \theta_i) p_{-i}(\theta_{-i}|\theta_i) = w_i(\theta_i) \quad \text{for all } i = 1, \ldots, n, \text{ and } \theta_i \in \Theta_i. \qquad (25)$$

In words, the system of transfers $\tau(\theta)$ is ex-post budget balanced and ensures that any agent-type θ_i gets the interim expected payoff equal to $w_i(\theta_i)$. Eq. (24) implies that coefficients $w_i(\theta_i)$ represent a redistribution of surplus.

Proof of Lemma A3. The proof is by construction. First, let us fix some $\bar{\theta}_1 \in \Theta_1$ and $\bar{\theta}_2 \in \Theta_2$. By the regularity condition, for any two agent-types $\theta_i \in \Theta_i$ and $\theta_j \in \Theta_j$, $i, j \in \{1, \ldots, n\}$, there is a profile of other agents' types $\hat{\theta}_{-i-j}(\theta_i, \theta_j) \in \Theta_{-i-j}$ such that $p(\theta_i, \theta_j, \hat{\theta}_{-i-j}(\theta_i, \theta_j)) > 0$. Using this property, set the following transfers between agent-types $\theta_1 \in \Theta_1 \backslash \bar{\theta}_1$, and $\bar{\theta}_2$:

$$\tau_1(\theta_1, \bar{\theta}_2, \hat{\theta}_{-1-2}(\theta_1, \bar{\theta}_2)) = -\tau_2(\theta_1, \bar{\theta}_2, \hat{\theta}_{-1-2}(\theta_1, \bar{\theta}_2)) = \frac{w_1(\theta_1) p_1(\theta_1)}{p(\theta_1, \bar{\theta}_2, \hat{\theta}_{-1-2}(\theta_1, \bar{\theta}_2))}.$$

Next, set the following transfers between each agent-type $\theta_i \in \bigcup_{i \in \{2, \ldots, n\}} \Theta_i \backslash \bar{\theta}_2$, and $\bar{\theta}_1$:

$$\tau_i(\bar{\theta}_1, \theta_i, \hat{\theta}_{-1-i}(\bar{\theta}_1, \theta_i)) = -\tau_1(\bar{\theta}_1, \theta_i, \hat{\theta}_{-1-i}(\bar{\theta}_1, \theta_i)) = \frac{w_i(\theta_i) p_i(\theta_i)}{p(\bar{\theta}_1, \theta_i, \hat{\theta}_{-1-i}(\bar{\theta}_1, \theta_i))}.$$

Finally, set the transfers between agent-types $\bar{\theta}_1$ and $\bar{\theta}_2$ as follows:

$$\begin{aligned}\tau_2(\bar{\theta}_1, \bar{\theta}_2, \hat{\theta}_{-1-2}(\bar{\theta}_1, \bar{\theta}_2)) &= -\tau_1(\bar{\theta}_1, \bar{\theta}_2, \hat{\theta}_{-1-2}(\bar{\theta}_1, \bar{\theta}_2)) \\ &= \frac{w_2(\bar{\theta}_2) p_2(\bar{\theta}_2) + \sum_{\theta_1 \in \Theta_1 \backslash \bar{\theta}_1} w_1(\theta_1) p_1(\theta_1)}{p(\bar{\theta}_1, \bar{\theta}_2, \hat{\theta}_{-1-2}(\bar{\theta}_1, \bar{\theta}_2))}.\end{aligned}$$

The system of transfers $\tau_i(.)$ that we have constructed above is such that condition (25) holds for every $\theta_i \in \bigcup_{i \in \{1, \ldots, n\}} \Theta_i \backslash \bar{\theta}_1$. This is immediate for any agent-type in $\bigcup_{i \in \{1, \ldots, n\}} \Theta_i \backslash \{\bar{\theta}_1, \bar{\theta}_2\}$ and can be ascertained for $\bar{\theta}_2$ by direct computation.

Finally, let us show that condition (25) also holds for agent-type $\bar{\theta}_1$. Since our system of transfers $\tau_i(.)$ is budget balanced and condition (24) holds, we have

$$\begin{aligned}\sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \sum_{\theta_{-i} \in \Theta_{-i}} & \tau_i(\theta_i, \theta_{-i}) p_{-i}(\theta_{-i}|\theta_i) p_i(\theta_i) \\ &= \sum_{\theta \in \Theta} p(\theta) \sum_{i=1,\ldots,n} \tau_i(\theta) = 0 = \sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} p_i(\theta_i) w_i(\theta_i).\end{aligned} \tag{26}$$

Since condition (25) holds for every $\theta_i \in \bigcup_{i \in \{1, \ldots, n\}} \Theta_i \backslash \bar{\theta}_1$, Eq. (26) can hold only if $\sum_{\theta_{-1} \in \Theta_{-1}} \tau_1(\bar{\theta}_1, \theta_{-1}) p_{-1}(\theta_{-1}|\bar{\theta}_1) = w_1(\bar{\theta}_1)$. $\square$

To complete the proof of sufficiency, let us fix an arbitrary *EASR* decision rule $x(.)$ and a feasible profile of net payoffs $\{v_i(\theta_i) \geqslant 0 | i = 1, \ldots, n; \theta_i \in \Theta_i\}$. We will now combine the results of Lemmas A2 and A3 to construct a *BB*, *IR* mechanism which implements the decision rule $x(.)$ and in which agent-type θ_i obtains net payoff $v_i(\theta_i)$.

Note that by budget-balancing, feasible net payoffs have to satisfy

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

Let $w_i(\theta_i) = v_i(\theta_i) - \sum_{\theta_{-i} \in \Theta_{-i}} u_i(x(\theta_{-i}, \theta_i), (\theta_{-i}, \theta_i)) p_{-i}(\theta_{-i}|\theta_i)$. Then $\sum_{i \in \{1, \ldots, n\}} \sum_{\theta_i \in \Theta_i} p_i(\theta_i) w_i(\theta_i) = 0$. Next, using the method of Lemma A3, construct a budget-balanced system of transfers $\tau(.) = (\tau_1(.), \ldots, \tau_n(.))$ $(\tau \in \mathbf{R}^{nL}$ in vector notation) satisfying

$\sum_{\theta_{-i} \in \Theta_{-i}} \tau_i(\theta_{-i}, \theta_i) p_{-i}(\theta_{-i} | \theta_i) = w_i(\theta_i)$ for all $i = 1, \ldots, n$, and $\theta_i \in \Theta_i$. Using the vector notation introduced above, the latter condition can be rewritten as follows:

$$\mathbf{p}_{\theta_i \theta_i} \cdot \boldsymbol{\tau} = w_i(\theta_i) \quad \text{for all } i = 1, \ldots, n \text{ and } \theta_i \in \Theta_i.$$

Further, take a vector of transfers $\mathbf{t} \in \mathbf{R}^{nL}$ satisfying (12)–(14) and consider the mechanism with decision rule $x(.)$ and aggregate system of transfers $\boldsymbol{\tau} + b\mathbf{t}$ for some $b \in \mathbf{R}_+$. This mechanism is ex-post budget balanced by construction. The payoff to any agent-type θ_i in this mechanism is equal to $w_i(\theta_i) + \sum_{\theta_{-i} \in \Theta_{-i}} u_i(x(\theta_{-i}, \theta_i), (\theta_{-i}, \theta_i)) p_{-i}(\theta_{-i} | \theta_i) = v_i(\theta_i) \geqslant 0$, so all *IR* constraints hold. It remains to choose b appropriately to ensure that *IC* constraints hold, i.e. that for all i, $\theta_i, \theta_i' \in \Theta_i, \theta_i \neq \theta_i'$:

$$(\mathbf{p}_{\theta_i \theta_i} - \mathbf{p}_{\theta_i \theta_i'}) \cdot (\boldsymbol{\tau} + b\mathbf{t}) \geqslant - \sum_{\theta_{-i} \in \Theta_{-i}} (u_i(x(\theta_{-i}, \theta_i), (\theta_{-i}, \theta_i)) - u_i(x(\theta_{-i}, \theta_i'), (\theta_{-i}, \theta_i)) p_{-i}(\theta_{-i} | \theta_i).$$

$$(27)$$

Since $\mathbf{p}_{\theta_i \theta_i} \cdot \mathbf{t} = 0$ and $\mathbf{p}_{\theta_i \theta_i'} \cdot \mathbf{t} < 0$ by construction, rearranging terms in (27) yields that (27) holds for all $i, \theta_i, \theta_i' \in \Theta_i$ if we choose b to exceed:

$$\max_{i \in \{1, \ldots, n\}, \theta_i, \theta_i' \in \Theta_i : \theta_i' \neq \theta_i} \left\{ \frac{\sum_{\theta_{-i} \in \Theta_{-i}} u_i(x(\theta_{-i}, \theta_i), (\theta_{-i}, \theta_i)) - u_i(x(\theta_{-i}, \theta_i'), (\theta_{-i}, \theta_i)) p_{-i}(\theta_{-i} | \theta_i) + (\mathbf{p}_{\theta_i \theta_i} - \mathbf{p}_{\theta_i \theta_i'}) \cdot \boldsymbol{\tau}}{\mathbf{p}_{\theta_i \theta_i'} \cdot \mathbf{t}} \right\}.$$

Necessity: Now let us prove the necessity part of the Theorem. As shown in Step 2 of Lemma A2, if $p(.)$ is not identifiable or Crémer–McLean condition fails for some agent $i \in \{1, \ldots, n\}$, then there exists a collection of coefficients $\{\mu_\theta | \theta \in \Theta\}$ and $\{\gamma_{\theta_i \theta_i'} \geqslant 0 | i = 1, \ldots, n, \theta_i, \theta_i' \in \Theta_i\}$, with $\gamma_{\hat{\theta}_j \tilde{\theta}_j} > 0$ for some $\hat{\theta}_j, \tilde{\theta}_j \in \Theta_j, \hat{\theta}_j \neq \tilde{\theta}_j$, such that

$$\sum_{i=1, \ldots, n} \sum_{\theta_i \in \Theta_i} \sum_{\theta_i' \in \Theta_i} \gamma_{\theta_i \theta_i'} \mathbf{p}_{\theta_i \theta_i'} = \sum_{\theta \in \Theta} \mu_\theta \mathbf{e}_\theta. \tag{28}$$

Below, we will construct a profile of utility functions and an efficient decision rule in such a way that agent-type $\hat{\theta}_j$ would have a strong incentive to deviate and imitate type $\tilde{\theta}_j$. Yet, this deviation cannot be punished with a sufficiently negative expected transfer. Indeed, if Crémer–McLean condition fails due to $\gamma_{\hat{\theta}_j \tilde{\theta}_j} > 0$, then j can report $\tilde{\theta}_j$ nontruthfully but in a nondetectable way i.e., so that the probability distribution of the other agents' reported type profiles conditional on j's reported type $\tilde{\theta}_j$ will be the same as $p_{-j}(.|\tilde{\theta}_j)$. But a truthtelling agent-type $\tilde{\theta}_j$ should not get a large negative expected transfer.

Alternatively, if Identifiability fails because of $\gamma_{\hat{\theta}_j \tilde{\theta}_j} > 0$, then every agent can induce by a unilateral deviation the same probability distribution that is induced when j unilaterally commits a deviation that involves $\hat{\theta}_j$ reporting $\tilde{\theta}_j$. Therefore, giving some other agent the proceeds from a negative expected transfer to agent j under this probability distribution, would create an incentive for the former to deviate in a way that induces this distribution.

Hence, the only way to prevent $\hat{\theta}_j$ from deviating is to give her a sufficiently large informational rent for reporting truthfully. Yet, in the example that we construct below, the available social surplus is not sufficiently large to cover the required informational rent, and so a budget-balanced mechanism fails to exist.

Specifically, suppose that $X \equiv \{x_1, x_2\}$ and consider a profile of the utility functions $u_i(x, \theta)$, $i = 1, \dots, n$, such that for some scalars $a > 0$ and $B > 0$ we have:

(i) $u_i(x_1, \theta) = a$ for all $i = 1, \dots, n$ and $\theta \in \Theta$, except for $i = j$ and $\theta = (\theta_{-j}, \tilde{\theta}_j)$; $u_j(x_1, (\theta_{-j}, \tilde{\theta}_j)) = 0$ for all $\theta_{-j} \in \Theta_{-j}$;

(ii) $u_i(x_2, (\theta_{-j}, \theta_j)) = a$ if $\theta_j \neq \hat{\theta}_j$, $u_i(x_2, (\theta_{-j}, \hat{\theta}_j)) = a - 2B$ for all $i \neq j$ and $\theta_{-j} \in \Theta_{-j}$;

(iii) $u_j(x_2, (\theta_{-j}, \tilde{\theta}_j)) = a$, $u_j(x_2, (\theta_{-j}, \hat{\theta}_j)) = a + B$, $u_j(x_2, (\theta_{-j}, \theta_j)) = 0$ for all $\theta_{-j} \in \Theta_{-j}$ and $\theta_j \notin \{\hat{\theta}_j, \tilde{\theta}_j\}$;

for this utility profile the unique ex-post efficient decision rule is:

$$x^*(\theta_{-j}, \theta_j) \equiv \begin{cases} x_1 & \text{if } \theta_j \neq \tilde{\theta}_j \text{ and any } \theta_{-j} \in \Theta_{-j}, \\ x_2 & \text{if } \theta_j = \tilde{\theta}_j \text{ and any } \theta_{-j} \in \Theta_{-j}. \end{cases}$$

Note that, with the decision rule $x^*(.)$, $u_i(x^*(\theta), \theta) = a$ for all $i \in \{1, \dots, n\}$ and $\theta \in \Theta$.

By definition, an *IR*, *BB* mechanism implementing the decision rule $x^*(.)$ exists if and only if there is a system of transfers $\tilde{\mathbf{t}} \in \mathbf{R}^{nL}$ (in vector notation) satisfying:

(i) BB: $\mathbf{e}_\theta \cdot \tilde{\mathbf{t}} = 0$ for all $\theta \in \Theta$; $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (29)

(ii) IR: $\mathbf{p}_{\theta_i \theta_i} \cdot \tilde{\mathbf{t}} \geqslant -a$ for all $i = 1, \dots, n$ and $\theta_i \in \Theta_i$; $\qquad\qquad\qquad$ (30)

(iii) IC: $\mathbf{p}_{\theta_i \theta_i} \cdot \tilde{\mathbf{t}} - \mathbf{p}_{\theta_i \theta_i'} \cdot \tilde{\mathbf{t}} \geqslant \varphi(i, \theta_i, \theta_i')$ for all $i = 1, \dots, n$ and all $\theta_i, \theta_i' \in \Theta_i : \theta_i \neq \theta_i'$, (31)

where $\varphi(i, \theta_i, \theta_i') = 0$ for all $i \neq j$ and

$$\varphi(j, \theta_j, \theta_j') = \begin{cases} B & \text{if } \theta_j = \hat{\theta}_j \text{ and } \theta_j' = \tilde{\theta}_j, \\ -a & \text{if } \theta_j = \tilde{\theta}_j \text{ and } \theta_j' \neq \tilde{\theta}_j, \\ -a & \text{if } \theta_j \notin \{\hat{\theta}_j, \tilde{\theta}_j\} \text{ and } \theta_j' = \tilde{\theta}_j, \\ 0 & \text{otherwise.} \end{cases} \qquad (32)$$

In the rest of the proof we will show that if B, the utility gain which $\hat{\theta}_j$ gets by reporting $\tilde{\theta}_j$, is sufficiently large, while per-capita social surplus, a, is sufficiently small, then (29)–(31) are incompatible with (28) and so the desired mechanism fails to exist.

Formally, since $u_i(x^*(\theta), \theta) = a$ for all i and $\theta \in \Theta$, the expected transfer received by any agent-type θ_i cannot strictly exceed $na/p_i(\theta_i)$, because otherwise budget balance would require the transfer to some other agent-type θ_l to be strictly less than $-a$ which would violate this type's individual rationality constraint. [20] That is, we must have

$$\mathbf{p}_{\theta_i \theta_i} \cdot \tilde{\mathbf{t}} \leqslant \frac{na}{p_i(\theta_i)} \qquad \text{for all } i = 1, \dots, n, \text{ and } \theta_i \in \Theta_i. \qquad (33)$$

---

[20] Indeed, budget-balance requires that $\sum_{\theta \in \Theta} p(\theta) \sum_{i \in \{1, \dots, n\}} t_i(\theta) = 0$, which can be rewritten as $p_i(\theta_i) \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i, \theta_{-i}) p_{-i}(\theta_{-i}|\theta_i) + \sum_{\theta_i' \neq \theta_i} p_i(\theta_i') \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i', \theta_{-i}) p_{-i}(\theta_{-i}|\theta_i')$ $+ \sum_{j \neq i} \sum_{\theta_j \in \Theta_j} p_j(\theta_j) \sum_{\theta_{-j} \in \Theta_{-j}} t_j(\theta_j, \theta_{-j}) p_{-j}(\theta_{-j}|\theta_j) = 0$. So, if $p_i(\theta_i) \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i, \theta_{-i}) p_{-i}(\theta_{-i}|\theta_i) > na$, then the expected transfer to some other agent-type has to be less than $-a$.

Combining (31) and (33) yields:

$$\mathbf{p}_{\theta_i \theta_i'} \cdot \tilde{\mathbf{t}} \leqslant p_{\theta_i \theta_i} \cdot \tilde{\mathbf{t}} - \varphi(i, \theta_i, \theta_i') \leqslant \frac{n}{p_i(\theta_i)} a - \varphi(i, \theta_i, \theta_i')$$

$$\text{for all } i = 1, \ldots, n \text{ and all } \theta_i, \theta_i' \in \Theta_i : \theta_i' \neq \theta_i. \tag{34}$$

Using (32)–(34), we obtain

$$\sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \sum_{\theta_i' \in \Theta_i} \gamma_{\theta_i \theta_i'} \mathbf{p}_{\theta_i \theta_i'} \cdot \tilde{\mathbf{t}}$$

$$\leqslant \sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \left\{ \sum_{\theta_i' \in \Theta_i} \gamma_{\theta_i \theta_i'} \frac{na}{p_i(\theta_i)} - \sum_{\theta_i' \in \Theta_i : \theta_i' \neq \theta_i} \gamma_{\theta_i \theta_i'} \varphi(i, \theta_i, \theta_i') \right\}$$

$$\leqslant a \left\{ n \sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \sum_{\theta_i' \in \Theta_i} \frac{\gamma_{\theta_i \theta_i'}}{p_i(\theta_i)} + \sum_{\theta_j' \in \Theta_j : \theta_j' \neq \tilde{\theta}_j} \gamma_{\tilde{\theta}_j \theta_j'} + \sum_{\theta_j \in \Theta_j : \theta_j \notin \{\hat{\theta}_j, \tilde{\theta}_j\}} \gamma_{\theta_j \tilde{\theta}_j} \right\} - B \gamma_{\hat{\theta}_j \tilde{\theta}_j}.$$
$$\tag{35}$$

Recall that $\gamma_{\hat{\theta}_j \tilde{\theta}_j} > 0$. So, when a is sufficiently small and B is sufficiently large, the right-hand side of (35) is strictly negative. On the other hand, multiplying both sides of (28) by $\tilde{\mathbf{t}}$ and using (29), we obtain: $\sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \sum_{\theta_i' \in \Theta_i} \gamma_{\theta_i \theta_i'} \mathbf{p}_{\theta_i \theta_i'} \cdot \tilde{\mathbf{t}} = \sum_{\theta \in \Theta} \mu_\theta \mathbf{e}_\theta \cdot \tilde{\mathbf{t}} = 0$. This contradicts the fact that the right-hand side of (35) is strictly negative. $\quad\square$

Proof of Corollary 2. By Lemma A2, it is sufficient to show that, when $n = 3$ and $\Theta_i = \{\theta_i^1, \theta_i^2\}$ for $i \in \{1, 2, 3\}$, $S \cup \Omega = \emptyset$ implies that inequalities (7) and (8) hold for some relabeling of agents and their types. The proof if by contrapositive, i.e. we will show that $S \cup \Omega \neq \emptyset$ if (7) and (8) do not hold for any relabeling of agents and their types.

By Eq. (18), $S \cup \Omega \neq \emptyset$ if there exists a profile of coefficients μ_{jkh}, $\gamma_{\theta_i^1 \theta_i^1}$, $\gamma_{\theta_i^2 \theta_i^2}$ $\gamma_{\theta_i^1 \theta_i^2} \geqslant 0$ and $\gamma_{\theta_i^2 \theta_i^1} \geqslant 0$, where $i \in \{1, 2, 3\}$ and $j, k, h \in \{1, 2\}$, such that the following equation holds: [21]

$$\sum_{j,k,h \in \{1,2\}} \mu_{jkh} \mathbf{e}_{\theta_1^j \theta_2^k \theta_3^h} + \sum_{i \in \{1,2,3\}} \sum_{j \in \{1,2\}} \gamma_{\theta_i^j \theta_i^j} \mathbf{p}_{\theta_i^j \theta_i^j} = \sum_{i \in \{1,2,3\}} (\gamma_{\theta_i^1 \theta_i^2} \mathbf{p}_{\theta_i^1 \theta_i^2} + \gamma_{\theta_i^2 \theta_i^1} \mathbf{p}_{\theta_i^2 \theta_i^1}) \neq 0.$$
$$\tag{36}$$

Simple computation establishes that system (36) has the following solution for any $v \in \mathbb{R}$:

$$\mu_{111} = v p_{211}(p_{111} p_{122} - p_{112} p_{121}), \quad \mu_{112} = v p_{212}(p_{111} p_{122} - p_{112} p_{121}),$$

$$\mu_{121} = v p_{221}(p_{111} p_{122} - p_{112} p_{121}), \quad \mu_{122} = v p_{222}(p_{111} p_{122} - p_{112} p_{121}),$$

$$\mu_{211} = -v p_{211}(p_{121} p_{212} + p_{112} p_{221} - p_{122} p_{211}) + v p_{111} p_{212} p_{221},$$

$$\mu_{212} = -v p_{212}(p_{121} p_{212} - p_{122} p_{211} - p_{111} p_{222}) - v p_{112} p_{211} p_{222},$$

$$\mu_{221} = -v p_{221}(p_{112} p_{221} - p_{122} p_{211} - p_{111} p_{222}) - v p_{121} p_{211} p_{222},$$

$$\mu_{222} = -v p_{222}(p_{121} p_{212} + p_{112} p_{221} - p_{111} p_{222}) + v p_{122} p_{212} p_{221},$$

$$\gamma_{\theta_1^1 \theta_1^1} = 0, \quad \gamma_{\theta_1^2 \theta_1^2} = v p(\theta_1^2)(p_{121} p_{212} + p_{112} p_{221} - p_{122} p_{211} - p_{111} p_{222}),$$

---

[21] Recall that in the case of $n = 3$, the vectors $\mathbf{e}_\theta$, $\mathbf{p}_{\theta_i \theta_i}$ and $\mathbf{p}_{\theta_i \theta_i'}$ are illustrated in footnote 18.

$$\gamma_{\theta_2^1 \theta_2^1} = vp(\theta_2^1)\left(p_{121}p_{212} - p_{211}p_{122}\right), \quad \gamma_{\theta_2^2 \theta_2^2} = vp(\theta_2^2)\left(p_{221}p_{112} - p_{111}p_{222}\right),$$

$$\gamma_{\theta_3^1 \theta_3^1} = vp(\theta_3^1)\left(p_{112}p_{221} - p_{211}p_{122}\right), \quad \gamma_{\theta_3^2 \theta_3^2} = vp(\theta_3^2)\left(p_{212}p_{121} - p_{111}p_{222}\right),$$

$$\gamma_{\theta_1^1 \theta_1^2} = vp(\theta_1^1)(p_{212}p_{221} - p_{211}p_{222}), \quad \gamma_{\theta_1^2 \theta_1^1} = vp(\theta_1^2)(p_{111}p_{122} - p_{112}p_{121}),$$

$$\gamma_{\theta_2^1 \theta_2^2} = vp(\theta_2^1)(p_{122}p_{221} - p_{121}p_{222}), \quad \gamma_{\theta_2^2 \theta_2^1} = vp(\theta_2^2)(p_{111}p_{212} - p_{112}p_{211}),$$

$$\gamma_{\theta_3^1 \theta_3^2} = vp(\theta_3^1)(p_{122}p_{212} - p_{112}p_{222}), \quad \gamma_{\theta_3^2 \theta_3^1} = vp(\theta_3^2)(p_{111}p_{221} - p_{121}p_{211}). \tag{37}$$

Let us show that if (7) and (8) do not hold for any relabeling of agents and their types, then all the coefficients $\gamma_{\theta_1^1 \theta_1^2}, \gamma_{\theta_2^2 \theta_2^1}, \gamma_{\theta_2^1 \theta_2^2}, \gamma_{\theta_2^2 \theta_2^1}, \gamma_{\theta_3^1 \theta_3^2}, \gamma_{\theta_3^2 \theta_3^1}$ are (weakly) of the same sign, i.e. either all nonnegative or nonpositive.

First, since (7) and (8) do not hold plainly without relabeling, $\gamma_{\theta_1^2 \theta_1^1}$ and $\gamma_{\theta_2^2 \theta_2^1}$ are of the same sign (i.e. $\gamma_{\theta_1^2 \theta_1^1} \gamma_{\theta_2^2 \theta_2^1} \geqslant 0$). Second, if we relabel players 2 and 3 as 3 and 2, respectively, then the left-hand side of (7) does not change, while the left-hand side of (8) becomes equal to $p_{111}p_{221} - p_{121}p_{211}$. So $\gamma_{\theta_1^2 \theta_1^1}$ and $\gamma_{\theta_3^2 \theta_3^1}$ are of the same sign.

Third, if we relabel only agent 1's types, i.e. relabel her type 1 as 2 and vice versa, then the left-hand side of (7) becomes $p_{211}p_{222} - p_{212}p_{221}$ and the left-hand side of (8) becomes $p_{112}p_{211} - p_{111}p_{212}$. So, $\gamma_{\theta_1^1 \theta_1^2}$ and $\gamma_{\theta_2^2 \theta_2^1}$ are of the same sign.

Fourth, if we relabel only agent 2's types, i.e. relabel her type 1 as 2 and vice versa, then the left-hand side of (7) becomes $p_{121}p_{112} - p_{122}p_{111}$ and the left-hand side of (8) becomes $p_{121}p_{222} - p_{122}p_{221}$. So, $\gamma_{\theta_1^2 \theta_1^1}$ and $\gamma_{\theta_2^1 \theta_2^2}$ are of the same sign.

Fifth, consider the following two-step relabeling. At first, relabel agent 2's types i.e. relabel her type 1 as 2 and vice versa. After this, relabel players 2 and 3 as 3 and 2, respectively. Then the left-hand side of (7) becomes $p_{112}p_{121} - p_{122}p_{111}$ and the left-hand side of (8) becomes $p_{112}p_{222} - p_{122}p_{212}$. So, $\gamma_{\theta_1^2 \theta_1^1}$ and $\gamma_{\theta_3^1 \theta_3^2}$ are of the same sign.

Altogether, the above five steps establish that all coefficients $\gamma_{\theta_1^1 \theta_1^2}, \gamma_{\theta_1^2 \theta_1^1}, \gamma_{\theta_2^1 \theta_2^2}, \gamma_{\theta_2^2 \theta_2^1}, \gamma_{\theta_3^1 \theta_3^2}, \gamma_{\theta_3^2 \theta_3^1}$ are of the same sign. Further, we can choose $v \in \mathbb{R}$ so that these coefficient are all nonnegative. If there is at least one nonzero coefficient among them, then Eq. (36) holds and hence $S \cap \Omega \neq \emptyset$. Finally, if $\gamma_{\theta_1^1 \theta_1^2} = \gamma_{\theta_1^2 \theta_1^1} = \gamma_{\theta_2^1 \theta_2^2} = \gamma_{\theta_2^2 \theta_2^1} = \gamma_{\theta_3^1 \theta_3^2} = \gamma_{\theta_3^2 \theta_3^1} = 0$, then it is easy to see that Crémer–McLean condition fails for all $i \in \{1, 2, 3\}$, i.e. $p_{-i}(.|\theta_i^1) = k_i p_{-i}(.|\theta_i^2)$ for some $k_i > 0$ and all $i \in \{1, 2, 3\}$. But then (36) holds if we take $\gamma_{\theta_i^1, \theta_i^1} = 1, \gamma_{\theta_i^2, \theta_i^1} = k_i$ for $i \in \{1, 2, 3\}$ and set all other coefficients in (36) to zero.

Proof of Theorem 2. Consider the alternative definition of Identifiability in Lemma 3. Using the vector notation introduced in the proof of Theorem 1, we can rewrite condition (10) in that Lemma as follows:

$$\text{for all } i, \sum_{\theta_i \in \Theta_i} \sum_{\theta_i' \in \Theta_i} s_{\theta_i \theta_i'} p_i(\theta_i) \mathbf{p}_{\theta_i, \theta_i'} + \sum_{\theta_i \in \Theta_i} b_i(\theta_i) p_i(\theta_i) \mathbf{p}_{\theta_i, \theta_i} = \mathbf{q} \in \mathbf{R}_+^L \tag{38}$$

with $\mathbf{q} \neq \alpha \mathbf{p}$ for any $\alpha \in \mathbf{R}_+$, where $\mathbf{p} \in \mathbf{R}_+^L$ is the vector of probabilities corresponding to $p(.)$. By Lemma 3, $p(.)$ is not identifiable iff (38) holds for some strategy profile $(s_1, \ldots, s_n)$ and some collection of functions $b_i(.) \colon \Theta_i \mapsto \mathbf{R}_+$ for $i = 1, \ldots, n$.

Let X_i^j be an $(m_j^2 + m_i^2) \times L$ matrix formed by stacking all vectors $\mathbf{p}_{\theta_j, \theta_j'}$ of player j on top of all player i's vectors $\mathbf{p}_{\theta_i, \theta_i'}$, and then reordering them so that the first column corresponds to θ_j, the second column corresponds to θ_j', and the relative ordering of the other columns remains unchanged.

Let us define **Condition G** as follows: *Suppose that $\mathbf{z} X_i^j = \mathbf{0}$ for some row vector $\mathbf{z} \in \mathbf{R}^{m_j^2 + m_i^2}$ s.t. at least one of the first m_j^2 entries of $\mathbf{z}$ and at least one of its last m_i^2 entries are nonzero. Let $\mathbf{z}_{(k)}$ denote the kth entry of $\mathbf{z}$. Then there exists $\beta \in \mathbf{R}$ s.t. $\mathbf{z}_{((h-1)m_j + h)} = \beta p_j(\theta_j^h)$ for all $h \in \{1, \ldots, m_j\}$, and $\mathbf{z}_{(m_j^2 + (k-1)m_i + k)} = -\beta p_i(\theta_i^k)$ for all $k \in \{1, \ldots, m_i\}$, and all other entries of $\mathbf{z}$ are zero.*

In words, only those elements of $\mathbf{z}$ that are multiplied by some $\mathbf{p}_{\theta_j, \theta_j}$ or $\mathbf{p}_{\theta_i, \theta_i}$ in the product $\mathbf{z} X_i^j$ are nonzero and are proportional to $p_j(\theta_j)$ and $p_i(\theta_i)$, respectively.

Comparing Condition G to (38), we conclude that Condition G implies that $p(.)$ is identifiable. Specifically, if Condition G holds, then (38) can hold only for $\mathbf{q}$ proportional to $\mathbf{p}$, in which case $p(.)$ is identifiable.

So, to prove the Lemma we will show that Condition G holds for a generic probability distribution $p(.) \in \mathcal{P}(\Theta)$ when we take $j \in \arg \min_{l \in \{1, \ldots, n\}} m_l$ and $i \in \arg \min_{l \in \{1, \ldots, n\}, l \neq j} m_l$. We will use the measure-theoretic notion of genericity. Specifically, consider a mapping $f : [0, 1]^L \backslash \mathbf{0} \mapsto \mathcal{P}(\theta)$ such that for any $q(.) \in [0, 1]^L \backslash \mathbf{0}$, we have: $f(q(\theta)) = \frac{q(\theta)}{\sum_{\theta \in \Theta} q(\theta)}$. This transformation is a continuous open map from $[0, 1]^L \backslash \mathbf{0}$ onto $\mathcal{P}(\theta)$. So, to establish genericity, we will consider that $p(.)$ lies in $[0, 1]^L \backslash \mathbf{0}$ (i.e. will not normalize the entries of $p(.)$ to sum up to 1) and show that Condition G fails if and only if $p(.)$ belongs to a subset of $[0, 1]^L \backslash \mathbf{0}$ with Lebesgue measure zero. Since the mapping $f(.)$ is continuous, open and onto, this would also imply that the subset of $\mathcal{P}(\theta)$ where Condition G fails has measure zero.

We will modify X_i^j using a series of the following rank-preserving elementary transformations: (i) interchanging its rows or columns; (ii) multiplying all entries in some row (or column) by a nonzero constant. First, for any $\theta_j \in \Theta_j$ and $\theta_i \in \Theta_i$ multiply the row equal to the vector $\mathbf{p}_{\theta_j, \theta_j'}$ by $p_j(\theta_j)$ and the row equal to the vector $\mathbf{p}_{\theta_i, \theta_i'}$ by $p_i(\theta_i)$. Second, reorder the columns of X_i^j by agent-types in the following sequence: $i, j, 1, \ldots, n$, so that the first L_{-i-j} columns correspond to type profiles $(\theta_{-i-j}, \theta_j^1, \theta_i^1)$ for all $\theta_{-i-j} \in \Theta_{-i-j}$, and so on. These two steps transform X_i^j into the following matrix $\bar{Y}_i^j$:

$$\bar{Y}_i^j = \begin{Vmatrix} M_{i1}^j & 0 & 0 & M_{i2}^j & 0 & 0 & \ldots\ldots & M_{im_i}^j & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \ldots\ldots & \cdots & \cdots & \cdots \\ 0 & 0 & M_{i1}^j & 0 & 0 & M_{i2}^j & \ldots\ldots & 0 & 0 & M_{im_i}^j \\ M_i^{j1} & \ldots & M_i^{jm_j} & 0 & 0 & 0 & \ldots\ldots & 0 & 0 & 0 \\ 0 & 0 & 0 & M_i^{j1} & \ldots & M_i^{jm_j} & \ldots\ldots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \ldots\ldots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots\ldots & M_i^{j1} & \ldots & M_i^{jm_j} \end{Vmatrix},$$

where M_i^{jh} (M_{ik}^j) is an $m_i \times L_{-i-j}$ $(m_j \times L_{-i-j})$ matrix given by

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

where $\mathbf{p}(.; \theta_i^k, \theta_j^h)$ is a row 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.

The matrices X_i^j and $\bar{Y}_i^j$ have the same rank because the employed elementary transformations are rank-preserving. Importantly, there is a one-to-one relationship between the set of solutions to $\mathbf{z}X_i^j = \mathbf{0}$ and the set of solutions $\left(\rho_1, \ldots, \rho_{m_j}, \xi_1, \ldots, \xi_{m_i}\right)$ to the following system:

$$
\left(\rho_1, \ldots, \rho_{m_j}, -\xi_1, \ldots, -\xi_{m_i}\right) \bar{Y}_i^j = \mathbf{0}, \tag{39}
$$

where ρ_h and ξ_k are row vectors of size m_j and m_i respectively. Equation (39) has the following 'basic' solution: for all $h \in \{1, \ldots, m_i\}$ and $k \in \{1, \ldots, m_j\}$, the hth (kth) entry of ρ_h (ξ_k) is nonzero and is equal to some $v \in \mathbf{R}$; all other entries of ρ_h (ξ_k) are equal to zero. This solution corresponds to the solution to $\mathbf{z}X_i^j = \mathbf{0}$ described in the statement of Condition *G*. Therefore, Condition *G* holds if the only nonzero solution to (39) is the 'basic' solution.

To complete the proof let us show that for almost all $p(.)$ the only nonzero solution $\left(\rho_1, \ldots, \rho_{m_j}, \xi_1, \ldots, \xi_{m_i}\right)$ to (39) is the 'basic solution'.

First of all, let us show that for almost all $p(.)$ there do not exist two different nonzero solutions to (39), $\left(\hat{\rho}_1, \ldots, \hat{\rho}_{m_j}, \hat{\xi}_1, \ldots, \hat{\xi}_{m_i}\right)$ and $\left(\tilde{\rho}_1, \ldots, \tilde{\rho}_{m_j}, \tilde{\xi}_1, \ldots, \tilde{\xi}_{m_i}\right)$ such that $\left(\hat{\rho}_1, \ldots, \hat{\rho}_{m_j}\right) = \left(\tilde{\rho}_1, \ldots, \tilde{\rho}_{m_j}\right)$. For, suppose otherwise. Then $\left(0, \ldots, 0, \hat{\xi}_1 - \tilde{\xi}_1, \ldots, \hat{\xi}_{m_i} - \tilde{\xi}_{m_i}\right)$ is also a nonzero solution to (39). Substituting this solution into (39) and inspecting it, we find that this is equivalent to linear nonindependence of m_i vectors of conditional probability distributions $p_{-i}(.|\theta_i^1), \ldots, p_{-i}(.|\theta_i^{m_i})$. However, since $i \in \arg\min_{l \in \{1, \ldots, n\}, l \neq j} m_l$, we have $m_i \leqslant L_{-i}$, so these vectors are not linearly independent only for a set of $p(.) \in [0, 1]^L$ of measure zero.

This results implies that to establish the Lemma it is sufficient to show that the only solution to (39) such that $(\rho_1, \ldots, \rho_{m_j}) \neq \mathbf{0}$ is the 'basic solution.' The proof is given separately for two cases. In Step 2 below, we deal with the case of $n = 3$ and $m_1 = m_2 = m_3 = m \geqslant 3$, while in Step 1 we deal with all other cases. These steps rely on the following two facts:

Fact 1. *A set* $\{(x_1, \ldots, x_L) \in [0, 1]^L | (x_1, \ldots, x_L)$ *satisfies a finite number of polynomial equations} has measure zero.*

Fact 2. *Let* M_i^{jh} (M_{ik}^j) *be an* $m_i \times L_{-i-j}$ $(m_j \times L_{-i-j})$ *matrix s.t.*

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

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 kth row. Finally, let $Y_{i(k)}^j$ be an $(m_j^2 + m_i - 1) \times L_{-i}$ matrix such that:

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

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

(i) If $m_i \leqslant L_{-i-j}$ ($m_j \leqslant L_{-i-j}$), 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 nonsingular.

(ii) If $m_j^2 + m_i - 1 \leqslant L_{-i}$, 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 $\{U_1, U_2, \ldots, U_{m_i}\}$ be a collection of principal minors of M_{ik}^j. All such minors exist because $m_j \leqslant L_{-i-j}$, i.e. M_{ik}^j has more columns than rows. The determinant (det(.)) of U_l is a nondegenerate polynomial. The nondegeneracy 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 $l \in \{1, \ldots, m_j - 1\}$ is equal to zero on a subset of $[0, 1]^L$ of measure zero and consider the determinant of U_{l+1}. We have

$$
\det(U_{l+1}) = \sum_{t=1}^{l+1} (-1)^{l+1+t} p(\theta_{-i-j}^{(t)}, \theta_i^k, \theta_j^{l+1}) \det(U_{(\theta_{-i-j}^{(t)}, \theta_i^k, \theta_j^{l+1})}), \tag{40}
$$

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 equal to the determinants of the complementary minors. These determinants 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 $l \in \{1, \ldots, m_j\}$ for almost all $p(.)$.

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

Now suppose that $Y_{m_j^2+s}$, $s \in \{0, m_i - 2\}$, is nonsingular 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:

$$\det(Y_{m_j^2+s+1}) = \sum_{t=1}^{m_j^2+s+1} (-1)^{m_j^2+s+1+t} g^{(t)} \det(Y_{m_j^2+s+1}^{(t)}), \tag{41}$$

where $g^{(t)}$ is the t-th element of the last row of $Y_{m_j^2+s+1}$ equal to $(\mathbf{p}(.; \theta_i^r, \theta_j^1), \ldots, \mathbf{p}(.; \theta_i^r, \theta_j^{m_j}))$, where $r = s + 1$ if $s < k - 1$ and $r = s + 2$ if $s \geqslant k - 1$, 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 $t \in \{1, \ldots, m_j^2 + s + 1\}$, $Y_{m_j^2+s+1}^{(t)}$ does not contain any elements equal to the entries of m_j vectors $\mathbf{p}(\theta_{-i-j}, \theta_i^r, \theta_j^1), \ldots, \mathbf{p}(\theta_{-i-j}, \theta_i^r, \theta_j^{m_j})$. 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 $s' \leqslant m_i - 1$ for almost all $p(.)$. $\square$

Now we are ready to prove Steps 1 and 2 described above.

Step 1: Suppose that either $n \geqslant 4$, or $n = 3$ and it is not true that $m_1 = m_2 = m_3$.

Fix some type $k \in \{1, \ldots, m_i\}$, 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}-th 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 nonzero entries in these columns (which are the rows from 1st to m_j^2-th and from $(k-1)m_i$-th to km_i-th). Obviously, if $(\rho_1, \ldots, \rho_{m_j}, -\xi_1, \ldots, -\xi_{m_i})$ solves (39), then $(\rho_1, \ldots, \rho_{m_j}, -\xi_k)\hat{Y}_{ik}^j = 0$.

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 $j \in \arg\min_{h \in \{1,\ldots,n\}} m_h$ and $i \in \arg\min_{h \in \{1,\ldots,n\}, h \neq j} m_j$. Therefore, $m_j < L_{-i-j}$ and $m_j^2 + m_i - 1 \leqslant L_{-i}$. 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 2: $n = 3$ and $m_1 = m_2 = m_3 = m \geqslant 3$.

To show that system (39) has a unique nonzero solution, first, notice that (39) is equivalent to: $\rho_h M_{ik}^j = \xi_k M_i^{jh}$ for all $k \in \{1, \ldots, m\}$ and $h \in \{1, \ldots, m\}$. By Fact 2 and $m_1 = m_2 = m_3 = m$, M_{ik}^j and M_i^{jh} are $m \times m$ nonsingular matrices for almost all $p(.)$. Hence, we can assume their nonsingularity in the rest of the proof. Then, (39) is equivalent to the following

system: $\xi_k = \rho_h M_{ik}^j (M_i^{jh})^{-1}$ for all k and h, which implies that:

$$(\rho_1, -\rho_h) \begin{pmatrix} M_{i1}^j (M_i^{j1})^{-1} & M_{i2}^j (M_i^{j1})^{-1} & \cdots & M_{im}^j (M_i^{j1})^{-1} \\ M_{i1}^j (M_i^{jh})^{-1} & M_{i2}^j (M_i^{jh})^{-1} & \cdots & M_{im}^j (M_i^{jh})^{-1} \end{pmatrix} = 0$$

for all $h \in \{2, \ldots, m\}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (42)

So, (39) has a unique solution iff (42) has a unique solution. In turn, (42) has a unique solution if the following square submatrix of it has rank $2m - 1$ for almost all $p(.)$:

$$\begin{pmatrix} M_{i1}^j (M_i^{j1})^{-1} & M_{im}^j (M_i^{j1})^{-1} \\ M_{i1}^j (M_i^{jm})^{-1} & M_{im}^j (M_i^{jm})^{-1} \end{pmatrix}. \qquad\qquad\qquad\qquad\qquad\qquad\qquad (43)$$

To complete the proof, let us show that the principal minors $\{Z_m, Z_{m+1}, \ldots, Z_{2m-1}\}$ of (43) (Z_l is an $l \times l$ matrix consisting of the elements of the first l rows and l columns of (43)) have nonzero determinants for almost all $p(.)$. We will proceed by induction. First, $Z_m = M_{i1}^j (M_i^{j1})^{-1}$ is nonsingular by Fact 2. Next, suppose that $\det(Z_{m+s-1}) \neq 0$ for some $s \in \{1, \ldots, m-1\}$. Let us show that $\det(Z_{m+s}) \neq 0$ for almost all $p(.)$. We have

$$\det(Z_{m+s}) = \sum_{t=1}^{m+s} (-1)^{m+s+t} b_{m+s}^t \det(Z_{m+s}^{-t}), \qquad\qquad\qquad\qquad\qquad (44)$$

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 $t \in \{1, \ldots, m\}$, and $b_{m+s}^{m+t'} = \sum_{h=1}^{m} p(\theta_a^h, \theta_i^m, \theta_j^s) d_{ht'}$ for $t' \in \{1, \ldots, s\}$. These are nonzero for almost all $p(.)$. Since $\det(Z_{m+s}^{-(m+s)}) = \det(Z_{m+s-1}) \neq 0$, the last entry $b_{m+s}^{m+s} \det(Z_{m+s}^{-(m+s)})$ of (44) is nonzero. If all other entries are zero, then $\det(Z_{m+1}) \neq 0$ for almost all $p(.)$. If there are other nonzero entries in the summation in (44), then with the new notation we can rewrite it as follows:

$$\det(Z_{m+s}) = \sum_{h=1}^{m} p(\theta_a^h, \theta_i^1, \theta_j^s) \sum_{t=1}^{m} d_{ht} \det(Z_{m+s}^{-t})(-1)^{m+s+t}$$
$$+ \sum_{h=1}^{m} p(\theta_a^h, \theta_i^m, \theta_j^s) \sum_{t=1}^{s} d_{ht} \det(Z_{m+s}^{-(m+t)})(-1)^{s+t}. \qquad\qquad (45)$$

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 $h, h' \in \{1, \ldots, m\}$, $s \neq m$, 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 $s \neq m$, (45) 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 matrix (43) is $2m - 1$ for almost all $p(.) \in [0, 1]^L$.

Proof of Lemma 1. The proof is by contrapositive. So, suppose that $p(.)$ is not identifiable, i.e. there exists $q(.) \in \mathcal{P}(\Theta)$, $q(.) \neq p(.)$, such that for all $i \in \{1, \ldots, n\}$ and $\theta_i' \in \Theta_i$, with $q_i(\theta_i') > 0$, we have: $q_{-i}(.|\theta_i') = \sum_{\theta \in \Theta_i} c_{\theta_i \theta_i'} p_{-i}(.|\theta_i)$ for some collection of nonnegative coefficients $c_{\theta_i \theta_i'}$.

Fix some $i \in \{1, \ldots, n\}$. We need to show that there exists $\hat{\theta}_i \in \Theta_i$ such that for any agent $j \neq i$ and any $\hat{\theta}_j \in \Theta_j$, the collection of $m_i + m_j - 1$ vectors of conditional probability distributions $p_{-i-j}(.|\theta_i, \hat{\theta}_j)$, $p_{-i-j}(.|\hat{\theta}_i, \theta_j)$, $\theta_i \in \Theta_i$, $\theta_j \in \Theta_j$, $\theta_j \neq \hat{\theta}_j$, is not linearly independent.

By the rules of conditional expectation, $p_{-i}(\theta_{-i-j}, \theta_j|\theta_i) = p_{-i-j}(\theta_{-i-j}|\theta_i, \theta_j) p_j(\theta_j|\theta_i)$ and $p_{-j}(\theta_{-i-j}, \theta_i|\theta_j) = p_{-i-j}(\theta_{-i-j}|\theta_i, \theta_j) p_i(\theta_i|\theta_j)$ for any $\theta_{-i-j} \in \Theta_{-i-j}$, $\theta_i \in \Theta_i$ and $\theta_j \in \Theta_j$. Therefore, for any $\theta_{-i-j} \in \Theta_{-i-j}$ we have

$$
\begin{aligned}
q(\theta_{-i-j}, \hat{\theta}_i, \hat{\theta}_j) &= q_i(\hat{\theta}_i) \sum_{\theta_i \in \Theta_i} c_{\theta_i \hat{\theta}_i} p_j(\hat{\theta}_j|\theta_i) p_{-i-j}(\theta_{-i-j}|\theta_i, \hat{\theta}_j) \\
&= q_j(\hat{\theta}_j) \sum_{\theta_j \in \Theta_j} c_{\theta_j \hat{\theta}_j} p_i(\hat{\theta}_i|\theta_j) p_{-i-j}(\theta_{-i-j}|\hat{\theta}_i, \theta_j).
\end{aligned}
\tag{46}
$$

Since (46) holds for all $\theta_{-i-j} \in \Theta_{-i-j}$, the collection of vectors of conditional probability distributions $p_{-i-j}(.|\theta_i, \hat{\theta}_j)$, $p_{-i-j}(.|\hat{\theta}_i, \theta_j)$, with $\theta_i \in \Theta_i$, $\theta_j \in \Theta_j$, $\theta_j \neq \hat{\theta}_j$, is not linearly independent, if either (i) there exists $\theta_i \in \Theta_i$, $\theta_i \neq \hat{\theta}_i$ s.t. $c_{\theta_i \hat{\theta}_i} p_j(\hat{\theta}_j|\theta_i) > 0$ or (ii) there exists $\theta_j \in \Theta_j$, $\theta_j \neq \hat{\theta}_j$ s.t. $c_{\theta_j \hat{\theta}_j} p_i(\hat{\theta}_i|\theta_j) > 0$.

To see that either (i) or (ii) is true, first note that by our regularity assumption $p_j(\hat{\theta}_j|\theta_i) > 0$ and $p_i(\hat{\theta}_i|\theta_j) > 0$ for any $\theta_i \in \Theta_i$ and $\theta_j \in \Theta_j$. Next, recall that $\hat{\theta}_i$ can be chosen arbitrarily. So, if there exist θ_i and $\theta_i' \in \Theta_i$, s.t. $\theta_i' \neq \theta_i$ and $c_{\theta_i \theta_i'} > 0$, then (i) holds for $\hat{\theta}_i = \theta_i'$. On the other hand, if $c_{\theta_i \theta_i'} = 0$ for all $\theta_i, \theta_i' \in \Theta_i$ s.t. $\theta_i' \neq \theta_i$, then by Lemma A1 for all $\theta_j' \in \Theta_j$ there exists θ_j s.t. $\theta_j \neq \theta_j'$ and $c_{\theta_j \theta_j'} > 0$. So, in this case (ii) is true for any $\hat{\theta}_i \in \Theta_i$. $\square$

Proof of Lemma 2. The proof consists of two main steps.

Step 1: Condition C is equivalent to the following Condition C':

Consider any collection of scalars $\{\mu_\theta, \gamma_{\theta_i \theta_i'} \geqslant 0 | i = 1, \ldots, n, \theta_i, \theta_i' \in \Theta_i, \theta_i' \neq \theta_i\}$ such that

$$
\sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \sum_{\theta_i' \in \Theta_i : \theta_i' \neq \theta_i} \gamma_{\theta_i \theta_i'} \left(\mathbf{p}_{\theta_i \theta_i} - \mathbf{p}_{\theta_i \theta_i'} \right) + \sum_{\theta \in \Theta} \mu_\theta \mathbf{e}_\theta = \mathbf{0}.
\tag{47}
$$

Then $\mu_\theta = 0$ for all $\theta \in \Theta$.   (48)

Proof of Step 1. In the vector notation introduced in the proof of Theorem 1, Condition C says the following: For any function $R(.) : \Theta \mapsto \mathbf{R}$ there exists a vector $\mathbf{t} \in \mathbf{R}^{nL}$ such that:

$$
(\mathbf{p}_{\theta_i \theta_i} - \mathbf{p}_{\theta_i \theta_i'}) \cdot \mathbf{t} \geqslant 0 \quad \text{for all } i = 1, \ldots, n, \ \theta_i, \theta_i' \in \Theta_i, \ \theta_i' \neq \theta_i,
\tag{49}
$$

$$
\mathbf{e}_\theta \cdot \mathbf{t} = R(\theta) \quad \text{for all } \theta \in \Theta.
\tag{50}
$$

Let us fix some $R(.) : \Theta \mapsto \mathbf{R}$. Then by the Theorem of Alternative (see Mangasarian [15, p. 34]) there exists $\mathbf{t} \in \mathbf{R}^{nL}$ solving (49)–(50) if and only if for any collection of scalar coefficients $\{\mu_\theta, \gamma_{\theta_i \theta_i'} \geqslant 0 | i \in \{1, \ldots, n\}, \theta_i, \theta_i' \in \Theta_i, \theta_i' \neq \theta_i\}$ satisfying (47) we have

$$
\sum_{\theta \in \Theta} \mu_\theta R(\theta) \geqslant 0.
\tag{51}
$$

Obviously, if Condition C' holds, then (51) is equal to zero for all $R(\theta)$. So Condition C' implies Condition C.

To show that Condition C implies Condition C', suppose that the latter fails, i.e. (47) holds for some collection of coefficients such that $\mu_\theta \neq 0$ for some $\theta \in \Theta$. Then (51) fails if we take $R(\theta) = -\mu_\theta$ for all $\theta \in \Theta$, so Condition C also fails. Thus, Conditions C and C' are equivalent.

Step 2: Condition C' is equivalent to Weak Identifiability.

Proof of Step 2. If Weak Identifiability fails, then there exists a profile of agents' nontruthful strategies $(s_1, \ldots, s_n)$ such that the profile of induced probability distributions $\pi(.|s_i, s^*_{-i})$, $i \in \{1, \ldots, n\}$, satisfies the following equation for all $\theta' \in \Theta$ and $i \in \{1, \ldots, n\}$:
$\pi(\theta'|s_i, s^*_{-i}) - p(\theta') = \Delta\pi(\theta')$, with $\Delta\pi(\theta') \neq 0$ for some $\theta' \in \Theta$. Then,

$$
\begin{aligned}
\Delta\pi(\theta') = \pi(\theta'|s_i, s^*_{-i}) - p(\theta') &= \sum_{\theta_i \in \Theta_i} s_{\theta_i \theta'_i} p(\theta'_{-i}, \theta_i) - p(\theta') \\
&= \sum_{\theta_i \in \Theta_i : \theta_i \neq \theta'_i} s_{\theta_i \theta'_i} p(\theta'_{-i}, \theta_i) + (s_{\theta'_i \theta'_i} - 1) p(\theta') \\
&= \sum_{\theta_i \in \Theta_i : \theta_i \neq \theta'_i} s_{\theta_i \theta'_i} p(\theta'_{-i}, \theta_i) - \sum_{\theta_i \in \Theta_i : \theta_i \neq \theta'_i} s_{\theta'_i \theta_i} p(\theta').
\end{aligned}
\tag{52}
$$

Let us set $\gamma_{\theta_i \theta'_i} = s_{\theta_i \theta'_i} p_i(\theta_i)$ for all $i \in \{1, \ldots, n\}$ and $\theta_i, \theta'_i \in \Theta_i$, $\theta'_i \neq \theta_i$. Then the $\mathbf{R}^{nL}$-vector $\sum_{\theta_i \in \Theta_i} \sum_{\theta'_i \in \Theta_i : \theta'_i \neq \theta_i} \gamma_{\theta_i \theta'_i} \left(\mathbf{p}_{\theta_i \theta_i} - \mathbf{p}_{\theta_i \theta'_i} \right)$ is such that its entry corresponding to agent i and some $\theta' \in \Theta$ is equal to $-\Delta\pi(\theta')$ and all other entries (i.e. the ones corresponding to $j \in \{1, \ldots, n\}$, $j \neq i$) are equal to zero. Therefore, setting $\hat{\mu}_{\theta'} = \Delta\pi(\theta')$ for all $\theta' \in \Theta$, yields $\sum_{i=1,\ldots,n} \sum_{\theta_i \in \Theta_i} \sum_{\theta'_i \in \Theta_i : \theta'_i \neq \theta_i} \gamma_{\theta_i \theta'_i} \left(\mathbf{p}_{\theta_i \theta_i} - \mathbf{p}_{\theta_i \theta'_i} \right) + \sum_{\theta \in \Theta'} \hat{\mu}_{\theta'} \mathbf{e}_{\theta'} = \mathbf{0}$. So, Condition C' fails.

In the opposite direction, suppose that Condition C' fails, i.e. there is a collection of coefficients $\{\mu_\theta, \gamma_{\theta_i \theta'_i} \geqslant 0 | i \in \{1, \ldots, n\}, \theta \in \Theta, \theta_i, \theta'_i \in \Theta_i, \theta'_i \neq \theta_i\}$ satisfying (47), with $\mu_\theta \neq 0$ for some $\theta \in \Theta$. Since multiplying all coefficients by some positive number preserves the failure of Condition C', we can without loss of generality assume that $\sum_{\theta'_i \in \Theta_i, \theta'_i \neq \theta_i} \gamma_{\theta_i \theta'_i} \leqslant p_i(\theta_i)$ for all i and θ_i. Let us now construct a profile of strategies $(s_1, \ldots, s_n)$ as follows. Set $s_{\theta_i \theta'_i} = \gamma_{\theta_i \theta'_i} / p_i(\theta_i)$ and $s_{\theta_i \theta_i} = 1 - \sum_{\theta'_i \in \Theta_i, \theta'_i \neq \theta_i} s_{\theta_i \theta'_i}$ for all i and $\theta_i, \theta'_i \in \Theta_i, \theta_i \neq \theta'_i$. Substituting these values into (52) and using (47), we obtain that for all $i \in \{1, \ldots, n\}$ and all $\theta' \in \Theta$,

$$
\pi(\theta'|s_i, s^*_{-i}) - p(\theta') = \sum_{\theta_i \in \Theta_i : \theta_i \neq \theta'_i} \gamma_{\theta_i \theta'_i} p_{-i}(\theta'_{-i}|\theta_i) - \sum_{\theta_i \in \Theta_i : \theta_i \neq \theta'_i} \gamma_{\theta'_i \theta_i} p_{-i}(\theta'_{-i}|\theta'_i) = \mu_{\theta'}.
$$

Since $\mu_{\theta'} \neq 0$ for some $\theta' \in \Theta$, it follows that $\pi(.|s_i, s^*_{-i}) = \pi(.|s_j, s^*_{-j}) \neq p(.)$ for all $i, j \in \{1, \ldots, n\}$. So, $p(.)$ is not weakly identifiable. $\quad\square$

Proof of Lemma 3. Suppose that $p(.)$ is not identifiable, i.e. for some $q(.) \in \mathcal{P}(\Theta)$, $q(.) \neq p(.)$, there exists a collection of coefficients $\{c_{\theta_i \theta'_i} \geqslant 0 | i \in \{1, \ldots, n\}, \theta_i, \theta'_i \in \Theta_i\}$ s.t. for all i and $\theta'_i \in \Theta$ we have $q_{-i}(.|\theta'_i) = \sum_{\theta_i \in \Theta_i} c_{\theta_i \theta'_i} p_{-i}(.|\theta_i)$.

Let $F = \max_{i \in \{1,\dots,n\}} \left\{ \max_{\theta_i' \in \Theta_i} \frac{\sum_{\theta_i \in \Theta_i} c_{\theta_i \theta_i'}}{p_i(\theta_i)} \right\}$. Next, for every i, and $\theta_i, \theta_i' \in \Theta_i, \theta_i \neq \theta_i'$, set:

$$s_{\theta_i \theta_i'} = \frac{c_{\theta_i \theta_i'} q_i(\theta_i')}{p_i(\theta_i) F}, \quad s_{\theta_i \theta_i} = 1 - \sum_{\theta_i' \in \Theta_i : \theta_i' \neq \theta_i} s_{\theta_i \theta_i'} \geqslant 0,$$

$$b_i(\theta_i) = (1 - s_{\theta_i \theta_i}) + \frac{c_{\theta_i \theta_i} q_i(\theta_i)}{F p_i(\theta_i)} \geqslant 0.$$

Then for all $i \in \{1, \dots, n\}$ and all $\theta' \equiv (\theta_1', \dots, \theta_n') \in \Theta$ we have

$$\pi(\theta' | s_i, s_{-i}^*) + b_i(\theta_i') p(\theta') = \sum_{\theta_i \in \Theta_i} \frac{q_i(\theta_i') c_{\theta_i \theta_i'}}{F} p_{-i}(\theta_{-i}' | \theta_i) + p(\theta') = \frac{q(\theta')}{F} + p(\theta').$$

The first equality follows from Definition 3 of induced probability distribution $\pi(. | s_1, \dots, s_n)$ and the definition of $b_i(.)$. The second equality holds because $q_{-i}(. | \theta_i') = \sum_{\theta_i \in \Theta_i} c_{\theta_i \theta_i'} p_{-i}(. | \theta_i)$.

Since the right-hand side of this equation $\frac{q(\theta')}{F} + p(\theta')$ is independent of i and $\frac{q(.)}{F} + p(.) \neq \alpha p(.)$ for all $\alpha \geqslant 0$, we have constructed a strategy profile $(s_1, \dots, s_n)$ and a collection of functions $b_i(.): \Theta_i \mapsto \mathbf{R}_+$ for $i = 1, \dots, n$ satisfying (10).

Conversely, suppose that for some $\hat{q}(.): \Theta \mapsto \mathbf{R}_+$, s.t. $\hat{q}(.) \neq \alpha p(.)$ for any $\alpha \geqslant 0$, there is a strategy profile $(s_1, \dots, s_n)$, $s_i \in S_i$, and a collection of functions $b_i(.): \Theta \mapsto \mathbf{R}_+$, for $i \in \{1, \dots, n\}$, satisfying $\pi(\theta' | s_i, s_{-i}^*) + b_i(\theta_i') p(\theta') = \hat{q}(\theta')$ for all $\theta' \in \Theta$. Let $\tilde{q}(.) \equiv \frac{\hat{q}(.)}{\sum_{\theta \in \Theta} \hat{q}(\theta)} \in \mathcal{P}(\Theta)$. Then:

$$\tilde{q}(\theta') = \sum_{\theta_i \in \Theta_i} \frac{s_{\theta_i \theta_i'} p(\theta_{-i}', \theta_i)}{\sum_{\theta \in \Theta} \hat{q}(\theta)} + \frac{b_i(\theta_i') p(\theta')}{\sum_{\theta \in \Theta} \hat{q}(\theta)} \quad \text{for all } \theta' \in \Theta.$$

Using the identity $\hat{q}_i(\theta_i') \equiv \left(\sum_{\theta \in \Theta} \hat{q}(\theta) \right) \tilde{q}_i(\theta_i')$ we can rewrite the above expression as follows:

$$\tilde{q}_{-i}(. | \theta_i') = \sum_{\theta_i \in \Theta_i : \theta_i \neq \theta_i'} \frac{s_{\theta_i \theta_i'} p_i(\theta_i)}{\hat{q}_i(\theta_i')} p_{-i}(. | \theta_i) + \frac{p_i(\theta_i')(s_{\theta_i' \theta_i'} + b_i(\theta_i'))}{\hat{q}_i(\theta_i')} p_{-i}(. | \theta_i').$$

Since $\tilde{q}(.) \neq p(.)$, $p(.)$ is not identifiable. $\quad \square$

References

[1] M. Aoyagi, Correlated types and Bayesian incentive compatible mechanisms with budget balance, J. Econ. Theory 79 (1998) 142–151.

[2] C. d'Aspremont, J. Crémer, L.-A. Gérard-Varet, Incentive and the existence of pareto-optimal revelation mechanisms, J. Econ. Theory 51 (1990) 233–254.

[3] C. d'Aspremont, J. Crémer, L.-A. Gérard-Varet, Correlation, independence and Bayesian incentives, Soc. Choice Welfare 21 (2003) 1–30.

[4] C. d'Aspremont, J. Crémer, L.-A. Gérard-Varet, Balanced Bayesian mechanisms, J. Econ. Theory 115 (2004) 385–396.

[5] C. d'Aspremont, L.-A. Gérard-Varet, Incentive and incomplete information, J. Public Econ. 11 (1979) 25–45.

[6] K.-S. Chung, A note on Matsushima's regularity condition, J. Econ. Theory 87 (1999) 429–433.

[7] J. Cremer, R. McLean, Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent, Econometrica 53 (1985) 345–361.

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

 [9] D. Fudenberg, D. Levine, E. Maskin, The Folk theorem with imperfect public information, Econometrica 62 (1994) 997–1039.
[10] D. Fudenberg, D. Levine, E. Maskin, Balanced-budget mechanisms with incomplete information, mimeo, Department of Economics, UCLA, 1996.
[11] A. Heifetz, Z. Neeman, On the generic (im)possibility or full surplus extraction in mechanism design, Econometrica 117 (2006) 213–233.
[12] P. Jehiel, B. Moldovanu, Efficient design with interdependent valuations, Econometrica 69 (2001) 1237–1261.
[13] V. Krishna, M. Perry, Efficient mechanism design, mimeo, Department of Economics, Pennsylvannia State University, 2002.
[14] G. Mailath, A. Postlewaite, Asymmetric information bargaining problems with many agents, Rev. Econ. Stud. 57 (1990) 351–367.
[15] L. Makowski, C. Mezzetti, Bayesian and weakly robust first best mechanisms: characterizations, J. Econ. Theory 64 (1994) 500–520.
[16] O. Mangasarian, Nonlinear Programming, McGraw-Hill, New York, 1969.
[17] H. Matsushima, Incentive compatible mechanisms with full transferability, J. Econ. Theory 79 (1991) 198–203.
[18] H. Matsushima, Mechanism design with side payments: individual rationality and iterative dominance, J. Econ. Theory 133 (2007) 1–30.
[19] P. McAfee, P. Reny, Correlated information and mechanism design, Econometrica 60 (1992) 395–421.
[20] R. McLean, A. Postlewaite, Informational size and incentive compatibility, Econometrica 70 (2002) 2421–2455.
[21] R. McLean, A. Postlewaite, Informational size and incentive compatibility with aggregate uncertainty, Games Econ. Behav. 45 (2003) 410–433.
[22] R. McLean, A. Postlewaite, Informational size and efficient auctions, Rev. Econ. Stud. 71 (2004) 809–827.
[23] C. Mezzetti, Mechanism design with interdependent valuations: efficiency, Econometrica 72 (2004) 1617–1626.
[24] C. Mezzetti, Mechanism design with interdependent valuations: surplus extraction, Econ. Theory 31 (2007) 473–488.
[25] R. Myerson, M. Satterthwaite, Efficient mechanisms for bilateral trading, J. Econ. Theory 28 (1983) 265–281.
[26] Z. Neeman, The relevance of private information in mechanism design, J. Econ. Theory 117 (2004) 55–77.
[27] R. Rob, Pollution claim settlements under private information, J. Econ. Theory 47 (1989) 307–333.
[28] S. Williams, A characterization of efficient, Bayesian incentive compatible mechanisms, Econ. Theory 14 (1999) 155–180.